DNA REPLICATION IN ARCHAEA: PRIMING, TRANSFERASE, AND ELONGATION ACTIVITIES

by

Zhongfeng Zuo

Bachelor degree, Beijing Technology and Business University, 1999

Submitted to the Graduate Faculty of

the Kenneth P. Dietrich School of Arts and Sciences

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2012

UNIVERSITY OF PITTSBURGH

THE KENNETH P. DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Zhongfeng Zuo

It was defended on

January 27, 2012

and approved by

Stephen G. Weber, Professor, Department of Chemistry

Billy Day, Professor, Department of Chemistry, Department of Pharmacy

Renã A. S. Robinson, Assistant Professor, Department of Chemistry

Dissertation Advisor: Michael A. Trakselis, Assistant Professor, Department of Chemistry

DNA REPLICATION IN ARCHAEA: PRIMING, TRANSFERASE, AND ELONGATION ACTIVITIES

Zhongfeng Zuo, Ph.D

University of Pittsburgh, 2012

DNA REPLICATION IN ARCHAEA: PRIMING, TRANSFERASE, AND ELONGATION ACTIVITIES

Zhongfeng Zuo, Ph.D

University of Pittsburgh, 2012

We have biochemically characterized the bacterial-like DnaG primase contained within the hyperthermophilic crenarchaeon *Sulfolobus solfataricus* (*Sso*) and compared *in vitro* priming kinetics with those of the eukaryotic-like primase (PriS&L) also found in *Sso*. *Sso*DnaG exhibited metal- and temperature-dependent profiles consistent with priming at high temperatures. The distribution of primer products for *Sso*DnaG was discrete but highly similar to the distribution of primer products produced by the homologous *Escherichia coli* DnaG. The predominant primer length was 13 bases, although less abundant products of varying sizes are also present. *Sso*DnaG was found to bind DNA cooperatively as a dimer with a moderate dissociation constant. Mutation of the conserved glutamate in the active site severely inhibited priming activity, showing functional homology with *E. coli* DnaG. *Sso*DnaG was also found to have a greater than four-fold faster rate of DNA priming over that of *Sso*PriS&L under optimal *in vitro* conditions. The presence of both enzymatically functional primase families in archaea suggests that the DNA priming role may be shared on leading or lagging strands during DNA replication.

DNA replication polymerases have the inherent ability to faithfully copy a DNA template according to Watson Crick base pairing. The primary B-family DNA replication polymerase (Dpo1) in *Sso* is shown here to possess a remarkable DNA stabilizing ability for maintaining weak base pairing interactions to facilitate primer extension. This thermal stabilization by *Sso*Dpo1 allowed for template-directed synthesis at temperatures more than 30 °C above the melting temperature of naked DNA. Surprisingly, *Sso*Dpo1 also displays a terminal deoxynucleotide transferase (TdT) activity unlike any other B-family DNA polymerases. *Sso*Dpo1 is shown to elongate single stranded DNA in template-dependent and template-independent manners. The multiple activities of this unique B-family DNA polymerase make this enzyme an essential component for DNA replication and DNA repair for the maintenance of the archaeal genome at high temperatures.

Preliminary results of primer transfer studies in *Sso* show that *Sso*Dpo1 can elongate DNA primers *de novo* synthesized by *Sso*PriS&L, but elongation of RNA primers synthesized by both *Sso*PriS&L and *Sso*DnaG was not observed for *Sso*Dpo1. Future studies to untangle the primer transfer mechanism in archaea are discussed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

This thesis was formulated from years of research that has been done since I came to the Trakselis Lab. In the past few years, I have met and worked with a number of people who have contributed to the success of my research in assorted ways. It's a great pleasure to convey my deepest gratitude to them all in this humble acknowledgement.

In the first place, I would like to record my gratitude to Professor Michael Trakselis for his supervision, advice, and guidance from the early stage of my research to the final thesis writing. His genuine scientific intuition has been a constant source for ideas, passion, and encouragement, which has exceptionally inspired me and enriched my growth as a student, a researcher, and a scientist. The joy and enthusiasm he has for the research is contagious and motivational for me especially during the tough times of my Ph.D study. It's been an honor to be his first Ph.D student. He taught me how rigorous biochemical experiments should be designed, executed, observed and recorded. I greatly appreciate the time and efforts that he put to correct my writing, which I have improved considerably and will benefit me in my future. I am also thankful for the excellent example he provided being a successful and respectable young professional.

Many thanks go to the Trakselis group. I must pay tribute to the former members, Dr. Andry Mikhekin for the EMSA experiment in chapter 2, and Cory Rodgers for making mutated *Sso*DnaG and the fluorescence anisotropy assays in chapter 2. I am grateful to my classmate

Hsiang-Kai Lin who determined the melting temperatures of DNA in chapter 3 and provided purified *Sso*RFC and *Sso*PCNA. I'd like to thank Robert Bauer for generously providing the purified *Sso*Dpo1. I am also grateful for the advice, discussion, collaboration, and any kind of interactions with the whole group. Without any of them, I would not achieve this much in Chevron in the past few years.

I gratefully acknowledge the committee members, Professor Billy Day, Professor Steve Weber and Professor Robinson for their time, interest, and insightful questions. Special thanks go to Professor Billy Day for his care of me in the first summer at Pitt, and Professor Steve Weber for the core course he taught on chemical separations which not only deepened my knowledge on chemical separations but also further inspired my enthusiasm in this field. I also thank Dr. Robinson for her willingness to help me on the committee and the comments she provided during the discussions.

I greatly appreciate the supportive sources from the Department of Chemistry. I am grateful for the funding support from my advisor, the department, and the American Cancer Society. I also feel thankful for the working opportunities in mass spectrometry facilities in the department, for the teaching experience, and for the excellent service from the department main office, stock room, mechanic shop, glass shop, and electronic shop. Without the support from them, my Ph.D would become impossible.

Lastly, I would send thanks to my family for their love and encouragement. I tearfully thank my understanding parents and siblings in China for whom I have not seen face to face for more than five years mainly due to financial pressure and worries about visa application that prevented me from returning to my homeland. I thank them for their understanding and the optimistic attitude toward life that make me feel less guilty. Words fail me to express my

appreciation to my lovely wife Nan whose dedication, love, and persistent confidence in me, have encouraged me to finally reach where I am. I owe her for her sacrifice of an excellent career in software engineering in China and staying with me here while investing her coding talent in the kitchen room for the first three years.

