Role of replicative primase in lesion bypass during DNA replication

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

Farimah Borazjani Gholami

February 2017

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

Farimah Borazjani Gholami

University of Sussex

Farimah. Borazjani Gholami

Doctor of Philosophy Biochemistry

Role of replicative primase in lesion bypass during DNA replication

Summary

Maintenance of genome integrity and stability is fundamental for any form of life. This is complicated as DNA is highly reactive and always under attack from a wide range of endogenous and exogenous sources which can lead to different damages in the DNA. To preserve the integrity of DNA replication, cells hav evolved a variety of DNA repair pathways. DNA damage tolerance mechanisms serve as the last line of defence to rescue the stalled replications forks. TLS, error-prone type of DNA damage tolerance, acts to bypass DNA lesions and allows continuation of DNA replication. Surprisingly majority of archaeal species lack canonical TLS polymerases. This poses a question as to how archaea restart stalled replication in the absence of TLS or lesion repair pathways. This thesis establishes that archaeal replicative primase (PriS/L), a member of the archaeo-eukaryotic primase (AEP) superfamily, possessing both primase and polymerase activities, is able to bypass the most common oxidative damages and highly distorting lesions caused by UV radiation. It has been postulated that archaeal replicative polymerases (Pol B and Pol D family Pols) can bind tightly to the deaminated bases uracil and cause replication fork stalling four bases prior to dU. A specific mechanism for resuming replication of uracil containing DNA by PriS/L is suggested in this thesis.

In this thesis, we also reported how the enzymatic activities of archaeal PriS/L are regulated. Here, it is demonstrated that in contrast to archaeal replicative polymerases, single-strand binding proteins (RPA) significantly limit the polymerase activity of PriS/L. The remaining results chapter interrogates the possible interactions between PriS/L and RPA. Finally, the attempts to reconstitute an archaeal CMG complex *in vitro*, with the aim of shedding light on the role of archaeal replicative primase in replication-specific TLS are described.

Acknowledgments

Firstly, I would like to express my sincere gratitude to my supervisor, Prof. Aidan Doherty, for giving me the opportunity to do my PhD in his laboratory on such an interesting project and for his continued support and guidance in all the time of research. Beside my supervisor, I would like to thank all the current and former members of Doherty laboratory, including, Stanislaw, Laura, Nigel, Ben, Tom, Pierre, Peter, Alfredo, Przemek and Katie for their invaluable scientific expertise and guidance.

A special thanks to Stanislaw who taught me so much and also for his insightful comments and suggestions. I would like to thank Laura for helping me and showing me how to function in a lab.

I would like to express my special appreciation to my committee members, Dr. Jo Murray, Dr. Jon Baxter and Dr. Antony Oliver for their helpful advice.

Thank you to all people of the GDSC, for providing a good atmosphere to work. I would like to thank my recently graduated friends and also Sahar, my late night lab buddy.

My deepest thank goes to my mother, father, brother and my aunt and rest of my family. Their patient and support during these past years enabled me to complete this thesis. A special word of thanks also goes to my husband, Masoud Hasrati for his encouragement and continued support.

I would like to dedicate this work to my parents

Table of Contents

Abbreviations

(6-4)PP Pyrimidine	(6-4) pyrimidone photoproducts
8-oxo-G	8-oxo-2'-deoxyguanosine
AEP	Archaeo-eukaryotic primase
AP	Apurinic/apyrimidinic
APE1	AP endonuclease 1
APS	Ammonium persulphate
ATP	Adeonsine triphosphate
BER	Base excision repair
bp	Base pairs
BSA	Bovine serum albumin
CAK	Cyclin-dependent kinase-activating kinase
CDC	Cell division cycle
CDK	Cyclin-dependent kinase
cDNA	Complementary DNA
CMG	CDC45, MCM, GINS
CPD	Cyclobutane pyrimidine dimer
dATP	Deoxyadenosine triphosphate
dCTP	Deoxycytosine triphosphate
DDK	Dfb4-dependent kinase
dGTP	Deoxyguanosine triphosphate
DNA	Deoxyribonucleic Acid
dNTP	Deoxynucleoside triphosphate
DSB	Double-stranded DNA break
DSBR	Double strand break repair
dsDNA	Double-stranded DNA
DTT	Dithiothreitol

dTTP	Deoxythymidine triphosphate
dUTP	Deoxyuridine triphosphate
EMSA	Electrophoretic mobility shift assay
exo	Exonuclease
FEN1	Flap Endonuclease 1
GINS	Go-Ichi-Ni-San
HR	Homologous recombination
HRP	Horseradish peroxidase
IMAC	Immobilised metal affinity column
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB	Lysogeny broth
MCM	Mini-chromosome maintenance
MGMT	O6-methylguanine-DNA-methyltransferase
MMEJ	Microhomology-mediated end joining
MMR	Mismatch Repair
MRN	Mre11, Rad50, Nbs1
mRNA	Messenger RNA
mtDNA	Mitochondrial DNA
mtSSB	Mitochondrial single-stranded binding protein
NCLDV	Nucleo-cytoplasmic large DNA virus
NER	Nucleotide Excision Repair
NHEJ	Non-homologous end-joining
Ni-NTA	Nickel-nitrolotriacetic acid
OB	Oligosaccharide/Oligonucleotide-binding
ODx	Optical density at wavelength x nanometres
ORC	Origin of replication complex
ORF	Open reading frame
PAD	Polymerase associated domain
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction

pI	Isoelectric point
PMSF	Phenylmethanesulfonylfluoride
pre-RC	Pre-replication complex
PVDF	Polyvinyladine fluoride
RBM	RPA-binding motif
RecJdbd	RecJ-like DNA-binding domain
RFC	Replication factor C
rNTP	Ribonucleoside triphosphate
RPA	Replication protein A
x g	Revolutions per minute
SDS	Sodium dodecyl sulphate
SDS-PAGE	SDS polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
SOB	Super optimal broth
SSB	Single-stranded DNA break
ssDNA	Single-stranded DNA
TCEP	Tris(2-carboxyethyl)phosphine
TEMED	N,N,N',N'-Tetramethylethylenediamine
Tg	Thymine glycol
TLS	Translesion DNA synthesis
Topo	Topoisomerase
tRNA	Transfer RNA
UV	Ultraviolet
XP	Xeroderma pigmentosum

List of Figures

Chapter 1

List of Tables

Chapter 1

Introduction

1.1 DNA

1.1.1. Biological function of DNA

The most fundamental and radical feature of all living organisms is the ability to reproduce giving rise to progeny and this would not be possible without the existence of a genetic make-up containing all the instructions about the organism's structure, metabolism, and reproduction. Although Oswald T. Avery and his co-workers reported DNA as the source of genetic material in 1940s, most scientists believed that proteins were the site of the gene rather than DNA. Avery expanded the earlier work of Griffith who demonstrated that the non-pathogenic strain of *Streptococcus pneumonia* could be transformed into a pathogenic strain. Therefore, he discovered the process of transformation in 1928 (Griffith, 1928). Sixteen years later Oswald T.Avery and colleagues investigated that the chemical nature of this transforming factor is most likely DNA (Avery *et al*., 1944). This discovery was then confirmed by other scientists in 1952 (Hershey and Chase, 1952). Soon after, Watson and Crick published their revolutionary discovery of DNA based on the X-ray data provided by Wilkins and Franklin and information of the base ratio data provided by Chargaff (Chargaff *et al*., 1952; Watson and Crick, 1953; Zamenhof *et al*., 1952).

1.1.2. Structure of DNA

Early in 1953, Watson and Crick published the double-helix structural model of DNA based on the diffraction data of Rosalind Franklin, Ray Gosling and Maurice Wilkins. DNA structure consists of two helical strands and each chain is coiled around the same axis (Figure 1.1). The two DNA stands are anti-parallel and complementary. The DNA consists of four deoxynucleotides, which covalently connected to one another. Each nucleotide possesses five-carbon sugar, 2' deoxyribose. At the 5' position of the sugar one phosphate group is esterified and one nitrogen group is bound to the 1' position. Each chain contains two type of nitrogen bases, purines (A-adenine, G-Guanine) and pyrimidines (C-cytosine and T-thymine). Complementary bases from each DNA strand are linked by hydrogen bonds. The chains are connected by a dyad perpendicular to the fibre axis. The nitrogen bases are situated on the inside of the helix while the phosphates are situated on the outside (Watson and Crick, 1953).

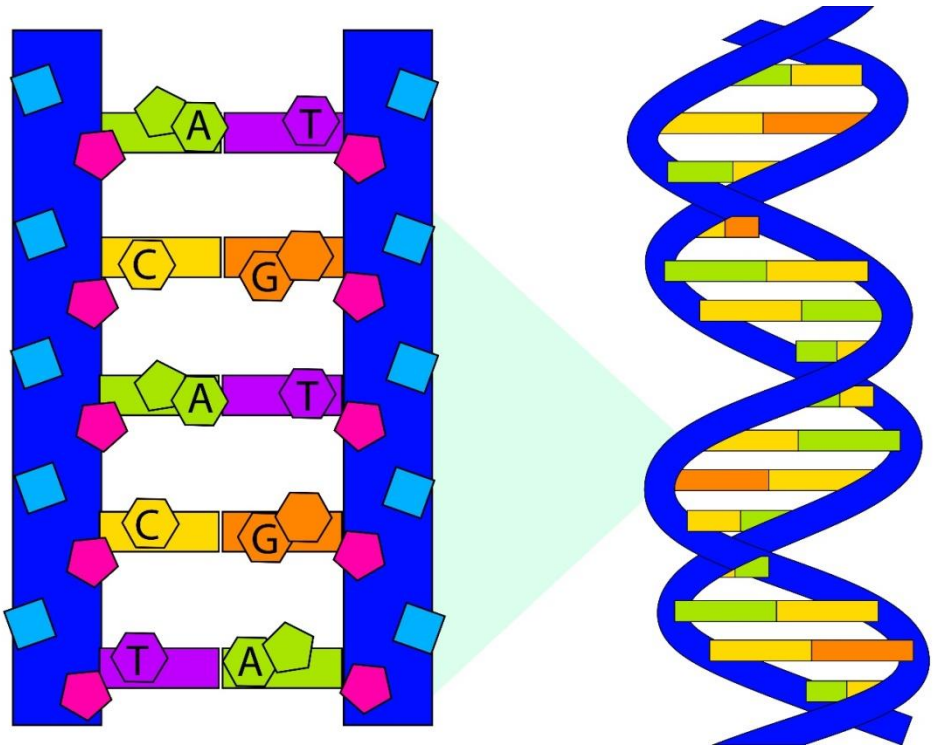


Figure 1.1. Structure of DNA double helix

DNA consists of two antiparallel complimentary strands, made of sugars (pink) and phosphates (light blue) twisted around each other and connected with nitrogen bases (A, T, C, G) (Watson and Crick, 1953).

Following to this discovery, significant number of studies described how genes encode proteins. This process has two major steps. Firstly, the genetic information in DNA is transferred into a messenger RNA (mRNA) through a process known as transcription. In the second step, the produced mRNA, which is a copy of the gene, is translated into a protein molecule (Crick, 1970). Messenger RNA (mRNA) contains a sequence of bases. A set of three bases called a 'codon' encodes for one amino acid and the next three bases code other amino acids and so on; the codons do not overlap (Crick *et al*., 1961).

1.2. Genome replication

Genome replication is a fundamental process occurring in all organisms to copy their DNA before each cell division. The semi-conservative nature of DNA replication, a process in which each parental strand serves as a template for the synthesis of a new complementary daughter strand (Figure 1.2) was first suggested by Watson and Crick and subsequently confirmed by Meselson and Stahl (Meselson and Stahl, 1958). The two strands of DNA helix align in an anti-parallel fashion. During DNA replication, all of double-stranded (ds) DNA is duplicated to generate a second identical DNA double helix and it must do so faithfully to prevent any genome instability and tumorigenesis. DNA replication *in vivo* is a much-regulated process involving a number of replication factors. This process occurs in three stages: initiation, elongation and termination.

The initiation of DNA replication is a three-step process: (i) Loading of initiator proteins to the specific sites of origin of replication (*ori*); (ii) recruitment of other replication components by initiator proteins for fork assembly; (iii) unwinding the origin by DNA helicase (Baker and Bell, 1998; Mackiewicz *et al*., 2004). Bacterial DNA replication initiates by DnaA initiator protein. In bacteria, DNA replication initiation of *Escherichia coli* (*E. coli*) is the best characterized.

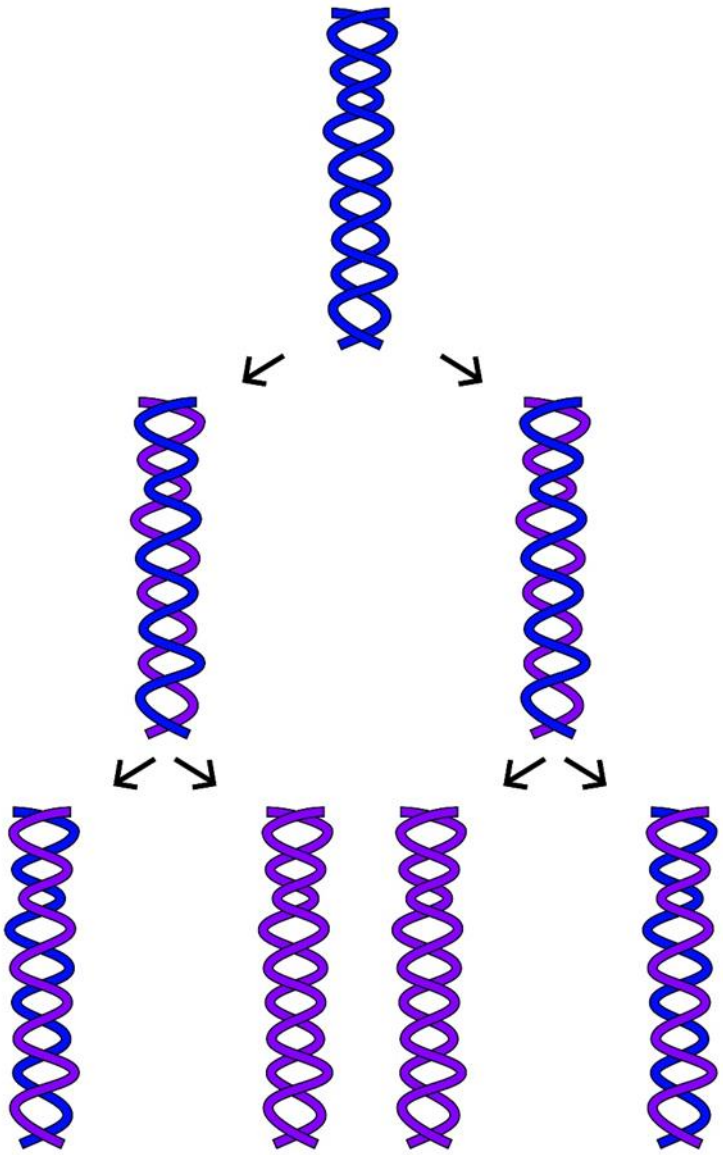


Figure 1.2. Schematic illustration of semiconservative DNA replication

The parental DNA double helix (blue) is separated, creating two single strands. Each strand acts as a template for the complementary strand (purple) to produce two progeny DNA molecules (blue and purple) (Meselson and stahl, 1958).

Bacteria which have a single, circular chromosome possess a single origin of replication, OriC, consisting of an A-T enriched region and an array of several DnaA boxes (Messer, 2002; Mackiewicz *et al*., 2004). In most bacteria, the DnaA protein recognizes a cluster of four DnaA boxes. Once the DnaA binds to the replication recognition sequence, the A-T rich region is melted, resulting in ssDNA bubble onto which the replicative helicase, DnaB in the case of *E. coli* loads (Kaguni, 2006). Following recruitment of proteins required for elongation of DNA, the multiprotein complex of replisome machinery is assembled on the parental DNA template (Messer, 2002). Bacterial helicase moves along the strand onto which it is bound in 5'→3' direction and unwinds the dsDNA. Following unwinding, bacterial single subunit primase, DnaG, binds to DnaB helicase and form primosome which moves away from the origin and makes an RNA primer that starts synthesis of first DNA strand (Alberts *et al*., 2002). A specialized enzyme called DNA polymerase, which moves along the fork in a 5' to 3' direction, carries out the elongation step. Interestingly, DNA polymerase increases the unwinding activity of the DNA helicase more than 10-fold (Kornberg and Baker, 1987). Due to the 5' to 3' direction of DNA polymerase, replication of one of the strands, called leading, is continuous while the other strand, termed as lagging, is synthesised discontinuously in segments (Okazaki fragments) that are 1000-2000 base pairs in length (Okazaki *et al*., 1968). A set of proteins is required to remove RNA primers, synthesis new DNA across the resulting gap, and DNA ligation (Baker and Bell, 1998).

1.3. DNA damage

Maintenance of genome integrity and stability is fundamental for any form of life. This is complicated as DNA is highly reactive and always under attack from endogenous and exogenous environmental agents, which can lead to ~ 50,000-100,000 different lesions in the DNA. If unrepaired, this can lead to mutations and disease (Hübscher and Maga, 2011). The current section will introduce commonest causes of DNA damage and in the next section (1.4), the ways in which cells deal with DNA damage will be discussed.

1.3.1. Ionizing radiation-induced DNA damage

For the first time, in 1927 Muller proved that ionizing radiation causes genetic mutations in living organisms, using *Drosophila melanogaster* as a model. Muller proposed that mutation rates were significantly higher in X-irradiated Drosophila, comparing to unirradiated flies (Muller, 1927). Ionizing radiation is a radiation with enough kinetic energy to cause ionization (loss of an electron) in atoms or molecules. However, not all electromagnetic (EM) radiation is ionizing. Among all electromagnetic (EM) radiations, only the high frequency portion of the electromagnetic spectrum including, gamma rays and X-rays are ionizing. Ionizing radiation is capable of producing reactive oxygen species (ROS), including peroxide, superoxide, and hydroxyl radical through interacting with intracellular molecules (oxygen and water). ROS can result in significant damage DNA. Radiochemical damage can occur by either direct action or indirect action. Alpha particles, beta particles or X-rays can directly break one or both of the sugar phosphate backbones.

Ionizing radiation leads to variety of damages to individual DNA bases including, single-stranded DNA breaks (SSBs) and double-stranded DNA breaks (DSBs). The most severe generated lesions are DSBs that are mainly repaired via either homologous recombination (HR) or non-homologous end-joining (NHEJ) pathways (Olive, 1998).

1.3.2. UV-induced DNA damage

DNA is very vulnerable to UV-induced damage in all forms of life. UV radiation induces two major classes of mutagenic DNA lesions known as cyclobutane pyrimidine dimers (CPDs) and pyrimidine (6-4) pyrimidone photoproducts ((6-4) PPs), and their Dewar valence isomers. Photoisomerization of 6-4PPs by wavelengths longer than 290 nm results in formation of Dewar valence isomers (Figure 1.3). Although CPDs are the most abundant lesions after UV irradiation, 6-4PPs may also have lethal and mutagenic effects. Translocation of an OH group at the carbon at the 4’ position of the 3’-nucleobase to the carbon in position 6’ of the 5’ nucleobase leads to generation of 6-4PPs.

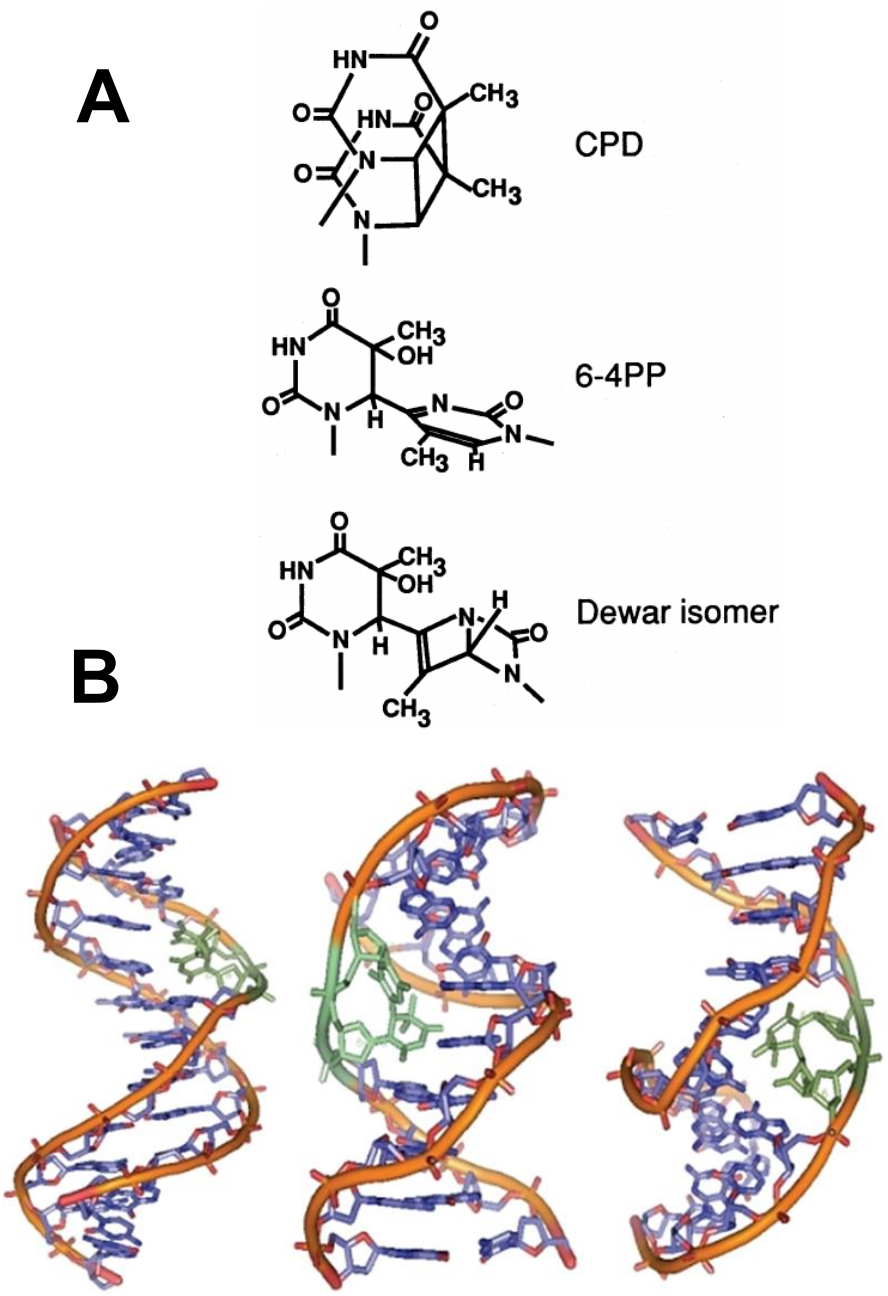


Figure 1.3. Structure of DNA lesions induced by UV light

(A) Two major types of DNA damage are developed as a result of absorption of UV light (CPD and (6-4) PPs) through irradiation of two adjacent (thymine) pyrimidines. (6-4)PPs can subsequently be converted into the Dewar valence isomer (panel A taken from Kobayashi *et al*., 2001) **(B)** Structures of DNA duplexes showing the presence of UV lesions (in green) of CPD, (6-4) PP, and (6-4) PP Dewar isomer (panel B taken from Rastogi *et al*., 2010).

It was proposed that all members of (6-4) photoproducts class are converted to their Dewar valence isomers upon exposure to UV light (Taylor *et al*., 1990). Both lesions can distort the DNA helix by inducing a kink or bend of 44° and 7-9°, respectively.

CPDs can form in ssDNA and at flexible ends of poly (dA)-(dT) tracts. On the other hand, CPD formation is limited in the rigid centre of poly (dA)-(dT) and also when there is bending of the DNA towards the minor groove (Becker and Wang., 1989; Lyamichev, 1991). Different studies on various organisms have been postulated that CPD is able to block DNA replication by inhibiting the activity of DNA polymerases. In addition, the transcription activity of RNA polymerase II from mammalian is eliminated in the presence of both CPDs and 6-4PPs (Mitchell *et al*., 1989; Protic-Sabljic and Kraemer, 1986). Therefore, DNA replication and transcription can be affected by UV-induced lesions and if not repaired, will result in cell death

1.3.3. Oxidative DNA damage

There has to be a balance between oxidants and antioxidants in cells and any disruption in this balance will results in oxidative stress. Reactive oxygen species (ROS) is known as the main source of oxidative stress in all living organisms. Among the characterized oxidative DNA modifications in mammalian genetic makeup, guanine (G) is believed to be the most vulnerable base because of its low redox potential (Halliwell and Aruoma., 1991). When the guanine (G) base is oxidised, it modified to 8-oxo-2'-deoxyguanosine (8-oxo-G) (Figure 1.4). 8-oxo-G is able to functionally mimic thymine in its *syn* conformation, which is the most mutagenic conformation of this lesion. It has been shown that formation of 8-oxo-G during DNA replication results in generation of double-strand breaks (DSBs) (Cheng *et al*., 1992).

The DNA double helix structure consists of a π-stack array that is formed by the heterocyclic base pairs. Since this structure is suitable medium for the migration of charge across long distances, it is suggested that oxidation of guanine to 8-oxo-G occurs up to 37 Å away from the site of reactive species, thereby, oxidative damage does not depend on distance (Hall *et al*., 1996).

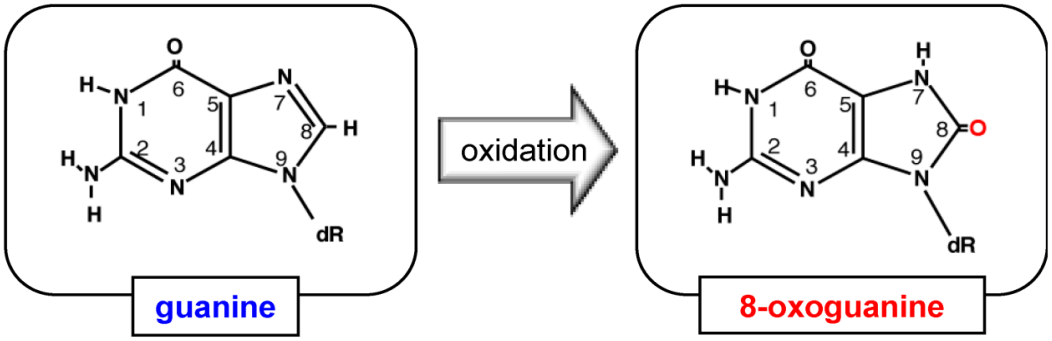


Figure 1.4. Oxidation of guanine by reactive oxygen species (ROS).
Reactive oxygen species like superoxide molecules can form 8-oxoguanine (8oxoG) from guanine (Gerald *et al*.,2004). Oxidative phosphorylation in mitochondria or ionising radiation produce superoxide radicals in the cell. If superoxide radicals metabolised to hydrogen peroxide, they might be used to oxidise iron molecules and results in the release of hydroxide ions and hydroxyl radicals. These hydroxyl radicals react with guanine to produce 8-oxoguanine (Taken from Nakabeppu *et al*., 2014).

Notably, oxidative of guanine to 8-oxo-G does not happen randomly. It is experimentally proven that between the two adjacent guanine moieties, 5' guanine is the most easily oxidized. There is a lot of evidence that 8-oxoG is responsible for G–C to T–A transversions, however, it is not responsible for G–C to C–G transversions as guanine is not inserted opposite 8-oxoG. However, there is some evidence to suggest that other oxidative lesion products of guanine including, imidazolone (Iz), guanidinohydantoin (Gh) and spiroiminodihydantoin (Sp) can cause G–C to C–G transversions. Considering difficulty of guanine incorporation, it is suggested that incorporation of guanine in the G–C to C–G transversions is achieved by specific DNA polymerases through formation of hydrogen bond (Kino and Sugiyama, 2001; 2005).

1.3.4. Alkylating agents and their effects on DNA

The alkylating agent bis (2-chloroethyl) sulphide (mustard gas) is another chemical agent that can induce DNA damage. Following the use of this chemical agent as a poison gas during World War I in 1917, the toxic, mutagenic and carcinogenic effects of this agent were recognised (Haddow, 1973; Philips, 1950). Nitrogen analogues of this compound showed more significant effects on DNA as it could be absorbed through the skin more quickly. The biological effects of this alkylating agent and its eighteen different analogues were studies by Koller in 1958. This study postulated that these lesions could inhibit cell growth either temporary or permanent (Koller, 1958). In addition, they evolved cytotoxic effects including, mitotic suppression; chromosomal breaks and bridges, and excessive fragmentation with loss of relationship to the spindle apparatus. These molecules also cross-link DNA (Rink *et al*., 1993; Rosenberg *et al*., 1969).

Alkylating agents generate different types of adducts based on their nucleophilicity. Alkylating agents with high nucleophilicity will undergo bimolecular S_N2 substitutions, like Methyl methanesulfonate (MMS). On the other hand, those with low nucleophilicity will undergo unimolecular (S_N1) substitutions, like N-ethyl-N-nitrosurea (ENU) (Hoffmann, 1980). Those adducts generated by high nucleophilicity are located at the nitrogen's position 7 of guanine and position 3 of adenine which leads to development of methylated adducts N7-methylguanine (7meG) and N3-methyladenine (3meA) respectively. Around 60-80% of methylated adducts are 7meG and approximately 20% of them are 3meA.

7meG is not cytotoxic; however, it is vulnerable to the formation of apurinic/apyrimidinic (AP) sites (section.1.3.5). In contrast, 3meA is highly cytotoxic since it can stop the DNA polymerase activity (Drabløs *et al*., 2004). S_N1 substitutions are not as abundant as S_N2. They can generate adduct lesions at position 7 and 6 of guanine (Hoffmann, 1980).

1.3.5. Abasic sites in DNA

One of the most common DNA damages in DNA is Apurinic/apyrimidinic (AP) sites. AP sites are formed as a consequence of spontaneous hydrolysis of the *N*-glycosylic bond (Nakamura *et al*., 1998). Notably, apurinic sites are more common than apyrimidinic sites. Studies on mammalian indicated that 10,000 bases are lost per cell per day (Lindahl and Nyberg., 1974). In addition, excision of damaged or inappropriate bases by DNA *N*-glycosylases results in formation of AP sites (Krokan *et al*., 1997). Genetic studies on yeast cells indicated that incorporation of dUTP instead of dTTP during DNA synthesis, leads to formation of the most spontaneous AP sites. Such that when the uracil is removed by Uracil DNA glycolase (Ung1), a potentially toxic AP site is created. If the Ung1-induced AP site lesion is not repaired, it will be bypassed by an error-prone Polζ DNA polymerase.

AP sites are known as cytotoxic and mutagenic since they stall DNA replication, transcription, and yield single base-pair substitution. Furthermore, cleavage of AP sites by AP endonucleases or AP lyases generates DNA single-strand breaks (SSBs) with 5'- or 3'-blocked ends, respectively (Demple and Harrison, 1994; Krokan *et al*.,1997). Since AP sites are mostly found as intermediates in base excision repair (BER) (see section 1.4.2.), this machinery often repairs them.

1.3.6. Cytosine deamination

A growing body of evidence indicate that cytosine is especially prone to deamination. Deamination of cytosine occurs by removal of exocyclic amino group and converting it into uracil. Uracil is normally found in RNA but not in DNA, therefore it can be detected and removed by DNA repair mechanisms. If the deamination remains unrepaired, it can base pair with adenine during DNA replication. As a consequence of this adenine base pairing, deamination of cytosine leads to C to T transitions (where G is replaced by A on the other strand

of DNA) mutation (Figure 1.5) (Duncan and Miller, 1980). Uracil can also found in DNA as a result of dUMP misincorporation in the place of dTMP during replication. It is proposed that the number of cytosine deaminations is around 60-500 per genome per day. Moreover, cytosine can also be deaminated by enzymatic activity. (Cytosine-5)-methyltransferase (MT) is able to deaminate cytosine when the content of S-adenosylmethionine (SAM) within the cell is low, or when the concentration of the MT is high. Uracil DNA glycolase can recognise and remove the inappropriate uracil and leaves an abasic site (AP) in the DNA (Krokan *et al*., 2002).

1.4. DNA repair pathways

Some vital processes in the cell, such as DNA replication and transcription can be interfered by divers DNA lesions. Therefore, damaged DNA must be repaired to ensure cell viability. To prevent genomic alterations that causes cancer or other diseases, cells have developed a network of DNA repair mechanisms to remove various DNA damages (Bertram, 2000; Hoeijmakers, 2009). Regardless of the type of lesion, cells have developed a highly coordinated cascade of events known as the DNA damage response (DDR). When genomes encounter damaged DNA, the DDR senses the damage and signals cells to activate DNA repair mechanisms (Zhou and Elledge, 2000). DNA repair deficiencies result in number of hereditary diseases such as, Ataxia-telangiectasia, Fanconi anemia, and Xeroderma pigmentosum (McKinnon, 2009). In this section, different types of DNA repair pathways will be introduced (Figure 1.6).

In general, cells have evolved two classes of DNA repair mechanisms to maintain their genomes integrity; (1) direct reversal repair and (2) excision of the damaged bases followed by replacement of the new DNA. Moreover, in rare cases when the damage avoids these repair systems, additional mechanisms are developed by cells to tolerate the DNA lesion and restart the stalled replication fork.

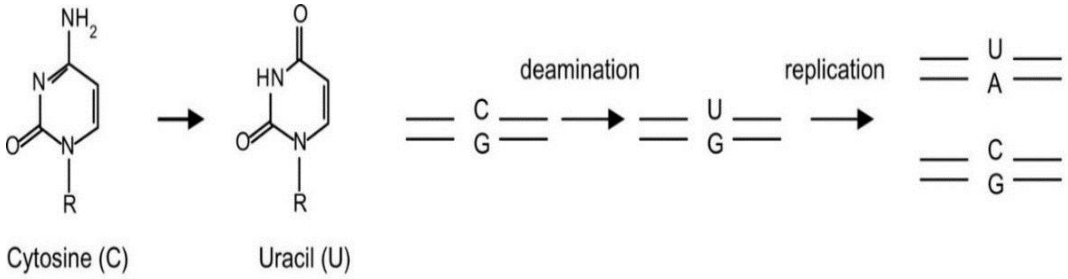


Figure 1.5. Consequence of cytosine deamination to uracil.

Cytosine undergoes spontaneous hydrolysis and is deaminated to a uracil base. The immediate product of this deamination in a double-stranded DNA is U:G mispair and if the uracil bases persist into replication, 50% of the progeny inherit a transition mutation (taken from Connolly, 2009).

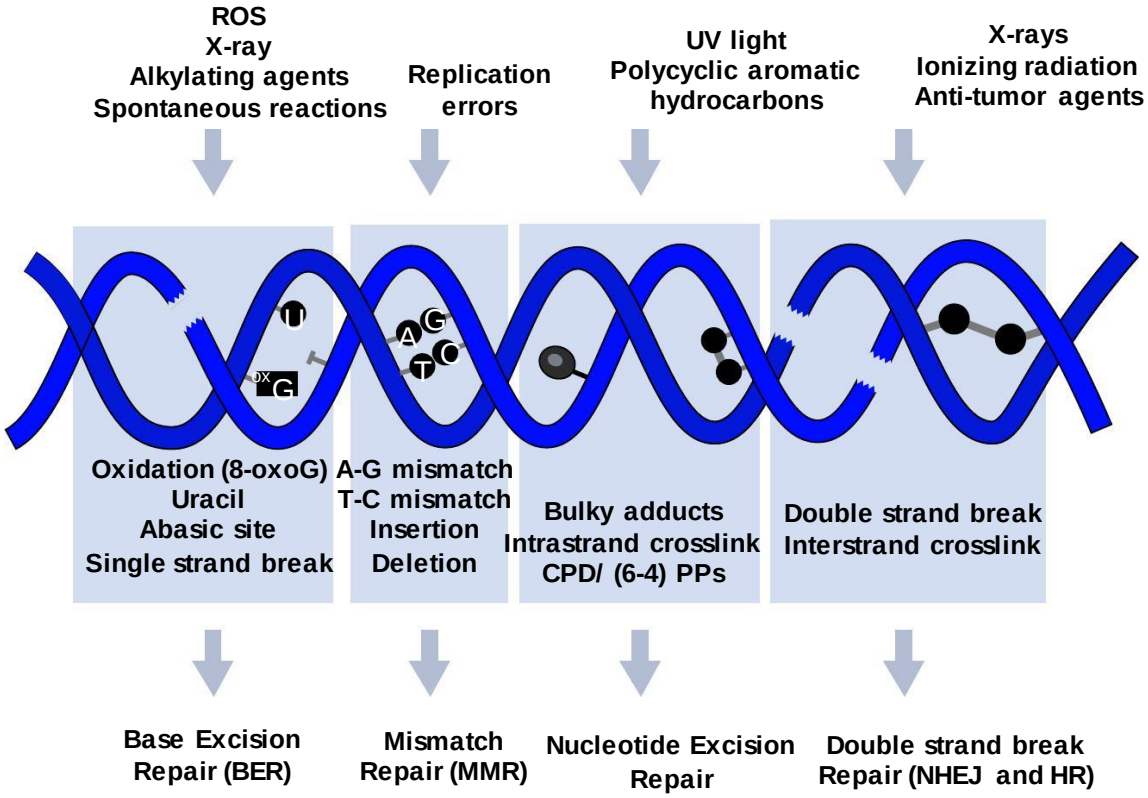


Figure 1.6. DNA repair pathway and mechanism.

Wide variety of exogenous and endogenous agents causes different types of DNA lesions. Several repair mechanisms are used to remove and repair damaged DNA including base excision repair (BER), nucleotide excision repair (NER), mismatch repair (MMR), double-strand break repair and direct reversal. Nonhomologous end-joining (NHEJ) and homologous recombination (HR) are two mechanisms through which double strand breaks are repaired.

1.4.1. Direct reversal repair

Only a few types of DNA damage, like UV induced pyrimidine dimers and alkylated guanine residues, are repaired by direct reversal of the damage. This mechanism is the simplest and the most energy efficient form of DNA repairs. UV induced pyrimidine dimers and methylation of guanine can be reversed directly by specialized enzymes photolyase and O6-methylguanine-DNA-methyltransferase (MGMT), respectively.

To recover UV induced damage of DNA, a light dependent process known as photo-reactivation is used. An enzyme known as photolyase facilitates this process. Photo-reactivation process removes the pyrimidine dimers without altering other nucleotides. This process uses visible light (blue) as source of energy. Photoreactivation has been found in bacteria, archaea and eukaryotes, however, some species such as humans do not possess this DNA repair pathway (Minato and Werbin, 1972; Todo *et al*, 1996).

1.4.2. Excision Repair

Excision repair mechanism is considered to be one of the most efficient way to repair a wide variety of chemical alterations to DNA, in both prokaryotes and eukaryotes. There are three types of excision repair pathways. These are: (1) base excision repair (BER), (2) nucleotide excision repair (NER) (3) mismatch repair (MMR). These pathways exist to recognise and remove the damage, either as free bases or as nucleotides following by replacing them with undamaged counterparts.

Base excision repair (BER) can recognise and remove different abnormal bases such as hypoxanthine, uracil-containing DNA, pyrimidine dimers and bases damaged by oxidation and ionizing radiation (Almeida and Sobol, 2007; Hitomi *et al*., 2007). Notably, deamination of adenine leads to formation of hypoxanthine.

The excision of the uracil, which arises in DNA either by incorporation of dUMP in place of dTMP or deamination of cytosine to uracil, is catalysed by DNA glycosylase. DNA glycosylase cuts the N-glycosylic bond between the base and the deoxyribose of the DNA backbone releasing a uracil base and leaving an AP site in the DNA. (Duncan and Miller, 1980; Krokan *et al*., 2002).

In general, DNA glycosylases are categorized into two groups: (1) monofunctional glycosylases, which only performs glycosylase activity before strand incision by AP endonuclease (APE1) (2) bifunctional glycosylases, which are glycosylase/β lyases these group are able to perform both strand incision and glycosylase activity (Fortini *et al*., 1999; Hitomi *et al*., 2007)

Following DNA glycosylase activity, the resulting AP site is cleaved by an AP endonuclease at the 5' side, resulting in a nick with a 3' hydroxyl and a 5' phosphate group which is then processed by two pathways: the short patch (1-nucleotide gap filling) and the long patch (2–6 nucleotide resynthesize) BER. During short patch pathway, DNA polymerase (pol) β displaces the AP site, resynthesize DNA and fill in the gap. The process is followed by DNA ligase III activity that forms a phosphodiester bond and completes the repair pathway. The long patch pathway, requires polymerases δ, ε, or β, combined with PCNA, flap structure-specific endonuclease 1 (FEN1), and DNA ligase I. Pol β-PCNA through strand displacement and polymerase activities displaces a DNA flap of up to 13 bases in length which then removed by FEN1 and finally DNA ligase I ligates the DNA. (Fortini *et al*., 1998; Gary, 1999; Prasad., 2001; Prasad *et al*., 2000; Stucki *et al*., 1998).

Nucleotide excision repair (NER) is a mechanism to recognize and repair bulky DNA adducts, such as those induced by UV lesions (CPD and (6-4) PPs), crosslinking lesions like cisplatin. NER removes the damaged bases as part of an oligonucleotide patch containing the lesion. In common with BER, several enzymes are required to fulfil NER pathway (Hess *et al*., 1997; Kuraoka *et al*., 2000; Reardon *et al*., 1999).

Nucleotide excision repair is more complicated in eukaryotes than in prokaryotes. In prokaryotes, the products of three genes facilitate NER (UvrA, UvrB, and UvrC). UvrA initially detects a damaged DNA and recruits UvrB to the lesion. UvrB separates the two DNA strands to verify the position of the lesion. The reaction is followed by releasing of UvrA and recruiting UvrC. UvrC contains two nuclease domains that incise the 3' and 5' sides of the damaged site and excise an oligonucleotide consisting of 12 or 13 bases. DNA helicase II (also known as UvrD) then comes in, cuts the hydrogen bonds between the complementary bases, and removes the excised segment. The gap created by this process is

filled by DNA polymerase I and DNA ligase (Memisoglu and Samson, 2000; Moolenaar *et al*., 2001).

In mammalian cells, initiation of NER pathway requires XPC-RAD23B initiator protein that binds nondamaged strand opposite the lesion. XPC-RAD23B interacts with TFIIH protein that pries the DNA open with its XPB subunit. XPB is known as helicases that unwind the damaged DNA. The XPD subunit of TFIIH track along DNA until stalls at the site of damage and verifies the chemical modifications of the damage. Stalling of XPD leads to recruitment of XPA, RPA, and XPG. The XPF/ERCC1 complex then cuts in the 5' direction, and XPG cuts in a 3' direction, resulting in formation of a gap measuring 24-32 nucleotides in length (Huang *et al*., 1992; Matsunaga *et al*., 1996; O'Donovan *et al*., 1994). This gap then appears to be filled in by DNA polymerase δ or ε, in association with PCNA and RPA and subsequently, sealed by either DNA ligase I or a complex of XRCC1 and Ligase III (Moser *et al*., 2007; Popanda and Thiemann, 1992; Shivji *et al*., 1992).

DNA polymerases perform proofreading activity to remove most of the mismatch bases (section 1.8.5.) but some mismatch bases can skip the editing step and stay in the genome. To remove the remaining ones, mismatch repair system (MMR) scans the newly synthesised DNA and remove the mismatch base specifically from the newly synthesised DNA. In eukaryotes, in comparison with regular DNA, the mismatched DNA can be detected and bound by MUtS homologue with a higher affinity (Gradia, 2000; Schofield *et al*., 2001). MMR is an ATP-dependent pathway that requires activation of MutS and MutL. Heterodimeric complex of MutS-MutL related proteins initiate MMR (Iyer *et al*., 2006). Mutation of MutS and MutL homologues increase cancer susceptibility in both mice and humans. Since a mismatched base itself contains no signal to activate the MMR, in *E.coli*, the absence of methylation at the restriction site is used to direct the repair to the error-containing strand. MutH, a methylation specific endonuclease, interacts with MutS via MutL and in the presence of ATP nicks the unmethylated strand at the hemimethylated GATC site to introduce an entry point for the excision reaction. After nicking the daughter strand by MutH, DNA helicase unwinds the DNA and exonuclease I excises the DNA

bidirectionally (Acharya *et al*., 2003). Finally, DNA polymerase III fills in the resulting gap and DNA ligase ligates the strand (Iyer *et al*., 2006).

1.4.3. Homologous recombination

Homologous recombination (HR) is a DNA metabolic process that conserved in all living organisms. This process performs high-fidelity template-dependent repair or tolerance of different DNA lesions, such as double-stranded breaks (DSBs) and DNA inter-strand crosslinks. This process is also vital for telomere maintenance and chromosome segregation (Game and Mortimer, 1974; Li and Heyer, 2008). Interestingly, HR competes with NHEJ and translesion DNA synthesis (TLS) pathways in the repair of DSBs and damage tolerance, respectively.

The Mre11, Rad50, and Nbs1 complex (MRN) is known as the key HR associated factors that performs important functions at stalled replication forks. The MRN complex binds to DSBs and initiates resection prior to repair by HR (van den Bosch *et al*., 2003). In the first stages of HR, RecA/Rad51 assembles onto ssDNA, by nucleation of a short tract of RecA protein onto ssDNA and displacement of SSB protein. Assembly of RecA on ssDNA results in formation of helical presynaptic filaments (Sung and Klein, 2006). While ATP binding is required for assembly of presynaptic filaments, ATP hydrolysis is not essential. After binding, presynaptic filaments stretch as much as 50% of the length of the duplex molecule in order to search for homology (Klapstein *et al*., 2004). After finding the homolgy, the single-stranded pre-synaptic filament performs strand invasion to invade the identical recipient duplex DNA and forms a displacement loop (D-loop) between the invading 3’ overhang strand and the homologous chromosome.

After strand invasion, HR occurs by either double-strand break repair (DSBR) or synthesis-dependent strand annealing (SDSA) (Helleday *et al*., 2007; Sung and Klein, 2006). In DSBR, the second 3’ end of DSB can be captured and forms a Holliday junction with the homologous chromosome. Using nicking endonucleases, the Holliday junction structure is resolved into crossover or non-crossover recombination products. In SDSA, a DNA polymerase extends the invaded 3’ strand. The newly synthesis 3’ invading strand is then able to anneal

to the second end of the DSB. This is then followed by gap-filling DNA synthesis and ligation. The SDSA always forms a non-crossover repair product (Helleday *et al*., 2007; Sung and Klein, 2006) (Figure 1.7).

Notably, homologues of Rad51 and Mre11 proteins are present in archaea. It has been shown that in yeast, Rad52 stimulates the binding specificity and strand invasion activity through interaction with Rad51 and RPA (New *et al*., 1998). On the other hand, in higher eukaryotes, tumour suppressor protein (BRCA2) assists in the assembly of Rad51 onto ssDNA (Yang *et al*., 2005). Due to the mediatory function of BRCA2 in HR, mutation of this protein leads to cancer formation.

1.5. DNA damage tolerance

Even though cells develop a variety of repair mechanisms that target and repair a vast array of DNA modifications, it is inevitable some DNA damages cannot be removed by DNA repair pathways therefore will persist in the genome and leads to replication fork stalling. DNA damage tolerance mechanisms have evolved to restart stalled replication forks and achieve the high fidelity for genome duplication. These include translesion DNA synthesis (TLS) mediated by specialized polymerases and error-free recombination-mediated restart and template switching (TS) (Figure 1.8).

Although the molecular mechanisms of TS are not fully understood, studies in yeast and mammalian cells suggested that repriming is able to restart the stalled replication fork and leaves gaps opposite the lesions. These gaps are then filled-in using newly synthesised strand as template. Template switching can be promoted by the RecQ helicase Sgs1 (BLM in mammalian cells) or Werner syndrome ATP-dependent helicases (WRN) (Machwe *et al*., 2006). BLM, WRN, and RAD51 are able to facilitate fork regression into a holiday junction. When the chicken foot Holiday junction is formed, Rad51 protein allows the double-stranded DNA end of the chicken foot Holiday junction to undergo strand invasion and form a D-loop (Petermann *et al*., 2010). D-loop formation leads to re-loading of the replication machinery. Re-loading of replication machinery results in formation of Holliday junction structure which prevents recombinant proteins being generated. Therefore, this structure requires dissolution (Petermann and Helleday, 2010).

Translesion DNA synthesis (TLS) is a process carried out by specialized DNA polymerases that synthesis short tracts of DNA opposite lesions, therefore making restart of replication possible. Since TLS polymerases perform low fidelity during synthesis of undamaged DNA, it is vital that they are only recruited when they are needed (Sale *et al*., 2012). TLS polymerases include pols η, ι, κ and REV1 from the Y family, pol μ from the X family and pol ζ from the B family of polymerases (Hübscher and Maga, 2011).

In 1996, it was found that, in yeast, the product of *REV1* and *REV3* genes are dCMP transferase and DNA pol ζ, respectively (Nelson *et al*.,1996b). It was also discovered that pol ζ could bypass cyclobutane pyrimidine dimer (CPD) (Nelson *et al*., 1996a). A few years later, it was shown that Rad30 is a *bona fide* DNA polymerase (termed Pol η) that is capable of bypassing *cis-syn* cyclobutane T-T dimers (Johnson *et al*., 1999). By the end of 1999, a human homologue of Rad30 was identified (Masutani *et al*., 1999). Soon after, *E.coli* DinB and UmuD′2C complex were characterized as *bona fide* TLS polymerases called *E.coli* Pol IV and Pol V, respectively (Reuven *et al*., 1999; Wagner *et al*., 1999). Together, these discoveries led to the defining of TLS as an enzymatic process, which is facilitated by conserved specialized polymerases. The mechanism of TLS involves several 'polymerase switching'. Once the replicative polymerase encounters damaged DNA, it stalls at site of damage and must be displace and replace with a TLS polymerase. Subsequently, the TLS polymerase incorporates either correct or incorrect nucleotides opposite a lesion and then either the same or the second TLS polymerase extends from the (mis)incorporated nucleotide. Finally, after bypassing the lesion, the TLS polymerase is quickly replaced by a replicative polymerase to complete genome duplication (Figure 1.9) (Johnson *et al*., 2000; Shachar *et al*., 2009). TLS can either be regulated during replication with the non-catalytic activity of REV1, or post-replicatively through monoubiquitination of PCNA.

The ability of REV1 to interact with all TLS polymerases, suggests, that REV1 acts as a scaffold for TLS polymerase and allows TLS polymerase switching

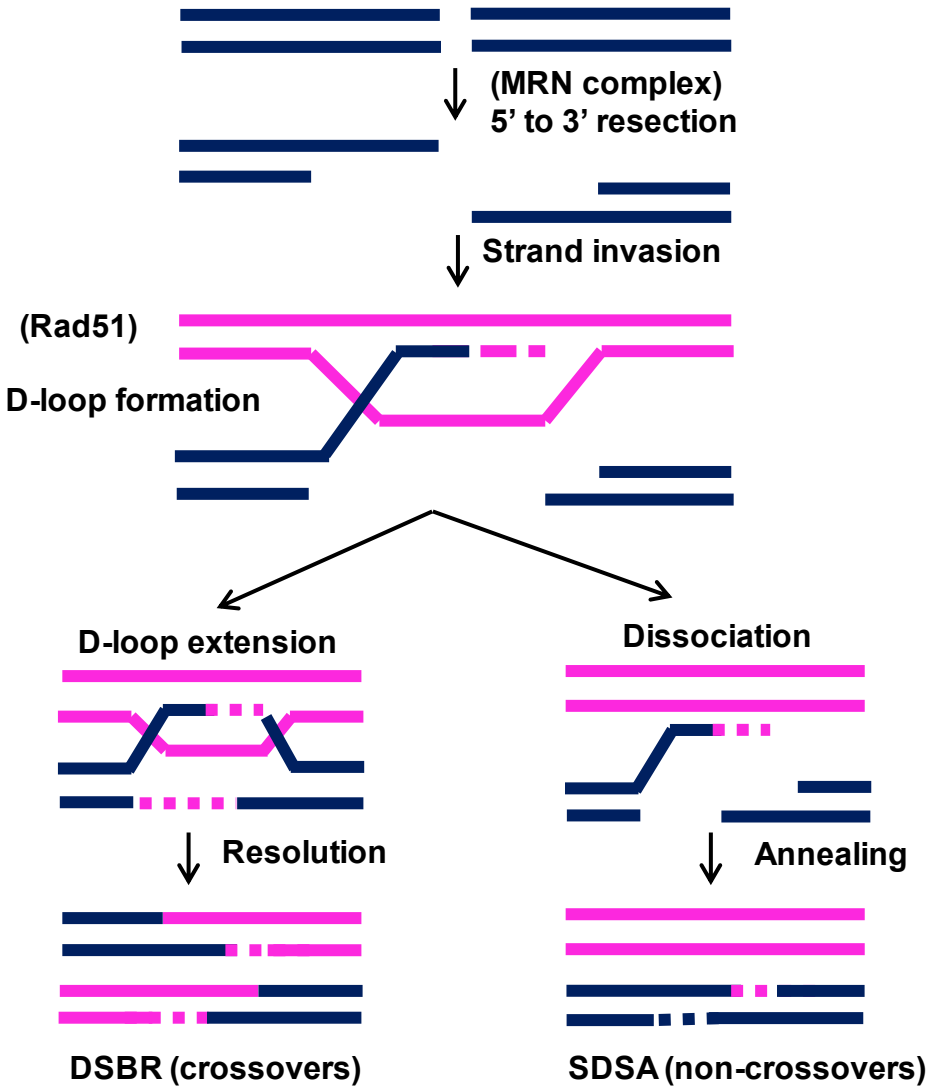


Figure 1. 7. Homologous recombination pathways

Repair of double strand breaks by homologous recombination initiates through generation of 3' single-stranded DNA (ssDNA) tails by MRN complex. Subsequently a Rad51 binds to the ssDNA and promotes invasion of the DNA ends into the homologous duplex DNA which leads to formation of D-loop structure. The D-loop is then used as both a template and primer for DNA synthesis. HR continues either through the DSBR mechanism (left), leading to crossovers or through the SDSA mechanism (right), which results in non-crossovers.

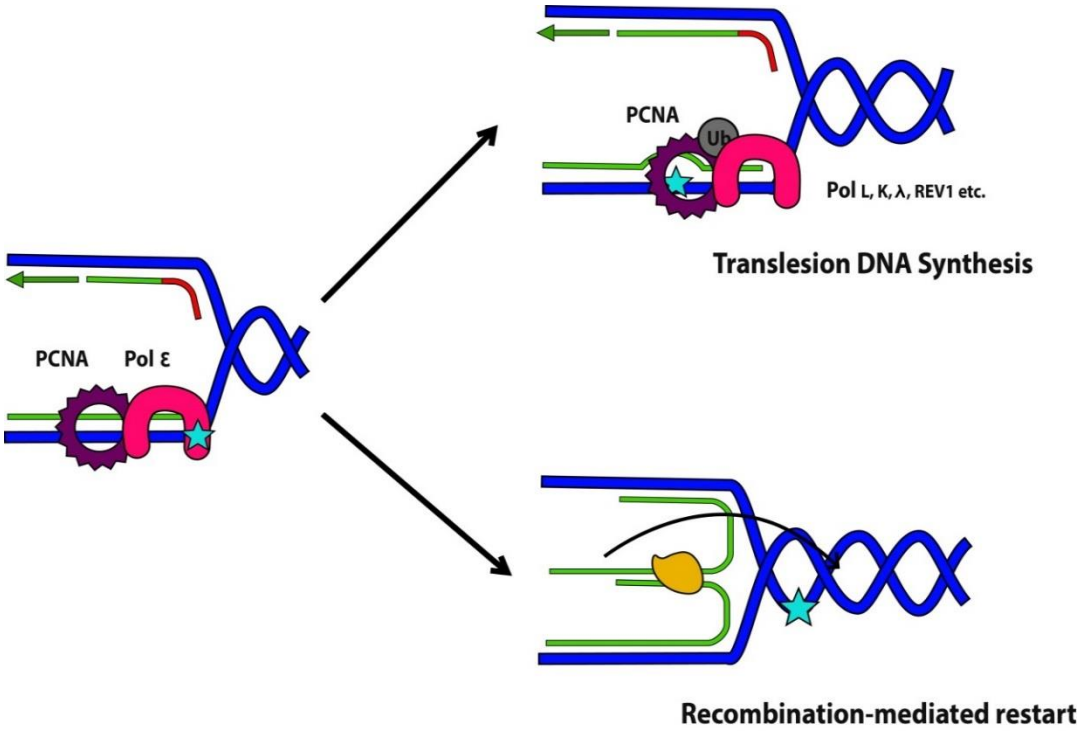


Figure 1.8. DNA damage tolerance mechanism

Various repairing mechanisms have evolved to tolerate damage during replication. Including, translesion DNA synthesis (TLS) catalysed by TLS DNA polymerases that synthesize short tracts of DNA opposite lesions and error-free bypass mechanisms, mediated by homologous recombination, utilize an alternative undamaged template to restart the stalled replisome.

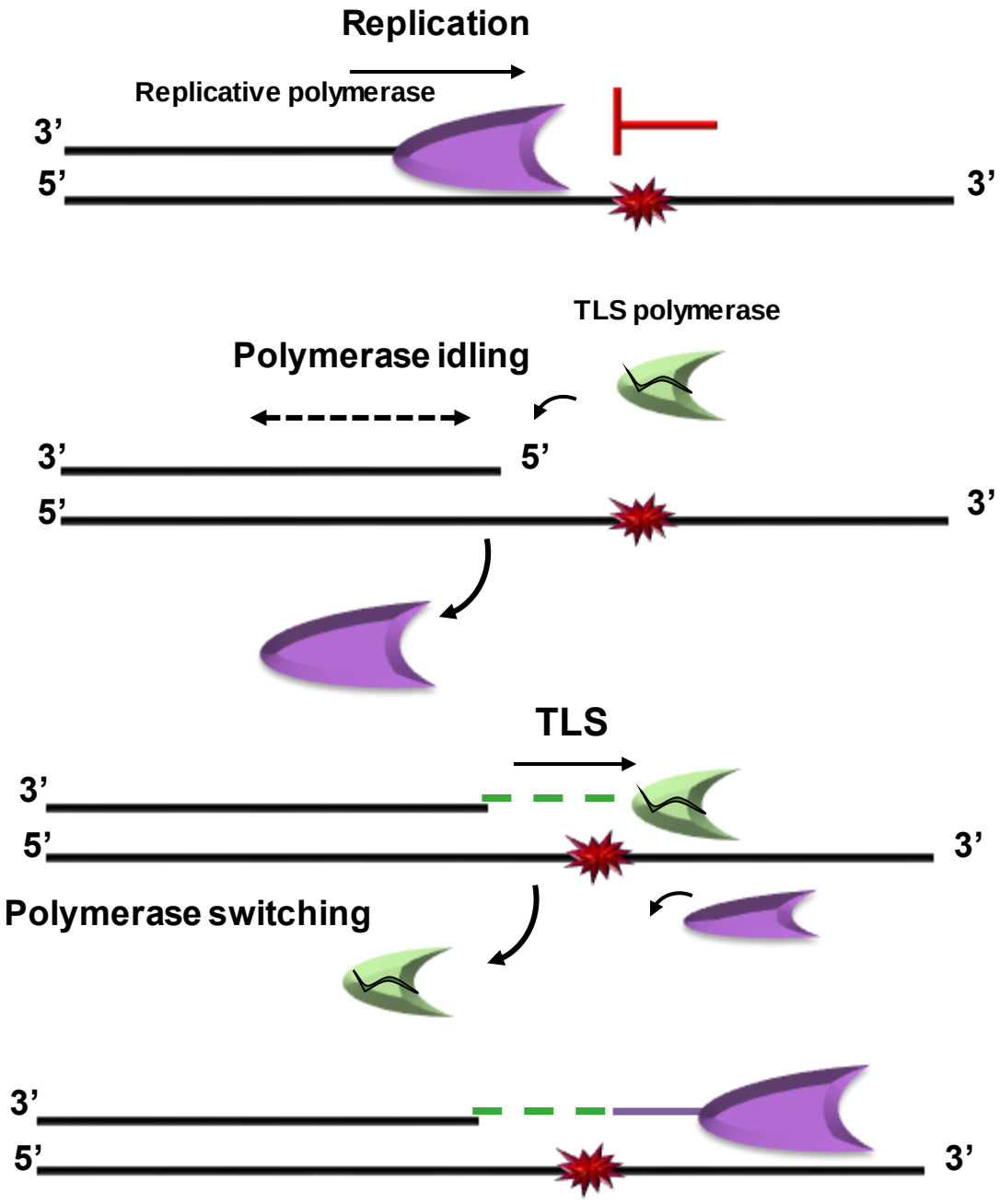


Figure 1.9. The basic mechanism of translesion DNA synthesis

Replication continues until it encounters the damage on the ssDNA. The replicative DNA polymerase stalls at the site of damage , following by polymerase idling. TLS DNA polymerase replicates over the lesion and before further synthesis replaced with a replicative polymerase to complete genome replication.

(Friedberg *et al*., 2005). PCNA mono-ubiquitylation plays a key role in TLS regulation by recruitment of the TLS polymerases to the site of damage when the replicative polymerases stall and leave ssDNA gaps.

All eukaryotic TLS polymerases contain specialised ubiquitin-binding motifs (UBM, UBZ), which stimulate their affinity for ubiquitinated PCNA (Lehmann *et al*., 2007). The existing theory for which polymerase is recruited to specific lesion is that the selection depends on the lesion itself, which means, if the polymerase can accommodate the damaged template-primer in its active site, it is able to complete polymerisation before dissociating, otherwise, another polymerase is then recruited to bypass the damage (Sale *et al*., 2012).

In bacteria, the product of *E.coli dinB,* DNA polymerase IV, a member of Y family DNA polymerase is able to incorporate nucleotides opposite different DNA lesions through TLS. This enzyme can bypass various DNA damages such as those induced by nitrofurazone (NFZ), benzo(a)pyrene, alkylating agents and oxidative damage. Similar to other Y family polymerases, Pol IV interacts with the homo-dimeric β-clamp through two distinct modes. First mode (CTP,) relies on the interaction of a C-terminal peptide of Pol IV with a hydrophobic pocket on the surface of β. Second mode (LF), relies on the interaction of the little finger of Pol IV and the edge of the β ring at the dimer interface (Bunting *et al*., 2003). It has been shown that the Pol IV LF-β interaction is required for Pol IV replicative activity. However, unlike CTP-β, LF-β interaction is not required for translesion synthesis activity of Pol IV (Wagner *et al*., 2009).

1.6. Bacterial DNA replication

Understanding of DNA replication comes largely from studies of *E.coli*. DNA replication in bacteria is a bidirectional and semi-conservative process and begins at specific sequences called origins of replication (*oriC*). Remarkably, bacterial chromosome has single origin of replication. In general, DNA replication is a three-step process: initiation, elongation, and termination. In bacteria, replication initiates through binding of initiation factor (DnaA) to the origin of replication which promotes the unwinding of DNA at *oriC*. The DnaB helicase unwinds origin of replication and extend the ssDNA for copying (Mott and Berger, 2007). As the duplex DNA is unwound by helicase, single strand binding proteins (SSB) cover

the single stranded regions to stabilize and protect ssDNA from possible threats including, nuclease attack and chemical modifications (Iftode *et al*., 2008; Wold, 1997). Subsequently, the DNA primase binds to the DNA and initiates synthesise of leading strand through synthesise of short RNA segments called RNA primers. During elongation step, DNA polymerase III extends the RNA primer to start synthesise of new DNA strand in the 5' to 3' direction. Specifically, DNA polymerase III holoenzyme consists of a β-subunit clamp which ensures that the polymerase attaches on the DNA. Bacterial beta clamp, a processivity-promoting factor, is assembled around the DNA by ATP hydrolysis and gamma subunit (Bloom, 2009). Following assembly around the DNA, the affinity of beta subunit for gamma subunit is replaced by an affinity for alpha and epsilon subunits, leading to formation of the complete holoenzyme. Once the β-subunit dimer tightly binds the DNA, it functions as a ''clamp'' that slide along the DNA and improves the processivity of polymerase. On the lagging strand, the free ssDNA is copied into short RNA primers by DNA primase. Pol III binds the 3'-OH group of each primer and elongates them through addition of deoxyribonucleotides. The produced short fragments are called Okazaki fragments. Subsequently, the RNA primers of the adjacent fragments are removed by Pol I which unlike Pol III has 5' to 3' exonuclease activity. The Pol I then fills in the gap between Okazaki fragments. Finally, DNA segments are joint together by another critical enzyme called DNA ligase (Garg *et al*., 2004; Lehman, 1974).

1.7. Eukaryotic DNA replication

Compared with bacteria, eukaryotic chromosomes are much larger (up to a billion bp) and linear and if eukaryotic DNA replication initiates at only one origin of replication (*oriC*) then it would take weeks to fully replicate a chromosome. Therefore, eukaryotic chromosomes have multiple origins of replication to replicate the entire genome significantly faster. In mammalian cells, the number of these specific sites is around 30,000-50,000(Cairns, 1966; Huberman and Riggs, 1968). Replication of DNA in eukaryote is performed using a similar process of bacterial replication. It starts with recognition of origin of replication by the eukaryotic initiator Origin Recognition Complex (ORC) and continues with the assembly of pre-replication complex (pre-RC) (Rao and Johnson, 1970), and the activation of pre-RC. Binding of initiator protein results in partial untwisting of the

duplex DNA following by further unwinding in the presence of helicase (Figure 1.10). DNA replication licensing, the first in DNA replication, occurs by the sequential recruitment of ORCs, cell division cycle 6 (Cdc6), CDC10 Target 1 (CDT1), to the origins at the late M and G1 phases. These two factors in concert with MCM9 can load the minichromosome maintenance (MCM) helicase complex MCM2-7 in an ATP-dependent manner to replication origins (Bell and Dutta 2002; Remus *et al*., 2009). Upon entry into S phase other replisome components required for activation of the replication origin including, Cdc45 and GINS are loaded to the replication origins. This loading is dependent on Dpb11, Sld2 and Sld3. These proteins are regulated by two kinases known as cyclin-dependent kinase (CDK) and Dfb4-dependent kinase (DDK) which phosphorylate MCM and Cdc6, respectively. The phosphorylated MCM2-7 helicase together with Cdc45 and GINS forms the CMG complex. To facilitate origin unwinding, a marked remodelling of CMG complex after loading onto ds DNA is required (Huang and Zhang, 2011). In order to prevent DNA re-replication within one cell cycle, geminin binds to Cdt1 and blocks binding of MCM to Cdt1 (Wohlschlegel *et al*.,2000). DNA unwinding by MCM2-7 leads to formation of the replisome at the bubble. Two steps facilitate synthesis of new DNA during S phase, synthesis of a short RNA primer by a DNA primase (PriS/PriL) and extension of the primer by specialized enzymes called DNA polymerases.

In eukaryotic cells the three main DNA polymerases involved in DNA replication are DNA polymerases α, δ and ε,which are able to synthesise DNA strands in 5'⟶3' direction. Replicative primases typically form a heterodimeric complex consisting of a small catalytic subunit (Prim1/PriS) and a large regulatory subunit (PriL). This heterodimer together with DNA Pol α subunits (A and B) forms a complex in eukaryotes which can initiate DNA replication. Following synthesis of short RNA primer of 10-15 nucleotides in length by primase, Pol α elongates the primer. The longer DNA primer produced by Pol α is subsequently extended by Pol δ and ε (Burgers, 1998; Frick and Richardson, 2001).

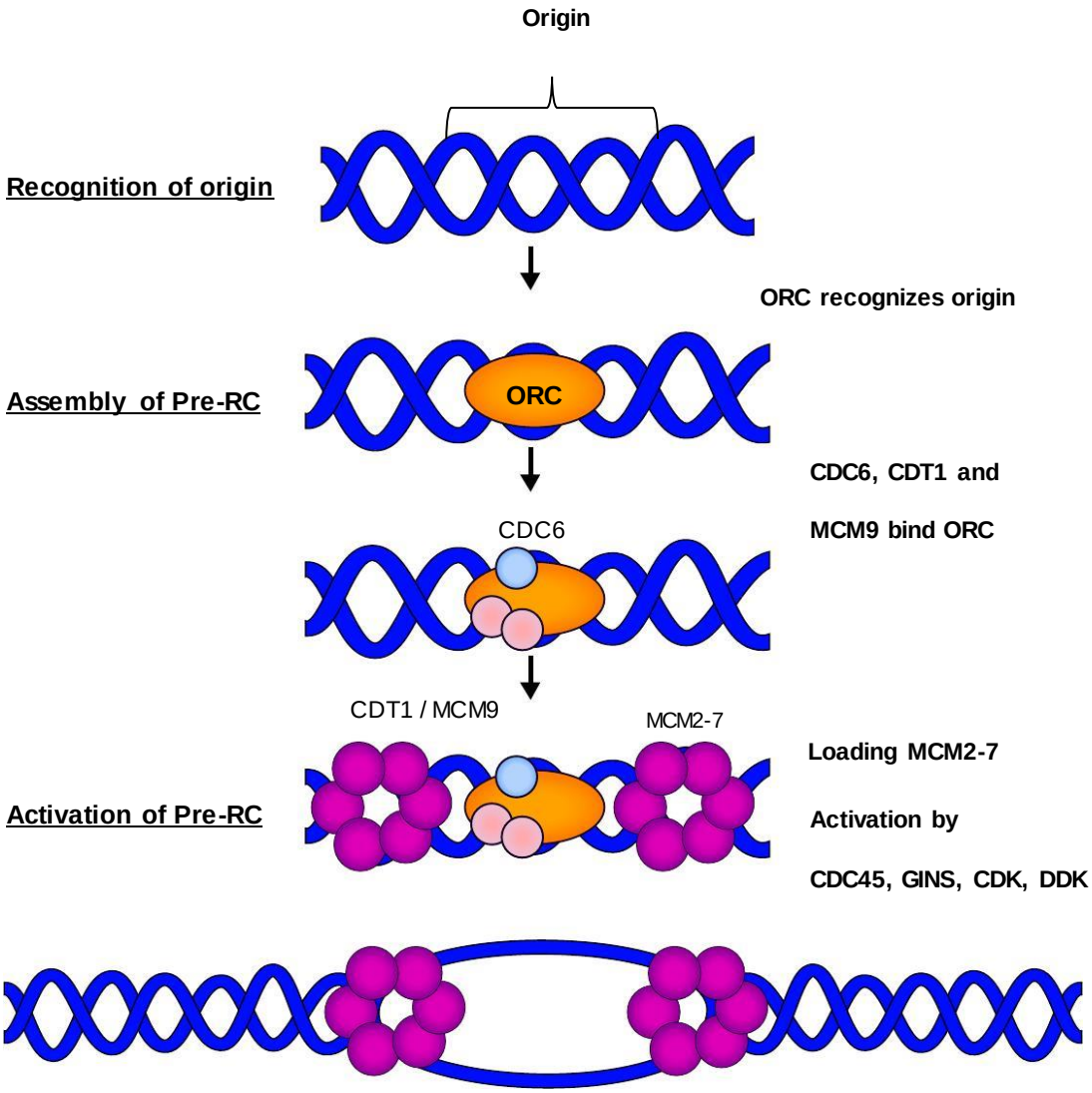


Figure 1.10. Assembly of eukaryotic pre-replication complex at replication origins

Origin recognition complex (ORC) recognises replication origins and recruits CDC6, CDT1 to the origins. These two factors in concert with MCM9 can load the MCM2-7 in an ATP-dependent manner to replication and form the pre-RC complex. The phosphorylated MCM2-7 together with Cdc45 and GINS form the CMG complex. Activation of pre-RC requires CDC45, GINS, CDK, DDK and other factors.

Pol δ and ε are two processive polymerases with 3'→5' exonuclease activity which are used for high fidelity replication of the DNA (M. Simon *et al*., 1991). Different studies suggesting that MCM10, a conserved eukaryotic protein essential for chromosomal replication, in budding yeast is required for loading and stability of pol α and pol δ (Rick and Bielinsky, 2004; 2006). A recent study proposed the importance of MCM10 in origin unwinding and in functioning of CMG complex (Watase *et al*., 2012). Cdc45 is also required for recruitment of pol α to the replication fork (Aparicio *et al*., 1999; Kukimoto *et al*., 2001). During elongation, the primase (PriS/ Prim 1) subunit of the Pol α complex synthesises short RNA primer of 10-15 nucleotides in length that prime DNA synthesis and extends from this about 20 nucleotides of DNA, from which other polymerases can extend (Conaway and Lehman, 1982).

Replication protein A (RPA), a heterotrimeric ssDNA-binding protein, is required during DNA replication to stabilize unwound single stranded conformations. This enzyme also prevents the formation of secondary structures straightening the single stranded DNA regions in lagging strand (Iftode *et al*., 2008; Wold, 1997). Other essential proteins during replication include proliferating cell nuclear antigen (PCNA), replication factor C (RFC), topoisomerase I and, topoisomerase II. The eukaryotic PCNA sliding clamp is a toroidal-shaped homotrimer which encircles DNA and slides along the duplex while binding to polymerases (Kelman, 1997). The RFC clamp loader is a heteropentameric complex that is responsible to open and close the circular PCNA in an ATP-dependent manner (Jonsson *et al*., 1997; Li and Burgers, 1994) PCNA sliding clamp plays an important role in improving the processivity of DNA polymerases during chain extension.

A model regarding replisome elongation initially suggested by Sugino lab and supported by the Kunkel lab. This model proposes following replication initiation, pol ε synthesises a continuous strand DNA opposite the leading template strand and pol δ produces discontinuous DNA strands opposite the lagging template strand (Kunkel and Burgers, 2008; Stillman, 2008). The synthesis of short segments on the lagging strand is named Okazaki fragments in honour of their discoverer Reiji Okazaki. Recently, a new model has been proposed, which is not compatible with the old model (Kunkel and Burgers. 2008; Stillman, 2008).

This model suggests following elongation, Pol ε is replaced by Pol δ as pol ε moves away from origin of replication and disconnected from the DNA, subsequently replication is continued by Pol δ (Pavlov and Shcherbakova, 2010).

Maturation of Okazaki fragments is a highly coordinated process which requires interaction of Pol δ, FEN1 and DNA ligase I with proliferating cell nuclear antigen. Polymerase δ reaches the previously produced Okazaki fragment. After gap filling, Pol δ is replaced by FEN1 to cleave short flap structures subsequently, DNA ligase I facilitates joining of Okazaki fragments in an ATP-dependent manner (Barnes *et al*., 1990; Garg *et al.*, 2004; Lehman, 1974).

During replication, DNA unwinding and progression of the replication fork generates positive supercoils ahead of the replisome that can causes supercoiling tension in DNA that must be relaxed. Topoisomerases are important molecular machines employed to release this torsional stress. Topoisomerases are classified into two groups based on the number of strands that they can break. Topoisomerase type I subfamily (IA, IB) breaks one strand of a DNA helix. In order to relax the helix DNA, the broken strand is rotated around the intact strand and subsequently the ends of broken strand are released. Topoisomerase type II (IIA, IIB) solves the topological problems associated with DNA replication by breaking and rejoining double-stranded DNA. Topo II unlike topo I is ATP dependent (Champoux, 2001).

1.8. Archaeal replisomes

DNA replication is a conserved process which exists in all domains of life (bacteria, eukaryote and archaea). In 1977, a phylogenetic analysis of ribosomal RNA sequences of a group of prokaryotes demonstrated remarkable differences between this group and bacteria, which led to classify this group as a separate domain: called Archaea (Woese and Fox, 1977). Archaea the third domain of life is superficially similar to bacteria in size and shape. Both bacteria and archaea are unicellular with no nucleus but both have cell walls and use flagella to swim. Despite the similarity of archaea and bacteria in terms of structure, bioinformatics and genetic studies have demonstrated that proteins involved in the very important biological processes, including, DNA replication, transcription, and translation are more similar to their eukaryotic counterparts than bacteria (Table

1.1). Identification of eukaryotic homologues in archaeal replication machinery including ORC (origin recognition complex), Cdc6, GINS (Sld5-Psf1-Psf2-Psf3), MCM (minichromosome maintenance), RPA (replication protein A), PCNA (proliferating cell nuclear antigen), RFC (replication factor C), FEN1 (flap endonuclease 1), DNA primase, DNA polymerase, and DNA ligase suggested that eukarya and archaea share a common ancestor (Figure 1.11) (Grabowski and Kelman, 2003; Kelman and Kelman, 2003). Since archaeal replication machinery is a simplified form of that in eukaryotes, archaea has been used as a model organism since 1990 to elucidate the functions of each of the components of the eukaryotic-type replication complex.

1.8.1. Archaeal replication origin

In 2001, Matsunaga and colleagues identified the first archaeal replication origin in *Pyrococcus abyssi*, an anaerobic hyperthermophile, by using pulsed-field gel electrophoresis and two-dimensional gel analysis (Figure 1.12). This origin was located upstream of a gene encoding the archaeal homologue of both eukaryotic Cdc6 and Orc1. Due to this sequence similarity and binding of the protein to *oriC* region, this protein was named Cdc6/Orc1. This single origin of *P. abyssi* contained one A-T-rich region referred to as duplex unwinding elements (DUE) (Matsunaga *et al*., 2001). In contrast to *P.abyssi*, three origins of replication were identified in *Sulfolobus acidocaldarius* and *Sulfolobus solfataricus* which belong to Crenarchaeota phylum. Characterization of *Sulfolobus* origins confirmed the presence of the conserved 13 bp repeats (DUE) and few longer repeated sequences called ORBs (Origin Recognition Box) (Lundgren *et al*.,2004; Robinson *et al*., 2004). Several studies subsequently postulated that similar to eukaryotes other archaeal species contain multiple origins. In general, archaeal origins have one or more DUE elements, which are surrounded by ORBs and these ORBs were shown to be the binding site for Cdc6 proteins. In many species, Cdc6 protein also binds to a consensus sequence named as mini-ORB that has sequence similarity with ORBs (Bell, 2012).

Attribute	Bacteria	Eukarya	Archaea
Chromosome	Circular	Linear	Circular
Replication origin	Single	Multiple	Single or Multiple
Pre-replication complex(pre-RC)			
Origin recognition	DnaA (1)	ORC (6)	Cdc6 (≥1)
Helicase	DnaB (1)	MCM (6)	MCM (≥1)
Helicase loader	DnaC (1)	Cdc6 (1) Cdt1(1)	Cdc6 (≥1)
Pre-initiation complex (pre-IC)			
Cdc45	_	Cdc45 (1)	Cdc45 (1)
GINS	_	GINS (4)	GINS (1-2)
RPA	SSB (1)	RPA (3)	RPA(1-3)
Elongation complex			
Primase	DnaG (1)	Polα/primase (4)	Primase (2)
Sliding clamp	β-subunit (1)	PCNA(1)	PCNA (1)
Clamp loader	Ƭ-complex (5)	RFC (5)	RFC (2)
DNA polymerase	PolC (3)	PolB (1)	PolB/PolD (1/2)
Okazaki fragment maturation			
Primer removal	PolI (1)	Fen1 (1)	Fen1 (1)
DNA ligase	NAD^+-dependent (1)	ATP-dependent (1)	ATP-dependent (1)

Table 1.1. A summary of DNA replication features in three domains of life

Bacterial, eukarya, and arcahaeal-specific proteins are shown. Parentheses represent the number of homologs identified in different species.

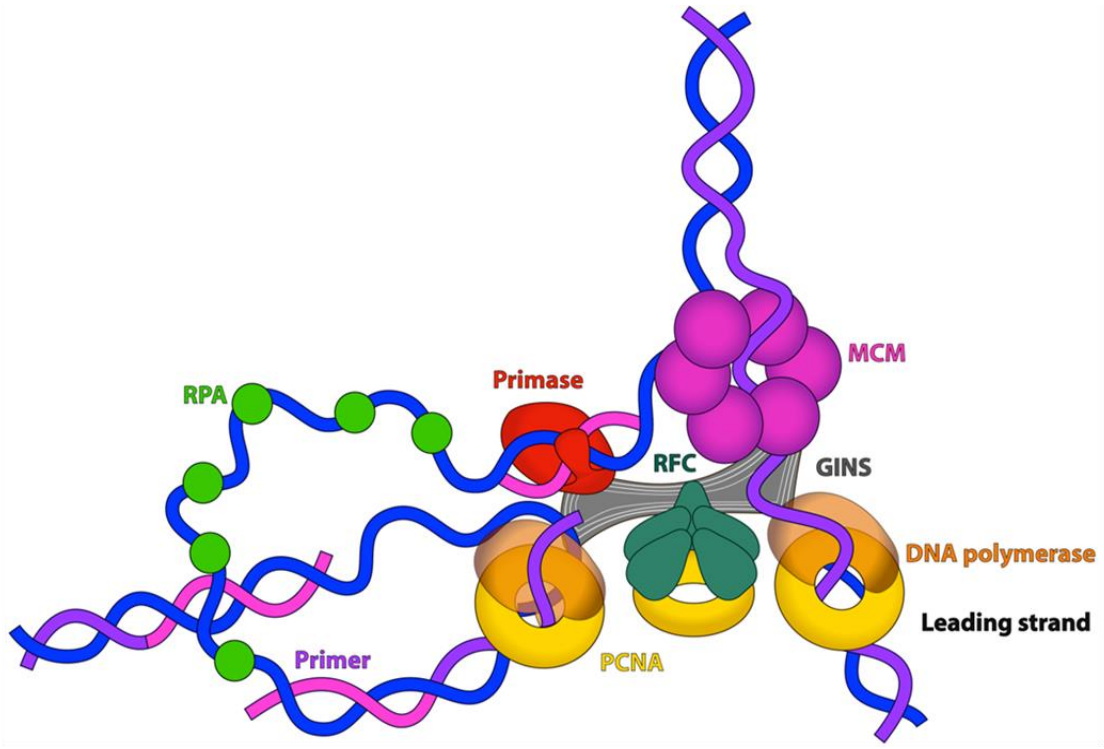


Figure 1.11. The overall architecture of archaeal replisome.

The role of hexameric MCM, at the replication fork is to unwound the duplex DNA into two ssDNA. The ssDNA can be protected by replication protein A (RPA) from endonucleolytic degradation. The pentameric replication factor C (RFC) loads the trimeric PCNA onto the DNA in an ATP-dependent manner. The processivity factor (PCNA) encircles the DNA and attaches the DNA polymerase to the template. Heterodimeric DNA primase synthesis a short strand of RNA or DNA primer, subsequently the primer is extended by DNA polymerase (Kelman and Kelman., 2014).

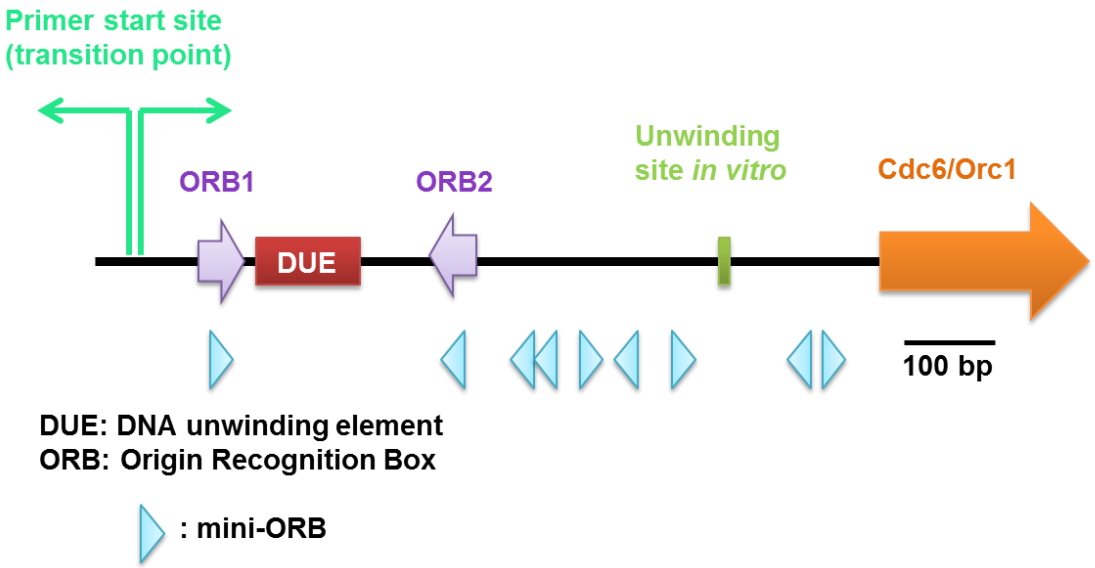


Figure 1.12. The *oriC* region in *Pyrococcus* genome.

oriC is a conserved region with 13 bp repeats. DUE, is surrounded by two origin recognition boxes (ORB). The 13 base repeat is called a mini-ORB. Small arrowheads indicate two mimi-ORB and the ORB1 and ORB2 are indicated by large arrow. Green arrows indicate the transition site. The orange arrow indicatest he cdc6/orc1 that binds to the oriC located in downstream (Ishino and Ishino, 2013).

Amino acid sequence alignments revealed that archaeal Cdc6/Orc1 proteins belong to AAA+ family of protein. In addition, crystal structure study of the Cdc6/Orc1 protein from *Pyrobaculum aerophilum* and *Aeropyrum pernix* demonstrated that Cdc6/Orc1 proteins consist of an N-terminal AAA+ ATPase domain and a C-terminal domain with a conserved winged helix fold. Both domains are shown to be involved in origin DNA binding (Singleton *et al*., 2004). Following the binding of Cdc6/Orc1 to the ORB sequence, the protein changes its conformation in the presence of ATP. It is believed that this conformational change is important for replication initiation, similar to that of bacterial DnaA protein (Akita *et al*., 2010).

Notably, loading of the replicative DNA helicase, another important protein involved in the initiation of DNA replication, at oriC is Cdc6/Orc1-dependent, but not in an ATP-dependent manner (Akita *et al*., 2010). Archaeal pre-recombination complex (pre-RC) consists of the assembly of the Cdc6/Orc1 and MCM helicase at the replication origins. However, until recently the archaeal homolog of eukaryotic Cdt1 has been identified only in some species belonging to crenarchaeota phylum.

1.8.2 MCM helicase

The first biochemically characterized archaeal MCM homologue was from *Methanothermobacter thermautotrophicus* (Kelman *et al*., 1999). Since then, many remarkable studies have been done to characterize other archaeal MCM proteins. Most archaeal species have a single gene encoding MCM. However, in those consisting more than one MCM genes, only one is important for viability (Ishino *et al*., 2011). Unlike eukaryotic MCM, archaeal MCM is a homohexamer or homo double hexamer, which shows DNA helicase activity on its own *in vivo* and *in vitro*. However, in *P. furiosus,* similar to eukaryotes, efficient MCM helicase activity was observed in the presence of other accessory factors e.g. GINS (Yoshimochi *et al*., 2008). In common with archaeo-eukaryotic replicative helicases archaeal MCM shows 3'→5' helicase activity (Barry *et al*., 2007; Grainge *et al*., 2003). As with other helicases, archaeal MCM translocates along ssDNA and dsDNA in an ATP-dependent manner. Characterization of *Archaeoglobus fulgidus* pre-replication complex demonstrated that both Cdc6/Orc1 and MCM proteins have DNA substrate specificity. Based on this data,

these proteins prefer a single stranded bubble that mimics early replication intermediates (Grainge *et al*., 2003). Different studies have been characterized the function of Cdc6/Orc1-MCM interaction in various archaeal species. While in both *M. thermautotrophicus* and *S. solfataricus* Cdc6/Orc1 inhibits MCM activity, Cdc6/Orc1 from *Thermoplasma acidophilum* stimulates the helicase activity of MCM (De Felice *et al*., 2003; Haugland *et al*.,2006).

Structurally, MCM contains two distinct (N-terminal and C-terminal) domains and a catalytic region. These domains are interacting with one another through a conserved long loop named as allosteric control loop (ACL). This conserved loop can facilitates the N-terminal and C-terminal interaction in response to ATP hydrolysis. The N- terminal domain of MCM possesses C_4-type zinc finger and β-hairpin motifs, essential for DNA binding (Brewster and Xiaojiang, 2010). It was proposed that addition of N-domain can regulates processivity and DNA substrate specificity of the MCM (Barry *et al*., 2007). The C-terminal is an AAA+ helicase domain and consists of a small winged helix bundle. This conserved region involves in ATP hydrolysis and DNA unwinding. The central part is the catalytic region of MCM contains three β-hairpins that are required for helicase activity (Fletcher *et al*., 2003).

Recruitment of MCM to *oriC* has been extensively studied in recent years. An *in vitro* study on *P. furiosus* illustrated the importance of Cdc6/Orc1 in MCM helicase assembly at *oriC.* Based on this study, preloading of Cdc6/Orc1 onto the ORBs significantly reduced MCM assembly onto the DNA, which suggested that free Cdc6/Orc1 is needed for MCM recruitment (Akita *et al*., 2010). Using protease sensitivity assay, Samson and colleagues investigated the relationship between ATP binding and hydrolysis activity of Cdc6/Orc1 and regulation of MCM recruitment onto oriC. It was reported that, in the presence of ATP, the conformation of Cdc6/Orc1 was changed which led to MCM recruitment. Interestingly, once the MCM is loaded, subsequent hydrolysis of ATP to ADP prevents further MCM recruitment by Cdc6/Orc1 (Samson and Bell, 2013).

1.8.3 GINS

Although archaeal MCM is able to act on its own without need to bind other proteins. However, the CMG complex also exists in archaea (Xu *et al*., 2016).

Eukaryotic GINS is a ring-shaped heterotetramer, comprising the Sld5, Psf1, Psf2 and Psf3 subunits (GINS is derived from the Japanese go-ichi-ni-san, meaning 5-1-2-3) (Figure 1.13). As well as interacting with MCM and Cdc45, the GINS complex also shows interaction with Polα-primase and Pol ε DNA polymerases. This suggests the role of GINS complex in both initiation and elongation in eukaryotes DNA replication (Labib and Gambus, 2007; Takayama *et al*., 2003).

Bioinformatics analyses, using archaea and eukaryotic genomic information, led to the proposal that archaeal GINS complexes are simplified versions of the eukaryotic complexes. The first archaeal GINS homologue was biochemically detected in *Solfolobus solfataricus* and due to the sequence similarity to Psf2 and Psf3 it named as Gins23. Yeast two-hybrid analyses indicated an interaction between Gins23 and MCM in *S. solfataricus* (Marinsek *et al*., 2006). Although some species consist of a single homologue, the Crenarchaeota and Euryarchaeota have another homologue with sequence similarity to Sld5 and Psf1, referred to as Gins51 (called Gins15). In general, in common with human, archaeal GINS proteins are formed in tetrameric complexes, either homotetramers or a heterotetramer of two copies of Gins15 subunit and twocopies of Gins23 subunit (Ishino *et al*., 2011; Marinsek *et al*., 2006) (Figure 1.13). Structural studies suggested that each of the Gins15 and Gins23 subunits are made of two domains with different orientations: one large A and one smaller B domain. In Gins15 subunit the A domain is located at the N-terminal and the B domain is at the C-terminal (AB type) whereas, in Gins23 the B domain is at the N –terminal and the A domain is at the C-terminal (BA type). Notably, the archaeal and human GINS show a difference in the way the two subunits connect with each (Kamada, 2012; Oyama *et al*., 2011). A crystal structure of *T. kodakarensis* GINS revealed that the protein is a tetramer consisting of Gins51 Gins23, in such a way that two Gins15 proteins are on top of the two Gins23 proteins forming a trapezoid shape with a narrow cavity in the centre. In Gins51 protein, there is a long disordered region. This region enables the conformation of C-terminal domains to be more flexible. Subsequently this creates asymmetrical homotetramer, rather than a symmetrical formation of the *T. kodakarensis* GINS (Oyama *et al*., 2011).

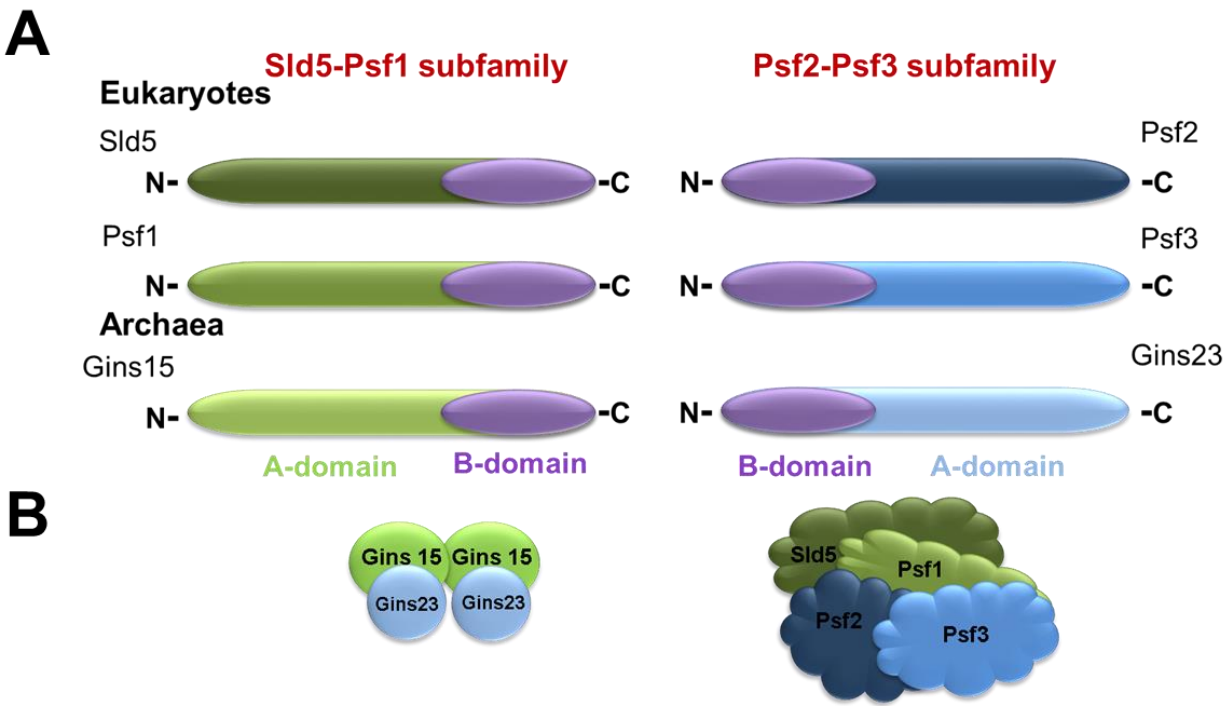


Figure 1.13. The GINS family.