1.0 INTRODUCTION

1.1 DNA REPLICATION

DNA replication is a fundamental biochemical process during which the parental duplex DNA is duplicated into two identical daughter DNA molecules each of which contains a strand from the parental molecule. This semi-conservative process, as proposed by Watson & Crick in 1953 (*1*), faithfully transfers the genetic information to the offspring. Accurate DNA replication involves a series of coordinated biochemical events that are completed by a variety of proteins. First, the origin recognition proteins bind to initiation sites of DNA replication and load the helicase (*2*). Once loaded to the replication origin, the helicase unwinds the double strand DNA and a replication fork forms as indicated in Figure 1-1. The resulting single strands serve as templates for the duplication of DNA. The exposed strands are coated by single-strand binding proteins (SSB) to prevent strand reannealing and to facilitate primase loading. The primase binds a specific priming initiation site on the template and synthesizes a short RNA strand called a primer. Primer synthesis is complimentary to the template and has a free 3'- hydroxyl group (*2*). The DNA polymerase hydrolyzes deoxyribonucleoside 5'-triphosphates releasing pyrophosphates and adds deoxyribonucleoside mono-phosphates to the growing chain at the 3' end of the primer guided by the template via the Watson-Crick base-pairing rules (*2*). The requirement of a free 3'-hydroxyl group for the polymerase to elongate the nascent strand

requires that DNA must be synthesized in different directions but with the same polarity on the two opened antiparallel (5'→ 3' and 3'→ 5') strands at the replication fork (*3*). The leading strand is synthesized continuously opposite the 3'→ 5' strand. while the lagging strand is synthesized discontinuously opposite the 5' → 3' stand (*4*). The primase synthesizes a primer *de novo* (i.e. without a free 3'-hydroxyl) which is then elongated by a lagging strand polymerase in a 5'→ 3' direction on the lagging strand resulting in small DNA units, termed Okazaki fragments. These fragements are ligated together later by a DNA ligase after removal of the RNA primer by a 5'→ 3' endonuclease.The overall synthesis direction of DNA lagging strand is 3'→ 5', and both nascent strands are synthesized with the same polarity as the replication fork advances.

The essential process of DNA replication is conserved through all domains of life, but the mechanistic details of enzymatic function and interactions has diverged between bacteria and eukaryotes. Intriguingly, the organisms within the third domain of life, termed Archaea, share properties to both bacteria and eukarya (*5*). Archaea are single-celled prokaryotic organisms which lack a nucleus similar to bacteria, but the DNA replication machinery is extremely homologous to that of higher eukaryotes (*6-8*). This homology allows us to probe DNA replication mechanisms in a relatively simple but highly relevant biological system where analogous replication processes in humans can be drawn.

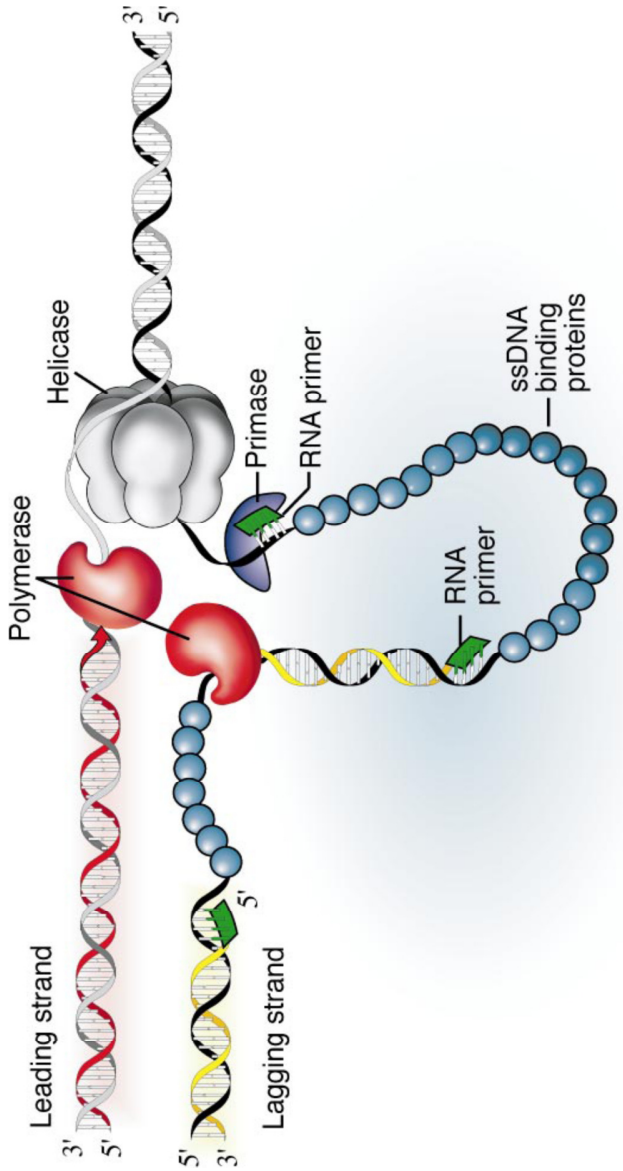

Figure 1-1. DNA replication fork (Reprinted from Reference 3)

1.2 PRIMASES AND PRIMING IN DNA REPLICATION

1.2.1 Primases and primers

DNA priming to initiate DNA replication is conserved throughout all three domains of life, although two separate families of primases have evolved: one group (DnaG family) contains all the primases from bacteria and their phages, while the other group consists of eukaryotic primases (Pri family). Curiously, phylogenetic analysis has revealed that both families of primases are conserved in archaea (3), but until now, only the PriS&L primase has been shown to have enzymatic activity (4, 5). All primases share many of the same biochemical properties (Table 1), but the enzymes in the two primase classes differ in sequence, structure, and with their interactions with other replication proteins.

1.2.1.1. Pri-type primases. The eukaryotic family primases are two-subunit complexes, including a small catalytic subunit (Pri1 in Figure 1-2A) and a large regulatory subunit (Pri2 in Figure 1-2A) that modulates both substrate binding and enzymatic activity. They are typically found complexed with two polymerase subunits: DNA polymerase α (Polα in Figure 1-2A) and polymerase B subunit (B in Figure 1-2A) to form the polymerase α/primase complex in eukaryotes. (*3*). The genes encoding all of the subunits of the Pol α/primase complex have been identified in humans, rats, mice, *Drosophila*, the yeasts *Sachharomyces cerevisiae* and *Schizosaccharomyces prombe*, (Table 1-1). This complex can synthesize RNA primers *de novo* and then extend them further by incorporating dNTPs into the primer by Pol α and Pol B. The

Table 1-1. Biochemical properties of purified DNA primases

Organism	Primase Gene	Mr (kD)	Primer length (nt)	Primer Sequence	Initiation Site[1]
Prokaryotic (DnaG-type)					
E. coli(9)	dnaG	65.6	9-14	5'-AG$(N)_{7-12}$	3'-<u>G</u>TC-5'
T7 (*10, 11*)	gene 4	62.7	4-5	5'-AC$(N)_{2-3}$	3'-<u>C</u>TC-5'
T4 (*12-14*)	gene 61	39.8	4-5	5'-AC$(N)_{2-3}$	3'-<u>T</u>TG-5'
Eukaryotic (Pri-type)					
S. cerevisiae	PRI 1	47.6	8-10	5'-(A/G)$(N)_{7-9}$	N/A
(15, 16)	PRI 2	62.3			
D.melanogaster	PRI 1	50.2	8-15	5'-(A/G)$(N)_{7-14}$	N/A
(17)	PRI 2	61.4			
Homo sapiens	PRI 1	49.9	11-14	5'-(A/G)$(N)_{8-13}$	N/A
(18)	PRI 2	58.8			
Archaeal (Pri-type)					
S. solfataricus	Pri L	35.7	<7000 [2]	N/A	N/A
(19)	Pri S	37.6	<1000[3]		
P. furiosus	p41	40.8	<700[3]	N/A	N/A
(20)	p46	46.0			