(A). Heterotetrameric eukaryotic GINS is made up of Sld5, Psf1, Psf2 and Psf3 subunits. Archaeal GINS protein is a tetrameric complex consisting of two copies of Gin15 subunit and two copies of Gin23 subunit. Both eukaryotic and archaeal GINS proteins contain an 'A-domain' that is rich in α-helices and a 'B-domain' that is rich in β-strands. **(B).** This carton illustrates organization of the subunits within eukaryotic and archaeal GINS complex (Labib and Gambus., 2007).

Recently it has been shown there is a similarity in structure of the B domain of archaeo-eukaryotic GINS protein and the C-terminal region of the catalytic subunit of archaeal primase (PriS). This can shed a light on the evolution of these proteins (Swiatek and MacNeill, 2010). A number of studies have been suggested that the GINS complex from archaea interacts with some of the fundamental components of relisome, including primase, MCM, DNA polymerase D (polD), PCNA, and RecJ-like proteins, this supports the idea that GINS functions as a scaffolding complex. Using yeast two-hybrid analyses, interactions between Gins23 and both small and large subunits of primase were detected in *S. solfataricus* (Marinsek *et al*., 2006). Li and colleagues successfully co-purified Gins15 and Gins23 from *Thermococcus kodakarensis* with both subunits of PolD and PCNA1/PCNA2 proteins, respectively (Li *et al*., 2010). Although in both *S. solfataricus* and *P. furiosus*, Gins23 interacts with MCM, the effect of the interaction on the helicase activity of MCM is different in each species. The interaction of *S. solfataricus* Gins23 with MCM does not affect the helicase activity. However, in the case of *P. furiosus*, the MCM helicase activity significantly is stimulated by the interaction with Gins23 (MacNeill, 2011).

1.8.4. Archaeal Cdc45

Although Cdc45 is a key DNA replication factor, until recently no counterpart for Cdc45 protein was identified in archaea. Interaction studies using two-hybrid studies indicated that there is an interaction between *S. solfataricus* GINS and a protein known as RecJdbd (RecJ-like DNA-binding domain), which is homologous to the C-terminal domain of bacterial RecJ. In addition, similar to eukaryotic Cdc45, RecJdbd protein lacks the nuclease domain (Marinsek *et al*., 2006). A recent study on *T. kodakarensis* demonstrated that the GINS complex interacts with a novel protein known as GINS-associated nuclease (GAN), which is a *bona fide* orthologous of bacterial RecJ with a DHH phosphoesterase domain. It has been reported that this protein is a processive exonuclease that degrades DNA in the 5'→3' direction (Li *et al*., 2010). Another related report showed sequence similarity between human Cdc45 and a putative ssDNA-specific exonuclease (RecJ protein) belonging to DHH family of phosphoesterases. These studies suggests that Cdc45 and archaeal RecJ-like proteins originated from an ancestral 5'-3' exonuclease that during evolution

underwent dramatic evolutionary events and lost its catalytic activity, but still retained for its ssDNA binding role (Krastanova *et al*.,2012; Makarova *et al*., 2012). Considering that archaea share the core eukaryotic replisome components, and that Cdc45 (the apparent RecJ orthologous) plays a significant role in replication, it is proposed that archaea possess a homolog of CMG complex consisting of RecJ-MCM-GINS.

1.8.5. Single-stranded DNA binding protein

Single-stranded DNA binding protein is one of the critical replication components in all domains of life. In bacteria, single-stranded DNA protein is referred as SSB, whereas in eukaryotes and archaea is known as replication protein A or RPA. Following unwinding the ds DNA by MCM, SSB proteins bind to stabilize and protect ssDNA from possible threats including, nuclease attack and chemical modifications during DNA replication, repair and recombination process (Prakash and Borgstahl, 2012). Significant structural similarities between the bacterial, eukaryotic, and archaeal SSBs, suggests that these proteins may derived from a common ancestor (Kelly *et al*., 1998).

SSB and RPA interact with ssDNA using a common fold, called the OB (oligonucleotide/oligosaccharide binding)-fold (Murzin, 1993). Bacterial SSB is a homotetramer in which each monomer consists of one N-terminal domain with a single OB fold and a flexible acidic C-terminal domain (Kelman *et al*., 1998). In contrast, the eukaryotic RPA is a stable heterotrimer comprises of RPA70, RPA32, and RPA14 proteins. RPA70 contains two OB folds and a conserved C_4-type zinc-finger motif. Only one of the OB folds from RPA70 is involved in ssDNA binding. On the other hand, RPA32 and RPA14 each contain a single OB fold. In addition, RPA32 possess a winged helix domain (WHD) near the C-terminal tail. It is postulated that the C_4-type zing-finger and winged helix motifs facilitates protein-protein interactions (Prakash and Borgstahl, 2012).

Following the identification of homologous SSBs encoded by *Methanobacterium thermoautrophicum* and *Archaeoglobus fulgidus* genomes (Klenk *et al*., 1997), in 1998, the first homologue of SSB from *Methanocaldococcus jannaschii* was characterized. There is a similarity of the amino acid sequence between M. jannaschii RPA and the eukaryotic RPA70. The amino acid sequence contains

four OB folds and one zinc finger motif. *M. jannaschii* RPA is a monomer and in common to eukaryotic RPA, binds to ssDNA with high affinity (Kelly *et al*., 1998). In contrast to methanogenic archaea, *P. furiosus* RPA is a heterotrimer in solution, comprising of three subunits, RPA41, RPA14, and RPA32 similar to the eukaryotic RPA. It was shown that *Pfu*-RPA strikingly stimulates *in vitro* strand-exchange reaction through interaction with RadA recombinant protein (Komori and Ishino, 2001).

It is believed that the RPA proteins from different archaeal lineages are divers in structure and domain organization. While in euryarchaea single, multiple and complex forms of the RPA have been identified, in crenarchaeal more similar to the bacterial, RPA proteins consist of single OB fold and a flexible C-terminal tail (Ishino and Ishino, 2013).

Methanosarcina acetivorans RPA has a unique feature. Two of the *M. acetivorans* RPAs (MacRPA2 AND MacRPA3) have two OB folds in the N-terminal region and a zinc finger motif in the C-terminal part. On the other hand, MacRPA1 is composed of four OB folds and lacks the conserved zinc finger domain. Interestingly, each of the *M. acetivorans* RPA is a putative SSB protein and can stimulate the primer extension activity of DNA polymerase BI from *M. acetivorans* (Lin *et al*., 2008).

From eukaryotic studies, it has been shown that replication protein A can modulate primase activity (Guilliam *et al*, 2015). Previous study from our lab has established the interaction of RPA70 and human PrimPol (Guilliam *et al*., 2015). Unlike SSBs, which stimulate replicative polymerase activity (Braun *et al*., 1997), RPA70 significantly limits the primase and polymerase activities of PrimPol (Guilliam *et al*., 2015). However, it was recently reported that the primase activity of PrimPol can alsp stimulates by RPA70 if it titrates at 1:1 levels. (Guilliam *et al*., under review).

1.9. DNA polymerases

1.9.1. Discovery of DNA polymerases

The term ‘‘polymerase’’ was used for the first purification and characterisation of an enzyme system from *Escherichia coli* in 1958, for the first time. This enzyme was capable to incorporate deoxynucleotides (dNTPs) into DNA in the presence

of Mg^{2+} and all four deoxynucleotides in the form of triphosphate (Lehman *et al*., 1958). Soon after, Kornberg and collaborators showed that polymerases catalyse DNA polymerization of deoxynucleoside triphosphates by liberation of an inorganic pyrophosphate. They also suggested that polymerase activity requires a DNA primer. The polymerization mechanism involves transferring an "activated" nucleoside triphosphate onto the growing strand of DNA following nucleophilic attack by the 3'-hydroxyl group of the primer terminus on the α and β phosphate groups of the incoming dNTP, resulting in liberation of pyrophosphate (PPi) (Figure 1.14). One of the characteristic features of DNA is to have complementary base pairs, i.e. adenine is paired with thymine by two hydrogen bonds and cytosine with guanine by three hydrogen bonds. This is achieved through the large difference in enthalpy between matched base pairs and mismatched base pairs (Petruska *et al*., 1988).

Probably the most remarkable feature of DNA polymerases is the rapid and accurate rate of DNA replication. DNA replication high fidelity is dependent on the correct geometry of base pairs and proofreading mechanisms. DNA polymerases catalyse the first step of proofreading just before a new nucleoside triphosphate is added to the growing chain. The next step, which is referred as 3' exonuclease proofreading, occurs once the incorrect nucleoside is covalently added to the growing strand of DNA. It is worth noting that DNA polymerases cannot synthesise novel strands of DNA *de novo*. Instead, this activity needs a primer with a free 3'-hydroxyl group that produced by specialised proteins known as DNA primases (section 1.9.).

1.9.2. Classification of DNA polymerases

DNA polymerases are the main enzymes responsible for of DNA replication in all three domains of life. Discovery of the first DNA polymerase, named as DNA polymerase I and isolated from *E.coli,* was followed by isolation and characterisation of many different DNA polymerases (Lehman *et al*., 1958). With the growing number of DNA polymerases, it became necessary to classify the enzymes groups sharing similar structural and functional features. DNA polymerases are categorized into seven families, A, B, C, D, E, X, and Y based

on the amino acid sequence similarity (Table 1.2) (Bebenek and Kunkel, 2004; Joyce and Benkovic, 2004).

Although the ability of DNA polymerases to synthesise DNA strand is conserved among all DNA polymerase families, other features, such as processivity, fidelity, and substrate nucleotide selectivity are different in each case. Enzymes within the same family often exhibit similar properties. It is believed that *E.coli* contain five DNA polymerases; Pol I, Pol II, and Pol III belong to families A, B, and C, respectively, Pol IV and Pol V, which belongs to Y family DNA polymerases that are specialize in translesion synthesis and bypassing damaged bases. (Guo *et al*., 2009; Ohmori *et al*.,2001).

It is reported that family B DNA polymerases contain the highest number of polymerases in viruses, bacteria, archaea and eukaryotes (Braithwaite and Ito, 1993). These polymerases are highly accurate and show high processivity when they bind to the sliding clamp cellular processivity factor, PCNA (Capson *et al*.,1992). In eukaryotes, replicative polymerases α, δ, ε belong to family B DNA polymerases.

Eukaryotic DNA polymerase ζ is a unique family B polymerase, which in contrast to the other members does not possess a 3'→5' exonuclease activity but is involved in translesion DNA synthesis. Some bacteria possess a family B DNA polymerase, and Pol II from *E.coli* is the most characterised example (Lovett, 2011). In common with bacteria and eukaryotes, archaea possess several DNA polymerases. Unlike crenarchaeota which possess at least two family B polymerases, in euryarchaeota only one family B polymerase exists (Uemori *et al*., 1993; Uemori *et al*., 1995).

Examples of each family and their putative functions *in vivo* are listed in Table 1.2. Features of each DNA polymerase are also summarized.

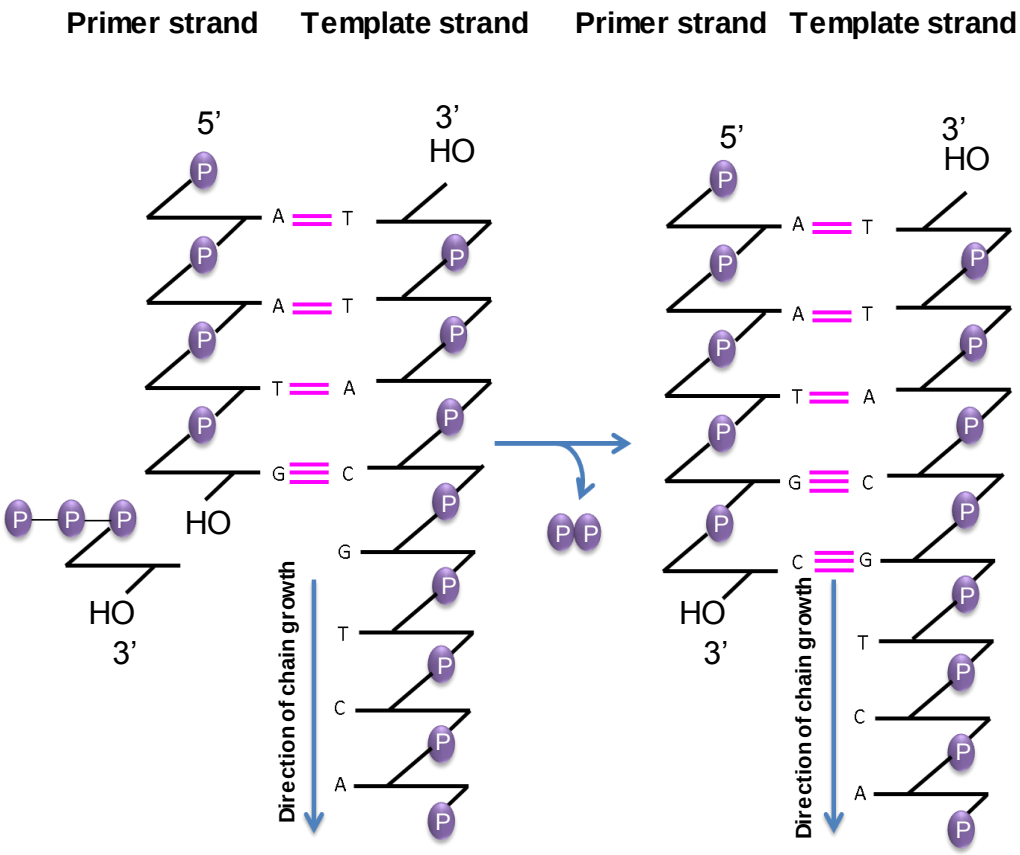


Figure 1.14. General scheme of DNA polymerisation.

The reaction is catalysed by DNA polymerases. Nucleophilic attack of the 3'-OH group of the primer on the α-phosphate group of the incoming dNTP forms a phosphodiester bond and releases a pyrophosphate. The pink lines indicate hydrogen bonds between complementary bases (Taken from Cooper, 2001).

Family	Examples	Functions	Features
A	T7 DNA polymerase, E.coli pol I, Taq pol I Eukaryotic pols γ, ν, θ	Chromosomal and mitochondrial DNA replication. Okazaki fragment processing. NER	5'→3'exonuclease, 3'→5; exonuclease
B	Eukaryotic pols α, δ, ε, ζ, T4 DNA polymerase, E.coli pol II, Some Viral and archaebacterial polymerase.	Chromosomal DNA replication These enzymes are generally tethered by a processivity factor Pol ζ used in DNA damage bypass	3'→5'exonuclease
C	Bacterial DnaE, E.coli pol III, Bacillus subtilis pol III	Chromosomal DNA replication DNA damage bypass	3'→5'exonuclease
D	Euarchaeota polymerases Including Pyroccocus furiosus polymerase II	suggested to be a replicative polymerase	3'→5'exonuclease
X	Eukaryotic pols β, μ, λ, Yeast pol IV, African Swine Fever Virus polX, Human Terminal Deoxynucleotide Transferase	Involved in DNA damage processing including base excision repair and double strand break repair	5'-deoxyribose phosphate lyase
Y	Eukaryotic pols η, κ, ι and REV1 E.coli DinB, UmuC	DNA damage bypass. V(D)J	DNA damage bypass
RT	HIV RT, Moloney murine Leukaemia virus RT Hepatitis B RT, Telomerase	Reverse transcription of viral genomes.Telomerase has an innate RNA component to synthesise GC-rich repeats at telomeres	RNase H

Table 1.2. Representative members of seven families of DNA polymerase.

Examples of each family and their putative functions *in vivo* are listed. Features of each DNA polymerase are also summarized.

Interestingly, euryarchaeal genomes encode a family D polymerase, which is unique to the archaea and has never been found in other domains (Cann *et al*., 1998). However, novel phyla in archaea including, Thaumarchaeota, Korarchaeota, and Aigarchaeota, harbour the genes encoding Pol D. Ishino and colleagues first discovered heterodimeric family D polymerase (Imamura *et al*., 1995). The family D polymerases comprise a small subunit (DP1), with sequence similarity to the non-catalytic B subunit of eukaryotic family B DNA polymerases. On the other hand the larger subunit of this family (DP2) shows no remarkable homology to other DNA polymerases (Cann *et al*., 1998). It appears that DP1 possess 3'→5' exonuclease activity and DP2 is involved in polymerisation (Cann and Ishino., 1999; Cann *et al*., 1998). Discovery of the 3'→5' exonuclease activity of family D polymerases and association with the sliding clamp (PCNA) suggests the involvement of this family in chromosomal DNA replication and DNA repair in archaea. A recent study of the biochemical behaviour of *Thermococcus kodakarensis* D and B polymerases postulated, unlike Pol B, Pol D is adequate for genome replication. Therefore, Pol D may be the essential replicative DNA polymerase in this species (Cubonova *et al*., 2013).

In addition, a few studies proposed that *T.kodakarensis* Pol D could form a stable complex *in vivo* with other replisome components including, Cdc6, GINS, MCM1, PCNA1, PCNA2, and GAN. Conversely, Pol B has not been present as a complex with these proteins. Based on these data, it can be concluded that Pol D, and not Pol B, functions as the genome-replicating DNA polymerase in *T.kodakarensis.* On the other hand, a biochemical study suggested that *Pyrococcus abyssi* Pol D is the major lagging strand polymerase and Pol B synthesizes the leading strand (Henneke *et al*, 2005). In order to complete Okazaki fragment maturation on the lagging strand, PolD requires other accessory proteins such as RPA, PCNA and RFC. Pol B is also required to fill the single-stranded gaps and displaces the 5'-end of the downstream fragment and form a flap structure which is subsequently cleaved by Fen1 (Greenough *et al*, 2015).

Unlike replicative polymerases from bacteria and eukaryotes, the archaeal replicative polymerase PolB recognizes uracil in DNA templates and stalls replication encountering these bases. Specifically family B DNA polymerases

through a specialized binding pocket in their N-terminal domain can bind tightly to uracil containing DNA templates (Greagg *et al*., 1999). Uracil commonly arises in DNA by deamination of cytosine to uracil, and replication of the deaminated bases results in transition mutations (C:G → A:T) in 50% of progeny (Lindahl., 1993). This unique property of Pol B family polymerases is not restricted to hyperthermophilic archaea. In fact, family B polymerases from mesophilic archaea, like *M. acetivorans* (optimum growth temperature 35-40 °C) can bind to template-strand uracil as firmly as Pfu-Pol B (Kelman and White, 2005;Wardle *et al*., 2008). A recent study proposed that the presence of dU base in DNA significantly reduces the polymerization activity of archaeal Pol D (Richardson *et al*., 2013). Together, the ability of the two polymerases to bind and respond to uracil can shed a light on the importance of these polymerases in DNA replication, as previously suggested.

Family Y DNA polymerases are relatively new DNA polymerases that shows little homology to any of the five previously discovered family polymerases. This family was first described as UmuC/DinB/Rev1/Rad30 superfamily and has been identified in all three domains of life. The most remarkable feature of this family is performing low fidelity while copying undamaged DNA templates. If members of A, B, C, D, and X family of polymerases block replication at sites of damage, Y family polymerases can bypass the lesion and rescue the stalled replication fork. Phylogenetic analysis reveals that the DinB family is widely distributed across three domains of life. In *E.coli,* Pol IV belongs to DinB family and in human Pol ɵ and Pol κ are products of *DINB1* gene (Ohmori *et al*., 2001). The first homologue of UmuC in archaea was identified in *Sulfolobus solfataricus* using the motif PCR approach. This homologue with high sequence similarity to DinB was designated as *dbh* (*dinB* homolog), which is also known as Dpo4 (Kulaeva *et al*., 1996). It has been reported that Dpo4 is an error-prone enzyme which can read through an abasic site, a *cis-syn* thymine–thymine dimer, as well as acetyl aminofluorene adducted- and cyclobutane pyrimidine dimer (Boudsocq *et al*., 2001; Ling *et al*., 2003). Y family DNA polymerases are 100-1000 fold less accurate than polymerases from other families. Among Y family polymerases, DinB is the most accurate enzyme, making about one mistake per 10^3 to 10^4 bases copied (McCulloch and Kunkel, 2008).

1.9.3. Structure and function of DNA polymerases

The carboxyl-terminal domain of DNA polymerase I isolated from *E. coli* retains DNA polymerase and 3'→ 5' exonuclease activities. This domain, known as the Klenow fragment (Klenow and Overgaard-Hansen, 1970), was the first DNA polymerase structure resolved by X-ray crystallography (Ollis *et al*., 1985). All DNA polymerases whose structures are solved to date, share a common overall architecture. The overall structure resembles that of a right hand with the palm domain at the bottom and thumb and fingers domains at each side which form a ''U'' shaped cleft (Figure 1.15). The function of each domain appears to be conserved in all polymerases. The palm domain that contains the polymerase active site involved in catalysis. The function of fingers domain is to interact with the incoming nucleoside triphosphates and on the other hand, the thumb domain is important for binding to template DNA and also plays a role in processivity and translocation (Steitz, 1999). Y family DNA polymerases possess conserved sequences at the N-terminal domain, in which all the essential catalytic amino acids reside. On the other hand, C-terminal domain does not contain the conserved primary motifs, but contains a conserved tertiary structure, consisting of four-stranded β sheet with two long α helices on one side. This domain is referred as PAD (polymerase-associated domain) in eukaryotes and little finger domain in archaeal and bacterial system. In addition, in Y family, the fingers and thumb subdomains appear to be smaller which makes the palm domain more open and solvent accessible (Pata, 2010, Johnson *et al*., 2001).

Based on solution studies, the DNA polymerisation mechanism consists of series of reactions. Polymerisation starts with binding of DNA substrate followed by addition of deoxynucleoside triphosphates to the polymerase-DNA complex. Binding of DNA results in minimal changes in the finger and palm domain but leads to significant conformational change in a helix-loop-helix motif in the thumb domain. If the incoming dNTP that binds to fingers domain is correct, a rate-limiting conformational change forms a tight ternary complex. Subsequently, a chemical reaction leads to formation of tightly bound enzyme-

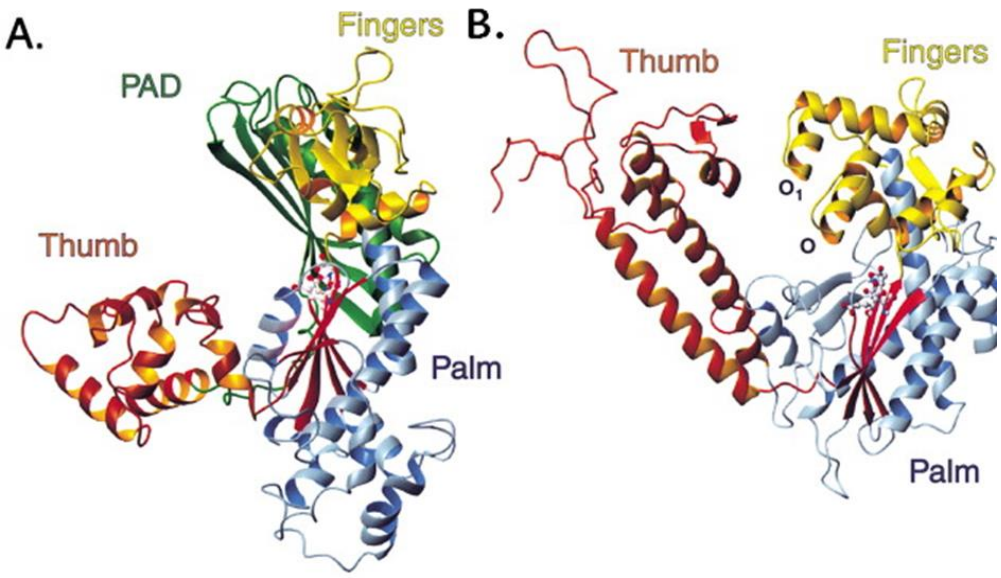


Figure 1. 15. Structures of Polη and T7 DNA Polymerases

(A) Polη and **(B)** T7 DNA polymerases share the common right hand shape architecture. In both polymerases, the palm domain is at the bottom and fingers and thumb domains are at each site, which form a U shape cleft. As indicated, the Polη fingers and thumb domains are smaller than the equivalent domains in T7 polymerase (Taken from Johnson *et al*., 2001).

product complex. In the next step, another conformational change relaxes the tightly bound enzyme-product complex. This step results in formation of phosphodiester bond and releasing pyrophosphate ion, which helps translocation of the DNA for the next nucleotide addition (Li *et al*., 1998).

DNA polymerase requires two metal ions for its activity that involves a nucleophillic attack by the 3'-OH group of the primer on the α-phosphate of the incoming dNTP. Both metal ions are located in the active site. While deoxynucleoside triphosphate (dNTP) and hydroxyl group of the primer are bound by one metal ion, the other one interacts only with the hydroxyl group (Figure1.16) (Steitz, 1993). Carboxyl groups of two aspartate residues in the palm domain connect the two metal ions. Both metal ions are presumed to stabilise the transition state through stabilising the extra negative charges that build up on the trigonal-bipyramidal transition state. Interestingly, a new study through using in crystal reaction and time resolved X-ray diffraction analysis discovered, in contrast to the two-metal-ion theory, an additional metal ion is required for the polymerisation mechanism. It is suggested that unlike the two metal ions, this third metal ion is not coordinated by DNA polymerase and may facilitates the phosphoryltransfer reaction with or without enzyme catalysis (Figure1.17) (Yang *et al*., 2016).

1.9.4. Fidelity of DNA polymerase

Since high accuracy of DNA replication is an essential factor for maintaining genomic stability, DNA polymerases are required to ensure faithful DNA replication through multiple steps. In general this is performed with approximately one error is generated during replication of each 10^9 to 10^{10} bases (Echols and Goodman, 1991). There are multiple steps to achieve the fidelity of DNA replication, including the ability of DNA polymerase to read a template strand, and select the correct nucleoside to insert at the 3' primer terminus. At this point, replication accuracy is estimated to be approximately one error per 10^3 to 10^5 bases. If an incorrect nucleotide being inserted into the growing chain, DNA polymerase can remove the incorrectly incorporated nucleotide by the action of 3'-5' exonuclease, referred to as ''proofreading'' (section 1.8.5). This can increase the fidelity by a factor of several hundred fold. In the case of escaping selection and proofreading steps, incorporation of the wrong nucleotide results in

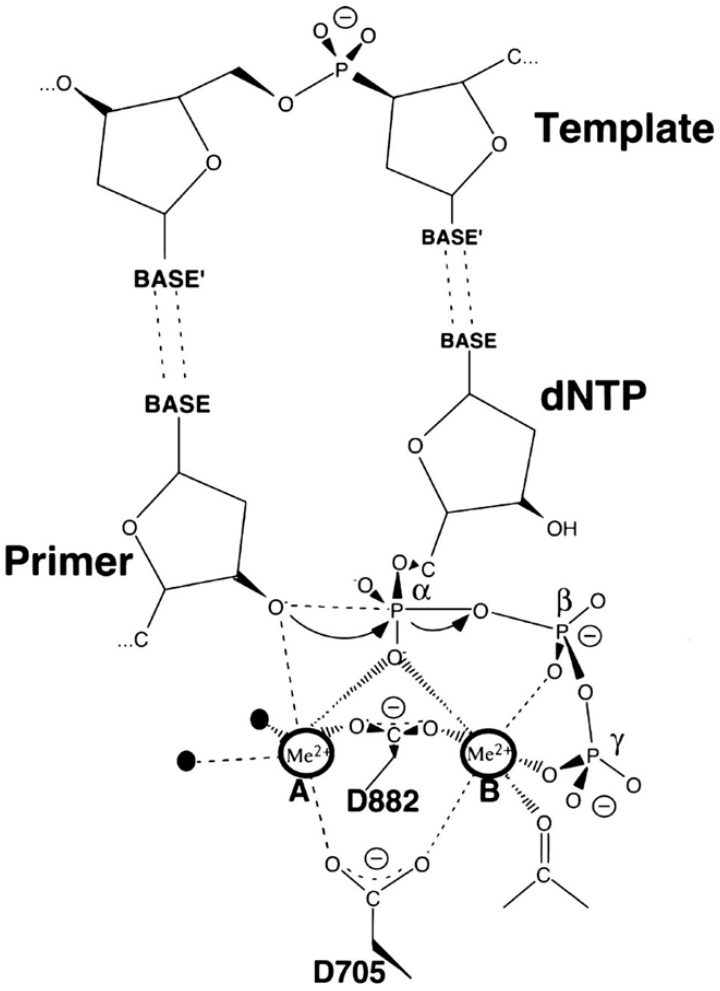


Figure 1.16. The two metal ion mechanism of the T7 DNA polymerase
Aspartic-acid residues 705 and 882 coordinate the two divalent metal ions A and B to T7 DNA polymerase. The metal ion A binds to water molecules, which are shown in black circles (Taken from Steitz.,1998).

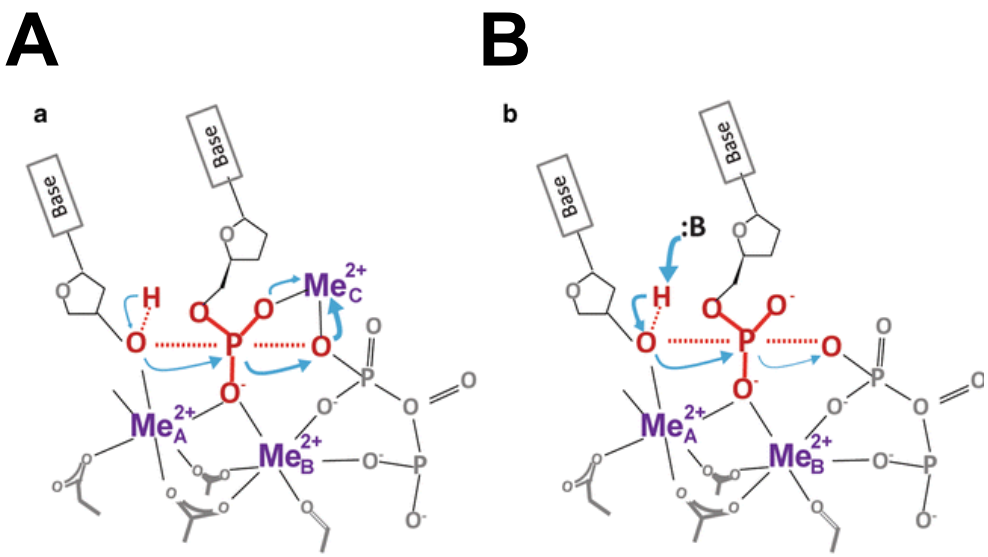


Figure 1.17. Comparison of two versus three metal ion mechanism in DNA polymerisation

(a) In three-metal ion catalysis, phosphoryltransfer reaction initiates by cleavage of phosphodiester bond in dNTP by the C-site metal ion. Deprotonation of the C-site metal ion is the result of this reaction. **(b)** In two-metal ion catalysis, initiation of phosphoryltransfer reaction is done through deprotonation of 3'-OH group. A comparison of the initiation of phosphoryltransfer in two- versus three-metal ion catalysis is shown (Taken from Yang *et al*., 2016).

a mismatch.In this case, the damage can be removed by replicative excision repair pathways, such as MMR (See section 1.4.2.) leading to increasing the fidelity few hundred times (Rothwell and Waksman, 2005).

The first proposed model for the deoxynucleotide incorporation pathway was based on the structure of DNA (Watson, 1953a; Watson, 1953b). One way to determine selection of correct over incorrect nucleotides is by the A-T and G-C hydrogen bond–mediated base pairing. However, the error rate limiting of DNA polymerases for nucleotide insertion is 10^{-3}–10^{-6} per base incorporation and the difference in free energy between correct and incorrect base pairs is sufficient only for an error rate of 1 per 2 x 10^{2} (Loeb and Kunkel, 1982). The second model was proposed based on the similar geometry of the correct Watson-Crick base pairs. This model relies on the assumption that the shape of DNA polymerase active site can accommodate only correct Watson-Crick base pairs and non-Watson-Crick base pairs might be rejected (Bruskov and Poltev, 1979; Engel and von Hippel, 1978). Recent study on DNA polymerase β postulated that, binding of the correct Watson-Crick base pair induces chemical shift changes resulting in DNA polymerase rotation leading to a switch from an open to a closed conformation. In contrast, incorporation of the wrong nucleotide does not affect the enzyme conformation and remains it in an open unstable conformation. This suggested, only in the presence of correct nucleotide the DNA polymerase switches to a closed ternary complex conformation (Moscato *et al*., 2015).

1.9.5. Proofreading mechanisms

As discussed earlier, many polymerases achieve high-fidelity DNA replication through their 3′→ 5′ exonuclease activity, often referred to proofreading. In 1972, two similar proofreading activities by two different DNA polymerases were discovered. Brutlag and Kornberg reported that *E.coli* Pol I DNA polymerase possess an exonuclease active site, which plays a significant role in error avoidance (Brutlag and Kornberg, 1972). It was also suggested, in common with Pol I, bacteriophage T4 DNA polymerase performs similar 3′→ 5′ exonuclease activity (Muzyczka *et al*., 1972). It was revealed that the replication fidelity can be affected by the T4 alleles through altering the spontaneous mutation frequency (Speyer, 1965). A structural study of Klenow fragment of DNA polymerase I has shown that the DNA polymerase domain and exonuclease domain are separated

from each other by a distance of 33 Å (Beese and Steitz, 1991). When misincorporation of a base occurs then further polymerisation becomes very slow, followed by translocation of the 3'-end of the primer to the proofreading active site through a DNA binding cleft between the polymerase and exonuclease domains.

The first step in proofreading is to recognise the misincorporated nucleotide at the primer-end and then to decide whether to continue mismatch extension or to transfer the primer-end from the polymerase to the exonuclease active site. In T4 and RB69 DNA polymerases, there is a β hairpin structure. This structure places itself between the template and primer to separate them. Once the incorrectly incorporated nucleotide at the primer-end is transferred into the exonuclease active site, it will be removed rapidly from the primer-end. Finally, the trimmed primer-end will be transferred back to the polymerase active site. Therefore, proofreading is a switching process between polymerase-to-exonuclease and exonuclease-to-polymerase active site (Baker and Reha-Krantz, 1998; Darmawan *et al*., 2015).

1.10. DNA Primase

As previously discussed, DNA polymerases are unable to initiate *de novo* polymerase synthesis reactions. Early studies suggested that DNA replication initiation requires RNA transcription (Brutlag *et al*., 1971). It was believed that DNA replication initiation proceed in two separate stages. The first step requires ribonucleoside triphosphates (rNTPs), a single stranded DNA, and enzymes. A DNA polymerase can continue DNA synthesis using the product of the first step, a short RNA chain, and dNTPs (Wickner *et al*., 1972). Soon after, Bouche and colleagues discovered that *E.coli* DnaG gene has a role in the replication initiation of nascent (Okazaki) fragments (Bouche *et al*., 1975). DNA-dependent RNA polymerases, distinct from classical RNA polymerases, which are capable of priming, were named as DNA primase (Rowen and Kornberg, 1978).

Until relatively recently, it was believed that the role of DNA primases was only limited to initiation of DNA replication. However, a large body of recent evidence has established that the activities and cellular roles of this family of polymerases extends much further than simply the initiation of DNA replication. Over recent years, different studies established the importance of primases in wide range of

biological processes including, DNA replication, repair, damage tolerance and transcription.

1.10.1. DNA primases- evolution and structure

Nature has evolved several distinct mechanisms for generating primers with a free 3' hydroxyl group to be subsequently elongated by DNA polymerases: (1) In all known systems and some DNA viruses, phages and plasmids generation of a short RNA primer with a free 3'-OH group is required. The synthesised primer is subsequently elongated by a DNA polymerase (Frick and Richardson, 2001). (2) Retroviruses initiate DNA replication using a tRNA with a free 3'-OH group which can be elongated using a reverse transcriptase (Mak and Kleiman, 1997). (3) In adenoviruses, Ser580 of the terminal protein provides a free 3'-OH group, which can be elongated by DNA polymerase (Pronk and van der Vliet, 1993). (4) Several DNA viruses, like parvoviruses, utilize a rolling-circle replication (RCR) system. During RCR, an endonuclease nicks one strands of DNA. The 5' end of nicked DNA is transformed onto a tyrosine residue on the nuclease and the 3' end with free OH group is elongated by a DNA polymerase to synthesis new strand (Noirot-Gros and Ehrlich, 1996).

Generally, primases are structurally different in two ways, either the primase activity requires one or two subunits, or the primase contains archaeo-eukaryotic primase (AEP) fold or a toprim architecture, which exists in bacterial and phage primases (Iyer *et al*., 2005). Bacterial primase is a single subunit protein belonging to the superfamily of DnaG. This primase consists of a common catalytic, Toprim (topoisomerase-primase) domain. The toprim domain of bacterial DnaG is also conserved in archaeal DnaG orthologous, topoisomerase Ia and topoisomerase II. This suggests the presence of toprim domain in the last universal common ancestor (LUCA) (Aravind *et al*., 1998). Eukaryotes and archaea contain of two subunit primases that are typically referred to as the small, catalytic subunit (PriS/Prim1) and a large accessory subunit (PriL/Prim2). The two-subunit primases are also found in herpes viruses, which contain a primase subunit and a helicase subunit (Iyer *et al*., 2005; Kuchta and Stengel, 2010). The superfamily of proteins that contain an archaeal / eukaryotic fold are known as archaeo-eukaryotic primases, or AEPs. The AEP superfamily proteins are composed of a catalytic core with an N- terminal $(\alpha/\beta)_2$ unit which is not structuraly

identical to other proteins in structural database (PDB). The C-terminal domain of the catalytic core contains a highly derived version of the RNA recognition motif (RRM), conserved in viral RNA-dependent RNA polymerases and A-B-Y family DNA polymerases (Iyer *et al*., 2005).

1.10.2. Evolutionary history of AEPs

The absence of homology between bacteria and eukarya/archaea is not limited to DNA primase superfamilies. The core components of transcription and translation including, DNA ligases, PCNA, RNAse HII, clamp-loaders, etc that are highly conserved in all divisions of life. Conversely, the core components of replication are unrelated or only distantly related between bacteria and eukarya/archaea (Edgell and Doolittle, 1997; Leipe *et al*.,1999). Two major scenarios are suggested to explain the existence of two groups of replicative enzymes. In the first model, bacteria and archaeal/eukaryotic replication machineries evolved separately twice from a last universal common ancestor, which carried out reverse transcription to replicate an RNA/DNA genome (Leipe *et al* 1999) as still happens in retroviruses, e.g HIV. It is believed that by the evolution of core components of DNA replication system, selective pressure eradicates the reverse transcription pathway from most organisms. This scenario is supported by the reverse transcription ability of some DNA primases and polymerases (Jozwiakowski and Connolly, 2011; Gill *et al*, 2014; Jozwiakowski *et al*, 2015). This model can clarify the similarity of transcription and translation core components between bacteria and archaea/eukaryotes (Figure 1.18).

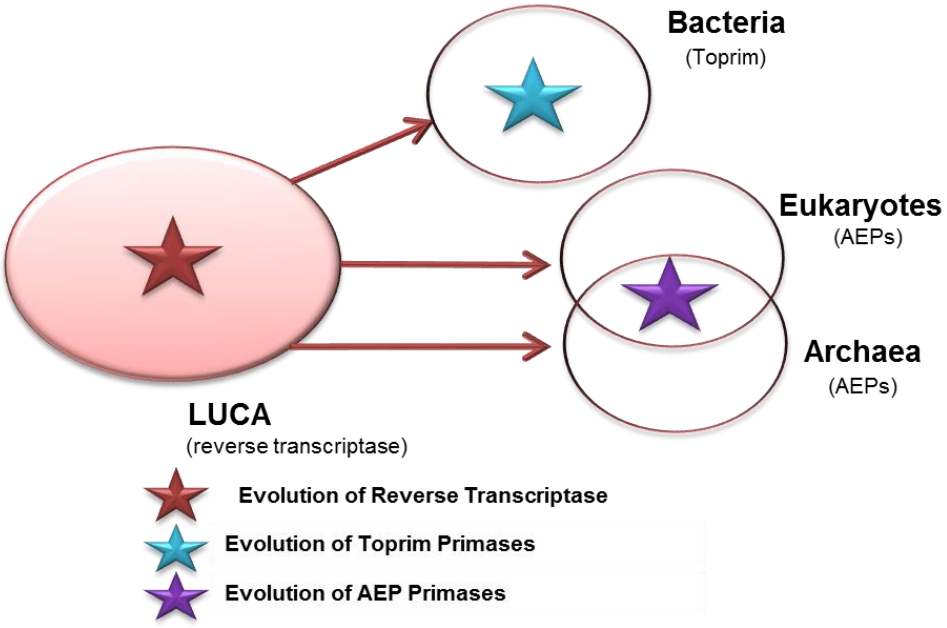


Figure 1.18. Evolution of AEP and Toprim primases (1)

The first model proposes that two replicative systems evolved twice separately from a common ancestor (LUCA) which has reverse transcriptase activity. Bacterial ancestors evolved toprim primases and archaeal and eukaryotic ancestors evolved AEP primases. Horizontal gene transfer between two lineages results in AEP primase's role in NHEJ in bacteria and toprim-type primase's in role in archaeal RNA degradation.

The second model proposes that the last universal common ancestor (LUCA) employed a dual-primase system consisting of both toprim (DnaG) and AEP (PriSL) primase (Hu *et al*., 2012). Apart from the dual-primase role, both toprim and AEP primases in the last universal common ancestor were able to perform RNA degradation and NHEJ processes respectively. Based on this hypothesis, selective pressure eliminated AEPs as replicative primases in bacteria and retained their role in NHEJ (Weller and Doherty, 2001). While in archaea, selective pressure led to the loss of toprim primase's replicative primase activity but the non-essential role in RNA degradation is reserved. Moreover, this model predicts that in eukaryote, the (DnaG) primase is lost and other proteins have replaced to serve its primase activity (Figure 1.19) (Hu *et al*., 2012).

There could be an alternative hypothesis to these two models. In this scenario, the last universal common ancestor consisted of either a toprim primase or AEP and following dramatic evolutionary events, the selective pressure caused the emergence of second primase superfamily to be revealed in either bacteria or archaea/eukarya respectively (Figure 1.20) (Guilliam *et al*., 2015).

1.10.3. Structural analysis of the AEP superfamily

The AEP superfamily shares a catalytic core (Figure 1.21) with an N-terminal $(\alpha/\beta)_2$ unit. This has no structural homology with other proteins and a C-terminal unit which is a highly derived version of the RNA recognition motif (RRM). This motif is also found in viral RNA-dependent RNA polymerases and the catalytic palm domain of A-B-Y family DNA polymerases. The active sites residues of the catalytic core are placed in between the two units, which are packed against each other. The conserved structure consists of six strands (1-6) (two tp belong to N-terminal unit and four to C-terminal domain) and four helices (1-4) (two from each unit). This catalytic core harbours three conserved motifs. Motif I, a hhhD motif in strand 3. Motif II, an sxH motif in strand 5 and motif III, an h- in strand 6 ('h' is a hydrophobic residue,'s' refers to small residue, 'x' refers to any residue and '–' is an acidic residue) (Iyer *et al*., 2005).

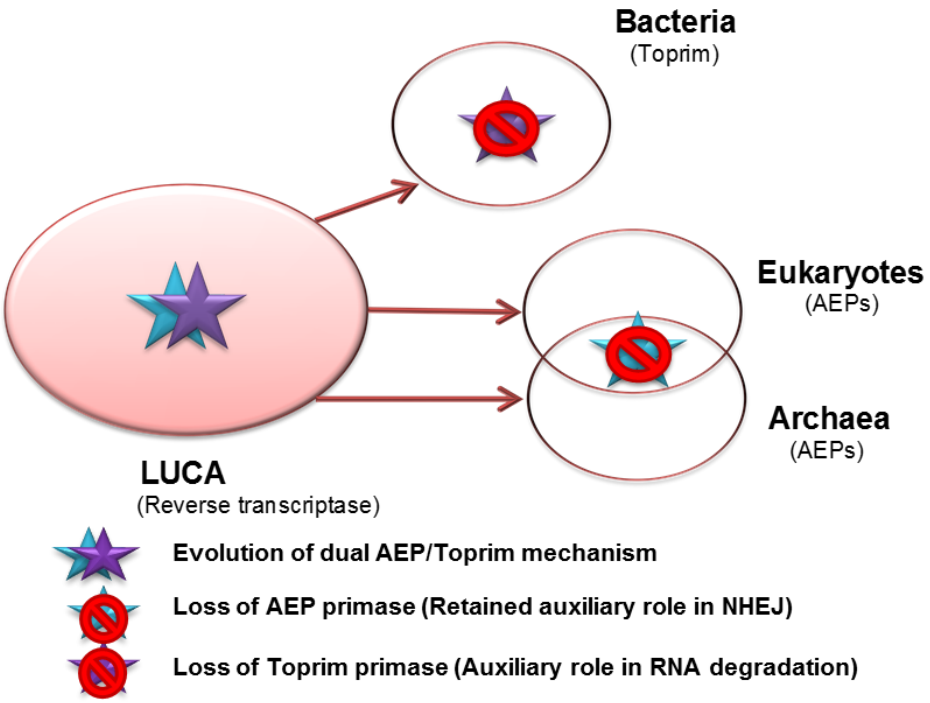


Figure 1.19. Evolution of AEP and Toprim primases (2)

The second model suggests that the last universal common ancestor (LUCA) employed a dual-primase system consisting of both toprim and AEP primases. Selective pressure during evolution led to elimination of replicative function of AEP primases in bacteria but retained them for the auxiliary function of NEHJ DNA repair. During evolution of archaea and eukaryotes selective pressure led to the loss of toprim primase's replicative primase activity but retained the none-essential role of these primases in RNA degradation.

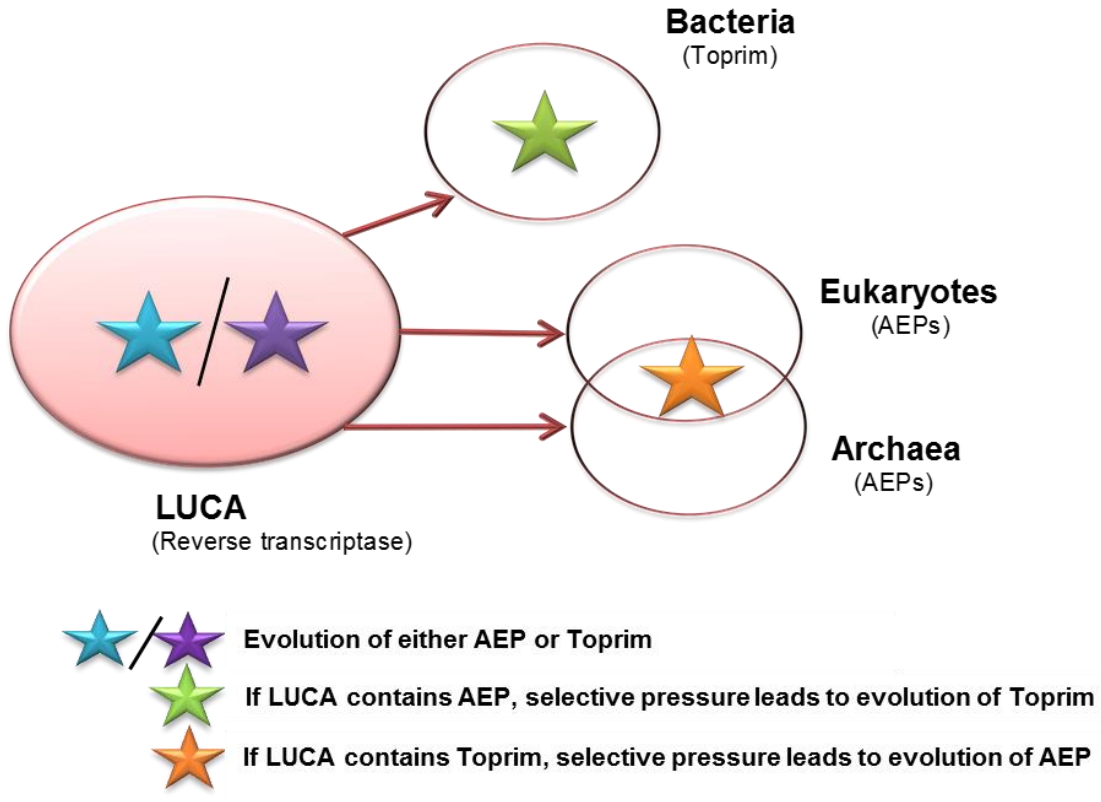


Figure 1.20. Evolution of AEP and Toprim primases (3)

The third model of primase evolution proposed that LUCA was consisting of either AEP or toprim-like primases. Subsequently, unknown evolutionary pressure could have forced the second class of primase to become another type.

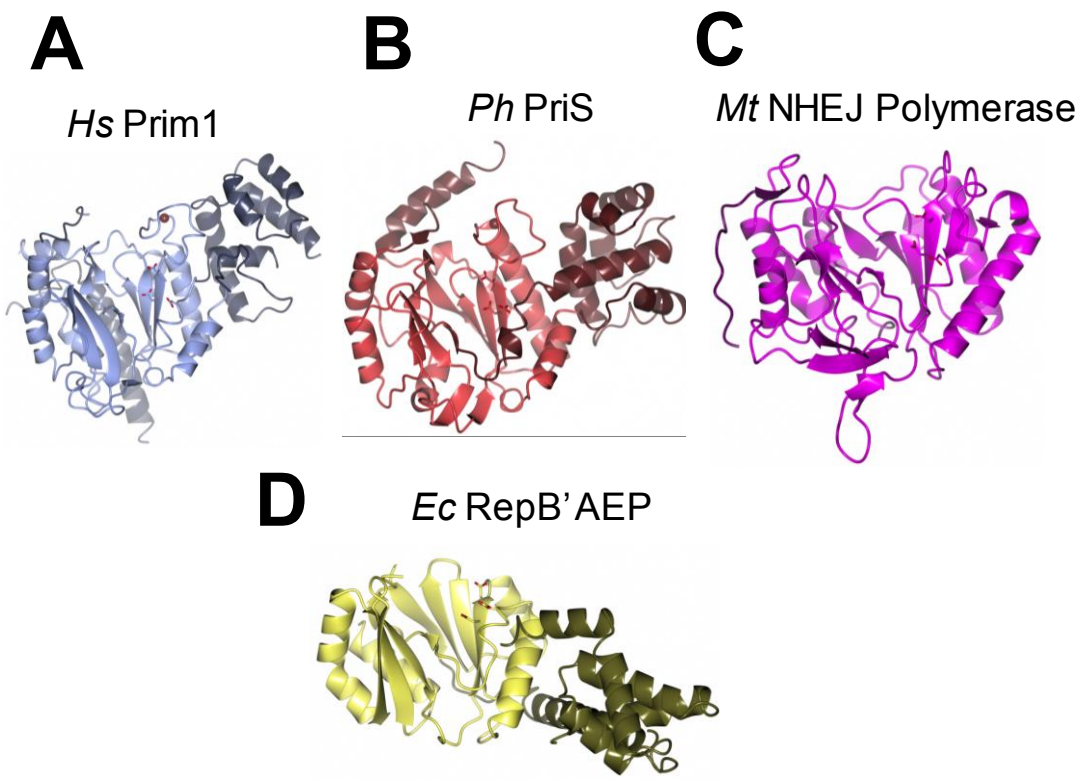


Figure 1.21. AEP catalytic subunit crystal structures from three domains of life

(A) Crystal structure of human primase (Prim1). **(B)** Crystal structure of small catalytic subunit from *Prococcus horikoshi*, a hypethermophillic archaea. **(C)** Crystal structure of PolDom from of *Mycobacterium tuberculosis*. **(D)** The AEP domain of RepB' gene from *Escherichia coli***.** In each structure, the shared catalytic core is demonstrated with lighter colour and catalytic groups are shown as sticks with the acidic oxygen in red. Zinc atoms in the zinc finger domains are shown in tan. (taken from Guilliam *et al* .,2015).

Site-directed mutagenesis studies have revealed the importance of these acidic residues for catalysis (Bianchi *et al*., 2013; Lao-Sirieix and Bell, 2004; Lipps *et al*., 2004; Iyer *et al*., 2005). Lipps and colleagues proposed that mutation of acidic residues in Prim-Pol, a member of AEP superfamily, which will be discussed later, abolishes both polymerase and primase activities (Lipps *et al*., 2004). Comparison of the amino acid sequences, proposed a significant sequence homology between AEP superfamily and PolX superfamily which includes nucleotidyl transferases (Aravind and Koonin., 1999; Kirk and Kuchta., 1999). To date, all characterized DNA primases consist of a zinc binding motif which plays a key regulatory role in sequence-specific recognition of the DNA template by primase (Ilyina *et al*., 1992; Pan and Wigley, 2000). Some AEPs in addition to the AEP polymerase domain contain zinc-binding domains and helicases domains. Several studies suggested variety of roles for AEPs including, terminal transferase, strand displacement, translesion DNA synthesis (TLS) and gap-filling synthesis (Figure 1.22) (Della *et al*., 2004; Pitcher *et al*., 2007; Zhu *et al*., 2006; Guilliam *et al*, 2015).

1.10.4. Eukaryotic AEP primases

The eukaryotic primase is a heterodimer consisting of a small catalytic subunit, 49 KDa (PriS, Prim1 or p48) and a large non-catalytic subunit, 58 KDa (PriL, Pri2 or p58). Eukayotic primase together with DNA polymerase α (Pol α or p180) and the B subunit (p70) form a polα/primase complex (Conaway and Lehman, 1982b; Wang,1991).The Polα-Primase complex encoding genes have been identified in humans, rats, mice, Drosophila, and yeast (Frick and Richardson, 2001).

Although no enzymatic activity has been characterized for PolB subunit, PriS synthesises RNA primers and Polα extends these primers with DNA. PriL is a non-catalytic subunit, which is predicated to regulate the PriS primer synthesis activity (Arezi *et al*., 1999; Zerbe and Kuchta, 2002). This subunit also transfers a primer of a 10 -15 nt in size to Polα. Subsequently, the Pol α extends the RNA primer and generates a RNA-DNA primer of ~30-40 nt in length. Co-immunoprecipitation studies revealed that the interaction between primase and Polα directly mediated by PriL subunit (Longhese *et al*., 1993).

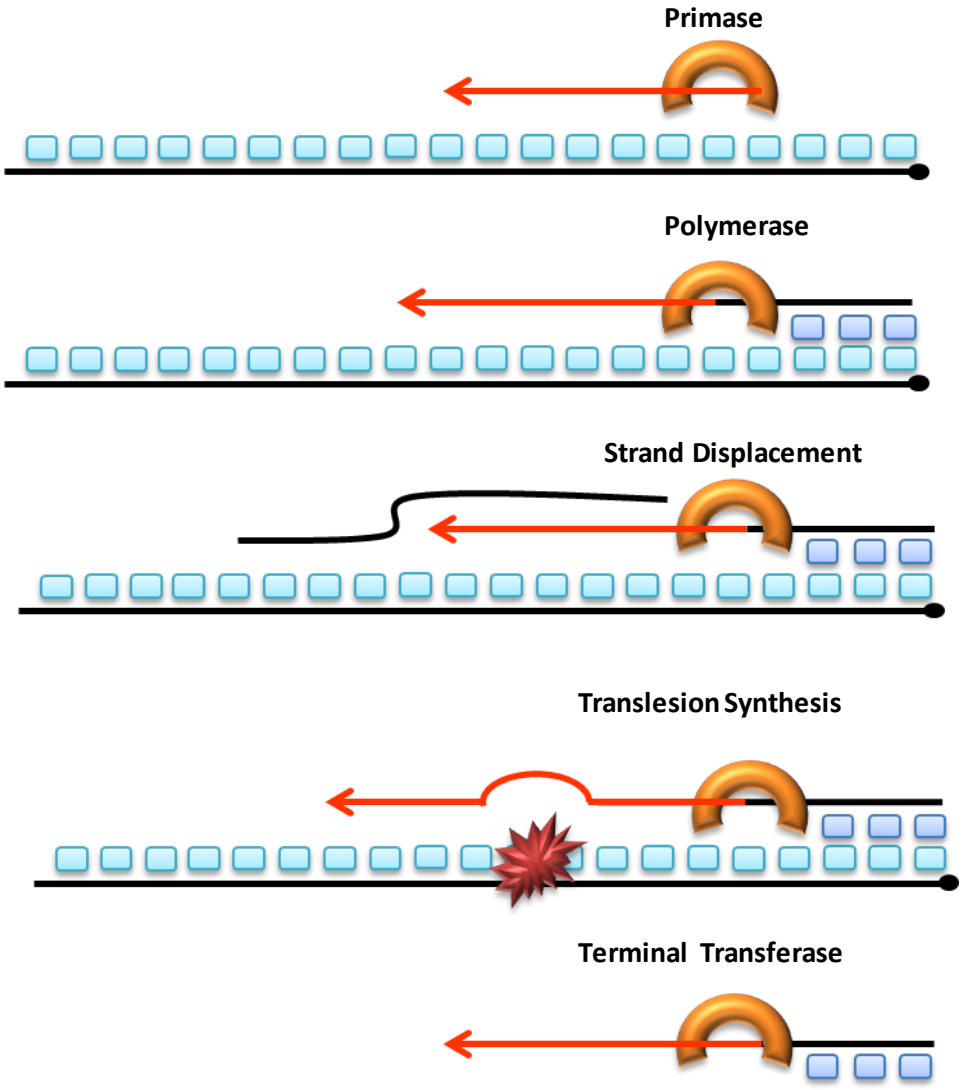


Figure 1.22. AEP primase-polymerases possess a variety of nucleotidyl transferase activities.

In addition to priming DNA, distinct AEPs can perform a range of nucleotidyl transferase activities such as DNA-dependent DNA polymerase, DNA-dependent RNA polymerase, strand displacement, translesion synthesise and terminal transferase activities.

The interaction between PriL and PriS is enabled by a conserved protein-protein interface, located in N-terminal domain of PriL (Lao-Sirieix *et al*., 2005). There is a [4Fe-4S] cluster in the C-terminal domain of PriL, coordinated by four conserved cysteine ligands. Based on biochemical studies, while single cysteine substitutions only slightly decrease the Fe binding and iron-sulfur cluster content, mutation of two cysteines out of four prohibits iron binding and can significantly affect the initiation of RNA synthesis *in vitro* (Liu and Huang.,2015). Iron-sulphur cluster binding domains are also found in other DNA processing enzymes including XPD (Rad3 in *Saccharomyces cerevisiae*) and FancJ helicases of Nucleotide Excision Repair (NER) (Liu *et al*., 2008) and DNA glycosylases of Base Excision Repair (BER) pathway (Messick *et al*., 2002).

1.10.5. AEPs involve in NHEJ

Several studies have identified eukaryotic homologues of the Ku protein in prokaryotes. Ku is a heterodimer that binds to the ends of DNA double strand breaks and is required for end-joining repair in eukaryotes. Soon after the characterisation of archaeal primases as template-dependent polymerases, the AEP homologues was discovered in bacteria (Koonin *et al*., 2000; Doherty, *et al*., 2001; Weller and Doherty, 2001; Weller *et al*., 2002). Interestingly, these AEP genes were often co-operonically associated with the Ku gene (Koonin *et al*., 2000), which suggested the existence of a conserved NHEJ pathway in prokaryotes. Homologues of NHEJ pathway in bacteria and archaea consists of another protein known as Ligase D (LigD). In mycobacteria, LigD consist of a phosphoesterase (PE) domain, LigD domain, and an AEP primase- polymerase (PolDom) domain. Together, these findings shed a light on our understanding regarding additional and unexpected roles of AEPs in DNA metabolism (Guilliam *et al.*, 2015). In contrast to replicative AEPs, PolDom is not able to synthesis *de novo* primers from single nucleotides opposite a DNA template (Brissett *et al*., 2007).

Although both NHEJ AEPs and replicative primase (PriS) share a conserved catalytic architecture, NHEJ AEPs show some distinctive adaptations that distinguishes them from related enzymes. These primase-polymerases have some distinctive DNA binding modes that assist them to function even at the extreme ends of DNA. They are also able to bind specifically to a 5’ phosphate

at the end of DSB using a positively charged surface pocket to accomplish the end processing. Crystal structures of *Mycobacterium tuberculosis* PolDom uncovered the ability of this domain to mediate the synapsis of two non-complementary DNA ends through variety of interactions to form a stable microhomology-mediated end joining (MMEJ) (Figure 1.23) intermediate. (Brissett *et al*., 2007, 2011, 2013; Bartlett *et al*., 2016). During MMEJ process, two PolDom polymerases bind to each side of the break and specific surface loops then promote break annealing. Eukaryotes family X polymerases perform gap-filling strand displacement and, terminal transferase activities. They are also involved in some repair pathways such as base excision repair and NHEJ. AEPs share similarities with catalytic and regulatory regions of family X polymerases, specifically pol β (Kirk and Kuchta., 1999). Recently, it was shown that, Pol θ belongs to family A polymerase performs microhomology-mediated end joining (MMEJ) in eukaryotes (Kent *et al*., 2015).

1.10.6. Viral AEP primases

Prokaryotic, eukaryotic, and archaeal AEPs appear to have viral origins. Viruses encode a range of AEPs including, UL52-like primases from herpes simplex viruses, D5-like primases from NCLDVs and Lef-1 primases from phage and baculoviruses.

Herpes virus type 1 encodes a heterotrimeric protein complex consisting of UL5-UL52-UL8. This three-subunit complex shows both helicase and primase activities. The complex may prime lagging strand synthesis while unwinding the viral DNA replication fork (Crute and Lehman, 1991). UL5 is required for the helicase activity (Gorbalenya *et al*., 1989) while UL52 is an AEP that facilitates priming activity (Crute and Lehman, 1991). In contrast to two other subunits, the UL8 subunit that is related to B-family polymerases appears to have lost its catalytic activity (Marsden *et al*., 1997; Sherman *et al*., 1992). This can be due to the evolution as it was suggested that evolution led to elimination of active site and most of DNA-binding motifs of UL8 (Kazlauskas and Venclovas, 2014). Interestingly, there are some evolutionary connections between UL5-UL52 components and a newly identified human primase (PrimPol), which belongs to AEP superfamily. This primase is discussed in next section (Section 1.9.7).

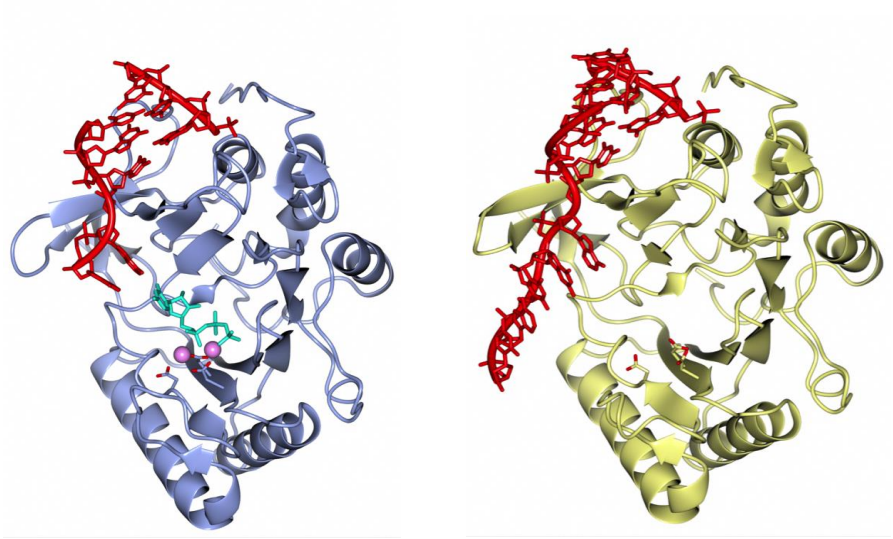


Figure 1.23. Examples of AEPs crystal structures bound to DNA substrates.

(A). Crystal structure of NHEJ repair polymerase (PolDom/LigD-Pol) in a pre-ternary catalytic conformation which is attached to a double-stranded DNA break with a 3' overhanging from *Mycobacterium tuberculosis*. The cyan colour represents UTP and pink represents manganese cofactors. **(B).** micro-homology mediated end-joining (MMEJ) intermediate crystal structure representing synapsis of DBS ends by NHEJ repair polymerase. (Taken from Guilliam *et al.,* 2015).

UL5-Ul52 possess polymerase and the Pif1 helicase activates (Iyer *et al*., 2005). Similar to prokaryotic and eukaryotic primases, UL52 protein possess a putative zinc finger at its C terminal that is required for catalytic activity *in vivo* (Biswas and S. K. Weller, 1999).

Another group of viral AEPs are present in poxviruses. DNA replication of this group of viruses takes place within the infected cells of cytoplasm. D5 is an AEP/helicase fusion protein which has been characterised from vaccine virus (VACV). This enzyme which is conserved across all viruses has a C-terminal domain in charge of helicase activity and an N-terminal domain with sequence and structural similarities with AEPs. *In vitro* studies indicated primase activity for this enzyme during DNA replication (De Silva *et al*., 2009).

Baculoviruses encode another clade of viral AEPs known as Lef-1-like primases. These enzymes are more similar to the "proper" clade that includes replicative and NHEJ AEPs (Iyer *et al*., 2005). Like archaeal replicative primase (PriS) (section 1.9.8.), Lef-1-like primases are able to produce RNA primers that are extended up to several kilobases in length (Mikhailov and Rohrmann, 2002). Unlike previously discussed viral AEPs, the gp43-like proteins from corynephage BFK20, do not show RNA primase activities. These proteins are only capable of deoxyribonucleoside triphosphate incorporation (Halgasova *et al*., 2012).

1.10.7. Discovery of second eukaryotic AEP, PrimPol

The replicative primase, PriS / Prim1 is not the only AEP in the eukarya. Bioinformatics searches and *in silico* analysis of the AEP superfamily detected a second AEP-like protein in eukaryotes, called Primase-Polymerase, or PrimPol (also known as CCDC111, FLJ33167, EukPrim2 or hPrimpol1) (Iyer *et al*., 2005). Previously, PrimPol was reported as a novel uncharacterized member of NCLDV-that herpes virus clade of AEPs (Iyer *et al*., 2005). PrimPol exists in most of of unicellular and multicellular eukaryotes, from planktonic, protists and algae to plants, animals, and mammals. However, PrimPol has not been found in some eukaryotes including *Caenohabditis elegans* and *Drosophila melanogaster*. Based on this distribution it has been suggested that PrimPol was obtained from a viral source through horizontal gene transfer and lost in several occasions in several organisms during evolution (Iyer *et al*., 2005).

PrimPol has an N-terminal AEP domain and a conserved C-terminal UL52-like zinc finger domain. The N-terminal domain is consisting of three distinct catalytic motifs of an AEP, motif I shares the consensus sequence LYFDLE, motif II shares the consensus sequence SxH and motif III the consensus sequence VD (Iyer *et al*., 2005). To reflect the ability of this enzyme to catalyse both primase and polymerase activities, it was named as PrimPol (Bianchi *et al*.,2013; Garcia-Gomez *et al*., 2013; Keen *et al*., 2014; Rudd *et al*.,2014).

A recent study has shown that PrimPol is also a competent TLS polymerase. It was indicated that PrimPol efficiently bypass a number of replication stalling damages including 6-4 photoproducts, an UV-induced lesion, and 8-oxo-dG, an oxidative lesion (Bianchi *et al*., 2013). Similar to many diverse DNA primases, PrimPol requires a zinc finger domain for its primase activity but notably this conserved element is not essential for PrimPol's polymerase activity (Keen *et al*., 2014). Although PrimPol polymerase activity is not dependent on the zinc finger domain, it has been shown that this element can affect the fidelity and processivity of DNA synthesis (Keen *et al*., 2014). A catalytically active fragment of human PrimPol, which contains the AEP domain ($PrimPol_{1–354}$), unlike the wild type enzyme, is able to bypass CPD lesions (Keen *et al*., 2014).

Due to the presence of PrimPol in both nucleus and mitochondria, it has been believed that it might play a role in damage tolerance in both cellular compartments. Deletion of PrimPol encoding gene leads to mitochondrial DNA defects (Bianchi *et al*., 2013). Furthermore, it was reported that the primase activity of PrimPol during replication plays a significant role in re-initiation of stalled replication fork, as the enzyme is able to bypass most UV irradiation-derived lesion through repriming event (Keen *et al*., 2014; Mouron *et al*., 2013). PrimPol is also able to bypass G4 structures and chain terminating nucleosides analogues by repriming downstream of these blockages (Schiavone *et al*., 2016; Kobayashi *et al* 2016).

Recent work has shown that human PrimPol interacts with the major cellular SSBs, RPA and mtSSB (Guilliam *et al*., 2015). This investigation showed that unlike replicative polymerases, which are regulated with mono-ubiquitinated PCNA, both the primase and polymerase activities of PrimPol are significantly modulated by these SSBs (Guilliam *et al*.,2015). PrimPol is a highly mutagenic

enzyme with a strong bias toward insertion-deletion (INDEL) errors therefore it likely to be tightly regulated in the cell and it has been proposed that SSBs may help to regulate its recruitment and activities at sites of stalled replication.

1.10.8. Archaeal AEP primases

As discussed earlier, archaea and eukaryotes share most of the core components of DNA replication, and DNA primases are no exception. The first archaeal primase to be biochemically characterized was from the thermophilic euryarchaeon *Methanococcus jannaschii* (Mjpri). Genome sequence analysis identified an open reading frame encoding a homologue of eukaryotic DNA primase (Desogus *et al*., 1999). This study demonstrated the ability of Mjpri primase to synthesise *de novo* oligoribonucleotides on homopyrimidine ss DNA templates [poly(dT) and poly(dC)] in the presence of divalent cations such as Zn^{2+}, Mg^{2+} and Mn^{2+} (Desogus *et al*., 1999). Soon after, a homologue of p48 eukaryotic primase (*Pfu*p41) was found in a hyperthermophilic euryarchaeote *Pyrococcus furiosus* (Bocquier *et al*., 2001). This primase was shown to prefer using dNTPs as substrates to synthesis DNA primer on non-synthetic template, the M13mp18 (Bocquier *et al*., 2001). Identification of p58-like protein (*Pfu*p46) in *P. furiosus* led to biochemical characterization of the *Pfu*p41-*Pfu*p46 as a stable primase complex in archaea. In contrast to *Pfu*p41, which can only synthesis oligodeoxynucleotide using dNTPs on ssDNA, the Pfup41-Pfup46 complex were able to synthesis RNA primer on M13 single-stranded DNA. In addition, the *Pfu*p41-*Pfu*p46 complex showed more efficient DNA binding activity, higher rates of DNA synthesis and shorter DNA fragments, compared to the catalytic *Pfu*p41 subunit alone (Liu *et al*., 2001). The properties of DNA primase complexes isolated from different archaea vary widely. Similar findings to *P. furiosus* primase have been reported for the primase complex isolated from *Pyrococcus abyssi* (Le Breton *et al*., 2007). Previous study on *Thermococcus kodakaraensis* suggested poor RNA synthesis with rNTPs by the catalytic subunit alone. This primase produced longer DNA fragments than those formed by the complex. *T. kodakaraensis* primase complex showed a preference for dNTPs over rNTPs to synthesis primer (Chemnitz Galal *et al*., 2012). While binding constant is in a same range for both dNTPs and rNTPs in Pyrococcus abyssi, primases from other archaeal species such as *Pyrococcus horikoshii, Solfolobus solfataricus*

and Pyrococcus furiosus prefer using dNTPs rather than rNTPs to form primer (Lao-Sirieix and Bell, 2004; L. Liu *et al*., 2001). Rest of the euryarchaeal organisms such as *Methanothermobacter thermoautotrophicus* (*Methanobacterium thermoautotrophicum*), *Archaeoglobus fulgidus*, and *Halobacterium*, consist of open reading frames (ORFs) homologous to *Pfu*p41-*Pfu*p46.

It is important to note that, unlike eukaryotic primases which form a primase-ssDNA-NTP-NTP quaternary complex to initiate primer synthesis, some of the archaeal primase complexes initiate chains using dNTPs and are able to extend primers using both rNTPs and dNTPs, with a marked preference for dNTPs. The size of DNA primer extension product by archaeal primases can vary between >500 bases in length to <6 kilobases (Bocquier *et al*., 2001; Chemitz Galal *et al*., 2012b; Lao-Sirieix and Bell, 2004; Le Breton *et al*., 2007; Liu *et al*., 2001). These studies showed the p41-p46 complex, in different species of archaea, is responsible for primer formation during DNA replication as in the eukaryotic polymerase α-primase complex (Liu *et al*., 2001).

Different studies have suggested that archaeal DNA primases in addition to primer synthesis activity, fulfilling multifunctional roles from DNA replication to DNA repair. The primase from crenarchaeon *Solfolobus solfataricus* showed the ability to synthesis greater than template length primers via terminal transferase-like activity (Lao-Sirieix and Bell, 2004; Guilliam *et al*., 2015). Terminal transferase is an enzyme which adds nucleotides to the 3' end of a DNA in a template-independent manner. This activity was also observed in *M. jannaschii*. There is now significant evidence accruing that AEPs play roles in DNA damage tolerance and repair, including the role of closely related NHEJ AEPs in break repair (Weller *et al*., 2002; Della *et al*., 2007; Bartlett *et al*, 2013, 2016). It has been reported that *Solfolobus solfataricus* primase shares a common active site architecture with family X DNA polymerase, the family that possess strand displacement and gap-filling activities (Lao-Sirieix and Bell, 2004). Given that the family X DNA polymerase is absent in archaea, it is not implausible that PriSL could have a role in DNA damage repair. As discussed in section 1.4, cells develop different cellular pathways to rescue the stalled replication fork, including translesion DNA Synthesis (TLS) or error-free bypass mechanisms. Despite the

high degree of conservation between eukaryotic and archaeal replisome machinery, the majority of archaeal species lack canonical TLS enzymes (e.g., Y family DNA polymerases). Among those archaeal species, which encode translesion DNA synthesis polymerases the best characterized example is Dpo4 (DinB), a member of Y family DNA polymerases from *S. solfataricus* (Kulaeva *et al*., 1996; Kelman and White., 2005). Furthermore, the existence of a nucleotide excision repair (NER) or photolyase pathways in archaea is still unclear. Only some mesophilic archaea like *M. thermautotrophicus,* possess a complete set of UvrABC (bacterial excision pathway) genes, but most archaeal species lack one or more eukaryal-type NER enzymes (Table 1.3) (Kelman and White., 2005). In addition, an alternative pathway analogous to mismatch excision repair that exist in bacteria and eukarya has not yet been identified in archaea. These differences lead us to question how archaeal species, specifically those reside in extreme environmental conditions (such as high temperature) tolerate DNA damage in the absence of TLS or lesion repair pathways. Deamination of cytosine to uracil can lead to G.C to A. T transition mutation if not repaired (Greagg *et al*., 1999). Since this process sped up in high temperature, hyperthemophilic archaeae are in high risk of cytosine deamination (Lindahl and Nyberg, 1974). Despite the lack of base excision repair pathway in archaea, which repair cytosine deamination by a uracil-DNA glycosylase (UDG) or a G·U/T mismatch-specific DNA glycosylase (TDG/MUG) in bacteria and eukarya (Greagg *et al*., 1999), archaea utilize a specific uracil detection system to stop replication. In Archaea, replicative polymerases (B and D Pols) stall at deoxyuracil containing templates due to the presence of a uracil binding pocket located in their amino-terminal domain (Richardson *et al*., 2013; Firbank., 2008). This poses a question as to how archaea restart replication after dU-induced replisome stalling.

The overall aims of this thesis were to investigate the biochemical properties of a number of archaeal PriS/L complexes, predominantly from *Archaeoglobus fulgidus,* and investigate the roles this enzyme system play in standard DNA replication. The thesis commences with the characterization of enzymatic and DNA-binding activities of the DNA primase complex (PriS/L) isolated from three different archaeal species, two hyperthemophiles and one mesophile, using biochemical methods (Chapter 3). In Chapter 4, lesion bypass proficiencies of archaeal replicative primases on damaged DNA templates as well as, measuring

the fidelity of PriS/L using standard primer-extension assay are explored. Chapter 5 reports the regulatory roles of single-stranded binding proteins on polymerase activity of PriS/L and interrogates the possible interactions between RPA and PriS/L. Chapter 6 describes attempts to reconstitute an archaeal CMG complex *in vitro*, with the aim of shedding light on the role of archaeal replicative primase in replication-specific TLS. This chapter also introduces initial observations regarding the potential interaction between PriS/L and RadA recombinase.

Species	XPF	XPB	XPD	Fen1	Dpo4	Photolyase	UvrABC	Muts/MutL	Temp(°C)
Aeropyrum pernix	1	1	1	1					95
Pyrobaculum aerophilum	1	2	1	1					95
Sulfolobus solfataricus	1	2	1	1	1	1			80
Sulfolobus tokodaii	1	2	2	1	1	1			80
Sulfolobus acidocaldarius	1	2	1	1	1	1			75
Methanopyrus kandleri	1			1					100
Pyrococcus furiosus	1	2	1	1					100
Pyrococcus horikoshii	1	2	1	1					100
Pyrococcus abyssi	1	2	1	1					100
Thermococcus	1	1	1	1					85
Kodakarensis									
Nanoarchaeum equitans	1	1	1	1					85
Methanococcus jannaschii	1		1	1					85
Archaeoglobus fulgidus	1	1		1					85
Methanothermobacter	1			1		1	*		70
thermautotrophicus									
Picrophilus torridus		1	1	1	1	1			60
Thermoplasma		1	1	1					60
Acidophilum									
Thermoplasma volcanium		1	1	1					60
Methanosarcina	1	1	1	1	1	1	*	*	<40
Acetivorans									
Methanosarcina mazei	1	1	1	1	1	1	*	*	<40
Haloarcula marismortui	1	2		1	1	3	*	*	<40
Halobacterium NRC-1	1	2	1	1	1	2	*	*	<40

Table 1.3. Distribution of archaeal DNA repair proteins.

The table is started with crenarchaea, followed by euryarchaeal species, ranked by growth temperature. The circles represent the presence of bacterial repair genes in archaeal genomes and the numbers indicate number of each repair gene in each species (Kelman anad white., 2005).

Chapter 2

Materials and Methods

2.1. Molecular Biology Methods

2.1.1. Preparation of competent *E.coli* DH5α

Lysogeny broth agar plates (1 % (w/v) tryptone, 0.5 % (w/v) yeast extract, 1 % (w/v) NaCl, pH 7) were used for plating DH5α cells. One colony of DH5α cells selected from LB plates was inoculated into 3 mL of LB medium and incubated overnight at 37°C with shaking at 180 x *g*. 250 mL super optimal broth (SOB) medium (2% (w/v) tryptone, 0.5 % (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 Mm $MgSO_4$ and 10 mM $MgCl_2$) was then used to dilute the culture. Diluted culture was incubated at 18°C with shaking. The optical density measured at 600 nm (OD_{600}) was 0.4. Cells were cooled down on ice for 10 minutes and then collected by centrifugation (4,000 x *g*, 10 minutes, 4°C, Sorvall Legend RT, 75006445 rotor). 80 mL precooled transformation buffer (100 mM PIPES (pH 6.7), 15 mM $CaCl_2$, 250 mM KCl, 55 mM $MgCl_2$) was used to resuspend the cells. After incubation on ice for 10-30 minutes, centrifugation carried out and then cells resuspended in 20 mL pre-cooled transformation buffer supplemented with 7% (v/v) dimethyl sulphoxide (DMSO). After 10 minutes incubation on ice, 50-100 µL aliquots of DH5α cells were frozen in liquid nitrogen and stored at -80°C.

2.1.2. Transformation of competent DH5α

50 µL of chemically competent DH5α cells (section 2.1.1) was defrosted on ice. 1 µL of plasmid DNA from a Miniprep (~ 50-100 ng) (section 2.1.3) or few µL of a ligation reaction (section 2.1.7) was mixed and incubated with 50 µL competent cells. After 20 minutes incubation on ice, the mixture was heat shocked at 42°C for 35 second and then returned to ice for further 2 minutes. 500 mL SOB medium was added to the cell and DNA mix following by one hour incubation at 37°C with shaking. One LB agar plate containing proper antibiotics was pre-warmed and then 100 µL of the cell growth was plated on it. Subsequently, the plate was incubated at 37°C overnight to allow colonies to grow. 100 µg/mL and 34 µg/mL final concentrations of ampicillin and kanamycin antibiotics were used.

2.1.3. Plasmid DNA amplification and purification

After successful transformation of plasmid DNA into DH5α cells (section 2.1.2) one colony was selected and inoculated into 5 mL of LB medium with appropriate antibiotic. Overnight incubation at 37°C with shaking at 180 x *g* was followed by purification of the plasmid DNA using the QIAprep Spin Miniprep Kit (Qiagen) according to the manufacturer's instructions. Plasmid DNA was eluted in 50 µL. ND-1000 NanoDrop spectrophotometer (Thermo Scientific) was used to measure DNA yields at a wavelength of 260 nm. Finally, the DNA stored at -20°C.

2.1.4. Agarose gel electrophoresis of DNA

DNA samples were loaded in DNA loading buffer (5% (v/v) glycerol, 3.3 mM Tris pH 8, 0.04% bromophenol blue) and then resolved on a 1% (w/v) agarose TAE (200 mM Tris, 100 mM Acetic Acid, 5 mM EDTA) gel supplemented with ~0.3 µg mL^{-1} ethidium bromide. DNA samples were resolved alongside a 1 kb (New England Biolabs) or a 100 bp (New England Biolabs) DNA ladder and electrophoresed at 120 V for~ 20-30 minutes. To visualise the DNA, UV illuminator (Syngene InGenius Gel Analysis System) was used and analyses of images was carried out using GeneSnap (Syngene).

2.1.5. Polymerase Chain Reaction (PCR)

PCR was performed using ~5 ng of template plasmid or 100-200 ng of genomic DNA, 1 mM dNTPs, 10 mM $MgCl_2$, 20 % DMSO, 1 U Phusion high-fidelity DNA polymerase (New England Biolabs) and 1 µM of each of a forward and reverse primer. The primers were designed typically to have a Tm of~ 55-65°C. Primers were flanked by relevant restriction digestion sites. Reactions were assembled into 200 µL tubes in 50 µL final volume. Using a Professional Trio Thermocycler (Biometra) the reactions were initially denatured for 1 minutes at 98°C before 30 cycles of 95°C for 30 seconds, annealing temperature, which was typically 3°C above the Tm's of the primers, for 30 seconds, elongation at 72°C for 1 minutes and a final elongation of 72°C for 3 minutes. The elongation time was depended on the desired DNA's size. To confirm the PCR, agarose gel electrophoresis (section 2.1.4) was carried out and PCR products were detected using a UV transilluminator (UVP). Confirmed PCR product was purified using the QIAquick

Gel Extraction kit (Qiagen) according to the manufacturer's instructions. The nucleotide sequence of primers used can be found in Table.2.1.

2.1.6. Restriction Digestion

In order to generate a construct, 20 µL of Miniprep plasmid DNA (~2 µg) or 30 µL of PCR product was digested with 10 U of the appropriate restriction enzymes (New England Biolabs) in a 50 µL reaction containing 100 µg/mL bovine serume albumin (BSA), and the proper reaction buffer. Following incubation at 37°C for 3-4 hours, to prevent recircularisation of the plasmid, the digested plasmid DNA was treated with Antarctic phosphatase (New England Biolabs) and incubated with Antarctic phosphatase buffer (50 mM Bis-Tris-propane-HCl pH6, 1 mM $MgCl_2$, 0.1 mM $ZnCl_2$) for further 30 minutes. The digested plasmid DNA was purified using the OIAquick PCR purification kit (Qiagen) according to the manufacturer's instructions.

2.1.7. Ligation of DNA

In a final volume of 20 µL reaction, restriction digested plasmid DNA and PCR products were ligated using 200 units T4 DNA ligase, and T4 DNA ligase buffer (50 mM Tris HCL pH 7.5, 10 mM $MgCl_2$, 1 Mm ATP, 10 mM DTT). Following 1 hour or overnight incubation at room temperature or 12°C respectively, 2 µL of reaction was transformed into DH5α (section 2.1.2)..

To determine the concentration of digested DNA the ND-1000 NanoDrop spectrophotometer (Thermo Scientific) was used, and concentration measured at a wavelength of 260 nm. Ligation reactions were set up at a ratio (n) of 6:1 or 3:1 of insert (I) to vector (V) using the following equation: $I_{(ng)} = nx [I_{(bp)} / V_{(bp)}] x V_{(ng)}$

2.1.8. Sequencing of DNA products

To confirm successful cloning, sequencing of plasmid DNA was performed by GATC biotech using universal primers or gene-specific primers. The sequencing chromatogram was read using 4peaks (Mekentosj).

#	Primer		Sequence
1	Afu-PriS	FP	5-GTTTCTTCATATGGCAGCAGGTTGTGATTATCAACTTCG-3'
2	Afu-PriS	RP	5-GTTTCTTCTCGAGTTAGGAATCGTAGCTTGCATCCCCTCTGCAAATC-3'
3	Afu-PriL	FP	5-GTTTCTTGAATTCGATGAAATACCTACCCCTTTACCCAATTTTG-3'
4	Afu-PriL	RP	5-GTTTCTTGCGGCCGCTCAACTTTTCGATTTACATTTTTTATAATAAATTAAG-3'
5	Pfu-PriS	FP	5-GTTTCTTCATATGCTGATGAGGGAAGTGACAAAGG-3'
6	Pfu-PriS	RP	5-GTTTCTTCTCGAGTTATTCATATTCCAAGGACTCTCTCCACAGTTTATAAG-3'
7	Pfu-PriL	FP	5-GTTTCTTGGATCCGTTAACTTCCATTCTCCCACCTCCATTAAAG-3'
8	Pfu-PriL	RP	5-GTTTCTTGCGGCCGCTTACTGTAGAATTCGCTCCTTTCCTCCTTTG-3'
9	Mmp-PriS	FP	5-GTTCTTTCATATGAACGACAATCCTGCCACAGATAAG-3'
10	Mmp-PriS	RP	5-GTTTCTTGCGGCCGCTTATTCTAAAATTTCAAAGTCGATACC-3'
11	Mmp-PriL	FR	5-GTTTCTTGAATTCGATGATTTTAGTGTTTTTAGACAATATGG-3'
12	Mmp-PriL	RP	5-GTTTCTTGCGGCCGCTTATTCTGATTTTTTTACATGTCTGTTAATG-3'
13	Afu-PolB	FR	5-GTTTCTTCATATGGAAAGAGTTGAGGGCTGGCTCATCG-3'
14	Afu-PolB	RP	5-GTTTCTTGCGGCCGCTTAGCTGAAGAATGAATCCAGGCTCATCTG-3'
15	Afu-PolD1	FP	5-GTTTCTTGAATTCGATGGATGCAACTCTTGACAGGTTCTTC-3'
16	Afu-PolD1	RP	5- GTTTCTTGCGGCCGCTCAAACGAAATCGGATATTGACACTTG-3'
17	Afu-PolD2	FP	5- GTTTCTTCATATGGTAATTAAAAATATCGATGCCGCAACAG-3'
18	Afu-PolD2	RP	5- GTTTCTTCATATGGTAATTAAAAATATCGATGCCGCAACAG-3'
19	Afu-MCM	FP	5-GTTTTCTTGGATCCATGGGTATAAGCAGTCCGGCACTTTG-3'
20	Afu-MCM	RP	5-GTTTCTTCTCGAGTTAAAGTTTGCTTACCAATTTGTAATG-3'
21	Afu-GINS	FP	5-GTTTCTTCATATGAACCTGGATGAGCTTCTG-3'
22	Afu-GINS	RP	5-GTTTCTTCTCGAGTTAACCTTCTTGGGATACTTC-3'
23	Afu-RecJ99	FP	5-GTTTCTTCATATG – AGTTCGGGCTCGGCAATTAACCTGC-3'
24	Afu-RecJ99	RP	5-GTTTCTTCTCGAG-CTAAGCGAGCATGGCACAGCAAAGCC-3'
25	Afu-RecJ98	FP	5-GTTTCTTGGATCCGC-TTGCTGTGCCATGCTCGCTTAGAAGTAGAATCC-3'
26	Afu-RecJ98	RP	5-GTTTCTTGCGGCCGC-TCAGCCGAGAACCTCAAGAACTCCTTCAC-3'
27	Afu-RPA780	FP	5-GTTTCTGGATCCATGTATAATCCTTTTCACTTGCATCCGGGGGTCGAAG-3'
28	Afu-RPA780	RP	5-GTTTCTTGCGGCCGCTCAGAACACCTCCAGCGTGAGCTGTCCAAC-3'
29	Afu-RPA-382	FP	5-GTTTCTTGGATCCATGGGGGCTGTTTCTGCTTTTTCTTGGGCTTC-3'
30	Afu-RPA-382 PR		5-GTTTCTTGCGGCCGCCTACTCCCCAACACTGAACCTCGAATTTCTATC-3'
31	Afu-RadA	FP	5-GTTTCTTCATATGAGTGAAGAGAGCAACGAAGAG-3'
32	Afu-RadA	RP	5-GTTTCTTCTCGAGTTATTTTTTCCTTTTCTTTTTGTCTTTCTCTTCAGCATCCTCTATTC-3'

Table 2.1. Primers used in PCR to produce expression vectors.

The designed primers listed here are used in this thesis for the construction of expression vectors listed in Table 2.3.

2.1.9. Yeast two hybrid methods

To detect interacting proteins Yeast two-hybrid system (Y2H) was performed. The *S. cerevisiae* Y190 strain was used in this system. pGBKT7 (containing the GAL4 DNA binding domain) and pGADT7 (containing the GAL4 DNA activation domain) were used as the vectors in the two-hybrid assay. Plasmid constructs used in this thesis are detailed in Table 2.2.

2.1.9.1. Yeast culture

Y190 yeast strain was grown in Yeast Extract medium (YE: 0.5% (w/v) yeast extract, 3% (w/v) glucose, 0.02% (w/v) adenine, and 0.1% (w/v) uracil, histidine, arginine and leucine) or on YEA plates (YE medium solidified with 2.5% granulated agar) at 30°C. Yeast plate was stored at 4°C for short time. For longer storage, one colony from the yeast plate was collected and resuspended in growth media and then supplemented with glycerol to a final concentration of 25% and finally stored at -80°C. To make a new culture, a few microliter of glycerol stock was streaked onto YEA plates following by incubated at 30°C for 3-5 days.

2.1.9.2. Yeast transformation

Few colonies of Y190 were inoculated into 1 mL of YE, vortexed vigorously to disperse clumps. 1 mL culture transferred into 50 mL of YE and incubated overnight at 30°C (250 x *g*) until growth reached stationary phase (OD600 > 1.5). 2 mL of an overnight cell culture was used to inoculate 100 ml of pre-warmed YE medium and incubated at 30ºC for 3-4 hours. After centrifugation, cells were washed with 25ml of sterile water and again subjected to centrifugation. The cells were then resuspended in 1mL of LiOAc buffer (100mM LiOAc, 100mM Tris-HCl, and 10mM EDTA) and incubated at 30ºC for 30 minutes. The yeast transformation competent cells were aliquot in 100µl samples, and spun down again. Next, each pellet was gently resuspended with 50µL LiOAc buffer and mixed with 1-5µg of plasmid DNA, 150µg of single-stranded salmon sperm DNA. Finally the cells and DNA suspension was treated with 700µl of PEG/LiOAc buffer (40% PEG 3,350, 100mM LiOAc, 100mM Tris-HCl (pH 7.5), and 10mM EDTA) and incubated at 30ºC for 30 minutes.

#	Primer	Sequence
1	**Afu-PriS FP**	5'-GTTTCTTCATATGGCAGCAGGTTGTGATTATCAACTTCG -3'
2	**Afu-PriS RP**	5'-GTTTCTTCTCGAGTTAGGAATCGTAGCTTGCATCCCCTCTGCAAATC-3'
3	**Afu-KU FP**	N/A
4	**Afu-KU RP**	N/A
5	**Afu-LigD FP**	N/A
6	**Afu-LigD RP**	N/A

Table 2.2. Primers used to generate plasmid constructs for two-hybrid assay

The designed primers listed here are used in this thesis to generate plasmid constructs listed in Table 2.4. for yeast-two hybrid analysis

The incubation of yeast with plasmid DNA was followed with heat shock step at 42ºC for 20 minutes. Finally, cells were spun briefly, resuspended in 80 µL of sterile water and plated on yeast minimal medium plates (YMM; yeast nitrogen base (YNB), 10 % glucose, 2.5 % granulated agar) supplemented with 20 mL adenine and histidine. After 2-3 days incubation in a 30°C incubator, colonies appear on the plate.

2.1.9.3. Detection of interaction

Colonies of both pGBKT7 and pGADT7 transformed cells were resuspened in 20 µL sterile water. 10 µL of cells were plated on YMM plates containing adenine and histidine and the rest 10 µL were spread on YMM plates containing adenine and 25 mM 3-amino-1, 4-triazole (3-AT). 3-AT is a heterocyclic organic compound which limits biosynthesis of HIS3 protein, therefore inhibiting low-level leaky expression (Durfee *et al*., 1993). The plates were incubated at 30°C for 2-3 days. Cell growth in plates with no histidin (3-AT plates) suggested a potential interaction of the proteins fused in pGBKT7 and pGADT7. Plasmid constructs used in two-hybrid assay are listed in Table 2.3.

2.2. Purification of recombinant proteins

2.2.1. Preparation of chemically competent E.coli strains

In this thesis, chemically competent BL21, SHuffle and Rosetta *E.coli* strains were used for purification of recombinant proteins. All of the chemically competent *E.coli* expression strains were prepared in the same manner as DH5α (section 2.1.1).

2.2.2. Strain optimisation of protein expression

A 50 µL aliquot of chemically competent *E.coli* expression strain was transformed with 2 µL of appropriate expression construct (Table 2.4). A fresh single colony was inoculated into 5 mL of LB medium containing the antibiotic resistance present on the expression construct. The culture was incubated at 37°C overnight with shaking at 180 x *g*. 1 mL of overnight culture was added to 50 mL of LB medium containing the appropriate antibiotic resistance and incubated at 37°C until exponential growth (OD_{600} = 0.4-0.6).

.

#	Gene product	Vector	source	Primers
A1	Afu-PriS	pET28a	PCR-Genomic DNA	1,2
A2	Afu-PriL	pETduet-1	PCR-Genomic DNA	3,4
A3	Afu-PolB	pET28a	Dr. Stanislaw Jozwiakowski	13,14
A4	Afu-PolD1	pETduet-1	Dr. Stanislaw Jozwiakowski	16,17
A5	Afu-MCM	pET28a	PCR-Genomic DNA	19,20
A6	Afu-GINS	pET28a	PCR-Genomic DNA	21,22
A7	Afu-RecJ99	pETduet-1	PCR-Genomic DNA	23,24
A8	Afu-RecJ99/98	pETduet-1	PCR-Plasmid A7	25,26
A9	Afu-RPA-780	pET28a	PCR-Genomic DNA	27,28
A10	Afu-RPA-382	pGEX-6p-1	PCR-Genomic DNA	29,30
A11	Afu-RadA	pET28a	PCR-Genomic DNA	31,32
P1	Pfu-PriS	pET28a	Dr. Stanislaw Jozwiakowski	5,6
P2	Pfu-PriL	pETduet-1	Dr. Stanislaw Jozwiakowski	7,8
M1	Mma-PriS	pET28a	Dr. Stanislaw Jozwiakowski	9,10
M2	Mma-PriL	pETduet-1	Dr. Stanislaw Jozwiakowski	11,12

Table 2.3. Plasmid expression vectors and their associated gene products.

pET28a and pETduet-1 plasmids containing a T7 promoter and a *lac* operon are used as the expression vectors in this thesis. Plasmid A10 is located within the polylinker of the pGEX plasmid for expression. The source is given with polymerase chain reaction (PCR) products indicated and the plasmid that was used as a template. Other scientist donated plasmids A1, A2, A3, and A4. Primers column is the list of primer numbers (#) from Table 2.1 that used to generate constructs.

#	Gene product	Vector	Source	Yeast strain	Genotype	Primers
A12	Afu-PriS	pGBKT7-pGADT7	PCR-Plasmid A1	Y190	MATa, ura3-52, his3-D200, lys2-801, ade2-101,trp1-901, leu2-3, 112, gal4D, gal80D, URA3::$GAL1_{UAS}$-$GAL1_{TATA}$-lacZ, cyh^r2, LYS2::GAL_{UAS}-$HIS3_{TATA}$-HIS3	1,2
A13	Afu-PriL	pGBKT7-pGADT7	PCR-Plasmid A2	Y190	MATa, ura3-52, his3-D200, lys2-801, ade2-101,trp1-901, leu2-3, 112, gal4D, gal80D, URA3::$GAL1_{UAS}$-$GAL1_{TATA}$-lacZ, cyh^r2, LYS2::GAL_{UAS}-$HIS3_{TATA}$-HIS3	3,4
A14	Afu-KU	pGBKT7-pGADT7	Dr. Nigel Brissett	Y190	MATa, ura3-52, his3-D200, lys2-801, ade2-101, trp1-901, leu2-3, 112, gal4D, gal80D, URA3::$GAL1_{UAS}$-$GAL1_{TATA}$-lacZ,cyh^r2, LYS2::GAL_{UAS}-$HIS3_{TATA}$-HIS3	N/A
A15	Afu-LigD	pGBKT7-pGADT7	Dr. Nigel Brissett	Y190	MATa, ura3-52, his3-D200, lys2-801, ade2-101, trp1-901, leu2-3, 112, gal4D, gal80D, URA3::$GAL1_{UAS}$-$GAL1_{TATA}$-lacZ,cyh^r2, LYS2::GAL_{UAS}-$HIS3_{TATA}$-HIS3	N/A

Table 2.4. Plasmid constructs for yeast-two hybrid

Constructs used and generated in this thesis for expression of proteins in *S.cereviase* for yeast-two hybrid assay are listed. The genotype of Y190 yeast strain is also indicated. pGBKT7 (containing the GAL4 DNA binding domain) and pGADT7 (containing the GAL4 DNA activation domain) are used. Plasmids A14 and A15 were donated by other scientist. Primers column is the list of primer numbers (#) from Table 2.1 which used to generate constructs.

The culture was then cooled on ice for 10 minutes. Few µL of uninduced sample was collected and remainder of expression was induced for expression through addition of 400 µM isopropyl β-D-1-thiogalactopyranoside (IPTG). The induced culture was returned to incubator for the desire amount of time. Cells were collected by centrifugation and resuspended in water or lysis buffer (300 mM NaCl, 50 mM Tris-HCl pH 7.5, 30 mM imidazole, 10% (v/v) glycerol, 1 protease inhibitor cocktail tablet EDTA-free (Roche), and 5 mM β-mercaptoethanol). The lysate was then sonicated (Vibrs-cell sonicator) on ice, 3 rounds for 20 second pulses at 30% amplitude. After sonication, cells were centrifuged at 13,000 x *g* for 1 hour at 4°C to clarify the cell lysate. Clarified cell lysate was then suspended in sample buffer, incubated at 95°C for 5 minutes and analysed by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) (section 2.3.1).

2.2.3. Expression of recombinant proteins

Chemically competent *E.coli* expression strain was transformed with appropriate expression construct. One colony was added into 50 mL of LB medium supplemented with appropriate antibiotic resistance and incubated at 37°C overnight. The overnight incubation followed by the 1:100 dilution of the culture into a large volume (1-6 Litres) of fresh LB medium containing appropriate antibiotic resistance and incubated at 37°C with shaking until exponential growth (OD_{600} = ~ 0.4). Growths were then incubated at 4°C for 30 minutes and subsequently 400 µM of IPTG was added for expression induction. Growths were then incubated at temperatures ranging from 16-25°C for 3-16 hours. The induction conditions for each of the proteins in this thesis are listed in Table 2.5.

Cells were spun down by centrifugation (4,000 x *g*, 15 minutes, 4°C, Sorvall Evolution, SLC-6000 rotor) and resuspended in water or lysis buffer (15 mL per litre of culture grown). Lysate was then supplemented with 1 mg/mL of chicken egg lysozyme and stir for 30-45 minutes. Following sonication using a Vibra-Cell sonicator (Sonics) on ice at 30% amplitude for 3 second pulses, with a 10 second rest between pulses, for a total 6 minutes sonication time, the soluble lysate was separated from insoluble cell debris through centrifugation (18,000 x *g*, 1 hour, 4°C, JA 25-50 rotor).

Protein	*E.coli* strain	Genotype	Growth Medium	Induction Temperature and Length
Afu-PriS/L	BL21	B F– ompT gal dcm lon hsdSB(rB–mB–) [malB+]K-12(λS)	LB	20°C overnight
Afu-PolB	Rosetta	F- ompT hsdSB(rB- mB-) gal dcm (DE3) pRARE (CamR)	LB	20°C 4 hours
Afu-PolD1	Rosetta	F- ompT hsdSB(rB- mB-) gal dcm (DE3) pRARE (CamR)	LB	N/A
Afu-MCM	Rosetta	F- ompT hsdSB(rB- mB-) gal dcm (DE3) pRARE (CamR)	LB	25°C overnight
Afu-GINS	Rosetta	F- ompT hsdSB(rB- mB-) gal dcm (DE3) pRARE (CamR)	LB	25°C overnight
Afu-RecJ99	BL21/Rosetta	″	LB	25 /20/16 °C overnight
Afu-RecJ99/98	BL21/Rosetta	″	LB	25 /20/16 °C overnight
Afu-RPA-780	Rosetta	F- ompT hsdSB(rB- mB-) gal dcm (DE3) pRARE (CamR)	LB	25°C overnight
Afu-RPA-382	Rosetta	F- ompT hsdSB(rB- mB-) gal dcm (DE3) pRARE (CamR)	LB	25°C overnight
Afu-RadA	Rosetta	F- ompT hsdSB(rB- mB-) gal dcm (DE3) pRARE (CamR)	LB	25°C overnight
Pfu-PriS/PriL	BL21	B F– ompT gal dcm lon hsdSB(rB–mB–) [malB+]K-12(λS)	LB	20°C overnight
Mmp-PriS/PriL	BL21,Rosetta	″	LB	20°C overnight

Table 2.5. Growth conditions for recombinant proteins

The *E.coli* strains used in the growth of recombinant proteins are listed. The genotype of each strain is also indicated. Lysogeny Broth (LB) was used as the growth medium. *E.coli* growths were induced at 16 °C, 20°C and 25°C.

The soluble cell lysate was next passed through a 0.45 µM filter (Milipore) and then subjected to chromatography columns for protein purification.

2.2.4. Immobilised metal affinity chromatography

In this thesis, the recombinant proteins fues to a 6xHis tag were purified using Immobilised metal affinity chromatography (IMAC). A column with 5 mL or 25 mL of nickel-nitrolotriacetic acid (Ni^{2+}-NTA) agarose resin (Qiagen) was subjected to chromatography using ÄKTAprime (GE Healthcare) system. The column was pre-equilibrated in IMAC buffer A (500 mM NaCl, 20 mM Tris-HCl pH 7.5, 20 mM imidazole, 10% (v/v) glycerol, 5 mM β-mercaptoethanol). The soluble cell lysate was then loaded onto the column at a flow rate of 2 mL/miN. Next, the column was washed with buffer A until all the unbound proteins were eluted which was measured by the absorbance at a wavelength of 280 nm (A_{280}) returning to basal levels. The column was then washed with 10% IMAC buffer B (buffer A with 300 mM imidazole) in order to eliminate weakly bound proteins. The bound proteins were eluted at 2 mL/min with 100% IMAC buffer B. -IPTG, Insoluble fraction, load, flow through, wash, 10% buffer B wash and 100% B elution samples were analysed by SDS-PAGE with Coomassie staining (section 2.3.1 and 2.3.2).

2.2.5. Heparin affinity chromatography

HiTrap Heparin HP (GE Healthcare) column was prepacked with Heparin Sepharose. Heparin binds to a wide range of biomolecules via both affinity and ion exchange mechanism. Heparin sepharose column was pre-equilibrated in Heparin buffer A (150mM NaCl, 20 mM Tris-HCl pH 7.5, 10% (v/v) glycerol, 5 mM β-mercaptoethanol). The IMAC elute was diluted 1:10 into Heparin buffer A to reduce the salt concentration to improve binding. Following loading the protein onto the column at a flow rate of 2 mL/min, the column was washed extensively with Heparin buffer A to elute unbound proteins. Proteins were eluted by gradient elution up to 50% Heparin buffer B (Heparin buffer A with 1 M NaCl). Fractions containing protein were collected and analysed by SDS-PAGE with Coomassie staining (section 2.3.1 and 2.3.2).

2.2.6. Ion-Exchange chromatography recombinant proteins

Ion-exchange chromatography is a chromatography process that separates ions and polar molecules based on their affinity to the ion exchanger. Recombinant proteins were subjected to either 5 mL HisTrap Q Sepharose FF columns (GE Healthcare) or HisTrap SP sepharose FF (GE Healthcare) based on their predicted isoelectric point (pI). Typically, HisTrap Q Sepharose FF column is used as an anion exchanger and HiTrap SP Sepharose FF is used as a cation exchanger. To purify proteins using Ion-exchange chromatography similar process as HiTrap Heparin HP column was performed (section 2.2.5).

2.2.7. Affi-Gel Blue chromatography

Some proteins (Table 2.6) were subjected to Affi-Gel Blue as an affinity chromatography resin. Affi-Gel Blue is a cross-linked agarose bead to which the dye Cibacron Blue F3GA has been covalently attached. The Cibacron Blue dye has ionic, hydrophobic, and aromatic character therefore displays affinity for many type proteins. 5 mL Affi-Gel Blue (Bio-Rad) column was pre-equilibrated with buffer A (50 mM HEPES-NaOH, pH 7.5, 10% glycerol, 10 mM β-mercaptoethanol, and 500 mM NaCl). The column was first washed with buffer A, and then with buffer A containing 1 M NaCl, then with buffer A containing 500 mM NaSCN. Following loading the protein into the column, the protein was eluted with 50 mL of buffer A containing 1.5 M NaSCN. Eluted fractions were pooled and dialyzed against buffer A with 100 mM NaCl, and 0.1% Tergitol-type NP-40. Precipitated materials were removed by centrifugation and the sample was analysed by SDS-PAGE with Coomassie staining (section 2.3.1 and 2.3.2).

2.2.8. Hydrophobic interaction chromatography (HIC)

Hydrophobic interaction chromatography is a technique that separates proteins from one another based on hydrophobicity. HIC is mostly suitable for the the intermediate steps of purification. HiTrap Phenyl FF (GE Healthcare) column was prepacked with phenyl sepharose. The column was pre-equilibrated in HIC buffer A (40 mM Tris-HCl pH 7.5, 500 mM NaCl, 1 M ammonium sulfate and 5 mM β-mercaptoethanol).

Protein	Affinity			Ion exchange	Hydrophobicity	Gel filtration
	Ni-NTA IMAC	Heparin	Affi-Gel Blue	Q		
Afu-PriSL	✓	✓				✓
Pfu-PriSL	✓	✓				
Mmp-PriSL	✓	✓				✓
Afu-PolB	✓	✓				
Afu-MCM	✓			✓		✓
Afu-GINS	✓			✓		
Afu-RecJ99/98	✓			✓		
Afu-RPA780	✓	✓			✓	
Afu-RPA780/382			✓	✓		
Afu-RadA	✓					✓
Afu-RadA/PriSL	✓					✓

Table 2.6. Purification steps for recombinant proteins.

The purified recombinant proteins are listed along with the chromatography columns used in this thesis.

Eluted protein from previous step was first concentrated using Vivaspin concentrators with an appropriate molecular weight cut off membrane (GE Healthcare) and diluted 1:1 into HIC buffer A with 1 M NaCl and 2 M ammonium sulfate. Following loading protein onto the column at a flow rate of 2 mL/min, the column was washed extensively with HIC buffer A and the protein eluted from Phenyl sepharose column with HIC buffer B (40 mM Tris-HCl pH 7.5, 500 mM NaCl, 5 mM β-mercaptoethanol). The fractions containing proteins were pooled and analysed by SDS-PAGE with Coomassie staining (section 2.3.1 and 2.3.2).

2.2.9. Size-exclusion chromatography

Some proteins in this thesis were subjected to size exclusion chromatography (SEC) for the final step of purification. (SEC) is a chromatographic method in which proteins are separated by their size. Based on the size of each protein, either an S75 or S200 10/300 GL gel-filtration column (GE Healthcare) was used. The column was subjected to ÄKTAprime system and pre-equilibrated with SEC buffer (50 mM Tris-HCl pH 7.5, 300 mM NaCl, 10% (v/v) glycerol and 0.5 mM Tris (2-carboxyethyl) phosphine (TCEP)) which had been sterile-filtered using a 0.2 µM pore size vacuum filtration system (Nalgene). 5 mL loop was pre-filled with SEC buffer and then the concentrated protein sample was loaded onto the loop. The protein was concentrated to ~ 4 mL using a Vivaspin sample concentrator. The SEC column was run at a flow rate of 2 mL/min and fractions were collected following 100 mL of flow-through.Fractions containing protein were analysed by SDS-PAGE with Coomassie staining(section 2.3.1 and 2.3.2).

2.2.10. Cleavage of His-tagged proteins with Thrombin

Since the hexahistidine tag linker in pET28a contains a thrombin cleavage site, the his tag at the N-terminal was removed using thrombin. The thrombin concentration used for cleavage is typically 1 unit per mg purified protein. Following mixing thrombin with purified protein, the mixture was incubated overnight at room temperature. To separate the thrombin and cleaved His-tag from the purified protein, 1 mL HisTrap HP column FF (GE Healthcare) connected in series to a HiTrap Benzamidine FF column (GE Healthcare) was used. The cleaved protein alongside the uncut protein was analysed by SDS-PAGE (section 2.3.1).

2.2.11. Storage of purified recombinant proteins

For storage, the purified proteins were concentrated using a Vivaspin sample concentrator (GE Healthcare) with the appropriate molecular weight filter. Concentrations of proteins were measured at A_{280} using ND-1000 NanoDrop (Thermo Scientific). Purified proteins were aliquoted into small volumes (50-100 µL) and stored -80°C.

2.3. Biochemistry Methods

2.3.1. SDS polyacrylamide gel electrophoresis

Protein samples were prepared through the addition of Laemmli sample buffer (2% (w/v) SDS, 10% v/v β-mercaptoethanol, 20% (v/v) glycerol, 0.02% bromophenol blue, 125 mM Tris-HCl pH 6.5) (Laemmli, 1970). SDS-PAGE gels were prepared in either 1 or 1.5 mm Novex Gel cassettes (Invitrogen) consisting of a stacking gel layer cast over resolving gel (Sambrook and Russell, 2006). The resolving gel was usually 8-12% acrylamide/bisacrylamide 30% (37.5:1) mix (National Diagnostics) 375 mM Tris-HCl (pH 8.8), 0.1% (w/v) SDS, 0.1% ammonium persulphate (APS), 0.04% N, N, N', N'-Tetramethylethylenediamine (TEMED). Isopropanol and water were added on top the resolving gel mix, after one hour removed, and topped with a layer of stacking gel mix (5% acrylamide mix, 125 mM Tris pH 6.8, 1% (w/v) SDS, 0.1 (w/v) APS, 0.04% (v/v) TEMED. The gel allowed setting after adding a well comb. After boiling at 95°C for 5 minutes, protein samples were run alongside a molecular weight marker, Precision Plus (Bio-Rad) for Coomassie staining (section 2.3.2) or Precision Plus Dual Marker (Bio-Rad) for western blotting (section 2.3.4). The XCell Surelock Mini-Cell Electrophoresis System (Invitrogen) was utilized and samples were electrophoresed at 170 V in SDS running buffer until the dye reached the bottom of gel.

2.3.2. Coomassie blue staining

To detect protein bands SDS gels were stained in Coomassie blue solution (50% (v/v) methanol, 10% (v/v) acetic acid, 0.5% (w/v) Commassie brilliant blue) with shaking on a rocking table for ~10-20 minutes. Coomassie blue solution was then

discarded and gels were rinse briefly with water. Destaining solution (10-20% (v/v) methanol, 10% (v/v) acetic acid) was then added to cover the gels .Gels were incubated in destain solution for 2-16 hours on the rocking table until adequate decolouration of the dye background. Gels were then air-dried using the GelAir Drying System (Bio-Rad) following the manufacturer's instructions.

2.3.3. In vitro pull down assays

To detect protein-protein interaction we took advantage of pull down assays. For His-pull down assay, 400 nM of the His-tagged protein and 400 nM of the non His-tagged protein were mixed with binding buffer (40 mM Tris pH 7.5, 100 mM NaCl and 30 mM imidazole) in 50 µL total volume. After 30 minutes incubation on ice protein mixture was added to pre-equilibrated Ni^{2+}-NTA Agarose resin and mixed on spinning wheel at 4ºC for 1 hour. The supernatant was removed after spinning at 2,000 x *g* for 2 min. The resin was then rinsed with 100 µL of wash buffer (40 mM Tris pH 7.5, 100 mM NaCl and 30 mM imidazole) for three times. 25 µL of elution buffer (40 mM Tris pH 7.5, 250 mM imidazole) was added to the protein mix and incubated for 15 minutes on a rocking table. After 2 minutes spinning at 2,000 x *g*, eluted protein was collected and analysed by 12% SDS.PAGE (section 2.3.1).

In GST-pull down assay, a GST-fused protein (bait protein) (either purified protein or cell extract) was mixed with a non-GST-tagged protein (Prey protein). The mixture was then loaded onto 30µL GST magnetic beads and incubated at 4°C for one hour. Following incubation, beads were washed three time in a buffer containing 40 mM Tris-HCL, pH 7.5, 300 mM NaCl, 5 mM β-mercaptoethanol, 0.1% NP-40. Next, beads were eluted, boiled and then run on a SDS-PAGE gel to visualize the interacting proteins.

2.3.4. Western blot analysis

Samples to be analysed by western blot were first resolved with SDS-PAGE. Following activating a polyvinyladine fluoride (PVDF) membrane (Milipore) with 100% methanol, the electrophoresed gel was washed with water and equilibrated in transfer buffer (20 mM Tris, 50 mM glycine, 10% (v/v) methanol). In order to transfer the proteins to the PVDF membrane the XCell II Blot Module (Invitrogen) was used according to the manufacturer's instructions. Transfer was carried out

at 25 V for 60 minutes in transfer buffer. Blocking buffer (Tris buffered saline) (TBS: 280 mM NaCl, 20 mM Tris) supplemented with 0.05% (v/v) Tween 20 and 5% (w/v) non-fat dried milk (Marvel) was used for blocking the membrane for at least 1 hour at room temperature. Membrane was incubated in 5 mL fresh blocking buffer containing primary antibody for 1 hour at room temperature or at 4°C overnight. The membrane was then rinsed 2 times in TBS containing 0.05% (v/v) Tween 20 and 5% (w/v) non-fat dried milk and one time in TBS supplemented only with 0.05 (v/v) Tween 20. Secondary antibody was diluted into 5 mL of blocking buffer and incubated for 1 hour at room temperature on rocking table. Following three times washes, Amersham ECL Western Blotting Detection reagent (GE Healthcare) was used for chemiluminescent detection according to the manufacturer's instructions. Light emission was captured with Amersham Hyperfilm (GE Healthcare) autoradiography film and developed using a Xograft compact X4 developer. Anti-6X His was used as the primary antibody at dilution of 1:3000 and Rabbit anti-mouse IgG H&L secondary antibody conjugated with horseradish peroxidase (HRP) (Abcam) was used at dilution of 1:5000.

2.3.6. Annealing of primer-template substrates

The DNA oligomers used to prepare the synthetic primer-template substrates were designed by Dr. Stanislaw Jozwiakowski and were HPLC grade manufactured by ATDBio. All DNA primers were labelled with hexachloroflurescein at 5' end. The sequences of all the oligonucleotides used in primer extension assays, DNA primase assays, electrophoretic mobility shift assays (EMSAs) and helicase assay are detailed in this thesis are in Table 2.7. To anneal primer-template substrates equimolar ratio of the labelled DNA primer and DNA templates were mixed to a final concentration of 200 nM in annealing buffer (50 mM NaCl, 1 mM EDTA, 20 mM Tris-HCl pH 7.5). Oligonucleotides were heated at 95 °C for 5 minutes. Annealed primer-template substrates were incubated at room temperature for 1 hour to allow annealing and then stored at -20°C.

Primer	Direction	Label	Sequence
Undamaged substrate(50/20mer)	5'-3' 3'-5'	5'-Hex None	TGTCGTCTGTTCGGTCGTTC ACAGCAGACAAGCCAGCAAGCCAGAAGTTCCGACA ACACGCGGGACGCGC
50mer ssDNA	5'-3'	5'-Hex	TGTCGTCTGTTCGGTCGTTCGGTCTTCAAGGCTGTT GTGCGCCCTGCGCG
50mer dsDNA	5'-3' 3'-5'	5'-Hex None	TGTCGTCTGTTCGGTCGTTCGGTCTTCAAGGCTGTT GTGCGCCCTGCGCG ACAGCAGACAAGCCAGCAAGCCAGAAGTTCCGACA ACACGCGGGACGCGC
45mer ssDNA	5'-3'	5'-FAM	AGTCGCATAGTGTAGTCGGTCTTGTTCGGTCATAGC TCATCGTGG
45mer dsDNA	5'-3' 3'-5'	5'-FAM None	AGTCGCATAGTGTAGTCGGTCTTGTTCGGTCATAGC TCATCGTGG CCACGATGAGCTATGACCGAACAAGACCGACTAACA TATGCGACT
65mer ssDNA	5'-3'	5'- Biotin	GTCTTCTATCTCGTCTATATTCTATTGTCTCTATGAAT ACCTTCATTCATTCTCACATAGATGCATC
Poly(dA)	5'-3	5'- Biotin	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
Poly(dC)	5'-3'	5'- Biotin	CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
Poly(dG)	5'-3'	5'- Biotin	GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
Poly(dT)	5'-3'	5'- Biotin	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
Fidelity substrate (AA)	5'-3' 3'-5'	5'-Hex None	TGTCGTCTGTTCGGTCGTTCGGTCTTC ACAGCAGACAAGCCAGCAAGCCAGAAGAACCGACA ACACGCGGGACGCGC
Fidelity substrate (CC)	5'-3' 3'-5'	5'-Hex None	TGTCGTCTGTTCGGTCGTTCGGTCTTC ACAGCAGACAAGCCAGCAAGCCAGAAGCCTTGACA ACACGCGGGACGCGC
Fidelity substrate (GG)	5'-3' 3'-5'	5'-Hex None	TGTCGTCTGTTCGGTCGTTCGGTCTTC ACAGCAGACAAGCCAGCAAGCCAGAAGGGCCGACA ACACGCGGGACGCGC
Fidelity substrate (TT)	5'-3' 3'-5'	5'-Hex None	TGTCGTCTGTTCGGTCGTTCGGTCTTC ACAGCAGACAAGCCAGCAAGCCAGAAGTTCCGACA ACACGCGGGACGCGC
Mismatch substrate (A/AA)	5'-3' 3'-5'	5'-Hex None	TGTCGTCTGTTCGGTCGTTCGGTCTTCA ACAGAGACAAGCCAGCAAGCCAGAAGAACCGACAA CACGCGGGACGCGC
Mismatch substrate (C/AA)	5'-3' 3'-5'	5'-Hex None	TGTCGTCTGTTCGGTCGTTCGGTCTTCC ACAGCAGACAACCAGCAAGCCAGAAGAACCGACAA CACGCGGGACGCGC
Mismatch substrate (G/AA)	5'-3' 3'-5'	5'-Hex None	TGTCGTCTGTTCGGTCGTTCGGTCTTCG ACAGCAGACAAGCCAGAAGCCAGAAGAACCGACAA CACGCGGGACGCGC

Mismatch substrate (T/AA)	5'-3' 3'-5'	5'-Hex None	TGTCGTCTGTTCGGTCGTTCGGTCTTCT ACAGCAGACAACCAGCAAGCCAGAAGAACCGACAA CACGCGGGACGCGC
Mismatch substrate (A/CC)	5'-3' 3'-5'	5'-Hex None	TGTCGTCTGTTCGGTCGTTCGGTCTTCA ACAGCAGACAGCCAGCAAGCCAGAAGCCCCGACAA CACGCGGGACGCGC
Mismatch substrate (C/CC)	5'-3' 3'-5'	5'-Hex None	TGTCGTCTGTTCGGTCGTTCGGTCTTCC ACAGCAGAAAGCCAGCAAGCCAGAAGCCCCGACAA CACGCGGGACGCGC
Mismatch substrate (G/CC)	5'-3' 3'-5'	5'-Hex None	TGTCGTCTGTTCGGTCGTTCGGTCTTCG ACAGCAGACAGCCAGCAAGCCAGAAGCCCCGACAA CACGCGGGACGCGC
Mismatch substrate (T/CC)	5'-3' 3'-5'	5'-Hex None	TGTCGTCTGTTCGGTCGTTCGGTCTTCT ACAGCGACAAGCCAGCAAGCCAGAAGCCCCGACAA CACGCGGGACGCGC
Mismatch substrate (A/GG)	5'-3' 3'-5'	5'-Hex None	TGTCGTCTGTTCGGTCGTTCGGTCTTCA ACAGCAGACAAGCCAGCAAGCCAGAAGGGCCGACA ACACGCGGGACG
Mismatch substrate (C/GG)	5'-3' 3'-5'	5'-Hex None	TGTCGTCTGTTCGGTCGTTCGGTCTTCC ACAGCAGACAAGCCAGCAAGCCAGAAGGGCCGACA ACACGCGGGACG
Mismatch substrate (G/GG)	5'-3' 3'-5'	5'-Hex None	TGTCGTCTGTTCGGTCGTTCGGTCTTCG ACAGCAGACAAGCCAGCAAGCCAGAAGGGCCGACA ACACGCGGGACG
Mismatch substrate (T/GG)	5'-3' 3'-5'	5'-Hex None	TGTCGTCTGTTCGGTCGTTCGGTCTTCT ACAGCAGACAAGCCAGCAAGCCAGAAGGGCCGACA ACACGCGGGACG
Mismatch substrate (A/TT)	5'-3' 3'-5'	5'-Hex None	TGTCGTCTGTTCGGTCGTTCGGTCTTCA ACAGCAGACAAGCCGCAAGCCAGAAGTTCCGACAA CACGCGGGACGCGC
Mismatch substrate (C/TT)	5'-3' 3'-5'	5'-Hex None	TGTCGTCTGTTCGGTCGTTCGGTCTTCC ACAGCAGACAAGCCACAAGCCAGAAGTTCCGACAA CACGCGGGACGCGC
Mismatch substrate (G/TT)	5'-3' 3'-5'	5'-Hex None	TGTCGTCTGTTCGGTCGTTCGGTCTTCG ACAGCAGACAAGCAGCAAGCCAGAAGTTCCGACAA CACGCGGGACGCGC
Mismatch substrate (T/TT)	5'-3' 3'-5'	5'-Hex None	TGTCGTCTGTTCGGTCGTTCGGTCTTCT ACAGCAGACAGCCAGCAAGCCAGAAGTTCCGACAA CACGCGGGACGCGC
CPD Substrate	5'-3' 3'-5'	5'-Hex None	TGTCGTCTGTTCGGTCGTTC ACAGCAGACAAGCCAGCAAGCCAGAAGT=TCCGACA ACACGCGGGACGCGC
8-oxo-G Substrate	5'-3' 3'-5'	5'-Hex None	TGTCGTCTGTTCGGTCGTTC ACAGCAGACAAGCCAGCAAGCCAGAAG8oGCCGAC AACACGCGGGACGCGC
Uracil Substrate	5'-3' 3'-5'	5'-Hex None	TGTCGTCTGTTCGGTCGTTC ACAGCAGACAAGCCAGCAAGCCAGAAGUCCGACAA CACGCGGGACGCGC

MCM helicase substrate 1	5'-3' 3'-5' 5'-3' 3'-5'	5'-Hex None None None	GACGCTGCCGAATTCTACCAGTGCCTTGCTAGGACA TCTTTGCCCACCTGCAGGTTCACCC ATCGATAGTCGGATCCTCTAGACAGCTCCATGTAGC AAGGCACTGGTAGAATTCGGCAGCGT TGGGTGAACCTGCAGGTGGGCAAAGATGTCC CATGGAGCTGTCTAGAGGATCCGACTATCGA
MCM helicase substrate 2	5'-3' 3'-5' 5'-3'	5'-Hex None None	GACGCTGCCGAATTCTACCAGTGCCTTGCTAGGACA TCTTTGCCCACCTGCAGGTTCACCC ATCGATAGTCGGATCCTCTAGACAGCTCCATGTAGC AAGGCACTGGTAGAATTCGGCAGCGT TGGGTGAACCTGCAGGTGGGCAAAGATGTCC
MCM helicase substrate 3	5'-3' 3'-5' 5'-3'	5'-Hex None None	GACGCTGCCGAATTCTACCAGTGCCTTGCTAGGACA TCTTTGCCCACCTGCAGGTTCACCC ATCGATAGTCGGATCCTCTAGACAGCTCCATGTAGC AAGGCACTGGTAGAATTCGGCAGCGT CATGGAGCTGTCTAGAGGATCCGACTATCGA
MCM helicase substrate 4	5'-3' 3'-5'	5'-Hex None	GACGCTGCCGAATTCTACCAGTGCCTTGCTAGGACA TCTTTGCCCACCTGCAGGTTCACCC ATCGATAGTCGGATCCTCTAGACAGCTCCATGTAGC AAGGCACTGGTAGAATTCGGCAGCGT
8-oxo-G substrate	5'-3' 3'-5'	5'-Hex None	TGTCGTCTGTTCGGTCGTTCGGTCTTCA CGCGCAGGGCGCACAACAGCC(G)TGAAGACCGAAC GACCGAACAGACGACA
CPD substrate	5'-3' 3'-5'	5'-Hex None	TGTCGTCTGTTCGGTCGTTCGGTCTTCA CGCGCAGGGCGCACAACAGCC(T=T)TGAAGACCGA ACGACCGAACAGACGACA
Uracil substrate	5'-3' 3'-5'	5'-Hex None	TGTCGTCTGTTCGGTCGTTCGGTCTTCA CGCGCAGGGCGCACAACAGCC(U)TGAAGACCGAAC GACCGAACAGACGACA
Tg substrate	5'-3' 3'-5'	5'-Hex None	CACTGACTGTATGATGA CTCGTCAGCATC(T/T)CATCATACAGTCAGTG
AP substrate	5'-3' 3'-5'	5'-Hex None	CACTGACTGTATGATGC CTCGTCAGCATC(T/T)CATCATACAGTCAGTG

Table 2.7. List of oligonucleotides used in this thesis

Primer, template oligonucleotides used in this thesis are listed. Mismatched bases are colored in blue (i.e. none Watson-Crick base pairs). DNA lesions in templates are indicated in red: T=T, cyclobutane pyrimidine dimers; 8oG, 8-oxo-guanine; U, uracil.

2.3.7. Primer extension assays

The annealed substrate (20 nM) was mixed with 1x reaction buffer (20 mM Tris, pH 8.8, 10 mM KCl, 10 mM $(NH_4)_2SO_2$, 2 mM $MgSO_4$) and 200 µM dNTPs. Reactions were incubated at 55°C for *Archaeoglobus fulgidus* and *Pyrococcus furiosus* or at 37°C for *Methanococcus maripalidus*. Protein concentration used in the reactions was 200-300 nM. Reactions were quenched at multiple time points (30 second, 1, 3, 5, 30 minutes) using equal volume of 2x stop buffer (95% formamide, 0.09% xylene cyanol, 0.05% bromophenol blue, 200 nM competitor oligonucleotide). The reason to use the competitor oligonucleotide in the stop buffer, which is complementary to the template strand, is to prevent re-annealing of the labelled reaction products with the template strand, which would affect analysis of gel electrophoresis. The reactions were then boiled at 95°C for 5 minutes and resolved by electrophoresis on a 15% polyacrylamide gel (37.5% (v/v) acrylamide/bisacrylamide (19:1) 40% mix (National Diagnostics), 7 M urea, TBE (89 Mm Tris, 2 mM EDTA, 0.89 M boric acid, pH 8.3) 0.1% (w/v) APS, 0.03% (v/v) TEMED). Prior to 3 hours electrophoresis at 15-20 watts in TBE, samples were pre-run at 15 W for 1 hour. To detect the hexachlorofluorescein label of the reaction products, the gel was scanned (Cy3, 532 nm) using a FLA-1500 scanner (FUJI).

2.3.8. DNA primase assays

The non-radioactive primase assay was performed using 1 µM homopolymeric single stranded templates with biotin labelling at the 5' end (Table 2.7). In a final volume of 20 µL, the DNA was mixed with 250 µM rNTPs (Invitrogen) or dNTPs (Roche), 40 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$ or 4 mM $MnCl_2$, 2 mM Dithiothreitol (DTT) and 20 mM NaCl. Following adding 200 µM of protein the reaction was supplemented with 2.5 µM FAM dNTPs (Jena Bioscience) and incubated at 55°C for *Archaeoglobus fulgidus* and *Pyrococcus furiosus* (hyperthermophile species of archaea) and 37°C for *Methanococcus maripalidus* (methanogen species of archaea) for 1 hour. The primer synthesis/labelling enzymatic reactions were terminated by adding 400 µl of binding-washing (B-W) buffer (10 mM Tris-HCl (pH 8.0), 500 mM NaCl, 10 mM EDTA). 20 µL of streptavidin-coated beads (Invitrogen) was mixed with the quenched reaction on spinning wheel for 1 hour at 4°C to allow the ssDNA binds to the beads. The

supernatants were removed using magnetic separation rack and then the beads were washed three times with 1 mL B-W buffer. The beads were then suspended in 20 µl of the B-W buffer supplemented with equal volume of loading buffer (8 M Urea, 10 mM EDTA). After heating at 95°C for 5 minutes the reaction were resolved by DNA-PAGE in TBE buffer. The gels were scanned for fluorescent signal detected by a Fujifilm FLA-5100 image reader.

The radioactive primase assay was performed using 250 ng M13mp18 ssDNA in a 20 µL reaction volume. The DNA mixture contained 40 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 4 mM $MnCl_2$, 2 Mm Dithiothreitol (DTT), 20 mM NaCl , 250 µM rNTPs or dNTPs including [α-^{32}P]ATP. The reactions were supplemented with 500 nM primase and incubated at 55 °C for *Archaeoglobus fulgidus* and *Pyrococcus furiosus* (hyperthermophile species of archaea) and 37°C for *Methanococcus maripalidus* (methanogen species of archaea) for 30 minutes. The mixtures were then stopped by the addition of equal volume of stop buffer (95% formammide, 10 mM EDTA and 0.05% bromophenol blue), boiled at 95°C for 5 minutes and subjected to electrophoresis through a 15% polyacrylamide gel containing 7 M urea in 1× TBE. The gel was exposed to a phosphor screen film and scanned and analysed using a FLA-1500 scanner (FUJI).

2.3.9. Electrophoretic mobility shift assays

Electrophoretic mobility shift assays (EMSA) were performed using 60nM of either hex-labelled single-stranded DNA or double-stranded DNA (Table 2.7). DNA was added to 1x NEB buffer (10 mM Bis-Tris-Propane-HCl pH 7.0, 10 mM $MgCl_2$, 100 µg/mL BSA) unless otherwise specified and supplemented with different concentrations of proteins. After 1 hour incubation of samples at room temperature, 2 µL of 25% (w/v) ficoll was added to reactions. Reaction were resolved on a 5% polyacrylamide gel (12.5% (v/v) acrylamide/bisacrylamide (19:1) 40% mix (National Diagnostics), 0.5x TBE (44.5 mM Tris, 1 mM EDTA, 0.445 M boric acid, pH 8.3) 0.1% (w/v0 APS, 0.03% (v/v) TEMED) cast in a 1 mm Novex Gel cassette (Invitrogen). To run the gel, the XCell SureLock Mini-Cell Electrophoresis System (Invitrogen) was used. To detect the hexachlorofluorescein label of the reaction products, the gel was scanned (Cy3, 532 nm) using a FLA-1500 scanner (FUJI).

2.3.10. DNA helicase assays

Reactions prepared using 500 nM of helicase enzyme mixed with 20 nM Tris (pH 7.5), 20 mM $MgCl_2$,1 mM EDTA and 200 nM of each substrate in a total volume of 20 µL. Substrates contained a 5' end labelled oligonucleotides annealed to other oligonucleotides. After 5 minute incubation at 50°C, ATP was added in different concentrations. Reactions were then stopped using stop buffer (10% glycerol, 1% SDS, 10 mM EDTA and bromophenol blue). Products were separated by electrophoresis on 15% polyacrylamide native gel run in 1x TBE buffer at 10 W for 2 hours. To detect the hexachlorofluorescein label of the reaction products, the gel was scanned (Cy3, 532 nm) using a FLA-1500 scanner (FUJI).

2.4. Bioinformatics and analytic tools

2.4.1. Computation of physical and chemical parameters of proteins

The information regarding chemical and physical properties of proteins were provided by the ProtParam program, from the ExPASy server. These parameters were determined from the primary protein sequence including pI, molecular weight and molecular extension coefficient.

2.4.2. Multiple sequence alignment

The gene sequences required for alignment were gathered using the National Centre for Biotechnology Information from GenBank (Benson *et al*., 2009; Sayers *et al*., 2010). To design and analyse multiple sequence alignments, Clustal Omega and JalView were used respectively (Sievers *et al*., 2011;Waterhouse *et al*., 2009).

Chapter 3

Biochemical characterisation of archaeal replicative primases

3.1. Introduction

Archaea encode both bacterial DnaG-type and AEP primases (Liu *et al*., 2015). All experimentally characterized AEP primases from archaea form heterodimers consisting of PriS and PriL, which are termed PriSL (Bocquier *et al*., 2001; Guilliam *et al*., 2015; Kelman and Kelman, 2014; Liu *et al*., 2001). The archaeal PriSL has been suggested to function as the replicative primase. Archaeal replicative primase not only capable of synthesizing primers but also is able to elongate them. The primer elongation capacity of replicative primase in various archaeal species ranges from less than 500 bases in length to >7 kb (Bocquier *et al*., 2001). In contrast with the majority of eukaryotic primases, which only synthesise RNA primers, archaeal PriSL can perform *de novo* synthesis of both DNA and RNA primers *in vitro*, using either dNTPs or rNTPs (Liu *et al*., 2001). Open reading frames (ORFs) that show sequence similarity to the eukaryotic DNA primase subunit p48 have been identified and characterized from different archaeal species including, *Methanococcus jannaschii*, *Pyrococcus furiosus* and *Archaeoglobus fulgidus* (Kirk and Kuchta., 1999). Moreover, comparative analysis identified the presence of an ORF overlapping with *P.furiosus* primase small subunit (*Pfup41*) which showed sequence similarity to eukaryotic DNA primase large subunit (p58) (Liu *et al*.,2001).

Three-dimensional structures of archaeal DNA primase catalytic subunits (PriS) revealed that PriS comprises a large catalytic domain and a small α-helical domain. In addition, the catalytic domain possesses a zinc-finger motif that is conserved in eukaryotes (Sarmiento *et al*., 2013). It has been suggested that the zinc-finger motif facilitates interaction of the primase with DNA (Iyer *et al*., 2005). There is an iron-sulfur domain in the C-terminal domain of PriL that might be important for retaining the correct three-dimensional structure of this domain (Lao-Sirieix *et al*., 2005). Since there is no direct interaction between PriL and the active site of PriS subunit, it is believed that PriL subunit may contact with the synthesized primer and regulate its length (Ishino and Ishino, 2013). Studies on the *Pyrococcus furiosis* PriSL complex indicated that although PriL increases RNA polymerase activities, it decreases the DNA polymerase activity. This subunit can also reduce the length of the DNA oligonucleotides synthesised by

PriS. Therefore, it has been concluded that PriL acts as a regulator in the primase complex (Liu *et al*., 2001).

The available evidence seems to suggest that archaeal replicative primases possess an additional role, extending primers with dNTPs in a polymerase mode that is similar to eukaryotic Pol α, therefore establishing that archaeal replicative primases can be categorized as primase-polymerases (Prim-Pols).

In this chapter, the biochemical properties of archaeal replicative primases, PriSL, from three different species, *Archaeoglobus fulgidus*, *Pyrococcus furiosus* and *Methanococcus maripaludis* are characterised. Here, both primase and polymerase activities of PriSL complex from these archaeal species are illustrated (Figure 3.1 and 3.2). A two-hybrid system was used to confirm the interaction between PriS and PriL in *Archaeoglobus fulgidus.* In addition, the ability of archaeal replicative primase to perform both DNA and RNA primase activities is demonstrated in this chapter. We also showed that the PriSL complex from *Archaeoglobus fulgidus* binds both double-stranded and single-stranded DNA. In general, the aim of this chapter was to characterise and compare the *in vitro* activities of PriSL from different archaeal species.

3.2. Identification of protein partners using the Yeast Two-Hybrid assays

To identify interactions between the DNA primase small subunit (PriS) and large subunit (PriL) from the archaeon *Archaeoglobus fulgidus*, the yeast two-hybrid (Y2H) system was chosen. In this genetic assay, transcriptional factors are used. In particular GAL4 protein of the yeast *Saccharomyces cerevisiae*, a transcriptional activator which promotes transcription of genes containing Gal4 is utilized. This protein consists of two discrete domains: An N-terminal domain that binds to the upstream activation domain (UAS) and a C-terminal domain required for transcription activation. In practice, the DNA binding domain (BD) of the GAL4 protein is fused to a protein X (bait) and the transcriptional activation domain (AD) is fused to a protein Y (prey). If bait and prey interact with each other, the two domains of the GAL4 protein are brought together and transcription of the specially designed reporter genes occurs which provides evidence for protein-protein interactions (Figure. 3.3) (Fields and Song,1989).

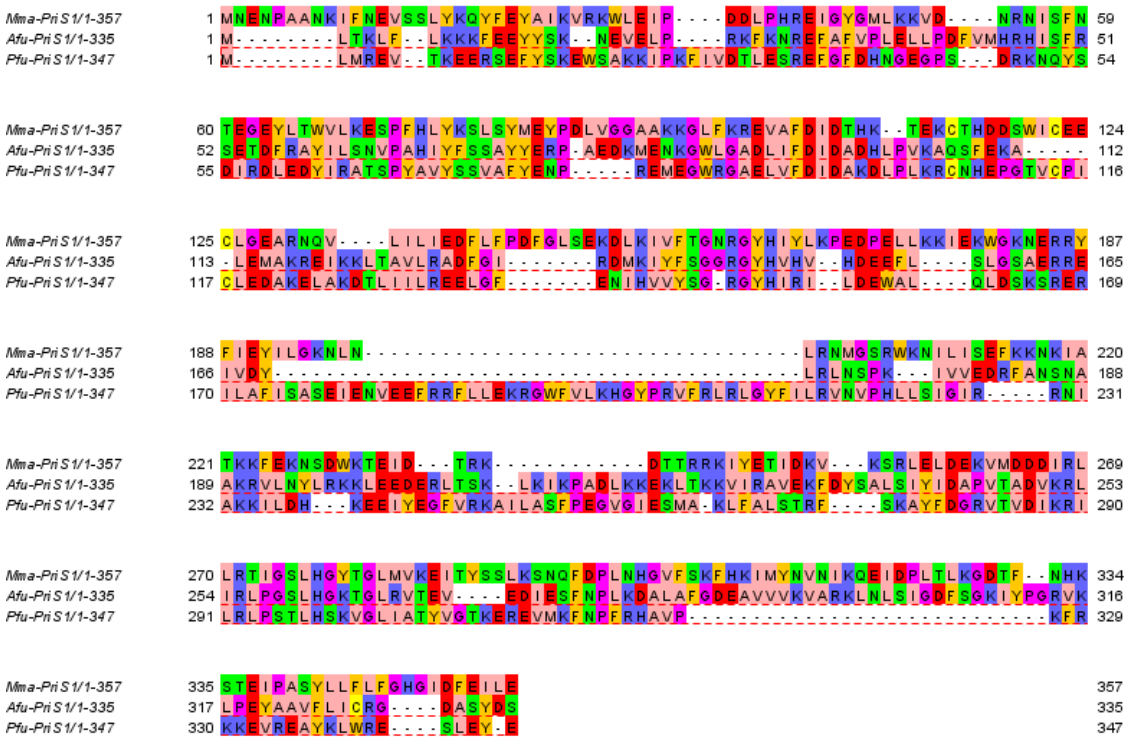


Figure 3.1. Alignment of archaeal primase small subunit (PriS1) homologues. The alignment of the three archaeal PriS1 homologues that are studies biochemically in this thesis from *Archaeoglobus fulgidus*, *Pyrococcus furiosus*, and *Methanococcus maripaludis C5*.

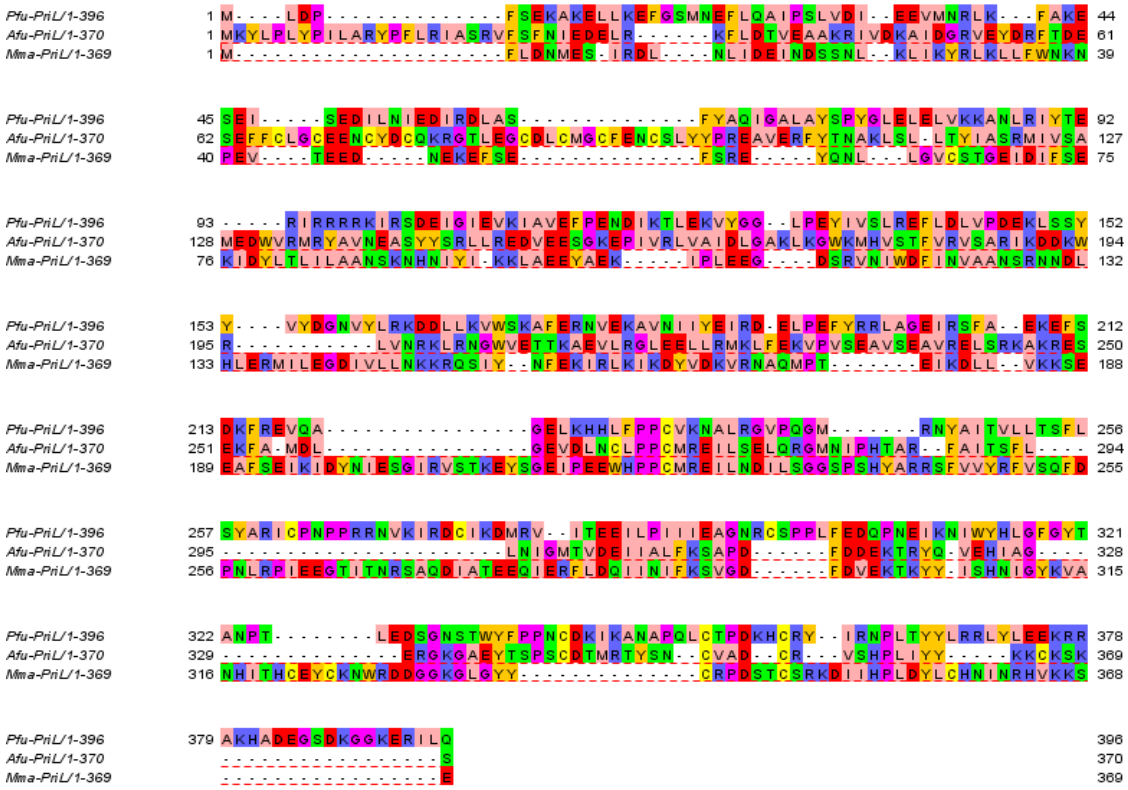


Figure 3.2. Alignment of archaeal primase large subunit (PriL) homologues.

The alignment of the three archaeal PriL homologues that are analysed biochemically in this thesis from *Archaeoglobus fulgidus*, *Pyrococcus furiosus*, and *Methanococcus maripaludis C5*.

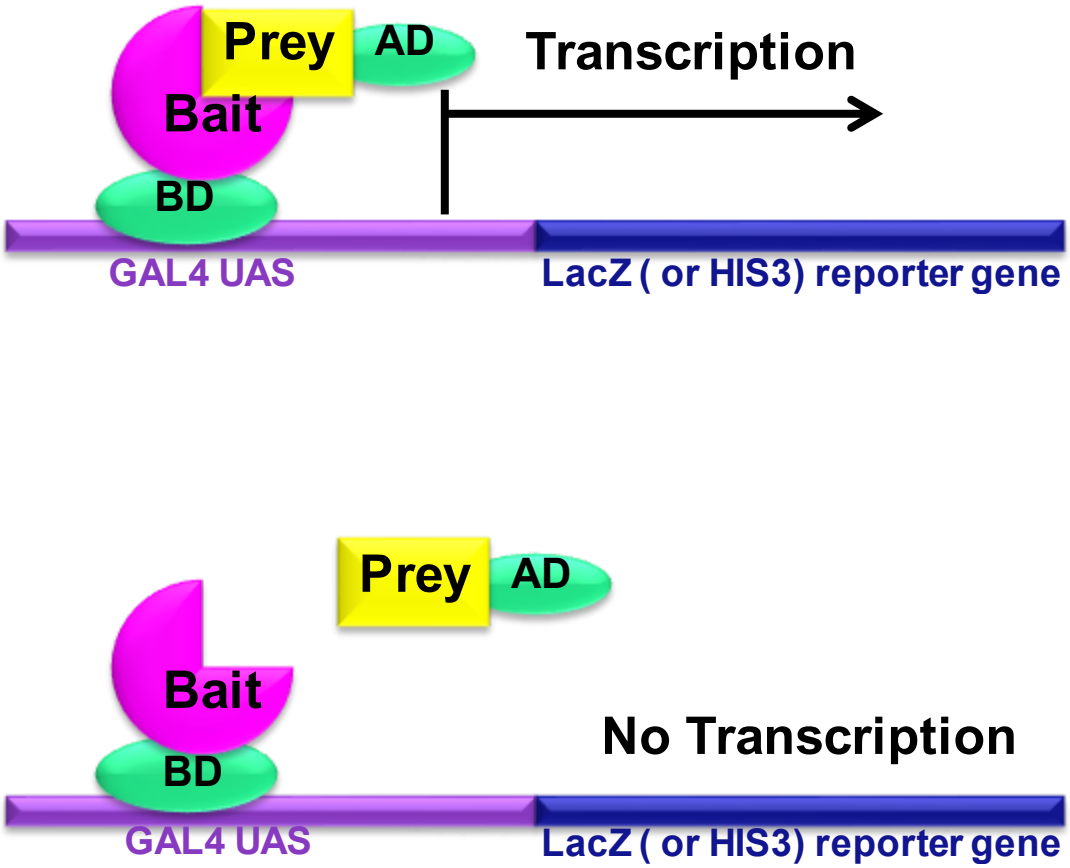


Figure 3.3. Principles of the yeast two-hybrid system

(A) Two-hybrid screening is designed to fuse the bait protein to the GAL4 DNA binding domain (BD) to screen against a library of the prey protein which is fused to the GAL4 transcriptional activation domain (AD). The BD binds the upstream activating sequence (UAS) and the AD domain activates transcription. When the bait protein interacts with the prey protein GAL4 transcription factor is reconstructed and allows expression of the reporter genes. **(B)** When bait and prey proteins do not interact, the GAL4 transcription factor is not reconstructed and so the reporter genes are not expressed.

Whilst searching for archaeal PriS orthologous in *A.fulgidus*, in addition to PriS1, we identified a second primase (AF_RS11580) which we have called PriS2. The genes encoding PriS1, PriS2 and PriL were cloned in both the pGBKT7 (bait vector) and pGADT7 (prey vector) in frame with the sequences encoding the GAL4 BD and AD domains, respectively (Figure 3.4) (Table 2.2). All combinations of bait and prey plasmids were successively transformed into the yeast strain Y190 with *LacZ* and *HIS3* reporter genes (Van Creikinge and Beyaert, 1999). The transformed cells were plated onto media lacking His, Arg, and Leu (SD-His-Leu-Arg) supplemented with optimal concentration of 3-AT, a competitive inhibitor of the yeast HIS3 protein that inhibits low-level leaky expression (Durfee *et al*., 1993), and incubated for 2-3 days at 30°C.

Y190 strain was first co-transformed with each pGBKT7:PriS1, pGBKT7:PriS2, and pGBKT7:PriL and empty pGADT7 vector in order to check whether expression of PriSL in the Y190 strain resulted in auto-activation of the *LacZ* and *HIS3* reporter genes. *LacZ* transcription results in expression of β-galactosidase and transcription of *HIS3* allows growth on media containing no histidine. In the absence of prey protein, expression of all three bait proteins (pGBKT7:PriS, pGBKT7:PriS2, and pGBKT7:PriL) did not result in expression of reporter genes. Since expression of these fusion proteins did not lead to reporter genes auto-activation therefore they could be used in the yeast two-hybrid assay (Figure 3.5. and 3.6). On the other hand, co-transormation of *S. cerevisae* Y190 with pGBKT7:PriS1 and pGADT7 PriL constructs, on a media containing 3-AT and lacking His, Leu, and Arg ended up with cell growth (Figure 3.5). This procedure was also carried out with a reverse order of bait and prey vectors. Co-transformation of Y190 with pGBKT7:PriL (bait) and pGADT7:PriS1 (prey) also showed detectable growth on the same media (Figure 3.6). Together, these results suggest a strong interaction of PriS1 with PriL. Interestingly, no interaction was observed between PriS2 and PriL. To ensure that the detected interaction was not a false positive, pGADT7:Ku and pGBKT7:LigD constructs were created (gift from Dr.Nigel Brissett, Doherty lab) (Table 2.2). Different studies have been demonstrated that in many prokaryotes, Ku and LigD proteins interact and form a complex to facilitate re-joining of DSBs (Guilliam *et al*., 2015; Bartlett *et al*., 2013).

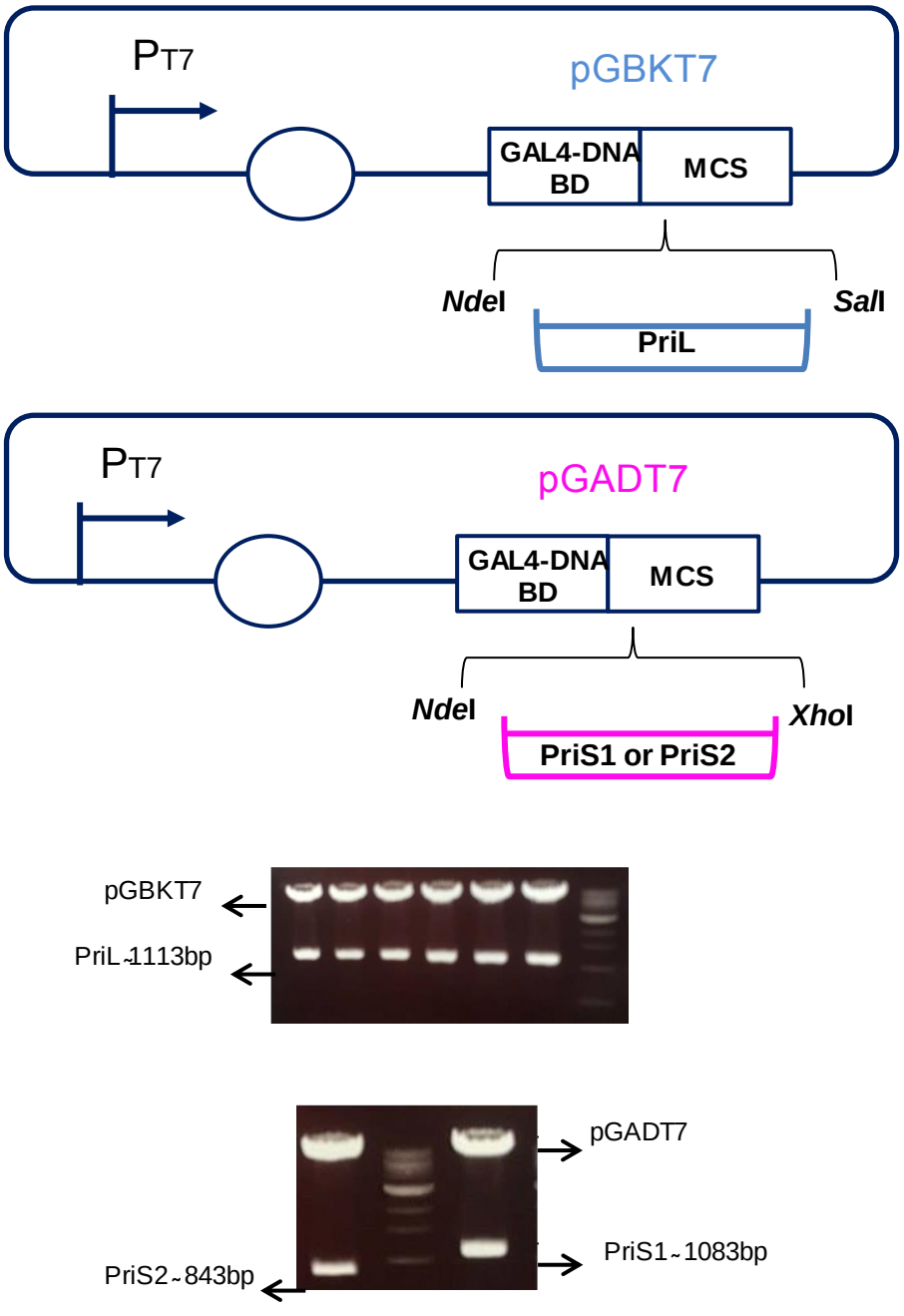
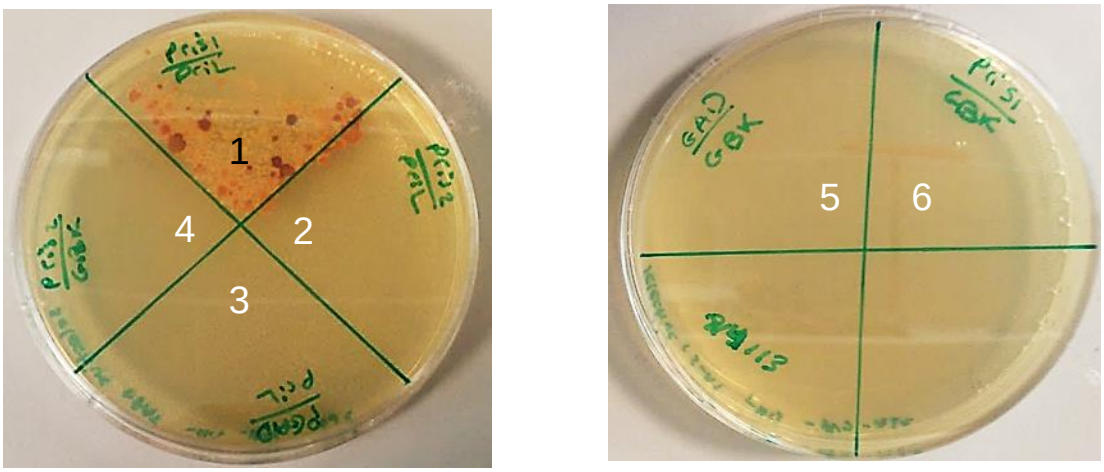


Figure 3.4. Cloning of PriS1, PriS2, and PriL into pGBKT7 and pGADT7 vectors

The open reading frame of PriS1, PriS2, and PriL primase subunits were amplified from *A fulgidus* genomic DNA (Table 2.2) introducing the relevant restriction sites to allow insertion into the multiple cloning site (MCS) of the yeast expression vector (either pGBKT7 or pGADT7). The PCR products were combined with 10x DNA loading dye and run on 1% agarose gels containing ethidium bromide.

A

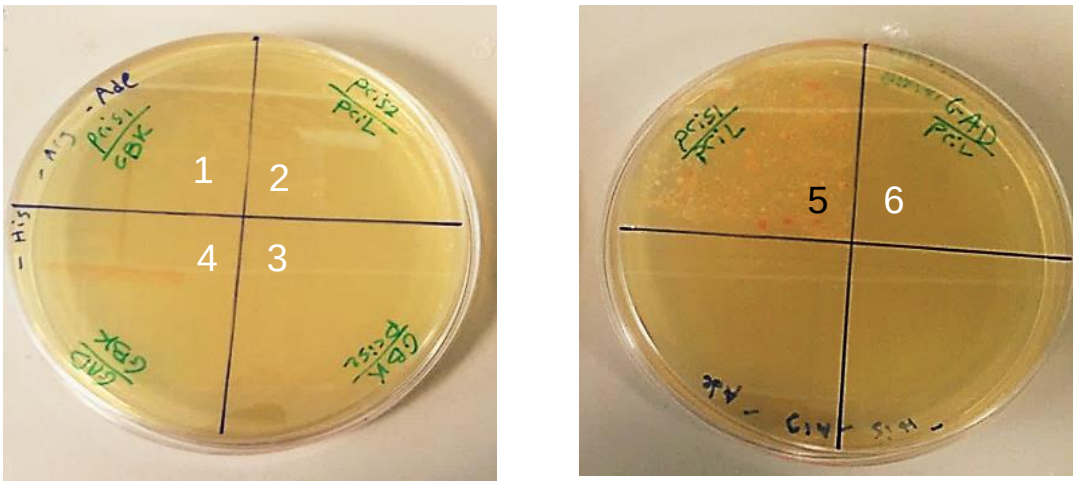


B

#	Bait/Prey pair	Induction	Interaction
1	PriS1-pGBK/priL-pGAD	-His/-Arg/-Leu +3AT	Strong
2	PriS2-pGBK +PriL-pGAD	-His/-Arg/-Leu +3AT	NO
3	pGBK+ PriL- pGAD	-His/-Arg/-Leu +3AT	NO
4	PriS2-pGAD+ pGBK	-His/-Arg/-Leu +3AT	NO
5	PriS1-pGAD+pGBK	-His/-Arg/-Leu +3AT	NO
6	pGAD+pGBK	-His/-Arg/-Leu +3AT	NO

Figure 3.5. Afu-PriS1 subunit interacts with Afu-PriL subunit.
(A) Results of two-hybrid assay on SD-His-Arg-Leu plates containing 3-AT,where genes encoding PriS1,PriS2 and PriL subunits were cloned in both the pGBKT7 (bait vector) and pGADT7 (prey vector). **(B)** Summary of bait and prey pairs used in the two-hybrid system. Strong interaction detected between PriS1-pGBKT7 and PriL-pGADT7.

A



B

#	Bait/Prey pair	Induction	Interaction
1	PriS1-pGAD/ pGBK	-His/-Arg/-Leu +3AT	NO
2	PriS2-pGAD/ PriL -pGBK	-His/-Arg/-Leu +3AT	NO
3	pGBK/PriS2-pGAD	-His/-Arg/-Leu +3AT	NO
4	pGAD/pGBK	-His/-Arg/-Leu +3AT	NO
5	PriS1-pGAD/PriL-pGBK	-His/-Arg/-Leu +3AT	Strong
6	pGAD/PriL-pGBK	-His/-Arg/-Leu +3AT	No

Figure 3.6. Afu-PriS1 and Afu-PriL subunits are interacting partners

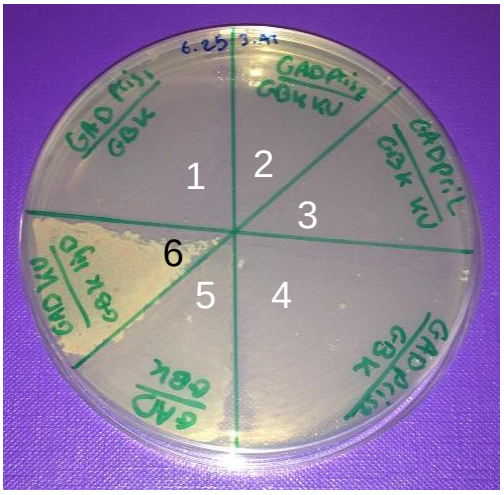
(A) Results of two-hybrid assay on SD-His-Arg-Leu plates containing 3-AT,where genes encoding PriS1,PriS2 and PriL subunits were cloned in both the pGBKT7 (bait vector) and pGADT7 (prey vector). **(B)** Summary of bait and prey pairs used in the two-hybrid system. (5) Strong interaction detected between PriL-pGBKT7 and PriS1-pGADT7.

Transformation of pGADT7: Ku and pGBKT7:LigD plasmids showed strong interaction (Figure 3.7). Thus, this result suggested that the yeast two-hybrid system worked properly and confirmed that the obtained interaction between PriS1 and PriL was true. This result is consistent with different studies in which it was reported that archaeal DNA primase is a heterodimer of PriS and PriL (Bocquier *et al*., 2001; Kelman and Kelman, 2014; Liu *et al*., 2001).

3.3. Cloning of the archaeal PriS/L genes into expression vectors

In order to characterise the enzymatic activities of archaeal PriS/L *in vitro*, the genes encoding PriS1 and PriL subunits from *Archaeoglobus fulgidus*, *Pyrococcus furiosus*, and *methanococcus maripaludis C5* were first cloned into appropriate expression vectors. The Pfu-PriS1, Pfu-PriL, Mma-PriS1and Mma-PriL expression constructs were provided by my colleague Dr. Stanislaw Jozwiakowski (Table 2.3). The ORF corresponding to Afu-PriS1 (AF_RS03760) and the ORF corresponding to Afu-PriL (AF_RS01565) were amplified from *A.fulgidus* genomic DNA (using primers in table 2.1) and cloned individually into the *E.coli* expression vector pET28a and pETduet-1, respectively, generating constructs pET28a:Afu-PriS1 and pETduet-1:Afu-PriL. PriS1 was cloned in-frame with an amino-terminal 6-histidine tag and expression of these fusions was under the control of T7 promoter (Figure 3.8). The open reading frame corresponding to Pfu-PriS1 (PF_RS00545) and the open reading frame corresponding to Pfu-PriL (PF_RS00550) were amplified from *P.furiosus* genomic DNA (using primers in table 2.1) and cloned individually into the *E.coli* expression vector pET28a and pETduet-1 respectively, generating constructs pET28a:Pfu-PriS1 and pETduet-1:Pfu-PriL. In parallel, the ORF corresponding to Mma-PriS1 (MMARC5_RS08185) and the open reading frame corresponding to Mma-PriL (MMARC5_RS08555) were amplified from *M.maripalidus* genomic DNA (Using primers in table 2.1) and cloned individually into the *E.coli* expression vector pET28a and pETduet-1 respectively, generating constructs pET28a:Mma-PriS1 and pETduet-1:Mma-PriL.

A



B

#	Bait/Prey pair	Induction	Interaction
1	pGAD-PriS1/ pGBK	-His/-Arg/-Leu +3AT	NO
2	pGAD-PriS2/pGBK-KU	-His/-Arg/-Leu +3AT	NO
3	pGAD-PriL/pGBK-KU	-His/-Arg/-Leu +3AT	NO
4	pGAD-PriS2/pGBK	-His/-Arg/-Leu +3AT	NO
5	pGAD/pGBK	-His/-Arg/-Leu +3AT	No
6	pGAD-KU/pGBK-ligD(positive control)	-His/-Arg/-Leu +3AT	Strong

Figure 3.7. Studying Protein-Protein interaction using yeast two-hybrid assays

(A) Results of two-hybrid assay on SD-His-Arg-Leu plates containing 3-AT, where genes encoding LigD and Ku proteins were cloned in both the pGBKT7 (bait vector) and pGADT7 (prey vector). **(B)** Summary of bait and prey pairs used in the two-hybrid system. (6) Strong interaction detected between pGBKT7-LigD and pGADT7-Ku.

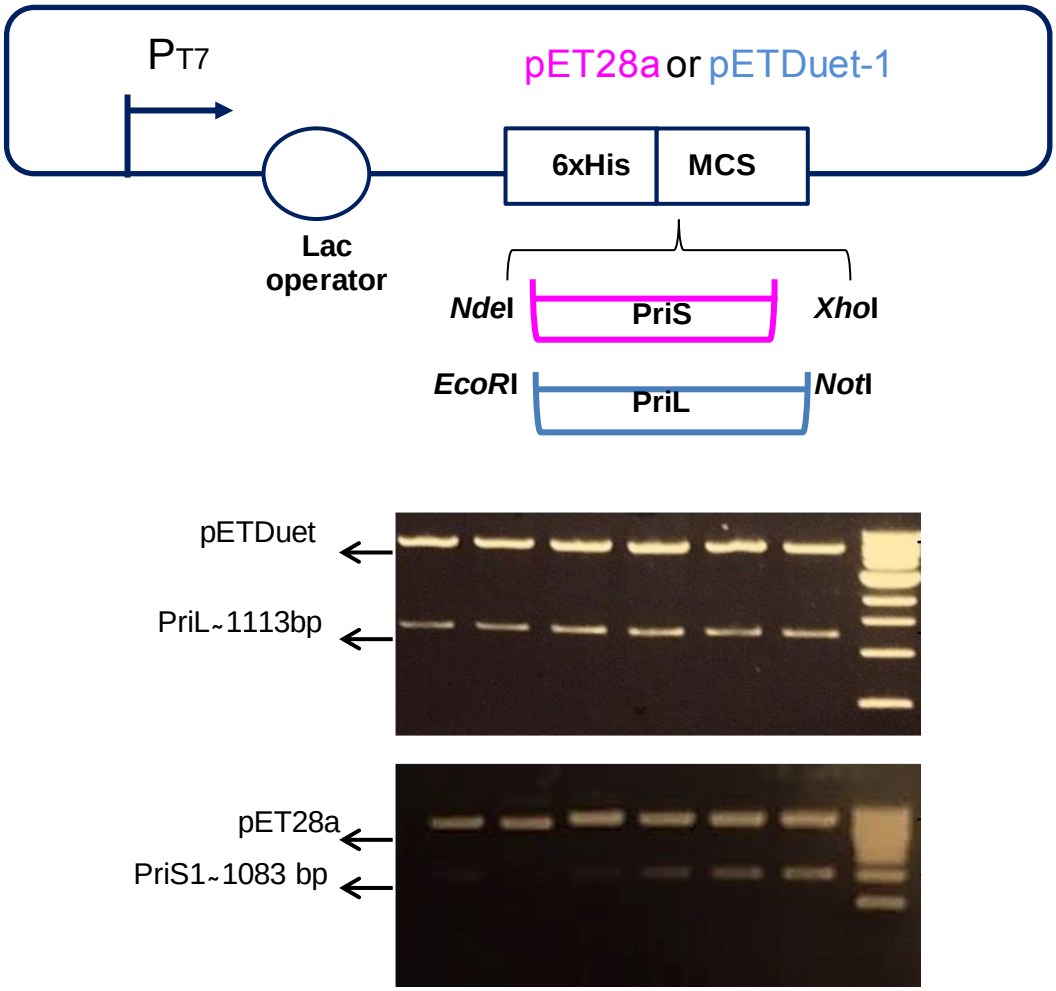


Figure 3.8. Cloning of Afu-PriS1 and PriL into pET28a and pETDuet-1 respectively

The open reading frames corresponding to PriS1 and PriL primase subunits were amplified from *A. fulgidus* genomic DNA (Table 2.3) introducing the applicable restriction sites to allow insertion into the multiple cloning site (MCS) and (MCS-2) of the pET28a and pETDuet-1 expression vectors respectively. A 6-histidine tag downstream of a T7 promoter (P_{T7}) was in frame with cloned PriS1 the PCR products were combined with 10x DNA loading dye and run on 1% agarose gels containing ethidium bromide.

3.4. Expression and purification of *A. fulgidus* PriS1/L

The Afu-PriS1 and Afu-PriL expression constructs were co-transformed into BL21 *E.coli*, which provides high-level expression of recombinant proteins and also improves cell lysis (Inouye *et al*.,1973;Moffatt and Studier,1987;Studier, 1991). To stabilise the zinc finger motif of the PriS1L, cultures were supplemented with the addition of 100mM zinc sulphate ($ZnSO_4$) to the media and expression was induced by addition of 1mM IPTG. The pET28a expression vector was utilised for expression as it allowed a (6x-His) affinity tag to be fused to the N-terminus of PriS1L thus allowing for purification by affinity of the imidazole rings to nickel affinity resin. Elution of recombinant PriS1L from the nickel affinity chromatography column was induced by addition of imidazole. Eluted PriS1L was resolved on an SDS-polyacrylamide gel to confirm the correct size. The result was in agreement with their predicted molecular masses. PriS1 has a predicted molecular mass of ~ 41 kDa and PriL has a predicted molecular mass of ~ 42 kDa (Figure 3.9). PriS1L was further purified using heparin column chromatography, a pseudo-affinity column that is commonly used for purification of proteins with affinity for DNA. The eluted peak fraction from the $Ni2^{+}$-NTA column was first diluted 10 fold into a buffer with 100mM NaCl to decrease the salt concentrations and then applied to a heparin chromatography column. The column was equilibrated with Heparin Buffer A (150 mM NaCl, 40 mM Tris-HCl pH 7.5, 10% (v/v) glycerol). After loading the $Ni2^{+}$-NTA eluted fractions into the column, buffer A was used to wash the column to elute any non-specifically bound proteins. Proteins were eluted by gradient elution up to 50% Heparin Buffer B (1 M NaCl). Fractions containing protein were resolved on SDS-PAGE gel and concentration was determined by absorbance at 280 nm. For the final step of purification, to remove low molecular weight contaminants such as salt, size-exclusion chromatography (Gel filtration) was performed (Figure 3.9). Gel filtration columns are used for separation of molecules with different molecular sizes. An S75 gel-filtration column was pre-equilibrated with gel filtration buffer (40 mM Tris-HCl (pH7.5), 100 mM NaCl, 10% (v/v) glycerol and 2mM β-mercaptoethanol) and then the concentrated protein was loaded onto the column. Finally, fractions were collected following 100 mL of flow-through and

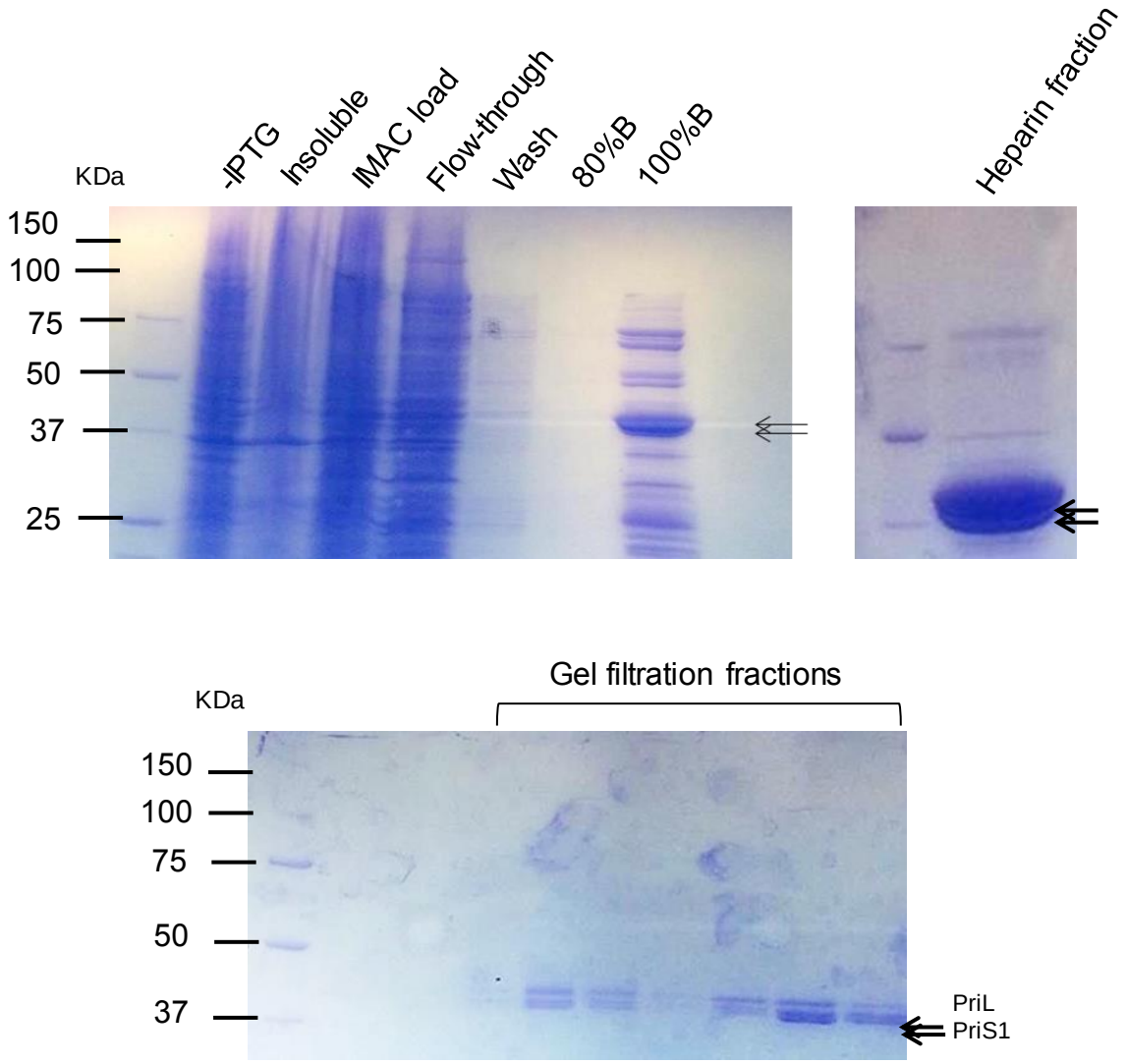


Figure 3.9. The chromatography purification of the Afu-PriS1/L complex

Purification of Afu-PriS1L complex (PriS1: 41 KDa and PriL: 42 KDa). A three litre culture of *E.coli* BL21 transformed with both pET28a: Afu-PriS1 and pETduet-1: Afu-PriL expression constructs was grown at 37°C for 3 hours then induced with 1 mM IPTG and incubated overnight. Following cell lysis, soluble cell lysate loaded onto Ni^{2+}-NTA chromatography. Bound Afu- PriS1/L was washed and then eluted with 300 mM imidazol. The 100% eluted peak fraction from the Ni2+-NTA column then subjected to heparin column, successfully eluting Afu-PriS1/L with 1M NaCl. In the last step of purification, protein containing fractions were subjected to S75 gel-filtration column which was pre-equilibrated with 300 mM NaCl. Finally, fractions were collected following 100 mL of flow-through and analysed by SDS-PAGE gel.

to confirm the correct size resolved on SDS-PAGE gel and determined by absorbance at 280 nm. We also attempted to co-transform the pET28a:Afu-PriS1 and pETduet-1:Afu-PriL expression constructs into Rosetta *E.coli* strains that possess additional tRNAs used in eukaryotes but rarely used in *E. coli* (data not shown). However, we observed that BL21 *E. coli* cells offered a significantly greater yield of PriS1/L. In order to compare the biochemical properties of PriS1/L complex with each subunit, pET28a: Afu-PriS1 and pETduet-1: Afu-PriL expression constructs were transformed into BL21 strain individually and purified in the same manner as the Afu-PriS1L complex (Figure 3.10). Purification of both Afu-PriS1 and Afu-PriL subunits started with the loading of the prepared soluble cell lysates onto a Ni2+-NTA agarose affinity chromatography that was pre-equilibrated with buffer A with 500 mM NaCl. The bound PriS1 and PriL eluted over a range of imidazole concentrations. Successfully eluted PriS1 and PriL was then analysed by SDS-PAGE to confirm the correct size of each primase subunit. Heparin affinity chromatography was then used to remove contaminating proteins. Both the Afu-PriS1 and Afu-PriL and *E.coli* contaminants were bound to the column. The contaminant was eluted with low salt concentration while both PriS1 and PriL eluted at 1 M NaCl. Following analysis of protein containing fractions using SDS-PAGE, some bands corresponding to *E. coli* contamination were observed in the case of PriL, therefore, size exclusion chromatography was utilized to remove the contaminants. The S75 gel-filtration column was used. After concentration, protein was loaded using a 5 mL loop. Finally, fractions were collected following 100 mL of flow-through and to confirm the correct size resolved on SDS-PAGE gel and determined by determined by absorbance at 280 nm.

.

3.5. Expression and purification of *P. furiosus* and *M. maripalidus* PriS1/L

Replicative primase (PriS1/L) from *P. furiosus* and *M.maripalidus* was purified to compare our results with PriS1/L from two other archaeal species, one hyperthermophile and one mesophile. These two complexes were purified in the same manner as the Afu-PriS1/L. Cultures of *E.coli* BL21 were transformed with either the *Mma*-PriS1/L or *Pfu*-PriS1/L (co-transformation of PriS1 and

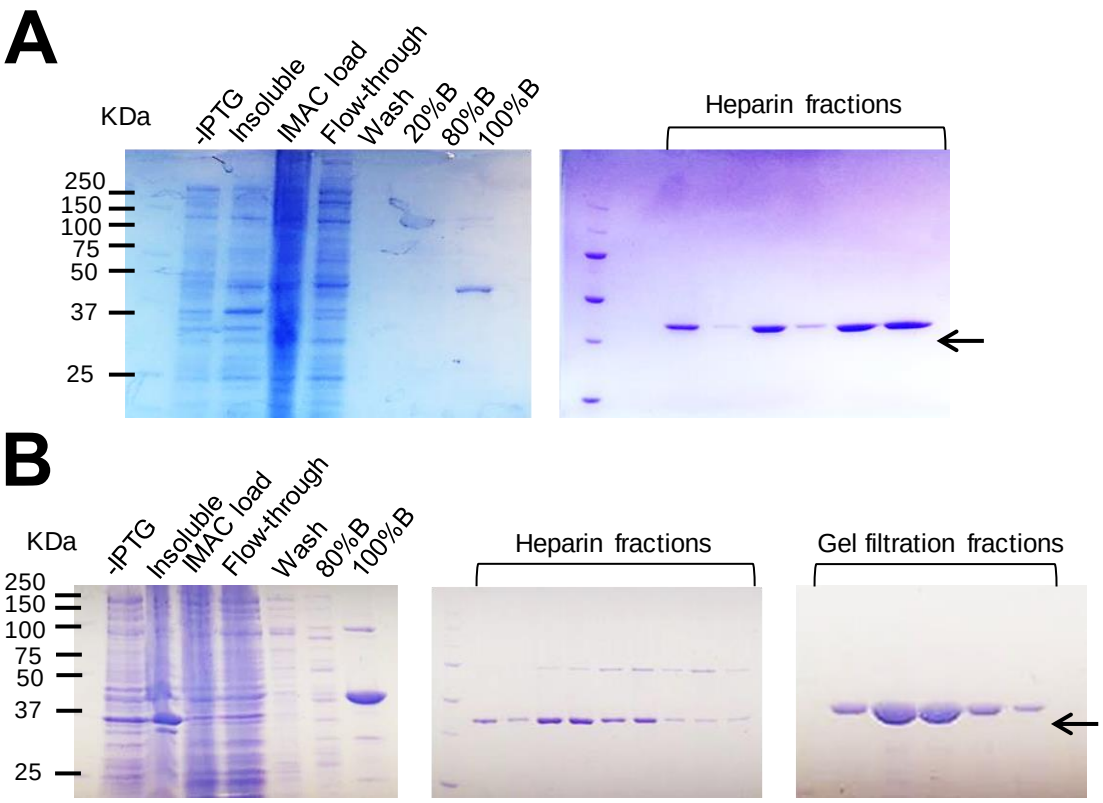


Figure 3.10. Purification of Afu-PriS1 and PriL subunits

Each pET28a: Afu-PriS1 and pETduet-1:Afu-PriL plasmid was purified individually. **(A)** Illustrates SDS-PAGE analysis of Afu-PriS1 (41 KDa) purification. **(B)** Illustrate SDS-PAGE analysis of the Afu-PriL (42KDa) purification. Soluble cell lysate from each PriS1 and PriL was subjected to Ni^{2+}-NTA chromatography. Bound Afu- PriS1 and PriL were washed and then eluted with 300 mM imidazol. The eluted peak fractions from the Ni2+-NTA column then subjected to heparin column, successfully eluting both Afu-PriS1 and PriL with 1M NaCl. SDS-PAGE analyses of heparin affinity purification showed *E.coli* contamination in case of Afu-PriL, so the PriL was subjected to S75 gel-filtration column which was pre-equilibrated with 300 mM NaCl to remove the contaminants. Finally, fractions were collected following 100 mL of flow-through and analysed by SDS-PAGE gel.

PriL) expression constructs, and then grown until exponential phase and protein expression was induced by the addition of 1 mM IPTG. The soluble cell lysatewas prepared and subjected to Ni^{2+}-NTA agarose affinity chromatography. To concentrate the eluted fractions as much as possible, a single elution of 300 mM imidazole was performed. This successfully eluted protein was resolved on an SDS-polyacrylamide gel to confirm the correct size and purity. The results were in agreement with their predicted molecular masses. Pfu-PriS1 has a predicted molecular mass of ~ 41 kDa and Pfu-PriL has a predicted molecular mass of ~ 46 kDa (Figure 3.11A). Mma- PriS1 has a predicted molecular masst of ~ 42 kDa and Mma-PriL has a predicted molecular mass of ~ 43 kDa (Figure 3.11B).

Heparin affinity chromatography was used to remove *E. coli* contaminants for both species. The eluted peak fraction from the Ni^{2+}-NTA column was first diluted 10 fold into a buffer with 100mM NaCl to decrease salt concentration and then applied to a heparin chromatography column. The column was equilibrated using pre-chilled Heparin Buffer A (150 mM NaCl, 40 mM Tris-HCl pH 7.5, 10% (v/v) glycerol). The Ni^{2+}-NTA eluted fractions were subjected to Heparin column and subsequently column was washed with Heparin Buffer A. Proteins were eluted by gradient elution up to 50% Heparin Buffer A with 1 M NaCl. Fractions containing protein were resolved on SDS-PAGE gel and determined by A_{280} level. Mma-PriS1/L was further purified using gel filtration column. S75 gel-filtration column was pre-equilibrated with gel filtration buffer (40 mM Tris-HCl (pH7.5), 300 mM NaCl, 10% (v/v) glycerol and 2mM β-mercaptoethanol). After concentration protein was loaded using a 5 mL loop. Finally, fractions were collected and determined by absorbance at 280 nm. The results were in agreement with their predicted molecular masses. Pfu-PriS1 has a predicted molecular weight of ~ 41 kDa and Pfu-PriL has a predicted molecular mass of ~ 46 kDa (Figure 3.11A). Mma- PriS1 had a predicted molecular mass of ~ 42 kDa and Mma-PriL had a predicted molecular mass of ~ 43 kDa (Figure 3.11B).

3.6. Archaeal replicative primase is an active polymerase

Previous studies indicated that PriS/L complexes from different archaea, including *Pyrococcus horikoshii*, *Pyrococcus furiosus*, *Sulfolobus solfactaricus* and *Thermococcus kodakaraensis* are active DNA polymerases (Bocquier *et al*., 2001; Lao-Sirieix *et al*., 2004). We examined the polymerase activity of

Archaeoglobus fulgidus PriS1/L using standard primer extension assays. Primer extensions employ a DNA oligonucleotide primer containing a fluorescent moiety (HEX) at the 5' end annealed to a longer DNA oligonucleotide template, yielding ds DNA with a 5' over-hang (Figure 3.12). In the presence of dNTPs, DNA polymerase replicates the template and produces labelled products. In the first experiment, the primer-template substrate was incubated with dNTPs, in a buffer containing magnesium and either Afu-PriS1 and Afu-PriL. PriL alone resulted in no extension of the labelled primer (Figure 3.12A). However, PriS1 alone exhibited limited extension of the primer (Figure 3.12A). However, incubation of the purified Afu-PriS1/L complex with the primer-template substrate and dNTPs resulted in full extension of the primer (Figure 3.12B). Similar to *A. fulgidus*, the PriS1/L complex from *M. maripalidus* was able to extend the primer in the presence of dNTPs and Mg^{2+} (Figure 3.12B). The DNA polymerase activity of the purified Pfu-PriS1/L complex was tested as a positive control enzyme (Figure 3.12B). However, our data suggested that the polymerase activity of Pfu and Mma PriS1/L was much slower than that of Afu-PriS1/L. In agreement with previous *in vitro* studies, these results indicate that the large subunit alone displays no polymerase activity and both small and large subunits interact and form a complex *in vitro* which is required for efficient polymerase activity.

3.7. Primase activity of archaeal replicative primase

It was next tested whether PriS1/L from *A.fulgidus* and *M.maripalidus,* similar to *P. furiosus* and some other archaea, could synthesise primers using both dNTPs or rNTPs. Different primase assays were carried out to confirm the primase activity of PriS1/L.

To determine whether PriS1/L, similar to eukaryotic primases, had DNA substrate preference, each of the four homopolymeric templates were used to study the primase activity of the PriS1/L. 60-mer 5'-biotinylated homopolymeric templates (poly-dA, poly-dC, poly-dG and poly-dT) were used (see sequences 7-10 in Table 2.7) to evaluate whether, in the presence of either rNTPs or dNTPs, primers were synthesised *de novo* on these templates. Primers were extended using Klenow

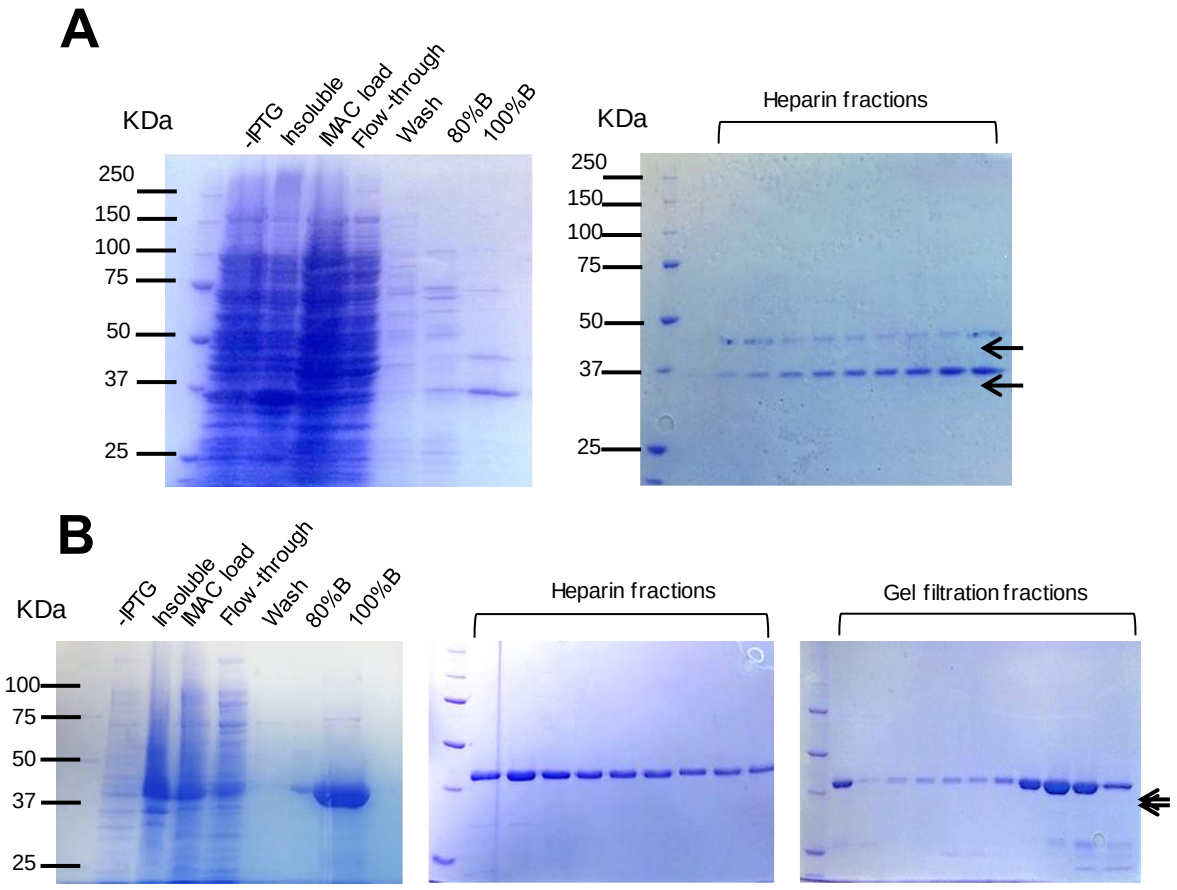


Figure 3.11. Purification of Pfu-PriS1L and Mma-PriS1L

(A) SDS-PAGE analysis of Pfu-PriS1L purification (PriS1: 41KDa and PriL: 46KDa). Samples from **(B)** SDS-PAGE analysis of Mma-PriS1L (PriS1: 42KDa and PriL: 43KDa) purification. Cultures of *E.coli* BL21 transformed with either Pfu-PriS1/L or Mma-PriS1/L expression (co-transformation of PriS1 and PriL). Cultures of both Pfu-PriS1/L and Mma-PriS1/L were grown at 37°C for 3 hours then induced with 1 mM IPTG and incubated overnight at 20°C. Following cell lysis, soluble fractions loaded to Ni^{2+}-NTA chromatography. Bound Pfu- PriS1/L and Mma-PriS1/L were washed and then eluted with 300 mM imidazol. The eluted peak fractions from the Ni2+-NTA column then subjected to heparin column, successfully eluting Pfu-PriS1/L and Mma-PriS1/L with 1M NaCl. Fractions containing protein were analysed by SDS-PAGE to confirm the size of proteins. Mma-PriS1/L was further purified using S75 size-exclusion chromatography which was pre-equilibrated with 300 mM NaCl to remove the contaminants. Finally, fractions were collected following 100 mL of flow-through and analysed by SDS-PAGE gel.