[1] Cryptic nucleotides not copied into the primer are underlined. Eukaryotic primases have not been shown to use specific recognition sites. [2] DNA primers [3] RNA primers

regulatory large subunit of the primase assists the priming activity of the small subunit by binding two initial ribonucleotides and ssDNA and counting the length of the synthesized RNA primer to generate primers of defined length (*21-24*). The "Pol β-like" domain in the large subunit of human primases was demonstrated to be responsible for this counting mechanism as the mutation of this domain severely interferes with the product length (*24*). The mature RNA primers (>7 nt) are transferred to Pol α for elongation into short stretch hybrid RNA-DNA oligonucleotides (~30nt) to form the pre-Okazaki fragments. In the absence of DNA polymerase α domain and dNTPs, the two subunits of the primase normally synthesize oligoribonucleotides with variable lengths ranging from 7 to 14 nt depending on the organism species (Table 1-1). In the presence of dNTPs, the complex can elongate the RNA primers into ssDNA up to several thousand nucleotides long. For instance, the human Polα/primase complex synthesizes RNA

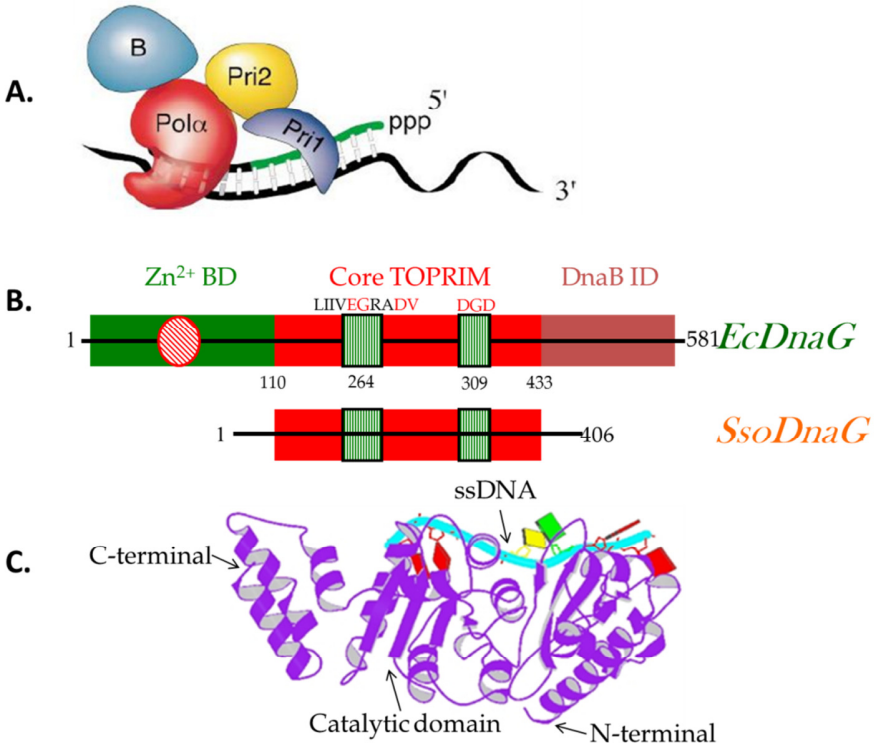

Figure 1-2. Primase Structures

A) The organization of the eukaryotic Pol α/primase complex. The Pol α/DNA polymerase (red) and its B subunit (blue) form a complex with the large primase subunit (yellow), encoded by the Pri2 gene, and the small primase subunit (purple), encoded by the Pri1 gene (Taken from Reference 3). B) Domain comparison of *Sso*DnaG with *Ec*DnaG with conserved TOPRIM domain labeled. *Sso*DnaG lacks both zinc binding domain and helicase binding domain C) Modeled Structure of *Ec*DnaG primase bound with ssDNA template ().

primers of 6-10 nucleotides long and then extends them to chains up to 3000 nucleotides long in the presence of dNTPs (*25*). Similar results have been reported for yeast, Drosophila and mice (*26-28*).

Eukaryotic-like primases from archaea contain only two of the four subunits of the polα primosome (small catalytic Pri S and large regulatory Pri L subunits) and lack additional polymerase subunits. In archaea, the small subunit has higher molecular weight than the large subunit, but the subunits are named according to the sequence homology in eukaryotes. These archaeal heterodimeric primases have been characterized in *Pyrococcus abyssi (Pab)*, *Pyrococcus horikoshii (Pho)*, *Pyrococcus furiosus (Pfu)* and *Sulfolobus solfataricus (Sso)*(*20, 29-32*). Archaeal primases don't follow the length-counting mechanism to synthesize primer products. Instead, they act like a hybrid of a primase and a polymerase to synthesize oligonucleotides up to 7 kb in length, though their ability to synthesize RNA products and DNA products varies (*19, 20, 33-36*). *Pab*PriS&L has a comparable affinity for dNTPs and NTPs and is able to synthesize a hybrid primer product (*8*). The *Pab*PriS subunit alone has no RNA synthesis activity but can synthesize long DNA strands *de novo* (up to 3 kb). Complexion with the *Pab*PriL subunit reveals the RNA synthesis capability (up to 200 bases) of the small subunit and increases the rate of DNA synthesis but decreases the length of the DNA products (less than 1kb) (*8, 20*). Similarly, *Pfu*PriS can also synthesize long DNA products (up to 7.2 kb), but can't synthesize RNA products. Like *Pab*PriS&L, *Pfu*PriS&L is able to synthesize RNA products (12-40 bases) and shorter DNA products (up to 700 bases) comparing with the *Pfu*PriS unit. However, RNA synthesizing activity of *Pfu*PriS&L is regulated by the concentration of dNTPs. When dNTPs are present at 10% relative to NTPs, the RNA products only account for 15% of the total products while equal concentration of dNTP and NTP completely abolishes RNA

synthesis of the enzyme (*20*). Likewise, *Sso*PriS&L can synthesize both RNA and DNA primer products, but it can synthesize 7-fold longer DNA products (up to 7kb) than RNA products (less than 1kb) at physiological temperatures (*19*). In addition, archaeal PriS&L is also able to extend a single strand DNA nonspecifically with a terminal transferase activity (*8, 19*) which is unknown to occur with any other eukaryotic primases.

1.2.1.2. DnaG-type primases. In contrast to multi-subunit eukaryotic primases, prokaryotic DnaG-type primases are single subunit enzymes that synthesize RNA primers ranging from 2 to 16 nucleotides long (Table 1-1) (*3, 37, 38*). The prokaryotic primase has three domains (Figure 1-2C); the N-terminal zinc binding domain, the central catalytic TOPRIM (TOPoisomerase-PRIMase) domain, and the C-terminal domain. The zinc binding domain was first identified from the primase of bacteriophage T7 (*39*). The 63-kDa full-length T7 primase contains zinc ions, while the 56-kDa N-terminal truncated primase is devoid of zinc ions. Mutation of any one of the four conserved cysteine residues in the zinc motif of the T7 primase abolishes primer synthesis ability. It was later found that the 63 kDa full-length T7 primase has both primase and helicase activity, but the 56 kDa truncated polypeptide only has helicase activity (*39*).