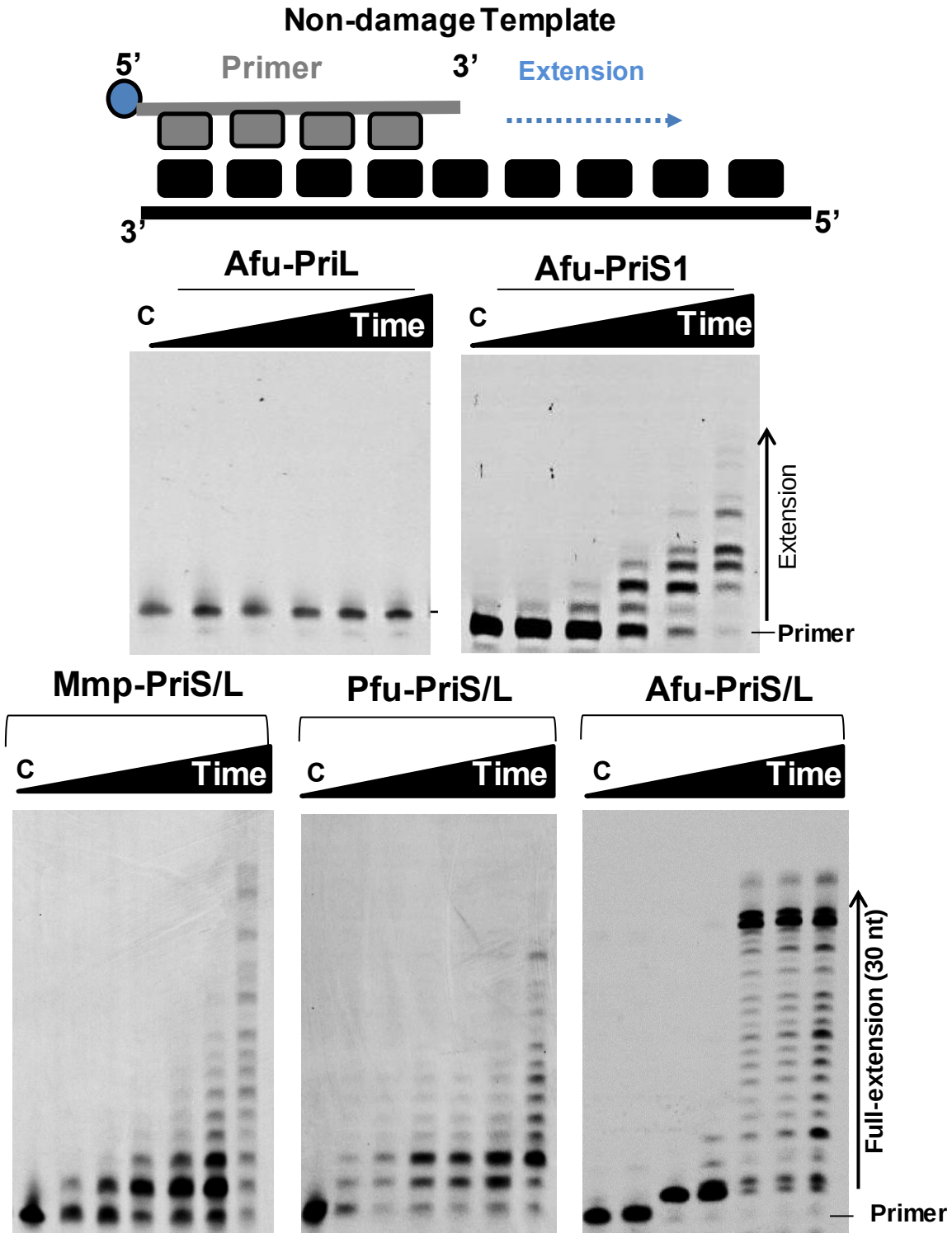


Figure 3.12. Afu-PriS/L is an active primase and polymerase.

(A) DNA primase was incubated with substrate 1 from Table 2.7. (Shown schematically) dNTPs and a reaction buffer containing $MgCl_2$ at 0.5 , 1, 3, 5, 10 and 30 minute time points alongside a control containing no protein. Afu-PriS1 showed limited polymerase activity on undamaged oligonucleotide substrate. No activity was observed with Afu-PriL. **(B)** Incubation of Afu-PriS1/L complex with dNTPs and reaction buffer resulted in significant extension of primer. Both Pfu-PriS1/L and Mma-PriS1/L exhibited efficient extension of undamaged DNA using dNTPs.

Taq supplemented with a fluorescent nucleotide analogue to allow for the fluorescent labelling of primers. The reactions were terminated and added to streptavidin-coated beads. After washing the beads to remove excess nucleotides, reactions were resolved on a polyacrylamide gel and detected by fluorescence (Keen *et al*., 2014).

As previously discovered, the archeal Pfu-PriS/L complex can synthesise primers using both dNTPs and rNTPs, but preferentially utilize deoxyribonucleotides (dNTPs) to synthesize DNA primers (Bocquier *et al*., 2001). Similar to *P. furiosus*, the PriS/L complex from *Sulfolobus solfataricus* is also capable of utilizing both ribonucleotides and deoxyribonucleotides. However, in contrast to *P. furiosus*, it has significantly higher affinity for rNTPs than dNTPs (Lao-Sirieix *et al*., 2004). Here, using a non-radioactivity primase assay, PriS1/L from all three species exhibits very limited primase activity on homopolymeric templates using rNTPs. However, Pfu-PriS1/L exhibited primase activity in the presence of both rNTPs and dNTPs (Figure 3.13). The four homopolymeric substrates were used with varying efficiencies by Afu-PriS1/L, with oligo (dG) being the least preferred substrate and oligo(dC) the most preferred one. In contrast, the preferred substrate of Pfu-PriS1/L was oligo (dA). Notably, the detected primase activity by Mma-PriS1/L was not as significant as activities observed by Afu and Pfu enzymes (Figure 3.13). It was previously shown that PrimPol, a newly discovered eukaryotic archaeo-eukaryotic primase, is able to catalyse *de novo* synthesis of primer strands on a 65-mer template containing G-quadruplex (G4) (Schiavone *et al*., 2016). We decided to examine the primase activity of archaeal PriS1/L using the same template. The PriS1/L proteins were incubated with 5' biotin-labelled template and reaction buffer. Reactions were incubated with either rNTPs or dNTPs and FAM dNTPs for 30 minutes. After quenching the reactions, DNA was bound to streptavidin-coated beads for 1 hour and then the beads were washed and suspended in stop buffer prior to resolving on a polyacrylamide gel and detection by fluorescence. Although Pfu-PriS1/L exhibited primase activity on the 65-mer template, no primase activity was detected by Afu-PriS1/L and Mma-PriS1/L (Figure 3.14). Notably, the observed primase activity of Pfu-PriS1/L was only in the presence of rNTPs. Our data suggested that maybe other assays are required to detect the primase activity. We therefore next

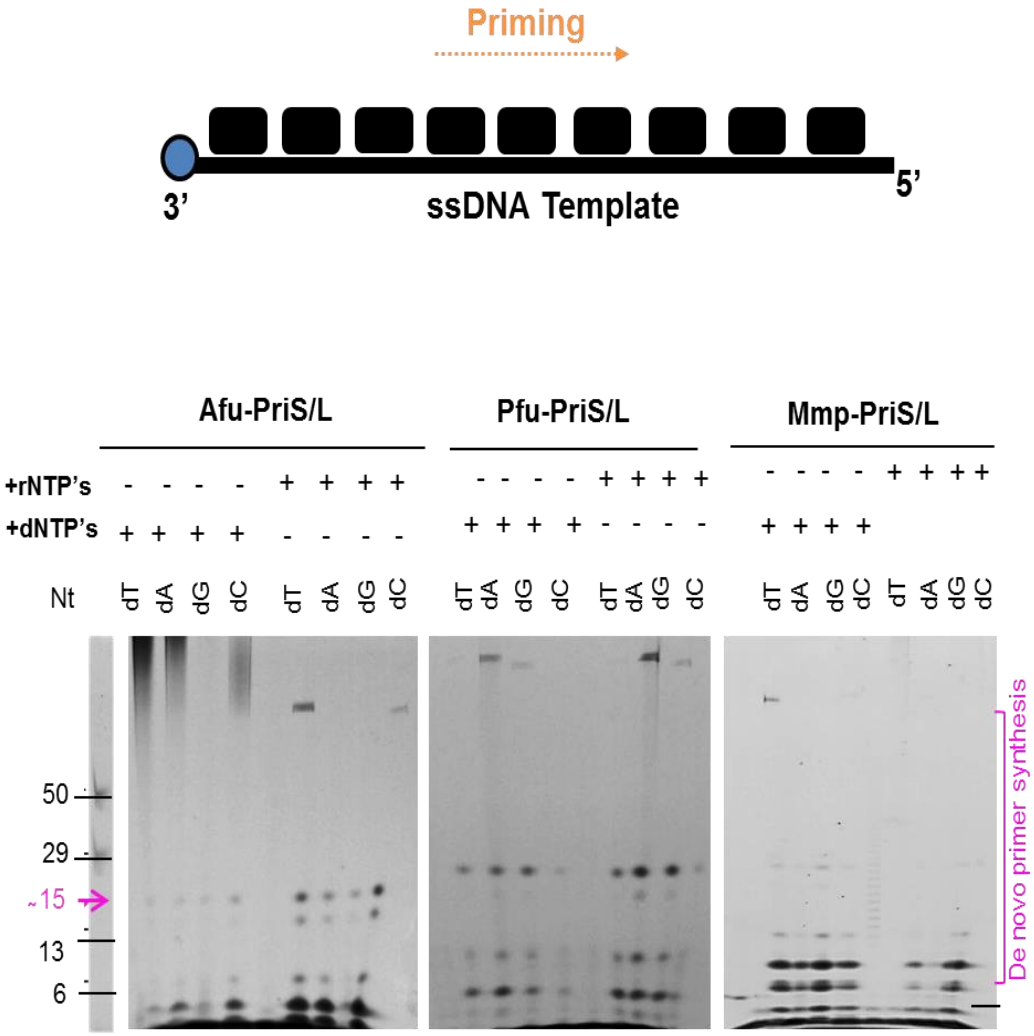


Figure 3.13. DNA primase activity of PriS1/L

Single-stranded homopolymeric templates (see sequences 7-10 in Table 2.7) were incubated with dNTPs or rNTP, 10 mM $MgCl_2$, 20 mM NaCl, and either Afu-PriS1/L or Pfu-PriS1/L or Mma-PriS1/L for 1 hour. Not very significant primase activity was observed with all three proteins. Afu-PriS1/L has primase activity similar to the Pfu-PriS1/L (around 15 Nt). However, the primase activity observed for Mma-PriS1/L was lower than the activity observed for Pfu-PriS1/L. The left panel indicates the oligonucleotide nucleotide (Nt) length markers.

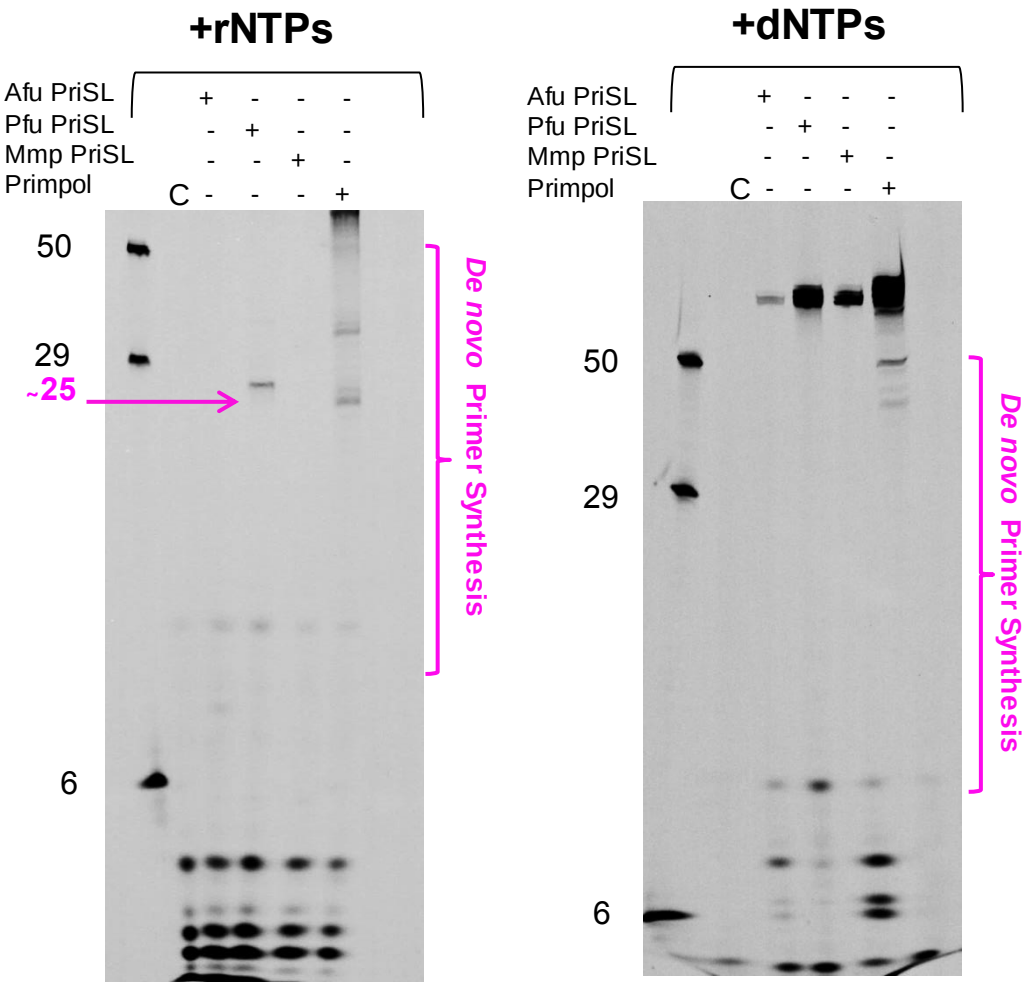


Figure 3.14. PriS1/L primase activity with on a synthetic DNA template

PriS1/L from three archaeal species were incubated for 30 minutes with rNTPs (250 µM) , FAM-dNTPs (dATP, dCTP, dUTP) (2.5 mM), 10 mM $MgCl_2$, and 65-mer biotin-labelled DNA template (1µM). Identical reactions were also performed with dNTPs (250 µM) instead of rNTPs. Oligonucleotide nucleotide (Nt) length markers are shown in left. 2 µM of PrimPol was used as the positive control. Only Pfu-PriS1/L shows *de novo* production activity using rNTPs. Neither AfuPriS1/L nor Mma-PriS1/L exhibit *de novo* primase activity.

tested whether PriSL complexes possessed primase activity on longer non-synthetic single-stranded M13 DNA templates. The M13 ssDNA was incubated with reaction buffer containing $MgCl_2$ and either Afu-PriS1/L, Pfu-PriS1/L or Mma-PriS1/L and rNTPs or dNTPs including [α-^{32}P] dATP/dCTP for 30 minutes. Identical reactions were also performed in the presence of $MnCl_2$ instead of $MgCl_2$.Reactions were then stopped and separated on a polyacrylamide gel. The gel was exposed to a phosphor screen and scanned and analysed using an FLA-1500 scanner (Fuji).

All three purified archaeal protein complexes displayed primase activity on a single-stranded M13 DNA. Afu-PriS1/L and Pfu-PriS1/L can synthesise primer strands in the presence of either Mg^{2+} or Mn^{2+}, however, both primases were more active in the presence of Mg^{2+} for RNA synthesis, yielding significant products that were not generated in the presence of Mn^{2+} (Figure 3.15). Detected primase activity by Mma-PriS1/L was only visible in the presence of Mg^{2+} (Figure 3.15). From this experiment alone, it was immediately evident that these enzymes showed a preference to prime using rNTPs over dNTPs.

Together, our data provide experimental evidence that the replicative primase complexes (PriS/L) from *A.fulgidus* and *M.maripalidus* can utilize deoxyribonucleotides and ribonucleotides for *de novo* primer synthesis. Similar to *T. kodakaraensis* and *S. solfataricus* these replicative primases select rNTPs in preference to dNTPs. Given that PriS1/L complexes from Afu, Pfu and Mma are able to initiate and elongate RNA and DNA strands, it seems reasonable to assume that, these enzymes can be categorized as primase-polymerases (Prim-Pols).

3.8. *A.fulgidus* PriS1/L can bind DNA

Since the PriS1/L complex is capable of extending primers at primer-template junctions, we tested whether the Afu-PriS1/L binds to ssDNA, dsDNA, or both. Gel electrophoresis mobility shift assays (EMSAs) were used with increasing concentrations of protein to examine the DNA binding ability of PriS1/L. Interestingly, PriS1/L was able to bind to both ssDNA and dsDNA with approximate comparable affinities. However, quantifications of protein-DNA binding suggest a little more preference for dsDNA over ssDNA by PriS1/L

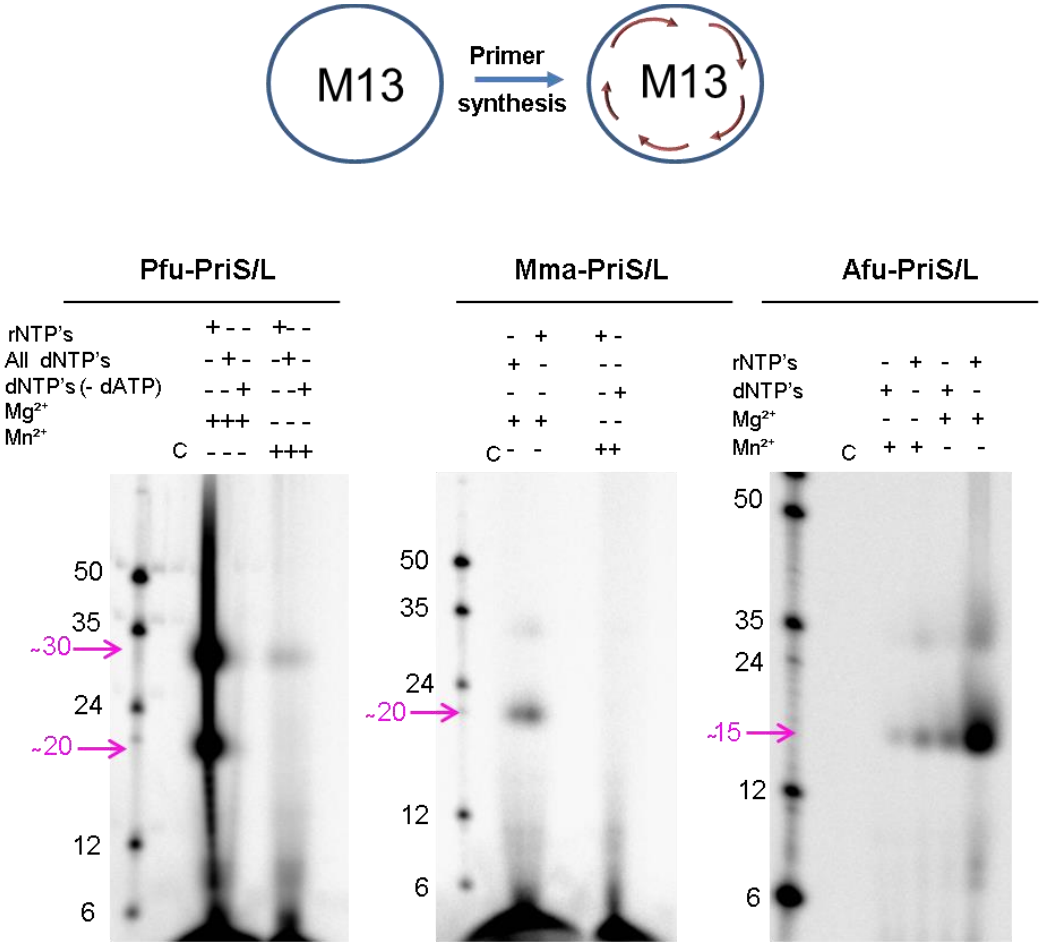


Figure 3.15. PriS1/L complex is a RNA/DNA primase

Primers synthesis on M13 ssDNA using either rNTPs or dNTPs. A scheme of primase reaction is shown. Afu-PriS1/L, Pfu-PriS1/L and Mma-PriS1/L were all analysed for RNA/DNA-dependent primase activity on M13 ssDNA. PriS1/L proteins were incubated with either 10 mM of $MgCl_2$ or 4 mM of $MnCl_2$ and 250 ng of M13 ssDNA. Samples were resolved on a polyacrylamide gel along with a control (C) containing no protein.

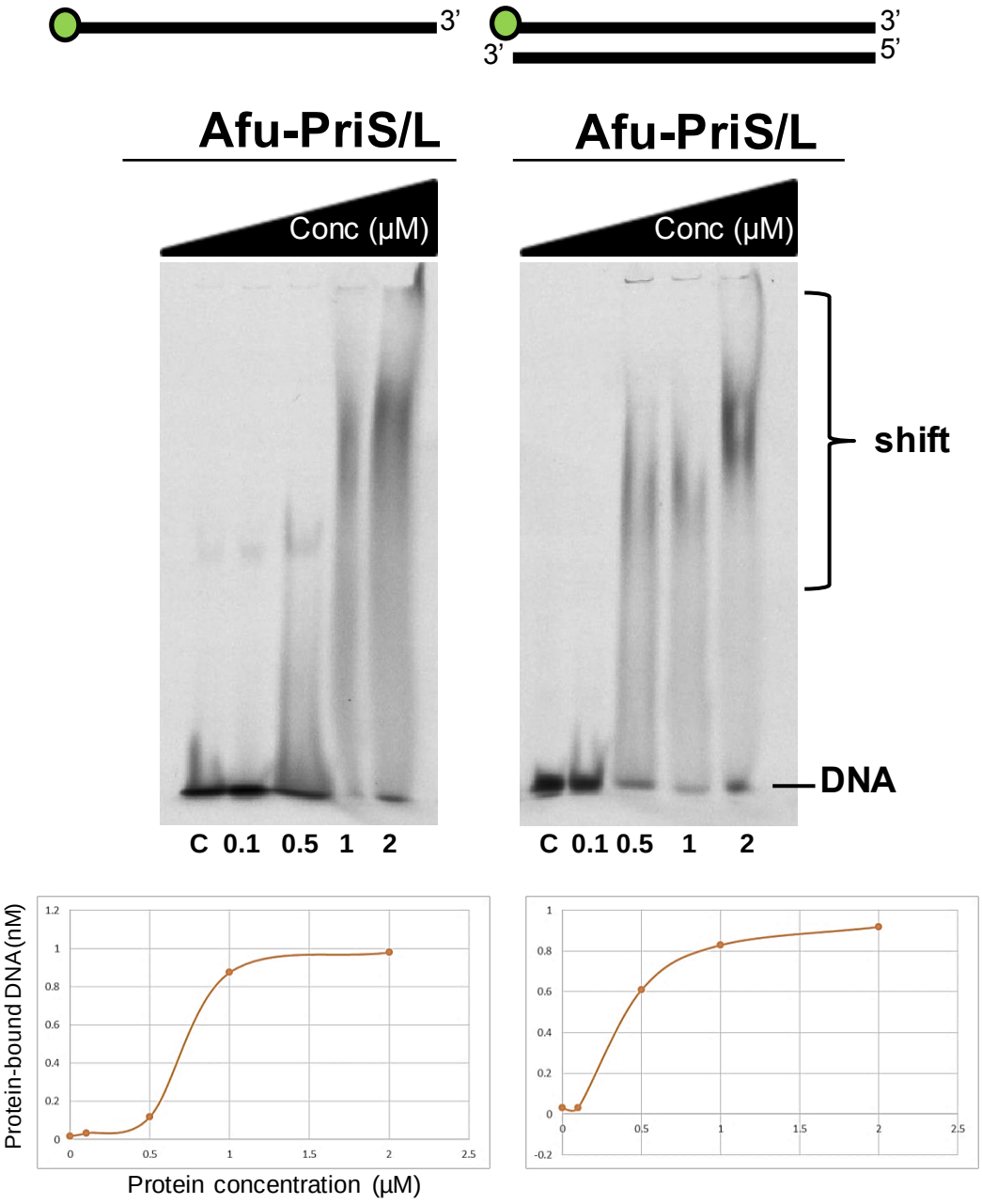


Figure 3.16. *A.fulgidus* PriS1/L binds to both ssDNA and dsDNA

DNA binding activity of PriS1/L complex was analysed on 60 nM of both single-stranded and double-stranded templates (Substrates 2 and 3 from Table 2.7) by EMSA. Increasing concentrations of protein were used as indicated. Although PriS1/L is able to bind both templates with approximate comparable affinities, the shifts of the fluorescent oligonucleotide templates on the polyacrylamide gel and quantifications suggest a little more binding affinity for dsDNA than ssDNA by PriS1/L. 2μM protein concentration showed highest DNA binding activity in both cases.

(Figure 3.16). Binding was apparent at ~0.5 µM and a complete binding was observed at 2 µM of PriS1/L (Figure 3.16). This DNA binding specificitymight suggest the presence of two closely clustered binding sites which allow the enzyme to recognize a primer-template junction. This data is consistent with the DNA binding ability of the archaeal primase complex isolated from *P. furiosus*, which is able to bind both ssDNA and dSDNA (Liu *et al*., 2001). It has also been previously demonstrated that the AEP polymerase domain of PrimPol also binds to both ssDNA and dsDNA (Keen *et al*., 2014).

3.9. *A.fulgidus* PriS1/L binds to homopolymeric ssDNAs

It was next tested whether Afu-PriS1/L has any binding preference for homopolymeric DNA templates. EMSAs were performed on 50-mer hexachlorofluorescein labelled single-stranded homopolymeric DNA substrates (poly-dA, poly-dC, poly-dG and poly-dT). The PriS1/L complex showed significant affinity to poly-dC and poly-dT (Figure 3.17). This interaction is consistent with Afu-PriS1/L's preference to synthesise *de novo* primers on polypyrimidine templates (dC and dT). PriS1/L showed affinity for poly dA, but it was not as significant as other homopolymeric templates. Interestingly, PriS1/L exhibited super-shifted intermediates with poly dG, which suggested that Afu-PriS1/L might have sequence or structural specificity for dG homopolymers (Figure 3.17). Poly dG sequences can form G-quadruplexes (G4), four-stranded helical structures, which have been shown to form just upstream of replication origin sites, which may reflect their regulatory role in activation of replication origins (Cayrou *et al*., 2012). Interestingly, it was recently demonstrated that PrimPol has an affinity for Poly-dG templates and can reprime downstream of G4 quadruplexes (Schiavone *et al*., 2016). Together, these findings shed a light on possible additional roles of archaeal PriS/L in repriming after G4 quadruplexes and origin firing.

3.10. Summary and discussion

In this chapter, we have characterized some of the biochemical properties of the DNA primase complex isolated from a hyperthermophilic archaeon, *Archaeoglobus fulgidus* and a mesophilic archaeon, *Methanococcus maripaludis*. We cloned, isolated, and characterised both *A.fulgidus* and *M.maripaludis* PriS1/L complexes.

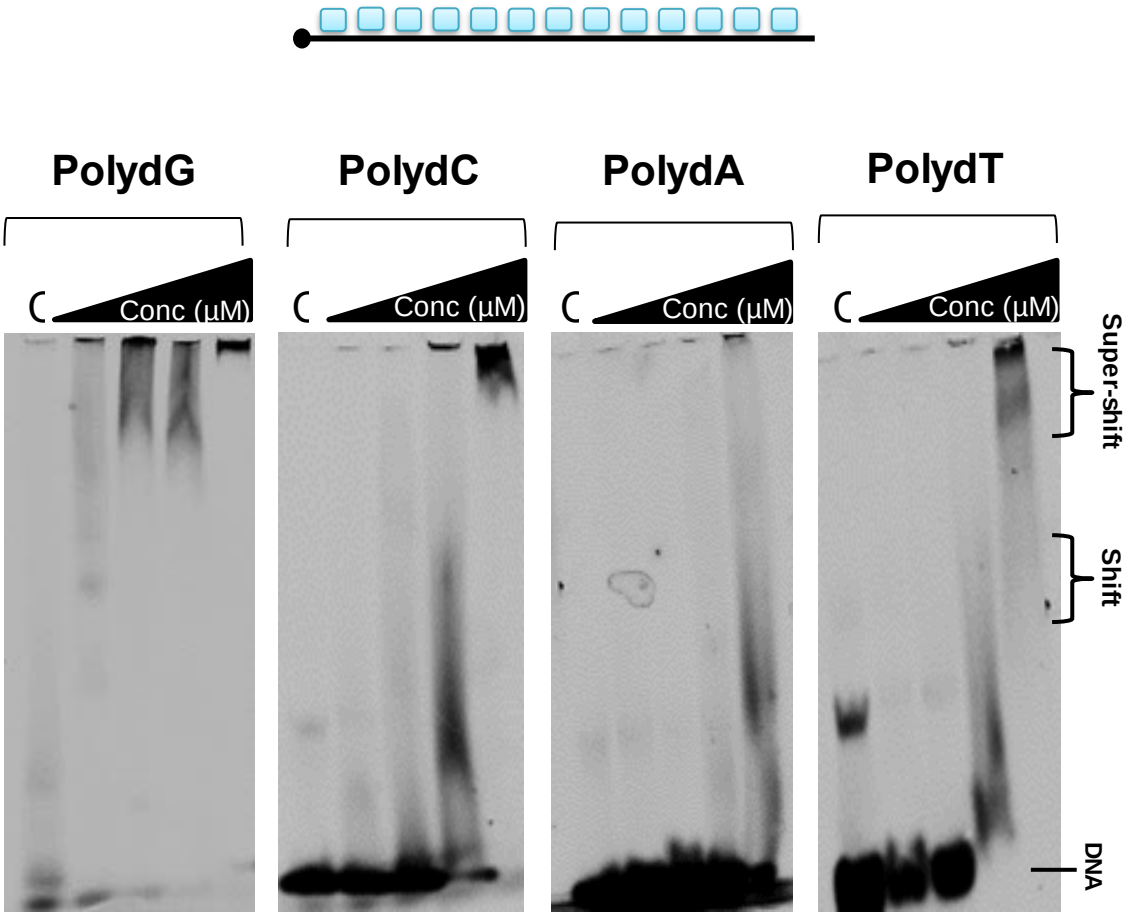


Figure 3.17. Afu-PriS1/L binds homopolymeric ssDNA templates

DNA binding preference of Afu-PriS1/L was analysed on 50-mer homopolymeric single-stranded templates. The EMSA binding patterns for increasing concentrations of Afu-PriS1/L (0.1, 0.5, 1 and 2µM) incubated with homopolymeric ((dA), (dC), (dG) and (dT)) is illustrated. PriS1/L showed higher DNA binding activity on poly-dC and dT templates. The DNA binding affinity of PriS1/L for poly-dA was not as significant as other templates. The highest DNA binding affinity PriS1/L was for poly-dG template.

Unlike bacterial monomeric DnaG, in archaea primases consist of two subunits that are shared by archaea and eukaryotes and compose of the core replicative primase.

In both archaea and eukaryotes, the small subunit of the primase performs the catalytic activity (Bocquier *et al*., 2001). Interestingly, in *A.fulgidus* the there are two small subunits (PriS1 and PriS2). Previous studies indicated that the interaction of the large subunit with the small subunit of DNA primase regulates and stabilizes the primase activity of small subunit (Liu *et al*., 2001). The data presented in this chapter demonstrates that PriS1 and PriL subunits from *Archaeoglobus fulgidus*, a genus of the phylum Euryarchaeota, interact and form a stable complex (PriS1/L). However, the possible interaction between PriS2 and PriL was not tested. To date, many studies have identified the interaction between PriS and PriL subunits in different species of archeae (Galal *et al*., 2012; Lao-sirieix and Bell, 2004; Liu *et al*., 2001). A growing body of evidence indicates that many archaeal PriS/L complexes display DNA-dependent DNA polymerase activity and also are able to synthesise primers using both dNTPs and rNTPs. It was reported that the primase catalytic subunit from *M. janaschii*, in common with the eukaryotic primase, could synthesise RNA primers (Desogus *et al*., 1999). However, the PriS (Pfu41) subunit from the hyperthermophile *Pyrococcus furiosus,* in contrast with eukaryotic primase, can synthesise primers using dNTPs (Liu *et al*., 2001). Subsequently, the primase complex (p41-p46) was characterized and the RNA primase activity which was not observed with the PriS (Pfu41) subunit was detected. Here, we observed that the PriS1 subunit of DNA primase in the absence of PriS2, together with PriL subunit is able to accomplish primase activity. We showed that the PriS1/L complex from *Archaeoglobus fulgidus* performs both priming and primer extension. We confirmed that PriS1/L complexes from *Archaeoglobus fulgidus* and *Methanococcus maripaludis* are DNA-dependent DNA polymerases *in vitro*. Interestingly, *A.fulgidus* replicative primase showed higher polymerase activity compared to the *M. maripaludis* and *P. furiosus* replicative primases. Although *A.fulgidus* DNA primase possess two small catalytic subunits (PriS1 and PriS2) there is less known about the roles and importance of the PriS2. This subunit maybe required for more efficient primase activity. It is also possible that PriS2 in addition to primase activity has an

additional role during DNA replication. However, in many species of archaea the PriS2 does not exist.

Our data suggest that *A.fulgidus* PriS1/L has a preference for polypyrimidine templates d(C) and d(T) for primase activity. This template specificity is a common feature among AEPs. As previously indicated, AEPs from *M. janaschii* and *T. kodakaraensis* have a preference for polypyrimidine templates (Bocquier *et al*., 2001; Desogus *et al*., 1999; Galal *et al*., 2012b). Moreover, a study on a primase from *S. solfataricus* demonstrated that it was able to synthesise primers on a synthetic thymine-containing bubble structure that mimics early replication intermediates (De Falco *et al*., 2004). Recently, PrimPol, a novel eukaryotic AEP superfamily primase-polymerase was found to preferentially prime on homopyrimidine (poly dT) but not homopurine templates (Keen *et al*., 2014). We found that similar to most species of archaea, including *Pyrococcus horikoshii*, *Sulfolobus solfactaricus* and *Pyrococcus furiosus,* the PriS1/L complex from *Archaeoglobus fulgidus* was also capable of synthesising primers utilizing both rNTPs and dNTPs. However, it showed a preference to prime using rNTPs over dNTPs. These observations are in keeping with properties of *Sulfolobus solfataricus* primase complex (Lao-Sirieix and Bell, 2004).This promiscuous feature of archaeal replicative primases mimics the dual activities of the Polα / primase complex in eukarya. Therefore, it is possible that DNA replication in archaea initiates with synthesis of an RNA primer and continues by extension of the primer by PriS/L. Furthermore, it is plausible that other replisome components are required to regulate primase-polymerase activities of PriS/L.

The intracellular concentration of nucleotides in archaea is still unknown but the average intercellular concentration of nucleotides in mammalian cells is ~ 0.3 to 3 mm for rNTPs and 10 to 50 µm for dNTPs (Traut, 1994). The higher levels of rNTPs compared to dNTPs suggests that DNA primases likely select rNTPs to initiate primer synthesise, as reported for several archaeal species. Surprisingly, our data showed that Mma-PriS1/L could only synthesis RNA primers on a non-synthetic M13 ssDNA. This difference may be due to the diverse environments in which these species grow. Unlike *A.fulgidus* that grows in extremely hot environments, *M. maripaludis,* a methanogen species, lives in temperatures that

are more moderate. Therefore, diverse environments have caused that archaea to adopt separate protein features that are customized for each environment.

Frick and Richardson have presented a simple mechanism for primer synthesis (Frick and Richardson, 2001; Frick *et al*., 1999). Based on this model, there are two nucleotide-binding sites in DNA primase; the initiation site and the elongation site. In the first step, the primase binds to ssDNA. Once the primase encounters an appropriate initiation site it binds to two nucleotides, forms a metal-dependent covalent dinucleotide linkage and releases inorganic pyrophosphate. Next, the dinucleotide is moved to the initiation site and allows addition of the third nucleotide in the free elongation site. It has been strongly indicated that zinc finger motifs play significant roles in primases. This structural element which has been found in most DNA primases, is not only required for DNA binding, but also is important for primer synthesis. Similar to herpes virus UL52, in PrimPol, the zinc finger domain exhibited an essential role in primase activity. Studies on AEP UL52 protein reported that mutation of the zinc finger led to lose of primase activity (Biswas and S. K. Weller, 1999; Chen *et al*., 2005). Deletion of the zinc finger domain in PrimPol totally inhibits the primase activity but not the polymerase activity (Keen *et al*., 2014). Therefore, since zinc finger motifs bind DNA and their deletion inhibits primer synthesis, it is believed that this element together with the polymerase mode stabilise the binding of the template, polymerase and nucleotides, which lead to generation of primer. Although this study was not focused on the role of the zinc finger domain in archaeal primase, it is known that PriS/L like other AEP primases contains a zinc finger motif and as discussed here, PriS/L can bind to ssDNA and catalyse primer synthesis. Therefore, it would be rational to assume that the zinc-finger motif in archaeal PriS/L is also important for the synthesis of initial primer. This primer is then extended in a more processive way by the polymerase to produce a longer product (Figure 3.18).

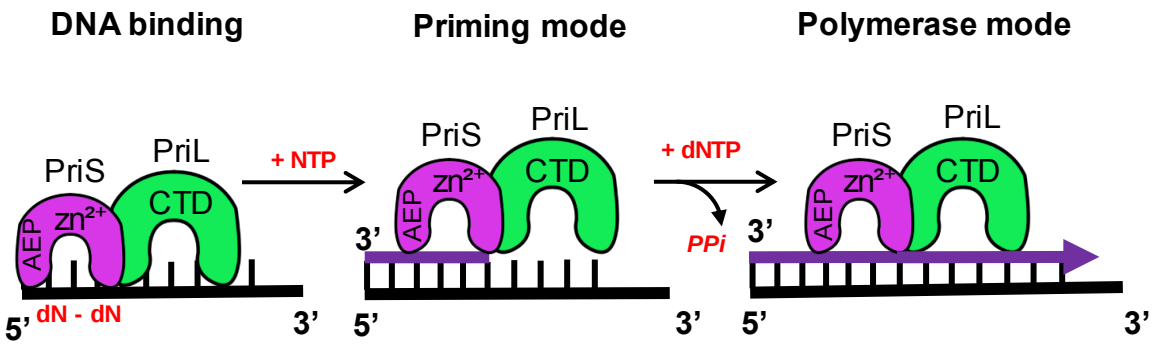


Figure 3.18. Archaea PriS/L performs primase and polymerase activities.

The model represents the catalytic activities of archaeal replicative primase. PriS/L complex performs both primase and polymerase activities. Upon binding the PriS/L complex to the specific ssDNA, the small catalytic subunit (PriS) which possess the zinc finger motif and AEP domain catalyses synthesis of a short RNA primer. The zinc finger domain highly likely stabilises the binding of the complex on the ssDNA also is required for synthesis of a dinucleotide. PriS together with the large non-catalytic accessory subunit (PriL) can bind to primer-template junction and in the presence of rNTPs or dNTPs are able to extend the short primer more processively to form a longer product. PriS/L Possesses two distinct synthesis modes, primase and polymerase, that are carried out under different DNA binding conditions.

Chapter 4

Archaeal replicative primases perform translesion DNA synthesis

4.1. Introduction

Although accuracy of DNA replication plays a crucial role in maintaining genome integrity in all living organisms, this process is continually threatened by various physical and chemical lesions and barriers (Aquilera and Gomez-Gonzales, 2008). Therefore, cells develop different repair pathways to rescue the stalled replication forks. Among these restart pathways translesion DNA synthesis (TLS), catalysed by specialized DNA polymerases, plays an important role in rescuing stalled replisomes (Sale *et al*, 2012). Another bypass mechanism is error-free recombination-mediated template switching, which is also able to rescue stalled replication forks (Li and Heyer, 2008). In addition, repriming downstream of lesions can also rescue progression of stalled replisome (Heller *let al*, 2006; Keen *et al*., 2014; Lopes *et al*, 2006; Schiavone *et al*., 2015). Specialized TLS DNA polymerases have been widely found in eukarya and bacteria, but surprisingly many archaeal species lack canonical TLS polymerases (Figure 4.1) (Kelman and White, 2006). Even though many key archaeal genes and several metabolic pathways are closely related to those of eukaryotes, majority of archaea lack canonical DNA repair pathways e.g. nucleotide excision or mismatch repair (Kelman and White, 2006). These differences lead us to question how archaeal species tolerate DNA damage in the absence of TLS or lesion repair pathways, specifically those that reside in extreme environmental conditions such as high temperature that can lead to deamination of cytosine into uracil. Archaeal replicative polymerase (PolB) stalls at deoxyuracil containing templates in order to avoid promutagenic bypass of these templates in which results in C–T transition (Richardson *et al*., 2013; Firbank, 2008). X-ray structural data has revealed that, a specific uracil binding pocket in the N-terminus of PolB mediates tight binding of PolB to the deaminated bases. In addition, inhibition of archaeal PolD activity by uracil was also suggested (Richardson *et al*., 2013). A recent study on family D DNA polymerases from *A.fulgidus* demonstrated that, uracil is able to limit the polymerase activity of Afu-PolD either in *cis*, when the deaminated base impedes replication of the strand on which it is situated, or in *trans*, when the deaminated base impedes replication of the alternate strand. However, the latter has not yet been observed with PolB (Abellon-Ruiz *et al*., 2016). Altogether, this poses a question as to how archaea restart replication after dU-induced replisome stalling?

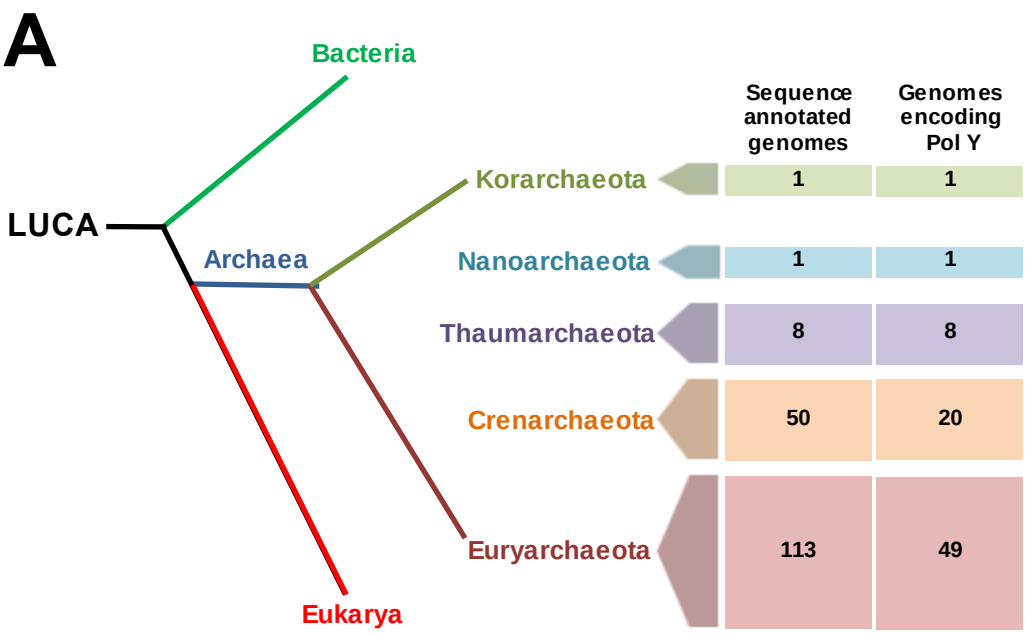
In this study, we demonstrate that archaeal replicative primases (PriS1 and PriS1/L) perform translesion DNA synthesis bypass of different DNA lesions. Based on our data which are discussed in this chapter it seems reasonable to assume that archaeal primases not only catalyse *de novo* primer synthesis, but are also involved in DNA damage tolerance and DNA repair.

The crystal structures of TLS (Y-family) polymerases suggest that, compared to replicative polymerases, the TLS polymerases possess smaller finger and thumb sub-domains, which leads to an open and accommodating active site (Yang, 2005). Since archaeal replicative primases are also competent TLS polymerases, it is expected that they have similar active sites. The more open, accessible active site can affect the polymerase's fidelity and its tolerance for mismatched bases. In this chapter, I explore these properties of PriS/L from three archaeal species.

As archaeal replicative primases have been found in all archaeal species, we chose this enzyme as a candidate to investigate if it is responsible for lesion bypass synthesis in archaea. Archaeal PriS catalytic subunit, a member of the AEP family, together with the PriL accessory subunit perform both primase and polymerase activities (Chapter 3). There is now significant evidence showing that AEPs play roles in DNA damage tolerance and repair, including the role of closely related NHEJ AEPs in DNA break repair (Bartlett *et al*., 2013., 2016; Pitcher *et al*., 2007b; Della *et al*., 2004; Weller *et al*., 2002). It has recently been shown that PrimPol is also a competent TLS polymerase. PrimPol efficiently bypasses a number of replication stalling lesions including 6-4 photoproducts, and 8-oxo-dG (Bianchi *et al*., 2013). Altogether, AEPs belong to a class of primase-polymerases called PrimPols to reflect their enzymatic activities and origins.

4.2. Purification of *Archaeoglobus fulgidus* PolB

In order to examine the TLS activity of archaeal replicative primase in the presence of stalled replicases, archaeal replicative polymerase (PolB) from *Archaeoglobus fulgidus* was expressed and purified.



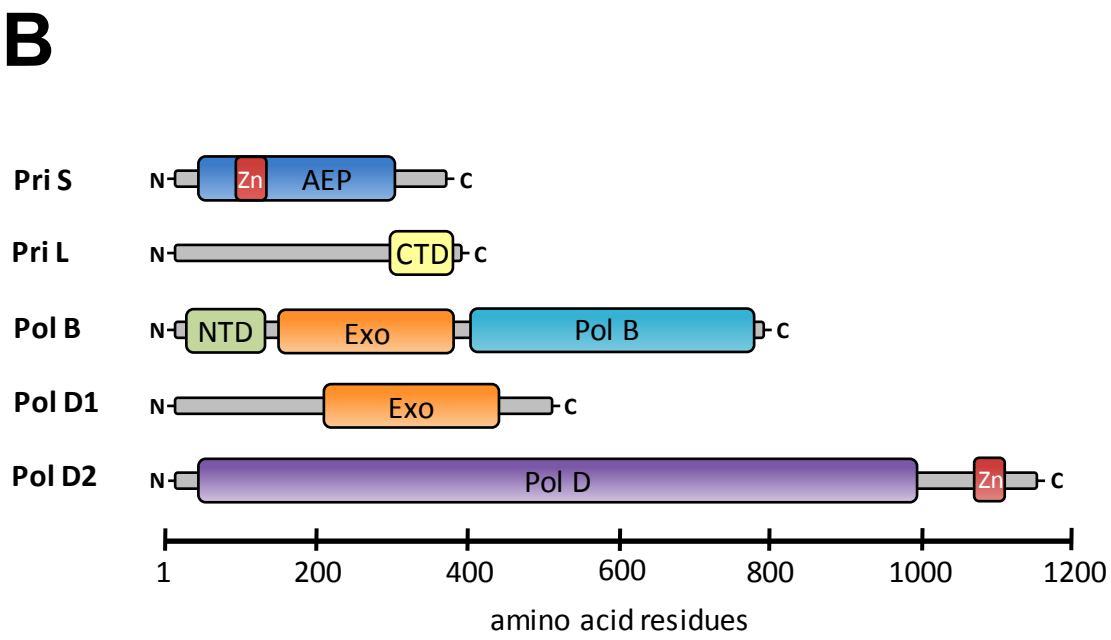


Figure 4.1. Many species of archaea lack canonical TLS polymerases. **(A).** The genes encoding Y-family polymerases are absent in majority of archaea**.** Only 79 archaeal genome among 173 encode canonical TLS DNA polymerases. **(B). The** structural elements of *A. fulgidus* replisomal enzymes analysed in this chapter.

pET28a:Afu-PolB and pET28a:Afu-PolD expression constructs were provided by my colleague Dr.Stanislaw Jozwiakowski (Table 2.3). Afu-PolB expression construct was transformed into the Rosetta *E.coli* strain. A culture of Afu-PolB was grown at 37°C until it reached exponential phase ($OD_{600=}$ ~0.6) followed by addition of 1 mM IPTG for 4 hours at 20°C. For thermostable enzymes like Afu-PolB, which can tolerate high temperature without denaturing, a heat step can be used to allow the enzyme to reform or redissolve. The prepared cell lysate was first subjected to heat denaturation for 20 minutes at 70°C and, after cooling down on ice, the denatured proteins can be removed by centrifugation. The cell lysate was then loaded onto a Ni^{2+}-NTA agarose affinity chromatography column. The bound PolB was eluted with 300 mM imidazole. The successfully eluted PolB was resolved on an SDS-polyacrylamide gel. Afu-PolB has a predicted molecular weight of 89 kDa (Figure 4.2). The Ni^{2+}-NTA eluate was diluted at least 10 times to reduce the salt concentration and then applied into a heparin chromatography column connected in series with 1 ml HiTrap DEAE FF column, a weak anion exchanger. The columns were equilibrated with 100 mM NaCl. Proteins were eluted with a 100–700 mM NaCl gradient in Tris pH 7.5. Collected proteins were resolved using SDS–Polyacrylamide gel. Proteins running at ~89 kDa were pooled and concentrated using a Vivaspin® centrifugal concentrator (Figure 4.2).

Afu-PolD was purified by my colleague Dr. Stanislaw Jozwiakowski.

4.3. Archaeal replicative primase can bypass 8-oxo-dG

Free-radical and non-radical oxidants which are common sources of endogenous DNA damage generate 8-oxo-dG lesions through altering individual bases. Oxidation of DNA under physiological conditions and environmental stress also leads to formation of 8-oxo-dG damage (Lindahl, 1993). Unlike UV irradiation damage, 8-oxo-dG does not majorly distort the DNA helix, however, it is able to stall replicative polymerases. In addition, the processivity of replicative polymerases can be reduced by 8-oxo-dG (McAuley-Hecht *et al*., 1994, Lipscombe *et al*., 1995). When replicative polymerases encounter a template containing 8-oxo-dG, they readily misincorporate dATP opposite this damage (Berquist and Wilson, 2012).

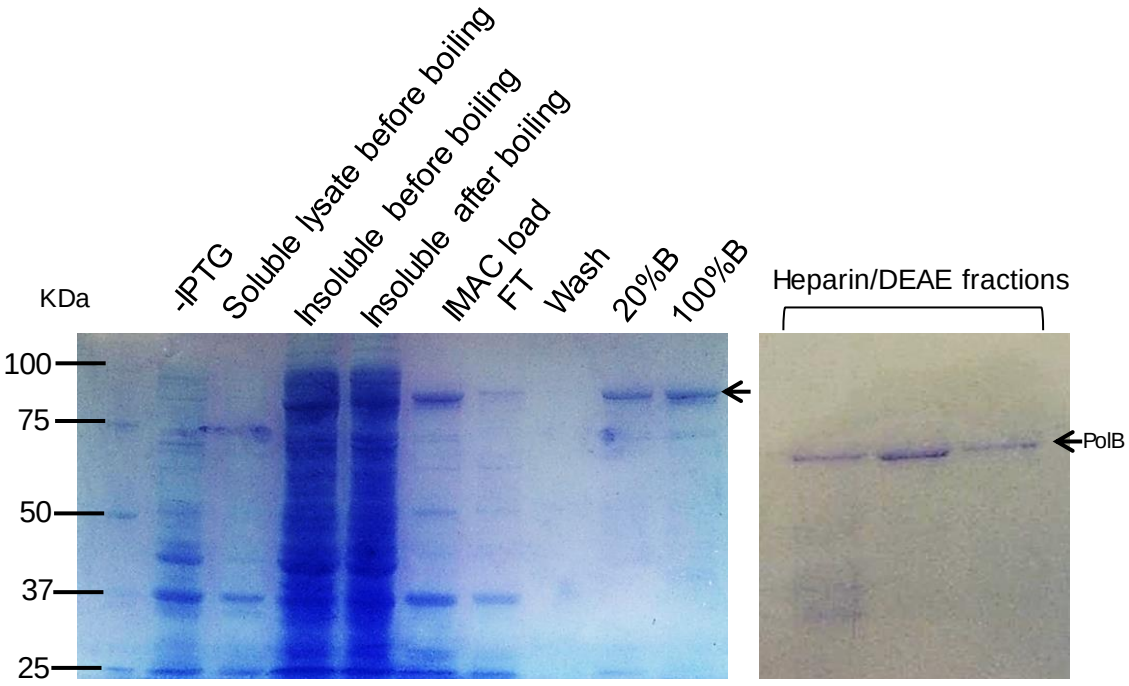


Figure 4.2. Purification of Afu-PolB

A culture of *E.coli* Rosetta transformed with pET28a:Afu-PolB was grown to exponential phase(OD_{600} =~ 0.6). After induction with 1 mM IPTG for 4 hours at 20°C, cells were lysed and subjected to heat-treatment for 20 minutes at 70°C. The denaturated proteins were removed by centrifugation and the soluble cell lysate was loaded to Ni^{2+}-NTA agarose column. Bound PolB was washed and eluted with 300 mM (100%B) imidazole and then subjected to heparin affinity chromatography connected in series with HiTrap DEAE FF column. The coulms were equilibrated with 100 mM NaCl and then developed with a 100–700 mM NaCl gradient in Tris pH 7.5. Fractions were analysed by SDS–Polyacrylamide gel and those containing a protein running at ~89 kDa were pooled and concentrated

We tested whether purified archaeal replicative polymerases (PolB and PolD) and replicative primase (PriS1/L) isolated from *Archaeoglobus fulgidus* (Afu) and *Pyrococcus furiosus* (Pfu), can read through a templated 8-oxo-dG. Standard primer extension assays were performed using a 50-mer template, containing a site-specific 8-oxo-dG at base 29, annealed to a 20-mer 5' flourescent labelled primer (Figure 4.3). A marked stalling before and after the 8-oxo-dG lesion was observed by Afu polymerases (PolB and PolD) and primase while bypassing the lesion. TLS activity was also observed for Pfu-PriS1/L (Figure 4.3).

It was next investigated whether bypass of a templated 8-oxo-dG by Afu-PriS1/L, Afu-PolB, Afu-PolD, and Pfu-PriS1/L was error-prone. Fidelity of 8-oxo-dG bypass was tested using single nucleotide incorporation assays. To generate the required substrate for the single nucleotide incorporation assays, the 3' end of the fluorescently labelled primer was annealed to the base 3' of the 8-oxo-dG containing template (Figure 4.4). Single nucleotide incorporation assays revealed that Afu-PolB incorporated dA opposite the lesion, while Afu-PolD and both primases (Afu-PiS1/L and Pfu-PriS1/L) incorporated dA and dC opposite 8-oxo-dG lesion with comparable efficiency (Figure 4.4). Our data suggested that bypass of a templated 8-oxo-dG by Afu-PolB is error-prone, but by Afu-PolD and both replicative primases can be both error-prone and error-free.

4.4. Error-free bypass of CPDs by Afu replicative primase

Despite an apparent lack of canonical TLS polymerase bypass mechanisms in the majority of archaea, hyerthermophilic archaea such as *Archaeoglobus fulgidus* and *Pyrococcus furiosus* can tolerate high doses of UV light and have mutagenic signatures indicative of the presence of TLS (Beblo *et al*., 2011; Watrin and Prieur, 1996). Cyclobutane pyrimidine dimers (CPDs) are one of the major types of UV-induced lesions (Sale *et al*., 2012; Sinha and Hader, 2002). This helix-distorting and replication blocking lesion is induced through cross-linking of adjacent pyrimidine bases in DNA by UV irradiation (Svoboda and Vos, 1995; Lopes *et al*., 2006). However, organisms can be threatened by similar blocking lesions even in the absence of UV light, e.g. produced by cross- linking with aldehyde (Voulgaridou *et al*. 2011).

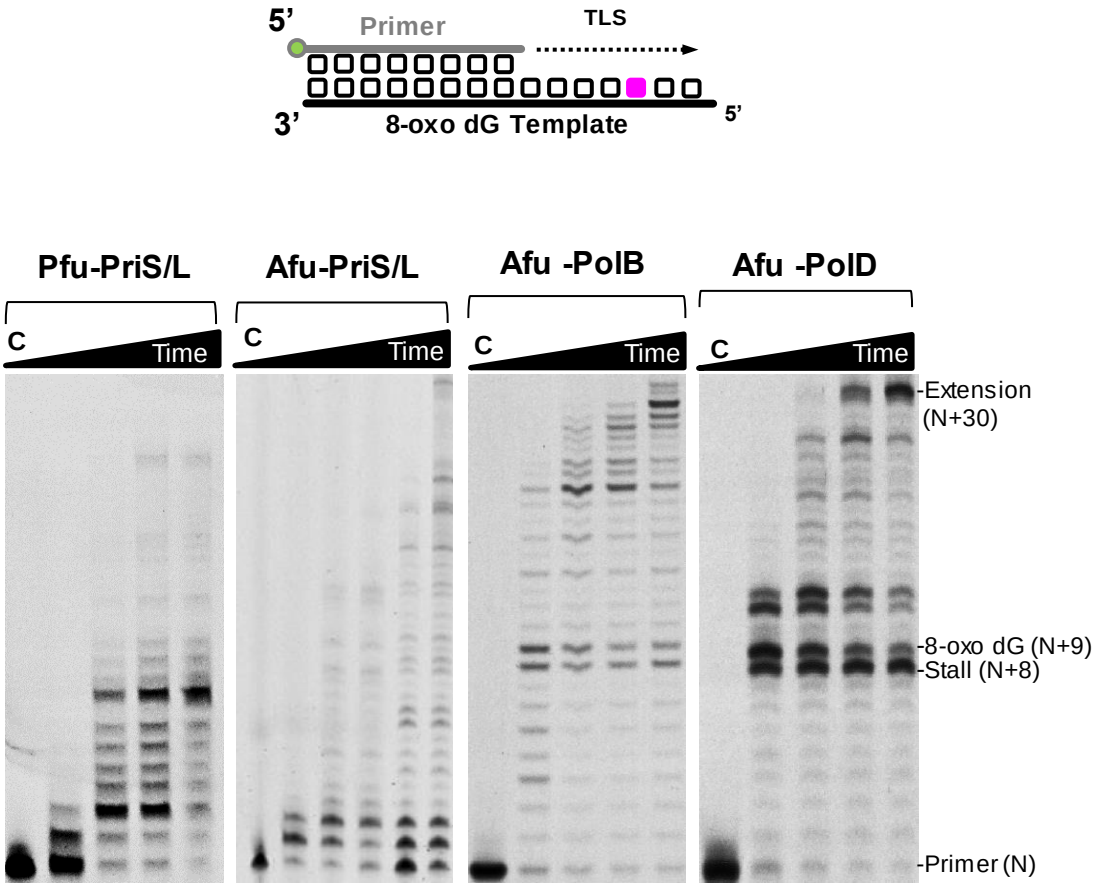


Figure 4.3. Archaeal replicative primase can bypass 8-oxo-dG.

The primer-template substrate used in the primer extension assay is shown schematically. A 50-mer template containing 8-oxo-dG was annealed to a 28-mer primer that was fluorescent labelled at the 5' terminus (Table 2.7). Primer-extension experiment was performed using 20nM of the substrate, 2 mM $MgSO_4$, 50 µM of each of the four dNTPs, and 100nM of the indicated primases and polymerases. The letter C denotes no enzyme control. All reactions were incubated at 50°C at 30 s, 1', 5', and 10'. The experiments showing the bypass of 8-oxo-dG by Afu-PolB and Afu-PolD were carried out by Dr. Stanislaw Jozwiakowski.

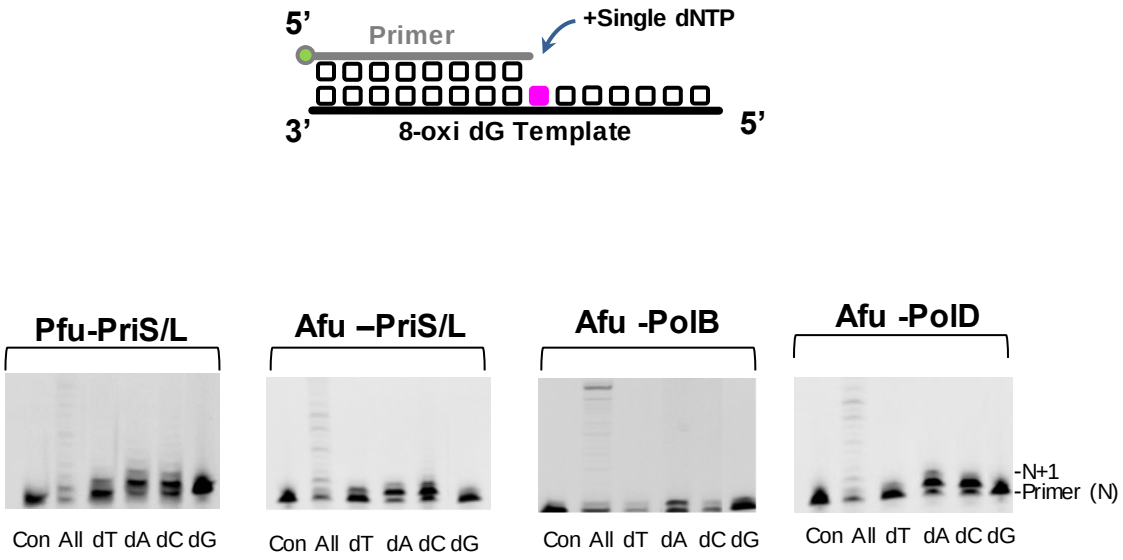


Figure 4.4. Single nucleotide incorporation opposite 8-oxo-dG.

The primer-template substrate used in the single nucleotide incorporation assay is shown schematically. A 50-mer template containing 8-oxo-dG was annealed to a 28-mer primer that was fluorescent labelled at the 5' terminus (Table 2.7). Single nucleotide incorporation experiment carried out using 20nM of the substrate, 2 mM $MgSO_4$, 50 µM of the particular dNTP under investigation, and 100nM of the indicated primases and polymerases. The letter C denotes no enzyme control. All reactions were incubated at 50°C for 10 minutes and then quenched and resolved by SDS-PAGE.

Since oxidative stress can speed up the formation of this type of DNA lesion, it is assuming that CPD-like lesions are abundant in hyperthemophilic archaea. First, we assayed Afu replicative polymerases for bypass of a CPD lesion by primer extension of a short, labelled primer annealed to a template containing T-T CPD at bases 28 and 29 (Table 2.7). Consistent with a CPD being a replication-blocking lesion, both replicative polymerases (PolB and PolD) were not able to fully extend the primer annealed to a CPD containing template (Figure 4.5). However, TLS activity across this UV-induced lesion was observed by Afu-PriS1/L. This indicated that Afu replicative primase is able to bypass a CPD lesion. In contrast to Afu, Pfu-PriS1/L cannot perform TLS opposite this lesion (Figure 4.5).

To determine whether replication of a CPD containing template by Afu replicative primase was error-prone, a single nucleotide incorporation assay was performed. Two dAs were incorporated by Afu-PriS1/L opposite both templating thymines of the CPD, revealing that the primase performs error-free bypass of this UV-induced lesion. Interestingly, error-free bypass of a CPD by Afu replicative primase is consistent with TLS activity of eukaryotic Pol η, which is able to bypass CPDs relatively accurately (Johnson *et al*., 1999). Incorporation of the correct dA base opposite the first thymine of the dimer was observed by Pfu-PriS1/L. (Figure 4.6).

4.5. Bypass of deoxyuracils by archaeal primase

Hydrolytic deamination of cytosine to uracil in DNA is one of the most abundant spontaneous changes, and can lead to mutations if left unrepaired. Since this mutagenic process is considerably enhanced at higher temperatures, it is believed that thermophiles are at significantly risk from hydrolytic deamination (Lindahl and Nyberg, 1974). Archaeal replicative PolB recognizes uracil in DNA templates and stalls replication when encountering these bases. Specifically, family B DNA polymerases can bind tightly to uracil in DNA template through a specialized binding pocket located in the N-terminal domain of the polymerase (Greagg *et al*., 1999). Remarkably, Pol B / PCNA binds tightly to dU and prevents the completion of DNA replication (Emptage *et al*., 2008). Therefore, it is believed

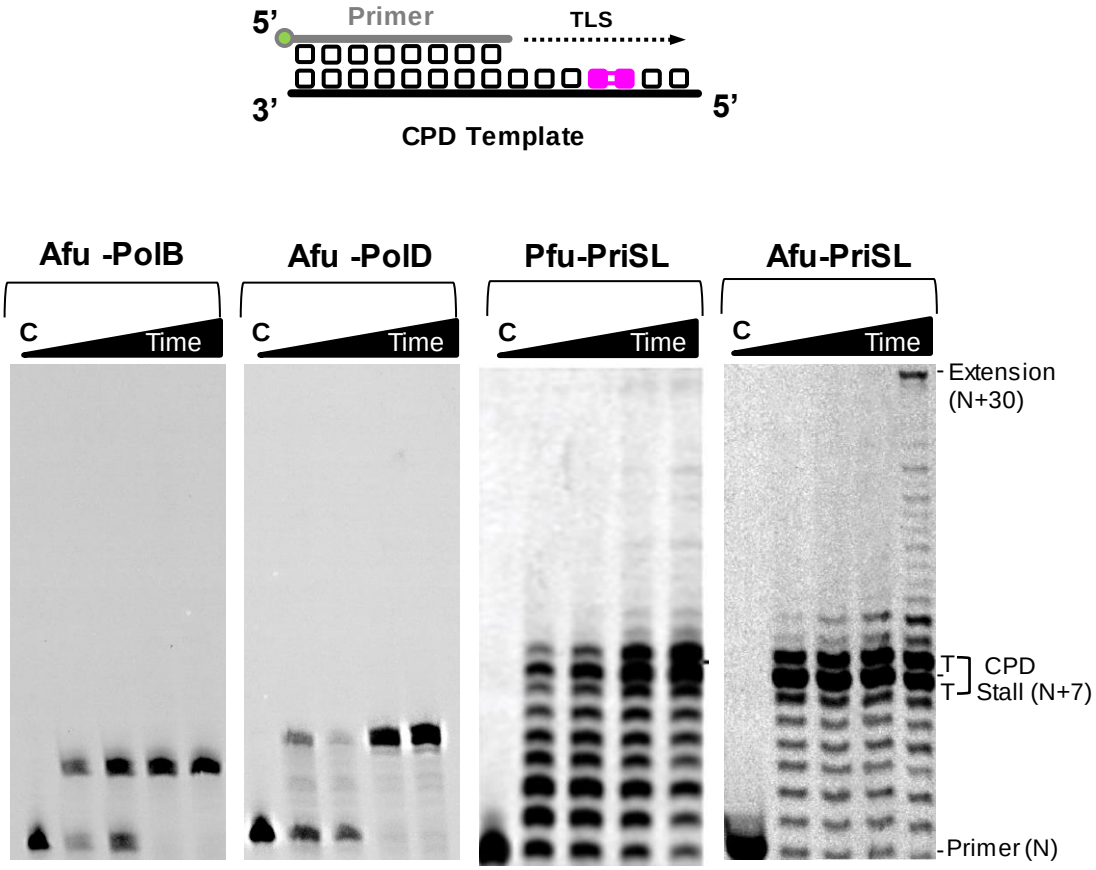


Figure 4.5. Afu-PriS1/L bypass CPD lesion

The primer-template substrate used in the primer extension assay is shown schematically. A 28-mer primer with a 5' fluorescent label was annealed to a 50-mer template containing a T-T CPD. Primer-extension experiment was performed using 20nM of the substrate, 2 mM $MgSO_4$, 50 µM of each of the four dNTPs, and 100nM of the indicated primases and polymerases. . The letter C denotes no enzyme control. All reactions were incubated at 50°C at 30 s, 1', 5', and 10'.Afu-PriS1/L is capable of inserting nucleotides opposite, and extending from, a CPD lesion *in vitro*. However, the Pfu-PriS1/L and both Afu replicative polymerses were unable to read through a CPD.

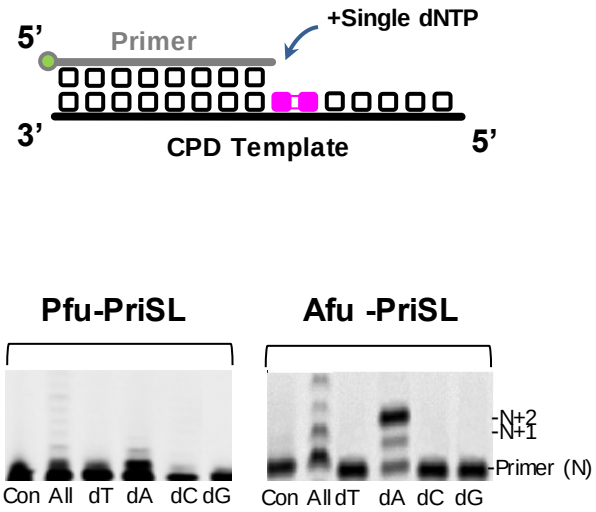


Figure 4.6. Single nucleotide incorporation opposite CPD

The primer-template substrate used in the single nucleotide incorporation assay is shown schematically. A short primer with a 5' fluorescent label was annealed to a 50-mer template containing a T-T CPD. Afu and Pfu replicative primases were analysed for their fidelity opposite a CPD. Afu-PriS1/L showed incorporation of dA nucleotides opposite both the 3'- and 5'-thymines of the CPD. Pfu-PriS1/L could incorporate a single dA opposite the first (3') base of the dimer.

that this mechanism is used to protect the integrity of dU-containing DNA during replication. In addition, polymerase activity of archaeal Pol D family polymerases is significantly reduced by the presence of uracil in the DNA template strand (Richardson *et al*., 2013). The ability of archaeal replicative primases to perform TLS on a dU-containing template was next tested. First, Afu replicative polymerases (PolB and PolD) were incubated with a template containing deoxyuracil to assess whether they could traverse du. Afu-PolB showed stalling four bases prior to the dU and the ability of Afu-PolD to extend the labelled primer annealed to the dU-containing template was significantly decreased (Figure 4.7).

Next, we investigated whether Afu-PriS1/L is able to replicate the DNA template containing deoxyuracil. Our data indicated that the PriS1/L from *A.fulgidus* significantly bypassed dU (Figure 4.7). A similar profile of bypassing dU was observed for Pfu-PriS1/L (Figure 4.7). These proteins were also assessed for their fidelity in incorporation opposite a dU lesion. Single nucleotide incorporation assays revealed that PolD and both replicative primases specifically incorporated a dA opposite dU (Figure 4.8) suggesting that the PriS1/L dependent bypass of dU is pro-mutagenic.

Although archaeal replicases bind tightly to a template-strand uracil, which leads to replication fork stalling, no restart mechanism has yet been investigated. In order to identify the possible role of replicative primases in restarting stalled DNA replication, TLS activities of Afu-PriS1/L on 30nt DNA templates containing one dU was tested. In this experiment, the substrate was pre-incubated with PolB/PCNA complex. The Afu-PriS1/L was able to fully extend the labelled primer annealed to a dU containing template, even in the presence of stalled PolB/PCNA complex (Figure 4.9). These data suggest that PriS1/L can read through dU-containing templates and also displace the uracil-bound polymerase thus restarting arrested replication forks.

Together, our results propose that archaeal replicative primases participate in both initiation and elongation of DNA synthesis by performing TLS bypass of DNA damages, therefore, preventing the archaeal replisome from arresting.

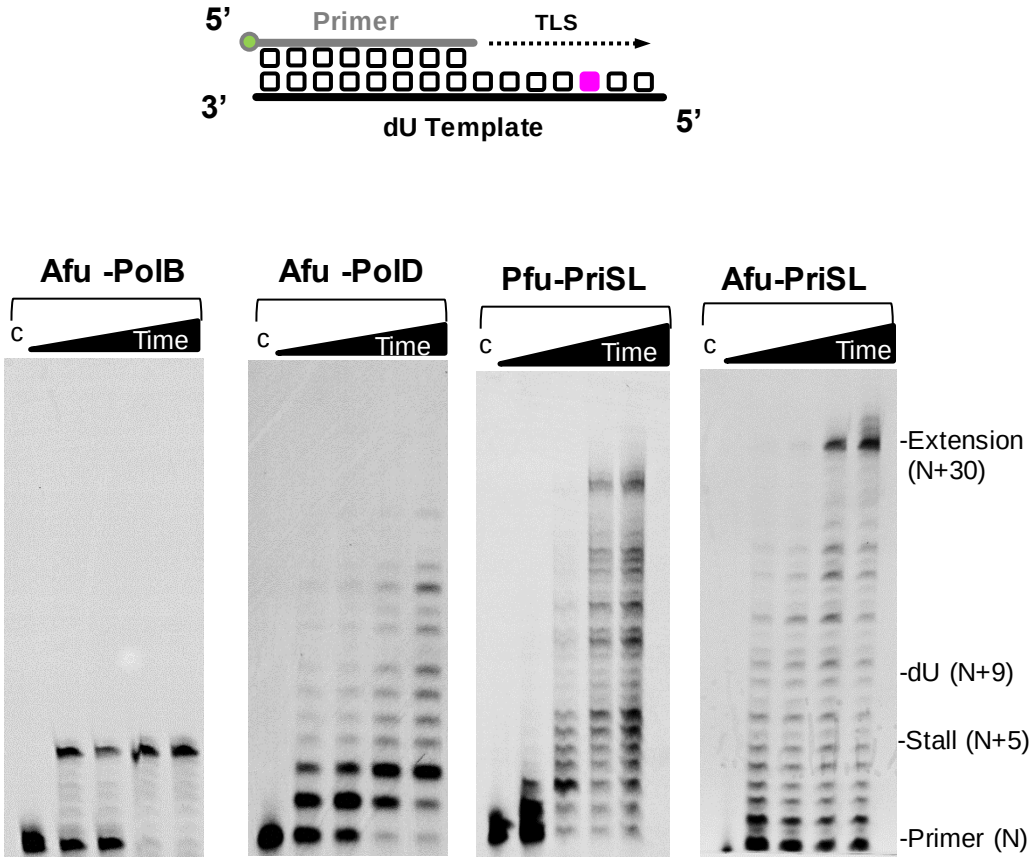


Figure 4.7. Archaeal replicative primase can bypass deoxyuracil

The primer-template substrate used in the primer extension assay is shown schematically. A 28-mer primer with a 5' fluorescent label was annealed to a 50-mer template containing dU. Archaeal replicative primase and polymerases were incubated with 20 nM of the substrate, 2 mM $MgSO_4$ (make sure you correct all such superscripts in your thesis) and, 50 µM of each of the four dNTPs at 50°C at 30 s, 1', 5', and 10' time points alongside a control containing no protein. Afu-PolB stalled four bases before dU and polymerase activity of PolD was significantly decreased. Both Afu and Pfu PriS1/L primases were able to read through a uracil base in the DNA without stalling.

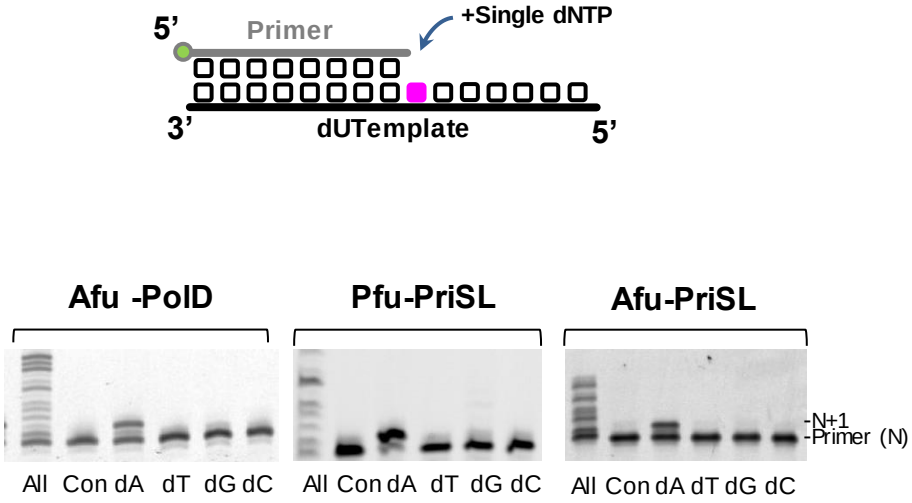


Figure 4.8. Single nucleotide incorporation opposite dU

The primer-template substrate used in the single nucleotide incorporation assay is shown schematically.A short primer with a 5' fluorescent label was annealed to a 50-mer template containing a dU. Single nucleotide incorporation experiment performed using 20nM of the substrate, 2 mM $MgSO_4$, 50 μM of the particular dNTP under investigation, and 100nM of the indicated primases and polymerases. The letter C denotes no enzyme control. All tested proteins showed incorporation of an adenine nucleotide opposite uracil.

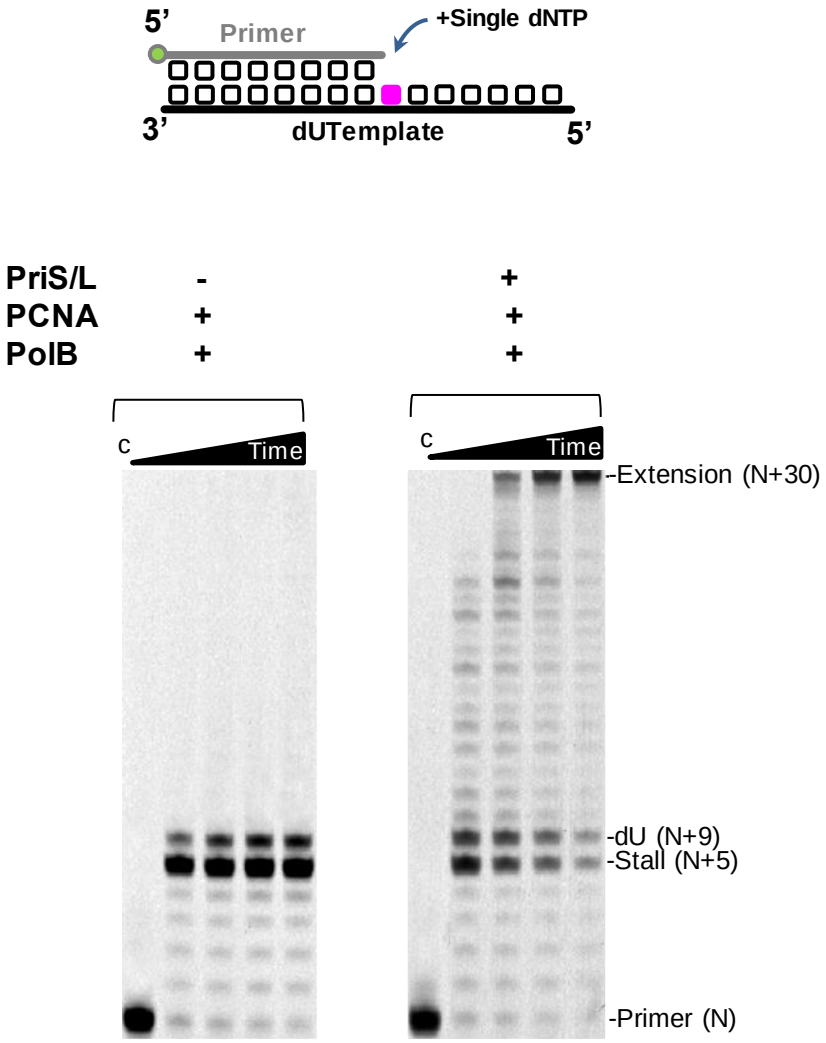


Figure 4.9. Stalled PolB/PCNA complex is rescued by Afu-PriS1/L TLS activity.

The primer-template substrate used in the primer extension assay is shown schematically. PolB/PCNA complex cannot read through a dU containing template and stalls four bases prior the dU. Afu-PriS1/L was incubated in reaction containing PolB/PCNA, dU containing template and. dNTPs at 50°C at 30 s, 1', 5', and 10' time points Afu-PriS1/L harbour the ability to read through dU and rescue stalled PolB/PCNA complex. The letter C denotes no enzyme control .The experiments shown here were carried out by Dr. Stanislaw Jozwiakowski.

4.6. Translesion DNA synthesis activity of *M.maripaludis* PriS1/L

Archaea exist in a diverse range of environmental conditions. Unlike hyperthermophiles, such as *A.fulgidus* and *P.furiosus* that thrive in extremely hot environments, mesophilic archaea like *M.maripaludis* (Mma) live in mild temperatures. Therefore, archaea have protein and cellular pathway adaptations to drastically varying biosystems. To investigate whether mesophilic replicative primases can also perform TLS, previously described primer extension assays were employed to examine TLS activity of purified Mma-PriS1/L using oligonucleotide templates containing site-specific replication-blocking DNA lesions (Figure 4.10). Our data indicate that Mma-PriS1/L has the capacity to replicate DNA templates containing deoxyuracils, however it is completely incapable of reading through 8-oxo-dG and CPD lesions. Upon testing the fidelity of Mma-PriS1/L when incorporating opposite dU, we found that, similar to Afu and Pfu primases, Mma-PriS1/L incorporates adenine opposite dU, which also suggests a pro-mutagenic bypass of dU by this PriS1/L complex (Figure 4.10). Together, our data suggest that compared to Afu and Pfu replicative primases and other TLS polymerases, Mma-PriS1/L is not efficient in replicating through 8-oxo-dG and CPD lesions, which could be due to protein adaptations to mild temperatures. Notably, organisms living in lower temperatures possess reduced enzyme activity, genetic expression and protein function (Reed *et al*., 2013; Yang *et al*., 2007). It is assumed that due to the moderate environmental conditions under which mesophilic archaea reside, they are threatened by less levels of DNA damage or they may use other damage tolerance mechanisms such as recombination. Therefore, the replicative primase in this archaea might not play an important role in translesion DNA synthesis. However, further investigation is required.

4.7. Detecting PriS/L's repriming activity

Since the PriS/L complex is known as a proficient DNA primase in archaea, it is highly likely that its primase activity as well as its TLS activity, is needed to tolerate a variety of DNA damages. Priming allows the enzyme to initiate DNA synthesis *de novo* downstream of a lesion following replication stalling at the

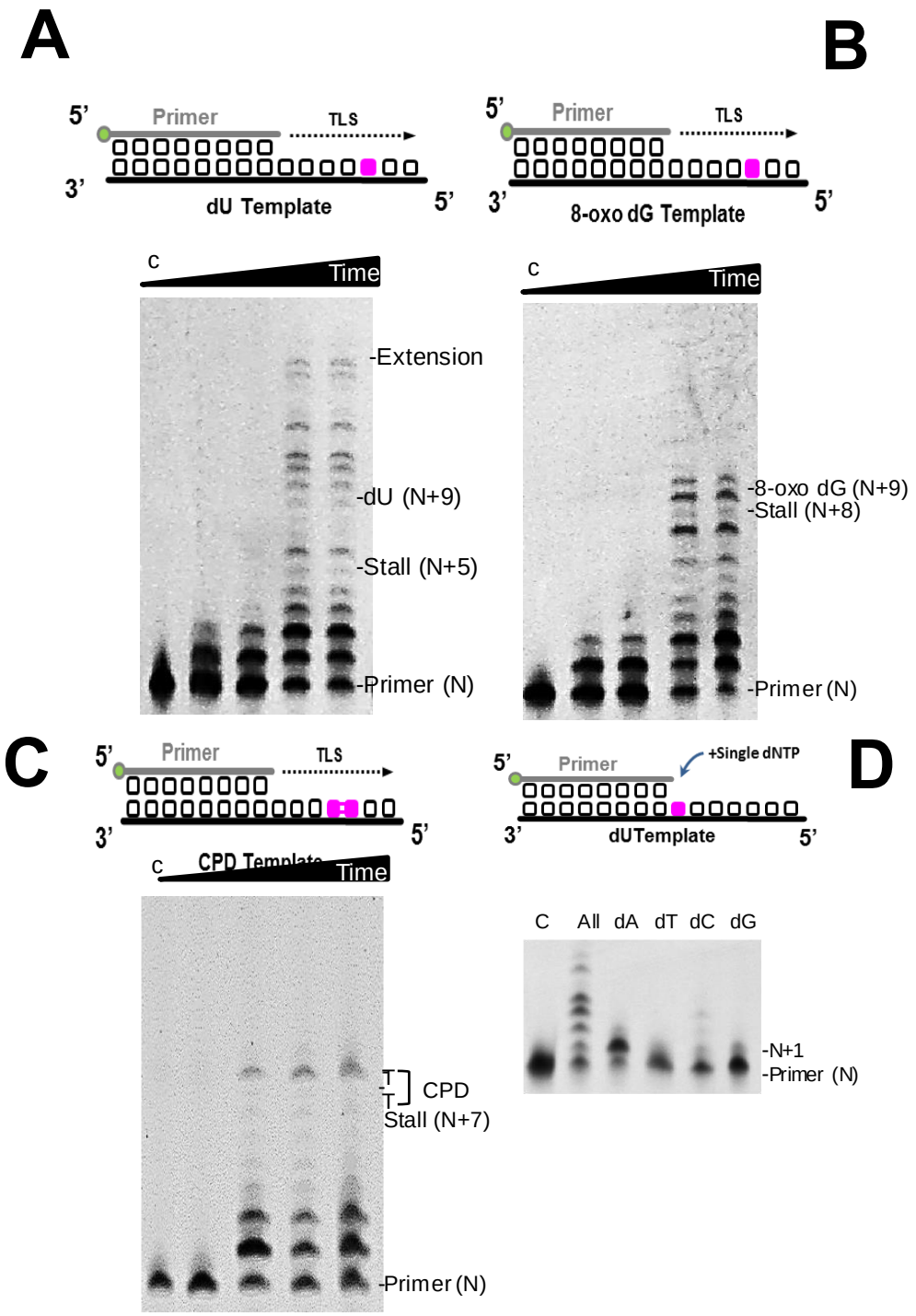


Figure 4.10. Mma-PriS1/L is not an efficient TLS polymerase

Schematic pictures of primer-template substrates used in primer-extension assay (A, B, C) and single nucleotide incorporation assay (D) are illustrated above each fluorescent gel. (A) Mma-PriS1/L was incubated with dU containing template. Mma-PriS1/L could bypass dU but with decreased processivity in comparision with Afu and Pfu primases. (B) Mma-PriS/L was not able to read through 8-oxo-dG lesion. (C) Mma-PriS1/L was incapable of extending the CPD lesion. (D) Mma-PriS1/L incorporated a dA opposite dU base.

site of damage, this leaves a ssDNA gap in the daughter strand (Bianchi *et al*., 2013 Keen *et al* 2014; Schiavone *et al*., 2015). It is believed that repriming is limited to lesions on the leading strand replication. Due to the discontinuous nature of lagging-strand replication, DNA blocking lesions or obstacles are likely to be inherently tolerated (Pagès and Fuchs, 2003; Svoboda and Vos, 1995). In the lagging strand, primers are constantly produced for Okazaki fragment synthesis. The newly generated primer is able to restart replication downstream of a lesion and generate a ssDNA gap. Repriming replication post-lesion has been reported in different organisms (Rupp and Howard -flanders, 1968). In mammalian cells, repriming downstream of UV lesions has been proposed (Lehman, 1972). Furthermore, repriming downstream of UV-irradiated stalled replication forks, which leave behind ssDNA gaps in the leading strand, has been detected in yeast (Lopes *et al*., 2006). Bacterial DnaG primase can also catalyse leading-strand repriming in *E. coli* (Heller and Marians, 2006; Yeeles and Marians, 2011). Remarkably, PrimPol, a novel eukaryotic DNA primase-polymerase can perform repriming downstream of different blocking lesions and obstacles and it was reported that PrimPol can facilitate close-couple repriming downstream of a G quadruplex (G4) structures (Schiavone *et al*., 2016). Moreover, in avian cells, the exhibited repriming activity by PrimPol could restart stalled replication forks following UV damage (Keen *et al*., 2014). Recently, PrimPol exhibited downstream repriming of apurinic/apyrimidinic site (Ap site) and thymine glycol (Tg) lesions in the template strand. (Kobayashi *et al*., 2016).

In order to assess the capacity of archaeal PriS/L to reprime downstream of lesions, Afu-PriS1/L was incubated with a template containing either apurinic/apyrimidinic site (Ap site) or thymine glycol (Tg) lesion. In common with PrimPol, archaeal replicative primases are unable to bypass Ap and Tg lesions through TLS, in the presence of Mg^{2+} (Jozwiakowski *et al*., 2015). However, this inability does not preclude the possibility that PriS/L rescue the stalled replication through repriming activity. To test this, we used a primer containing a terminal 3' dideoxynucleotide moiety, which was annealed upstream of the templating lesions that represent the replication stalling site. This stops template-independent primer extension that impedes the assessment of PiS/L's repriming activity. As a control, PrimPol was also incubated with both templating lesions. As indicated, on the non-damaged template with no annealed primer, Afu-PriS1/L

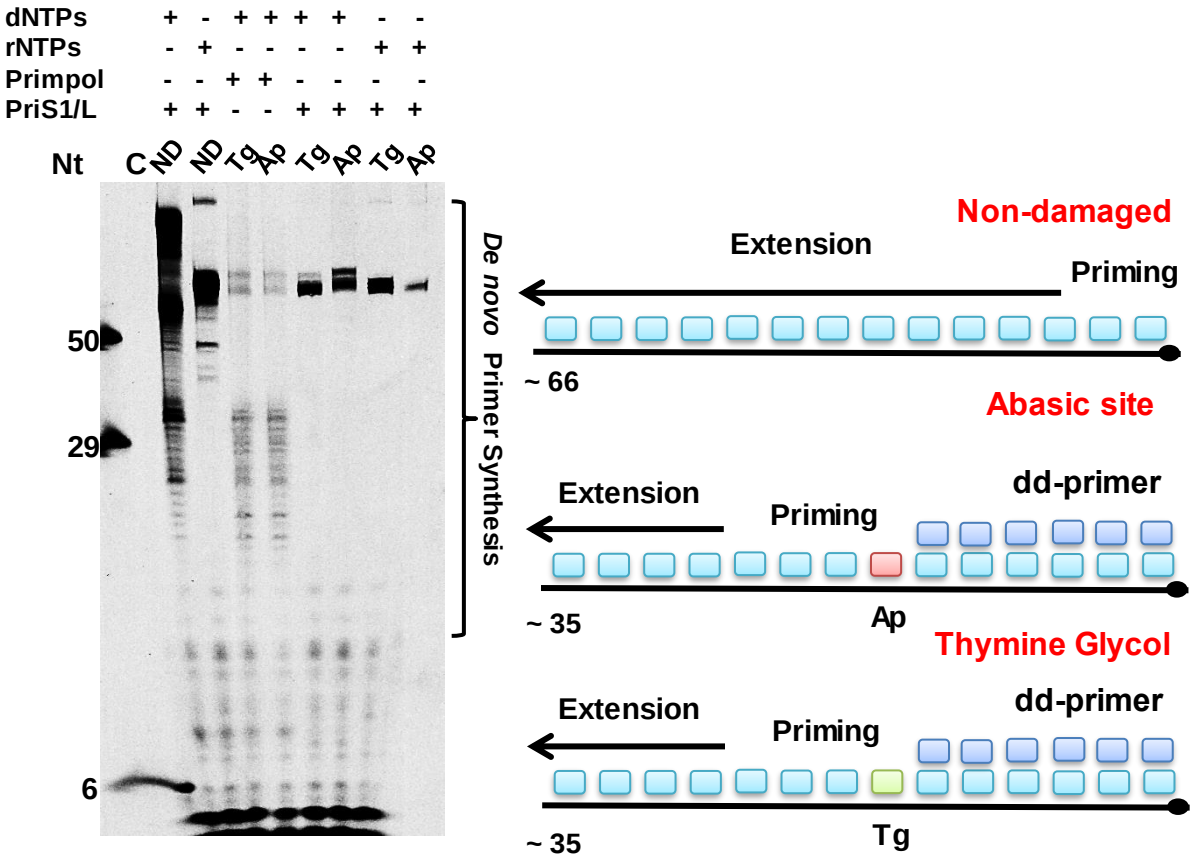


Figure 4.11. Archaeal PriS/L cannot catalyse repriming downstream of Ap and Tg

Reactions containing 1μM of Afu-PriS1/L, 250μM of dNTPs or rNTPs, 2.5μM of FAM dNTPs and 1μM of primer-templates (as indicated in schematic) were incubated for 15 minutes at 50°C and 37°C for Afu-PriS1/L and PrimPol, respectivelyIn order to test the repriming ability of PriS1/L rather than its TLS activity, primers with a 3′ dideoxynucleotide were annealed upstream of the lesion on templates containing Ap and Tg. PrimPol showed close-coupling repriming downstream of lesions as the length of primer extension products produced in each lesion was near identical. Afu-PriS1/L was unable to reprime downstream of templates containing Ap and Tg. Oligonucleotide nucleotide (Nt) length markers are indicated on the left. The letter C denotes no enzyme control and “ND” denotes non-damaged template without an annealed primer.

showed significant priming activity (Figure 4.11). Unlike PrimPol, which was able to perform close-couple repriming downstream of an Ap site and Tg blocking lesions, no repriming activity downstream of these lesions was observed by Afu-PriS1/L, either in the presence of rNTPs or dNTPs. We observed efficient priming by AfuPriS1/L complex with M13 ssDNA using a radioactivity assay in chapter 3. Hence, considering that PriS1/L is a genuine primase in archaea and maintaining replication following potential lethal DNA lesions can be facilitated by repriming activity of DNA primases in different organisms, the inability of Afu-PriS1/L in repriming downstream of Ap and Tg lesions cannot be conclusive. Unlike higher eukaryotes, in other organisms such as budding yeast, which lack PrimPol, the leading strand repriming is dependent on the replicative DNA primase (Prim1) of the Pol α complex (Iyer *et al*.,2005). This is in common with *E. coli* DnaG primase that is required for leading strand repriming. This suggests that in archaea, instead of PriS/L, the non-essential bacterial-like DnaG primase might be required for repriming activity. Although this is just speculation and further research is necessary.