Like the T7 primase, the *E. coli* DnaG primase also contains a zinc motif in the N-terminus. Although removal of zinc ions from the protein doesn't affect the priming activity, the absence of zinc ions renders the DnaG primase more sensitive to oxidative agents and less stable in primase activity (*40*). The zinc-binding domain has been implicated in binding single-stranded DNA and may influence the primer length by interacting with polymerase domain of another primase molecule in the primase-helicase complex as determined for *E. coli* DnaG and DnaB (*3, 37, 41*).

All primases previously characterized contain a metal binding site. In prokaryotic primases, this binding site is located at the N-terminus, at the C-terminus in viral primases, and in the central region in the small subunit of eukaryotic primases (*3*). TOPRIM is a signature motif conserved in bacterial DnaG-type primases and topoisomerases (*42*). It has two conserved motifs, one of which centers at a conserved glutamate (E) and the other one at two conserved aspartates (DxD) (Figure 1-2B). The conserved glutamate residue in primases acts as a critical base for nucleotide polymerization. DxD motif in primases is responsible for Mg^{2+} or Mn^{2+} binding which is required for primase activity. The C-terminal domains of prokaryotic primases are not evolutionarily conserved.

The C-terminus of T7 has helicase activity and can interact with T7 DNA polymerase (*43, 44*). In contrast, the C-termini of both *E. coli* DnaG primase and T4 phage primase lack helicase activity but have helicase interacting domains that facilitate association with the cognate helicase to form a primosome complex which in turn activates priming activity (*45, 46*). The C-terminus of bacterial DnaG also directly regulates primer synthesis and the DNA unwinding activity of the DnaB helicase (*47, 48*). Primers synthesized by *E.coli* DnaG are typically 11 nucleotides in length but can range from 9-14 nucleotides (*9*). A comparison between bacterial and archaeal genomes revealed a TOPRIM domain in archaeal genome, but both N-terminal zinc binding domain and C-terminal helicase binding domain are absent (*42, 49*) (Figure 1-2B). The archaeal DnaG-type primase has little known about its structure and enzymatic function. Previously, archaeal DnaG has been linked to the RNA exosome complex by association in a pull-down assay performed in *Sulfolobus solfataricus* (*Sso*) (*50*). Recently, it has been shown that the *Sso* exosome is localized at the periphery of the cell and this localization is most likely mediated by *Sso*DnaG (*51, 52*).

1.2.1.3. Primer initiation sites. For all prokaryotic primases analyzed to date, primer synthesis doesn't happen randomly but instead at initiation sites on DNA templates. Primases initiate primer synthesis at specific template trinucleotides, 3'-XN_1N_2-5', where N1 and N2 are the first two nucleotides on the template directing the primer synthesis, while X is not replicated and only acts to direct the primase to the initiation site DnaG-type primases always utilize purine (A or G) as the 5'-terminal nucleotide, hence N1 on the template is always a pyrimidine. The compelling reason for this is because of the inherent difficulty of annealing the short primers to the template. The intial GC pairing during the initiation stage stabilizes the primers by the extra hydrogen bond. The primases that follow this initiation strategy include most DnaG-type primases such as those from *E. coli*, T7 and T4 phages as well as some Pri-type primases such as herpes primase (*37*). However, most of the Pri-type primases have minimal template specificity and only need a pyrimidine on the template to code for a purine at the 5'-end of the primers (*37*).

1.2.2 Mechanism of DNA priming

Frick and Richardson have proposed a plausible model for the mechanism of primer synthesis (*3*). According to this model, primers can be synthesized in a number of steps in all domains of life: template binding, NTP binding, initiation, extension to a functional primers, and primer transfer to DNA polymerase (*3, 37*) (Figure 1-3).

The primase first binds the DNA template and slides along the strand to locate priming initiation site (*53, 54*). Then the primase must bind two NTP substrates to catalyze the formation of a dinucleotide with inorganic pyrophosphate as a byproduct. In Frick's model, there are two NTP binding sites on the primase, an initiation binding site which binds an NTP that will be

placed at the 5' end of the primer and an elongation binding site which binds another NTP that will be added to the 3' end of the primer (*3*).

During each elongation step of primer synthesis, the synthesized oligonucleotide must be handed to the initiation binding site from the elongation binding site to make room for the incoming ribonucleotides. The first two NTPs may bind the primase in a particular order. It has been shown that the first NTP that binds the calf thymus primase becomes the second nucleotide in the primer (*54*). Based on this finding, Frick & Richardson proposed that the first NTP binds the elongation site on the primase to facilitate the search for priming recognition site on the DNA template, and the second NTP binds the initiation site in the primase, which will become the 5' end of the primer. T7 primase can synthesize primers that begin with 5'-AC at a rate of four dinucleotides per second, a rate sufficient for T7 DNA synthesis(*55*). Like calf thymus primase, T7 primase places the first bound CTP as the second base in the primers (*56*). Indeed, the binding of CTP and/or ATP to primase increases the affinity of T7 primase for DNA templates (*57*), which also suggests that the binding of the first nucleotide may help to locate the initiation sequence on the template as which was determined for calf thymus primase.

Following the dinucleotide formation, the primases elongate the initial dimer by continuously adding NTPs to the 3' end. The primases have evolved an ability to count and then produce primers with defined lengths (*3*). Structural studies from T7 and *E. coli* primases suggest that the zinc binding domain of one primase interacts with the catalytic domain of the neighboring primase to give this primase the ability to count the product length (*58, 59*). Similarly in humans, the Pol β-like domain in the large regulatory p58 subunit is able to modulate the length of primers. This Pol β-like domain binds the melted ssDNA functioning as a "hinge" to limit the movement of p58 subunit on the template and thus block further primer

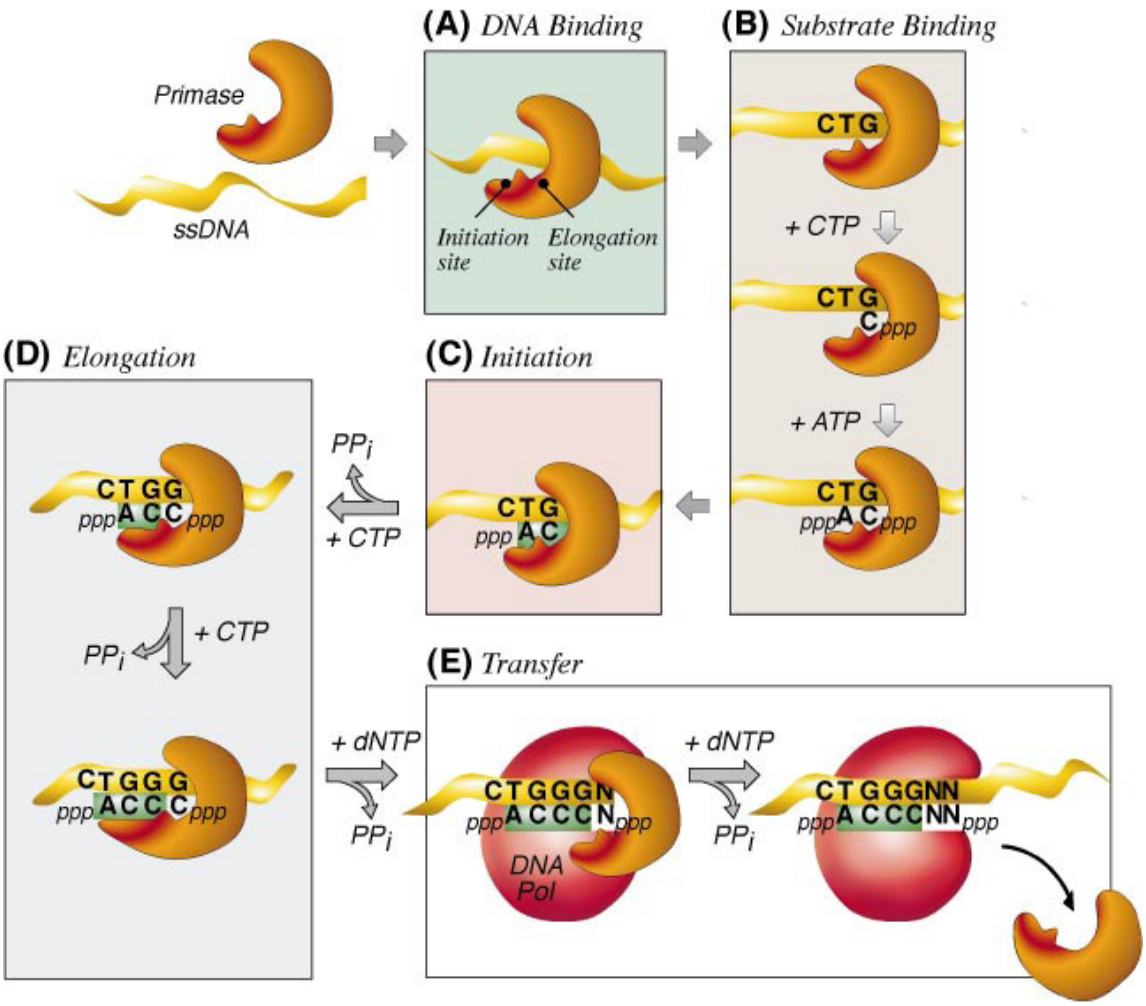

Figure 1-3. Model for Primer Synthesis

(*A*) DNA primase (*orange*) binds to ssDNA (*yellow*). (*B*) When the primase identifies an appropriate initiation site, it binds two NTPs. The first NTP binds to the elongation site, eventually becoming the second nucleotide in the primer. The second NTP binds to the initiation site, and is incorporated at the 5' end of the primer. (*C*) Primer synthesis is initiated by the formation of a dinucleotide and inorganic pyrophosphate from the two bound NTPs. (*D*) The growing oligoribonucleotide is transferred to the initiation site while additional NTPs bind to the elongation site; nucleotides are incorporated at the 3' end of the primer. (*E*) Primer RNA is transferred to the replicative DNA polymerase that adds deoxynucleotides derived from dNTPs to their 30 ends. (Reprinted from Reference 3)

synthesis by the small catalytic p49 subunit (*60*). Unlike all other primases, eukaryotic-like primases in archaea don't follow this counting mechanism (*6, 8, 19*). They can synthesize polynucleotides with wide distributions up to several thousand nucleotides (*37*). Once the primers mature to a certain length, they will be transferred to replicative DNA polymerases that add deoxynucleotide derived from dNTPs to their 3' ends.

1.2.3. Primer handoff

The primary biological function of the primase is to synthesize RNA primers for the DNA polymerase to elongate into DNA. There exist several primer-handoff models. In eukaryotic cells, the primase and the DNA polymerase form a tight holoenzyme complex. Therefore, the primase and polymerase activities are interconnected. In the presence of ssDNA template, NTPs, and dNTPs, the eukaryotic primase transfers the synthesized RNA primers directly to Pol α subunit without dissociating from the templates for dNTP extension. However, in the prokaryotic systems, the primase and polymerase transiently interact at the replication fork, and multiple primer handoff mechanisms have been evolved. The T7 primase directly transfers the primers as short as 4 nt to the T7 DNA polymerase (*61*). To circumvent the unstable annealing of these short primers with the template, the T7 DNA polymerase and T7 primase bind to each other and form a stable complex with the primed template, and thus the primer can be directly handed to the polymerase from the primase (*61*). However, the primer handoff event in *E. coli* is more complicated than the simple direct handoff in T7 phage. In these cells, efficient transfer of the primed template to DNA polymerase III requires two other cofactors: Pol III's processivity factor β (β sliding clamp) and γ clamp-loading complex (*62*). In this model, the primase has to interact with single strand binding protein (SSB) to remain on the primed template. Then the

clamp loader makes contacts with SSB and disrupts the primase-SSB interaction, followed by primase dissociation from the primed template. The γ complex loads β sliding clamp onto the primed template and then DNA Pol III core associates with β sliding clamp to form a processive DNA polymerase. The primer transfer in T4 phages is similar to the model for *E. coli* system in that both the clamp loader and clamp in T4 phages are involved in the transfer process, but the T4 primase can remain associated with replisome or dissociate from replisome depending on the rate of the clamp loader and clamp to enter the replisome. The T4 primase favors association with the replisome in the presence of the functional interaction between the primase and the single strand binding protein (gp32) (*63*). Unfortunately, the primer transfer mechanism in archaea from either the PriSL or the DnaG primases is currently unclear.

1.3 DNA POLYMERASES, DNA REPLICATION AND DNA REPAIR

A variety of DNA polymerases play individual roles in DNA replication and repair in all domains of life in an effort to maintain the genomic integrity of the cell. All of these DNA polymerases are classified into six main groups based upon sequence homology and structure similarities with *E. coli* Pol I (family A), *E. coli* Pol II (family B), *E. coli* Pol III (family C), Euryarchaeotic Pol II (family D), human Pol β (family X), and *E. coli* UmuC/DinB and eukaryotic RAD30/xeroderma pigmentosum variant (family Y) (*64*). Polymerases in each family play common roles in DNA replication or DNA repair in their respective organisms. DNA polymerases from the three domains of life are summarized in Table 1-2.

Table 1-2. DNA Polymerase Families

Family	Main enzymatic function	Domain		
		Bacteria	Eukarya	Archaea
A	*DNA replication/DNA repair*	*Pol I*	*ϑ, γ*	
B	*DNA replication/DNA repair*	*Pol II*	*α, δ, ε, ζ*	*Dpo1, Dpo2, Dpo3*
C	*Bacterial DNA replication*	*Pol III*		
D	*Archaeal DNA replication*			*Pol D (phyla specific)*
X	*DNA repair/immune diversity*		*β, λ, μ, TdT*	
Y	*Lesion bypass synthesis*	*Pol IV, Pol V*	*η, ι, κ, Rev1*	*Dpo4*