4.8. Nucleotide insertion fidelity of the PriS1/L complex

TLS is a two step process, first a polymerase (mis)incorporates a nucleotide oppsite a DNA lesion and then subsequently extends past this lesion (Sale *et al*., 2012). Similar to PrimPol and other TLS polymerases, archaeal replicative primase (PriS/L) lacks a 3'-5' exonuclease activity (Liu *et al*., 2001; Guilliam *et al*., 2015; Keen *et al*., 2014). This suggests that archaeal replicative primases have low fidelity and such low fidelity is also a hallmark of canonical TLS polymerases. In order to measure the fidelity of PriS1/L, we utilized primer extension assays based on single incorporation of correct or incorrect nucleotides. To analyse the single base incorporation, a 27-mer primer was annealed with a 50-mer template with either adenine (dA), cytosine (dC), guanine (dG) or thymine (dT) interrogated as the primary templating base (N+1 position) and also at the N+2 position. In three substrates, one templating cytosine (dC) was located following the first and second templating bases (N+3 position). However, in one substrate where C was at N+1 and N+2 positions, dT was the following templating bases (N+3 position). In this experiment, PriS1/L from three archaeal species (*A.fulgidus, P.furiosus and M.maripalidus*) was incubated with

each primer-template substrate and either dATP, dCTP, dGTP or dTTP as the incoming bases.

4.8.1. Afu-PriS1/L shows a tendency for misincorporation

Single-base misincorporation assays using Afu-PriS1/L suggested that these enzymes have a strong tendency to misincorporate some bases opposite templating bases (Figure 4.12). Our data showed that Afu-Pris1/L misicorporated dGTP as the incoming base opposite templating dG and dT (Figure 4.12). This result was consistent with the reported ability of PrimPol to misincorporate dGTP opposite templating dG (Guilliam *et al*., 2015). In addition to dGTP, PriS1/L also displayed misincorporation of dCTP and dATP opposite templating A and C, respectively (Figure 4.12). Single misincorporation of dATP opposite templating C was also observed by PrimPol (Guilliam *et al*., 2015). When dATP was used as the correct incoming base opposite two templating dTs incorporation of dATP opposite the third cytosine nucleotide at the N+3 position was visible (Figure 4.12). In addition, incorporation of dG opposite two templating cytosines was followed by incorporation of a dG opposite the third nucleotide base, thymine, at the N+3 position (Figure 4.12). Therefore, these results suggested that PriS1/L could misincorporate dA and dG opposite templating dC and dT, respectively. Another possibility would be that PriS1/L scrunches the template strand and one of the thymidines and cytosines is read twice as a templating base. Together, these observations propose that Afu-PriS1/L has a propensity for nucleotide misincorporation.

4.8.2. Nucleotide misincorporation by Pfu-PiS1/L

Similar single incorporation assays were employed to analyse the base substitution fidelity of Pfu-PriS1/L. Similar to Afu-PriS1/L, Pfu-PriS1/L exhibited a tendency to misincorporate dG, especially opposite templating dT (Figure 4.13). In addition, when PriS1/L incorporated dGTP opposite two templating cytosine followed by two templating thymines (N+3 and N+4 positions), significant product bands were observed at N+3 and N+4 positions (Figure 4.13). Again, these correspond to the propensity of PriS1/L to misincorporate dGTP opposite a templating dT. Single misincorporations, albeit not very significant, were visible opposite templating dC when dATP was used as the incoming base. Additionally, there was some evidence for the incorporation of dA opposite templating

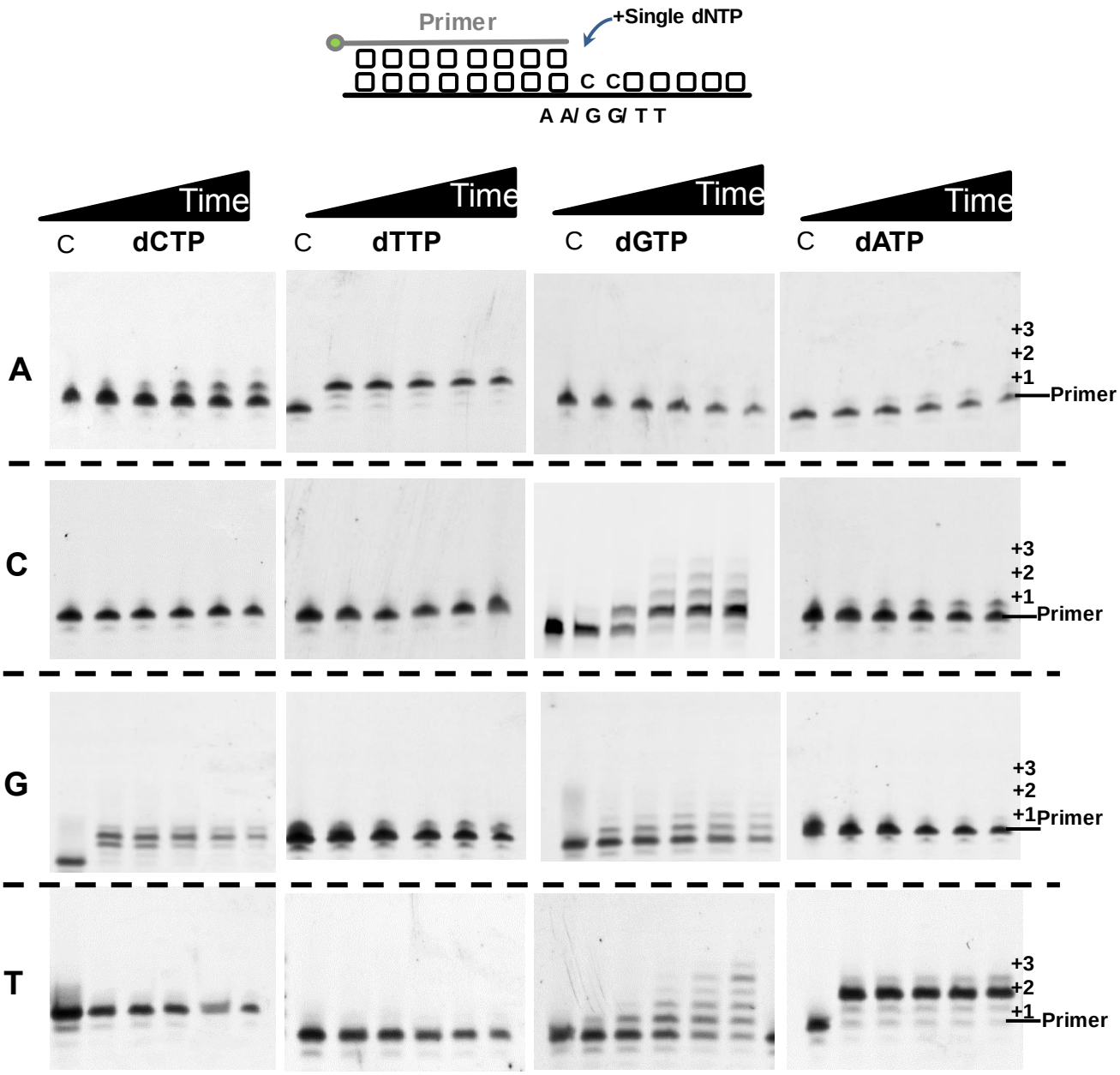


Figure 4.12. Single nucleotide incorporation fidelity of Afu-PriS1/L

Fidelity of Afu-PriS1/L was measured using single-base incorporation assay. The primer-template substrate used in the single incorporation assay is shown schematically. PriS1/L was incubated with 20 nM primer-template substrate (Table 2.7.) containing either AA, CC,GG or TT as two templating bases and supplemented with 100µM of either dCTP, dTTP, dGTP or dATP at 50°C for 30 seconds, 1, 5, 10, 20 minutes. The templating bases are shown on the left and the incoming dNTPs are indicated above. The letter C denotes no enzyme control.

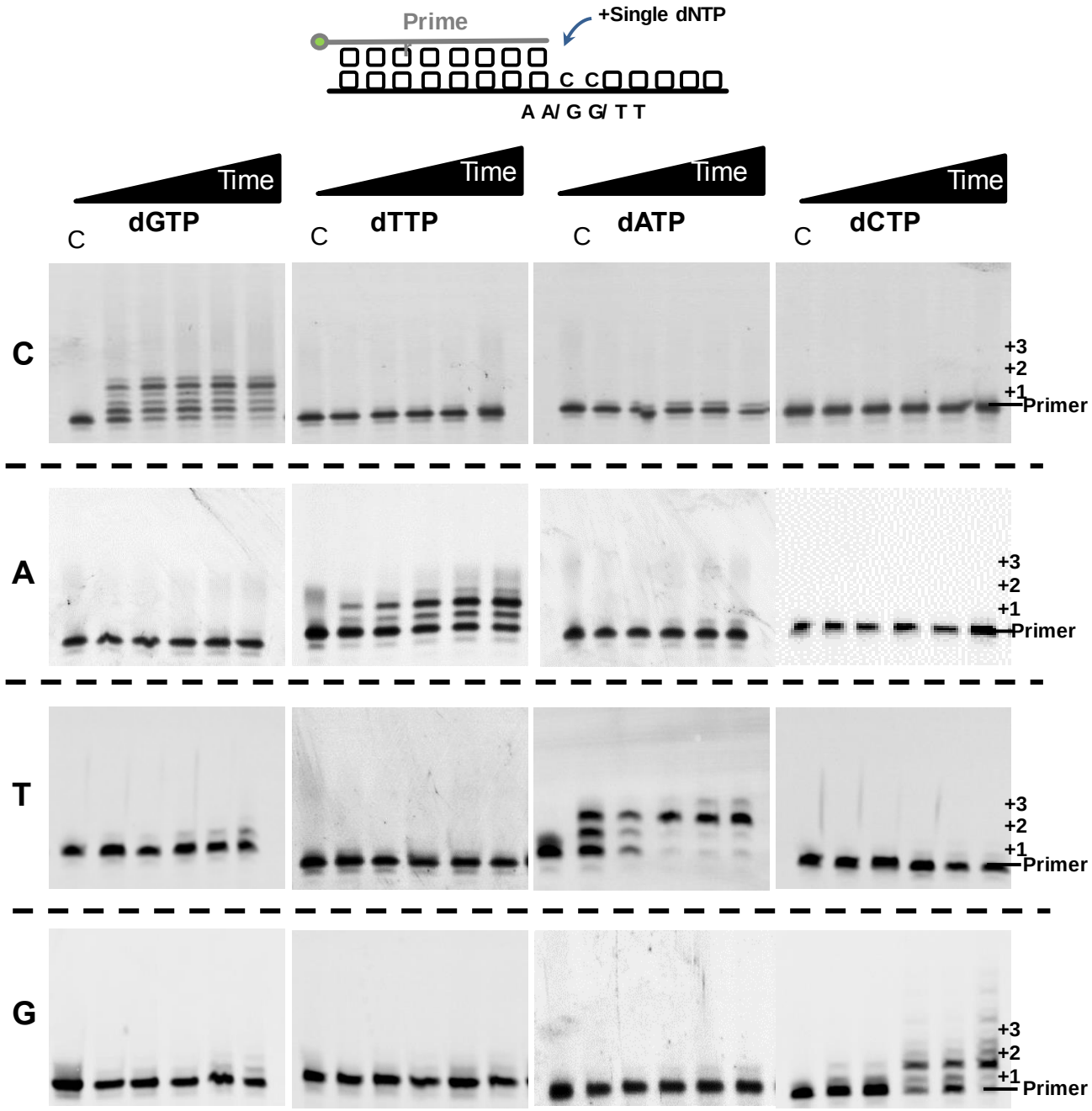


Figure 4.13. Analysis of Pfu-PriS1/L fidelity

Fidelity of Pfu-PriS1/L was measured in the same manner as Afu-PriS1/L. PriS1/L was incubated with individual dNTPs and substrates from table 2.7 at 50°C for 30 seconds, 1, 5, 10, 20 minutes. Watson-Crickbase pairing opposite each templating base is observed. Although, the enzyme shows erroneously incorporation of dCTP opposite templting C (N+3 and N+4 positins) when incorporating cytocine opposite templating G. In addition, when PriS1/L incorporating dGTP opposite two templating cytocine followed by two templating thymines (N+3 and N+4 positions), significant product bands were observed at N+3 and N+4 positions. The templating bases are shown on the left and the incoming dNTPs are indicated above. The letter C denotes no enzyme control.

dT at the N+3 position, which also suggests that Pfu-PriS1/L can misincorporate dA opposite templating dC. This misincorporation was also observed in the case of Afu-PriS1/L. A similar result was detected on a dG template with the correct incoming base (dCTP) and misincorporation of dC opposite templating dC was observed at the N+3 and N+4 positions (Figure 4.13). This increased N+3 and N+4 incorporation could be due to the over incubation of PriS1/L, rather than through misincorporation, as it was visible after 20 minutes incubation. These results suggest that, similar to Afu-PriS1/L, Pfu-PriS1/L has a tendency for nucleotide misincorporation.

4.8.3. Nucleotide misincorporation by Mma-PiS1/L

As previously indicated (section 4.6.), in comparison to PriS1/L complexes isolated from *A.fulgidus* and *P.furiosus* that showed bypass of different lesions through TLS, PiS1/L isolated from *M.maripalidus* was only able to read through dU containing templates. Therefore, this suggested that Mma-PriS1/L might not be an efficient TLS DNA polymerase. To measure the fidelity of Mma-PriS1/L, single-base incorporation assay was used. Although Mma-PriS1/L only showed incorporation of correct nucleotides opposite most templating bases, a small amount of aberrant incorporation of nucleotides was also observed (Figure 4.14). Our data indicated that, consistently, product bands at N+3 and N+4 positions were detectable on the templating C when dGTP was the incoming base which suggested misincorporation of dGTP opposite templating T (Figure 4.14). Since all three examined PriS1/L complexes carried out this type of misincorporation, it can be concluded that archaeal PriS/L has a strong tendency to misincorporate dGTP, mostly opposite templating T. This could be the error signature of PriS1/L. Furthermore, very weak misincorporation of dCTP and dTTP opposite templating T by Mm-PriS1/L was observed (Figure 4.14).

4.9. PriS/L has the capacity for mismatch extension

Alteration of the coordination of incoming bases for strand extension can also change the coordination of primer DNA. DNA polymerases differentiate against mismatches at the binding and insertion steps (Johnson and Beese, 2004), so the ability of archaeal PriS1/L to tolerate mismatches at primer-template

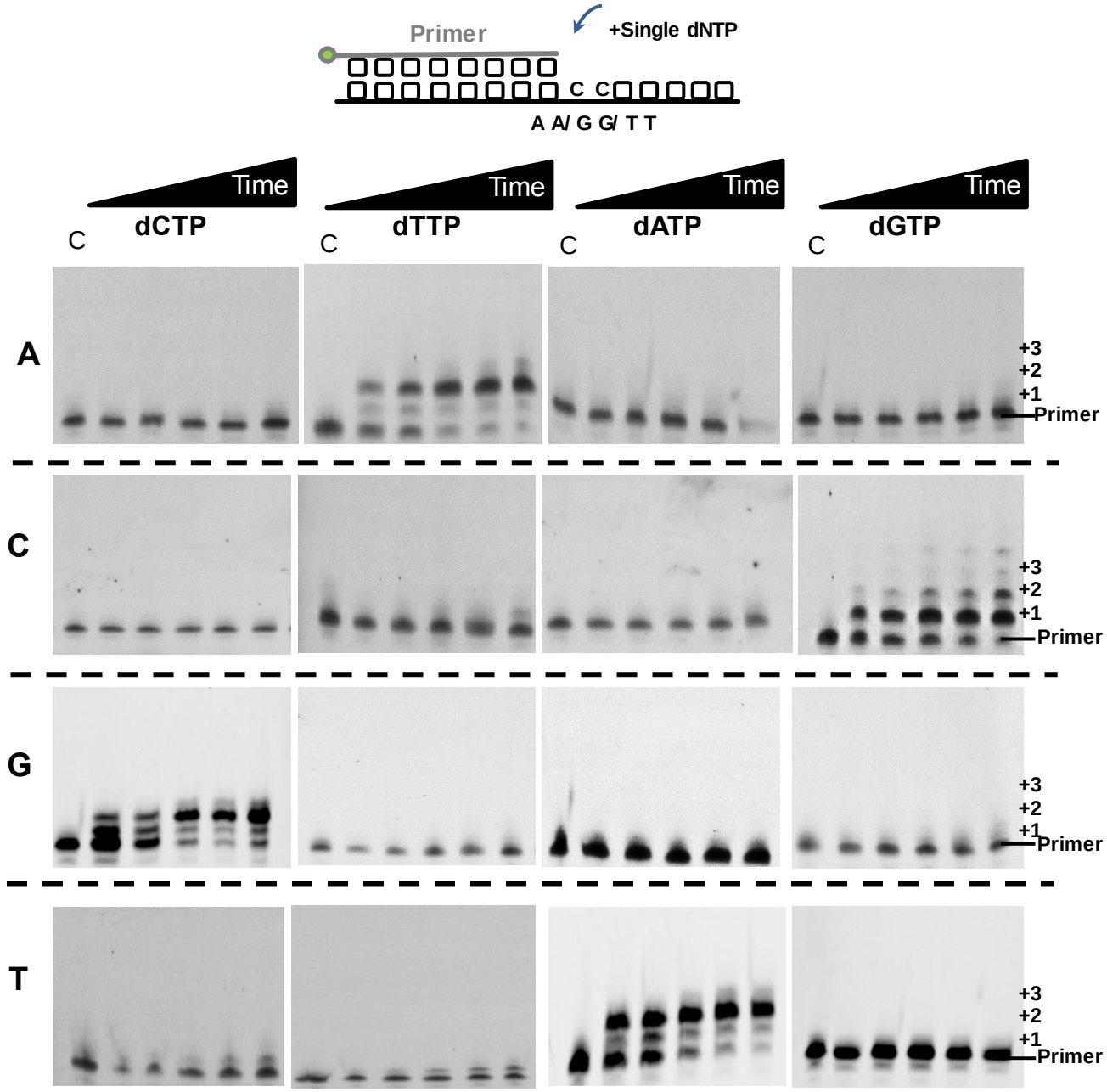


Figure 4.14. Single nucleotide incorporation fidelity of Mma-PriS1/L

Fidelity of Mma-PriS1/L was also examined in the same manner as Afu and Pfu PriS1/L complexes. The primer-template substrate used in the single incorporation assay is shown schematically. Mma-PriS1/L shows less misincorporation relative to two other replicative primases. The templating bases are shown on the left and the incoming dNTPs are indicated above. The letter C denotes no enzyme control.

junctions suggests reduced catalytic efficiency for this enzyme. Therefore, next, we tried to analyse the mismatch extension abilities of PriS1/L on different templates with a mismatch at the primer-template junction. Following incubation of Afu-PriS1/L with each primer-template substrate, a proficient extension from a G-G mismatch was detected. Moreover, this enzyme exhibited a capacity to extend A-C, A-G, C-C and T-G mismatches (Figure 4.15). However, incubation of Pfu-PriS1/L with these Primer-template substrates showed only very weak extension from G-T and T-C mismatches (Figure 4.16). In contrast, Mma-PriS1/L was not able to extend from any of the mismatched bases (Figure 4.17). The inability of Mma-PriS1/L to extend terminal mismatched base pairs is consistent with the lack of ability of this replicative primase to extend beyond some lesions.

The results obtained from mismatch extension experiments are summarized in Table 4.1. In some cases, PriS1/L was able to form a Watson-Crick base pair with the next templating base (N+1 position) following accommodation of the mismatch in the active site. Therefore, it could extend the mismatch in a template-dependent fashion. However, there is one case in which dGTP is incorporated where the next templating base is not cytosine but rather the cytosine is located at N+2 position (Table 2.7). This can happen by scrunching the template by Afu-PriS1/L and incorporating dGTP opposite the downstream C, which would in turn leads to deletions *in vivo*.

4.10. Summary and discussion

This chapter reports that archaeal replicative primases are able to perform translesion DNA synthesis in order to bypass replication-blocking lesions and rescue stalled replication forks *in vitro*. Our data indicated that PriS1/L complexes isolated from hyperthermophilic archaea (*A.fulgidus* and *P.furiosus*) are capable of traversing 8-oxo-dG containing templates. In addition, unlike archaeal replicative polymerases, both Afu-PriS1/L and Pfu-PriS1/L show the ability to replicate past dU bases. Furthermore, both primases exhibit correct incorporation of dC, through Watson-Crick base pairing and incorrect incorporation of dA by forming Hoogsteen base pairing opposite an 8-oxo-dG lesion. This propensity of incorporating C, as well as A, opposite 8-oxo-dG is evident in some other polymerases. However, presence of some proteins such as PCNA and RPA

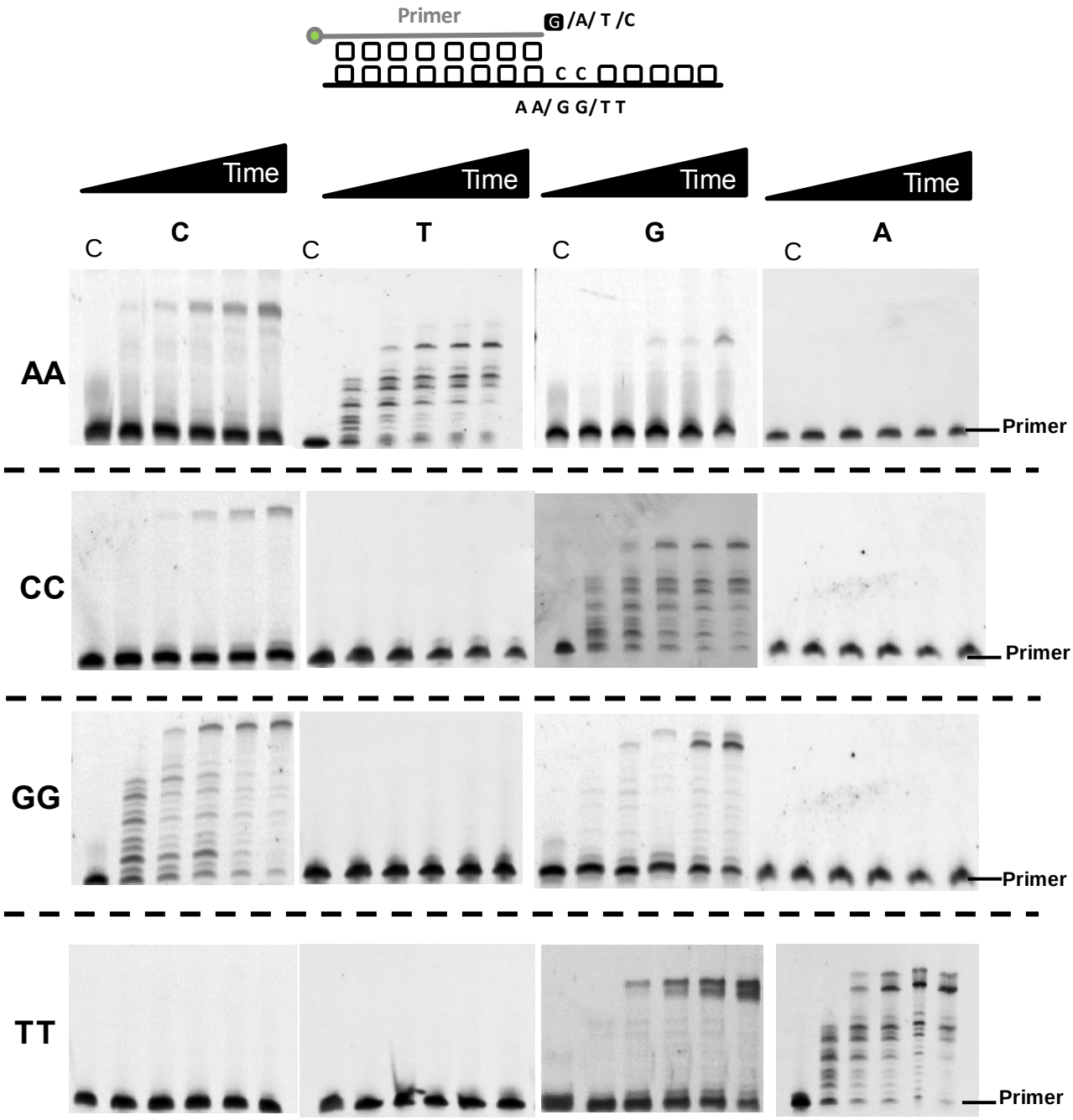


Figure 4.15. Afu-PriS1/L performs mismatch extension

Mismatch extension ability of Afu-PriS1/L was measured in the presence of all four dNTPs and primer-template substrates from table 2.7. Each primer-template substrate consists of a mismatch base at the 3' terminal of the primer. Four variants of primer-template substrates show Watson-Crick base pairs, while, 12 substrates represent other possible combinations of base pairs. The templating bases are shown on the left and each mismatched base at the 3' end of the annealed primer is illustrated above.

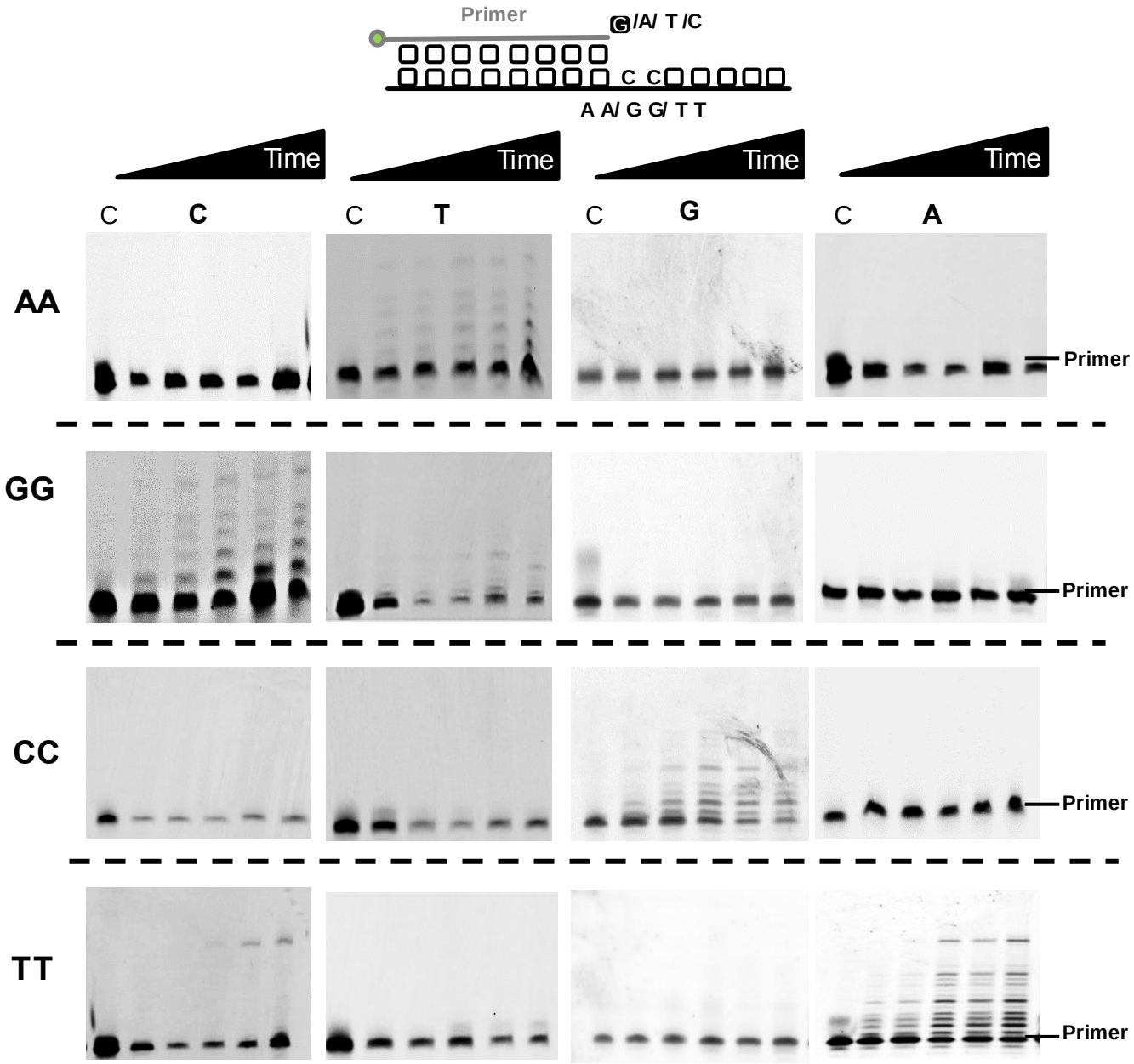


Figure 4.16. Analyses of Pfu-PriS1/L base mismatch tolerance

Mismatch extension by Pfu-PriS1/L was studied in the same manner as Afu-PriS1/L. Incubation of Pfu-PriS1/L with 16 different primer-template substrates indicated tolerance for G-T and T-C mismatches by this enzyme. The templating bases are shown on the left and each mismatched base at the 3' end of the annealed primer is illustrated above.

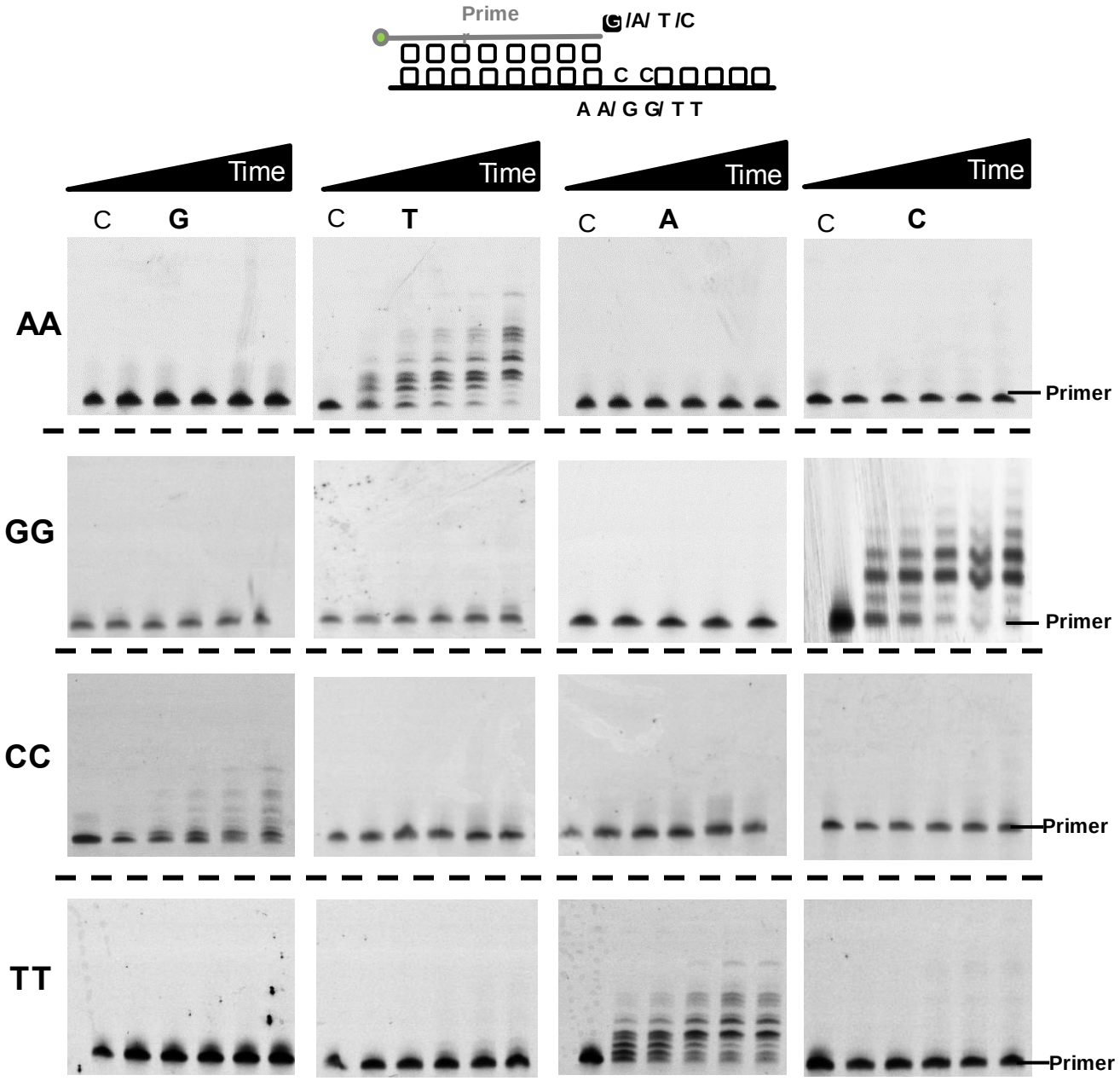


Figure 4.17. Analyses of Mma-PriS1/L base mismatch tolerance

Mismatch extension by Mma-PriS1/L was analysed in the same manner as Afu and Pfu PriS1/L proteins. In contrast to Afu and Pfu, Mma-PriS1/L was not able to extend from any of none Watson-Crick base pairs.

Sequence	Pairing conformation	Afu-PriSL		Pfu-PriSL		Mmp-PriSL	
		Tolerance	Nucleotide incorporation	Tolerance	Nucleotide incorporation	Tolerance	Nucleotide incorporation
3'---AACC---5' 5'---C	Disordered	+	C	-		-	
3'---AACC---5' 5'---G	Frayed	+	G	-		-	
3'---AACC---5' 5'---T	Watson-Crick	+++	T	+++	T	+++	T
3'---CCCC---5' 5'---A	Disordered	-		-		-	
3'---CCCC---5' 5'---C	Frayed	+	C	-		-	
3'---CCCC---5' 5'---G	Watson-Crick	+++	G	+++	G	+++	G
3'---CCCC---5' 5'---T	Disordered	-		-		-	
3'---GGCC---5' 5'---A	anti-anti	-		-		-	
3'---GGCC---5' 5'---C	Watson-Crick	+++	C	+++	C	+++	C
3'---GGCC---5' 5'---G	syn-anti	+	G	-		-	
3'---GGCC---5' 5'---T	Wobble	-		+	T	-	
3'---TTCC---5' 5'---A	Watson-Crick	+++	A	+++	A	+++	A
3'---TTCC---5' 5'---C	open	-		+	C	-	
3'---TTCC---5' 5'---G	wobble	+	G	-		-	
3'---TTCC---5' 5'---T	wobble	-		-		-	

Table 4.1. Summary of mismatch tolerance of archaeal replicative primase.

Previously, confirmation of 12 possible combinations of DNA mispairs was structurally characterised (Johnson and Bees, 2004). Although both Afu-PriS1/L and Pfu-PriS1/L are tolerated to some mismatches, Mma-PriS1/L cannot tolerate the extension of all 12 mismatch combinations.

can drastically increase the preference of Pols λ and η for adenine over cytosine (Maga *et al*., 2007; van Loon *et al*., 2010). However, this has not yet been observed for archaeal PriS/L. Apurinic/apyrimidinic (AP Sites) which are produced as a result of oxidative stress, can also block replication. Similar to PrimPol and Pol γ, the mitochondrial replicative polymerase, archaeal PriS/L cannot bypass this lesion (Graziewicz *et al*., 2007; Keen *et al*., 2014; Pinz *et al*., 1995). We also demonstrated that Afu-PriS1/L could perform TLS activity on a template containing a CPD photo-lesion. Afu-PriS1/L can catalyse error-free bypass of CPDs through incorporating two dAs opposite a T-T CPD.

In contrast to the tested hyperthermophilc archaea, the PriS1/L complex isolated from *M.maripalidus*, which is a mesophile, could not read-through either 8-oxo-dG or CPD lesions. However, similar to Afu and Pfu primases it can replicate past dU bases. The inability of Mma PriS1/L to bypasss CPD and 8-oxo-dG lesions may reflect the difference in environmental conditions under which these organisms reside. Since *M.maripalidus* grows in moderate environmental conditions, comparing to two other species, it may be exposed to lower levels of genotoxic DNA lesions, such as cross-links. Another possibility could be the usage of other repair / damage tolerance pathways in this species to overcome such damage during replication.

As PriS/L is able to insert nucleotides opposite, and extend from bulky and distorting lesions, most probably the active site of this polymerase, which is involved in phosphodiester bond formation, is large and flexible enough for DNA lesions to be accommodated. Therefore, it is not suprising that PriS/L displays low fidelity when incorporating bases opposite some lesions. Notably, this hallmark is also exhibited by PrimPol, which has low fidelity when bypassing highly distorting lesions (Bianchi *et al*., 2013; Keen *et al*., 2014; Rudd *et al*., 2013). Accommodation of lesions, e.g.8-oxo-dG, at the active site of PriS/L decreases the fidelity of the polyemerase, hence PriS/L is a low fidelity polymerase that can connect non-specifically to replicating base pairs.

The presence of different DNA lesions and obstacles in the genome is pro-mutagenic, therefore, to overcome this problem cells have developed various DNA repair pathways. However, in some cases, repair mechanisms are not able to remove the damage, therefore, the lesion stays in the genome and if left

unrepaired it can lead to problems, such as replisome stalling. In this situation, DNA damage tolerance mechanisms have been evolved by cells to rescue stalled replication forks. DNA damage tolerance can either occur at the fork by TLS or template switching, or it can take place after replication by repriming. During repriming on the leading strand, a primase initiates DNA synthesis *de novo* downstream of lesion following replication stalling at the sites of damage, which leaves a ssDNA gap opposite the lesion. Repriming downstream of lesions has been proposed for different organisms, such as, mammalian cells and budding yeast (Rupp and Flander, 1968; Lehman, 1972; Lopes *et al*.,2006;). Although we have discovered that archaeal replicative primases (PriS/L) can act as TLS DNA polymerases that bypass DNA lesions and rescue stalled replication forks, we were unable to show repriming downstream of lesions by archaeal PriS/L. Since these enzymes are proficient DNA primases, it is conceivable that they perform repriming downstream of lesions but our assays are not sensitive enough to detect this. Therefore, further studies are required to investigate this.

Until recently, it was strongly believed that DNA primases were a class of proteins that were only required for initiation of DNA replication through synthesis of short RNA primers. However, a growing body of evidence has indicated that the enzymatic activity of these diverse enzymes is not limited to producing RNA primers. Indeed, this class of proteins are capable of performing a wide range of roles from DNA replication to DNA damage tolerance and DNA repair. It is now known that AEP-like primases, which were initially identified in eukaryotes, first arose in prokaryotes and bacteriophage to facilitate DNA repair mechanisms (Weller *et al*., 2002; Della *et al*., 2004). These studies highlighted the necessity for a re-evaluation of both the origin and roles of AEP-like primases in biology. DnaG is a bacterial DNA primase which fulfils the role of primer synthesis to initiate replication in bacteria and bacteriophage. Soon after the discovery of the polymerase activity of archaeal primases, AEP orthologous were also found in prokaryotes. This discovery suggests an early diversification in AEP protein's roles. It is believed that the last universal common ancestor (LUCA) employed a dual-primase system consisting of both DnaG and AEP (PriS/L) primases (Guilliam *et al*., 2015; Hu *et al*., 2012). Based on this hypothesis, selective pressure eliminated AEPs as replicative primases in bacteria and retained their

role in NHEJ (Weller and Doherty, 2001). While in archaea, the replicative primase activity of DnaG was supressed by selective pressure.

Archaeal DnaG comprises N- and C- terminal domains flanking a TOPRIM domain. While NT and TOPRIM domains are conserved among all archaea, conservation of CTD correlates with the presence of the exosome. Recently it was discovered that the NT is a novel RNA-binding domain. It was also shown that degradation of A-rich RNA by the exosome can be increased by a protein containing the eukaryotic homologue of Cs14 and the NT domain of DnaG. Together these findings suggested that archaeal DnaG is an RNA-binding protein and this enzyme in the context of the exosome degrades stable RNA (Hou *et al*., 2014).

In 2005, following division of the AEP superfamily into three distinct clades, using in silico analyses, a second primase called PrimPol was discovered as a novel AEP in higher eukaryotes (Iyer *et al*., 2005). PrimPol is assigned to the NCLDV-herpes virus clade. Interestingly, the primase and polymerase activities described in this thesis (Chapter 3) for archaeal replicative primases (PriS/L) mirror primase and polymerase functions of PrimPol. PriS, and the NHEJ AEPs, belong to the more ancient “proper” clade of AEPs that also includes the eukaryotic PriS / Prim1.

Previously, it was assumed that Dpo4, a defining member of the Y family DNA TLS polymerases, which exists in some species of archaea and has lesion-bypass properties, would be important for damage tolerance. However, it was recently been shown that in the absence of Dpo4, *Sulfolobus* strains do not show increased sensitivity to damaging agents, including UV radiation (Sakofsky *et al*., 2012). This study suggests the presence of other TLS pathways for traversing replication blocking damages in archaea. Notably in this regard, another study on *Sulfolobus* reported that UV induction could decrease expression of DNA replication genes, notably with the exception of Pris, and increase expression of some genes encoding unknown proteins (Gotz *et al*., 2007). The mutagenic consequences of ethyl methanesulfonate (EMS) and UV on *Pyrococcus abyssi* has been tested. This study indicated both EMS and UV could significantly increase the rates of mutagenicity, such that the spontaneous mutation frequency elevated ~150-fold after EMS treatment and ~400-fold after UV exposure (Watrin

and Prieur, 1996). This finding supports the possible existence of DNA damage tolerance mechanisms in archaea. Furthermore, fractionation experiments of *Pyrococcus furiosus* whole cell extracts, which showed polymerase activities in three main fractions, also supported this prediction (Ishino and Ishino, 2006). Notably, TLS activity was observed in one fraction and, even though the polymerase that catalysed this activity was not identified, it was reconcilable with the elution peak containing the PriS/L complex although the significance of this escaped the notice of the authors.

Together, these findings by other groups add additional support to our presented model in this thesis, which postulates that archaeal replicative primases play important roles in DNA damage tolerance in some archaea by performing translesion synthesis and also probably by repriming replication restart downstream of lesions or secondary structures (Figure 4.18). Since replicative primases are core components of the replisome, it asserts that in many living organisms the DNA replication machinery is inherently TLS proficient and that other DNA damage tolerance mechanisms may act as “back-ups” to rescue more profoundly stalled replication forks. These findings shed new light on the further roles of DNA primases during DNA replication and the subsequent evolution of related PrimPol-centric TLS/ repriming pathways in eukaryotic cells.

Taken together, the data presented in this chapter indicates that, similar to other TLS polymerases, PriS/L has a lower fidelity than replicative polymerases. Hence, PriS/L is a relative error-prone polymerase and the enzymatic activities of this protein need to be tightly regulated. Otherwise, its unscheduled or deregulated activities would synthesise long tracts of DNA in an error-prone fashion, which could lead to the accumulation of multiple genetic errors.

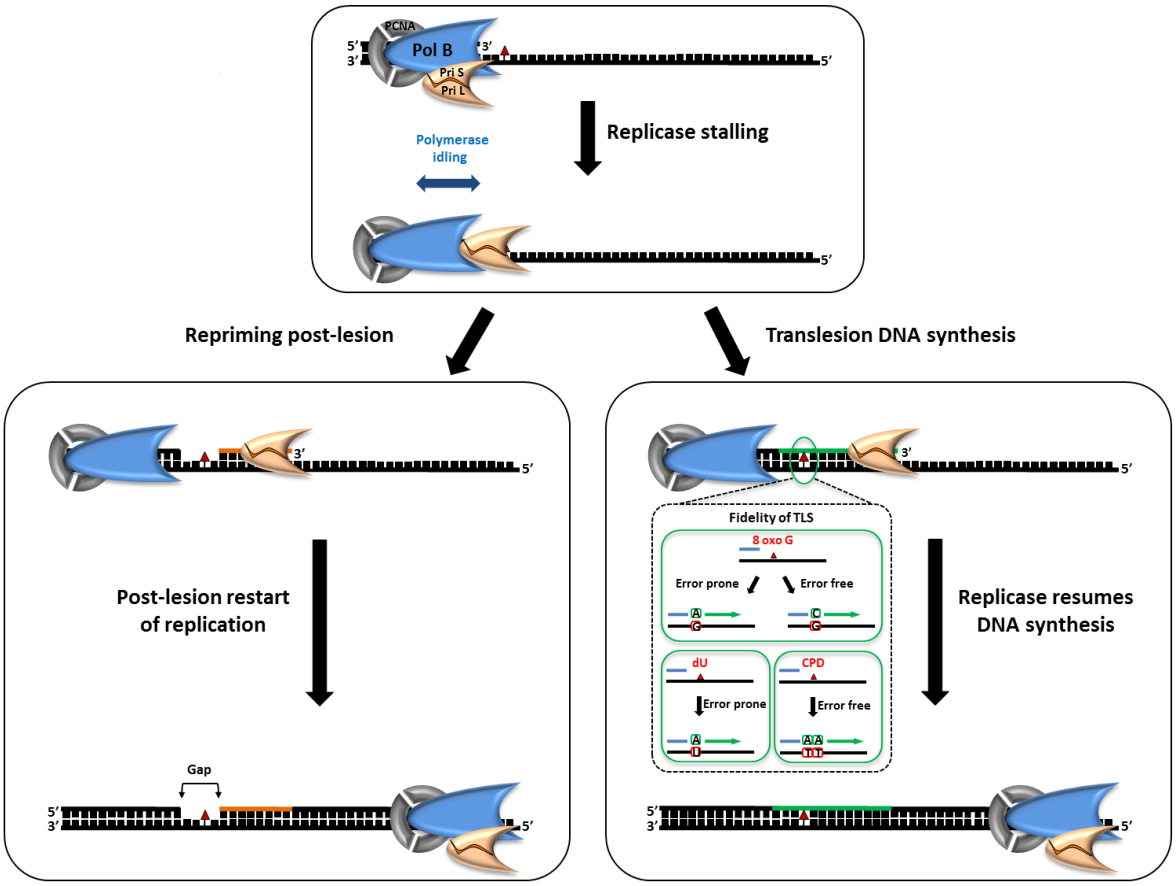


Figure 4.18. The roles of archaeal replicative primases in DNA damage tolerance

The picture on the top illustrates the polymerase activity of replicative polymerase (Pol B, blue) with sliding clamp (PCNA). Encountering of the PolB/PCNA complex with a blocking lesion leads to polymerase idling and recruitment of the primase (PriS/L complex, yellow). In the next step, depending on the type of damage, PriS/L rescues the stalled replication employing either translesion DNA synthesis (TLS) activity or re-priming. The picture on the right shows a situation where the blocking lesion is rather small (e.g., 8-oxo dG, dU, or CPD) and primase (PriS/L) can bypass the damage through TLS activity, this leads to resuming the DNA replication by PolB/PCNA complex. The fidelity of TLS carried out by *A.fulgidus* PriS/L is also illustrated. The picture on the left demonstrates the re-priming activity of PriS/L downsteam of relatively larg lesions which leaves a gap on the damaged DNA and synthesis a short primer after the lesion so that PolB/PCNA complex can restart the DNA replication.

Chapter 5

Characterisation of the regulation of archaeal replicative primases by RPA

5.1. Introduction

Unlike replicative polymerases, specialised TLS DNA polymerases display low fidelity when copying undamaged DNA templates. Therefore, these polymerases need to be regulated at several different levels, otherwise their deregulated activity gives rise to mutagenesis (Sale *et al*., 2012). One way to regulate the enzymatic activity of proteins, which are required in response to specific stresses, such as TLS polymerases, is to control their intracellular concentration. In *E. coli*, the concentration of these proteins is under the control of the SOS response to ensure that their concentration is low in undamaged cells and induced in the presence of damage (Michel, 2005). In vertebrate cells, access of TLS proteins to the replisome is regulated through post-translational modification of PCNA. Replication fork stalling that occurs in the presence of damage stimulates mono-ubiquitination of PCNA which in turn increases its affinity for TLS polymerases. This leads to recruitment of TLS enzymes to the stalled replication fork. Following bypass of the lesion by a TLS polymerase, TLS polymerase idling occurs. The TLS polymerase is then displaced and replaced with a high fidelity replicative DNA polymerase which continues DNA replication (Friedberg *et al*., 2005). This polymerase switching mechanism plays a key role in limiting the progression of replication by a low fidelity DNA polymerase. This allows TLS polymerases access to the replisome only when DNA damage tolerance is required. Recently, it was shown that there is no interaction between PrimPol, which is involved in DNA damage tolerance through TLS or re-priming, and PCNA. Indeed, it was discovered that, human PrimPol interacts with the key cellular single-stranded DNA binding proteins, RPA, and mitochondrial SSB (mtSSB) (Guilliam *et al*., 2015). Both RPA and mtSSBs increase the activity of their respective replicative Pols, δ and γ (Oliveira and Kaguni, 2010; Tsurimoto and Stillman, 1989). By contrast, polymerase activity of PrimPol is restricted by both RPA and mtSSB (Guillium *et al*., 2015). In addition, these SSBs can also inhibit the primase activity of pol α-primase (Collins and Kelly, 1991). More recently, it has been demonstrated in our lab that, depending on the concentration of RPA, the effect of RPA on the primase activity of PrimPol can be changed. While sub-saturation of RPA significantly stimulates primer synthesis, increasing RPA's concentration restricts this activity (Guilliam *et al*., under review). Chapter 4 established that archaeal PriS1/L is an error-prone DNA polymerase, capable of bypassing uracil

and base lesions. Therefore, strict regulation is probably required during its participation in DNA replication. The aim of this chapter was to study the regulatory effect of RPA on *A.fulgidus* PriS1/L. In particular, to study whether RPA was able to affect the enzymatic activities of PriS/L, which would be consistent with the effect of RPA and mtSSB on human PrimPol. In addition, in this chapter we also discuss our attempts to identify the potential interaction between RPA and PriS1/L in *A.fulgidus.*

5.2. Cloning the archaeal PriS/L genes into expression vectors

Eukaryotic RPA is a heterotrimer consisting of three subunits (RPA70, RPA32 and RPA14). Although in general archaeal RPAs are similar to their eukaryotic counterpart, members of archaea possess various forms of single-stranded binding proteins (Kelman and Kelman, 2014). In euryarchaeota, different forms of RPA, such as single, multiple and complex have been identified. Unlike *Methanococcus jannaschii* and *Methanothermobacter thermoautotrophicus*, RPA from *P.furiosus* forms a hetero-oligomeric complex composing of three subunits including, RPA41, RPA32 and RPA14 (Komori and Ishino, 2001). Two RPA41 homologues have been identified in *A.fulgidus*, RPA-780 which contains a zinc finger motif and RPA-382. Since RPA-780 and RPA-382 are ordered in a tandem arrangement in *A.fulgidus* genome, it is believed that these RPAs are functionally associated (Komori and Ishino, 2001). In order to characterise the effect of archaeal RPA on the polymerase and primase activities of the *A.fulgidus* replicative primase, two ORFs corresponding to RPA orthologous in *A.fulgidus* (AF0-780 and AF0-382) were amplified from *A.fulgidus* genomic DNA (using primers in table 2.1) and cloned into the pET28a and pGEX-6p-1 *E.coli* expression vectors, respectively. Therefore, pET28a:Afu-RPA-780 and pGEX-6P-1:Afu-RPA-382 expression constructs were generated (Figure 5.1).

5.3. Expression and purification of A.fulgidus RPA-780 and RPA-382

The Afu-RPA-780 and Afu-RPA-382 expression constructs were transformed into Rosetta *E.coli* strain. 3 litre cultures were grown at 37°C for 3 hours until exponential phase and induced with 1mM IPTG at 25°C overnight. First, we

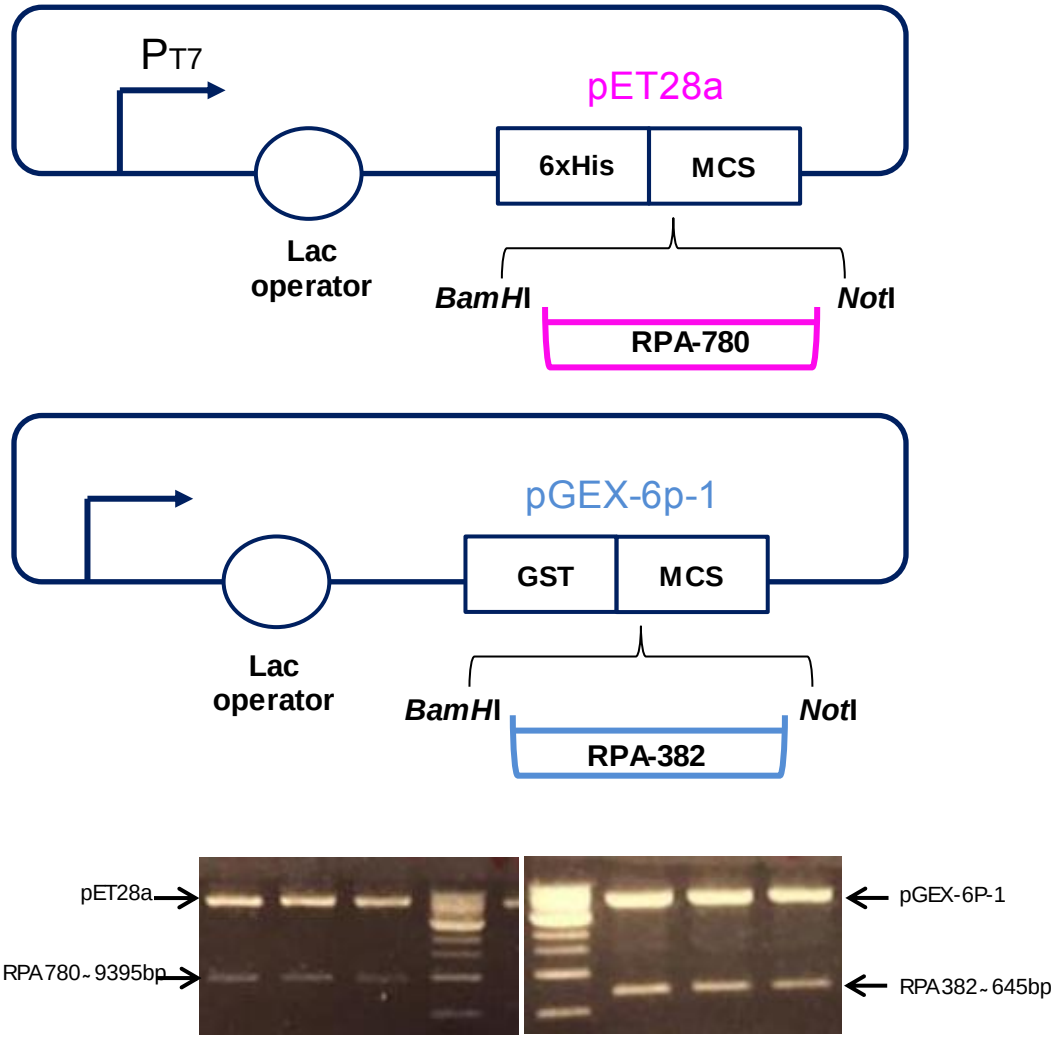


Figure 5.1. Cloning of the *A. fulgidus* RPA genes, Afu-RPA-780 and RPA-382

The open reading frames corresponding to RPA-780 and RPA-382 were PCR amplified from *A.fulgidus* genomic DNA (Table 2.3), introducing the applicable restriction sites to allow insertion into the multiple cloning site (MCS) of the pET28a and pGEX-6p-1 expression vectors, respectively. RPA-780 and RPA-382 were cloned in-frame with 6-histidine and GST tags downstream of the promoter. The PCR products were combined with 10x DNA loading dye and run on 1% agarose gels containing ethidium bromide.

attempted to purify RPA-780, which was cloned into the pET28a expression vector using Ni^{2+}-NTA agarose affinity chromatography, which specifically binds the fused 6X His tag. The prepared soluble cell lysate was loaded onto the Ni^{2+}-NTA chromatography, followed by elution of bound RPA-780 from the nickel column with a range of imidazole concentration. Eluted RPA-780 was resolved on a SDS-polyacrylamide gel to confirm the correct size. The result was in agreement with its predicted molecular mass. RPA-780 has a predicted molecular mass of ~35 kDa (Figure 5.2). To remove *E.coli* contaminants, the Ni^{2+}-NTA eluate was subjected to purification on a Heparin affinity column. The column was equilibrated with Heparin Buffer A (150 mM NaCl, 40 mM Tris-HCl pH 7.5, 10% (v/v) glycerol). After 1:10 dilution with Heparin Buffer A (100mM NaCl), the Ni^{2+}- NTA eluate was slowly loaded onto the Heparin column and eluted with 500 mM NaCl. For the last step of purification, Hydrophobic Interaction Chromatography (HIC) was employed. HIC is usually used to separate protein based on hydrophobicity. To decrease the availability of water molecules in solution and enhance hydrophobic interactions, salting-out ions (e.g. sodium chloride, potassium chloride or ammonium sulphate) are used typically. Phenyl Sepharose 6 Fast Flow, a standard aromatic hydrophobic interaction chromatography (HIC) medium was chosen for this purification. Heparin fractions containing protein were first concentrated using a Vivaspin® centrifugal concentrator. Protein sample was prepared with a 1:1 volume of protein and HIC buffer with high-salt ammonium sulfate (2M $(NH_4)_2SO_4$) to make the salt concentration of the protein roughly equivalent to the starting buffer used for column equilibration. After loading the sample onto the column and washing it with sodium chloride and 1M ammonium sulfate, the bound RPA-780 was eluted in a buffer with 500 mM sodium chloride and no ammonium sulfate. Collected fractions were monitored at 280 nm and analysed by SDS-PAGE. A band corresponding to RPA-780 was detected on the gel (Figure 5.2).

Afu-RPA-382 was cloned into the pGEX-6P-1 vector with a GST tag located at the N-terminus. Since GST has a high affinity to glutathione coupled to a Sepharose matrix, we attempted to purify RPA-382 using Glutathione Sepharose 4 Fast Flow medium. Prepared cell lysate was mixed with glutathione sepharose and incubated at 4°C for 2 hours. The cell lysate and resin mix was packed into a column.

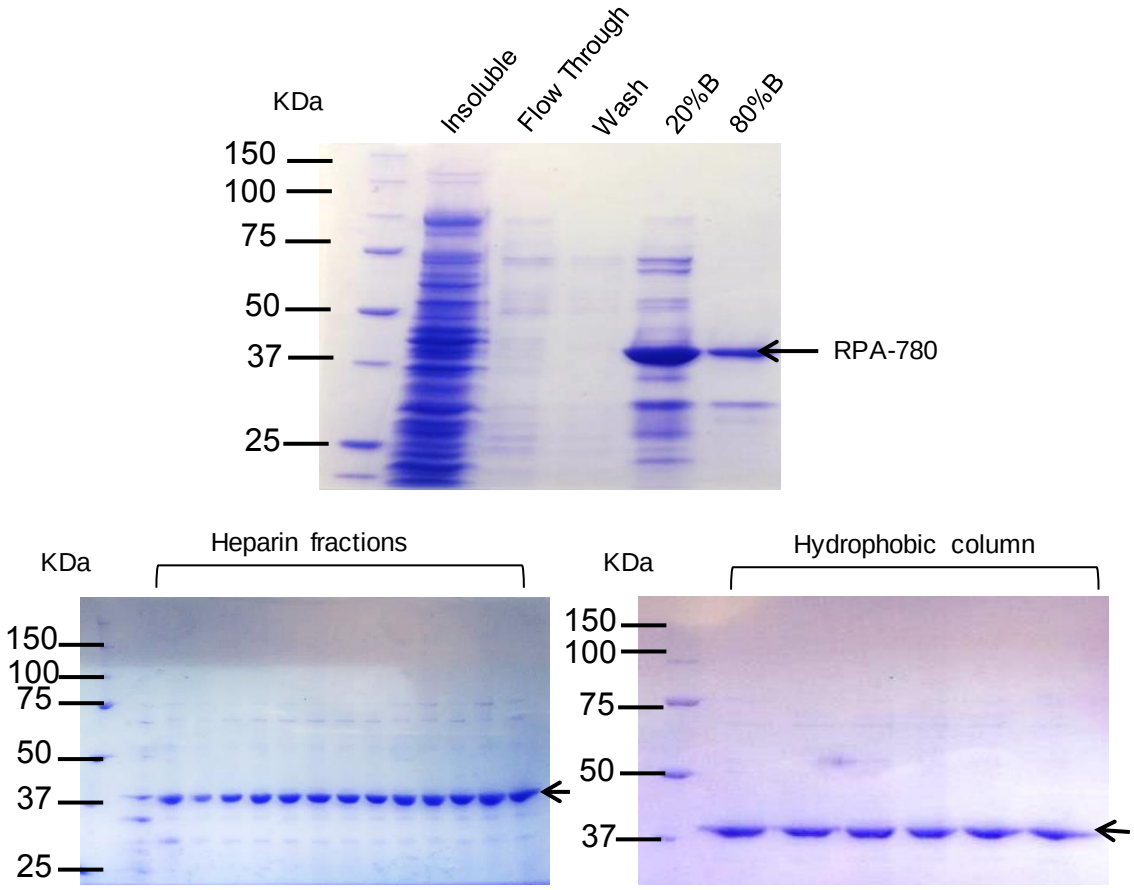


Figure 5.2. Chromatography purification of Afu-RPA-780

A culture of *E.coli* strain Rosetta transformed with pET28a:Afu-RPA-780 expression construct was grown at 37°C for 3 hours then induced with 1 mM IPTG and incubated overnight. Prepared cell lysate was subjected to Ni^{2+}-NTA chromatography. Bound Afu- PriS1/L was washed and then eluted with range of imidazole concentration. The 80% B eluted peak fraction was then subjected to heparin affinity chromatography to remove *E.coli* contaminants. The bound RPA-780 was eluted with 500 mM NaCl. Heparin fractions were loaded into Phenyl Sepharose 6 Fast Flow, a standard aromatic hydrophobic interaction chromatography and washed with 1 M ammonium sulfate. the bound RPA-780 was eluted in a buffer with sodium chloride and no ammonium sulfate. Collected fractions were monitored at 280 nm and analysed by SDS-PAGE. As indicated, the size of obtained bands was in agreement with RPA-780 predicted molecular mass (~ 35 KDa).

The column was subjected to ÄKTA prime system and then washed with 100 mM NaCl. For elution, buffer containing 40mM Tris pH 7.5, 50 mM NaCl, 0.5 mM TCEP, 20 mM containing 500 mM NaSCN, reduced glutathione and, 10% glycerol was used. Eluted proteins were analysed by SDS-PAGE gel. Although RPA-382 was expressed in soluble form, the protein did not elute from the GST column. Unexpectedly, the remained beads after elution showed no protein on the SDS gel. This result suggested that the protein did not bind to GST beads (data not shown), an observation that could potentially be explained if the fusion protein caused misfolding. We decided to remove the GST tag from the RPA-382 expression construct and reverted to co-purification of RPA-780 with RPA-382 using Affi-Gel Blue affinity chromatography (see below).

5.4. Co-purification of Afu-RPA-780/RPA-382 using Affi-Gel Blue chromatography

Affi-Gel Blue is a cross-linked agarose resin to which the dye Cibacron Blue F3GA has been covalently attached. The Cibacron Blue dye has ionic, hydrophobic, and aromatic characteristics and it therefore displays affinity for many types of proteins. In order to prevent the potential disrupting effect of the GST tag on the formation of Afu-RPA-780/382 complex, as in some cases the fused tag can cause insolubility of proteins, we first tried to remove the GST tag from pGEX-6P-1:Afu-RPA-382 through PCR amplification with new primers complementary to the ends of the GST sequence. After ligation in order to ensure that the GST sequence was removed from the construct, double digestion using the relevant restriction enzymes was carried out. Finally, deletion was confirmed by sequencing.

The pET28a:Afu-RPA-780 expression construct was co-transformed with pGEX-6P-1:Afu-RPA-382 construct (with no tag) into Rosetta *E.coli* strain. Three litre cell culture was grown to exponential phase (OD_{600}=~0.6) at 37°C for 3 hours and then induced for expression with addition of 1 mM IPTG at 25°C overnight. Cell lysate was prepared using lysis buffer (50 mM HEPES-NaOH, pH 7.5, 0.1mM EDTA, 10 mM β-mercaptoethanol, 500 mM NaCl, 100 mM spermidine, 4 mg/ml lysosome and 1 mM phenylmethylsulfonyl fluoride (PMSF). A 5 mL Affi-Gel Blue (Bio-Rad) column was pre-equilibrated with buffer A (50 mM HEPES-NaOH, pH

7.5, 10% glycerol, 10 mM β-mercaptoethanol, and 500 mM NaCl). After loading the clarified cell lysate onto the Affi-Gel Blue column, the column was washed first with buffer A, and then with buffer A containing 1 M NaCl, then with buffer A eluted with 50 mL of buffer A containing 1.5 M NaSCN. Eluted proteins were collected and dialyzed against buffer A with 300 mM NaCl, 50 mM HEPES-NaOH, and 0.1% NP-40. Precipitated materials were removed by centrifugation. RPA-780/RPA-382 complex was co-eluted with considerable amounts of *E. coli* contaminants (Figure 5.3). Therefore an additional step of purification was required. To remove *E.coli* contaminants, the purified Afu-RPA-780/RPA-382 complex was subjected to an anion exchange column. Afu-RPA-780/RRPA-382 complex was eluted with 300 mM NaCl, while *E.coli* contaminants were washed away with low salt. The successfully purified RPA complex, with an apparent molecular mass of ~ 35 and ~25 kDa corresponding to RPA-780 and RPA-382 respectively resolved on a SDS-PAGE which was in agreement with their predicted molecular masses (Figure 5.3).

5.5. Examination of DNA binding affinity of *A.fulgidus* RPA

Eukaryotic RPA, a heterotrimeric complex, has high affinity for binding to ssDNA (Broderick *et al*., 2010). Additional studies have shown that archaeal RPA homologues also bind to ssDNA. The single-stranded DNA binding proteins from *Sulfolobus solfataricus* and *Methanococcus jannaschii* exhibited ssDNA binding with high affinity (Kernchen and Lipps, 2006; Kelly *et al*., 1998). Interestingly, in contrast to other archaeal and eukaryotic RPA proteins, each RPA subunit isolated from *Methanosarcina acetivorans* separately showed a distinct ssDNA binding ability (Robbins *et al*., 2003). To confirm the ability of recombinant *A.fulgidus* RPA-780 and RPA-780/RPA-382 complex to bind to ssDNA, we employed agarose gel mobility shift assays with 50-mer fluorescently labelled ssDNA (Table 2.7) and increasing concentrations of proteins. These assays showed that both RPA-780 and RPA-780/RPA-382 complex significantly bind to ssDNA. Our data suggested that Afu-RPA-780 has higher affinity for ssDNA binding when it is in complex with Afu-RPA-382, compared to Afu-RPA-780 alone (Figure 5.4). Moreover, in the case of Afu-RPA-780 protein alone, binding was apparent at 0.1 µM protein and a prominent super-shifted complex was observed at a protein concentration of 5 µM (Figure 5.4B), while in the case of the Afu-RPA-780/RPA-382 complex, binding was apparent at 0.05 µM protein and a

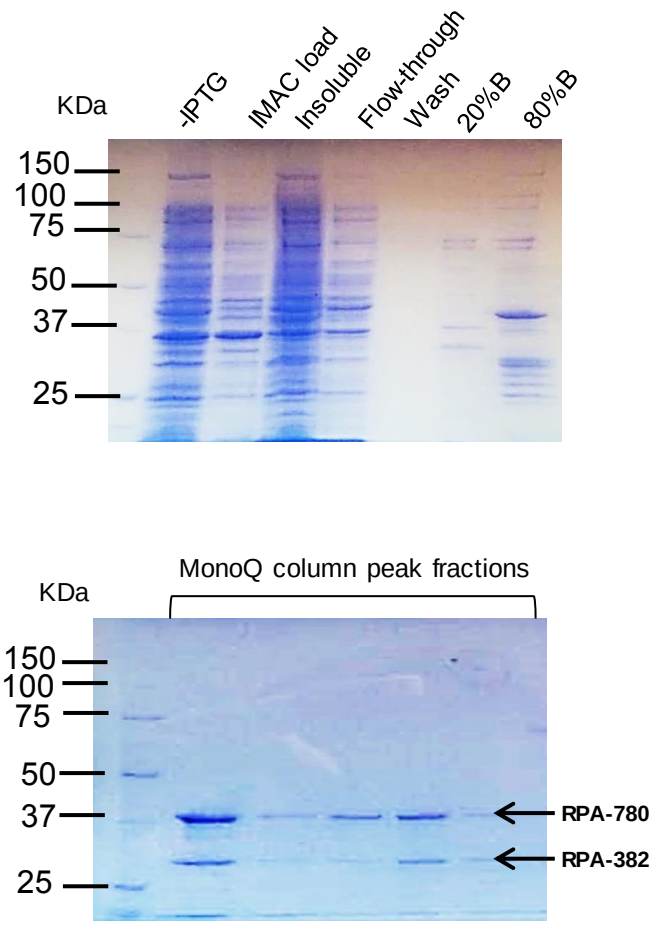


Figure 5.3. Co-purification of the Afu-RPA (RPA780 / RPA382) complex

A 3 litre culture of *E.coli* strain Rosetta transformed with Afu-RPA780/Afu-RPA382 was grown to exponential phase and then induced with 1 mM IPTG. A 5 ml pre-packed Affi-Gel Blue column. was washed sequentially with 500 mM NaCl, 1 M NaCl and, 500 mM NaSCN. Next, cell lysate was loaded into the column and the bound protein eluted with 1.5 M NaSCN. To remove *E.coli* contaminants, The 80% B eluted peak fraction from Affi-Gel Blue column was subjected to MonoQ column and after wash step, the bound RPA780/RPA382 complex was eluted with 500 mM NaCl. The successfully purified complex, with an apparent molecular mass of ~ 35 and ~25 kDa corresponding to RPA780 and RPA382, respectively, resolved on a SDS-PAGE which was in agreement with their predicted molecular masses

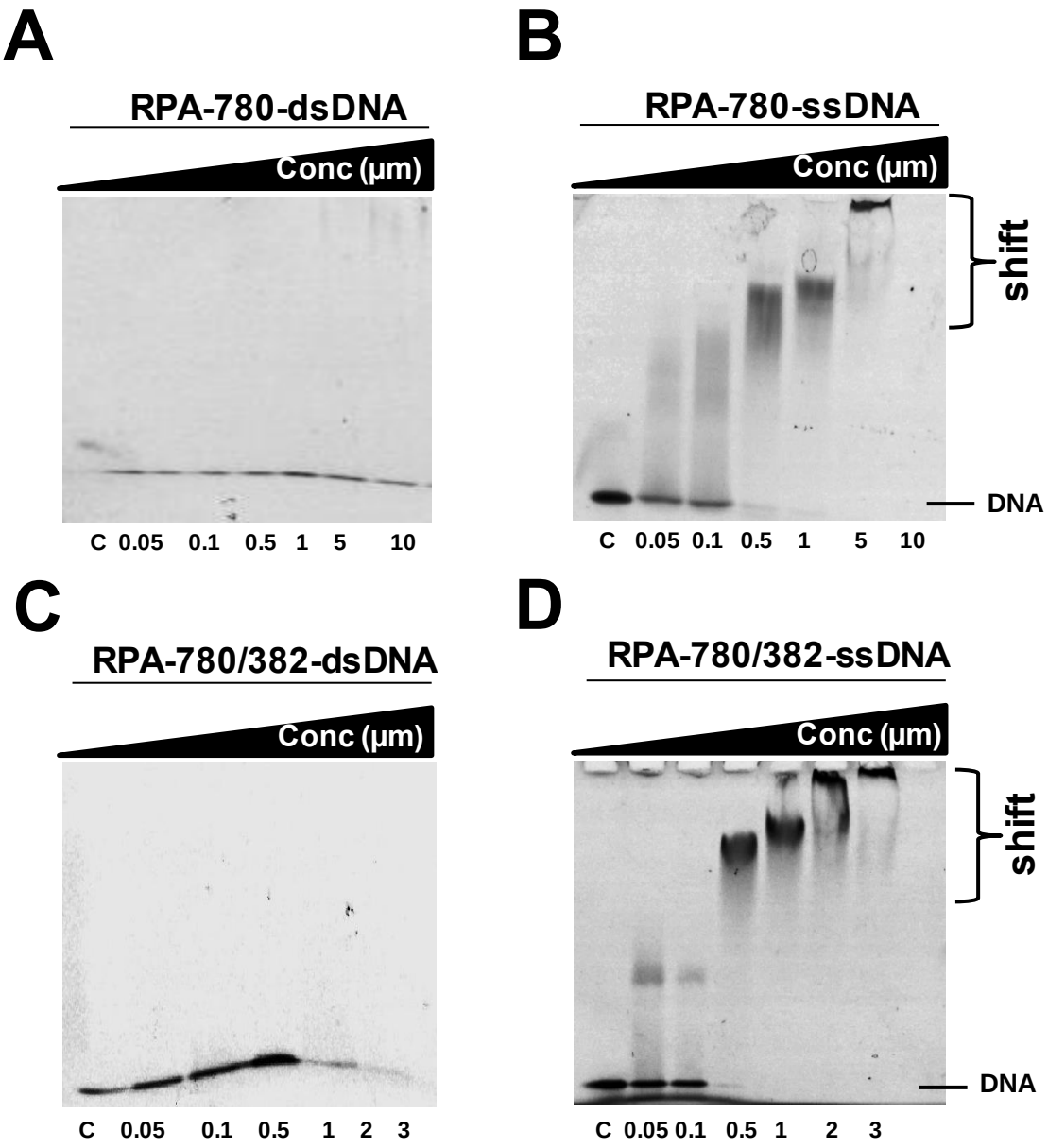


Figure 5.4. ssDNA binding activities of Afu-RPA proteins examined using EMSA

50-mer fluorescently labelled ssDNA (Table 2.7) was incubated with increasing concentration of either RPA-780 **(B)** or RPA-780/382 **(D)**. The concentrations of RPA used in these experiments are shown below each gel. In parallel, both RPA-780 and RPA-780/382 were incubated with dsDNA as the negative control **(A,C)**. The RPA-780/382 complex shows binding to ssDNA at lower concentration than RPA780 single protein.

super-shifted complex at 2 µM protein (Figure 5.4D). Identical reactions were also performed with dsDNA, instead of ssDNA, as a negative control. As expected, both Afu-RPA-780 and RPA-780/RPA-382 did not show any binding to dsDNA. We quantified our data to determine $K_{D\ (DNA)}$ values for Afu-RPA-780 and Afu-RPA-780/RPA-382 using Sigma plot. The dissociation constant of Afu-RPA-780/RPA-382 to ssDNA was nearly four times lower than Afu-RPA-780. The Afu-RPA-780/RPA-382 $K_{D\ (DNA)}$ was 0.25 µM, compared with Afu-RPA-780 of 0.97 µM (Figure 5.5). To confirm our results, we repeated these experiments using different concentrations of each protein (data not shown). The outcome was consistent with the results described earlier. Together, our data establish that purified RPA complex from *A.fulgidus* is a real single-stranded DNA binding protein that possesses high affinity for ssDNA.

5.6. Regulatory effect of RPA on enzymatic activities of Afu-PriS1/L

There is a direct interaction between RPA and the primase subunit of Polα/primase complex in eukaryotes (Dornreiter *et al*., 1992). Recently, it was demonstrated that there is a direct interaction between PrimPol and the N-terminal domain of RPA70, which suggested that PrimPol might be recruited to the replication fork by RPA (Guilliam *et al*., 2015; Wan *et al*., 2013). In addition, it was found that RPA can supress enzymatic activities of PrimPol (Guilliam *et al*., 2014). In archaea, immunoprecipitation experiments strongly supported the idea that RPA41 from *P. furiosus* interacts with the primase 41 subunit (Pfu41) (Bocquier *et al*., 2001; Komori and Ishino, 2001). However, the influence of RPA on enzymatic activities of archaeal primase has not been yet been evaluated.

5.6.1. The effects of RPA on the primase activity of Afu-PriS1/L

As priming by Afu-PriS1/L was much more efficient on ssM13 templates, shown in Chapter 3, we examined the ability of RPA to influence the primase activity of Afu-PriS1/L on ssM13 templates. In addition, since Afu-PriS1/L showed a preference to prime using rNTPs over dNTPs (Chapter 3), in this radioactivity-based ssM13 primase assay, the enzyme was incubated with ssDNA, rNTPs and Mg^{2+}. 500 nM, 1 µM and 2 µM of RPA-780/382 complex was added into the reactions before the addition of PriS1/L (1 µM).

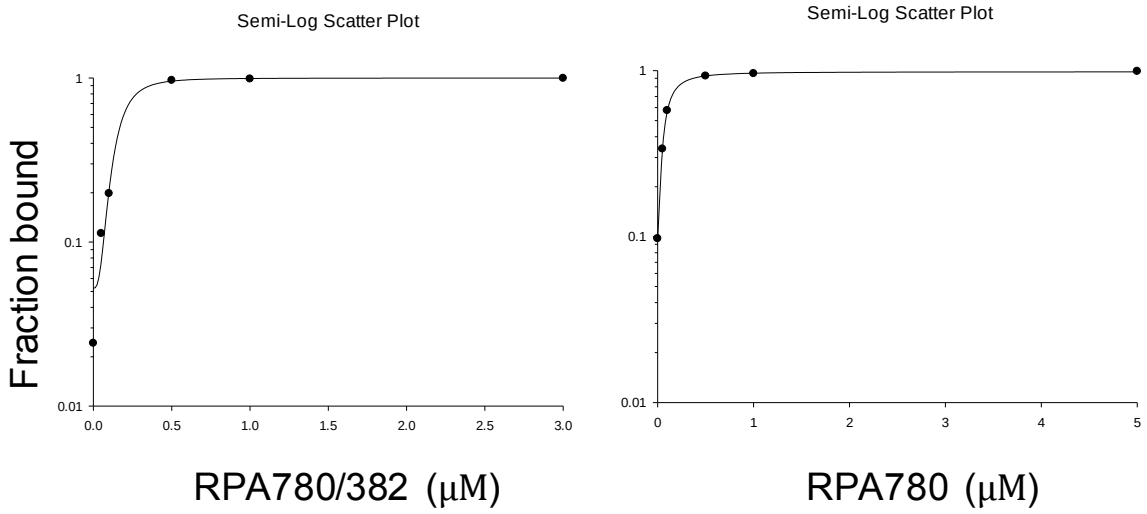


Figure 5.5. Quantification of DNA binding efficiency of Afu-RPA

The ssDNA binding affinity of Afu RPA, detected by gel mobility shift assays (Figure 5.4), was subsequently quantified using Sigmaplot to determine the K_D(DNA) values for RPA-780 alone and the RPA-780/382 complex. The binding efficiency of RPA-780/382 for ssDNA was nearly four times lower than RPA780. The RPA-780/382 complex has a calculated K_D (DNA) of 0.25 μM, compared with 0.97 μM for RPA780.

The RPA concentrations were chosen based on the obtained results from EMSA (Section 5.5). In order to ensure that the DNA is fully covered by RPA, excess amounts of RPA complex over ssDNA were used (Figure 5.6). As discussed previously, Afu-PriS1/L is able to synthesise primers on ssDNA in the absence of RPA. However, addition of RPA with increasing concentrations did not show any significant effect on the primase activity of Afu-PriS1/L (Figure 5.6). This observation was in contrast with studies of the Polα complex and PrimPol, which indicated that RPA suppresses *de novo* primer synthesis by these enzymes (Collins and Kelly, 1991; Guilliam *et al*., 2014).

5.6.2. The effects of RPA on primer extension by Afu-PriS1/L

Although RPA inhibits primer synthesis by Polα, it has been shown that this enzyme stimulates the polymerase activity of the Polα during elongation (Braun *et al*., 1997). This implicates a role for RPA in inhibiting Polα binding to ssDNA, while also acting to promote primer extension. Moreover, polymerase activities of both Polδ and Polγ are shown to be stimulated by RPA and mtSSB, respectively (Oliveira and Kaguni, 2010; Tsurimoto and Stillman, 1989). Next, the effect of RPA on the polymerase activity of Afu-PriS1/L was tested. T4 SSB was also used as a non-interacting control. Standard primer extension assays were carried out using a 50-mer primer-template substrate and increasing concentration of either RPA or T4. These experiments mimic the situation in which DNA polymerase and helicase are uncoupled because of the accumulation of DNA damage, which in turn leads to the formation of stretches of RPA-bound ssDNA (Byun *et al*., 2005; Lopes *et al*., 2006). Our data showed that, in the absence of RPA, Afu-PriS1/L could fully extend the primer by the last time point. Addition of RPA to the reaction led to significant inhibition of polymerase activity, especially when the concentration of RPA was >5-fold higher than the concentration of PriS1/L (Figure 5.7). Interestingly, the presence of T4 SSB also caused remarkable suppression of primer extension by Afu-PriS1/L (Figure 5.8). Restricted primer extension by PriS1/L in the presence of RPA proposed that PriS1 /L was not able to displace the RPA from ssDNA during primer extension. These findings suggest that free ssDNA ends adjacent to the primer-template junction are bound by PriS1/L until the PriS1/L encounters downstream RPA.

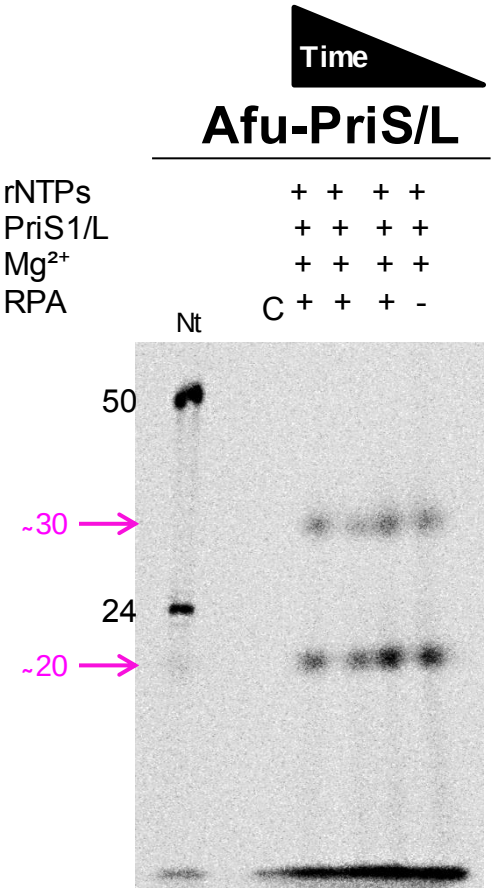


Figure 5.6. Afu-RPA does not affect the primase activity of Afu-PriS1/L

250ng of M13mp18 ssDNA was incubated with 10 mM $MgCl_2$, rNTPs and 100nM PriS/L, either alone or in the presence of RPA-780/382 complex. RPA was used at protein concentrations of 0, 100, 200 AND 500 µM. After 30 minutes incubation at 50°C, samples were subjected to electrophoresis. Primer synthesis by PriS1/L was detected in presence and absence of RPA. The left panel indicates the oligonucleotide nucleotide (Nt) length markers .

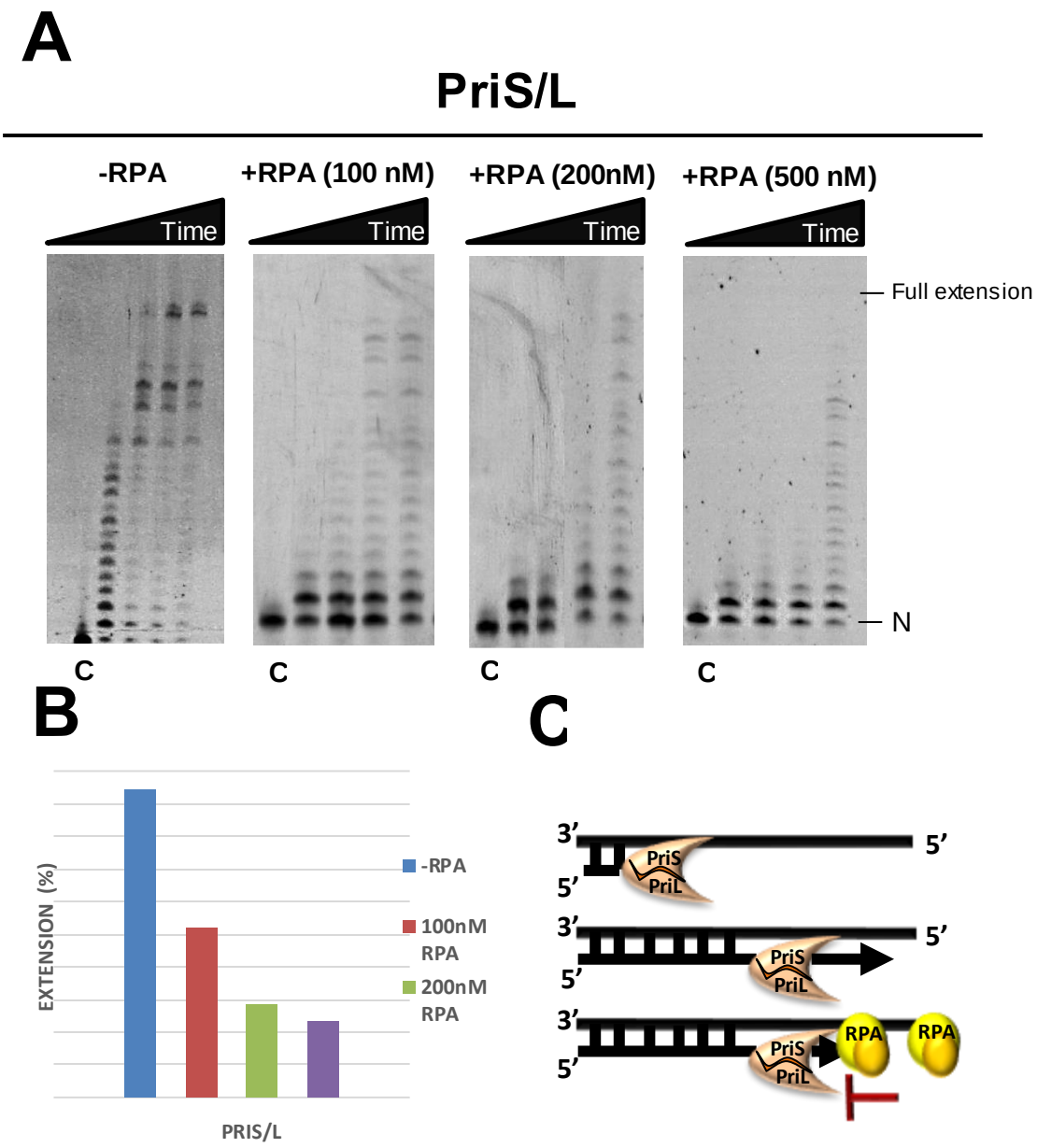


Figure 5.7. Afu-RPA limits the primer extension activity of the Afu-PriS1/L complex

20 nM of primer-template substrate with dNTPs and 100nM Afu-PriS1/L were incubated with increasing concentration of RPA as indicated above each gel, for 0.5,3,5,10 min. Identical reactions were also performed in the absence of RPA. **(A)** The polymerase activity of PriS1/L was significantly supressed with increasing concentrations of RPA. "N" denotes primer and "C" denotes no enzyme control. **(B)** Quantification of 10 minutes time-point of each gel confirmed limitation of PriS1/L primer extension by RPA. **(C)** Cartoon schematic showing the inhibition of PriS1/L primer extension activity by RPA.

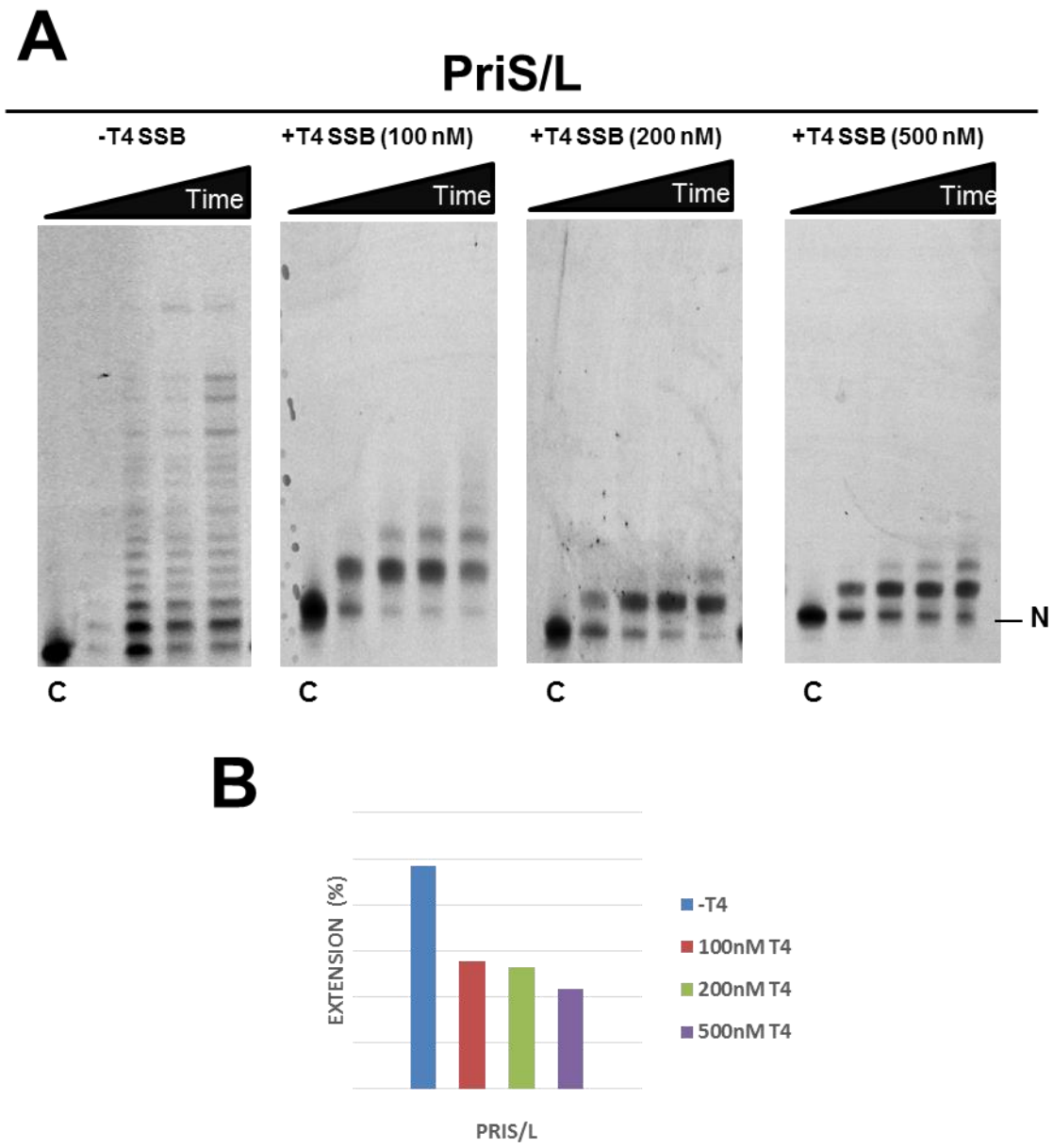


Figure 5.8. T4 limits the primer extension activity of Afu-PriS1/L

20 nM of primer-template substrate with dNTPs and 100nM Afu-PriS1/L were incubated with increasing concentration of T4 SSB as indicated above each gel, for 0.5,3,5,10 min. Identical reactions were also performed in the absence of T4 SSB. **(A)** The polymerase activity of PriS1/L was significantly inhibited with increasing concentrations of RPA. "N" denotes primer and "C" denotes no enzyme control. **(B)** Quantification of 10 minutes time-points of each gel confirmed the suppression of PriS1/L primer extension by T4 SSB.

This in turn leads to the dissociation of PriS1/L from DNA and due to presence of RPA the enzyme cannot bind to the ssDNA again. Notably, the varying levels of suppression by RPA and T4 SSB could be due to the different affinities of RPA and T4 SSB for the DNA template. Our data suggested that the polymerase activity of PriS1/L is also inhibited by T4 SSB. This effect is probably not as the result of PriS1/L interaction with SSBs. Notably, these observations are consistent with a recent study on the polymerase activity of human PrimPol, which also indicated a severe inhibition of PrimPol's polymerase activity by either RPA, mtSSB or T4 SSB (Guilliam *et al*., 2015). Archaeal replicative polymerases (PolB and PolD) were also incubated with RPA and T4 SSB in the same conditions in which the polymerase activity of PriS1/L was limited. Although the presence of RPA slightly decreased the polymerase activities of the replicative polymerases, both enzymes showed the ability to displace RPA (Figure 5.9) and T4 SSB (Figure 5.10) to extend most of the primers. This demonstrates the differences in the capacities of PriS1/L and the replicative polymerases to displace RPA.

Altogether, these results strongly indicate that RPA considerably inhibits polymerase activity of Afu-PriS1/L. As previously shown (Chapter 4), PriS1/L is a proficient TLS polymerase that can bypass different DNA lesions with relatively low fidelity. It is assumed that the inability of PriS1/L to displace RPA-bound ssDNA stretches, acts as a mechanism to restrict PriS1/L involvement in DNA replication.

5.7. Evalution of the interaction between Afu-PriS1/L and RPA

Uncoupling of priming and unwinding in response to replication fork stalling leads to fork uncoupling, which in turn results in the generation of ssDNA stretches in both DNA strands (Lopes *et al*., 2006). Although on the lagging strand synthesise of new Okazaki fragments reduces the impact of this process, on the leading, strand extended uncoupling can result in the production of long stretches of ssDNA. It has been reported that RPA can coat the resulting ssDNA and facilitate initiation of checkpoint responses (Zou *et al*., 2003).