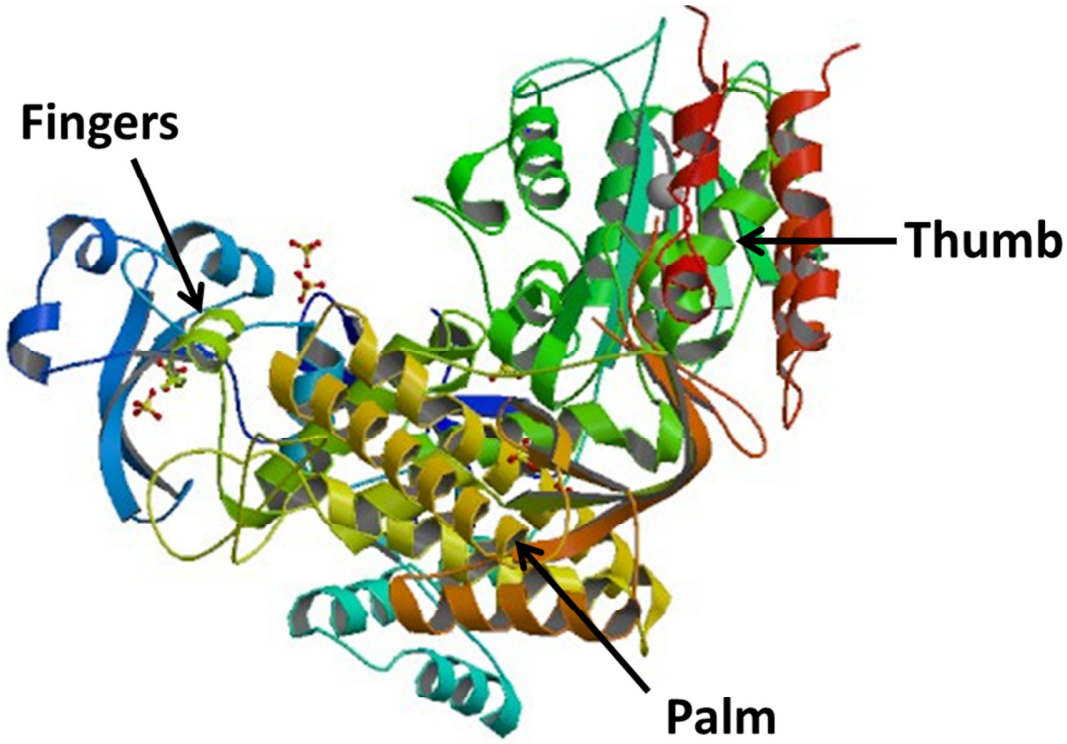

Figure 1-4. The Right Hand Conformation of *Sso*Dpo1

The fingers, palm and thumb domains are highlighted (PDBID: 1s5j).

It is well-known that all DNA polymerases harbor a conserved catalytic core which adopts a right-hand shape consisting of "thumb", "palm", and "fingers" domains (*65*) (Figure 1-4). The palm subdomain catalyzes the phosphoryl transfer reaction whereas the fingers domain is involved in correctly positioning the template and the incoming dNTPs. On the other hand, the thumb domain is important in DNA binding, processivity and translocation (*66*). Addition of different specific domains to the conserved core diversifies the DNA polymerases in terms of the cellular functions (*67*). In general, most polymerases in Families A and B are high fidelity enzymes primarily involved in faithful DNA replication and in repair of replication mistakes. High fidelity DNA polymerases have a proof-reading domain or 3'-exonuclease domain, which removes mismatched deoxyribonucleotide from the 3' end and effectively increases the DNA replication fidelity (*66, 67*). Family-Y polymerases lack an exonuclease domain, are less accurate, and typically participate in lesion bypass replication which allows cells to continue DNA replication by bypassing the lesion. Family-X polymerases are relatively inaccurate enzymes involved in a number of specific DNA repair processes, including base excision repair (BER) and double-strand break repair.

X-family polymerases in eukaryotic cells consist of DNA polymerase β, μ, λ and terminal deoxynucleotidyl transferase (TdT) (*68*). Though the polymerase domains of these proteins only range in primary sequence similarity from 23 to 44% (*69*), the structures of the X-family polymerases are remarkably similar. Like polymerases in other families, the four X-family polymerases each have a conserved polymerase domain which is comprised of thumb, palm and fingers domains. Within the polymerase domain of Pols β and λ, there is an extra 8-kD lyase subdomain which plays essential in BER. Pols λ, μ, and TdT each also have an N-terminal BRCT domain that is important for interaction with DNA substrates and end-binding proteins

during DNA double strand break repair and V(D)J recombination. V(D)J recombination is the rearrangement of antigen receptors to create unique antibody specificities to thwart biological attacks on the immune system (*70, 71*). The rearrangement requires the creation of double strand breaks within immunoglobin (Ig) genes and TdT activity to extend ssDNA, followed by reannealing of random gene regions to propagate immune diversity (*72, 73*). One of the key differences between family-X polymerases is the mechanism of template-dependent versus template-independent polymerization. Usually, TdT is a template-independent enzyme where nucleotides can be synthesized onto a 3' strand without direct base pairing. Pol β and Pol λ are primarily template-dependent enzymes, whereas Polμ has both template-dependent and template-independent activities (*69*). The gradient of template-dependence of X-family polymerases gives rise to the variation of the polymerizing fidelity. Pol β has the highest fidelity among them. Pol λ is slightly less accurate for single-base substitutions and much less accurate for single-base deletions (*74, 75*). Pol μ is highly error-prone for frame shifts and for substitutions via transient misalignment (*76, 77*). TdT has the lowest fidelity due to the template-independent nature. The relatively low fidelity of X-family polymerases is primarily because they lack the 3' → 5' exonuclease domain to proofread errors unlike the polymerases of families A and B.

Archaea lack X-family DNA polymerases. But the sequence comparison between eukayotic X-family polymerases and the small catalytic domain of archaeal eukaryotic-type primase has revealed some sequence homology (*78*). Structure comparison also showed that the highly conserved catalytic aspartate residue in the archaeal PriS&L primase can be superimposed on the catalytic core of Pol β (*79, 80*). The TdT activity of eukaryotic-like archaeal primases echoed the sequence and structure homology with X-family DNA polymerases and archaeal

PriS&L primases (*8, 19*). *Sso*PriS&L can utilize both NTPs and dNTPs to elongate ssDNA, but it shows higher TdT activity for NTPs substrates than dNTPs substrates. *Sso*PriS&L incorporates all NTPs to the 3' end of the four 20mer homo-oligomeric DNA templates, but it prefers to choose ribonucleotides complimentary to the templates as the substrates.

1.4 DNA REPLICATION IN ARCHAEA

Since the pioneering work from Carl Woese in 1970s (*81*), it has been well established that the archaea constitute a third domain of life distinct from eukaryotic and prokaryotic domains. With the availability of the genomic sequence information of archaeal organisms in late 1990s, it became apparent that archaea have strikingly similar DNA replication machinery to that in eukaryotes and clearly different from that in prokaryotes (*82-84*). Archaea have a simplified, and presumably ancestral, form of that in eukaryotic organisms. This simplicity yet similarity of the DNA replication machinery have evoked considerable interest to use archaea model organisms to study conserved DNA replication mechanism from eukaryotes.

In eukaryal primase/polymerase complex, the primase subunits synthesize short RNA primers (6–15 nt), which are then elongated by Polα to about 30 nt to form the pre-Okazaki fragment. This RNA-DNA hybrid is then extended by the replicative polymerase. Archaeal two-subunit primases are capable of initiating and synthesizing both RNA and DNA strands. In this light, the two subunits of the archaeal primase appear to combine the eukaryal primase/polymerase activities to synthesize RNA products and then elongate them into RNA-DNA hybrids, though evidence for this hypothesis is absent currently. With the finding of DnaG

primase genes in archaeal genomes (*49*), the DNA priming mechanism in archaea has become complicated as to which primase is functional *in vivo*.