A growing body of evidence indicates extensive interactions between RPA and other core components of the replisome in all domains of life. In eukaryotes, RPA interacts with the DNA polymerase α-primase complex directly via

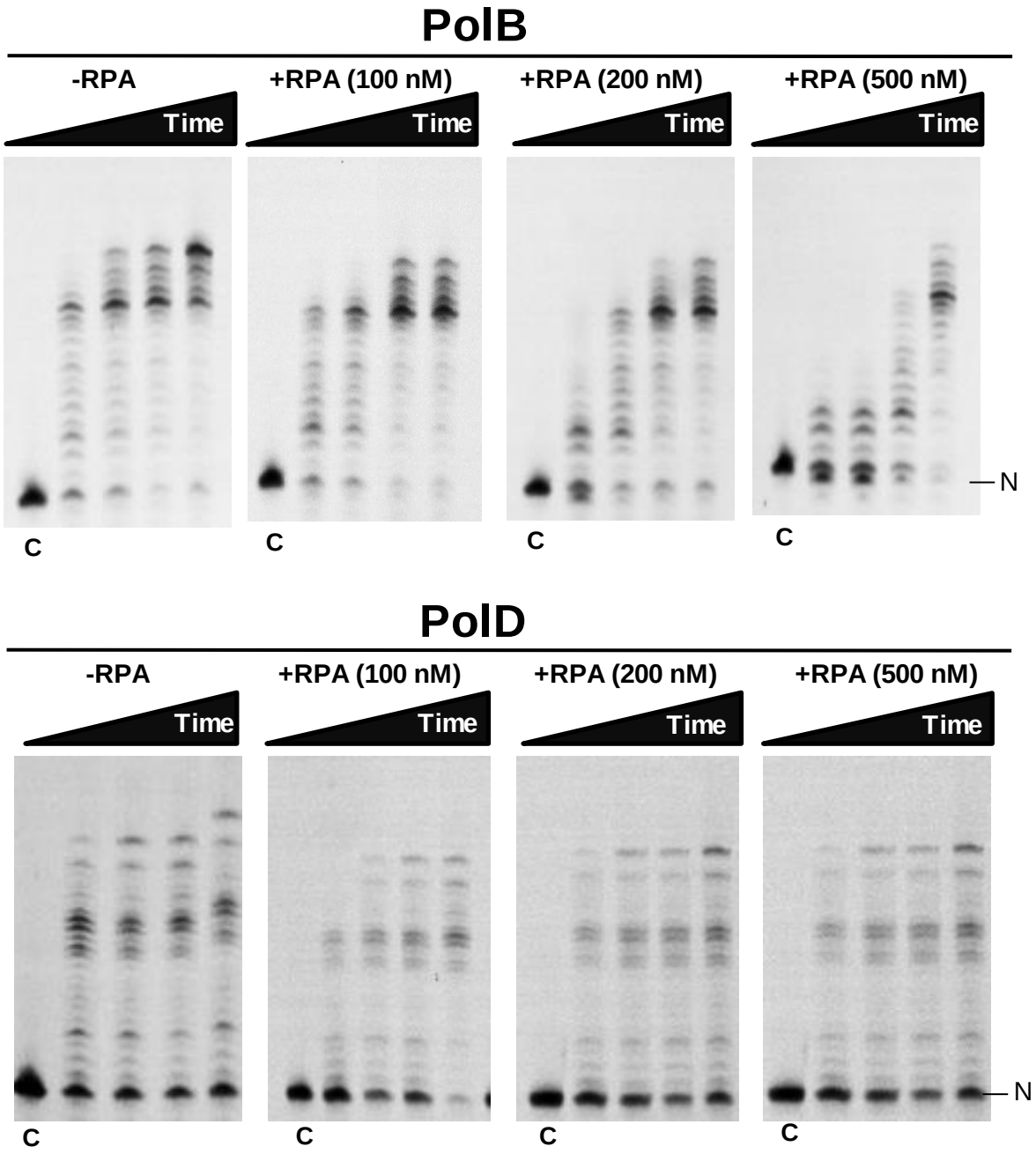


Figure 5.9. Archaeal replicative polymerases can displace Afu-RPA

20 nM of primer-template substrate with dNTPs and 100nM of either Afu-PolB or Afu-PolD were incubated with increasing concentrations of RPA as indicated above each gel, for 0.5,3,5,10 min. Identical reaction were also performed in the absence of RPA. Both PolB and PolD showed ability to displace RPA to extend primers. "N" denotes primer and "C" denotes no enzyme control.

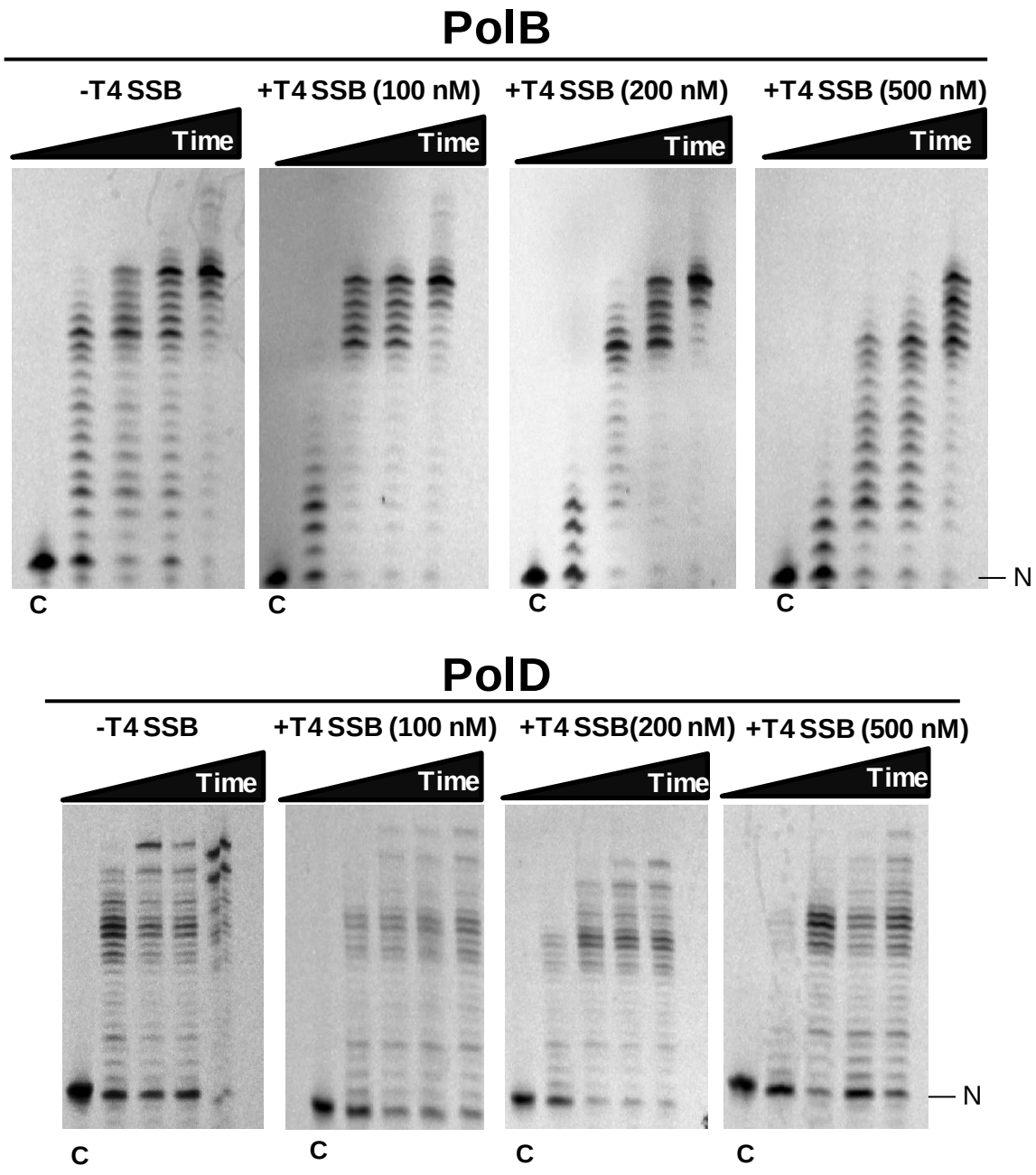


Figure 5.10. Archaeal replicative polymerases can displace T4 SSB

20 nM of primer-template substrate with dNTPs and 100nM of either Afu-PolB or Afu-PolD were incubated with increasing concentrations of T4 SSB as indicated above each gel, for 0.5,3,5,10 min. Identical reactions were also performed in the absence of T4 SSB. Both PolB and PolD showed ability to displace T4 SSB to extend primers. “N” denotes primer and “C” denotes no enzyme control.

interaction with the primase subunits (Dornreiter *et al*., 1992). Recent studies demonstrated the interaction between RPA70, the largest subunit of eukaryotic RPA, and human PrimPol (Wan *et al*., 2013; Guilliam *et al*., 2015). More recently, the structural basis for PrimPol's interaction with RPA has been characterised and it has been shown that disruption of this specific interaction can influence the cellular role of PimPol. In fact, mutation of the RBM-A motif in PrimPol, which is the main mediator of this interaction *in vivo,* leads to disruption of PrimPol's ability to restart stalled replication forks following UV-damage (Guilliam *et al.,* under review). Together, these findings suggest that RPA recruits PrimPol to the stalled replication fork by means of this interaction.

In archaea, immunoprecipitation experiments strongly supported the idea that RPA41 from *P. furiosus* interacts with the small catalytic subunit of primase (Pfu41 or PriS) (Bocquier *et al*., 2001; Komori and Ishino, 2001). Therefore, to test whether RPA is responsible for recruiting Afu-PriS1/L to the stalled replication fork, next, we tried to detect the potential interaction between RPA and DNA primase in *A.fulgidus*.

5.7.1. Studying PriS1/L's interaction with RPA by GST pull-down assays

To assess the interaction between PriS1/L and RPA, GST pull-down experiments were carried out. In theory, for GST pull-down, a GST-fusion bait protein, expressed in *E.coli* is required. The GST fusion protein can bind to glutathione (GSH)-coupled beads, which allows us to purify its interacting proteins (prey). In practice, prey protein is either cell lysates or recombinant purified proteins (Figure 5.11).

PGEX-6P-1:Afu-PriS1 and pGEX-6P-1:Afu-PriL expression constructs were provided by my colleague Dr. Stanislaw Jozwiakowski. Both constructs were transformed with BL21 strain *E.coli.* After incubation at 37°C, growth was induced with 1mM IPTG and then incubated at 20°C overnight. The cell lysate was subjected to sonication and then cleared by centrifugation. 30µL of GST magnetic beads were washed three times with buffer containing 40 mM Tris pH 7.5, 100 mM NaCl and 30 mM imidazole. The purified RPA-780/382 protein was incubated with 1ml of either PriS1 or PriL cell lysates in buffer containing 100

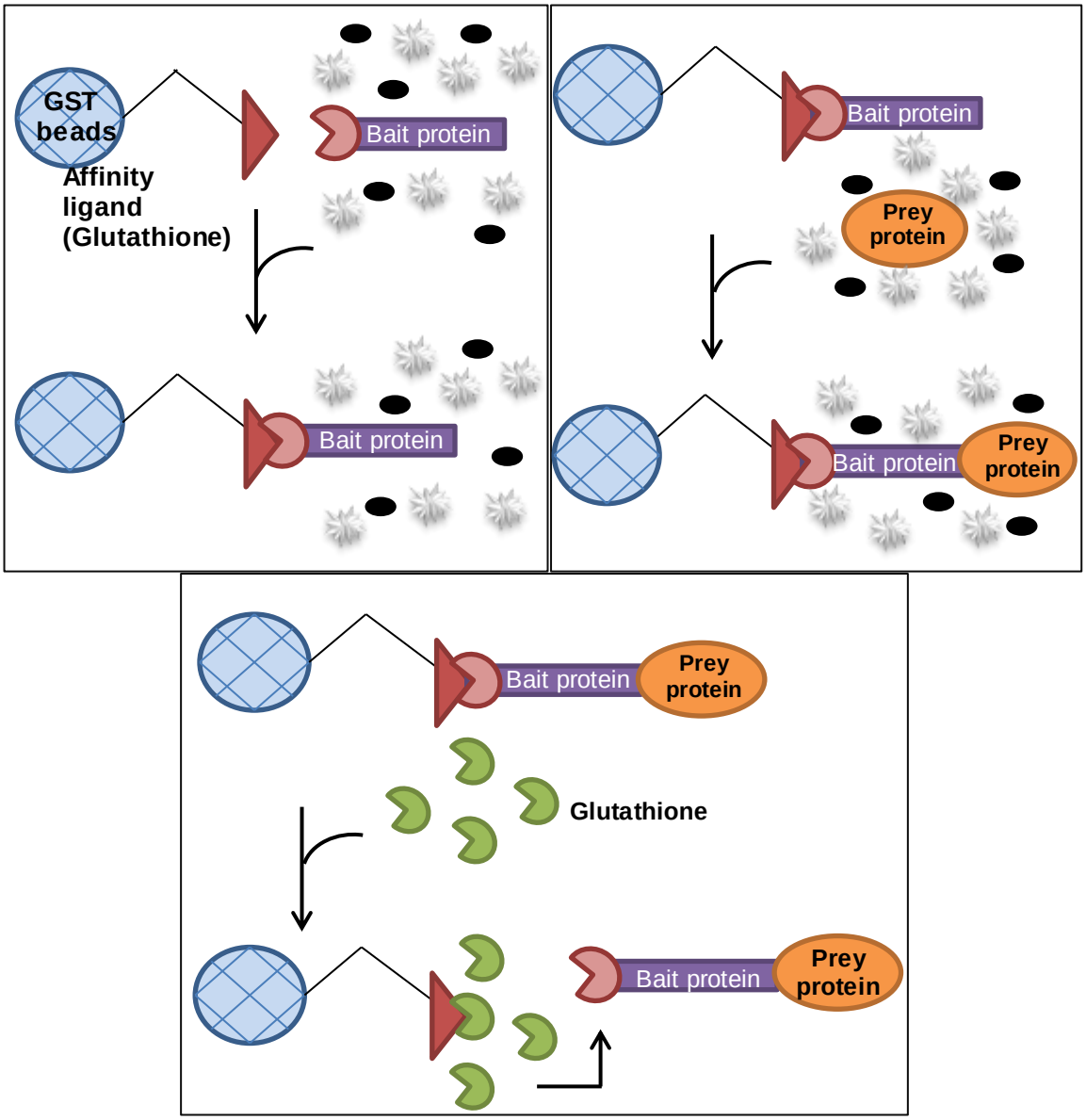


Figure 5.11. Schematic illustration of the principles of GST pull-down assays

The procedure starts with expression of GST fusion bait proteins in *E.coli* . Next, the GST fusion protein from cell lysate or other biological samples is captured by glutathione-agarose beads. The bait protein is incubated with the prey protein and subsequently, the interacting complex is eluted using reduced glutathione. Glutathione binds to the GST-agarose beads and displaces the interacting complex**.**

mM NaCl. The protein mix was then added to the GST beads and incubated for 1 hour at 4°C. After incubation, the mixed proteins were washed three times. 25 µL of elution buffer (40 mM Tris pH 7.5, 100 mM NaCl, 20mM glutathione, 0.5 mM TCEP) was added to the protein mix and incubated for 15 minutes on a rotating platform. Using a magnetic separator, eluted proteins were collected and then analysed by 12% SDS-PAGE gel (Figure 5.12). As expected, when pulling down the RPA complex (prey) with PriL (bait) protein, a band with an apparent molecular mass of ~ 68 kDa, corresponding to the fusion protein (PriL-GST), was observed in the eluate obtained from immobilized GST fusion protein beads (Figure 5.12A). We also observed a faint band with an apparent molecular mass of ~ 35 kDa consistent with RPA-780 prey protein. However, this band was also observed in the control sample, where only RPA complex (no PriL) was incubated with the beads, suggesting a false-positive result (Figure 5.12A). In addition, no band corresponding to RPA-382 protein was observed in the eluate obtained from the immobilized fusion protein sample (Figure 5.12A). On the other hand, when RPA was pulled down with PriS1, a band with an apparent molecular mass of~ 67 KDa corresponding to the fusion protein (PriS1-GST) in the eluate was obtained. However, no band corresponding to RPA780 and RPA-382 was detected. Although, the GST pull-down assay suggested there was no interaction between both PriS1 or PriL subunits and the RPA complex, further experiments were attempted to examine this possible interaction between PriS1/L and RPA-780/382 in *A.fulgidus*.

5.7.2. Identifying protein-protein interactions using His pull-down assays

In addition to GST pull-down experiments, to assess the interaction between PriS1/L and RPA, we also performed *in vitro* His tag pull-down assays. In general, pull-down assays are similar to immunoprecipitation but, instead of an antibody, a bait protein with affinity for the prey protein is used. To do these pull-downs, previously purified proteins (Afu-PriS1/L and Afu-RPA-780/382) were used.

As both purified proteins possessed a histidine (His) tag at the N-terminal, we first tried to cleave the His-tag from Afu-PriS1/L recombinant protein using

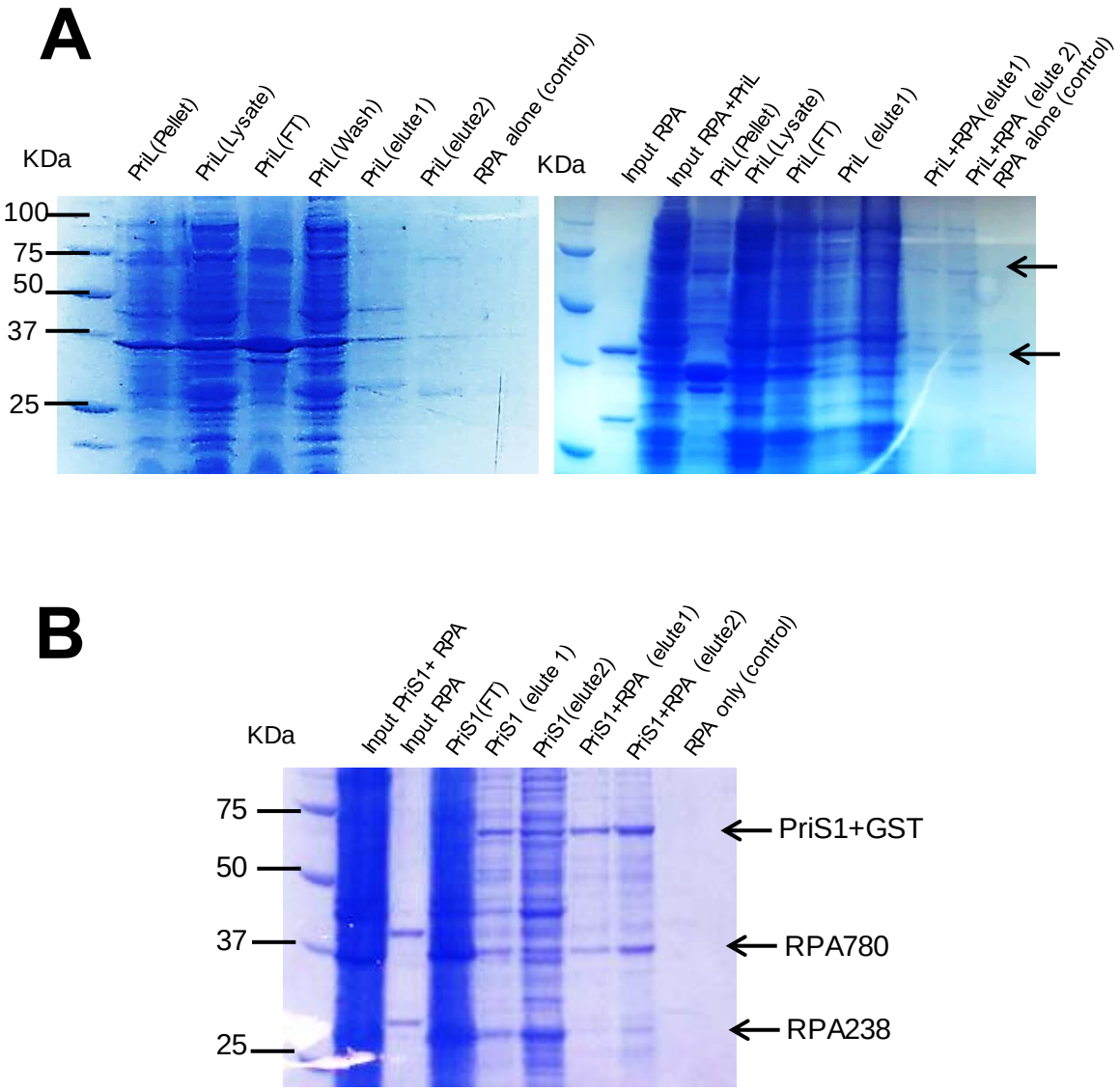


Figure 5.12. Detecting protein-protein interactions using GST-Tag pull-down assays

E.coli BL21 strain was transformed with either pGEX-6P-1:Afu-PriL **(A)** or pGEX-6P-1:Afu-PriS1 **(B)** expression constructs. Cultures were incubated at 37°C for 3 hours and then induced by 1 mM IPTG and incubated at 20°C overnight. 30µL of GST magnetic beads was washed with 100 mM NaCl and 30 mM imidazole. 1 ml of each cell lysate was incubated with RPA780/382 purified protein. The mix protein was then added to the GST magnetic beads and incubated for 1 hour at 4°C. After washing the beads the GST fusion protein was eluted with 25µL of elution buffer containing 100 mM NaCl and 20mM reduced glutathione and subsequently analysed by SDS-PAGE gel. **(A)** in the elute obtained from immobilized PriL-GST protein, two bands with apparent molecular masses of~ 68 KDa and~ 35 KDa corresponding to the bait protein (PriL-GST) and the prey protein (RPA780) respectively are shown. In control sample the band with molecular mass of~35 is also indicated. **(B)** In the elute obtained from immobilized PriS1-GST protein beads, one band, with an apparent molecular mass of~ 67 KDa corresponding to the bait protein (PriS1-GST) is shown, however, no band corresponding to RPA780 and RPA382 proteins can be detected.

thrombin as it contains a thrombin cleavage site after the tag. The thrombin concentration used for this cleavage was 1 unit per mg purified protein. Following cleavage, Afu-PriS1/L was separated from the cleaved His-tag by HisTrap and HiTrap Benzamidine columns connected in series. Cleaved PriS1/L was analysed by SDS-PAGE gel (Section 2.2.10) (Figure 5.13A). 400 nM of the His-tagged protein (RPA-780/382) and 400 nM of the non His-tagged protein (PriS1/L) were mixed with binding buffer (40 mM Tris pH 7.5, 100 mM NaCl and 30 mM imidazole) in a total reaction volume of 50 µL. After 30 minutes incubation on ice, the protein mixture was added to pre-equilibrated Ni-NTA Agarose resin and mixed on a spinning wheel at 4ºC for 1 hour. The supernatant was removed after spinning at 2,000 x g for 2 min. The resin was washed three times with 100 µL of wash buffer (40 mM Tris pH 7.5, 100 mM NaCl and 30 mM imidazole) and then 25 µL of elution buffer (40 mM Tris pH 7.5, 250 mM imidazole) was added to the protein mix followed by 15 minutes incubation on a rotating platform. After 2 minutes spinning at 2,000 x g, eluted protein was collected and analysed by SDS-PAGE gel. His-tag pull-down assay exhibited two bands with apparent molecular masses of ~ 35 and ~ 41/42 kDa corresponding to RPA-780 and PriS1/L, respectively. However, no band corresponding to was detected (Figure 5.13B). Subsequently, western blotting analysis using anti His-tag antibody was carried out. The protein samples were blotted on a membrane to confirm the interaction between RP780 and PriS1/L. However, only one band corresponding to RPA-780 was observed on the blot. No band corresponding to PriS1/L could be detected (Figure 5.13C). Consistent with the GST pull-down experiment (section 5.7.1), these data suggest that there is no physical interaction between RPA-780/382 and PriS1/L in *A.fulgidus.* However, these results are not consistent with previous studies by Ishino and colleagues, which reported that PriS subunit from *P. furiosus* interacts with RPA (Bocquier *et al*., 2001; Kayoko and Ishino, 2001). This interaction might facilitate the recruitment of the primase to stalled replication forks. Since our data do not fully rule out the possibility of the interaction between Afu-PriS1/L and Afu-RPA-780/382, further strategies such as mass spectrometry are required to test possible interactions between PriS1/L and RPA in *A.fulgidus*.

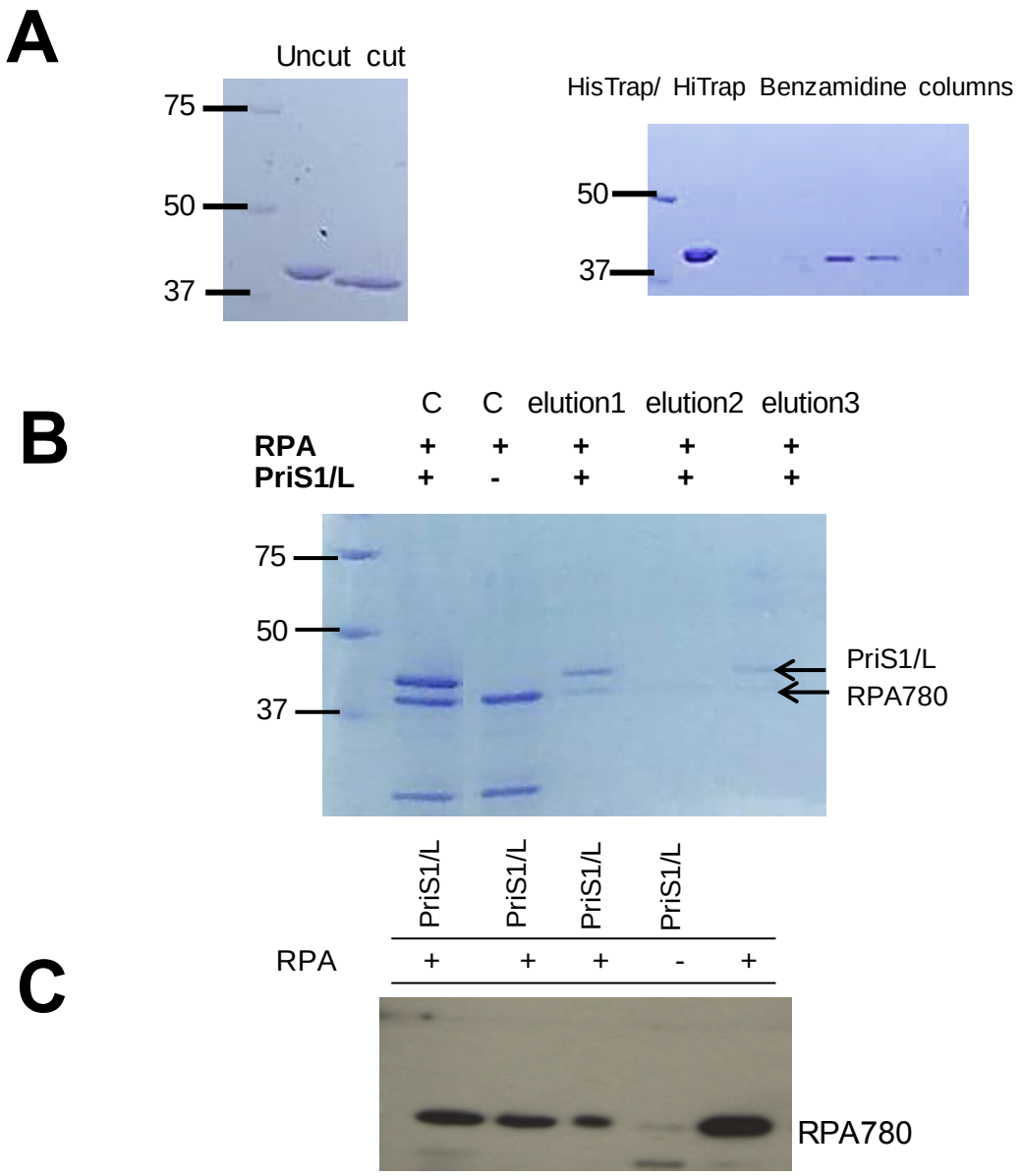


Figure 5.13. Detecting protein-protein interactions using His-Tag pull-down assays

(A) 1 unit thrombin per mg of Afu-PriS1/L protein was used to cleave the his tag at the N-terminal of protein. **(B)** 400 nM of RPA780/382 and 400 nM of PriS1/L were mixed with binding buffer containing 100 mM NaCl and 30 mM imidazole. The protein mixture was added to pre-equilibrated Ni-NTA Agarose resin and mixed for 1 hour. The resin was washed with 100 µL of wash buffer. 25 µL of elution buffer (40 mM Tris pH 7.5, 250 mM imidazole) was added to the protein mix and incubated for 15 minutes . Eluted protein was collected and analysed by 12% SDS PAGE. SDS-PAGE gel indicated presence of two bands with apparent molecular masses of ~35 and ~41/42 kDa corresponding to RPA-780 and PriS1/L respectively. No band corresponding to RPA-382 was detected. Letter C denotes no GST beads control samples. **(C)** Western blotting analysis using anti his-tag antibody was carried out. The protein samples were blotted on a membrane to confirm the interaction between RP780 and PriS1/L. Only one band corresponding to RPA780 was detected on the blot. No band corresponding to PriS1/L could be observed

5.8. Summary and Discussion

Archaeal PriS1/L complex is a DNA primase-polymerase enzyme that belongs to the AEP super-family (Guilliam *et al*, 2015). An additional role is discovered for PriS1/L in DNA damage tolerance through performing translesion DNA synthesis during DNA replication (Chapter 4). To accomplish this role, PriS1/L must be efficiently recruited to the ssDNA downstream of the lesion. Recent studies explored that human PrimPol, another member of the AEP super-family, which also performs TLS activity, interacts with SSBs. This suggested that PrimPol is recruited to the stalled replication fork by RPA. In addition, this interaction proposed a unique regulatory mechanism discrete from mechanisms used by other TLS polymerases (Guilliam *et al*., 2015; Lehmann *et al*., 2007; Wan *et al*., 2013). Subsequent studies illustrated the impact of SSBs on enzymatic activities of PrimPol. It was shown that, SSBs act to remarkably supress both primase and polymerase activities of PrimPol (Guilliam *et al*., 2015). This was in common with previous findings regarding the inhibitory role of RPA on primase activity of the Polα complex (Collins and Kelly, 1991). These studies implicated the important role of SSBs in regulating the non-specific primase and polymerase activities during replication. Since archaeal PriS1/L displays a relatively low fidelity and could be error-prone on undamaged templates, similar to other TLS polymerases strict regulation is likely required during its participation in DNA replication. Otherwise, its unscheduled or deregulated activities would synthesise long tracts of DNA in an error-prone fashion, which could be very mutagenic. Our results in this chapter, suggest a possible regulatory mechanism by RPA that limits the involvement of PriS1/L in DNA replication.

We initially showed that two RPA orthologous from *A.fulgidus* (RPA-780 and RPA-382) possess high affinity for ssDNA. Next, we demonstrated that the RPA-780/382 complex isolated from *A.fulgidus* has no regulatory effect on primer synthesis by Afu-PriS1/L, which was in contrast with the detected inhibitory role of RPA on *de novo* primer synthesis by Polα complex and PrimPol. As DNA primases are able to bind to and prime on ssDNA with high affinity, deregulated primase activity could lead to unscheduled priming whenever ssDNA is present and since the generated primers are subsequently extended by other polymerases, this deregulation could be detrimental to cells. Although our data

did not show any regulatory effect by RPA on Afu-PriS1/L pimase activity, this does not preclude the possibility that primase activity of Afu-PriS1/L is regulated by other components of replisome.

RPA stimulates the polymerase activity of Polα and Polδ (Kenny *et al*., 1989; Tsurimoto and Stillman, 1989). Stimulation of Polγ's polymerase activity by mtSSB was also demonstrated (Oliveira and Kaguni, 2010). On the other hand, both RPA and mtSSB inhibit the polymerase activity of PrimPol (Guilliam *et al*., 2015). In this chapter, it was shown that the Afu-RPA-780/382 complex significantly limits the primer extension activity of Afu-PriS1/L. In addition, Afu-PriS1/L's polymerase activity is considerably supressed by the non-interacting T4 SSB. This might suggest that the inhibitory role of RPA on polymerase activity of PriS1/L is not due to the physical interaction between these two proteins. Interestingly, in *E.coli*, the primer extension by Pol II and Pol IV TLS polymerases is also inhibited by SSB (Indiani *et al*., 2013). It is assumed that the inhibitory role of RPA on PriS1/L's polymerase activity is partly due to the inhibition of rebinding of the PriS1/L to ssDNA. In Chapter 4, we provided experimental evidence, which suggested that archaeal PriS1/L is a relative error-prone TLS DNA polymerase. Therefore, it can be concluded that, restriction of PriS1/L's DNA synthesis by RPA acts as a mechanism by which the contribution of PriS1/L in DNA replication is regulated to prevent its mutagenic effects, which could threaten genome stability.

We also demonstrated that archaeal replicative polymerases (PolB and PolD), in contrast to replicative primases, are able to displace both RPA and T4 SSB and extend most of the primers. Interestingly, RPA stimulates the polymerase activity of Polα and Polδ (Kenny *et al*., 1989; Tsurimoto and stillman, 1989). A subsequent study revealed thtat Polγ's polymerase activity is also stimulated by mtSSB (Oliveira and Kaguni, 2010). Stimulation of the polymerase activity of replicative polymerases by RPA implies the ability of replicative polymerases to readily displace RPA. Therefore, these enzymes continue DNA synthesis on RPA-coated ssDNA until they encounter a DNA lesion. In fact, the orientation of the interaction of replicative polymerases with RPA might explain this ability. It is believed that binding of RPA to ssDNA occurs with a defined polarity (Fan and Pavletich, 2012; Kolpashchikov *et al*., 2001). In eukaryotes, RPA70 binds to ssDNA in a tandem fashion. Initially RPA70 binds ssDNA through the DNA-

binding domain A (DBD-A) and DBD-B oligonucleotide binding (OB) folds, which forms an 8-nt interaction mode. The binding complex subsequently stretched to 20-30 NTs through the binding of DBD-C and DBD-D, which takes place in a 5'-3' direction on the DNA template (Brosey *et al*., 2013). Given that replicative polymerases bind RPA from the 3' end, they first encounter the weakly bound DBD-D and DBD-C domains. As the result, the 20-30-nt binding complex is shifted to the 8-nt mode, which is weaker in terms of binding. Therefore allows the displacement of RPA by replicative polymerases.

In general, replicative polymerases have high processivity. These enzymes are able to incorporate thousands of nucleotides without dissociation from the DNA template. However, in contrast to replicative polymerases and similar to TLS polymerases, AEP primases have low processivity. It has been reported that PrimPol shares the characteristic of low processivity with canonical TLS polymerases. This enzyme can incorporate only 1-4 nucleotides per binding event (Keen *et al*., 2004). Even though the processivity of archaeal priS/L has not been tested yet, it is highly likely that this enzyme, similar to PrimPol, is a low processive polymerase. This might explain the different capabilities of replicative polymerases and replicative primases in displacing RPA in archaea.

The possible different RPA regulatory mechanisms that influence archaeal replicative polymerases and primases are summarized in Figure 5.14.

PriS1/L has not yet been shown to perform repriming downstream of lesions. But since PriS1/L is a proficient primase, the possible ability of this enzyme to reprime replication post lesion cannot be excluded. Replication fork uncoupling as the result of leading-strand damage leads to the generation of ssDNA stretches on leading strand. RPA binds to ssDNA as replication progresses. PriS1/L is able to bind to free ssDNA. Recruitment of PriS1/L to ssDNA might be facilitated by RPA. Following binding to ssDNA, PriS1/L might perform repriming downstream of lesion until it encounters RPA which can inhibit primer extension activity of PriS1/L. Subsequently, the primer is handed-off to the PolB replicative polymerase to continue DNA replication. However, clearly further studies are required to explore this possible mechanism.

Despite the discovery of a possible RPA regulatory role on polymerase activity of Afu-PriS1/L, we were unable to demonstrate a direct interaction between RPA and PriS1/L in this study. Given that the interaction between RPA and DNA primase has been identified in some other species of archaea and also in eukaryotic cells, our data does not categorically rule out the possibility of an interaction between RPA and PriS1/L. Therefore, future interaction studies with RPA are required to definitively show whether a *bona fide* interaction with PriS1/L exists or not.

Notably, another possible regulatory mechanism can be achieved through tethering the PriS/L to the CMG complex. Previous studies indicated that in eukaryotic cells, a complex of GINS and Ctf4 trimer is required to couple Polα to MCM2-7 (Gambus *et al*., 2009). It was indicated that coupling of replicative polymerases to MCM plays a key role in regulation of DNA replication in eukaryotes. Interestingly, in *Sulfolobus*, GINS directly interacts with primase (Marinsek *et al*., 2006). This can suggest that in archaea GINS acts to tether the primase to the CMG complex. Moreover, previously, it was shown that archaeal PriS subunit has a GINS domain fused to its C-terminus. This relationship might play an important role for primase and GINS interaction (Swiatek and MacNeill, 2010). Thus, the possible interaction between GINS and PriS/L could allow the primase to prime following DNA damage. Future studies are required to test this hypothesis.

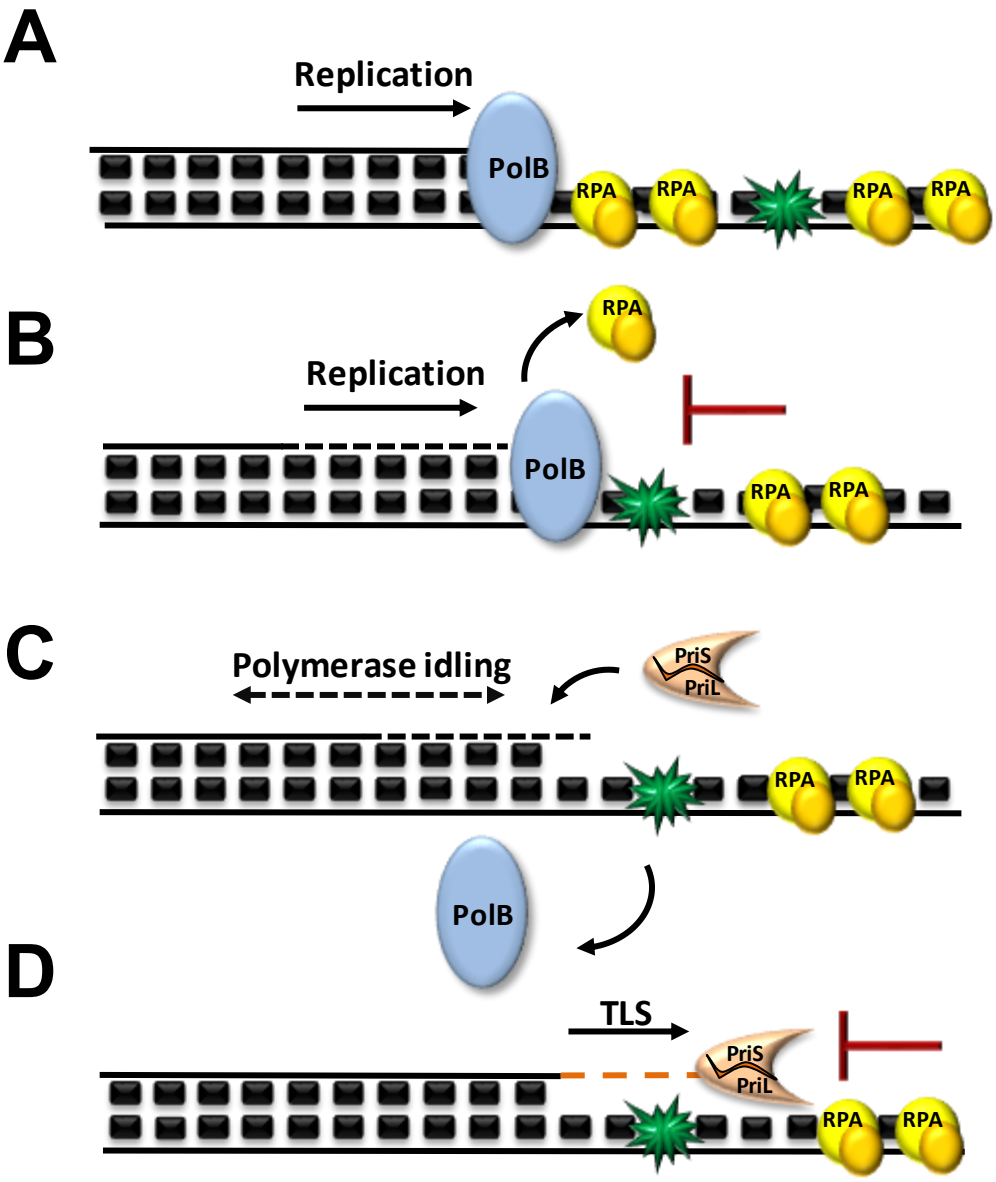


Figure 5.14. Regulation of PriS1/L DNA synthesis by RPA during DNA replication

(A) Replicative polymerases can synthesise DNA on RPA-coated ssDNA by displacing the bound RPA. **(B)** When the intolerant replicative polymerase encounters a DNA lesion, it stalls at the site of damage which leads to polymerase idling and displacing surrounding RPAs to generate an unbound ssDNA interface. **(C)** PriS1/L is then recruited to the free ssDNA, either upstream or downstream, of RPA. **(D)** PriS1/L performs TLS activity to bypass DNA lesion until is inhibited by downstream RPA. Replication is then continued with the replicative polymerase.

Chapter 6

Towards the reconstitution of archaeal CMG complex *in vitro*

6.1. Introduction

The CMG (Cdc45, MCM2-7, GINS) complex is pre-assembled at origins prior to the initiation of eukaryotic replication. MCM2-7 helicase activity requires efficient coordination of Cdc45 and GINS accessory proteins (Costa *et al*., 2011; Moyer *et al*., 2006). Thus, the assembly of CMG complex plays an important role in the initiation and progression of DNA replication. A growing body of evidence has demonstrated that archaeal proteins involved in the processes of DNA replication, transcription, and translation are much more closely related to their eukaryotic counterparts than bacteria, suggesting that eukarya and archaea share a common ancestor (Grabowski and Kelman,2003;Kelman and Kelman,2003). In some archaea, the active helicase is a simple homo-hexameric MCM (Kelman and Kelman, 2014). Unlike eukaryotes, most archaeal species possess a single MCM orthologous (Sakakibara *et al*., 2009). The archaeal homo-hexameric MCM has been characterised in several species, such as, *Archaeoglobus fulgidus*, *Thermococcus kodakarensis*, *Pyrococcus furiosus*, *Sulfolobus solfataricus*. In addition, similar evolutionary patterns, as with the MCM, were observed for archaeal GINS proteins. Most species of archaea possess a single GINS complex, containing two copies of a subunit related to eukaryotic Psf1 and Sld5, forming a homotetramer (Gins15). However, some archaea also possess a GINS complex containing two copies of a subunit related to eukaryotic Psf2 and Psf3 (Gins23) (Oyama *et al*., 2011). Notably, there is a direct physical interaction between archaeal GINS and MCM helicase. Interestingly, recent studies proposed the presence of archaeal CdC45 homologs in several species. Bell and colleagues were able to co-purify the *Sulfolobus* GINS complex with a protein, which showed homology with the DNA binding domain of a bacterial exonuclease, RecJ (Marinsek *et al*., 2006). Therefore, they initially named the protein as RecJdbh. Following recent structural studies which demonstrated that RecJ and eukaryotic CdC45 are related to each other. RecJdbh was renamed to CdC45 (Xu et al., 2016). Together, these findings confirm that the CMG complex is a conserved component in all archaeal and eukaryotic DNA replication machineries.

Interaction studies, using *S.solfataricus* as a model organism, detected direct interactions between Gin23 and MCM and also between Gins23 and the

heterodimeric core primase (PriS/L) (Marinsek *et al*., 2006). Since no direct interaction between MCM and primase was detected, it has been assumed that archaeal GINS acts to link the primase complex to the MCM, which in turn leads to coupling of helicase activity with priming on leading and lagging strands. This proposed that in archaea, GINS might act to tether the primase to the CMG complex. This possible interaction between GINS and PriS/L could allow the primase to either initiate priming at the initiation replication site or downstream of sites of damage. Therefore, assembly of the CMG complex might be required for the recruitment of PriS/L to the DNA. In this chapter, we describe our initial steps towards the long-term goal of reconstituting the archaeal CMG complex *in vitro*. An additional aim discussed in this chapter is shedding light on the role of archaeal replicative primase in replication-specific TLS. We also biochemically characterised the *in vitro* activities of MCM and GINS proteins from *A.fulgidus*, and examined if Afu-GINS was capable of modulating the helicase activity of Afu-MCM.

6.2. Cloning of the archaeal CMG complex genes into expression vectors

In order to reconstitute the archaeal CMG complex *in vitro* and to characterise the biochemical properties of the CdC45, GINS, and MCM protein complexes, first the genes encoding CdC45, GINS, and MCM from *Archaeoglobus fulgidus* were cloned into appropriate expression vectors (Table 2.3). *A. fulgidus* encodes a single homologue of eukaryotic MCM2-7 (Grainge *et al*., 2003). The ORF corresponding to Afu-MCM (AF_RS02630) was amplified from *A.fulgidus* genomic DNA (using primers in Table 2.1). *A.fulgidus* possesses a single GINS protein with sequence homology to those of eukaryotic Sld5/Psf1 (Yoshimochi *et al*., 2008). The ORF corresponding to Afu-GINS (AF1322) was amplified from *A.fulgidus* genomic DNA (using primers in Table 2.1.) (Figure 6.1). Recently, two RecJ homologs have been identified in *A.fulgidus* (Makarova *et al*., 2012). Similar to two other genes, the ORFs corresponding to Afu-RecJ-like (Cdc45) (Af0699, Af0698) were amplified from *A.fulgidus* genomic DNA and Af0699 PCR-Plasmid, respectively (Table 2.3) (using primers in Table 2.1).

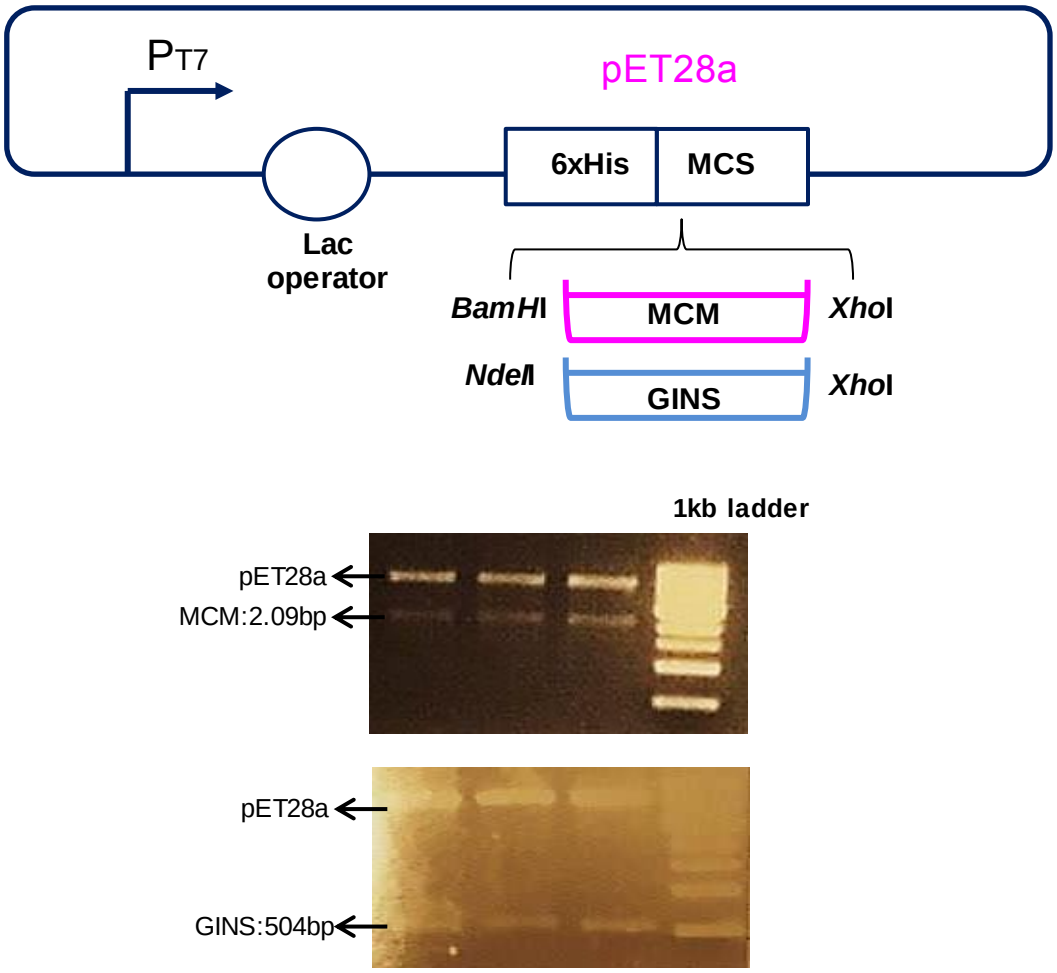


Figure 6.1. Cloning of Afu-MCM and GINS into pET28a

The open reading frames corresponding to MCM and GINS were amplified from *A. fulgidus* genomic DNA (Table 2.3.) introducing the applicable restriction sites to allow insertion into the multiple cloning site (MCS) of the Pet28a expression vectors. A 6-histidine tag downstream of a T7 promoter (P_{T7}) was in frame with cloned MCM and GINS. The PCR products were combined with 10x DNA loading dye and run on 1% agarose gels containing ethidium bromide.

Subsequently, the DNA fragments encoding MCM and GINS were cloned individually into the *E.coli* expression vector pET28a, generating pET28a:Afu-MCM and pET28a:Afu-GINS expression constructs, respectively. In parallel ,the Af0699 amplified gene was cloned into pETduet-1 expression vector to generate pETduet-1:Afu-RecJ699, and Af0698 amplified DNA was subcloned into the pETduet-1:Afu-RecJ699 construct to generate pETduet-1:Afu-RecJ699/698 expression construct (Figure 6.2).

6.3. Expression and purification of Afu-MCM

The pET28a:Afu-MCM expression construct was transformed into Rosetta *E.coli* and then incubated at 37°C for 3 hours. Following addition of 1 mM IPTG at 25°C overnight, the soluble cell lysate was prepared and loaded onto a Nickel-NTA agarose affinity chromatography column. MCM was eluted from the column by addition of 300 mM imidazole. Eluted MCM was resolved on an SDS-polyacrylamide gel to confirm the correct size. The result was in agreement with its predicted molecular mass. Afu-MCM has a predicted molecular weight of ~ 78 kDa (Figure 6.3). Since the MCM was co-eluted with a considerable amount of *E.coli* contaminants, anion exchange chromatography (Mono Q) was employed to remove the contaminants. The 80% B eluted peak fraction from nickel affinity column was subjected to MonoQ column and after a wash step, the bound MCM was eluted with 1M NaCl. A protein with an apparent molecular mass of ~ 78 kDa corresponding to Afu-MCM was observed on the SDS-PAGE (Figure 6.3). For the final step of purification, size-exclusion chromatography (Gel filtration) was performed. A S75 gel-filtration column was pre-equilibrated with 300 mM NaCl. Concentrated protein samples were loaded onto a 5 mL loop. Finally, fractions were collected following 100 mL of flow-through and to confirm the correct size resolved on a SDS-PAGE gel and determined by A_{280} level (Figure 6.3). The MCM eluted as a single peak and at a size consistent with MCM forming a hexameric complex.

6.4. Afu-MCM DNA binding activity

The DNA binding activity of archaeal MCM proteins requires the presence of two distinct types of structural motifs, a β-hairpin motif and a zinc finger motif

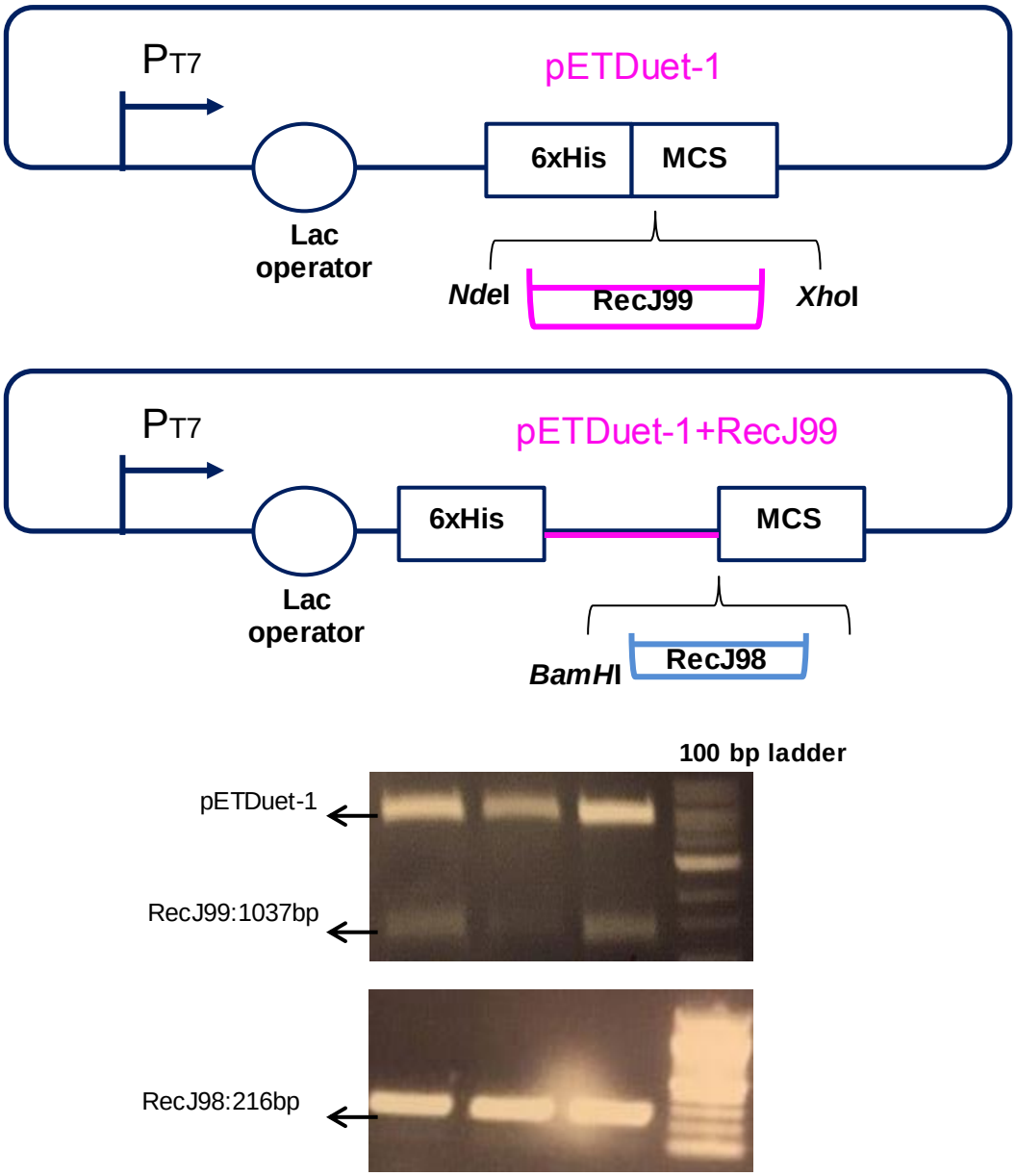


Figure 6.2. Cloning of Afu-RecJ99 and RecJ98 into pETDuet-1

The open reading frames corresponding to RecJ99 and RecJ98 were amplified from *A.fulgidus* genomic DNA and Af0699 PCR-Plasmid respectively (Table 2.3.) introducing the applicable restriction sites to allow insertion into the multiple cloning site (MCS) of the pETDuet-1 expression vector. A 6-histidine tag downstream of a T7 promoter (P_{T7}) was in frame with cloned RecJ99. The PCR products were combined with 10x DNA loading dye and run on 1% agarose gels containing ethidium bromide.

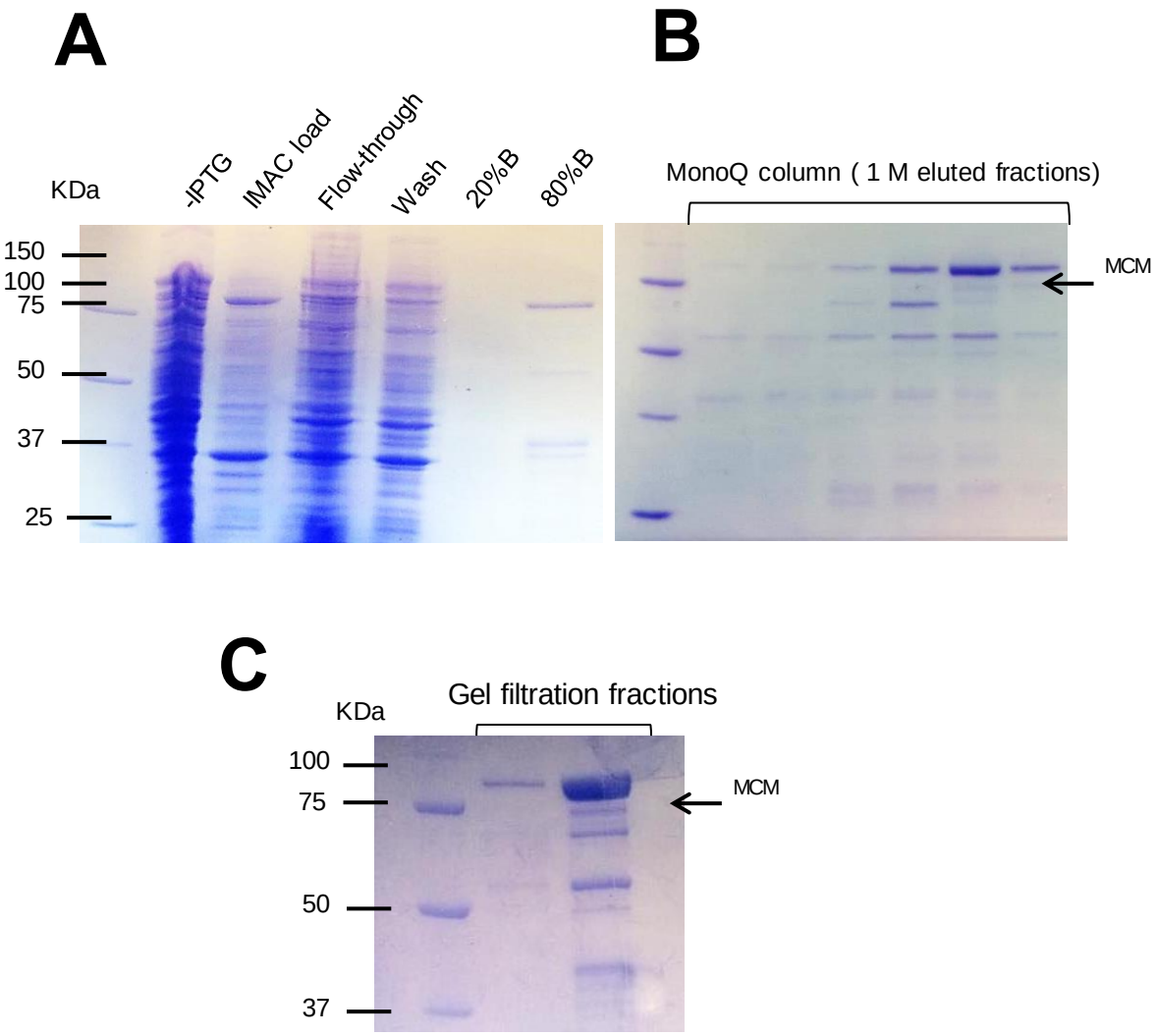


Figure 6.3. Purification of Afu-MCM using chromatography

Purification of Afu-MCM (78KDa). **(A)** A three litre culture of *E.coli* Rosetta transformed with pET28a: MCM expression construct was grown at 37°C for 3 hours and subsequently induced with 1 mM IPTG and incubated overnight. Following cell lysis, soluble cell lysate was loaded ontoNickle column. Bound Afu- MCM was washed and eluted with 300 mM imidazole. **(B)** The 80% B eluted peak fraction from nickel affinity column was subjected to MonoQ column and after wash step, the bound MCM was eluted with 1M NaCl. **(C)** In the last step of purification, protein containing fractions were subjected to S75 gel-filtration column which was pre-equilibrated with 300 mM NaCl. Finally, fractions were collected following 100 mL of flow-through and analysed by SDS-PAGE gel.

situated at the N-terminus of the protein and a β-hairpin motif situated at the AAA+ catalytic domain (Sakakibara *et al*., 2009). It was shown that, mutation of the zinc finger motif in *M. thermautotrophicus* and *Aeropyrum pernix* resulted in reduction of both ssDNA and dsDNA binding activities of these MCM proteins (Atanassova and Grainge, 2008;Kasiviswanathan *et al*.,2006). Interestingly, while mutation of the β-hairpin motif of the *M. thermautotrophicus* MCM strongly inhibited its DNA binding activity (Fletcher *et al*., 2003), the equivalent mutation in the *S. solfataricus* MCM showed no effect on DNA binding activity of this enzyme (McGeoch *et al*., 2005). Thus, these results suggested that differences exist in the importance of the β-hairpin motif on DNA binding activity of archaeal MCM proteins. One interesting feature of archaeal MCM proteins is the different binding affinities of these proteins for various DNA substrates. For instance, MCM proteins from *A.fulgidus*, *A. pernix* and *S. solfataricus* show higher binding affinity for bubble and forked DNA substrates over ssDNA or dsDNA (Atanassova and Grainge, 2008; Grainge *et al*., 2003; Rothenberg *et al*., 2007). These findings suggested the potential replicative helicase activity of archaeal MCM proteins at the replication fork.

To examine DNA binding activity of the purified MCM from *A.fulgidus*, standard electrophoretic mobility shift assays (EMSAs) was utilized. 45-mer labelled ssDNA and dsDNA were incubated with Mg^{2+} and increasing concentrations of MCM. ATP was not used in this DNA binding assay as previous studies proposed that ATP has no effect on the DNA binding activity of MCM from several archaeal species (Fletcher *et al*., 2003; Kelman and Hurwits, 2003; Grainge *et al*.,2003). In the EMSAs, it was observed that both ssDNA and dsDNA were bound by Afu-MCM and that the extent of binding was dependent on the concentration of protein (Figure 6.4). The presence of super-shifted MCM-DNA complexes suggested that the DNA was long enough for efficient binding. The binding of MCM to both ssDNA and dsDNA was apparent at a concentration of 0.1 µM (MCM) and complete shift of substrate exhibited at 0.2 µM. Although both ssDNA and dsDNA were shifted with approximately equal efficiency, quantifications of shifted MCM-DNA complexes showed little variation in binding efficiency between them (Fig 6.4). MCM exhibited slightly higher binding affinity for dsDNA over ssDNA. Since previous studies proposed higher binding affinity for forked DNA substrates by MCM isolated from some archaeal species (*A. fulgidus*, *A. pernix*

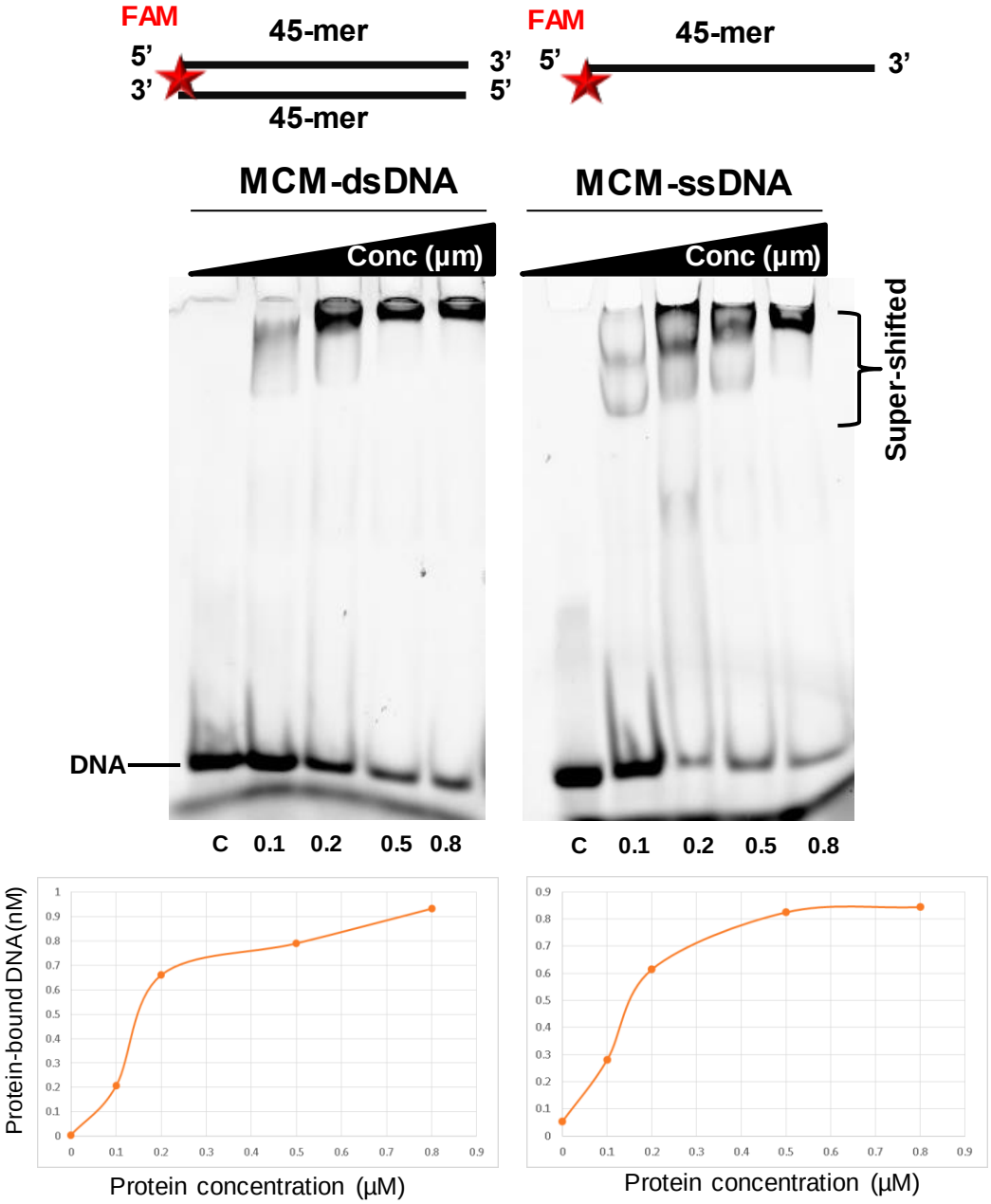


Figure 6.4. Afu-MCM binds both ssDNA and dsDNA

DNA binding gel of Afu-MCM with dsDNA and ssDNA. 45-mer 5' labelled ssDNA and dsDNA (Table 2.7) were incubated with 10 mM Mg^{2+},1 mM DDT and increasing concentration of MCM. The concentrations of MCM used in these experiments are shown below each gel. MCM binds to both templates with comparable affinities. In both cases, binding was detected at 0.1 µM of MCM concentration and super-shifted was appeared at 0.2 µM.

and *S. solfataricus)* (Atanassova and Grainge, 2008; Grainge *et al*., 2003; Rothenberg *et al*., 2007), we next examined the DNA binding activity of Afu-MCM using a 60-mer duplex DNA substrate with a 30-mer tail at both 3' and 5' ends producing a forked substrate. This EMSA experiment was carried out under the same conditions used for ssDNA and dsDNA. In this experiment MCM showed slower migration of the MCM-DNA complex compared to either the ssDNA or dsDNA shifted MCM species, which suggest that MCM has lower binding affinity for the forked substrate. (Figure 6.5). While the binding of MCM to the 45-mer ssDNA and dsDNA was apparent at a concentration of 0.1 µM (MCM) and a complete shift of substrate was observed at 0.2 µM, in the case of the forked DNA substrate, MCM binding was apparent at 0.2 µM and a complete shift of substrate at ~0.5-1 µM. Despite slower migration, efficient binding of MCM to the Y-shaped substrate may suggest a role for MCM as the replicative helicase at the replication fork. It is believed that binding of MCM to forked DNA is distinct. Binding takes place through encircling the 3'-tail of substrate by MCM, while 5'-tail interacts with the MCM surface (Rothenberg *et al*., 2007). This interaction might be important for substrate selectivity and MCM-forked DNA complex stability. In addition, this MCM and 5'-tail interaction might be required for DNA unwinding. Thus, it may explain why the MCM complex in both archaea and eukarya prefers unwinding the Y-shaped structures.

6.5. DNA helicase activity of Afu-MCM

Several studies demonstrated that the archaeal and eukaryotic replicative MCM complex share similar biochemical properties (Carpentieri *et al*., 2002; Kelman *et al*., 1999; Poplawski *et al*., 2001). To date, all biochemically characterised archaeal MCM proteins show an ATP-dependent 3'→5' helicase activity. Unlike eukaryotic MCM, some archaeal MCM proteins display helicase activity independent of accessory proteins *in vitro* (Duggin and Bell, 2006; Kelman and Kelman, 2003). Similar to eukaryotic MCM and *Escherichia coli* (DnaB), MCM proteins isolated from several species of archaea, including *M. thermautotrophicus* and *Thermoplasma acidophilum*, not only translocate along ssDNA, but also translocate along dsDNA (Haugland *et al*., 2006). Different models for the unwinding were suggested; a steric-exclusion model, a

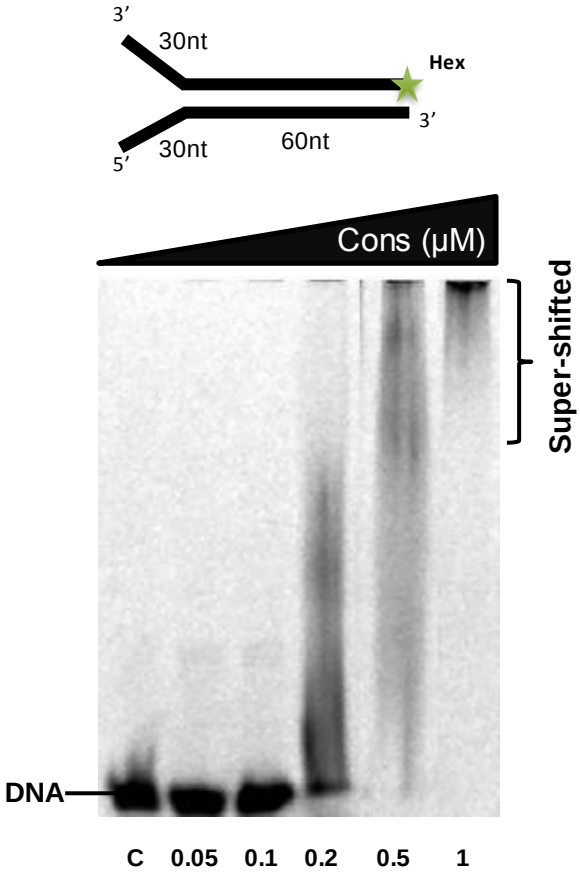


Figure 6.5. Binding of Afu-MCM to forked DNA substrate

DNA binding affinity of Afu-MCM to a forked DNA substrate was examined. The substrate used in the mobility shift assay is shown schematically. A 60-mer duplex DNA with 30-mer 3' and 5' tails (Table 2.7) was incubated with 10 mM Mg^{2+}, 1 mM DDT and increasing concentration of MCM. The concentrations of MCM used in these experiments are shown below the gel. Considerable DNA binding was observed by MCM in the presence of this substrate. Binding was apparent at 0.2 μM and a complete shift of substrate at 0.5 μM.

ploughshare model and a rotary pump model (Sakakibara *et al*., 2009). In the steric-exclusion model, the MCM encircles and moves only along ssDNA. A number of studies support this model in archaea. On the other hand, both the ploughshare and rotary pump models propose the ability of helicase to translocate along duplex DNA. Notably, in the rotary pump model, a large excess of MCM is required and since in archaea only six MCM subunits are present per replication fork (Matsunaga *et al*., 2001), this model might not be applicable to archaea.

In order to test the helicase activity of Afu-MCM, we utilized various Y-shaped substrates containing a 5' end labelled 60-mer duplex region with 30-mer 3' and 5' tails (Table 2.7). In substrate 1, both 3' and 5' tails were annealed to a complementary oligonucleotide, while in substrates 2 and 3, either the 5' tail or 3' tail was annealed with the complementary oligonucleotide. In substrate 4, none of the tails were annealed to an oligonucleotide. This experiment was performed to investigate requirements for forked DNA substrates. 500 nM of MCM was incubated with 200 nM of each substrate and 20 mM $MgCl_2$ at 50°C for 5 minutes (See methods 2.3.10). different range of ATP was then added to reactions. After 10 minutes incubation, reactions were terminated using stop buffer (See methods 2.3.10). Controls consisted of reactions lacking either ATP or MCM, where no activity was observed. No helicase activity was detected with 5 mM ATP but strand displacement activity was displayed with higher concentrations of ATP. Strand displacement starts with 10 mM of ATP, and at 25 mM ATP the most efficient helicase activity exhibited (Figure 6.6). Our data suggested that, among substrates used in this assay, the most efficient helicase activity by MCM was observed in the presence of the Y-shaped substrate with a 3'-ssDNA tail (Figure 6.6). No unwinding was observed when the substrate bearing both 3' and 5' tails annealed to an oligonucleotide (Figure 6.6). This could be due to the absence of a ssDNA tail in this substrate. Initiation of duplex unwinding by DNA helicases requires a flanking ssDNA tail. In the case of MCM helicases, which possess a 3'→5' helicase activity, the presence of a 3'-ssDNA end is essential. It has been suggested that unwinding of a forked DNA is achieved through encircling the 3'-ssDNA tail and interacting with the 5' tail by MCM helicases (Rothenberget al., 2007). Together, our data confirmed that, similar to most characterised archaeal helicases, MCM isolated from *A.fulgidus* displays helicase activity without the

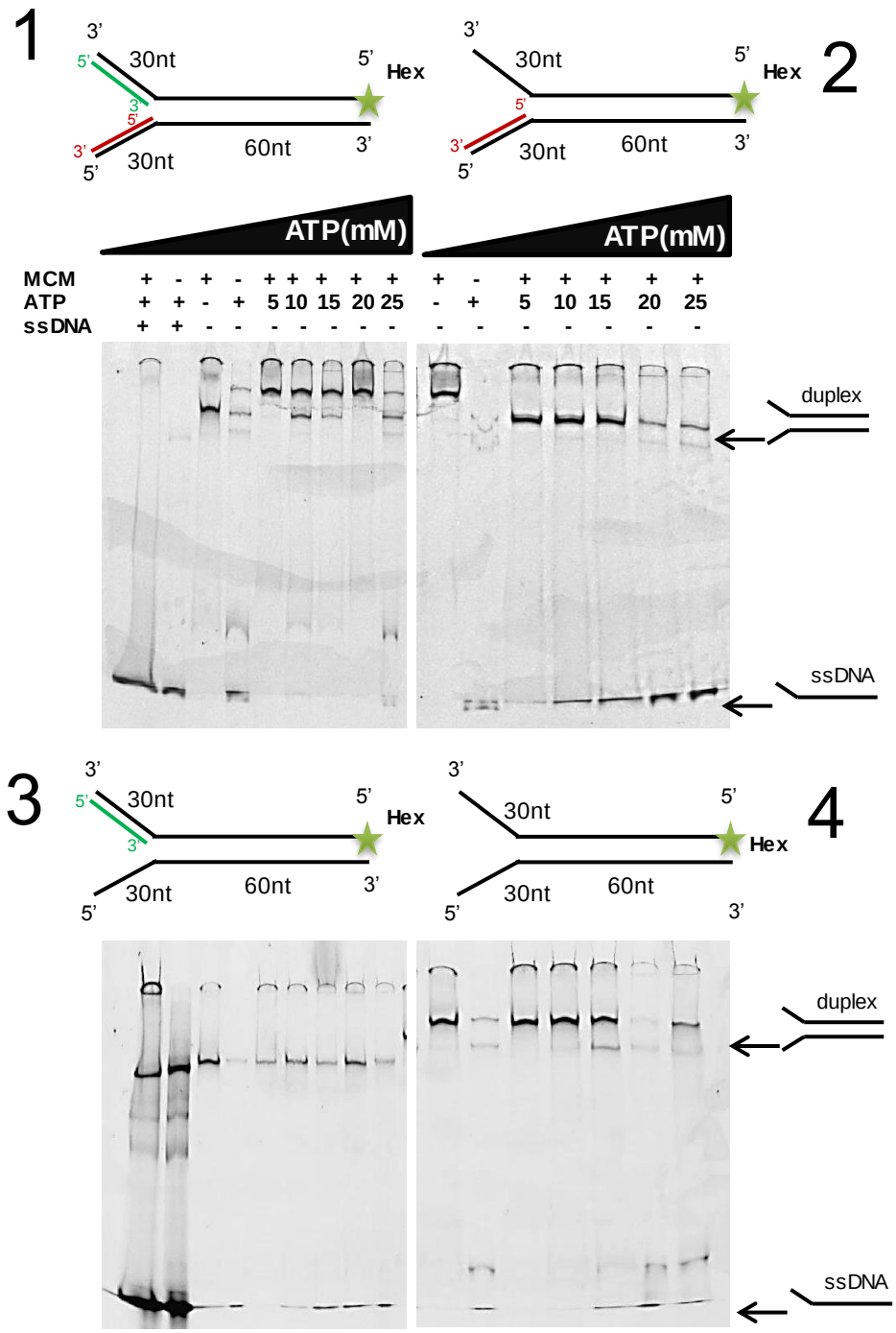


Figure 6.6. Afu-MCM is an ATP-dependent helicase

Y-shaped substrates, labelled on one strand, were incubated with 500 nM of MCM and 20 mM $MgCl_2$ at 50°C before adding various concentrations of ATP. The concentrations of ATP used in these experiments are shown on top of each gel. ssDNA was used as a comparison. The control lanes are reactions without either ATP or MCM, where no unwinding is observed. Arrows show the substrates (duplex DNA) and products (ssDNA). Among four substrates, substrate 2 was the best substrate for Afu-MCM helicase activity. The bands that are higher than the substrate in lanes are believed to be excessive amounts of MCM protein.

need for auxiliary factors *in vitro*. Moreover, we observe that *A.fulgidus* MCM has an ATP- dependent 3'→5' helicase activity. This is similar to eukaryotic MCM, but in contrast to the helicase activity of *E.coli* DnaB, which possesses a 5'→3' helicase activity, suggesting the requirement of both 3' and 5' tails for the unwinding activity of bacterial replicative helicases (Kaplan and Steitz, 1999).

6.6. Expression and purification of Afu-GINS

Toward our long-term goal of reconstituting an archaeal CMG complex *in vitro*, we next purified Afu-GINS using affinity chromatography. The pET28a:Afu-GINS expression construct was transformed into Rosetta *E.coli* and then incubated at 37°C for 3 hours. Following addition of 1 mM IPTG at 25°C overnight, the soluble cell lysate was prepared and loaded onto nickel-agarose affinity chromatography. Bound GINS was eluted from the nickel affinity chromatography column by addition of 300 mM imidazole. Eluted GINS was resolved on a SDS-polyacrylamide gel to confirm the correct size. The result was in agreement with its predicted molecular mass. Afu-GINS has a predicted molecular weight of ~ 25 kDa (Figure 6.7). As a second purification step, anion exchange chromatography (Mono Q) was employed to remove *E.coli* contaminants. The 100% B eluted peak fraction from nickel affinity column was subjected to MonoQ column and following a wash step, the bound GINS was eluted with 1M NaCl. The successfully purified protein resolved as a band with an apparent molecular mass of ~ 25 kDa on the SDS-PAGE (Figure 6.7).

6.7. Characterisation of biochemical properties of Afu-GINS

6.7.1. Afu-GINS DNA binding ability

Previously GINS complex demonstrated a high affinity for binding ssDNA or for binding dsDNA containing ssDNA stretches (Boskovic *et al*., 2007). It was shown that the GINS complex isolated from crenarchaeon *S. solfataricus* (SsoGINS), similar to human GINS, possesses DNA binding activity and preferentially binds to ssDNA over dsDNA (Lang and Haung, 2015). However, studies on the euryarchaea *P. furiosus* and *T. acidophilum* GINS complexes proposed no DNA binding activity by these proteins (Yoshimochi *et al*., 2008; Ogino *et al*., 2011). To determine if GINS protein isolated from *A.fulgidus* was capable of binding

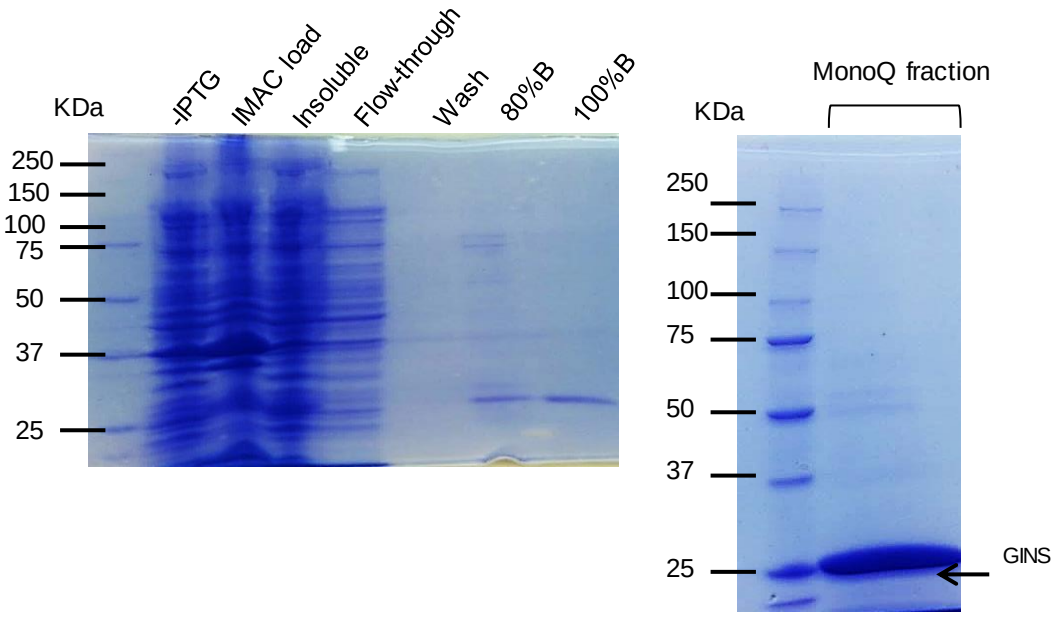


Figure 6.7. Purification of Afu-GINS using chromatography

A three litre culture of *E.coli* Rosetta transformed with pET28a: GINS expression construct was grown at 37°C for 3 hours and subsequently induced with 1 mM IPTG and incubated overnight. Following cell lysis, soluble cell lysate was loaded onto Ni $^{2+}$-NTA chromatography. Bound Afu- GINS was washed and eluted with 300 mM imidazole. The 100% B eluted peak fraction from nickel affinity column was subjected to MonoQ column and after wash step, the bound GINS was eluted with 1M NaCl. Fractions were collected and analysed by SDS-PAGE gel. The result was in agreement with Afu-GINS predicted molecular mass.(~ 25 KDa).

to DNA, gel mobility shift assays (EMSAs) using a 45-mer labelled ssDNA and dsDNA substrates were used. Based on our results, unlike SsoGINS but similar to GINS complexes isolated from *P. furiosus* and *T. acidophilum*, almost no band-shift was observed with Afu-GINS (Figure 6.8). This might suggest that Afu-GINS has little DNA binding ability by itself. We also incubated the Afu-GINS with forked DNA, which was used for MCM DNA binding and helicase activity assays (section 6.4 and 6.5.). Afu-GINS failed to bind the forked DNA as well (Figure 6.9). Using protein sequence analysis, Lang and his colleague suggested that all known crenarchaeal GINS proteins are distantly related to their homologs in euryarchaea (Lang and Haung, 2015). Therefore, this could explain why GINS complexes show different DNA binding ability in these species.

6.7.2. GINS does not stimulate MCM helicase activity

As previously shown (section 6.5.), MCM protein isolated from *A.fulgidus* resembles its homologues in *M. thermoautotrophicum* and some other species (Ishino *et al*., 2011; Grainge *et al*., 2003; Kelman and Hurwits, 1999) possessing an independent 3'→5' helicase activity. However, a growing body of evidence indicated that homohexameric MCM isolated from *Picrophilus torridus*, *S. solfataricus* and *P.furiosus* require GINS as an auxiliary factor for helicase activity (Goswami *et al*., 2015; Lang and Huanf, 2015; Yoshimochi *et al* 2008). To explore the effect of Afu-GINS on Afu-MCM, the helicase activity of Afu-MCM was measured in the presence and absence of Afu-GINS. Unexpectedly, the helicase activity of MCM was reduced by increasing concentrations of Afu-GINS (Figure 6.10). It is not likely that Afu-GINS inhibits loading of Afu-MCM on the DNA, as Afu-GINS did not show any DNA binding activity in contrast to Afu-MCM (Sections 6.4 and 6.7.1.) which suggests that Afu-MCM has a higher affinity for DNA compared to Afu-GINS. Since an interaction between GINS and MCM has been identified in different archaea, it is possible that the GINS-MCM interaction, similar to the Cdc6-MCM interaction, regulates the helicase activity of MCM. Previously, Kelman and colleagues indicated that DNA-Cdc6 interaction was not required for helicase activity inhibition. Notably, the MCM-Cdc6 interaction resulted in the inhibition of helicase activity by MCM in *Methanothermobacter thermautotrophicus* (Kasiviswanathan *et al*., 2005). In section 6.9 (below), detection of interaction of GINS and MCM in *A.fulgidus* is described.

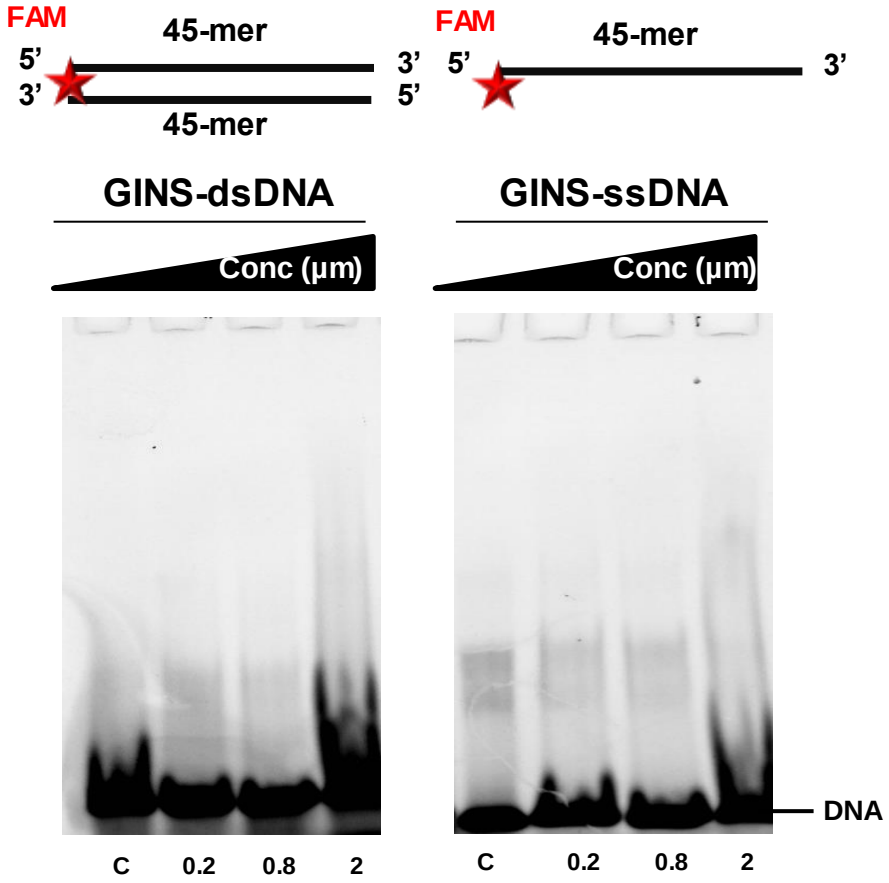


Figure 6.8. Afu-GINS shows no DNA binding activity

The substrates used in the mobility shift assay are shown schematically. Various concentrations of Afu-GINS were incubated with either 45-mer ssDNA or dsDNA (Table 2.7) and10 mM Mg^{2+}. The concentrations of GINS used in these experiments are shown below each gel. No DNA binding activity was observed by Afu-GINS in the presence of both ssDNA and dsDNA.

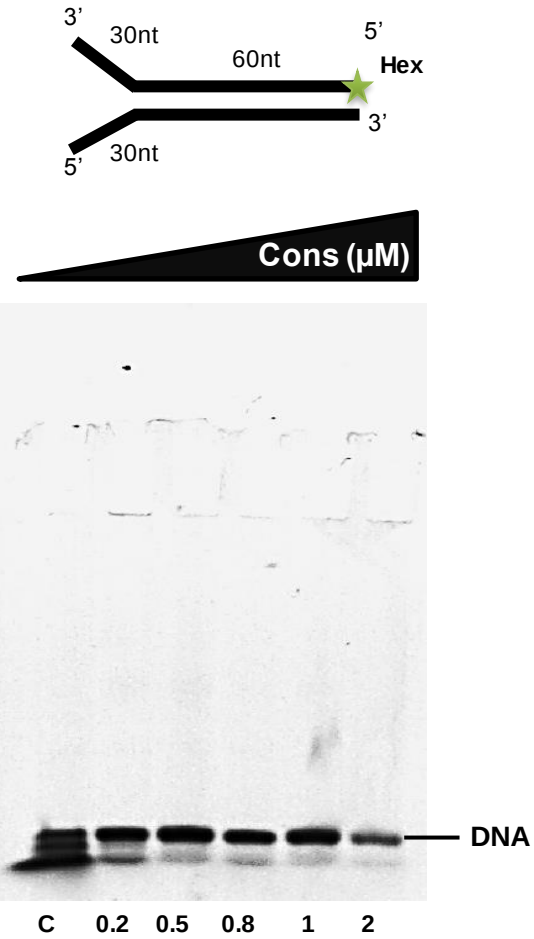


Figure 6.9. Afu-GINS does not bind to forked DNA

DNA binding affinity of Afu-GINS to a forked DNA substrate was examined. The substrate used in the mobility shift assay is shown schematically. A 60-mer duplex DNA with 30-mer 3' and 5' tails (Table 2.7) was incubated with 10 mM Mg^{2+} ,1 mM DDT and increasing concentration of GINS. The concentrations of GINS used in these experiments are shown below the gel. Again no DNA binding was observed by Afu-GINS.

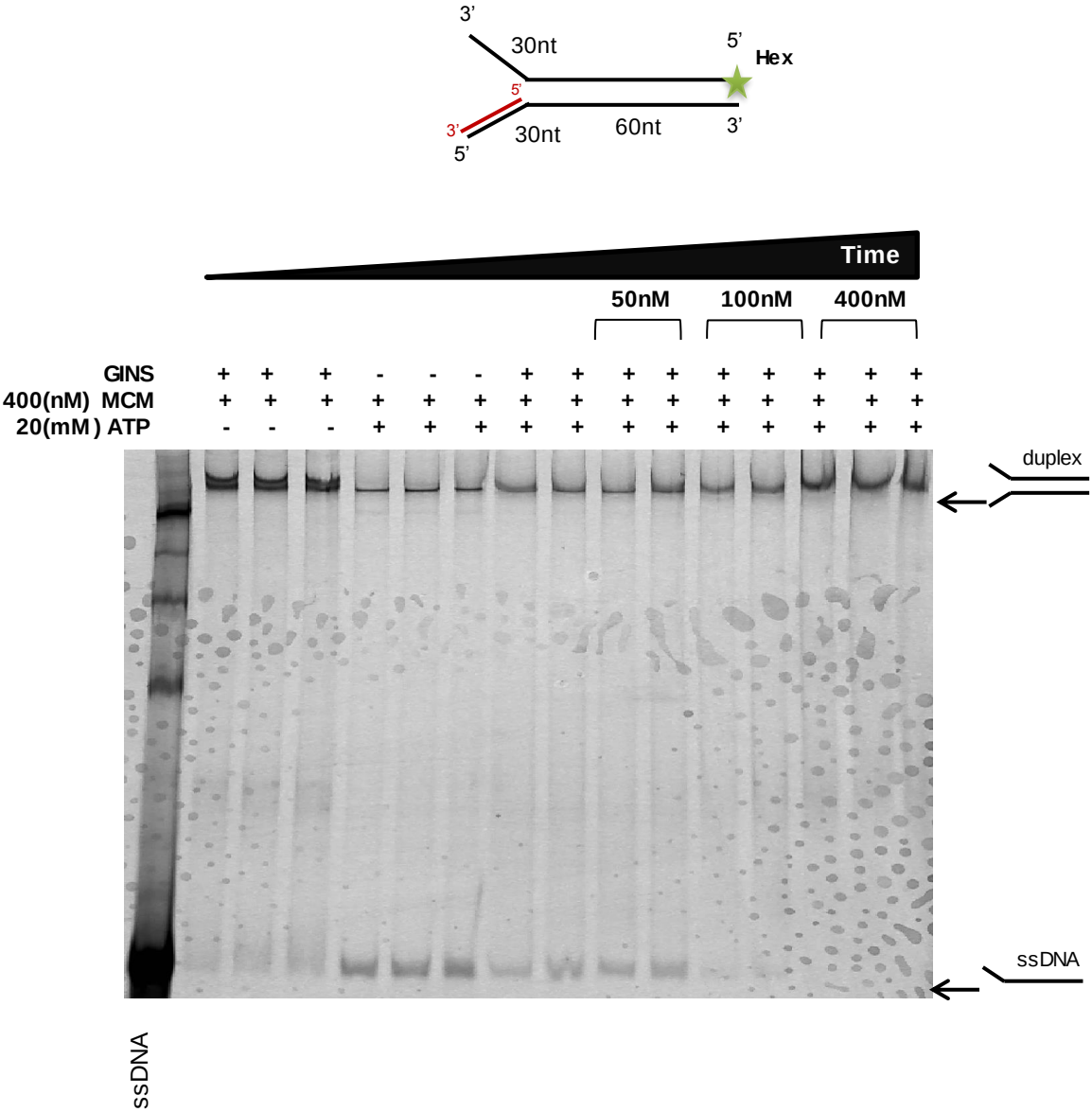


Figure 6.10. Afu-GINS does not affect the helicase activity by Afu-MCM at stoichiometry level.

The helicase activity of Afu-MCM was examined in the presence of GINS protein. The substrate used in DNA helicase assay is shown schematically. Y-shaped substrate, labelled on one strand, was incubated with 400 nM of MCM and 20 mM $MgCl_2$ and increasing amounts of GINS at 50°C before adding 20 nM ATP. The concentrations of GINS used in the experiment are shown on top of the gel. The ssDNA was used as a comparison. The control lanes are reactions without either ATP or GINS. The substrates (duplex DNA) and products (ssDNA) are shown by arrows. All reactions were incubated for 30 s, 5', and 10'.

In addition, it cannot exclude the possibility that Afu-GINS requires to form a complex with another protein(s), such as Cdc45, in order to stimulate the helicase activity of Afu-MCM. Therefore, further biochemical studies regarding the effects of Afu-GINS on Afu-MCM helicase activity are required.

6.8. Expression and purification of Afu-RecJ-like (Cdc45) homologue

Until relatively recently, a homologue of eukaryotic Cdc45 had not been identified in an archaeal genome. However, in 2006, an interaction study on the *Sulfolobus solfataricus* GINS complex showed the co-purification of SsoGINS with an additional polypeptide, which was initially named as RecJdbd (RecJ-like DNA-binding domain) due to its homology with the DNA binding domain of a bacterial exonuclease (RecJ) (Marinsek *et al*., 2006). Subsequent studies revealed that the GINS complex from euryarchaeon *T. kodakarensis* interacts with TK1252p nuclease protein known as GINS-associated nuclease (GAN) (Li *et al*., 2010; Li *et al*., 2011). Unlike RecJdbd, GAN possesses a distinguishable DHH phosphoesterase domain. Interestingly, sequence analysis studies demonstrated that Cdc45 possesses a DHH phosphoesterase domain at the N-terminus. These findings suggested that Cdc45 is the homologue of RecJ (Krastanova *et al*., 2012; Sanchez-Pulido *et al*., 2011). Subsequently, this homology was confirmed by further structural studies (Makarova *et al*., 2012; Yuan *et al*., 2016). Recently, RecJdbd was renamed as CdC45 (Xu *et al*., 2016). To date, except for *Caldivirga maquilingensis*, all sequenced archaeal genome possess at least one RecJ orthologous. This suggests the importance of this protein in archaeal replication machinery. In *A.fuligidus* genome two RecJ homologues have been identified using the arCOG database (Af0699 and Af0698) (Makarova *et al*., 2012).

To purify Afu-RecJ-like (Cdc45), the pETduet-1:Afu-RecJ699/698 expression construct (section 6.2) was transformed into *E.coli* Rosetta and grown at 37°C for 3 hours until late exponential phase (OD_{600}=0.6) and then induced for expression using 1mM IPTG at 25°C overnight. Cells were harvested and resuspended in 20 ml lysis buffer with 20 mM Tris (pH 7.5) and 200 mM NaCl. After sonication and centrifugation, whole cell lysate was subjected to heat denaturation at 70°C for 20 minutes. Clarified cell lysate was loaded to Nickel-NTA agarose affinity

chromatography. The column was washed with 20 mM imidazole before elution with 200 mM imidazole. The 100% B eluted peak fraction from the nickel affinity column was collected and diluted with buffer 10X to decrease the salt concentration. The eluate was then loaded onto a 5 ml MonoQ column, After washing the column with 200 mM NaCl, a linear salt gradient was applied to the column (200- 500 mM NaCl). RecJ99/98 eluted at ~ 100 mM NaCl. Fractions collected from both steps of purification were analysed on a 10% Tricine-SDS-PAGE gel, which can be used to detect proteins in the mass range of 1-100 KDa (Figure 6.11). The reason we used a Tricine-SDS gel was to detect faster migrating proteins, such, as the Afu-RecJ698 with molecular mass of only ~8 kDa. The 100% B eluted peak fraction from the nickel affinity column showed over-expression of some non-specific bands, which were likely *E.coli* contaminants. Although a band with a size similar to Afu-RecJ699 (48 kDa) was observed, unexpectedly, no band corresponding to Afu-RecJ698 was detected on the gel. Therefore, further optimisation of the purification protocol was required. Several strategies and induction temperatures for expressing Afu-RecJ99/98 were tested, including the use of *E.coli* strains BL21 and Shuffle to improve correct folding of proteins. In each case, cultures were incubated at either 16°C or 20°C overnight following IPTG induction. Only when the Shuffle strain was used could a low level of IPTG-induced expression of RecJ99 be observed at 20°C (Figure 6.12). Other conditions did not show any soluble RecJ99 protein. However, in none of the conditions tested could any soluble Afu-RecJ98 be produced.

Previously, Bell and colleagues were able to co-purify the GINS complex from *S. solfataricus* with RecJdbh in eight steps (Marinsek *et al*., 2006). They also purified *S. acidocaldarius* GINS-Cdc45 complex recently (Xu *et al*., 2016). Therefore, since we failed to purify the Afu-RecJ99/98 complex, it was decided to try the co-purification of Afu-RecJ99/98 with Afu-GINS. To enable this co-purification, pET28a:Afu-GINS was co-transformed with the pETduet-1:Afu-RecJ699/698 expression construct in a Rosetta *E.coli* strain and incubated at 37°C for 3 hours. The culture was then induced for expression using 1 mM IPTG followed by overnight incubation at 20°C. After sonication and centrifugation, harvested cells were subjected to the heat denaturation at 70°C for 20 minutes.

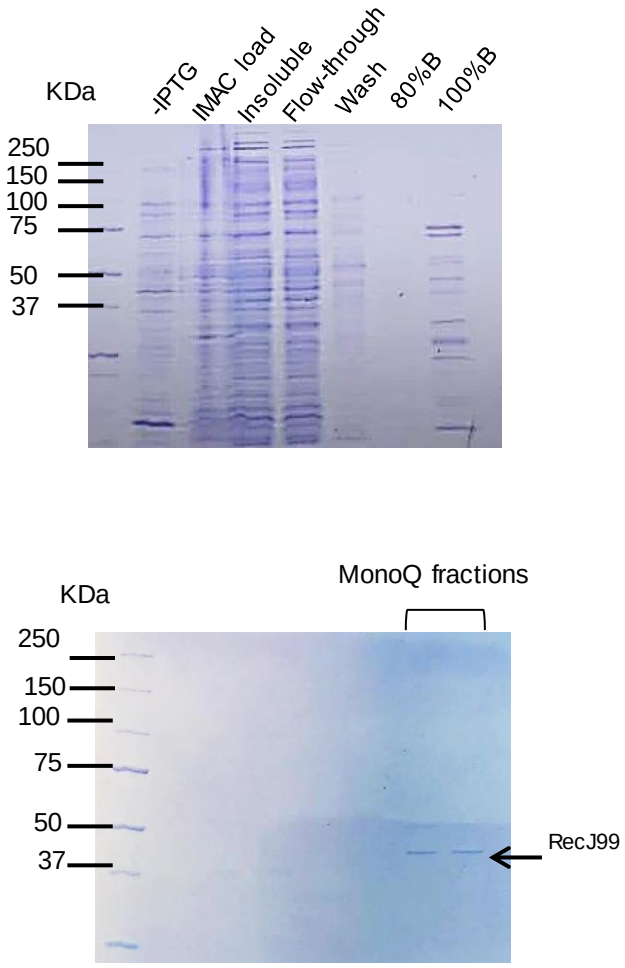


Figure 6.11. Purification of Afu-RecJ99/98 (Cdc45) complex

A three litre culture of *E.coli* Rosetta transformed with pETDuet: Afu-RecJ99/98 expression construct was grown at 37°C for 3 hours and subsequently induced with 1 mM IPTG and incubated overnight. Following cell lysis, clarified cell lysate was loaded onto Ni^{2+}-NTA chromatography. The column was washed with 20 mM imidazole before elution with 200 mM imidazole. The 100% B eluted peak fraction from nickel affinity column was collected and diluted at least 10X. The eluate was then subjected to MonoQ column and developed with a linear gradient to 20 mM Tris (pH 7.5) and 500 mM NaCl. RecJ99/98 was then eluted at ~ 100 mM NaCl. Fractions collected from both steps of purification were resolved by a 10% Tricine-SDS-PAGE and coomassie stained. Arrow indicates the protein band with an apparent molecular mass of ~48 KDa.

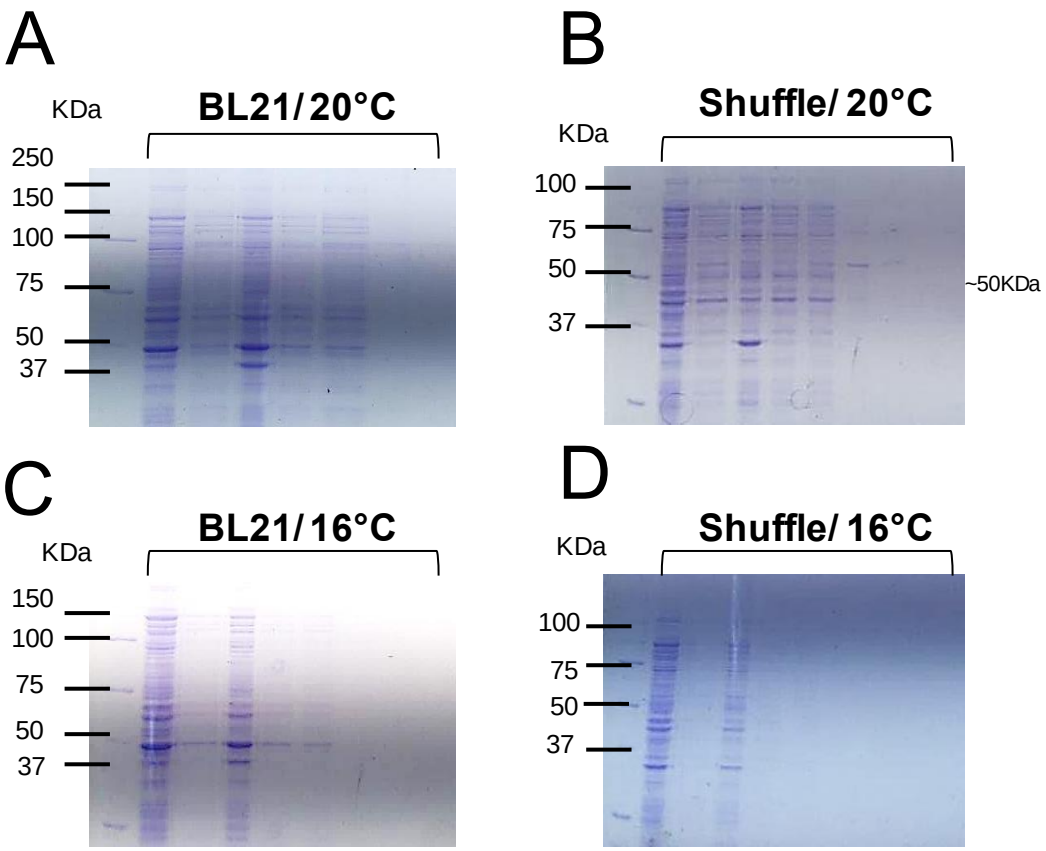


Figure 6.12. Optimisation of Afu-RecJ99/98 expression in *E.coli* strains

The pETDuet-1:Afu-RecJ99/98 construct was transformed into either BL21 or Shuffle *E.coli* strains and grown to exponential phase (OD_{600}=~0.6). Cultures were then induced with addition of 1 mM IPTG before incubation at either 20°C (**A**,**B**) or 16°C (**C,D**). Clarified cell lysate was loaded to the Ni2+-NTA agarose affinity chromatography. Eluted Samples along with –IPTG samples were analysed by SDS-PAGE gel. Arrow indicates a protein band with an apparent molecular mass of ~50 KDa.

After centrifugation, clarified cell lysate was loaded onto nickel-NTA agarose affinity chromatography. The column was washed with 20 mM Tris pH 7.5, 500 mM NaCl and 10 mM imidazole before elution with 300 mM imidazole. Eluted protein complexes were resolved on a Tricine-SDS-PAGE gel. One band with a molecular mass of 25 kDa corresponding to Afu-GINS and one band with size of ~ 50 kDa were detected on the gel (Figure 6.13). However, again no band corresponding to AfU-RecJ98 was observed. Anion exchange chromatography (MonoQ) was used as the second step of purification. The 100% B eluted peak fraction from nickel affinity column was collected and diluted 10X with buffer to decrease the salt concentration. The eluate was loaded onto a 5 ml MonoQ column, a linear salt gradient applied to the column (150 mM NaCl) and protein complex eluted at 1M NaCl. Collected fractions were analysed on a Tricine-SDS gel. In addition to GINS protein, another band with a molecular mass of ~ 50kDa was detected. To identify whether this protein was the RecJ-like protein from *A.fulgidus*, this protein band was analysed by mass spectrometry, which was carried out by my colleague Dr. Peter Kolesar. Even though the size of the band represented a protein with similar size to Afu-RecJ99, the mass spectrometry data indicated that the obtained protein was an *E.coli* contaminant (not shown). Therefore, a different approach for successful purification of Afu-RecJ99/98 complex was required.

Cloning and sequencing of the pETDuet-1Afu-RecJ99/98 expression construct was repeated and confirmed its identity. In addition, different ranges of salt concentration were applied to improve protein solubility (data not shown). However, despite the numerous attempts to express Afu-RecJ99/98 in different bacterial systems, we were unable to produce recombinant Afu-RecJ99/98 protein in an *E.coli* heterologous expression system. Although recombinant RecJ-like protein was not soluble in *E.coli* cells, it would be possible that this protein might be expressed from the endogenous locus. Recently, *S. islandicus* Cdc45 protein was purified by Bell and colleagues. For this purification they took advantage of the pSSR plasmid and TSVY medium for expression and cell growth (Zheng *et al*., 2011; Xu *et al*., 2016). Due to time limitations, we were unable to pursue the purification of Afu-RecJ-like protein, therefore, future purification studies are required.

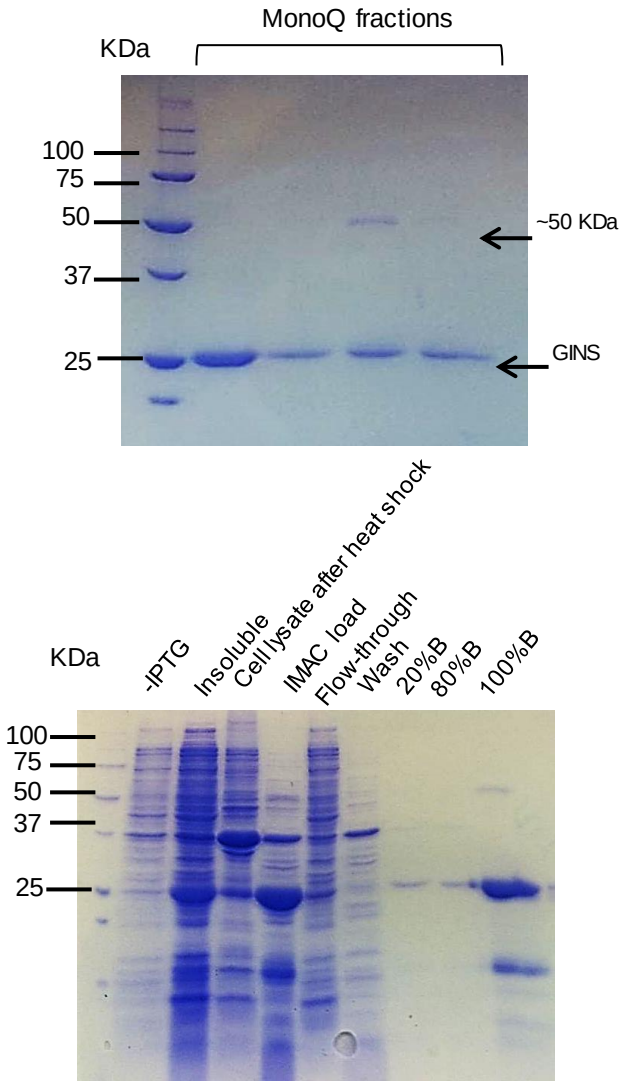


Figure 6.13. Co-purification of Afu-GINS with Afu-RecJ99/98

pET28a: Afu-GINS was co-transformed with the pETduet-1: Afu-RecJ699/698 expression construct into Rosetta *E.coli* and incubated at 37°C until exponential phase (OD_{600}=~0.6). Subsequently the culture was induced with addition of 1 mM IPTG and incubated at 20°C overnight. Following heat shock at 70°C, clarified cell lysate was loaded onto Ni $^{2+}$-NTA chromatography. The column was then washed with 10 mM imidazole before elution with 300 mM imidazole. The 100% B eluted peak fraction from nickel affinity column was collected and diluted at least 10X. The eluate was then subjected to MonoQ column and developed with a linear gradient to 20 mM Tris (pH 7.5) and 150 mM NaCl. The complex was then eluted at ~ 1 M NaCl. Fractions collected from both steps of purification were analysed by SDS-PAGE gel. Arrows indicate a protein band with an apparent molecular mass of ~50 KDa and GINS protein with molecular mass of 25 KDa. .

6.9. Towards reconstitution of an archaeal CMG complex in vitro

It is evident that replicative DNA polymerases in all domains of life are not able to initiate synthesis of new strands of DNA *de novo* and thus the presence of a primer synthesised by a DNA primase is essential for initiation of all cellular DNA replication (Frick and Richardson, 2001). Primases only bind to ssDNA to start priming before MCM loading and unwinding activity. It has been shown that the DnaG primase from *E.coli* binds to the DnaB helicase. This binding not only assists synthesis of short RNA primers by DnaG, but also facilitates release of the regulatory protein DnaC from the helicase. This suggests that there is co-ordination of priming and unwinding in bacteria (Kaguni, 2011). In eukaryotes, the DNA polymerase α-primase complex catalyses priming. The heterotetrameric GINS complex that binds both Cdc45 and MCM2-7 proteins to form the CMG complex, may bind the Pol α/primase and recruit it to the replication fork. In budding yeast, it was demonstrated that a complex of GINS and the Ctf4 trimer is required to link the Pol α/primase complex to MCM2-7 (Gambus *et al*., 2009; Masai *et al*., 2010; Simon *et al*., 2014). These studies proposed that coupling of Pol α/primase to MCM helicase might be required for efficient DNA replication in eukaryotes. The archaeal DNA replication machinery is significantly similar to that of eukaryotes and recently it was indicated that the CMG complex is a conserved component in all archaeal and eukaryotic DNA replication machineries. In archaea, a heterodimeric AEP primase complex (PriS/L) performs priming. The presence of a GINS domain fused to the C-terminal domain of PriS subunit (Swiatek and MacNeill, 2010) as well as the GINS interactions with primase and helicase (Marinsek *et al*., 2006) suggest that archaeal GINS might act to link the primase to the CMG complex. These findings point towards the possible coupling of priming and unwinding events in archaea. A recent study indicated that repriming activity by Pol α/primase coupling by MCM helicase activity plays an important role in genomic stability in eukaryote. Deregulated or limited repriming leads to replication fork uncoupling, that in turn generates long stretches of ssDNA at the fork (Fumasoni *et al*., 2015).

In all domains of life, there are distinct DNA damage tolerance mechanisms that stabilize and rescue stalled replication forks, including repriming and TLS mechanisms. It is believed that repriming is limited to lesions on the leading

strand replication (Guilliam et al., under review). Due to the discontinuous nature of lagging strand synthesis, the lagging-strand replication is inherently tolerant of DNA blocking lesions or obstacles (Pagès and Fuchs, 2003; Svoboda and Vos, 1995). During lagging strand synthesis, primers are produced constantly for Okazaki fragment synthesis. During repriming, the newly generated primer is able to restart replication downstream of the lesion. On the other hand, leading strand repriming past the lesion will generate ssDNA gaps opposite the lesion and subsequently these gaps can be filled by post-replicative TLS mechanisms. In this thesis, we have shown that archaeal replicative primase (PriS/L) plays a significant role in DNA damage tolerance by performing translesion synthesis (TLS). We discovered that PriS1/L is capable of bypassing some DNA lesions through TLS, which facilitates the resumption of stalled replication forks (Chapter 4). Repriming downstream of lesions by archaeal PriS1/L has not yet been shown. However, it is likely that they also perform repriming post-lesion. Here, we describe our attempts toward reconstitution of the archaeal CMG complex *in vitro* with the aim of shedding light on the role of archaeal replicative primase in replication-specific TLS or repriming.

To reconstitute the CMG complex from *A.fulgidus*, we first tried to express and purify components of CMG (MCM, GINS and Cdc45) separately, as discussed in sections 6.3, 6.6 and 6.8. Since we failed to purify Afu-Cdc45 either alone or as a complex with Afu-GINS (Section 6.8), we could only try to assemble the MCM and GINS complex onto DNA. We first loaded MCM onto a biotinylated dsDNA coupled to streptavidin magnetic beads (Yeeles *et al*., 2015). The protein recruitment assay was carried out using 1 µM DNA and low-salt buffer containing 20 mM Tris, pH 7.5, 20 mM $MgCl_2$, 1 mM EDTA and 2 mM DTT. 500 nM of MCM was added into the reaction prior to incubation with magnetic beads at room temperature for 30 minutes with agitation. Beads were washed three times with wash buffer (40m Mm Tris, pH 7.5, 20 mM $MgCl_2$, 1 mM EDTA and 2 mM DTT) and subsequently resuspended in SDS-loading buffer. Immunoblotting was deployed to analyse bound MCM (Figure 6.14). The control consisted of a sample with no DNA. As a positive control, a purified his-tagged protein was used. The his-tagged MCM was recognized by anti-his antibody, thus, a band corresponding to MCM was detected on the blot (Figure 6.14). When the protein

was incubated with reaction buffer and no DNA, no band was detected by western blotting. Therefore, this result confirmed an interaction of MCM with the DNA.

We next tried to assemble both MCM and GINS onto the DNA. Initially, MCM was loaded onto DNA and attached to the beads. Magnetic beads were then isolated and a low-salt buffer (20 mM Tris, pH 7.5, 20 mM MgCl2, 1 mM EDTA and 2 mM DTT) containing 1 µM GINS (his-tag was cleaved by thrombin) was added. After incubation at room temperature for 30 minutes, beads were washed three times using wash buffer. After resuspending in SDS-loading buffer, bound proteins were analysed by Immunoblotting. Controls consisted of a sample with no MCM and a sample with no DNA (Figure 6.15). We were able to assemble the MCM and GINS complex onto the DNA *in vitro*. As shown in figure 6.15, when loaded MCM was incubated with GINS, two bands corresponding to MCM and GINS were detected on the blot. In parallel, his-tagged GINS protein was used as the positive control and recognized by anti-his antibody. Interestingly, in the absence of MCM, GINS was not loaded onto the DNA (Figure 6.15). Based on the results, recruitment of GINS requires MCM. When MCM was omitted GINS was not recruited to the DNA. Therefore, in *A.fulgidus*, similar to *Sulfolobus solfataricus* (Marinsek et al., 2006) GINS and MCM form a complex.

In this chapter, we describe the cloning and purification of the MCM, GINS and Cdc45 proteins. We also described initial biochemical characterisations of MCM and GINS proteins. However, despite the numerous attempts, we were not able to purify RecJ-like (Cdc45) protein. Therefore, we could not reconstitute the whole CMG complex *in vitro*, we only showed assembly of MG (MCM and GINS) complex on the DNA. Further experiments, including optimizing conditions for the purification of Cdc45 and recruitment of Cdc45 and MG into a stable complex are required to continue this study. Additionally, once the CMG complex is reconstituted *in vitro*, an assay can be carried out to detect the ability of PriS1/L to replicate past a lesion by TLS or reprime downstream of a lesion following generation of ssDNAs by CMG complex, which can be bound by PriS1/L.

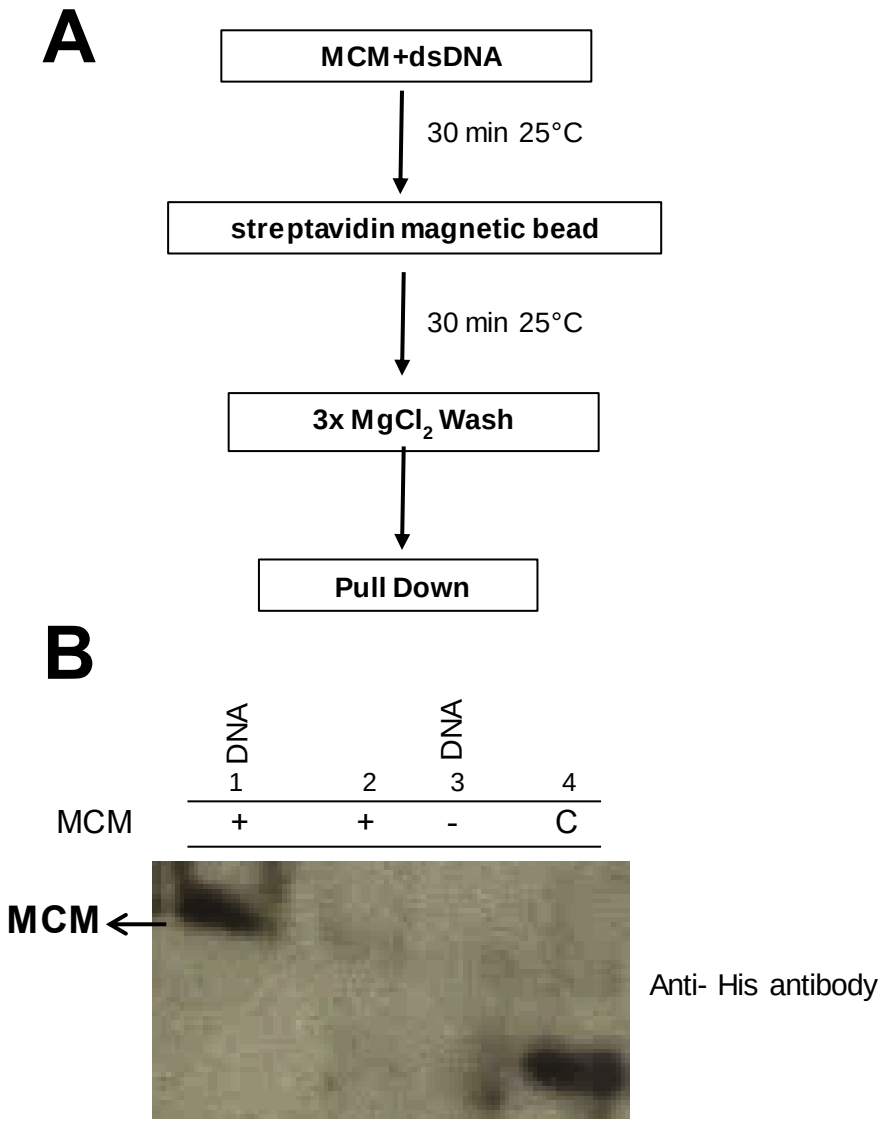


Figure 6.14. Afu-MCM recombinant protein loading onto DNA

(A) Reaction scheme for MCM loading onto DNA. **(B)** Afu-MCM (500nM) was incubated with 1 µM of double-stranded biotinylated DNA and low-salt buffer containing 20 mM Tris, pH 7.5, 20 mM $MgCl_2$, 1 mM EDTA and 2 mM DTT at 25°C for 30 minutes. The reaction was added to the streptavidin magnetic beads. Following more incubation, beads were washed three times with wash buffer containing 40m Mm Tris, pH 7.5, 20 mM $MgCl_2$, 1 mM EDTA and 2 mM DTT. Subsequently beads were resuspended in SDS-loading buffer. Bound MCM was visualized using Immunoblotting. Arrow indicates loaded MCM onto DNA in lane 1. Lane 2 shows reaction with no DNA and lane 3 illustrates reaction with no MCM. Lane 4 shows detection of a his-tagged protein by antibody which used as the positive control.

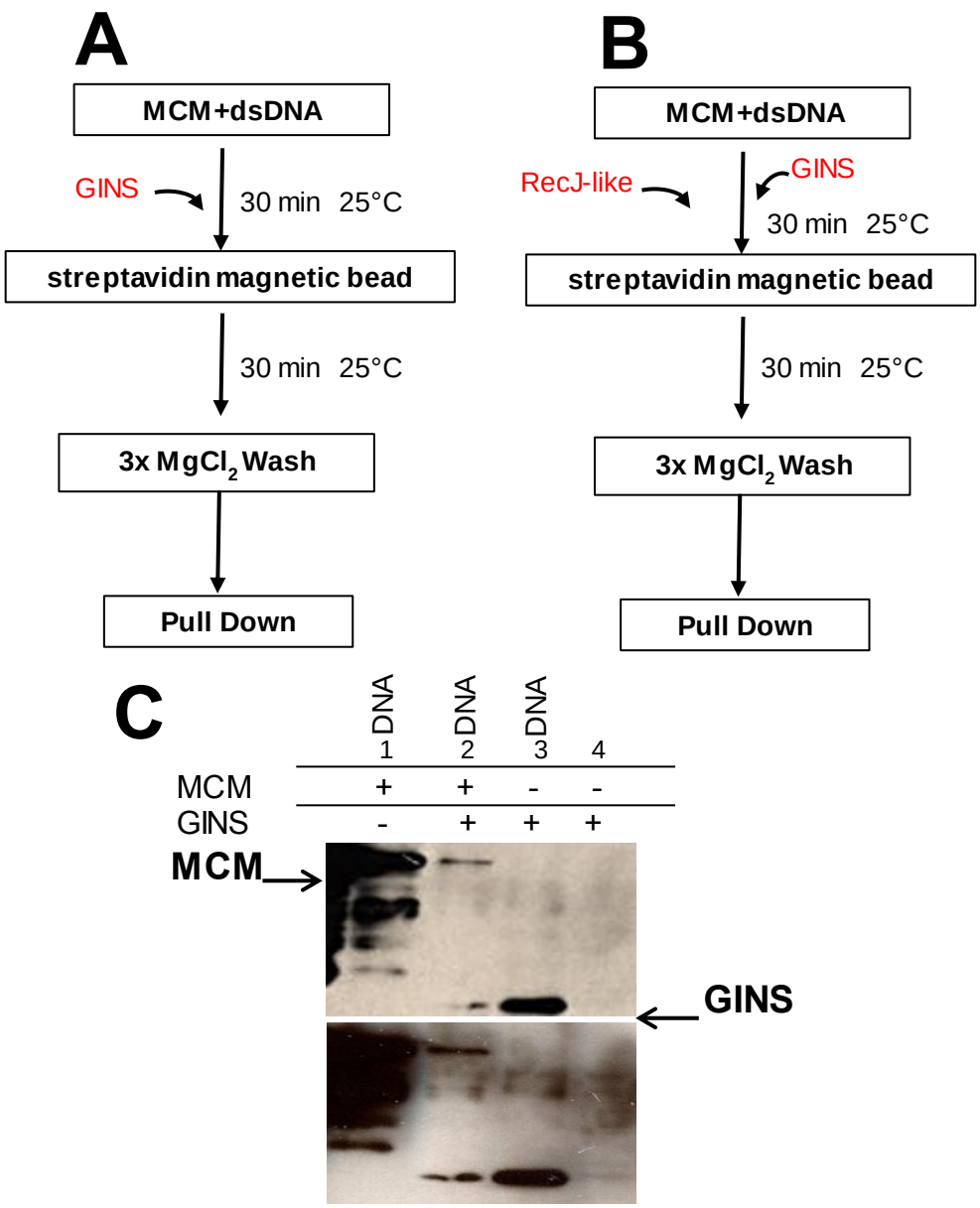


Figure 6.15. Assembly of MCM-GINS complex

(A) Reaction scheme for Afu-GINS recruitment onto loaded Afu-MCM. **(B)** Reaction scheme for GINS and RecJ-like proteins onto loaded Afu-MCM. **(C)** After loading the MCM onto the DNA, beads were isolated. Low-salt buffer (20 mM Tris, pH 7.5, 20 mM $MgCl_2$, 1 mM EDTA and 2 mM DTT) containing 1µM GINS was added to the beads. Reaction were incubated at 25°C for 30 minutes and then beads were washed with wash buffer containing 40m Mm Tris, pH 7.5, 20 mM $MgCl_2$, 1 mM EDTA and 2 mM DTT. Bound proteins were resuspended in SDS-loading buffer and visualized using Immunoblotting. DNA was added to reactions 1, 2 and 3 but not 4. Lane 1 shows reaction containing MCMC. Lane 2 shows reaction containing both MCM and GINS. Lane 3 shows purified his-tagged GINS and lane 4 illustrates reaction with no MCM.

The plan to accomplish this long-term aim is to recruit other replication proteins which were purified in this study (PriS/L, replicative polymerases (PolB, PolD) and RPA) successively to the CMG complex for reconstitution of a minimal replisome *in vitro*.

A possible model for the roles, recruitment, and regulation of PriS/L by other proteins during DNA replication in the presence of a lesion is discussed below (Section 6.14).

6.10. The possible role of RadA during replication in *A.fulgidus*

DNA replication in all domains of life initiates at specific sites known as origins. The number of origins differs in different organisms (DePamphilis, 1993). While bacteria and some archaea possess single origin, in eukaryotes and most archaea there are multiple origins. Notably, initiation of replication in bacteriophage T4 occurs through two mechanisms: origin-dependent replication and recombination-dependent replication (Dudas and Kreuzer, 2001). Recently, a recombination-dependent replication mechanism was also investigated in *Haloferax volcanii*, a member of the euryarchaeota, which contains four chromosomal replication origins (Hawkins *et al*., 2013). It was found that, deletion of all four origins in this species led to a significant increase in growth rate. Strains with no origins grew 7.5% faster comparing to the wild-type strain (Hawkins et al., 2013). This study also demonstrated that RadA, an archaeal homologue of RecA and Rad51 that catalyse homologous recombination, is required for initiation of replication in the absence of origins (Hawkins et al., 2013) suggesting the discovery of a recombination-dependent mechanism for initiating replication in this archaeal species. Although RadA-dependent replication has not yet been observed in other species of archaea, our primary results in this section might suggest a possibility that in *A.fulgidus*, RadA might be important for reinitiation of replication by homologous recombination downstream of lesion. We therefore set out to examine whether RadA interacts with PriS1/L in *A.fulgidus*. Cloning, expression and purification of Afu-RadA followed by GST pull down assays were carried out to test this interaction (see below).

6.11. Cloning of Afu-RadA gene into expression vectors

To detect if there is an interaction between RadA and PriS1/L, the ORF corresponding to RadA (AF0993) was PCR amplified from *A.fulgidus* genomic DNA (using primers in Table 2.1) and cloned into the pET28a *E.coli* expression vector, generating the pET28a:Afu-RadA expression construct (Figure 6.16).

6.12. Expression and purification of Afu-RadA

The pET28a:Afu-RadA expression construct was transformed into Rosetta *E.coli* and then incubated at 37°C for 3 hours. Following addition of 1 mM IPTG at 25°C overnight, the soluble cell lysate was prepared and loaded onto nickel-NTA agarose affinity chromatography. After washing with 500 mM NaCl and 10 mM imidazole, RadA was eluted from the nickel affinity chromatography column by addition of 300 mM imidazole. Eluted RadA was resolved on a SDS-polyacrylamide gel. The result was in agreement with its predicted molecular mass. Afu-RadA has a predicted molecular weight of 37 KDa (Figure 6.17). RadA co-eluted with considerable amounts of *E.coli* contaminants. Therefore, size-exclusion chromatography (Gel filtration) was performed. An S75 gel-filtration column was pre-equilibrated with gel filtration buffer (40 mM Tris-HCl (pH7.5), 300 mM NaCl, 10% (v/v) glycerol and 2mM β-mercaptoethanol). Concentrated protein samples were loaded onto the column via a 5 mL loop. Fractions were collected following 100 mL of flow-through and analysed by a SDS-PAGE gel and determined by absorbance at 280 nm (Figure 6.17).

6.13. Studying the interaction between Afu-PriS/L and Afu-RadA

To date, an interaction between the recombination protein, RadA and the replication primase, PriS/L has not yet reported in archaea. To test whether PriS1/L from *A.fulgidus* interacts with Afu-RadA, we utilised a GST pull-down assay using either pGEX-6P-1:Afu-PriS1 or pGEX-6P-1:Afu-PriL expression constructs. For the prey protein, we used either purified RadA protein or RadA cell extract. However, experiments with RadA cell extract provided us with better results.

The pET28a: Afu-RadA was co-transformed with either pGEX-6P-1:Afu-PriS1 or pGEX-6P-1:Afu-PriL expression constructs into Rosetta *E.coli* strain. Cell cultures were incubated at 37°C for 3 hours until late exponential phase

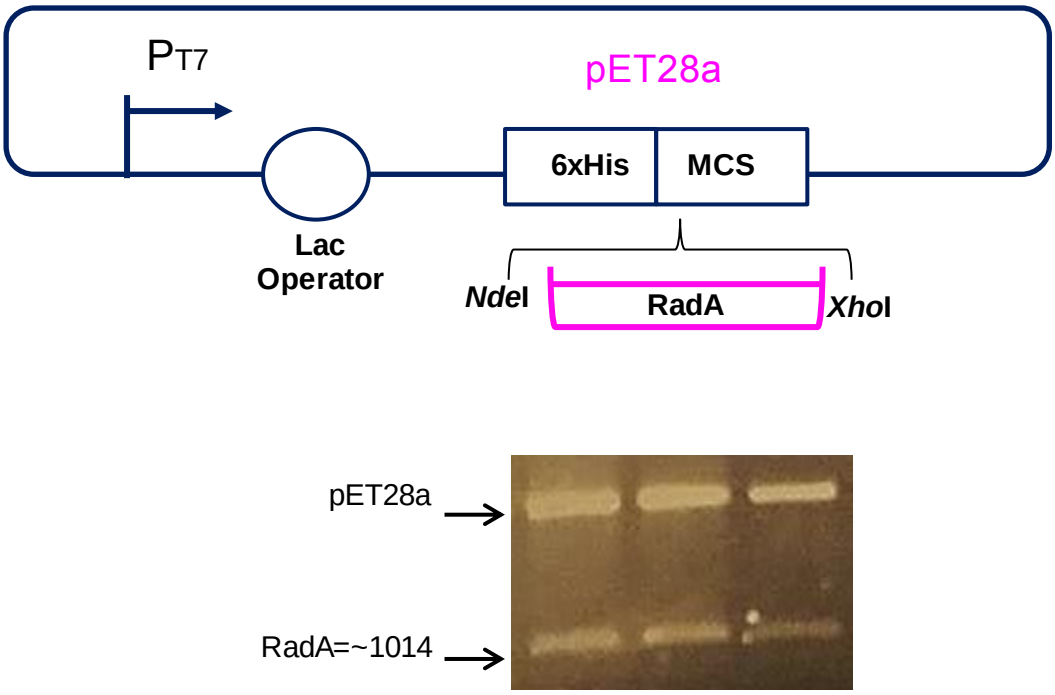


Figure 6.16. Cloning of Afu-RadA into pET28a

The open reading frame corresponding to RadA was amplified from *A. fulgidus* genomic DNA (Table 2.3.) introducing the appropriate restriction sites to allow insertion into the multiple cloning site (MCS) of the pET28a expression vector. A 6-histidine tag downstream of a T7 promoter (P_{T7}) was in frame with cloned RadA. 10x DNA loading dye was added to PCR samples and then samples were run on 1% agarose gels containing ethidium bromide.

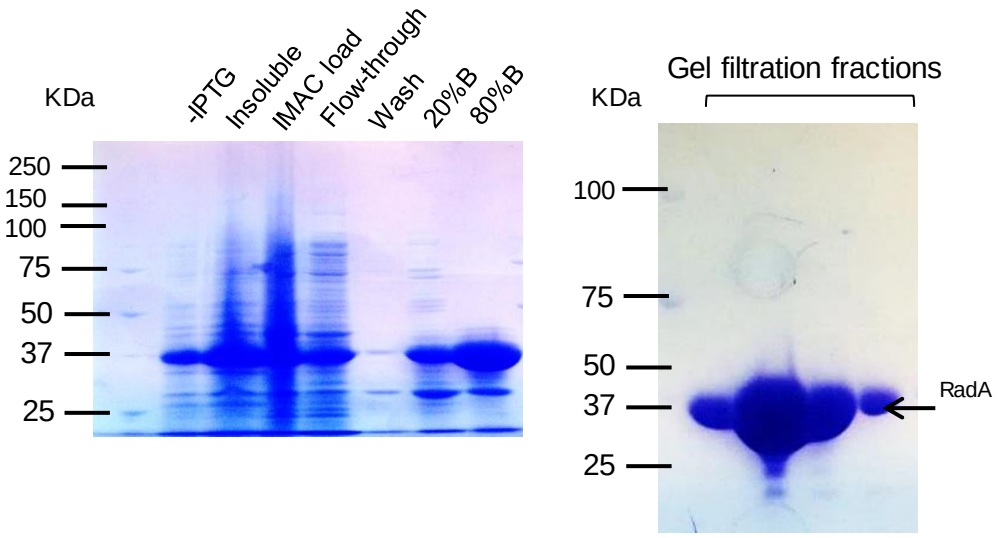


Figure 6.17. Purification of Afu-RadA

Transformation of pET28a:Afu-RadA expression construct into Rosetta E.coli followed by 3 hours incubation at 37°C. 1 mM IPTG was added to the growth for expression. Soluble cell lysate was prepared and loaded onto nickle column. After washing with 500 mM NaCl and 10 mM imidazole RadA was eluted through addition of 300 mM imidazole. Eluted RadA was resolved on a SDS-polyacrylamide gel. Arrow illustrates Afu-RadA with molecular weight of 37 KDa. In order to remove *E.coli* contaminants size-exclusion chromatography (Gel filtration) was performed. 75 gel-filtration column was pre-equilibrated with gel filtration buffer (40 mM Tris-HCl (pH7.5), 300 mM NaCl, 10% (v/v) glycerol and 2mM β-mercaptoethanol). Concentrated protein samples were loaded onto a 5 mL loop. Fractions were collected following 100 mL of flow-through and analysed by a SDS-PAGE gel and determined by A280 level.

(OD_{600}=0.6). Expression was induced using 1mM IPTG and cells were grown at 37°C for 2 hours. To lyse the cells, cell lysis buffer (40 mM Tris, pH 7.5, 300 mM NaCl, 5 mM β-mercaptoethanol, 0.1% NP-40 and 1 tablet protease inhibitor) was used. After sonication and centrifugation, clarified cell lysate was incubated and insoluble pellet were diluted with water and boiled for 10 minutes at 72°C. 30µL of GST magnetic beads at 4°C for one hour. In parallel, 1ml of soluble cell lysate and insoluble pellet were diluted with water and boiled for 10 minutes at 72°C. Beads were washed three times using lysis buffer. Subsequently, proteins were eluted and boiled in SDS-sample buffer for 5 minutes and visualized using 12% SDS-PAGE gel (Figure 6.18). Interestingly, in both pull-down assays (PriS1/RadA and PriL/RadA), a band with an apparent molecular mass of ~ 37 KDa corresponding to Afu-RadA was observed. In addition, a band of ~67 KDa corresponding to PriS1 and a very faint band of ~68 KDa corresponding to PriL were detected in PriS1/RadA and PriL/RadA pull-downs, respectively. Although further experiments are required to confirm this interaction, based on our data it can be proposed that RadA is a potential protein interaction partner of the replicative primase in *A.fulgidus.* This interaction suggests a potential coupling of replication and recombination in this species of archaea. Therefore, similar to bacteriophage T4 virus and *Haloferax volcanii*, origin-less firing mediated by RadA may also exist in *A.fulgidus*.

6.14. Summary and discussion

The critical preliminary step of DNA replication is unwinding of a duplex to two ssDNAs. In all three domains of life, helicases catalyse the unwinding process. In eukaryotic cells, MCM helicase is active in a complex with GINS and Cdc45 accessory proteins. However, in this chapter we could observe a robust 3'→5' unwinding activity by MCM protein isolated from *A.fulgidus* on its own. This data was in agreement with previous studies implicated that the helicase activity by MCM proteins from several archaeal species did not require the presence of other accessory factors (Grainge *et al*., 2003; Sakakibara *et al*., 2009). In contrast to some studies, which reported a stimulatory effect of GINS on the unwinding activity of MCM, our data showed a decrease in unwinding of forked-DNA by MCM in the presence of Afu-GINS. Since Afu-MCM showed a high binding affinity to DNA, and Afu-GINS did not bind to DNA, it is not likely that Afu-GINS could

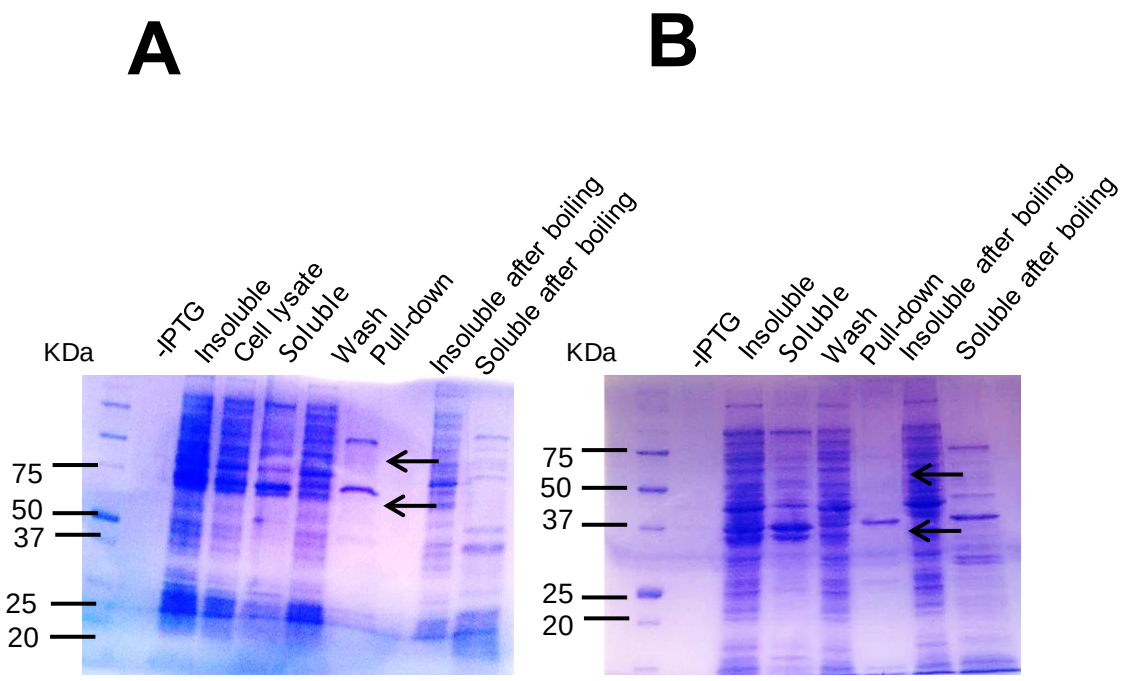


Figure 6.18. Potential interaction between Afu-RadA and Afu-PriS1/L

pET28a:Afu-RadA expression construct was co-transformed and co-expressed with either pGEX-6P-1:Afu-PriS1**(A)** and pGEX-6P-1:Afu-PriL**(B)** into Rosetta *E.coli* strain. Cultures were grown at 37°C for 3 hours followed by addition of IPTG for induction. 1 ml of clarified cell lysate was subjected to the heat shock and remains were added to GST beads. After one hour incubation, beads were washed and subsequently proteins were eluted and boiled in SDS-sample buffer for 5 minutes and visualized using 12% SDS-PAGE gel. **(A)**Arrows illustrate bands corresponding to RadA (37 KDa) and PriS1 (~67KDa). **(B)**Arrows indicate bands corresponding to RadA (37 KDa) and PriL (~68KDa).

inhibit loading of Afu-MCM onto the DNA. Although the exact reason for this inhibitory effect of GINS on MCMs helicase activity is not obvious, it is possible that the detected GINS-MCM interaction is responsible for this effect. This assumption is based on a study on *M. thermautotrophicus* that indicated that the inhibitory effect of Cdc6 on MCM helicase activity was not due to the Cdc6-DNA interaction, in fact, the MCM-Cdc6 interaction resulted in inhibition of helicase activity (Kasiviswanathan et al., 2005). Moreover, we cannot exclude the possibility that similar to *Sulfolobus acidocaldarius*, in which the GINS/CdC45 complex but not the GINS alone was required to stimulate the MCM helicase activity (Xu et al.,2016) Afu-GINS might need to form a complex with another protein, such as Cdc45, in order to stimulate the helicase activity of Afu-MCM. In addition, since Afu-MCM exhibited robust helicase activity alone, maybe no accessory protein is required for stimulation. Further biochemical studies regarding the effects of Afu-GINS on Afu-MCM helicase activity are required.

The experiments in this chapter reflect the initial attempts to reconstitute the archaeal CMG complex *in vitro* to investigate the role of archaeal replicative primase in replication-specific TLS or repriming. This work was designed to test the ability of PriS1/L to replicate past a lesion by TLS or reprime downstream of a lesion following generation of ssDNAs by CMG complex *in vitro*. To reconstitute the CMG complex, we first tried to express and purify each component (Cdc45-MCM-GINS) individually. However, our attempts to express and purify the archaeal homologue of Cdc45 from *A.fulgidus* (RecJ-like) were not successful. To purify the RecJ-like protein, we took advantage of the His-Tag. Since the histidine residues are capable of binding to different types of immobilized metal ions, such as nickel, purification of his-tagged fusion proteins can be straightforward. Despite several attempts and using different strategies to express the Afu-RecJ-like protein in different bacterial systems, we were not able to produce this recombinant protein from a heterologous system. Although recombinant RecJ-like protein was not soluble in *E.coli* cells, it may be possible to express this protein from its endogenous locus or maybe in the actual *A. fulgidus* cells using a fermenter. Alternatively, soluble Afu-RecJ-like protein may be obtained from *E.coli* after codon optimization of the native gene. Although unsuccessful, purification of Afu-RecJ-like (Cdc45) recombinant protein prevented us from reconstituting the whole CMG complex, we were able to

assemble the MCM and GINS complex onto the DNA in a stage controlled manner and show that the binding of GINS to DNA is MCM-dependent.

Maintenance of genome integrity and stability is fundamental for any form of life. However, it is inevitable that DNA replication will encounter DNA lesions. Even though cells have a variety of repair mechanisms that target and repair a vast array of DNA modifications, some types of DNA damage will persist long enough to be encountered by the DNA replication machinery. To ensure genome stability, cells use DNA damage tolerance mechanisms (DDT). These include error-prone translesion DNA synthesis (TLS), and error-free recombination-mediated restart and template switching (TS). DTT can also take place after replication by re-priming. In this thesis, we discovered that archaeal replicative primase (PriS/L) plays an important role in DNA damage tolerance through performing translesion DNA synthesis (Chapter 4). Although, we were not able to show repriming downstream of lesions by archaeal PriS/L, as these enzymes are proficient DNA primases, it is possible that they perform repriming downstream of lesions. In eukaryotes, the DNA polymerase α-primase complex is involved in origin-dependent replication initiation and also origin-independent replication initiation of lagging strands. This complex is also able to perform repriming downstream of lesions to restart stalled replication fork (Heller and Marians, 2006). Previously, it was postulated that in budding yeast, a complex of GINS and Ctf4 trimer is required to link the Pol α/pimase complex to MCM2-7 (Gambus *et al*., 2009; Masai *et al*., 2010;Simon *et al*., 2014). Remarkably, a recent study in *Saccharomyces cerevisiae* showed that, mutation of Primase/Ctf4, which is important for genome stability, resulted in deregulated or limited repriming which in turn can cause fork uncoupling and generation of ssDNA stretches at the fork. It was also suggested that uncoupling of MCM and repriming led to fork-reversal with ssDNA stretches (Fumasoni et al., 2015). In some archaeal species including *Picrophilus torridus*, *Sulfolobus solfataricus* and *S. acidocaldarius,* direct interaction between MCM and GINS has been shown (Goswami et al., 2015; Marinsek et al., 2006; Xu et al., 2006). In the present chapter, the physical interaction between MCM and GINS was also observed in *A.fulgidus*. Similar to eukaryotes, in some archaeal species interaction between GINS and primase has been identified (Marinsek et al., 2006; Swiatek and MacNeill, 2010). Although no direct interaction between MCM and primase has been discovered in archaea, it

is believed that MCM can interact with primase via GINS. This linkage might facilitate tethering of the primase to the CMG complex and allows coupling of MCM and primase on the fork. In the presence of a blocking lesion, uncoupling of unwinding and priming lead to uncoupling of leading and lagging strands and generation of long ssDNA stretches on the leading strand. ssDNAs are bound by RPA and provide a ssDNA-RPA interface. In humans, PrimPol is recruited to the ssDNA via interaction with RPA. Here, the regulatory effect of RPA on PriS/L has been identified (Chapter 5) but, detection of the possible interaction between RPA and PriS/L in *A.fulgidus* was not successful. However, interaction of RPA and primase in some archaeal species has been identified. This might suggest a possible role of RPA in the recruitment of PriS/L to the stalled replication fork. It is believed that unscheduled repriming would extend the internal gaps. These gaps can then be filled-in through a TLS mechanism. In archaea this process can be performed by replicative primase (PriS/L) until it is restricted by upstream RPA. RPA limits the involvement of PriS1/L in DNA replication to prevent any mutagenesis (Chapter 5). Altogether, it can be proposed that the coupling of helicase-priming events facilitated by the GINS complex plays a critical role in the maintenance of replication fork stability in archaea.

The possible leading strand repriming by archaeal PriS/L recruited by either CMG complex or RPA is summarized in figure 6.19.

It has been shown that DNA primases in bacteria and eukaryotes are involved in both origin-dependent replication and origin-independent replication. Previous studies on *E.coli* demonstrated the essential role of PriA, one of the components of the primosome, in damaged-inducible DNA replication (Kogoma *et al*., 1996; Kogomo, 1997). Notably, PriA protein is required for oriC-independent replication induced by DNA damage in addition to normal DNA replication. It was shown that *PriA* mutants are defective in homologous recombination and that they are also sensitive to gamma rays (Kogoma *et al*., 1995). PriA together with PriB, PriC, DnaB, DnaC, DnaG and DnaT proteins form the replication-restart primosome complex, which is involved in origin-independent restart of replication in bacteria (Masai, 2013). Interestingly, Branzei and colleagues in their recent study suggested a role for *Saccharomyces cerevisiae* polα/primase complex in DNA damage tolerance (Fumasoni *et al*., 2015). They indicated that, mutation of this

complex could disrupt recombination-mediated damage bypass mechanism. A recent genetic study on *Haloferax volcanii* reported the remarkable role of RadA in origin-less firing (Hawkins *et al*., 2013). This finding illustrates for the first time the presence of recombination-dependent replication in archaea. To date, origin-less firing has yet not been observed in other archaea. Interestingly, we could detect an interaction between PriS1 and PriL subunits of primase with RadA, using pull-down assays. This data might reflect the possible role of RadA in replication. RadA is the core protein of homologous recombination in archaea. This recombinase catalyses strand invasion by loading onto the DNA. In common with RecA and Rad51, RadA mutation led to sensitivity towards UV and ethylmethane sulfonate in *Haloferax volcanii* (Woods and Dyall-smith, 1997). In spite of the preliminary nature of the interaction between RadA and PriS1/L in *A.fulgidus*, it might be possible that recombination-dependent replication also exists in *A.fulgidus* and PriS1/L might play an important role in this process.. Future studies are required to establish if RadA, in addition to its post-replication role, plays an important role during on-going replication and also examine the requirement of the replicative primase in RadA-dependent replication.

.

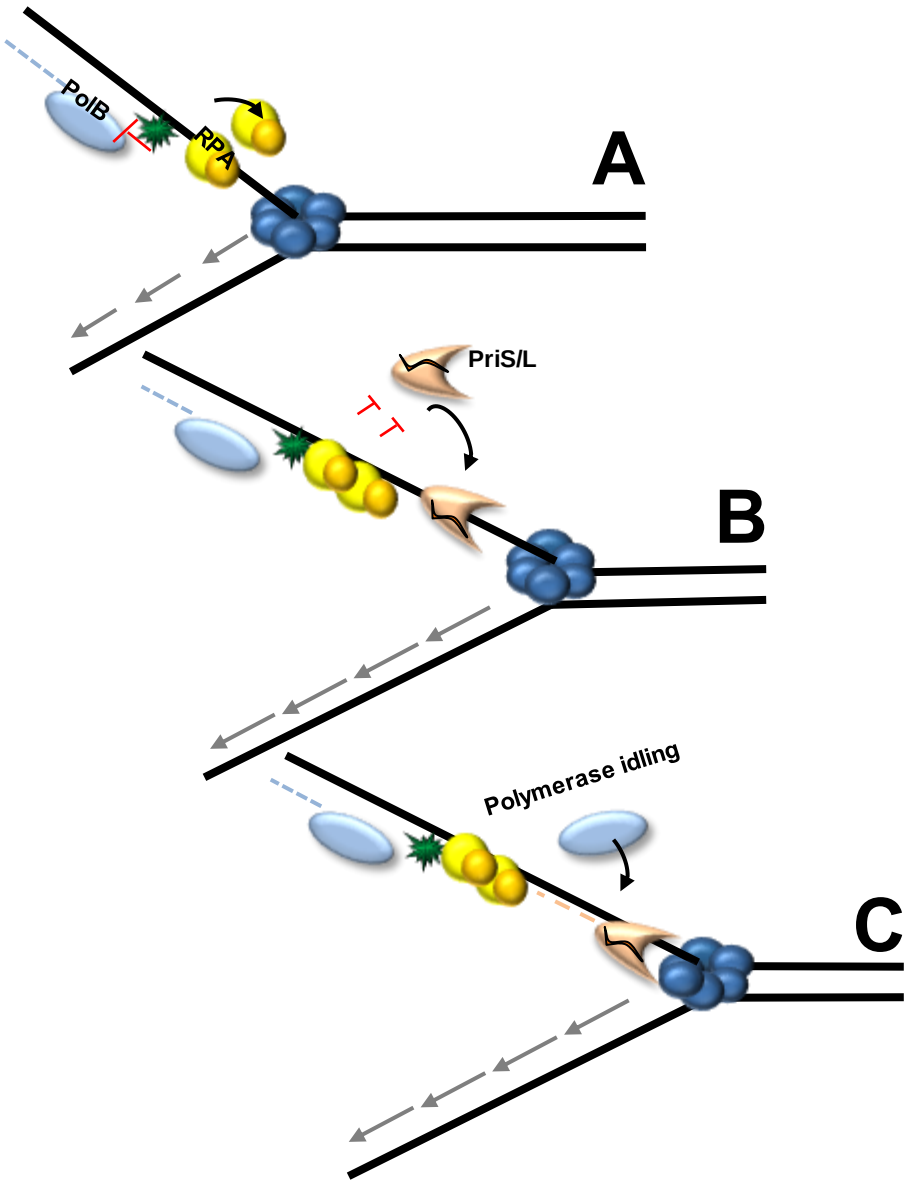


Figure 6.19. Role of replicative primase during DNA replication

(A) On the leading strand, the replicative polymerase (PolB) is stalled encountering a blocking lesion. On the lagging strand replication continues. Uncoupling of leading and lagging strands leads to generation of ssDNA on the leading strand. Unwinding continues by CMG complex and RPA bound to ssDNA. **(B)** Replicative primase (PriS/L) is recruited to the ssDNA downstream of lesion and ssDNA/RPA interface. PriS/L might either recruited by RPA or CMG complex. PriS/L might perform repriming on the leading strand. **(C)** Primer extension continues by PriS/L until it restricted by RPA. Therefore, polymerase idling occurs and a replicative polymerase continues DNA replication.

References

ABELLON-RUIZ, J., WALDRON, K. J. & CONNOLLY, B. A. 2016. Archaeoglobus Fulgidus DNA Polymerase D: A Zinc-Binding Protein Inhibited by Hypoxanthine and Uracil. *J Mol Biol,* 428**,** 2805-13.

ACHARYA, S., FOSTER, P. L., BROOKS, P. & FISHEL, R. 2003. The coordinated functions of the E. coli MutS and MutL proteins in mismatch repair. *Mol Cell,* 12**,** 233-46.

AGUILERA, A. & GOMEZ-GONZALEZ, B. 2008. Genome instability: a mechanistic view of its causes and consequences. *Nat Rev Genet,* 9**,** 204-17.

AKITA, M., ADACHI, A., TAKEMURA, K., YAMAGAMI, T., MATSUNAGA, F. & ISHINO, Y. 2010. Cdc6/Orc1 from Pyrococcus furiosus may act as the origin recognition protein and Mcm helicase recruiter. *Genes to Cells,* 15**,** 537-552.

ALBERTS, B. 2002. *Molecular biology of the cell,* New York, Garland Science.

ALMEIDA, K. H. & SOBOL, R. W. 2007. A unified view of base excision repair: lesion-dependent protein complexes regulated by post-translational modification. *DNA Repair (Amst),* 6**,** 695-711.

APARICIO, O. M., STOUT, A. M. & BELL, S. P. 1999. Differential assembly of Cdc45p and DNA polymerases at early and late origins of DNA replication. *Proc Natl Acad Sci U S A,* 96**,** 9130-5.

ARAVIND, L., LEIPE, D. D. & KOONIN, E. V. 1998. Toprim--a conserved catalytic domain in type IA and II topoisomerases, DnaG-type primases, OLD family nucleases and RecR proteins. *Nucleic Acids Res,* 26**,** 4205-13.

AREZI, B., KIRK, B. W., COPELAND, W. C. & KUCHTA, R. D. 1999. Interactions of DNA with human DNA primase monitored with photoactivatable cross-linking agents: implications for the role of the p58 subunit. *Biochemistry,* 38**,** 12899-907.

ATANASSOVA, N. & GRAINGE, I. 2008. Biochemical characterization of the minichromosome maintenance (MCM) protein of the crenarchaeote Aeropyrum pernix and its interactions with the origin recognition complex (ORC) proteins. *Biochemistry,* 47**,** 13362-70.

AVERY, O. T., MACLEOD, C. M. & MCCARTY, M. 1944. Studies on the

Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from Pneumococcus Type Iii. *Journal of Experimental Medicine,* 79**,** 137-158.

BAKER, R. P. & REHA-KRANTZ, L. J. 1998. Identification of a transient excision intermediate at the crossroads between DNA polymerase extension and proofreading pathways. *Proc Natl Acad Sci U S A,* 95**,** 3507-12.

BAKER, T. A. & BELL, S. P. 1998. Polymerases and the replisome: Machines within machines. *Cell,* 92**,** 295-305.

BAKER, T. A., FUNNELL, B. E. & KORNBERG, A. 1987. Helicase action of dnaB protein during replication from the Escherichia coli chromosomal origin in vitro. *J Biol Chem,* 262**,** 6877-85.

BARNES, D. E., JOHNSTON, L. H., KODAMA, K., TOMKINSON, A. E., LASKO, D. D. & LINDAHL, T. 1990. Human DNA ligase I cDNA: cloning and functional expression in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A,* 87**,** 6679-83.

BARRY, E. R., MCGEOCH, A. T., KELMAN, Z. & BELL, S. D. 2007. Archaeal MCM has separable processivity, substrate choice and helicase domains. *Nucleic Acids Res,* 35**,** 988-98.

BARTLETT, E. J., BRISSETT, N. C. & DOHERTY, A. J. 2013. Ribonucleolytic resection is required for repair of strand displaced nonhomologous end-joining intermediates. *Proc Natl Acad Sci U S A,* 110**,** E1984-91.

BARTLETT, E. J., BRISSETT, N. C., PLOCINSKI, P., CARLBERG, T. & DOHERTY, A. J. 2016. Molecular basis for DNA strand displacement by NHEJ repair polymerases. *Nucleic Acids Res,* 44**,** 2173-86.

BEBENEK, K. & KUNKEL, T. A. 2004. Functions of DNA polymerases. *Adv Protein Chem,* 69**,** 137-65.

BEBLO, K., DOUKI, T., SCHMALZ, G., RACHEL, R., WIRTH, R., HUBER, H., REITZ, G. & RETTBERG, P. 2011. Survival of thermophilic and hyperthermophilic microorganisms after exposure to UV-C, ionizing radiation and desiccation. *Arch Microbiol,* 193**,** 797-809.

BECKER, M. M. & WANG, Z. 1989. Origin of Ultraviolet Damage in DNA. *Journal of Molecular Biology,* 210**,** 429-438.

BEESE, L. S. & STEITZ, T. A. 1991. Structural basis for the 3'-5' exonuclease

activity of Escherichia coli DNA polymerase I: a two metal ion mechanism. *EMBO J,* 10**,** 25-33.

BELL, S. D. 2012. Archaeal orc1/cdc6 proteins. *Subcell Biochem,* 62**,** 59-69.

BELL, S. P. & DUTTA, A. 2002. DNA replication in eukaryotic cells. *Annu Rev Biochem,* 71**,** 333-74.

BENSON, D. A., KARSCH-MIZRACHI, I., LIPMAN, D. J., OSTELL, J. & SAYERS, E. W. 2010. GenBank. *Nucleic Acids Res,* 38**,** D46-51.

BERQUIST, B. R. & WILSON, D. M., 3RD 2012. Pathways for repairing and tolerating the spectrum of oxidative DNA lesions. *Cancer Lett,* 327**,** 61-72.

BERTRAM, J. S. 2000. The molecular biology of cancer. *Mol Aspects Med,* 21**,** 167-223.

BIANCHI, J., RUDD, S. G., JOZWIAKOWSKI, S. K., BAILEY, L. J., SOURA, V., TAYLOR, E., STEVANOVIC, I., GREEN, A. J., STRACKER, T. H., LINDSAY, H. D. & DOHERTY, A. J. 2013. PrimPol bypasses UV photoproducts during eukaryotic chromosomal DNA replication. *Mol Cell,* 52**,** 566-73.

BISWAS, N. & WELLER, S. K. 1999. A mutation in the C-terminal putative Zn2+ finger motif of UL52 severely affects the biochemical activities of the HSV-1 helicase-primase subcomplex. *J Biol Chem,* 274**,** 8068-76.

BISWAS, N. & WELLER, S. K. 1999. A mutation in the C-terminal putative Zn2+ finger motif of UL52 severely affects the biochemical activities of the HSV-1 helicase-primase subcomplex. *J Biol Chem,* 274**,** 8068-76.

BOCQUIER, A. A., LIU, L., CANN, I. K., KOMORI, K., KOHDA, D. & ISHINO, Y. 2001. Archaeal primase: bridging the gap between RNA and DNA polymerases. *Curr Biol,* 11**,** 452-6.

BOSKOVIC, J., COLOMA, J., APARICIO, T., ZHOU, M., ROBINSON, C. V., MENDEZ, J. & MONTOYA, G. 2007. Molecular architecture of the human GINS complex. *EMBO Rep,* 8**,** 678-84.

BOUCHE, J. P., ZECHEL, K. & KORNBERG, A. 1975. dnaG gene product, a rifampicin-resistant RNA polymerase, initiates the conversion of a single-stranded coliphage DNA to its duplex replicative form. *J Biol Chem,* 250**,** 5995-6001.

BOUDSOCQ, F., IWAI, S., HANAOKA, F. & WOODGATE, R. 2001. Sulfolobus solfataricus P2 DNA polymerase IV (Dpo4): an archaeal DinB-like DNA

polymerase with lesion-bypass properties akin to eukaryotic poleta. *Nucleic Acids Res,* 29**,** 4607-16.

BRAITHWAITE, D. K. & ITO, J. 1993. Compilation, alignment, and phylogenetic relationships of DNA polymerases. *Nucleic Acids Res,* 21**,** 787-802.

BRAUN, K. A., LAO, Y., HE, Z., INGLES, C. J. & WOLD, M. S. 1997. Role of protein-protein interactions in the function of replication protein A (RPA): RPA modulates the activity of DNA polymerase alpha by multiple mechanisms. *Biochemistry,* 36**,** 8443-54.

BREWSTER, A. S. & CHEN, X. J. S. 2010. Insights into the MCM functional mechanism: lessons learned from the archaeal MCM complex. *Critical Reviews in Biochemistry and Molecular Biology,* 45**,** 243-256.

BRISSETT, N. C., MARTIN, M. J., PITCHER, R. S., BIANCHI, J., JUAREZ, R., GREEN, A. J., FOX, G. C., BLANCO, L. & DOHERTY, A. J. 2011. Structure of a preternary complex involving a prokaryotic NHEJ DNA polymerase. *Mol Cell,* 41**,** 221-31.

BRISSETT, N. C., PITCHER, R. S., JUAREZ, R., PICHER, A. J., GREEN, A. J., DAFFORN, T. R., FOX, G. C., BLANCO, L. & DOHERTY, A. J. 2007. Structure of a NHEJ polymerase-mediated DNA synaptic complex. *Science,* 318**,** 456-9.

BRODERICK, S., REHMET, K., CONCANNON, C. & NASHEUER, H. P. 2010. Eukaryotic single-stranded DNA binding proteins: central factors in genome stability. *Subcell Biochem,* 50**,** 143-63.

BROSEY, C. A., YAN, C., TSUTAKAWA, S. E., HELLER, W. T., RAMBO, R. P., TAINER, J. A., IVANOV, I. & CHAZIN, W. J. 2013. A new structural framework for integrating replication protein A into DNA processing machinery. *Nucleic Acids Res,* 41**,** 2313-27.

BRUSKOV, V. I. & POLTEV, V. I. 1974. [Recognition by enzymes of nitrogenous base pairs and enhancement of specific interactions in matrix synthesis processes]. *Dokl Akad Nauk SSSR,* 219**,** 231-4.

BRUTLAG, D. & KORNBERG, A. 1972. Enzymatic synthesis of deoxyribonucleic acid. 36. A proofreading function for the 3' leads to 5' exonuclease activity in deoxyribonucleic acid polymerases. *J Biol Chem,* 247**,** 241-8.

BRUTLAG, D., SCHEKMAN, R. & KORNBERG, A. 1971. A possible role for RNA polymerase in the initiation of M13 DNA synthesis. *Proc Natl Acad Sci U*

S A, 68**,** 2826-9.

BURGERS, P. M. 1998. Eukaryotic DNA polymerases in DNA replication and DNA repair. *Chromosoma,* 107**,** 218-27.

BYUN, T. S., PACEK, M., YEE, M. C., WALTER, J. C. & CIMPRICH, K. A. 2005. Functional uncoupling of MCM helicase and DNA polymerase activities activates the ATR-dependent checkpoint. *Genes Dev,* 19**,** 1040-52.

CAIRNS, J. 1966. Autoradiography of HeLa cell DNA. *J Mol Biol,* 15**,** 372-3.

CANN, I. K. & ISHINO, Y. 1999. Archaeal DNA replication: identifying the pieces to solve a puzzle. *Genetics,* 152**,** 1249-67.

CANN, I. K., KOMORI, K., TOH, H., KANAI, S. & ISHINO, Y. 1998. A heterodimeric DNA polymerase: evidence that members of Euryarchaeota possess a distinct DNA polymerase. *Proc Natl Acad Sci U S A,* 95**,** 14250-5.

CAPSON, T. L., PELISKA, J. A., KABOORD, B. F., FREY, M. W., LIVELY, C., DAHLBERG, M. & BENKOVIC, S. J. 1992. Kinetic characterization of the polymerase and exonuclease activities of the gene 43 protein of bacteriophage T4. *Biochemistry,* 31**,** 10984-94.

CARPENTIERI, F., DE FELICE, M., DE FALCO, M., ROSSI, M. & PISANI, F. M. 2002. Physical and functional interaction between the mini-chromosome maintenance-like DNA helicase and the single-stranded DNA binding protein from the crenarchaeon Sulfolobus solfataricus. *Journal of Biological Chemistry,* 277**,** 12118-12127.

CAYROU, C., COULOMBE, P., PUY, A., RIALLE, S., KAPLAN, N., SEGAL, E. & MECHALI, M. 2012. New insights into replication origin characteristics in metazoans. *Cell Cycle,* 11**,** 658-667.

CHAMPOUX, J. J. 2001. DNA topoisomerases: structure, function, and mechanism. *Annu Rev Biochem,* 70**,** 369-413.

CHARGAFF, E., LIPSHITZ, R. & GREEN, C. 1952. Composition of the Desoxypentose Nucleic Acids of 4 Genera of Sea-Urchin. *Journal of Biological Chemistry,* 195**,** 155-160.

CHEMNITZ GALAL, W., PAN, M., GIULIAN, G., YUAN, W., LI, S., EDWARDS, J. L., MARINO, J. P., KELMAN, Z. & HURWITZ, J. 2012. Formation of dAMP-glycerol and dAMP-Tris derivatives by Thermococcus kodakaraensis DNA primase. *J Biol Chem,* 287**,** 16220-9.

CHEMNITZ GALAL, W., PAN, M., KELMAN, Z. & HURWITZ, J. 2012. Characterization of DNA primase complex isolated from the archaeon, Thermococcus kodakaraensis. *J Biol Chem,* 287**,** 16209-19.

CHEN, Y., CARRINGTON-LAWRENCE, S. D., BAI, P. & WELLER, S. K. 2005. Mutations in the putative zinc-binding motif of UL52 demonstrate a complex interdependence between the UL5 and UL52 subunits of the human herpes simplex virus type 1 helicase/primase complex. *J Virol,* 79**,** 9088-96.

CHENG, K. C., CAHILL, D. S., KASAI, H., NISHIMURA, S. & LOEB, L. A. 1992. 8-Hydroxyguanine, an abundant form of oxidative DNA damage, causes G----T and A----C substitutions. *J Biol Chem,* 267**,** 166-72.

COLLINS, K. L. & KELLY, T. J. 1991. Effects of T antigen and replication protein A on the initiation of DNA synthesis by DNA polymerase alpha-primase. *Mol Cell Biol,* 11**,** 2108-15.

CONAWAY, R. C. & LEHMAN, I. R. 1982. Synthesis by the DNA primase of Drosophila melanogaster of a primer with a unique chain length. *Proc Natl Acad Sci U S A,* 79**,** 4585-8.

CONNOLLY, B. A. 2009. Recognition of deaminated bases by archaeal family-B DNA polymerases. *Biochem Soc Trans,* 37**,** 65-8.

COSTA, A., ILVES, I., TAMBERG, N., PETOJEVIC, T., NOGALES, E., BOTCHAN, M. R. & BERGER, J. M. 2011. The structural basis for MCM2-7 helicase activation by GINS and Cdc45. *Nat Struct Mol Biol,* 18**,** 471-7.

CRICK, F. 1970. Central Dogma of Molecular Biology. *Nature,* 227**,** 561-&.

CRICK, F. H., BRENNER, S., WATSTOBI.RJ & BARNETT, L. 1961. General Nature of Genetic Code for Proteins. *Nature,* 192**,** 1227-&.

CRUTE, J. J., BRUCKNER, R. C., DODSON, M. S. & LEHMAN, I. R. 1991. Herpes simplex-1 helicase-primase. Identification of two nucleoside triphosphatase sites that promote DNA helicase action. *J Biol Chem,* 266**,** 21252-6.

CUBONOVA, L., RICHARDSON, T., BURKHART, B. W., KELMAN, Z., CONNOLLY, B. A., REEVE, J. N. & SANTANGELO, T. J. 2013. Archaeal DNA polymerase D but not DNA polymerase B is required for genome replication in Thermococcus kodakarensis. *J Bacteriol,* 195**,** 2322-8.

DARMAWAN, H., HARRISON, M. & REHA-KRANTZ, L. J. 2015. DNA

polymerase 3'-->5' exonuclease activity: Different roles of the beta hairpin structure in family-B DNA polymerases. *DNA Repair (Amst),* 29**,** 36-46.

DE FELICE, M., ESPOSITO, L., PUCCI, B., CARPENTIERI, F., DE FALCO, M., ROSSI, M. & PISANI, F. M. 2003. Biochemical characterization of a CDC6-like protein from the crenarchaeon Sulfolobus solfataricus. *J Biol Chem,* 278**,** 46424-31.

DE SILVA, F. S., PARAN, N. & MOSS, B. 2009. Products and substrate/template usage of vaccinia virus DNA primase. *Virology,* 383**,** 136-41.

DELLA, M., PALMBOS, P. L., TSENG, H. M., TONKIN, L. M., DALEY, J. M., TOPPER, L. M., PITCHER, R. S., TOMKINSON, A. E., WILSON, T. E. & DOHERTY, A. J. 2004. Mycobacterial Ku and ligase proteins constitute a two-component NHEJ repair machine. *Science,* 306**,** 683-5.

DEMPLE, B. & HARRISON, L. 1994. Repair of Oxidative Damage to DNA - Enzymology and Biology. *Annual Review of Biochemistry,* 63**,** 915-948.

DEPAMPHILIS, M. L. 1993. Origins of DNA replication that function in eukaryotic cells. *Curr Opin Cell Biol,* 5**,** 434-41.

DESOGUS, G., ONESTI, S., BRICK, P., ROSSI, M. & PISANI, F. M. 1999. Identification and characterization of a DNA primase from the hyperthermophilic archaeon Methanococcus jannaschii. *Nucleic Acids Res,* 27**,** 4444-50.

DORNREITER, I., ERDILE, L. F., GILBERT, I. U., VON WINKLER, D., KELLY, T. J. & FANNING, E. 1992. Interaction of DNA polymerase alpha-primase with cellular replication protein A and SV40 T antigen. *EMBO J,* 11**,** 769-76.

DRABLOS, F., FEYZI, E., AAS, P. A., VAAGBO, C. B., KAVLI, B., BRATLIE, M. S., PENA-DIAZ, J., OTTERLEI, M., SLUPPHAUG, G. & KROKAN, H. E. 2004. Alkylation damage in DNA and RNA--repair mechanisms and medical significance. *DNA Repair (Amst),* 3**,** 1389-407.

DUDAS, K. C. & KREUZER, K. N. 2001. UvsW protein regulates bacteriophage T4 origin-dependent replication by unwinding R-loops. *Mol Cell Biol,* 21**,** 2706-15.

DUGGIN, I. G. & BELL, S. D. 2006. The chromosome replication machinery of the archaeon Sulfolobus solfataricus. *J Biol Chem,* 281**,** 15029-32.

DUNCAN, B. K. & MILLER, J. H. 1980. Mutagenic Deamination of Cytosine

Residues in DNA. *Nature,* 287**,** 560-561.

DURFEE, T., BECHERER, K., CHEN, P. L., YEH, S. H., YANG, Y., KILBURN, A. E., LEE, W. H. & ELLEDGE, S. J. 1993. The retinoblastoma protein associates with the protein phosphatase type 1 catalytic subunit. *Genes Dev,* 7**,** 555-69.

ECHOLS, H. & GOODMAN, M. F. 1991. Fidelity Mechanisms in DNA-Replication. *Annual Review of Biochemistry,* 60**,** 477-511.

EDGELL, D. R. & DOOLITTLE, W. F. 1997. Archaea and the origin(s) of DNA replication proteins. *Cell,* 89**,** 995-8.

EMPTAGE, K., O'NEILL, R., SOLOVYOVA, A. & CONNOLLY, B. A. 2008. Interplay between DNA polymerase and proliferating cell nuclear antigen switches off base excision repair of uracil and hypoxanthine during replication in archaea. *J Mol Biol,* 383**,** 762-71.

ENGEL, J. D. & VON HIPPEL, P. H. 1978. Effects of methylation on the stability of nucleic acid conformations. Studies at the polymer level. *J Biol Chem,* 253**,** 927-34.

FAN, J. & PAVLETICH, N. P. 2012. Structure and conformational change of a replication protein A heterotrimer bound to ssDNA. *Genes Dev,* 26**,** 2337-47.

FIELDS, S. & SONG, O. 1989. A novel genetic system to detect protein-protein interactions. *Nature,* 340**,** 245-6.

FIRBANK, S. J., WARDLE, J., HESLOP, P., LEWIS, R. J. & CONNOLLY, B. A. 2008. Uracil recognition in archaeal DNA polymerases captured by X-ray crystallography. *J Mol Biol,* 381**,** 529-39.

FIRBANK, S. J., WARDLE, J., HESLOP, P., LEWIS, R. J. & CONNOLLY, B. A. 2008. Uracil recognition in archaeal DNA polymerases captured by X-ray crystallography. *J Mol Biol,* 381**,** 529-39.

FLETCHER, R. J., BISHOP, B. E., LEON, R. P., SCLAFANI, R. A., OGATA, C. M. & CHEN, X. S. 2003. The structure and function of MCM from archaeal M. Thermoautotrophicum. *Nat Struct Biol,* 10**,** 160-7.

FORTINI, P., PARLANTI, E., SIDORKINA, O. M., LAVAL, J. & DOGLIOTTI, E. 1999. The type of DNA glycosylase determines the base excision repair pathway in mammalian cells. *J Biol Chem,* 274**,** 15230-6.

FORTINI, P., PASCUCCI, B., PARLANTI, E., SOBOL, R. W., WILSON, S. H. &

DOGLIOTTI, E. 1998. Different DNA polymerases are involved in the short- and long-patch base excision repair in mammalian cells. *Biochemistry,* 37**,** 3575-80.

FRICK, D. N., KUMAR, S. & RICHARDSON, C. C. 1999. Interaction of ribonucleoside triphosphates with the gene 4 primase of bacteriophage T7. *J Biol Chem,* 274**,** 35899-907.

FRICK, D. N. & RICHARDSON, C. C. 2001. DNA primases. *Annu Rev Biochem,* 70**,** 39-80.

FRIEDBERG, E. C., LEHMANN, A. R. & FUCHS, R. P. P. 2005. Trading places: How do DNA polymerases switch during translesion DNA synthesis? (vol 18, pg 499, 2005). *Molecular Cell,* 19**,** 143-143.

FUMASONI, M., ZWICKY, K., VANOLI, F., LOPES, M. & BRANZEI, D. 2015. Error-free DNA damage tolerance and sister chromatid proximity during DNA replication rely on the Polalpha/Primase/Ctf4 Complex. *Mol Cell,* 57**,** 812-23.

GALAL, W. C., PAN, M., KELMAN, Z. & HURWITZ, J. 2012. Characterization of DNA Primase Complex Isolated from the Archaeon, Thermococcus kodakaraensis. *Journal of Biological Chemistry,* 287**,** 16209-16219.

GAMBUS, A., VAN DEURSEN, F., POLYCHRONOPOULOS, D., FOLTMAN, M., JONES, R. C., EDMONDSON, R. D., CALZADA, A. & LABIB, K. 2009. A key role for Ctf4 in coupling the MCM2-7 helicase to DNA polymerase alpha within the eukaryotic replisome. *EMBO J,* 28**,** 2992-3004.

GAME, J. C. & MORTIMER, R. K. 1974. A genetic study of x-ray sensitive mutants in yeast. *Mutat Res,* 24**,** 281-92.

GARCIA-GOMEZ, S., REYES, A., MARTINEZ-JIMENEZ, M. I., CHOCRON, E. S., MOURON, S., TERRADOS, G., POWELL, C., SALIDO, E., MENDEZ, J., HOLT, I. J. & BLANCO, L. 2013. PrimPol, an archaic primase/polymerase operating in human cells. *Mol Cell,* 52**,** 541-53.

GARG, P., STITH, C. M., SABOURI, N., JOHANSSON, E. & BURGERS, P. M. 2004. Idling by DNA polymerase delta maintains a ligatable nick during lagging-strand DNA replication. *Genes Dev,* 18**,** 2764-73.

GARY, R., PARK, M. S., NOLAN, J. P., CORNELIUS, H. L., KOZYREVA, O. G., TRAN, H. T., LOBACHEV, K. S., RESNICK, M. A. & GORDENIN, D. A. 1999. A novel role in DNA metabolism for the binding of Fen1/Rad27 to

PCNA and implications for genetic risk. *Mol Cell Biol,* 19**,** 5373-82.

GERALD, D., BERRA, E., FRAPART, Y. M., CHAN, D. A., GIACCIA, A. J., MANSUY, D., POUYSSEGUR, J., YANIV, M. & MECHTA-GRIGORIOU, F. 2004. JunD reduces tumor angiogenesis by protecting cells from oxidative stress. *Cell,* 118**,** 781-94.

GILL, S., KRUPOVIC, M., DESNOUES, N., BEGUIN, P., SEZONOV, G. & FORTERRE, P. 2014. A highly divergent archaeo-eukaryotic primase from the Thermococcus nautilus plasmid, pTN2. *Nucleic Acids Res,* 42**,** 3707-19.

GORBALENYA, A. E., KOONIN, E. V., DONCHENKO, A. P. & BLINOV, V. M. 1989. Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. *Nucleic Acids Res,* 17**,** 4713-30.

GOSWAMI, K., ARORA, J. & SAHA, S. 2015. Characterization of the MCM homohexamer from the thermoacidophilic euryarchaeon Picrophilus torridus. *Sci Rep,* 5**,** 9057.

GOTZ, D., PAYTUBI, S., MUNRO, S., LUNDGREN, M., BERNANDER, R. & WHITE, M. F. 2007. Responses of hyperthermophilic crenarchaea to UV irradiation. *Genome Biol,* 8**,** R220.

GRABOWSKI, B. & KELMAN, Z. 2003. Archeal DNA replication: eukaryal proteins in a bacterial context. *Annu Rev Microbiol,* 57**,** 487-516.

GRADIA, S., ACHARYA, S. & FISHEL, R. 2000. The role of mismatched nucleotides in activating the hMSH2-hMSH6 molecular switch. *J Biol Chem,* 275**,** 3922-30.

GRAINGE, I., SCAIFE, S. & WIGLEY, D. B. 2003. Biochemical analysis of components of the pre-replication complex of Archaeoglobus fulgidus. *Nucleic Acids Res,* 31**,** 4888-98.

GRAZIEWICZ, M. A., BIENSTOCK, R. J. & COPELAND, W. C. 2007. The DNA polymerase gamma Y955C disease variant associated with PEO and parkinsonism mediates the incorporation and translesion synthesis opposite 7,8-dihydro-8-oxo-2'-deoxyguanosine. *Hum Mol Genet,* 16**,** 2729-39.

GREAGG, M. A., FOGG, M. J., PANAYOTOU, G., EVANS, S. J., CONNOLLY, B. A. & PEARL, L. H. 1999. A read-ahead function in archaeal DNA

polymerases detects promutagenic template-strand uracil. *Proc Natl Acad Sci U S A,* 96**,** 9045-50.

GREENOUGH, L., KELMAN, Z. & GARDNER, A. F. 2015. The roles of family B and D DNA polymerases in Thermococcus species 9 degrees N Okazaki fragment maturation. *J Biol Chem,* 290**,** 12514-22.

GRIFFITH, F. 1928. The Significance of Pneumococcal Types. *Journal of Hygiene,* 27**,** 113-159.

GUILLIAM, T. A., JOZWIAKOWSKI, S. K., EHLINGER, A., BARNES, R. P., RUDD, S. G., BAILEY, L. J., SKEHEL, J. M., ECKERT, K. A., CHAZIN, W. J. & DOHERTY, A. J. 2015. Human PrimPol is a highly error-prone polymerase regulated by single-stranded DNA binding proteins. *Nucleic Acids Res,* 43**,** 1056-68.

GUILLIAM, T. A., KEEN, B. A., BRISSETT, N. C. & DOHERTY, A. J. 2015. Primase-polymerases are a functionally diverse superfamily of replication and repair enzymes. *Nucleic Acids Res,* 43**,** 6651-64.

GUO, C., KOSAREK-STANCEL, J. N., TANG, T. S. & FRIEDBERG, E. C. 2009. Y-family DNA polymerases in mammalian cells. *Cell Mol Life Sci,* 66**,** 2363-81.

HADDOW, A. 1973. On the biological alkylating agents. *Perspect Biol Med,* 16**,** 503-24.

HALGASOVA, N., MESAROSOVA, I. & BUKOVSKA, G. 2012. Identification of a bifunctional primase-polymerase domain of corynephage BFK20 replication protein gp43. *Virus Res,* 163**,** 454-60.

HALL, D. B., HOLMLIN, R. E. & BARTON, J. K. 1996. Oxidative DNA damage through long-range electron transfer. *Nature,* 382**,** 731-5.

HALLWELL, B. & ARUOMA, O. I. 1991. DNA damage by oxygen-derived species. Its mechanism and measurement in mammalian systems. *FEBS Lett,* 281**,** 9-19.

HAUGLAND, G. T., SHIN, J. H., BIRKELAND, N. K. & KELMAN, Z. 2006. Stimulation of MCM helicase activity by a Cdc6 protein in the archaeon Thermoplasma acidophilum. *Nucleic Acids Res,* 34**,** 6337-44.

HAWKINS, M., MALLA, S., BLYTHE, M. J., NIEDUSZYNSKI, C. A. & ALLERS, T. 2013. Accelerated growth in the absence of DNA replication origins. *Nature,* 503**,** 544-7.

HELLEDAY, T., LO, J., VAN GENT, D. C. & ENGELWARD, B. P. 2007. DNA double-strand break repair: from mechanistic understanding to cancer treatment. *DNA Repair (Amst),* 6**,** 923-35.

HELLER, R. C. & MARIANS, K. J. 2006. Replication fork reactivation downstream of a blocked nascent leading strand. *Nature,* 439**,** 557-62.

HENNEKE, G., FLAMENT, D., HUBSCHER, U., QUERELLOU, J. & RAFFIN, J. P. 2005. The hyperthermophilic euryarchaeota Pyrococcus abyssi likely requires the two DNA polymerases D and B for DNA replication. *J Mol Biol,* 350**,** 53-64.

HERSHEY, A. D. & CHASE, M. 1952. Independent Functions of Viral Protein and Nucleic Acid in Growth of Bacteriophage. *Journal of General Physiology,* 36**,** 39-56.

HESS, M. T., GUNZ, D., LUNEVA, N., GEACINTOV, N. E. & NAEGELI, H. 1997. Base pair conformation-dependent excision of benzo[a]pyrene diol epoxide-guanine adducts by human nucleotide excision repair enzymes. *Mol Cell Biol,* 17**,** 7069-76.

HITOMI, H., KIYOMOTO, H. & NISHIYAMA, A. 2007. Angiotensin II and oxidative stress. *Curr Opin Cardiol,* 22**,** 311-5.

HOEIJMAKERS, J. H. 2009. DNA damage, aging, and cancer. *N Engl J Med,* 361**,** 1475-85.

HOFFMANN, G. R. 1980. Genetic effects of dimethyl sulfate, diethyl sulfate, and related compounds. *Mutat Res,* 75**,** 63-129.

HU, J., GUO, L., WU, K., LIU, B., LANG, S. & HUANG, L. 2012. Template-dependent polymerization across discontinuous templates by the heterodimeric primase from the hyperthermophilic archaeon Sulfolobus solfataricus. *Nucleic Acids Res,* 40**,** 3470-83.

HUANG, J. C., SVOBODA, D. L., REARDON, J. T. & SANCAR, A. 1992. Human nucleotide excision nuclease removes thymine dimers from DNA by incising the 22nd phosphodiester bond 5' and the 6th phosphodiester bond 3' to the photodimer. *Proc Natl Acad Sci U S A,* 89**,** 3664-8.

HUBERMAN, J. A. & RIGGS, A. D. 1968. On the mechanism of DNA replication in mammalian chromosomes. *J Mol Biol,* 32**,** 327-41.

HUBSCHER, U. & MAGA, G. 2011. DNA replication and repair bypass machines. *Curr Opin Chem Biol,* 15**,** 627-35.

HUBSCHER, U. & MAGA, G. 2011. DNA replication and repair bypass machines. *Current Opinion in Chemical Biology,* 15**,** 627-635.

IFTODE, C., DANIELY, Y. & BOROWIEC, J. A. 1999. Replication protein A (RPA): The eukaryotic SSB. *Critical Reviews in Biochemistry and Molecular Biology,* 34**,** 141-180.

ILYINA, T. V., GORBALENYA, A. E. & KOONIN, E. V. 1992. Organization and evolution of bacterial and bacteriophage primase-helicase systems. *J Mol Evol,* 34**,** 351-7.

IMAMURA, M., UEMORI, T., KATO, I. & ISHINO, Y. 1995. A non-alpha-like DNA polymerase from the hyperthermophilic archaeon Pyrococcus furiosus. *Biol Pharm Bull,* 18**,** 1647-52.

INDIANI, C., PATEL, M., GOODMAN, M. F. & O'DONNELL, M. E. 2013. RecA acts as a switch to regulate polymerase occupancy in a moving replication fork. *Proc Natl Acad Sci U S A,* 110**,** 5410-5.

INOUYE, M., ARNHEIM, N. & STERNGLANZ, R. 1973. Bacteriophage T7 lysozyme is an N-acetylmuramyl-L-alanine amidase. *J Biol Chem,* 248**,** 7247-52.

ISHINO, S., FUJINO, S., TOMITA, H., OGINO, H., TAKAO, K., DAIYASU, H., KANAI, T., ATOMI, H. & ISHINO, Y. 2011. Biochemical and genetical analyses of the three mcm genes from the hyperthermophilic archaeon, Thermococcus kodakarensis. *Genes Cells,* 16**,** 1176-89.

ISHINO, S. & ISHINO, Y. 2006. Comprehensive search for DNA polymerase in the hyperthermophilic archaeon, Pyrococcus furiosus. *Nucleosides Nucleotides Nucleic Acids,* 25**,** 681-91.

ISHINO, S., KELMAN, L. M., KELMAN, Z. & ISHINO, Y. 2013. The archaeal DNA replication machinery: past, present and future. *Genes Genet Syst,* 88**,** 315-9.

IYER, L. M., KOONIN, E. V., LEIPE, D. D. & ARAVIND, L. 2005. Origin and evolution of the archaeo-eukaryotic primase superfamily and related palm-domain proteins: structural insights and new members. *Nucleic Acids Res,* 33**,** 3875-96.

IYER, R. R., PLUCIENNIK, A., BURDETT, V. & MODRICH, P. L. 2006. DNA mismatch repair: functions and mechanisms. *Chem Rev,* 106**,** 302-23.

JOHNSON, A. A. & JOHNSON, K. A. 2001. Exonuclease proofreading by human

mitochondrial DNA polymerase. *J Biol Chem,* 276**,** 38097-107.

JOHNSON, R. E., PRAKASH, S. & PRAKASH, L. 1999. Efficient bypass of a thymine-thymine dimer by yeast DNA polymerase, Poleta. *Science,* 283**,** 1001-4.