DNA polymerases are required for chromosomal replication as well as repair and recombination. Thus, a frequently asked question is what cellular functions DNA polymerases play in cells. To be an accurate replicative DNA polymerase, the polymerase must have a 3'→ 5' exonuclease proofreading activity and its polymerase activity should be stimulated by interacting with cognate accessory proteins such as the sliding clamp, the clamp loader and singe-stranded binding proteins (*85*). By now, five DNA polymerases have been identified in this third domain of life (Tables 1-2). Depending on species, one to three B-family DNA polymerases, a D-family polymerase and a Y-family polymerase can be found in archaeal organisms (*86*). Only the euryarchaea phyla of archaea have D-family polymerases(*85*). *Sso* has three B-family polymerases, Dpo1, Dpo2 and Dpo3, and one Y-family polymerase, Dpo4. (*87*). Based upon *in vitro* findings, it is conceivable that Dpo1 is the major replicative DNA polymerases due to its fast and accurate rate of synthesis. Dpo2 and Dpo3 also possess exonuclease proofreading domains but are unlikely to be the major DNA polymerases due to their slower rates of synthesis (*87*). It is possible that other unknown components present *in vivo* may stimulate the polymerase activity of Dpo2 and Dpo3 and make them suitable for DNA replication to survive harsh environments, archaeal organisms must contain auxiliary factors to assure that accurate DNA replication occurs. These factors can take the form of specialized DNA repair enzymes to maintain the genome even in the absence of DNA replication. Therefore, the definitive cellular roles of all cellular DNA polymerases will be revealed once suitable genetic approaches are available or a systematic characterization of their enzymatic abilities are deciphered.

REFERENCES

1. Watson JD, Crick FHC. (1953) Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. *Nature.* 171, 737-738.

2. Marians KJ. (1992) Prokayotic DNA replication. *Annu. Rev. Biochem.* 61, 673-719.

3. Frick DN, Richardson CC. (2001) DNA primases. *Annu. Rev. Biochem.* 70, 39-80.

4. Hubscher PT. and Thommes U. (1990) Eukaryotic DNA replication, *Eur. J. Biochem.* 194, 699-712.

5. Woese CR, Kandler O and Wheelis ML. (1990) Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya, *Proc.Natl. Acad. Sci. U S A.* 87, 4576-4579.

6. Liu L, Komori K, Ishino S, Bocquier AA, Cann IK, Kohda D, Ishino Y. (2001) The Archael DNA primase: biochemical characterization of the p41-p46 complex from *Pyrococcus furiosus*., *J. of Bio. Chem.* 276, 45484-45490.

7. Bocquier AA, Liu L, Cann IK, Komori K, Kohda D, Ishino Y. (2001) Archaeal primase: bridging the gap between RNA and DNA polymerases. *Curr. Biol.* 11, 452-456.

8. Le Breton M, Henneke G, Norais C, Flament D, Myllykallio H, Querellou J, Raffin JP. (2007) The heterodimeric primase from the euryarchaeon *Pyrococcus abyssi*: a multifunctional enzyme for initiation and repair? *J. Mol. Biol.* 374, 1172-1185.

9. Zechner EL, Wu CA and Marians KJ. (1992) Coordinated leading- and lagging-strand synthesis at the *Escherichia coli* DNA replication fork. III. A polymerase-primase interaction governs primer size. *J. Biol. Chem.* 267, 4054-4063.

10. Romano LJ, Richardson CC. (1979) Requirements for synthesis of ribonucleic acid primers during lagging strand synthesis by the DNA polymerase and gene 4 protein of bacteriophage T7. *J. Biol. Chem.* 254, (20), 10476-82.

11. Scherzinger E, Lanka E, Morelli G, Seiffert D, Yuki A. (1977) Bacteriophage-T7-induced DNA-priming protein. A novel enzyme involved in DNA replication. *Eur. J. Biochem.* 72, 543-558.

12. Nossal NG. (1980) RNA priming of DNA replication by bacteriophage T4 proteins. *J. Biol. Chem.* 255, 2176-2182.

13. Kurosawa Y, Okazaki T. (1979) Structure of the RNA portion of the RNA-linked DNA pieces in bacteriophage T4-infected *Escherichia coli* cells. *J. Mol. Biol.* 135, 841-861.

14. Liu CC, Alberts BM. (1980) Pentaribonucleotides of mixed sequence are synthesized and efficiently prime *de novo* DNA chain starts in the T4 bacteriophage DNA replication system. *Proc. Natl. Acad. Sci. U S A.* 77, 5698-5702.

15. Badaracco G, Bianchi M, Valsasnini P, Magni G, Plevani P. (1985) Initiation, elongation and pausing of in vitro DNA synthesis catalyzed by immunopurified yeast DNA primase: DNA polymerase complex. *EMBO.* 4, 1313-1317.

16. Badaracco G, Valsasnini P, Foiani M, Benfante R, Lucchini G, Plevani P. (1986) Mechanism of initiation of in vitro DNA synthesis by the immunopurified complex between yeast DNA polymerase I and DNA primase. *Eur. J. Biochem.* 161, 435-440.

17. Conaway RC, Lehman IR. (1982) A DNA primase activity associated with DNA polymerase alpha from *Drosophila melanogaster* embryos. *Proc. Natl. Acad. Sci. U S A.* 79, 2523-2527.

18. Wang TS, Hu SZ, Korn D. (1984) DNA primase from KB cells. Characterization of a primase activity tightly associated with immunoaffinity-purified DNA polymerase-alpha. *J. Biol. Chem.* 259, 1854-1865.

19. Lao-Sirieix SH, 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, 1251-1263.

20. Liu L, Komori K, Ishino S, Bocquier AA, Cann IK, Kohda D, Ishino Y. (2001) The archaeal DNA primase: biochemical characterization of the p41-p46 complex from Pyrococcus furiosus. *J. Biol. Chem.* 276, 45484-45490.

21. Sauguet L, Klinge S, Perera RL, Maman JD, Pellegrini L. (2010) Shared active site architecture between the large subunit of eukaryotic primase and DNA photolyase. *PLoS One 5*, e10083.

22. Vaithiyalingam S, Warren EM, Eichman BF, Chazin WJ. (2010) Insights into eukaryotic DNA priming from the structure and functional interactions of the 4Fe-4S cluster domain of human DNA primase. *Proc. Natl. Acad. Sci.* 107, 13684-13689.

23. Weiner BE, Huang H, Dattilo BM, Nilges MJ, Fanning E, Chazin WJ. (2007) An iron-sulfur cluster in the C-terminal domain of the p58 subunit of human DNA primase. *J. Biol. Chem.* 282, 33444-33451.

24. Zerbe LK, Kuchta RD. (2002) The p58 subunit of human DNA primase is important for primer initiation, elongation, and counting. *Biochem.* 41, 4891-4900.

25. Wang T, Hu SZ, Korn D. (1984) DNA primase from KB cells: Characterization of a primase activity tightly associated with immunoaffinity-purified DNA polymerase-alpha. *J. Biol. Chem.* 259, (3), 1854-1865.

26. Badaracco G, Bianchi M, Valsasnini P, Magni G, Plevani P. (1985) Initiation, elongation and pausing of in vitro DNA synthesis catalyzed by immunopurified yeast DNA primase: DNA polymerase complex. *EMBO.* 4 (5), 1313-1317.