JOHNSON, R. E., WASHINGTON, M. T., HARACSKA, L., PRAKASH, S. & PRAKASH, L. 2000. Eukaryotic polymerases iota and zeta act sequentially to bypass DNA lesions. *Nature,* 406**,** 1015-9.

JOHNSON, S. J. & BEESE, L. S. 2004. Structures of mismatch replication errors observed in a DNA polymerase. *Cell,* 116**,** 803-16.

JONSSON, Z. O., HINDGES, R. & HUBSCHER, U. 1998. Regulation of DNA replication and repair proteins through interaction with the front side of proliferating cell nuclear antigen. *Embo Journal,* 17**,** 2412-2425.

JOYCE, C. M. & BENKOVIC, S. J. 2004. DNA polymerase fidelity: kinetics, structure, and checkpoints. *Biochemistry,* 43**,** 14317-24.

JOZWIAKOWSKI, S. K., BORAZJANI GHOLAMI, F. & DOHERTY, A. J. 2015. Archaeal replicative primases can perform translesion DNA synthesis. *Proc Natl Acad Sci U S A,* 112**,** E633-8.

JOZWIAKOWSKI, S. K. & CONNOLLY, B. A. 2011. A modified family-B archaeal DNA polymerase with reverse transcriptase activity. *Chembiochem,* 12**,** 35-7.

KAGUNI, J. M. 2006. DnaA: Controlling the initiation of bacterial DNA replication and more. *Annual Review of Microbiology,* 60**,** 351-371.

KAGUNI, J. M. 2011. Replication initiation at the Escherichia coli chromosomal origin. *Curr Opin Chem Biol,* 15**,** 606-13.

KAMADA, K. 2012. The GINS complex: structure and function. *Subcell Biochem,* 62**,** 135-56.

KAPLAN, D. L. & STEITZ, T. A. 1999. DnaB from Thermus aquaticus unwinds forked duplex DNA with an asymmetric tail length dependence. *J Biol Chem,* 274**,** 6889-97.

KASIVISWANATHAN, R., SHIN, J. H. & KELMAN, Z. 2005. Interactions between the archaeal Cdc6 and MCM proteins modulate their biochemical properties. *Nucleic Acids Research,* 33**,** 4940-4950.

KAZLAUSKAS, D. & VENCLOVAS, C. 2014. Herpesviral helicase-primase subunit UL8 is inactivated B-family polymerase. *Bioinformatics,* 30**,** 2093-

7.

KEEN, B. A., JOZWIAKOWSKI, S. K., BAILEY, L. J., BIANCHI, J. & DOHERTY, A. J. 2014. Molecular dissection of the domain architecture and catalytic activities of human PrimPol. *Nucleic Acids Res,* 42**,** 5830-45.

KELLY, T. J., SIMANCEK, P. & BRUSH, G. S. 1998. Identification and characterization of a single-stranded DNA-binding protein from the archaeon Methanococcus jannaschii. *Proc Natl Acad Sci U S A,* 95**,** 14634-9.

KELMAN, L. M. & KELMAN, Z. 2003. Archaea: an archetype for replication initiation studies? *Mol Microbiol,* 48**,** 605-15.

KELMAN, L. M. & KELMAN, Z. 2014. Archaeal DNA replication. *Annu Rev Genet,* 48**,** 71-97.

KELMAN, Z. 1997. PCNA: structure, functions and interactions. *Oncogene,* 14**,** 629-40.

KELMAN, Z., LEE, J. K. & HURWITZ, J. 1999. The single minichromosome maintenance protein of Methanobacterium thermoautotrophicum DeltaH contains DNA helicase activity. *Proc Natl Acad Sci U S A,* 96**,** 14783-8.

KELMAN, Z. & WHITE, M. F. 2005. Archaeal DNA replication and repair. *Curr Opin Microbiol,* 8**,** 669-76.

KELMAN, Z., YUZHAKOV, A., ANDJELKOVIC, J. & O'DONNELL, M. 1998. Devoted to the lagging strand - the chi subunit of DNA polymerase III holoenzyme contacts SSB to promote processive elongation and sliding clamp assembly. *Embo Journal,* 17**,** 2436-2449.

KENNY, M. K., LEE, S. H. & HURWITZ, J. 1989. Multiple functions of human single-stranded-DNA binding protein in simian virus 40 DNA replication: single-strand stabilization and stimulation of DNA polymerases alpha and delta. *Proc Natl Acad Sci U S A,* 86**,** 9757-61.

KENT, T., CHANDRAMOULY, G., MCDEVITT, S. M., OZDEMIR, A. Y. & POMERANTZ, R. T. 2015. Mechanism of microhomology-mediated end-joining promoted by human DNA polymerase theta. *Nat Struct Mol Biol,* 22**,** 230-7.

KERNCHEN, U. & LIPPS, G. 2006. Thermodynamic analysis of the single-stranded DNA binding activity of the archaeal replication protein A (RPA) from Sulfolobus solfataricus. *Biochemistry,* 45**,** 594-603.

KINO, K. & SUGIYAMA, H. 2001. Possible cause of G-C-->C-G transversion mutation by guanine oxidation product, imidazolone. *Chem Biol,* 8**,** 369-78.

KINO, K. & SUGIYAMA, H. 2005. UVR-induced G-C to C-G transversions from oxidative DNA damage. *Mutat Res,* 571**,** 33-42.

KIRK, B. W. & KUCHTA, R. D. 1999. Arg304 of human DNA primase is a key contributor to catalysis and NTP binding: primase and the family X polymerases share significant sequence homology. *Biochemistry,* 38**,** 7727-36.

KLAPSTEIN, K., CHOU, T. & BRUINSMA, R. 2004. Physics of RecA-mediated homologous recognition. *Biophys J,* 87**,** 1466-77.

KLENK, H. P., CLAYTON, R. A., TOMB, J. F., WHITE, O., NELSON, K. E., KETCHUM, K. A., DODSON, R. J., GWINN, M., HICKEY, E. K., PETERSON, J. D., RICHARDSON, D. L., KERLAVAGE, A. R., GRAHAM, D. E., KYRPIDES, N. C., FLEISCHMANN, R. D., QUACKENBUSH, J., LEE, N. H., SUTTON, G. G., GILL, S., KIRKNESS, E. F., DOUGHERTY, B. A., MCKENNEY, K., ADAMS, M. D., LOFTUS, B., PETERSON, S., REICH, C. I., MCNEIL, L. K., BADGER, J. H., GLODEK, A., ZHOU, L., OVERBEEK, R., GOCAYNE, J. D., WEIDMAN, J. F., MCDONALD, L., UTTERBACK, T., COTTON, M. D., SPRIGGS, T., ARTIACH, P., KAINE, B. P., SYKES, S. M., SADOW, P. W., D'ANDREA, K. P., BOWMAN, C., FUJII, C., GARLAND, S. A., MASON, T. M., OLSEN, G. J., FRASER, C. M., SMITH, H. O., WOESE, C. R. & VENTER, J. C. 1997. The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon Archaeoglobus fulgidus. *Nature,* 390**,** 364-70.

KLENOW, H. & OVERGAARD-HANSEN, K. 1970. Proteolytic cleavage of DNA polymerase from Escherichia Coli B into an exonuclease unit and a polymerase unit. *FEBS Lett,* 6**,** 25-27.

KOBAYASHI, K., GUILLIAM, T. A., TSUDA, M., YAMAMOTO, J., BAILEY, L. J., IWAI, S., TAKEDA, S., DOHERTY, A. J. & HIROTA, K. 2016. Repriming by PrimPol is critical for DNA replication restart downstream of lesions and chain-terminating nucleosides. *Cell Cycle,* 15**,** 1997-2008.

KOGOMA, T. 1997. Stable DNA replication: interplay between DNA replication, homologous recombination, and transcription. *Microbiol Mol Biol Rev,* 61**,**

212-38.

KOGOMA, T., CADWELL, G. W., BARNARD, K. G. & ASAI, T. 1996. The DNA replication priming protein, PriA, is required for homologous recombination and double-strand break repair. *J Bacteriol,* 178**,** 1258-64.

KOLLER, P. C. 1958. Comparative effects of alkylating agents on cellular morphology. *Ann N Y Acad Sci,* 68**,** 783-801.

KOLPASHCHIKOV, D. M., KHODYREVA, S. N., KHLIMANKOV, D. Y., WOLD, M. S., FAVRE, A. & LAVRIK, O. I. 2001. Polarity of human replication protein A binding to DNA. *Nucleic Acids Res,* 29**,** 373-9.

KOMORI, K. & ISHINO, Y. 2001. Replication protein A in Pyrococcus furiosus is involved in homologous DNA recombination. *J Biol Chem,* 276**,** 25654-60.

KOONIN, E. V., WOLF, Y. I., KONDRASHOV, A. S. & ARAVIND, L. 2000. Bacterial homologs of the small subunit of eukaryotic DNA primase. *J Mol Microbiol Biotechnol,* 2**,** 509-12.

KRASTANOVA, I., SANNINO, V., AMENITSCH, H., GILEADI, O., PISANI, F. M. & ONESTI, S. 2012. Structural and functional insights into the DNA replication factor Cdc45 reveal an evolutionary relationship to the DHH family of phosphoesterases. *J Biol Chem,* 287**,** 4121-8.

KROKAN, H. E., DRABLOS, F. & SLUPPHAUG, G. 2002. Uracil in DNA--occurrence, consequences and repair. *Oncogene,* 21**,** 8935-48.

KROKAN, H. E., STANDAL, R. & SLUPPHAUG, G. 1997. DNA glycosylases in the base excision repair of DNA. *Biochemical Journal,* 325**,** 1-16.

KUCHTA, R. D. & STENGEL, G. 2010. Mechanism and evolution of DNA primases. *Biochim Biophys Acta,* 1804**,** 1180-9.

KUKIMOTO, I., IGAKI, H. & KANDA, T. 1999. Human CDC45 protein binds to minichromosome maintenance 7 protein and the p70 subunit of DNA polymerase alpha. *Eur J Biochem,* 265**,** 936-43.

KULAEVA, O. I., KOONIN, E. V., MCDONALD, J. P., RANDALL, S. K., RABINOVICH, N., CONNAUGHTON, J. F., LEVINE, A. S. & WOODGATE, R. 1996. Identification of a DinB/UmuC homolog in the archeon Sulfolobus solfataricus. *Mutat Res,* 357**,** 245-53.

KUNKEL, T. A. & BURGERS, P. M. 2008. Dividing the workload at a eukaryotic replication fork. *Trends Cell Biol,* 18**,** 521-7.

KURAOKA, I., KOBERTZ, W. R., ARIZA, R. R., BIGGERSTAFF, M.,

ESSIGMANN, J. M. & WOOD, R. D. 2000. Repair of an interstrand DNA cross-link initiated by ERCC1-XPF repair/recombination nuclease. *J Biol Chem,* 275**,** 26632-6.

LABIB, K. & GAMBUS, A. 2007. A key role for the GINS complex at DNA replication forks. *Trends in Cell Biology,* 17**,** 271-278.

LAEMMLI, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature,* 227**,** 680-5.

LANG, S. & HUANG, L. 2015. The Sulfolobus solfataricus GINS Complex Stimulates DNA Binding and Processive DNA Unwinding by Minichromosome Maintenance Helicase. *J Bacteriol,* 197**,** 3409-20.

LAO-SIRIEIX, S. H. & BELL, S. D. 2004. The heterodimeric primase of the hyperthermophilic archaeon Sulfolobus solfataricus possesses DNA and RNA primase, polymerase and 3'-terminal nucleotidyl transferase activities. *J Mol Biol,* 344**,** 1251-63.

LE BRETON, M., HENNEKE, G., NORAIS, C., FLAMENT, D., MYLLYKALLIO, H., QUERELLOU, J. & RAFFIN, J. P. 2007. The heterodimeric primase from the euryarchaeon Pyrococcus abyssi: a multifunctional enzyme for initiation and repair? *J Mol Biol,* 374**,** 1172-85.

LEHMAN, I. R. 1974. DNA ligase: structure, mechanism, and function. *Science,* 186**,** 790-7.

LEHMAN, I. R., ZIMMERMAN, S. B., ADLER, J., BESSMAN, M. J., SIMMS, E. S. & KORNBERG, A. 1958. Enzymatic Synthesis of Deoxyribonucleic Acid. V. Chemical Composition of Enzymatically Synthesized Deoxyribonucleic Acid. *Proc Natl Acad Sci U S A,* 44**,** 1191-6.

LEHMANN, A. R. 1972. Postreplication repair of DNA in ultraviolet-irradiated mammalian cells. *J Mol Biol,* 66**,** 319-37.

LEHMANN, A. R., NIIMI, A., OGI, T., BROWN, S., SABBIONEDA, S., WING, J. F., KANNOUCHE, P. L. & GREEN, C. M. 2007. Translesion synthesis: Y-family polymerases and the polymerase switch. *DNA Repair (Amst),* 6**,** 891-9.

LEIPE, D. D., ARAVIND, L. & KOONIN, E. V. 1999. Did DNA replication evolve twice independently? *Nucleic Acids Res,* 27**,** 3389-401.

LI, X. & BURGERS, P. M. 1994. Molecular cloning and expression of the Saccharomyces cerevisiae RFC3 gene, an essential component of

replication factor C. *Proc Natl Acad Sci U S A,* 91**,** 868-72.

LI, X. & HEYER, W. D. 2008. Homologous recombination in DNA repair and DNA damage tolerance. *Cell Res,* 18**,** 99-113.

LI, Y., KOROLEV, S. & WAKSMAN, G. 1998. Crystal structures of open and closed forms of binary and ternary complexes of the large fragment of Thermus aquaticus DNA polymerase I: structural basis for nucleotide incorporation. *Embo Journal,* 17**,** 7514-7525.

LI, Z., PAN, M., SANTANGELO, T. J., CHEMNITZ, W., YUAN, W., EDWARDS, J. L., HURWITZ, J., REEVE, J. N. & KELMAN, Z. 2011. A novel DNA nuclease is stimulated by association with the GINS complex. *Nucleic Acids Res,* 39**,** 6114-23.

LI, Z., SANTANGELO, T. J., CUBONOVA, L., REEVE, J. N. & KELMAN, Z. 2010. Affinity purification of an archaeal DNA replication protein network. *MBio,* 1.

LIN, Y., LIN, L. J., SRIRATANA, P., COLEMAN, K., HA, T., SPIES, M. & CANN, I. K. 2008. Engineering of functional replication protein a homologs based on insights into the evolution of oligonucleotide/oligosaccharide-binding folds. *J Bacteriol,* 190**,** 5766-80.

LINDAHL, T. 1993. Instability and decay of the primary structure of DNA. *Nature,* 362**,** 709-15.

LINDAHL, T. & NYBERG, B. 1974. Heat-induced deamination of cytosine residues in deoxyribonucleic acid. *Biochemistry,* 13**,** 3405-10.

LING, H., BOUDSOCQ, F., PLOSKY, B. S., WOODGATE, R. & YANG, W. 2003. Replication of a cis-syn thymine dimer at atomic resolution. *Nature,* 424**,** 1083-7.

LIPPS, G., WEINZIERL, A. O., VON SCHEVEN, G., BUCHEN, C. & CRAMER, P. 2004. Structure of a bifunctional DNA primase-polymerase. *Nat Struct Mol Biol,* 11**,** 157-62.

LIPSCOMB, L. A., PEEK, M. E., MORNINGSTAR, M. L., VERGHIS, S. M., MILLER, E. M., RICH, A., ESSIGMANN, J. M. & WILLIAMS, L. D. 1995. X-ray structure of a DNA decamer containing 7,8-dihydro-8-oxoguanine. *Proc Natl Acad Sci U S A,* 92**,** 719-23.

LIU, B., OUYANG, S., MAKAROVA, K. S., XIA, Q., ZHU, Y., LI, Z., GUO, L., KOONIN, E. V., LIU, Z. J. & HUANG, L. 2015. A primase subunit essential

for efficient primer synthesis by an archaeal eukaryotic-type primase. *Nat Commun,* 6**,** 7300.

LIU, H., RUDOLF, J., JOHNSON, K. A., MCMAHON, S. A., OKE, M., CARTER, L., MCROBBIE, A. M., BROWN, S. E., NAISMITH, J. H. & WHITE, M. F. 2008. Structure of the DNA repair helicase XPD. *Cell,* 133**,** 801-12.

LIU, L. & HUANG, M. 2015. Essential role of the iron-sulfur cluster binding domain of the primase regulatory subunit Pri2 in DNA replication initiation. *Protein Cell,* 6**,** 194-210.

LIU, L., KOMORI, K., ISHINO, S., BOCQUIER, A. A., CANN, I. K., KOHDA, D. & ISHINO, Y. 2001. The archaeal DNA primase: biochemical characterization of the p41-p46 complex from Pyrococcus furiosus. *J Biol Chem,* 276**,** 45484-90.

LOEB, L. A. & KUNKEL, T. A. 1982. Fidelity of DNA synthesis. *Annu Rev Biochem,* 51**,** 429-57.

LONGHESE, M. P., JOVINE, L., PLEVANI, P. & LUCCHINI, G. 1993. Conditional mutations in the yeast DNA primase genes affect different aspects of DNA metabolism and interactions in the DNA polymerase alpha-primase complex. *Genetics,* 133**,** 183-91.

LOPES, M., FOIANI, M. & SOGO, J. M. 2006. Multiple mechanisms control chromosome integrity after replication fork uncoupling and restart at irreparable UV lesions. *Mol Cell,* 21**,** 15-27.

LOVETT, S. T. 2011. The DNA Exonucleases of Escherichia coli. *EcoSal Plus,* 4.

LUNDGREN, M., ANDERSSON, A., CHEN, L., NILSSON, P. & BERNANDER, R. 2004. Three replication origins in Sulfolobus species: synchronous initiation of chromosome replication and asynchronous termination. *Proc Natl Acad Sci U S A,* 101**,** 7046-51.

LYAMICHEV, V. 1991. Unusual conformation of (dA)n.(dT)n-tracts as revealed by cyclobutane thymine-thymine dimer formation. *Nucleic Acids Res,* 19**,** 4491-6.

MACHWE, A., XIAO, L. R., GRODEN, J. & ORREN, D. K. 2006. The Werner and Bloom syndrome proteins catalyze regression of a model replication fork. *Biochemistry,* 45**,** 13939-13946.

MACKIEWICZ, P., ZAKRZEWSKA-CZERWINSKA, J., ZAWILAK, A., DUDEK, M.

R. & CEBRAT, S. 2004. Where does bacterial replication start? Rules for predicting the oriC region. *Nucleic Acids Research,* 32**,** 3781-3791.

MACNEILL, S. A. 2011. Protein-protein interactions in the archaeal core replisome. *Biochemical Society Transactions,* 39**,** 163-168.

MAGA, G., VILLANI, G., CRESPAN, E., WIMMER, U., FERRARI, E., BERTOCCI, B. & HUBSCHER, U. 2007. 8-oxo-guanine bypass by human DNA polymerases in the presence of auxiliary proteins. *Nature,* 447**,** 606-8.

MAK, J. & KLEIMAN, L. 1997. Primer tRNAs for reverse transcription. *J Virol,* 71**,** 8087-95.

MAKAROVA, K. S., KOONIN, E. V. & KELMAN, Z. 2012. The CMG (CDC45/RecJ, MCM, GINS) complex is a conserved component of the DNA replication system in all archaea and eukaryotes. *Biol Direct,* 7**,** 7.

MARINSEK, N., BARRY, E. R., MAKAROVA, K. S., DIONNE, I., KOONIN, E. V. & BELL, S. D. 2006. GINS, a central nexus in the archaeal DNA replication fork. *EMBO Rep,* 7**,** 539-45.

MARSDEN, H. S., MCLEAN, G. W., BARNARD, E. C., FRANCIS, G. J., MACEACHRAN, K., MURPHY, M., MCVEY, G., CROSS, A., ABBOTTS, A. P. & STOW, N. D. 1997. The catalytic subunit of the DNA polymerase of herpes simplex virus type 1 interacts specifically with the C terminus of the UL8 component of the viral helicase-primase complex. *J Virol,* 71**,** 6390-7.

MASAI, H. 2013. A personal reflection on the replicon theory: from R1 plasmid to replication timing regulation in human cells. *J Mol Biol,* 425**,** 4663-72.

MASAI, H., MATSUMOTO, S., YOU, Z., YOSHIZAWA-SUGATA, N. & ODA, M. 2010. Eukaryotic chromosome DNA replication: where, when, and how? *Annu Rev Biochem,* 79**,** 89-130.

MASUTANI, C., KUSUMOTO, R., YAMADA, A., DOHMAE, N., YOKOI, M., YUASA, M., ARAKI, M., IWAI, S., TAKIO, K. & HANAOKA, F. 1999. The XPV (xeroderma pigmentosum variant) gene encodes human DNA polymerase eta. *Nature,* 399**,** 700-4.

MATSUNAGA, F., FORTERRE, P., ISHINO, Y. & MYLLYKALLIO, H. 2001. In vivo interactions of archaeal Cdc6/Orc1 and minichromosome maintenance proteins with the replication origin. *Proc Natl Acad Sci U S*

A, 98**,** 11152-7.

MATSUNAGA, T., PARK, C. H., BESSHO, T., MU, D. & SANCAR, A. 1996. Replication protein A confers structure-specific endonuclease activities to the XPF-ERCC1 and XPG subunits of human DNA repair excision nuclease. *J Biol Chem,* 271**,** 11047-50.

MCAULEY-HECHT, K. E., LEONARD, G. A., GIBSON, N. J., THOMSON, J. B., WATSON, W. P., HUNTER, W. N. & BROWN, T. 1994. Crystal structure of a DNA duplex containing 8-hydroxydeoxyguanine-adenine base pairs. *Biochemistry,* 33**,** 10266-70.

MCCULLOCH, S. D. & KUNKEL, T. A. 2008. The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. *Cell Res,* 18**,** 148-61.

MCGEOCH, A. T., TRAKSELIS, M. A., LASKEY, R. A. & BELL, S. D. 2005. Organization of the archaeal MCM complex on DNA and implications for the helicase mechanism. *Nat Struct Mol Biol,* 12**,** 756-62.

MCKINNON, P. J. 2009. DNA repair deficiency and neurological disease. *Nat Rev Neurosci,* 10**,** 100-12.

MEMISOGLU, A. & SAMSON, L. 2000. Contribution of base excision repair, nucleotide excision repair, and DNA recombination to alkylation resistance of the fission yeast Schizosaccharomyces pombe. *J Bacteriol,* 182**,** 2104-12.

MESELSON, M. & STAHL, F. W. 1958. The Replication of DNA in Escherichia-Coli. *Proceedings of the National Academy of Sciences of the United States of America,* 44**,** 671-682.

MESSER, W. 2002. The bacterial replication initiator DnaA. DnaA and oriC, the bacterial mode to initiate DNA replication. *Fems Microbiology Reviews,* 26**,** 355-374.

MESSICK, T. E., CHMIEL, N. H., GOLINELLI, M. P., LANGER, M. R., JOSHUA-TOR, L. & DAVID, S. S. 2002. Noncysteinyl coordination to the [4Fe-4S]2+ cluster of the DNA repair adenine glycosylase MutY introduced via site-directed mutagenesis. Structural characterization of an unusual histidinyl-coordinated cluster. *Biochemistry,* 41**,** 3931-42.

MICHEL, B. 2005. After 30 years of study, the bacterial SOS response still surprises us. *PLoS Biol,* 3**,** e255.

MIKHAILOV, V. S. & ROHRMANN, G. F. 2002. Baculovirus replication factor LEF-1 is a DNA primase. *J Virol,* 76**,** 2287-97.

MINATO, S. & WERBIN, H. 1972. Excitation and fluorescence spectra of the chromophore associated with the DNA-photoreactivating enzyme from the blue-green alga Anacystis nidulans. *Photochem Photobiol,* 15**,** 97-100.

MITCHELL, D. L. 1988. The relative cytotoxicity of (6-4) photoproducts and cyclobutane dimers in mammalian cells. *Photochem Photobiol,* 48**,** 51-7.

MOFFATT, B. A. & STUDIER, F. W. 1987. T7 lysozyme inhibits transcription by T7 RNA polymerase. *Cell,* 49**,** 221-7.

MOOLENAAR, G. F., HOGLUND, L. & GOOSEN, N. 2001. Clue to damage recognition by UvrB: residues in the beta-hairpin structure prevent binding to non-damaged DNA. *EMBO J,* 20**,** 6140-9.

MOSCATO, B., SWAIN, M. & LORIA, J. P. 2016. Induced Fit in the Selection of Correct versus Incorrect Nucleotides by DNA Polymerase beta. *Biochemistry,* 55**,** 382-95.

MOSER, J., KOOL, H., GIAKZIDIS, I., CALDECOTT, K., MULLENDERS, L. H. & FOUSTERI, M. I. 2007. Sealing of chromosomal DNA nicks during nucleotide excision repair requires XRCC1 and DNA ligase III alpha in a cell-cycle-specific manner. *Mol Cell,* 27**,** 311-23.

MOURON, S., RODRIGUEZ-ACEBES, S., MARTINEZ-JIMENEZ, M. I., GARCIA-GOMEZ, S., CHOCRON, S., BLANCO, L. & MENDEZ, J. 2013. Repriming of DNA synthesis at stalled replication forks by human PrimPol. *Nat Struct Mol Biol,* 20**,** 1383-9.

MOYER, S. E., LEWIS, P. W. & BOTCHAN, M. R. 2006. Isolation of the Cdc45/Mcm2-7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. *Proc Natl Acad Sci U S A,* 103**,** 10236-41.

MULLER, H. J. 1927. Mutations Caused by X-Rays. *Scientific Monthly,* 25**,** 284-285.

MURZIN, A. G. 1993. OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J,* 12**,** 861-7.

MUZYCZKA, N., POLAND, R. L. & BESSMAN, M. J. 1972. Studies on the biochemical basis of spontaneous mutation. I. A comparison of the deoxyribonucleic acid polymerases of mutator, antimutator, and wild type

strains of bacteriophage T4. *J Biol Chem,* 247**,** 7116-22.

NAKAMURA, J., WALKER, V. E., UPTON, P. B., CHIANG, S. Y., KOW, Y. W. & SWENBERG, J. A. 1998. Highly sensitive apurinic/apyrimidinic site assay can detect spontaneous and chemically induced depurination under physiological conditions. *Cancer Res,* 58**,** 222-5.

NELSON, J. R., LAWRENCE, C. W. & HINKLE, D. C. 1996. Deoxycytidyl transferase activity of yeast REV1 protein. *Nature,* 382**,** 729-31.

NELSON, J. R., LAWRENCE, C. W. & HINKLE, D. C. 1996. Thymine-thymine dimer bypass by yeast DNA polymerase zeta. *Science,* 272**,** 1646-9.

NEW, J. H., SUGIYAMA, T., ZAITSEVA, E. & KOWALCZYKOWSKI, S. C. 1998. Rad52 protein stimulates DNA strand exchange by Rad51 and replication protein A. *Nature,* 391**,** 407-410.

NOIROT-GROS, M. F. & EHRLICH, S. D. 1996. Change of a catalytic reaction carried out by a DNA replication protein. *Science,* 274**,** 777-80.

O'DONOVAN, A., DAVIES, A. A., MOGGS, J. G., WEST, S. C. & WOOD, R. D. 1994. XPG endonuclease makes the 3' incision in human DNA nucleotide excision repair. *Nature,* 371**,** 432-5.

OGINO, H., ISHINO, S., MAYANAGI, K., HAUGLAND, G. T., BIRKELAND, N. K., YAMAGISHI, A. & ISHINO, Y. 2011. The GINS complex from the thermophilic archaeon, Thermoplasma acidophilum may function as a homotetramer in DNA replication. *Extremophiles,* 15**,** 529-39.

OHL, F. W., SCHEICH, H. & FREEMAN, W. J. 2000. Topographic analysis of epidural pure-tone-evoked potentials in gerbil auditory cortex. *J Neurophysiol,* 83**,** 3123-32.

OHMORI, H., FRIEDBERG, E. C., FUCHS, R. P., GOODMAN, M. F., HANAOKA, F., HINKLE, D., KUNKEL, T. A., LAWRENCE, C. W., LIVNEH, Z., NOHMI, T., PRAKASH, L., PRAKASH, S., TODO, T., WALKER, G. C., WANG, Z. & WOODGATE, R. 2001. The Y-family of DNA polymerases. *Mol Cell,* 8**,** 7-8.

OKAZAKI, R., OKAZAKI, T., SAKABE, K., SUGIMOTO, K. & SUGINO, A. 1968. Mechanism of DNA chain growth. I. Possible discontinuity and unusual secondary structure of newly synthesized chains. *Proc Natl Acad Sci U S A,* 59**,** 598-605.

OLIVE, P. L. 1998. The role of DNA single- and double-strand breaks in cell killing

by ionizing radiation. *Radiat Res,* 150**,** S42-51.

OLIVEIRA, M. T. & KAGUNI, L. S. 2010. Functional roles of the N- and C-terminal regions of the human mitochondrial single-stranded DNA-binding protein. *PLoS One,* 5**,** e15379.

OLLIS, D. L., BRICK, P., HAMLIN, R., XUONG, N. G. & STEITZ, T. A. 1985. Structure of large fragment of Escherichia coli DNA polymerase I complexed with dTMP. *Nature,* 313**,** 762-6.

OYAMA, T., ISHINO, S., FUJINO, S., OGINO, H., SHIRAI, T., MAYANAGI, K., SAITO, M., NAGASAWA, N., ISHINO, Y. & MORIKAWA, K. 2011. Architectures of archaeal GINS complexes, essential DNA replication initiation factors. *BMC Biol,* 9**,** 28.

PAGES, V. & FUCHS, R. P. 2003. Uncoupling of leading- and lagging-strand DNA replication during lesion bypass in vivo. *Science,* 300**,** 1300-3.

PAN, H. & WIGLEY, D. B. 2000. Structure of the zinc-binding domain of Bacillus stearothermophilus DNA primase. *Structure,* 8**,** 231-9.

PATA, J. D. 2010. Structural diversity of the Y-family DNA polymerases. *Biochim Biophys Acta,* 1804**,** 1124-35.

PAVLOV, Y. I. & SHCHERBAKOVA, P. V. 2010. DNA polymerases at the eukaryotic fork-20 years later. *Mutat Res,* 685**,** 45-53.

PETERMANN, E. & HELLEDAY, T. 2010. Pathways of mammalian replication fork restart. *Nat Rev Mol Cell Biol,* 11**,** 683-7.

PETERMANN, E., ORTA, M. L., ISSAEVA, N., SCHULTZ, N. & HELLEDAY, T. 2010. Hydroxyurea-stalled replication forks become progressively inactivated and require two different RAD51-mediated pathways for restart and repair. *Mol Cell,* 37**,** 492-502.

PETRUSKA, J., GOODMAN, M. F., BOOSALIS, M. S., SOWERS, L. C., CHEONG, C. & TINOCO, I., JR. 1988. Comparison between DNA melting thermodynamics and DNA polymerase fidelity. *Proc Natl Acad Sci U S A,* 85**,** 6252-6.

PHILIPS, F. S. 1950. Recent contributions to the pharmacology of bis(2-haloethyl) amines and sulfides. *J Pharmacol Exp Ther,* 99**,** 281-323.

PINZ, K. G., SHIBUTANI, S. & BOGENHAGEN, D. F. 1995. Action of mitochondrial DNA polymerase gamma at sites of base loss or oxidative damage. *J Biol Chem,* 270**,** 9202-6.

PITCHER, R. S., BRISSETT, N. C. & DOHERTY, A. J. 2007. Nonhomologous end-joining in bacteria: a microbial perspective. *Annu Rev Microbiol,* 61**,** 259-82.

PITCHER, R. S., BRISSETT, N. C., PICHER, A. J., ANDRADE, P., JUAREZ, R., THOMPSON, D., FOX, G. C., BLANCO, L. & DOHERTY, A. J. 2007. Structure and function of a mycobacterial NHEJ DNA repair polymerase. *J Mol Biol,* 366**,** 391-405.

POPANDA, O. & THIELMANN, H. W. 1992. The function of DNA polymerases in DNA repair synthesis of ultraviolet-irradiated human fibroblasts. *Biochim Biophys Acta,* 1129**,** 155-60.

POPLAWSKI, A., GRABOWSKI, B., LONG, S. E. & KELMAN, Z. 2001. The zinc finger domain of the archaeal minichromosome maintenance protein is required for helicase activity. *J Biol Chem,* 276**,** 49371-7.

PRAKASH, A. & BORGSTAHL, G. E. 2012. The structure and function of replication protein A in DNA replication. *Subcell Biochem,* 62**,** 171-96.

PRASAD, R., DIANOV, G. L., BOHR, V. A. & WILSON, S. H. 2000. FEN1 stimulation of DNA polymerase beta mediates an excision step in mammalian long patch base excision repair. *J Biol Chem,* 275**,** 4460-6.

PRASAD, R., LAVRIK, O. I., KIM, S. J., KEDAR, P., YANG, X. P., VANDE BERG, B. J. & WILSON, S. H. 2001. DNA polymerase beta -mediated long patch base excision repair. Poly(ADP-ribose)polymerase-1 stimulates strand displacement DNA synthesis. *J Biol Chem,* 276**,** 32411-4.

PRAT, F., HOUK, K. N. & FOOTE, C. S. 1998. Effect of guanine stacking on the oxidation of 8-oxoguanine in B-DNA. *Journal of the American Chemical Society,* 120**,** 845-846.

PRONK, R. & VAN DER VLIET, P. C. 1993. The adenovirus terminal protein influences binding of replication proteins and changes the origin structure. *Nucleic Acids Res,* 21**,** 2293-300.

PROTICSABLJIC, M. & KRAEMER, K. H. 1986. Reduced Repair of Non-Dimer Photoproducts in a Gene Transfected into Xeroderma Pigmentosum-Cells. *Photochemistry and Photobiology,* 43**,** 509-513.

RAO, P. N. & JOHNSON, R. T. 1970. Mammalian cell fusion: studies on the regulation of DNA synthesis and mitosis. *Nature,* 225**,** 159-64.

RASTOGI, R. P., RICHA, KUMAR, A., TYAGI, M. B. & SINHA, R. P. 2010.

Molecular mechanisms of ultraviolet radiation-induced DNA damage and repair. *J Nucleic Acids,* 2010**,** 592980.

REARDON, J. T., VAISMAN, A., CHANEY, S. G. & SANCAR, A. 1999. Efficient nucleotide excision repair of cisplatin, oxaliplatin, and Bis-aceto-ammine-dichloro-cyclohexylamine-platinum(IV) (JM216) platinum intrastrand DNA diadducts. *Cancer Res,* 59**,** 3968-71.

REED, C. J., LEWIS, H., TREJO, E., WINSTON, V. & EVILIA, C. 2013. Protein adaptations in archaeal extremophiles. *Archaea,* 2013**,** 373275.

REMUS, D., BEURON, F., TOLUN, G., GRIFFITH, J. D., MORRIS, E. P. & DIFFLEY, J. F. 2009. Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. *Cell,* 139**,** 719-30.

REUVEN, N. B., ARAD, G., MAOR-SHOSHANI, A. & LIVNEH, Z. 1999. The mutagenesis protein UmuC is a DNA polymerase activated by UmuD', RecA, and SSB and is specialized for translesion replication. *J Biol Chem,* 274**,** 31763-6.

RICHARDSON, T. T., GILROY, L., ISHINO, Y., CONNOLLY, B. A. & HENNEKE, G. 2013. Novel inhibition of archaeal family-D DNA polymerase by uracil. *Nucleic Acids Res,* 41**,** 4207-18.

RICKE, R. M. & BIELINSKY, A. K. 2004. Mcm10 regulates the stability and chromatin association of DNA polymerase-alpha. *Mol Cell,* 16**,** 173-85.

RICKE, R. M. & BIELINSKY, A. K. 2006. A conserved Hsp10-like domain in Mcm10 is required to stabilize the catalytic subunit of DNA polymerase-alpha in budding yeast. *J Biol Chem,* 281**,** 18414-25.

RINK, S. M., SOLOMON, M. S., TAYLOR, M. J., RAJUR, S. B., MCLAUGHLIN, L. W. & HOPKINS, P. B. 1993. Covalent Structure of a Nitrogen Mustard-Induced DNA Interstrand Cross-Link - an N7-to-N7 Linkage of Deoxyguanosine Residues at the Duplex Sequence 5'-D(Gnc). *Journal of the American Chemical Society,* 115**,** 2551-2557.

ROBBINS, J. B., MURPHY, M. C., WHITE, B. A., MACKIE, R. I., HA, T. & CANN, I. K. 2004. Functional analysis of multiple single-stranded DNA-binding proteins from Methanosarcina acetivorans and their effects on DNA synthesis by DNA polymerase BI. *J Biol Chem,* 279**,** 6315-26.

ROSENBERG, B., VANCAMP, L., TROSKO, J. E. & MANSOUR, V. H. 1969. Platinum Compounds - a New Class of Potent Antitumour Agents. *Nature,*

222**,** 385-+.

ROTHENBERG, E., TRAKSELIS, M. A., BELL, S. D. & HA, T. 2007. MCM forked substrate specificity involves dynamic interaction with the 5'-tail. *J Biol Chem,* 282**,** 34229-34.

ROTHWELL, P. J. & WAKSMAN, G. 2005. Structure and mechanism of DNA polymerases. *Adv Protein Chem,* 71**,** 401-40.

ROWEN, L. & KORNBERG, A. 1978. Primase, the dnaG protein of Escherichia coli. An enzyme which starts DNA chains. *J Biol Chem,* 253**,** 758-64.

RUDD, S. G., BIANCHI, J. & DOHERTY, A. J. 2014. PrimPol-A new polymerase on the block. *Mol Cell Oncol,* 1**,** e960754.

RUPP, W. D. & HOWARD-FLANDERS, P. 1968. Discontinuities in the DNA synthesized in an excision-defective strain of Escherichia coli following ultraviolet irradiation. *J Mol Biol,* 31**,** 291-304.

SAKAKIBARA, N., KELMAN, L. M. & KELMAN, Z. 2009. Unwinding the structure and function of the archaeal MCM helicase. *Mol Microbiol,* 72**,** 286-96.

SAKOFSKY, C. J., FOSTER, P. L. & GROGAN, D. W. 2012. Roles of the Y-family DNA polymerase Dbh in accurate replication of the Sulfolobus genome at high temperature. *DNA Repair (Amst),* 11**,** 391-400.

SALE, J. E., LEHMANN, A. R. & WOODGATE, R. 2012. Y-family DNA polymerases and their role in tolerance of cellular DNA damage. *Nature Reviews Molecular Cell Biology,* 13**,** 141-152.

SAMSON, R. Y. & BELL, S. D. 2013. MCM loading--an open-and-shut case? *Mol Cell,* 50**,** 457-8.

SANCHEZ-PULIDO, L. & PONTING, C. P. 2011. Cdc45: the missing RecJ ortholog in eukaryotes? *Bioinformatics,* 27**,** 1885-8.

SARMIENTO, F., MRAZEK, J. & WHITMAN, W. B. 2013. Genome-scale analysis of gene function in the hydrogenotrophic methanogenic archaeon Methanococcus maripaludis. *Proc Natl Acad Sci U S A,* 110**,** 4726-31.

SAYERS, E. W., BARRETT, T., BENSON, D. A., BOLTON, E., BRYANT, S. H., CANESE, K., CHETVERNIN, V., CHURCH, D. M., DICUCCIO, M., FEDERHEN, S., FEOLO, M., GEER, L. Y., HELMBERG, W., KAPUSTIN, Y., LANDSMAN, D., LIPMAN, D. J., LU, Z., MADDEN, T. L., MADEJ, T., MAGLOTT, D. R., MARCHLER-BAUER, A., MILLER, V., MIZRACHI, I., OSTELL, J., PANCHENKO, A., PRUITT, K. D., SCHULER, G. D.,

SEQUEIRA, E., SHERRY, S. T., SHUMWAY, M., SIROTKIN, K., SLOTTA, D., SOUVOROV, A., STARCHENKO, G., TATUSOVA, T. A., WAGNER, L., WANG, Y., JOHN WILBUR, W., YASCHENKO, E. & YE, J. 2010. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res,* 38**,** D5-16.

SCHIAVONE, D., JOZWIAKOWSKI, S. K., ROMANELLO, M., GUILBAUD, G., GUILLIAM, T. A., BAILEY, L. J., SALE, J. E. & DOHERTY, A. J. 2016. PrimPol Is Required for Replicative Tolerance of G Quadruplexes in Vertebrate Cells. *Mol Cell,* 61**,** 161-9.

SCHOFIELD, M. J., BROWNEWELL, F. E., NAYAK, S., DU, C., KOOL, E. T. & HSIEH, P. 2001. The Phe-X-Glu DNA binding motif of MutS. The role of hydrogen bonding in mismatch recognition. *J Biol Chem,* 276**,** 45505-8.

SHACHAR, S., ZIV, O., AVKIN, S., ADAR, S., WITTSCHIEBEN, J., REISSNER, T., CHANEY, S., FRIEDBERG, E. C., WANG, Z., CARELL, T., GEACINTOV, N. & LIVNEH, Z. 2009. Two-polymerase mechanisms dictate error-free and error-prone translesion DNA synthesis in mammals. *EMBO J,* 28**,** 383-93.

SHERMAN, G., GOTTLIEB, J. & CHALLBERG, M. D. 1992. The UL8 subunit of the herpes simplex virus helicase-primase complex is required for efficient primer utilization. *J Virol,* 66**,** 4884-92.

SHIVJI, K. K., KENNY, M. K. & WOOD, R. D. 1992. Proliferating cell nuclear antigen is required for DNA excision repair. *Cell,* 69**,** 367-74.

SIEVERS, F., WILM, A., DINEEN, D., GIBSON, T. J., KARPLUS, K., LI, W., LOPEZ, R., MCWILLIAM, H., REMMERT, M., SODING, J., THOMPSON, J. D. & HIGGINS, D. G. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol,* 7**,** 539.

SIMON, A. C., ZHOU, J. C., PERERA, R. L., VAN DEURSEN, F., EVRIN, C., IVANOVA, M. E., KILKENNY, M. L., RENAULT, L., KJAER, S., MATAK-VINKOVIC, D., LABIB, K., COSTA, A. & PELLEGRINI, L. 2014. A Ctf4 trimer couples the CMG helicase to DNA polymerase alpha in the eukaryotic replisome. *Nature,* 510**,** 293-7.

SIMON, M., GIOT, L. & FAYE, G. 1991. The 3' to 5' Exonuclease Activity Located in the DNA-Polymerase Delta-Subunit of Saccharomyces-Cerevisiae Is

Required for Accurate Replication. *Embo Journal,* 10**,** 2165-2170.

SINGLETON, M. R., MORALES, R., GRAINGE, I., COOK, N., ISUPOV, M. N. & WIGLEY, D. B. 2004. Conformational changes induced by nucleotide binding in Cdc6/ORC from Aeropyrum pernix. *J Mol Biol,* 343**,** 547-57.

SINHA, R. P. & HADER, D. P. 2002. UV-induced DNA damage and repair: a review. *Photochem Photobiol Sci,* 1**,** 225-36.

SPEYER, J. F. 1965. Mutagenic DNA polymerase. *Biochem Biophys Res Commun,* 21**,** 6-8.

STEITZ, T. A. 1999. DNA polymerases: structural diversity and common mechanisms. *J Biol Chem,* 274**,** 17395-8.

STEITZ, T. A. & STEITZ, J. A. 1993. A general two-metal-ion mechanism for catalytic RNA. *Proc Natl Acad Sci U S A,* 90**,** 6498-502.

STILLMAN, B. 2008. DNA polymerases at the replication fork in eukaryotes. *Mol Cell,* 30**,** 259-60.

STUCKI, M., PASCUCCI, B., PARLANTI, E., FORTINI, P., WILSON, S. H., HUBSCHER, U. & DOGLIOTTI, E. 1998. Mammalian base excision repair by DNA polymerases delta and epsilon. *Oncogene,* 17**,** 835-43.

STUDIER, F. W. 1991. Use of Bacteriophage-T7 Lysozyme to Improve an Inducible T7 Expression System. *Journal of Molecular Biology,* 219**,** 37-44.

SUGIMOTO, K., OKAZAKI, T. & OKAZAKI, R. 1968. Mechanism of DNA Chain Growth .2. Accumulation of Newly Synthesized Short Chains in E Coli Infected with Ligase-Defective T4 Phages. *Proceedings of the National Academy of Sciences of the United States of America,* 60**,** 1356-&.

SUNG, P. & KLEIN, H. 2006. Mechanism of homologous recombination: mediators and helicases take on regulatory functions. *Nat Rev Mol Cell Biol,* 7**,** 739-50.

SVOBODA, D. L. & VOS, J. M. 1995. Differential replication of a single, UV-induced lesion in the leading or lagging strand by a human cell extract: fork uncoupling or gap formation. *Proc Natl Acad Sci U S A,* 92**,** 11975-9.

SWIATEK, A. & MACNEILL, S. A. 2010. The archaeo-eukaryotic GINS proteins and the archaeal primase catalytic subunit PriS share a common domain. *Biol Direct,* 5**,** 17.

TAKAYAMA, Y., KAMIMURA, Y., OKAWA, M., MURAMATSU, S., SUGINO, A.

& ARAKI, H. 2003. GINS, a novel multiprotein complex required for chromosomal DNA replication in budding yeast. *Genes Dev,* 17**,** 1153-65.

TAYLOR, J. S. 1990. DNA, Sunlight, and Skin-Cancer. *Journal of Chemical Education,* 67**,** 835-841.

TODO, T., RYO, H., YAMAMOTO, K., TOH, H., INUI, T., AYAKI, H., NOMURA, T. & IKENAGA, M. 1996. Similarity among the Drosophila (6-4)photolyase, a human photolyase homolog, and the DNA photolyase-blue-light photoreceptor family. *Science,* 272**,** 109-12.

TRAUT, T. W. 1994. Physiological concentrations of purines and pyrimidines. *Mol Cell Biochem,* 140**,** 1-22.

TRINCAO, J., JOHNSON, R. E., ESCALANTE, C. R., PRAKASH, S., PRAKASH, L. & AGGARWAL, A. K. 2001. Structure of the catalytic core of S. cerevisiae DNA polymerase eta: implications for translesion DNA synthesis. *Mol Cell,* 8**,** 417-26.

TSURIMOTO, T. & STILLMAN, B. 1989. Multiple replication factors augment DNA synthesis by the two eukaryotic DNA polymerases, alpha and delta. *EMBO J,* 8**,** 3883-9.

UEMORI, T., ISHINO, Y., DOI, H. & KATO, I. 1995. The hyperthermophilic archaeon Pyrodictium occultum has two alpha-like DNA polymerases. *J Bacteriol,* 177**,** 2164-77.

UEMORI, T., ISHINO, Y., TOH, H., ASADA, K. & KATO, I. 1993. Organization and Nucleotide-Sequence of the DNA-Polymerase Gene from the Archaeon Pyrococcus-Furiosus. *Nucleic Acids Research,* 21**,** 259-265.

VAN CRIEKINGE, W. & BEYAERT, R. 1999. Yeast Two-Hybrid: State of the Art. *Biol Proced Online,* 2**,** 1-38.

VAN DEN BOSCH, M., BREE, R. T. & LOWNDES, N. F. 2003. The MRN complex: coordinating and mediating the response to broken chromosomes. *EMBO Rep,* 4**,** 844-9.

VAN LOON, B., MARKKANEN, E. & HUBSCHER, U. 2010. Oxygen as a friend and enemy: How to combat the mutational potential of 8-oxo-guanine. *DNA Repair (Amst),* 9**,** 604-16.

VOULGARIDOU, G. P., ANESTOPOULOS, I., FRANCO, R., PANAYIOTIDIS, M. I. & PAPPA, A. 2011. DNA damage induced by endogenous aldehydes: current state of knowledge. *Mutat Res,* 711**,** 13-27.

WAGNER, J., GRUZ, P., KIM, S. R., YAMADA, M., MATSUI, K., FUCHS, R. P. & NOHMI, T. 1999. The dinB gene encodes a novel E. coli DNA polymerase, DNA pol IV, involved in mutagenesis. *Mol Cell,* 4**,** 281-6.

WAN, L., LOU, J., XIA, Y., SU, B., LIU, T., CUI, J., SUN, Y., LOU, H. & HUANG, J. 2013. hPrimpol1/CCDC111 is a human DNA primase-polymerase required for the maintenance of genome integrity. *EMBO Rep,* 14**,** 1104-12.

WANG, T. S. 1991. Eukaryotic DNA polymerases. *Annu Rev Biochem,* 60**,** 513-52.

WARDLE, J., BURGERS, P. M. J., CANN, I. K. O., DARLEY, K., HESLOP, P., JOHANSSON, E., LIN, L. J., MCGLYNN, P., SANVOISIN, J., STITH, C. M. & CONNOLLY, B. A. 2008. Uracil recognition by replicative DNA polymerases is limited to the archaea, not occurring with bacteria and eukarya. *Nucleic Acids Research,* 36**,** 705-711.

WATASE, G., TAKISAWA, H. & KANEMAKI, M. T. 2012. Mcm10 plays a role in functioning of the eukaryotic replicative DNA helicase, Cdc45-Mcm-GINS. *Curr Biol,* 22**,** 343-9.

WATERHOUSE, A. M., PROCTER, J. B., MARTIN, D. M., CLAMP, M. & BARTON, G. J. 2009. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics,* 25**,** 1189-91.

WATRIN, L. & PRIEUR, D. 1996. UV and ethyl methanesulfonate effects in hyperthermophilic archaea and isolation of auxotrophic mutants of Pyrococcus strains. *Curr Microbiol,* 33**,** 377-82.

WATSON, J. D. & CRICK, F. H. 1953. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature,* 171**,** 737-8.

WATSON, J. D. & CRICK, F. H. C. 1953. Molecular Structure of Nucleic Acids - a Structure for Deoxyribose Nucleic Acid. *Nature,* 171**,** 737-738.

WELLER, G. R. & DOHERTY, A. J. 2001. A family of DNA repair ligases in bacteria? *FEBS Lett,* 505**,** 340-2.

WELLER, G. R., KYSELA, B., ROY, R., TONKIN, L. M., SCANLAN, E., DELLA, M., DEVINE, S. K., DAY, J. P., WILKINSON, A., D'ADDA DI FAGAGNA, F., DEVINE, K. M., BOWATER, R. P., JEGGO, P. A., JACKSON, S. P. & DOHERTY, A. J. 2002. Identification of a DNA nonhomologous end-joining complex in bacteria. *Science,* 297**,** 1686-9.

WICKNER, W., BRUTLAG, D., SCHEKMAN, R. & KORNBERG, A. 1972. RNA synthesis initiates in vitro conversion of M13 DNA to its replicative form. *Proc Natl Acad Sci U S A,* 69**,** 965-9.

WOESE, C. R. & FOX, G. E. 1977. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc Natl Acad Sci U S A,* 74**,** 5088-90.

WOHLSCHLEGEL, J. A., DWYER, B. T., DHAR, S. K., CVETIC, C., WALTER, J. C. & DUTTA, A. 2000. Inhibition of eukaryotic DNA replication by geminin binding to Cdt1. *Science,* 290**,** 2309-12.

WOLD, M. S. 1997. Replication protein A: A heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annual Review of Biochemistry,* 66**,** 61-92.

WOODS, W. G. & DYALLSMITH, M. L. 1997. Construction and analysis of a recombination-deficient (radA) mutant of Haloferax volcanii. *Molecular Microbiology,* 23**,** 791-797.

XU, Y., GRISTWOOD, T., HODGSON, B., TRINIDAD, J. C., ALBERS, S. V. & BELL, S. D. 2016. Archaeal orthologs of Cdc45 and GINS form a stable complex that stimulates the helicase activity of MCM. *Proc Natl Acad Sci U S A,* 113**,** 13390-13395.

YANG, H. J., LI, Q. B., FAN, J., HOLLOMAN, W. K. & PAVLETICH, N. P. 2005. The BRCA2 homologue Brh2 nucleates RAD51 filament formation at a dsDNA-ssDNA junction. *Nature,* 433**,** 653-657.

YANG, W. 2005. Portraits of a Y-family DNA polymerase. *FEBS Lett,* 579**,** 868-72.

YEELES, J. T., DEEGAN, T. D., JANSKA, A., EARLY, A. & DIFFLEY, J. F. 2015. Regulated eukaryotic DNA replication origin firing with purified proteins. *Nature,* 519**,** 431-5.

YEELES, J. T. & MARIANS, K. J. 2011. The Escherichia coli replisome is inherently DNA damage tolerant. *Science,* 334**,** 235-8.

YOSHIMOCHI, T., FUJIKANE, R., KAWANAMI, M., MATSUNAGA, F. & ISHINO, Y. 2008. The GINS complex from Pyrococcus furiosus stimulates the MCM helicase activity. *J Biol Chem,* 283**,** 1601-9.

YUAN, Z., BAI, L., SUN, J., GEORGESCU, R., LIU, J., O'DONNELL, M. E. & LI, H. 2016. Structure of the eukaryotic replicative CMG helicase suggests a pumpjack motion for translocation. *Nat Struct Mol Biol,* 23**,** 217-24.

ZAMENHOF, S., BRAWERMAN, G. & CHARGAFF, E. 1952. On the Desoxypentose Nucleic Acids from Several Microorganisms. *Biochimica Et Biophysica Acta,* 9**,** 402-405.

ZERBE, L. K. & KUCHTA, R. D. 2002. The p58 subunit of human DNA primase is important for primer initiation, elongation, and counting. *Biochemistry,* 41**,** 4891-900.

ZHENG, T., HUANG, Q., ZHANG, C., NI, J., SHE, Q. & SHEN, Y. 2012. Development of a simvastatin selection marker for a hyperthermophilic acidophile, Sulfolobus islandicus. *Appl Environ Microbiol,* 78**,** 568-74.

ZHOU, B. B. & ELLEDGE, S. J. 2000. The DNA damage response: putting checkpoints in perspective. *Nature,* 408**,** 433-9.

ZHU, H., NANDAKUMAR, J., ANIUKWU, J., WANG, L. K., GLICKMAN, M. S., LIMA, C. D. & SHUMAN, S. 2006. Atomic structure and nonhomologous end-joining function of the polymerase component of bacterial DNA ligase D. *Proc Natl Acad Sci U S A,* 103**,** 1711-6.

ZOU, L., LIU, D. & ELLEDGE, S. J. 2003. Replication protein A-mediated recruitment and activation of Rad17 comp

Appendix

PNAS PLUS

Archaeal replicative primases can perform translesion DNA synthesis

Stanislaw K. Jozwiakowski, Farimah Borazjani Gholami, and Aidan J. Doherty[1]

Genome Damage and Stability Centre, School of Life Sciences, University of Sussex, Brighton BN1 9RQ, United Kingdom

Edited by Kenneth J. Marians, Memorial Sloan-Kettering Cancer Center, New York, NY, and accepted by the Editorial Board January 8, 2015 (received for review July 9, 2014)

DNA replicases routinely stall at lesions encountered on the template strand, and translesion DNA synthesis (TLS) is used to rescue progression of stalled replisomes. This process requires specialized polymerases that perform translesion DNA synthesis. Although prokaryotes and eukaryotes possess canonical TLS polymerases (Y-family Pols) capable of traversing blocking DNA lesions, most archaea lack these enzymes. Here, we report that archaeal replicative primases (Pri S, primase small subunit) can also perform TLS. Archaeal Pri S can bypass common oxidative DNA lesions, such as 8-Oxo-2′-deoxyguanosines and UV light-induced DNA damage, faithfully bypassing cyclobutane pyrimidine dimers. Although it is well documented that archaeal replicases specifically arrest at deoxyuracils (dUs) due to recognition and binding to the lesions, a replication restart mechanism has not been identified. Here, we report that Pri S efficiently replicates past dUs, even in the presence of stalled replicase complexes, thus providing a mechanism for maintaining replication bypass of these DNA lesions. Together, these findings establish that some replicative primases, previously considered to be solely involved in priming replication, are also TLS proficient and therefore may play important roles in damage tolerance at replication forks.

archaea | replication | translesion synthesis | AEP | primase

The DNA replication machinery rapidly and accurately copies genomes but is prone to stalling at lesions and physical barriers (1). A variety of cellular pathways have evolved to restart stalled replication forks. These include translesion DNA synthesis (TLS) that is performed by specialized polymerases that synthesize short tracts of DNA opposite lesions, thus enabling reinitiation of replication (2). Error-free bypass mechanisms, mediated by homologous recombination, use an alternative undamaged template to rescue stalled replication forks (3). Stalled replisomes can also be rescued by repriming downstream of the blockage, leaving a gap opposite the lesion (4, 5).

Eukaryotes and prokaryotes encode distinct TLS polymerases required for DNA damage tolerance (e.g., Y-family Pols). Although much of our understanding of TLS mechanisms has come from studies of archaeal Y-family DNA polymerases, the majority of archaeal species lack canonical TLS enzymes (Fig. L4) (6), surprising given the otherwise high degree of conservation between eukaryotic and archaeal replisomes. Many archaea do not appear to encode nucleotide excision repair or photolyase pathways that remove UV light-induced damage (6). These anomalies pose the question as to how archaea, lacking canonical TLS or lesion repair pathways, tolerate the presence of lesions that stall replication. This is particularly pertinent to archaea because of the harsh environmental conditions under which many species reside, including extreme temperatures, which promote increased levels of DNA damage.

Archaeal replicases (B- and D-family Pols) specifically arrest at deoxyuracil (dU) (7, 8). This unique feature is limited to replicases from archaea (9). Two important questions regarding dU-induced stalling of archaeal replisomes remain unanswered. First, why do archaea stall replication in response to the template strand dU? Second, how are archaeal genomes containing dU copied? This stalling mechanism may have evolved to prevent promutagenic bypass of the template strand dU, resulting in C–T transition (7, 9). The mechanism used by archaea to resume replication after dU-induced replisome stalling has not been identified.

In this study, we report that archaeal replicative primases (primase small subunit, Pri S) can perform translesion DNA synthesis on damaged DNA templates. Pri S can bypass common DNA lesions, such as oxidative and UV damages, faithfully bypassing cyclobutane pyrimidine dimers (CPDs). Additionally, we report that Pri S can replicate past template strand dUs, even in the presence of stalled replicative polymerase B and proliferating cell nuclear antigen (Pol B/PCNA) complexes, thus providing a specific mechanism for maintaining timely replication of DNA containing dU lesions. Together, these findings establish that the archaeal primase is not only required for de novo primer synthesis during initiation of DNA replication but also actively participates during the elongation step by assisting the major DNA replicases in traversing DNA lesions.

Results and Discussion

Archaeal Primases Replicate Past 8-Oxo-2′-Deoxyguanosines. To address which enzymes are responsible for lesion bypass synthesis in archaea, we first considered which polymerases are present in all archaeal species that could facilitate tolerance of commonly occurring replicase-stalling lesions. One candidate is the replicative primase, a specialized DNA polymerase involved in de novo primer synthesis. In archaea and eukaryotes, the small catalytic subunit primase (Pri S) a member of the archaeo-eukaryotic primase (AEP) family (10), together with the large

Significance

DNA replicases stall at lesions during replication, potentially leading to genome instability. However, cells use specialized lesion bypass polymerases to restart stalled replisomes. Although most organisms possess these damage tolerance polymerases, capable of traversing blocking DNA lesions, many appear to lack these enzymes. We have discovered that replicative primases from archaea, previously considered to be solely involved in priming replication, are also capable of performing translesion DNA synthesis. This discovery has major implications for our understanding of additional roles of DNA primases during replication and the subsequent evolution of related lesion bypass pathways in eukaryotic organisms.

Author contributions: S.K.J. and A.J.D. designed research; S.K.J. and F.B.G. performed research; S.K.J. and A.J.D. analyzed data; and S.K.J. and A.J.D. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. K.J.M. is a guest editor invited by the Editorial Board.

Freely available online through the PNAS open access option.

[1]To whom correspondence should be addressed. Email: ajd21@sussex.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1412982112/-/DCSupplemental.

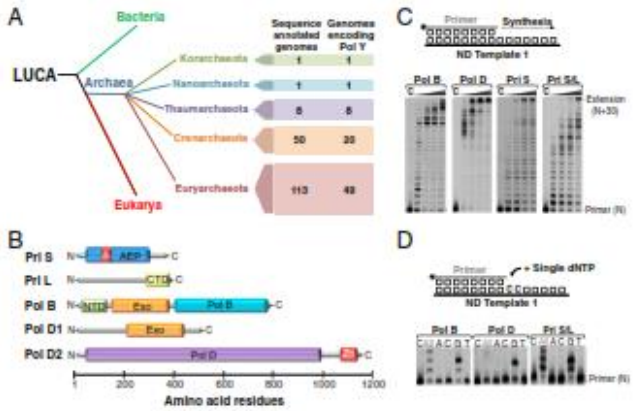


Fig. 1. *A. fulgidus* replisomal enzymes displaying DNA polymerase activity. Analysis of 173 archaeal genomes revealed that only 79 archaea encode canonical TLS DNA polymerases. DNA polymerization activities of *A. fulgidus* replisomal enzymes. (*A*) The absence of genes encoding Y-family DNA polymerases in most archaea is shown. (*B*) Structural elements present in *A. fulgidus* replisomal enzymes. Abbreviations: AEP, archaeo-eukaryotic primase; CTD, carboxy terminal domain; Exo, exonuclease; NTD, amino terminal domain; Pol, polymerase; and Zn, zinc binding site. (*C*) Polymerization on nondamaged templates. (*D*) Single nucleotide incorporation on nondamaged templates. The letter C denotes no enzyme control. The triangles above gel panels indicate time course of the polymerization (30 s, 1′, 3′, and 10′).

subunit (Pri L), is requisite for the initiation of DNA replication (11). Although Pri S has been considered to function exclusively in RNA primer synthesis during replication, archaeal Pri S and related plasmid-borne primases possess both primase and polymerase activities (12–14). In contrast to eukaryotic primases, archaeal Pri S possesses robust DNA polymerase activity (12). It also has terminal transferase activity, suggesting less stringent recognition of DNA substrates (14, 15) and potential involvement in DNA repair processes (16). Notably, closely related prokaryotic and archaeal AEPs, primase-polymerase domain of ligase D (PolDom/LigD) are polymerases involved in the repair of DNA breaks and possess a variety of DNA polymerization activities, including TLS (17, 18). PrimPol, a novel eukaryotic primase, was recently shown to assist in the bypass of lesions encountered during replication (19). Together, these enzymes belong to a growing class of primase polymerases known as PrimPols to reflect their enzymatic activities and origins.

To examine the potential TLS activities of archaeal primases, we cloned, expressed, and purified the catalytically active small primase subunit (Afu-Pri S), heterodimeric primase (Afu-Pri S/L), and the major replicases (Afu-Pol B and Pol D) from *Archaeoglobus fulgidus* (Afu) and the primase holoenzyme (Pfu-Pri S/L) from *Pyrococcus furiosus* (Pfu) (*SI Appendix*, Fig. S1*A*). The structural features of the Afu enzymes studied here are illustrated in Fig. 1*B*. First, we tested the DNA template-dependent polymerase activity of Afu replicative polymerases (Afu-Pol B and Afu-Pol D), primase subunit (Afu-Pri S), and the heterodimeric primase complex (Afu-Pri S/L). All enzymes exhibited robust and error-free DNA polymerase activity (Fig. 1 *C* and *D*). Notably, when we assayed for reverse transcriptase activity, this activity was observed for both primases but not the replicative polymerases (*SI Appendix*, Fig. S3*C*). Again, this suggests that archaeal primases have relaxed substrate specificity, a feature characteristic of TLS polymerases.

The most frequent type of DNA damage is induced by oxidative stress, resulting in the formation of 8-oxo-2′-deoxyguanosine (8-oxo dG) (20). Most replicative DNA polymerases misrecognize 8-oxo dG, resulting in incorrect deoxyadenosine (dA) incorporation opposite to this lesion (21). We tested whether Afu polymerases and primases were promutagenic while traversing 8-oxo dG and observed that all of these enzymes could bypass 8-oxo dG, showing marked stalling before and after the lesion (Fig. 2*A*). A similar profile of bypass past 8-oxo dG was observed for Pfu-Pri S/L (*SI Appendix*, Fig. S1*B*). Next, we investigated the fidelity of 8-oxo dG bypass by Afu and Pfu enzymes using single nucleotide incorporation assays. Family-B DNA polymerase (Afu-Pol B) displayed error-prone bypass of 8-oxo dG, incorporating dA opposite the lesion (Fig. 2*B*). Notably, Afu-Pol D and both primases (Afu-Pri S/L and Pfu-Pri S/L) incorporated the correct deoxycytosine (dC) and incorrect deoxyadenosine (dA) opposite the damage with comparable efficiency (Fig. 2*B* and *SI Appendix*, Fig. S1*C*). Thermophilic archaea are subjected to increased levels of oxidative stress, promoting depurination of 8-oxo dG to abasic site (Ab) and oxidation of thymine to thymine glycol (Tg). However, all of the tested enzymes were strongly blocked by Ab or Tg lesions (*SI Appendix*, Fig. S3 *A* and *B*, respectively), in common with PrimPol (19).

Archaeal Primases Replicate DNA Templates Containing Deoxyuracils. Hydrolytic deamination of deoxycytosine (dC) to dU frequently occurs in DNA (20) and is greatly accelerated by temperature. Therefore, thermophiles are at increased risk from this type of damage (22). Archaeal replicative polymerases (Pol B and D) evolved specifically to avoid replicating past dU and Pol B possesses a uracil-binding pocket (Fig. 1*B*) that scans the template for this lesion in advance of the replicase (23). When detected, the deaminated base is bound tightly and replication arrests four bases before dU (23). Notably, Pol B/PCNA binds dU tightly enough to prevent the lesion being removed by base excision repair enzymes (24). This suggests that the uracil-binding pocket may be important for protecting the integrity of dU-containing DNA during replication. DNA synthesis by Pol D is also markedly inhibited by dU on the template strand (8). To determine if dU inhibition was also evident in *Afu* replicases, we measured whether Afu-Pol B and D could traverse dU and observed that they both profoundly arrested at this lesion (Fig. 2*C*). Next, we assayed for synthesis opposite dU by Afu-Pri S and Afu-Pri S/L and observed that the primase readily bypassed dU (Fig. 2*C*).

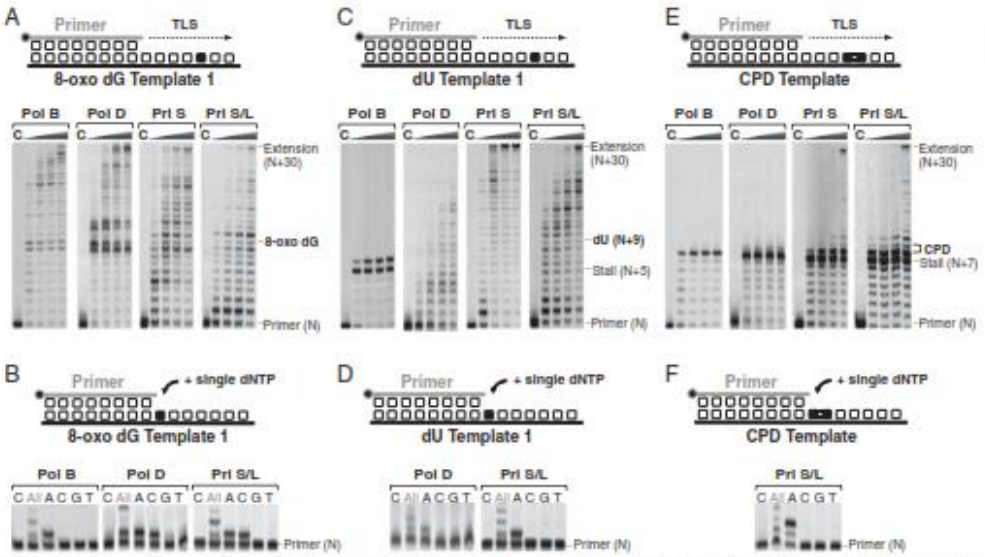


Fig. 2. Translesion synthesis past 8-oxo-dGs, dUs, and CPDs. (*A*) TLS performed on templates containing 8-oxo-dGs. (*B*) Single nucleotide incorporation opposite 8-oxo-dG. (*C*) TLS performed on templates containing dUs. (*D*) Single nucleotide incorporation opposite dU. (*E*) TLS performed on templates containing CPDs. (*F*) Single nucleotide incorporation opposite CPD. The letter C denotes no enzyme control. The triangles above gel panels indicate time course of the polymerization (30 s, 1′, 5′, and 10′).

Bypass of dUs was also observed for Pfu-Pri S/L (*SI Appendix*, Fig. S1*B*). Next, we measured the fidelity of the dU bypass using single nucleotide incorporation assay opposite the lesion. Archaeal primases, specifically incorporated dA opposite dU, indicating that bypass synthesis is promutagenic (Fig. 2*D* and *SI Appendix*, Fig. S1*D*).

Afu Replicative Primase Catalyzes Error-Free Bypass of Cyclobutane Pyrimidine Dimers. Thermophilic archaea, including *Archaeoglobus* and *Pyrococcus*, tolerate high doses of UV light (25, 26), notable given their apparent lack of recognizable lesion repair or bypass mechanisms. UV induces DNA lesions, particularly cross-links between adjacent pyrimidine bases including: cyclobutane pyrimidine dimers (CPDs) and 6-4 photoproducts (6-4PPs) modifications (27). These UV photoproducts distort DNA and act as potent replication blocking lesions (2, 21). Similar replicase stalling lesions can occur at significant rates in organisms not exposed to light, e.g., produced by cross-linking with aldehyde (28). Cross-linking rates are accelerated by oxidative stress; therefore, this type of DNA damage may be abundant in thermophilic archaea. First, we tested whether Afu-Pri S has TLS activity on templates containing CPDs. Whereas the replicative polymerases were incapable of bypassing CPDs, both Afu-Pri S and Afu-Pri S/L performed TLS across this UV-induced lesion (Fig. 2*C*), establishing that *Afu* replicative primase can also catalyze bypass of the CPDs. Second, we evaluated the fidelity of the CPD bypass, measuring single base incorporations opposite the damage. Afu-Pri S/L incorporated two dAs opposite both the 3′ and 5′ templating thymines of the CPD (Fig. 2*F*), establishing that archaeal primase catalyses error-free TLS past this UV damage. Notably, the bypass fidelity of Afu-Pri S/L mirrors the TLS activity of Pol η, a eukaryotic Y-family polymerase involved in error-free bypass of CPD (2). In contrast, *Pfu* primase (Pfu-Pri S/L) was unable to traverse the CPD, incorporating a single correct incoming base (dA) opposite the first (3′) base of the dimer (*SI Appendix*, Fig. S1 *B* and *E*). Next, we measured the capacity of archaeal enzymes to replicate past the 6-4PPs. The archaeal enzymes were unable to bypass this UV damage (*SI Appendix*, Fig. S3*D*). Although PrimPol can bypass this lesion (19), 6-4PPs cannot be traversed by other TLS DNA polymerases (2, 27). Bypass often requires the collaborative effort of two specialized enzymes, where the first performs insertion opposite the 6-4PP (an "inserter" polymerase) and the second extends the primer bearing 3′ terminal base aligned with the dimer (an "extender" polymerase) (2, 21). We therefore measured if the archaeal enzymes display TLS "extender" abilities on UV lesions and observed that both primases could extend primers containing 3′ terminal bases annealed to the CPDs and 6-4PPs (*SI Appendix*, Figs. S1*G*, S2 *A* and *B*, and S4 *B* and *D*); see *SI Appendix*, *SI Results* for details.

Afu Primase Rescues Pol B Stalled at the Template Strand Deoxyuracil. Although archaeal replicases bind dU with nanomolar affinity (29), a restart mechanism has not yet been identified. To address whether the replicative primase plays a role in restarting arrested replisomes, we assayed TLS activities of Afu-Pri S/L on short templates (30 nt), containing a single dU, preincubated with the Pol B/PCNA complex (Fig. 3*A*). We observed that Pri S/L retained robust TLS bypass activity on dU-containing templates, even in the presence of stalled replisome components. These data indicate that Pri S/L has an innate capacity to access the 3′ end of the primer, even when Pol B/PCNA complex has stalled at dU. This is a notable observation given that Pol B/PCNA stalling at dU abrogates detection and removal of deaminated bases by DNA glycosylases (24).

As Pri S/L assists in the bypass of replication blocking lesions, we next investigated whether archaeal primases function as a

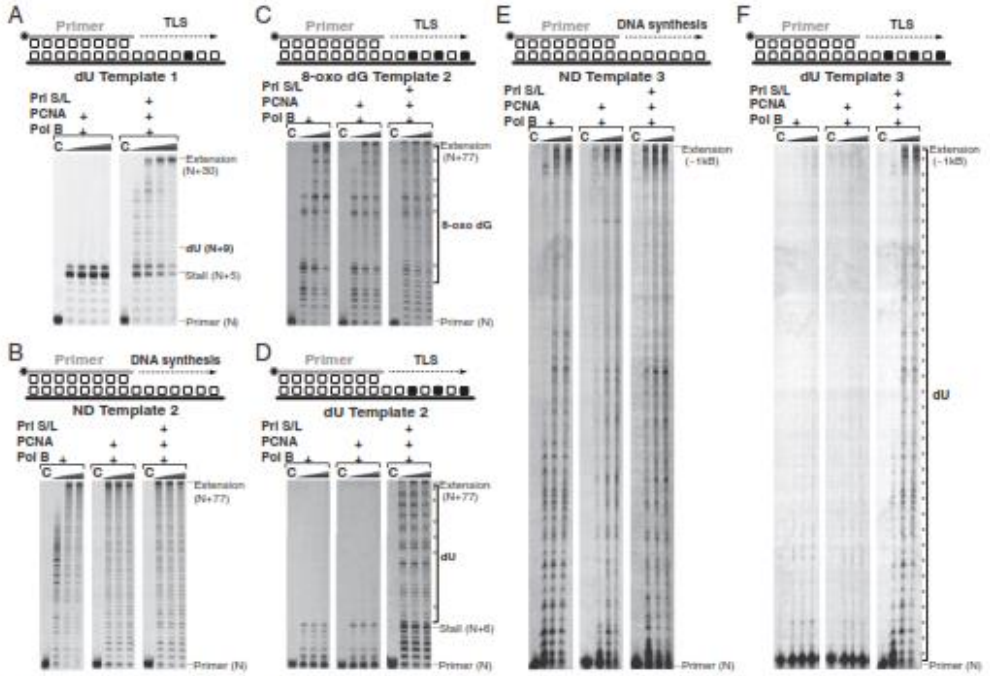


Fig. 3. Translesion synthesis past multiple DNA damage. (*A*) Pol B/PCNA stalls four bases before dU (*Left*). Stalled Pol B/PCNA is rescued when TLS past dU is performed by Pri S/L (*Right*). (*B*) Control primer extension on nondamaged substrates. Polymerization performed by Pol B (*Left*), Pol B/PCNA (*Middle*), and Pol B/PCNA and Pri S/L (*Right*). (*C*) Primer extension on templates containing seven 8-oxo dGs. TLS performed by Pol B and Pol B/PCNA results in full-length product (*Left* and *Middle*, respectively) but pronounced pausing pattern opposite to 8-oxo dG is observed. Polymerization performed by Pol B/PCNA and Pri S/L results in efficient primer extension. (*D*) Polymerization on a DNA template containing seven dUs performed by Pol B and Pol B/PCNA is strongly inhibited (*Left* and *Middle*, respectively). In contrast, DNA synthesis performed by Pol B/PCNA and Pri S/L resulted in bypass of the multiple dUs. (*E*) Primer extension on long nondamaged templates. Polymerization was performed by Pol B (*Left*), Pol B/PCNA (*Middle*), Pol B/PCNA and Pri S/L (*Right*). (*F*) Polymerization on long templates containing ~20 randomly distributed dUs. Strong inhibition of DNA synthesis is observed for Pol B and Pol B/PCNA (*Left* and *Middle*, respectively). Again, DNA synthesis performed by Pol B/PCNA and Pri S/L resulted in full-length primer extension. The letter C denotes no enzyme control. The triangles above gel panels indicate time course of the polymerization; *A* (30 s, 1′, 5′, and 10′), *B* and *C* (1′, 5′, and 10′), and *E* and *F* (3′, 10′, and 20′).

component of the replication machinery, required to maintain progression on templates containing multiple lesions. To address this possibility, we assayed the capacity of Afu-Pol B and Pol B/PCNA to replicate DNA templates (77 nt) containing multiple 8-oxo dG lesions, in the presence or absence of Pri S/L (Fig. 3 *B* and *C*). Although Pol B and Pol B/PCNA fully extended primers on templates containing multiple 8-oxo dG, a regular pausing pattern in close proximity to the lesions was observed, compared with nondamaged template extension (Fig. 3 *B* and *C*, *Left* and *Middle*). This indicates that a significant slowing of the replicase was occurring. Overall, levels of DNA synthesis were enhanced by the addition of PCNA on nondamaged templates but not on templates containing multiple 8-oxo dGs. However, addition of Pri S/L restored rates of primer extension to those observed on undamaged primer templates (Fig. 3 *B* and 3*C*, *Right*).

Next, we examined the capacity of Afu-Pol B to replicate DNA templates containing multiple dUs. In contrast with templates containing multiple 8-oxo dGs, Pol B alone or assisted by PCNA was unable to traverse multiple dUs (Fig. 3*D*, *Left* and *Middle*). Again, addition of Pri S/L restored efficient DNA synthesis on this heavily damaged template (Fig. 3*D*, *Right*), indicating that the archaeal primase and polymerase may collaborate to maintain efficient replication fork progression on DNA containing multiple dU "roadblocks." To address whether the primase facilitates the maintenance of robust and processive DNA replication on much longer DNA templates, a ~1-kb-long DNA template was prepared containing multiple dUs (~20 lesions per template). As anticipated, Afu-Pol B alone or assisted by PCNA displayed highly processive synthesis on the undamaged template (Fig. 3*E*) and profound stalling on dU-containing templates (Fig. 3*F*, *Left* and *Middle*). However, addition of Afu-Pri S/L again rescued DNA synthesis by Afu-Pol B/PCNA complex (Fig. 3*F*, *Right*). Together, these data provide evidence that primases, in addition to their role in initiation of DNA synthesis, also actively participate in the elongation phase of DNA replication by performing TLS bypass of lesions thus preventing the archaeal replisome from arresting (Fig. 4). This interplay between replicase and primase appears to be important for processive synthesis to ensure timely DNA replication.

Concluding Remarks

It has been widely assumed that AEP-like primases evolved as replication enzymes relatively late in evolution, based on their

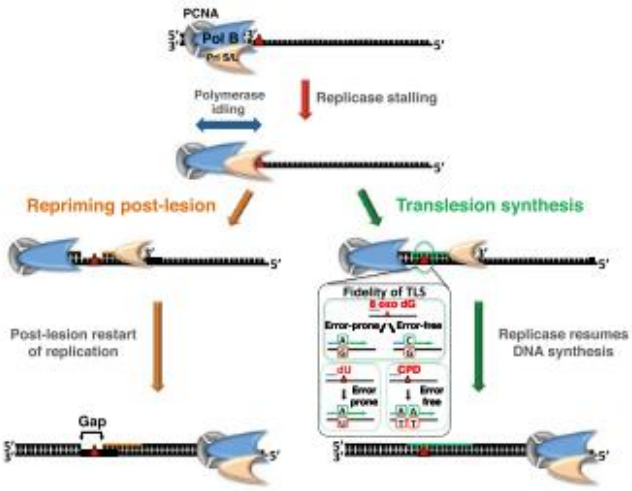


Fig. 4. Collaboration of the core components of the archaeal replisome results in bypass of DNA lesions. *Top* shows polymerizing replicase (Pol B, blue) with sliding clamp (PCNA, gray). The complex encounters blocking DNA lesion (red triangle) resulting in Pol B idling, which allows recruitment of the primase (Pri S/L, yellow). Depending on the type of the lesion, Pri S/L employs either primase or translesion synthesis (TLS) activity. *Left* illustrates a scenario where the blocking lesion is relatively large (i.e., large aromatic organic compound or protein covalently attached to DNA) and Pri S/L synthesizes a short primer (orange) after the damage so that the Pol B/PCNA complex can restart replication downstream from the block. Alternatively, when the blocking lesion is small (e.g., 8-oxo dG, dU, or CPD), TLS (green) is performed by Pri S/L so that the Pol B/PCNA complex resumes DNA synthesis. The fidelity of TLS performed by Afu-Pri S/L is shown.

initial discovery in eukaryotic organisms. However, it is now apparent that both the evolutionary origins and roles of enzymes belonging to the AEP superfamily must be significantly reevaluated. Primordial AEPs first originated in prokaryotes and bacteriophage to perform a diverse range of roles in DNA metabolism, including DNA repair (17, 18). Indeed, many bacterial species possess multiple AEP enzymes, in addition to possession of a replicative DnaG primase, suggesting an early diversification of functions within the AEP superfamily. During evolution of archaea from the last universal common ancestor (LUCA), both DnaG and AEPs were maintained but the essential role of DnaG as the replicative primase was superseded by AEP primases (30). Presumably, these enzymes offered particular advantages to these newly evolving organisms, including the capacity to traverse DNA lesions. Recently, a second AEP (PrimPol) has been identified in higher eukaryotes (10), whose primase and polymerase functions emulate those described here for archaeal replicative primases (19, 31). PrimPol was probably acquired from large cytoplasmic viruses during the early evolution of the eukarya (10), where it superseded Pri S's lesion bypass role during replication. This potentially allowed eukaryotic Pri S to assume more specialized roles in primer synthesis. This demarcation of the primase/polymerase activities between PrimPol and Pri S probably reflects the additional replication requirements for much larger genomes, such as more regulated lagging-strand synthesis.

Although some archaeal species possess canonical Y-family DNA polymerases (Dpo4), specialized in traversing DNA lesions, a recent study has reported that *Sulfolobus* strains lacking Dpo4 display no increased sensitivity to DNA damaging agents, including UV (32). This strain also exhibited no difference in rates of spontaneous mutagenesis, suggesting that other TLS pathways assist in bypassing replication-blocking lesions. Notably in this regard, although expression of the DNA replication genes is down-regulated after exposure of *Sulfolobus* cells to UV, Pri S/L was up-regulated after irradiation, supporting its proposed role in TLS (33). In addition, ethyl methanesulfonate (EMS) and UV treatments of *Pyrococcus* strains, which also lack canonical TLS polymerases, resulted in a strongly induced mutagenesis, with spontaneous mutation frequencies increased ~150-fold after EMS treatment and ~400-fold after UV exposure (26). The observed DNA damage induced mutagenesis indicates the possible existence of an active TLS pathway operating in these organisms. This predication is also supported by fractionation studies of whole cell extracts from *P. furiosus* that identified distinct polymerase activities in three major fractions (34). Notably, one of these fractions contained apparent TLS activity and, although the polymerase responsible for this was not identified, it was fully coincident with the elution of the Pri S/L complex. Together, with the findings presented here, these studies add support to our model postulating that archaeal replicative primases are involved in synthesis past DNA lesions during replication.

Since the discovery of canonical TLS polymerases, it has been widely assumed that these enzymes are largely responsible for TLS during replication (2). However, this study identifies that replicative primases can also act as proficient TLS polymerases that assist replicases in bypassing blocking DNA lesions during replication. As these enzymes are core replisomal factors, it argues that in many organisms the replisome is inherently TLS proficient and other damage tolerance mechanisms may provide additional assistance in a postreplicative manner. This discovery has major implications for our understanding of additional roles of DNA primases during replication and the subsequent evolution of related PrimPol-centric TLS pathways in eukaryotic cells (19, 31). Although this report has focused on the TLS activities of archaeal Pri S/L, these enzymes are also proficient DNA primases and therefore their ability to reprime replication postlesion (Fig. 4),

proposed for other organisms (4, 5, 19, 35), is also likely to be important.

Materials and Methods

Materials. Deoxynucleotide triphosphates (dNTPs) and deoxyuracil triphosphate (dUTP) were from by Roche and Jena Bioscience, respectively. Enzymes were provided by New England Biolabs.

Cloning of Archaeal Genes. *A. fulgidus* and *P. furiosus* genes were PCR amplified and cloned using standard molecular biology techniques. All of the PCR primers (Eurofins MWG Operon) sequences used in this study are listed in *SI Appendix*, Table S1.

Expression and Purification of Archaeal Proteins. Archaeal proteins were produced in *Escherichia coli* Rosetta strain (Novagene) grown at 37 °C. Typical purification procedure is composed of three chromatography steps: immobilized metal affinity, ion exchange, and gel filtration. Afu-PCNA was expressed and purified as previously described (36).

Synthetic Primer Templates. The oligodeoxynucleotides and oligoribonucleotides used to prepare primer templates were purchased from ATDBio and Eurofins MWG Operon, respectively. All primers were fluorescently labeled to aid visualization. Primer templates were annealed by heating equimolar amounts of the oligomers in 10 mM Tris, pH 7.5, 50 mM NaCl, and 0.5 mM EDTA at 95 °C for 5 min, followed by cooling slowly to room temperature. All primer-template sequences used in this study are listed in *SI Appendix*, Table S2.

Enzymatic Preparation of Long Single-Stranded DNA Templates. Approximately 1-kb-long single-stranded DNA (ssDNA) templates were prepared using PCR followed by strand-specific exonucleolytic degradation. Reaction mixtures contained primers listed in *SI Appendix*, Table S1. One of the primers was 5′ end phosphorylated to direct strand-specific degradation. The PCR comprised 50 μL of 1× Taq reaction, 200 μM of the four dNTPs, 1 μM of the forward and reverse primer, 50 ng of pUC18, and 20 units/mL of Taq DNA polymerase. To prepare ssDNA containing dUs, 20 μM of dUTP was added to the PCR mixture. A total of 30 PCR cycles (30 s at 95 °C, 35 s at 55 °C, and 1 min at 72 °C) were used. DNA was resolved and purified on 1% agarose/ethidium bromide gel. PCR product was excised and electroeluted. Recovered dsDNA was ethanol precipitated and the 5′ phosphorylated strand degraded using λ-exonuclease. ssDNA templates were annealed with 20-nt-long primer to yield nondamaged and dU-containing primer templates (*SI Appendix*, Table S3).

Primer Extension and Single Nucleotide Incorporation Assays. Reactions were performed in 20 μL volume containing 20 mM Tris, pH 8.8, 10 mM KCl, 10 mM $(NH_4)_2SO_2$, 2 mM $MgSO_4$, and 20 nM primer template. The polymerization was monitored for variety of archaeal enzymes, including 50 nM replicases and 50 nM of primases. Some of the reactions were supplemented with PCNA (200 nM). For running start extensions, 50 μM of each of the four dNTPs was used. In cases of single dNTP addition, 50 μM of the particular dNTP under investigation was added. All reactions were carried out at 50 °C and quenched by addition of an equal volume of 95% (vol/vol) formamide/5% (vol/vol) water containing 20 mM EDTA. Primer extensions were carried out for the time indicated in legends of Figs. 1–3 and *SI Appendix*, Figs. S1–S5. All single nucleotide incorporations were terminated after 5 min. Polymerization products were resolved on 15% (wt/vol) denaturing polyacrylamide gels containing 7 M urea. The gels were visualized using fluorescent scanner Fuji FLA-150.

ACKNOWLEDGMENTS. We thank Prof. B. Connolly and Dr. M. Reijns for providing us with archaeal genomic DNA, Prof. J. Tainer for the generous gift of the Afu-PCNA construct, and Prof. S. Iwai for providing the (6-4) photoproduct containing ssDNA. A.J.D.'s laboratory is supported by Grants BB/J018643/1 and BB/H019723/1 from the Biotechnology and Biological Sciences Research Council and Centre Grant G080130 from the Medical Research Council.

1. Aguilera A, Gómez-González B (2008) Genome instability: A mechanistic view of its causes and consequences. *Nat Rev Genet* 9(3):204–217.
2. Sale JE, Lehmann AR, Woodgate R (2012) Y-family DNA polymerases and their role in tolerance of cellular DNA damage. *Nat Rev Mol Cell Biol* 13(3):141–152.
3. Li X, Heyer W-D (2008) Homologous recombination in DNA repair and DNA damage tolerance. *Cell Res* 18(1):99–113.
4. Lopes M, Foiani M, Sogo JM (2006) Multiple mechanisms control chromosome integrity after replication fork uncoupling and restart at irreparable UV lesions. *Mol Cell* 21(1):15–27.
5. Heller RC, Marians KJ (2006) Replication fork reactivation downstream of a blocked nascent leading strand. *Nature* 439(7076):557–562.
6. Kelman Z, White MF (2005) Archaeal DNA replication and repair. *Curr Opin Microbiol* 8(6):669–676.
7. Greagg MA, et al. (1999) A read-ahead function in archaeal DNA polymerases detects promutagenic template-strand uracil. *Proc Natl Acad Sci USA* 96(16):9045–9050.
8. Richardson TT, Gilroy L, Ishino Y, Connolly BA, Henneke G (2013) Novel inhibition of archaeal family-D DNA polymerase by uracil. *Nucleic Acids Res* 41(7):4207–4218.
9. Wardle J, et al. (2008) Uracil recognition by replicative DNA polymerases is limited to the archaea, not occurring with bacteria and eukarya. *Nucleic Acids Res* 36(3):705–711.
10. Iyer LM, Koonin EV, Leipe DD, Aravind L (2005) Origin and evolution of the archaeo-eukaryotic primase superfamily and related palm-domain proteins: Structural insights and new members. *Nucleic Acids Res* 33(12):3875–3896.
11. Frick DN, Richardson CC (2001) DNA primases. *Annu Rev Biochem* 70:39–80.
12. Bocquier AA, et al. (2001) Archaeal primase: Bridging the gap between RNA and DNA polymerases. *Curr Biol* 11(8):452–456.
13. Lipps G, Röther S, Hart C, Krauss G (2003) A novel type of replicative enzyme harbouring ATPase, primase and DNA polymerase activity. *EMBO J* 22(10):2516–2525.
14. Lao-Sirieix S-H, Pellegrini L, Bell SD (2005) The promiscuous primase. *Trends Genet* 21(10):568–572.
15. Lao-Sirieix S-H, Bell SD (2004) The heterodimeric primase of the hyperthermophilic archaeon Sulfolobus solfataricus possesses DNA and RNA primase, polymerase and 3′-terminal nucleotidyl transferase activities. *J Mol Biol* 344(5):1251–1263.
16. Le Breton M, et al. (2007) The heterodimeric primase from the euryarchaeon Pyrococcus abyssi: A multifunctional enzyme for initiation and repair? *J Mol Biol* 374(5):1172–1185.
17. Weller GR, et al. (2002) Identification of a DNA nonhomologous end-joining complex in bacteria. *Science* 297(5587):1686–1689.
18. Della M, et al. (2004) Mycobacterial Ku and ligase proteins constitute a two-component NHEJ repair machine. *Science* 306(5696):683–685.
19. Bianchi J, et al. (2013) PrimPol bypasses UV photoproducts during eukaryotic chromosomal DNA replication. *Mol Cell* 52(4):566–573.
20. Lindahl T (1993) Instability and decay of the primary structure of DNA. *Nature* 362(6422):709–715.
21. Hübscher U, Maga G (2011) DNA replication and repair bypass machines. *Curr Opin Chem Biol* 15(5):627–635.
22. Lindahl T, Nyberg B (1974) Heat-induced deamination of cytosine residues in deoxyribonucleic acid. *Biochemistry* 13(16):3405–3410.
23. Fogg MJ, Pearl LH, Connolly BA (2002) Structural basis for uracil recognition by archaeal family B DNA polymerases. *Nat Struct Biol* 9(12):922–927.
24. Emptage K, O'Neill R, Solovyova A, Connolly BA (2008) Interplay between DNA polymerase and proliferating cell nuclear antigen switches off base excision repair of uracil and hypoxanthine during replication in archaea. *J Mol Biol* 383(4):762–771.
25. Bebio K, et al. (2011) Survival of thermophilic and hyperthermophilic microorganisms after exposure to UV-C, ionizing radiation and desiccation. *Arch Microbiol* 193(11):797–809.
26. Watrin L, Prieur D (1998) UV and ethyl methanesulfonate effects in hyperthermophilic archaea and isolation of auxotrophic mutants of Pyrococcus strains. *Curr Microbiol* 33(6):377–382.
27. Sinha RP, Häder DP (2002) UV-induced DNA damage and repair: A review. *Photochem Photobiol Sci* 1(4):225–236.
28. Voulgaridou GP, Anestopoulos I, Franco R, Panayiotidis MI, Pappa A (2011) DNA damage induced by endogenous aldehydes: current state of knowledge. *Mutat Res* 711(1-2):13–27.
29. Shuttleworth G, Fogg MJ, Kurpiewski MR, Jen-Jacobson L, Connolly BA (2004) Recognition of the pro-mutagenic base uracil by family B DNA polymerases from archaea. *J Mol Biol* 337(3):621–634.
30. Raymann K, Forterre P, Brochier-Armanet C, Gribaldo S (2014) Global phylogenomic analysis disentangles the complex evolutionary history of DNA replication in archaea. *Genome Biol Evol* 6(1):192–212.
31. Rudd SG, Glover L, Jozwiakowski SK, Horn D, Doherty AJ (2013) PPL2 translesion polymerase is essential for the completion of chromosomal DNA replication in the African trypanosome. *Mol Cell* 52(4):554–565.
32. Sakofsky CJ, Foster PL, Grogan DW (2012) Roles of the Y-family DNA polymerase Dbh in accurate replication of the Sulfolobus genome at high temperature. *DNA Repair (Amst)* 11(4):391–400.
33. Götz D, et al. (2007) Responses of hyperthermophilic crenarchaea to UV irradiation. *Genome Biol* 8(10):R220.
34. Ishino S, Ishino Y (2006) Comprehensive search for DNA polymerase in the hyperthermophilic archaeon, Pyrococcus furiosus. *Nucleosides Nucleotides Nucleic Acids* 25(4-6):681–691.
35. Keen BA, Jozwiakowski SK, Bailey LJ, Bianchi J, Doherty AJ (2014) Molecular dissection of the domain architecture and catalytic activities of human PrimPol. *Nucleic Acids Res* 42(9):5830–5845.
36. Dulbis JM, Kelman Z, Hurwitz J, O'Donnell M, Kuriyan J (1996) Structure of the C-terminal region of p21(WAF1/CIP1) complexed with human PCNA. *Cell* 87(2):297–306.