27. Conaway R, Lehman IR. (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, 2523-2527.

28. Prussak C, Almazan MT, Tseng BY. (1989) Mouse primase p49 subunit molecular cloning indicates conserved and divergent regions. *J. Biol. Chem.* 264, (9), 4957-4966.

29. Lao-Sirieix SH, 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, 1251-1263.

30. Ito N, Matsui I, Matsui E. (2007) Molecular basis for the subunit assembly of the primase from an archaeon *Pyrococcus horikoshii*. *FEBS.* 274, 1340-1351.

31. Kuchta RD, Stengel G. (2010) Mechanism and evolution of DNA primases. *Biochim. Biophys. Acta*. 1804, 1180-1189.

32. Rowen L, Kornberg A. (1978) A ribo-deoxyribonucleotide primer synthesized by primase. *J. Biol. Chem.* 253, 770-774.

33. Le Breton M., Henneke G, Norais C, Flament D, Myllykallio H, Querellou J and Raffin JP. (2007) The heterodimeric primase from the euryarchaeon *Pyrococcus abyssi*: a multifunctional enzyme for initiation and repair. *J. Mol. Biol.* 374, 1172-1185.

34. Ito N, Matsui I and Matsui E. (2007) Molecular basis for the subunit assembly of the primase from an archaeon *Pyrococcus horikoshii*. *FEBS.* 274, 1340-1351.

35. Matsui E, Nishio M, Yokoyama H, Harata K, Darnis S and Matsui I. (2003) Distinct domain functions regulating *de novo* DNA synthesis of thermostable DNA primase from hyperthermophile *Pyrococcus horikoshii*. *Biochem.* 42, 14968-14976.

36. Lao-Sirieix SH, Nookala RK, Roversi P, Bell SD and Pellegrini L. (2005) Structure of the heterodimeric core primase. *Nat. Struct. Mol. Biol.* 12, 1137-1144.

37. Kuchta RD, Stengel G. (2010) Mechanism and evolution of DNA primases, *Biochim. Biophys. Acta.* 1804, 1180-1189.

38. Rowen L, Kornberg A. (1978) A ribo-deoxyribonucleotide primer synthesized by primase. *J. Biol. Chem. 253*, 770-774.

39. Mendelman LV, Richardson CC. (1994) Requirement for a zinc motif for template recognition by the bacteriophage T7 primase. *EMBO. 13*, 3909-3916.

40. Griep MA, Lokey ER. (1996) The role of zinc and the reactivity of cysteines in *Escherichia coli* primase. *Biochem.* 35, 8260-8267.

41. Pan H, Wigley DB. (2000) Structure of the zinc-binding domain of *Bacillus stearothermophilus* DNA primase. *Structure.* 8, 231-239.

42. Aravind L, Leipe DD and Koonin EV. (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, 4205-4213.

43. Ziegelin G, Lurz R, Lanka E. (1997) The helicase domain of phage P4 alpha protein overlaps the specific DNA binding domain. *J. Bacteriol.* 179, 4087-4095.

44. Nakai H, Richardson CC. (1986) Interactions of the DNA polymerase and gene 4 protein of bacteriophage T7. Protein-protein and protein-DNA interactions involved in RNA-primed DNA synthesis. *J. Biol. Chem.* 261, 15208-15216.

45. Valentine AM, Ishmael FT, Shier VK and Benkovic SJ. (2001) A zinc ribbon protein in DNA replication: primer synthesis and macromolecular interactions by the bacteriophage T4 primase. *Biochem.* 40, 15074-15085.

46. Tougu K, Marians KJ. (1996) The extreme C terminus of primase is required for interaction with DnaB at the replication fork. *J. Biol. Chem.* 271, 21391-21397.

47. McMacken R, Kornberg A. (1977) Migration of Escherichia coli dnaB protein on the template DNA strand as a mechanism in initiating DNA replication. *Proc. Natl. Acad. Sci. U S A.* 74, (10), 4190-4194.

48. Bird LE, Soultanas P, Wigley DB. (2000) Mapping protein-protein interactions within a stable complex of DNA primase and DnaB helicase from Bacillus stearothermophilus, *Biochem.* 39, 171-182.

49. Koonin EV, Galperin MY, Walker DR. (1997) Comparison of archaeal and bacterial genomes: computer analysis of protein sequences predicts novel functions and suggests a chimeric origin for the archaea. *Mol. Microbiol.* 25, 619-637.

50. Walter P, Klein F, Lorentzen E, Ilchmann A, Klug G. and Evguenieva-Hackenberg E. (2006) Characterization of native and reconstituted exosome complexes from the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Mol. Microbiol.* 62, 1076-1089

51. Roppelt V, Albers SV, Lassek C, SchwarzH, Klug G. (2010) The archaeal exosome localizes to the membrane. FEBS Lett. 584:2791-2791.

52. Evguenieva-Hackenberg E, Lassek C and Klug G. (2011) Subcellular localization of RNA degrading proteins and protein complexes in prokaryotes. *RNA Biol.* 81, 49-54.

53. Masai H, Arai K. (1996) Mechanisms of primer RNA synthesis and D-loop/R-loop-dependent DNA replication in *Escherichia coli*. *Biochimie.* 78, 1109-1117.

54. Sheaff RJ, Kuchta RD. (1993) Mechanism of Calf Thymus DNA Primase: Slow Initiation, Rapid Polymerization, and Intelligent Termination. *Biochem.* 32, 3027-3037.

55. Lechner RL, Richardson CC. (1983) A preformed, topologically stable replication fork. Characterization of leading strand DNA synthesis catalyzed by T7 DNA polymerase and T7 gene 4 protein. *J. Biol. Chem.* 258, 11185-11196.

56. Kusakabe T, Richardson CC. (1996) The role of the zinc motif in sequence recognition by DNA primases. *J. Biol. Chem.* 271, 19563-19570.

57. Frick DN, Richardson CC. (1999) Interaction of bacteriophage T7 gene 4 primase with its template recognition site. *J. Biol. Chem.* 274, 35889-35898.

58. Qimron U, Hamdan SM, Richardson CC. (2006) Primer initiation and extension by T7 DNA primase. *EMBO.* 17, 2199-2208.

59. Corn JE, Hura GL, Berger JM. (2005) Crosstalk between primase subunits can act to regulate primer synthesis in trans. *Mol. Cell.* 20, 391-401.

60. Singhal RK, Wilson SH. (1993) Short gap-filling synthesis by DNA polymerase beta is processive. *J. Biol. Chem.* 268, 15906-15911.

61. Kato M, Ito T, Wagner G, Ellenberger T. (2004) A molecular handoff between bacteriophage T7 DNA primase and T7 DNA polymerase initiates DNA synthesis. *J. Biol. Chem.* 279, 30554-30562.

62. Yuzhakov A, Kelman Z, O'Donnell M. (1999) Trading places on DNA — a three-point switch underlies primer handoff from primase to the replicative DNA polymerase. *Cell* 96, 153-163.

63. Nelson SW, Kumar R, Benkovic SJ. (2008) RNA primer handoff in bacteriophage T4 DNA replication: the role of single-stranded dna-binding protein and polymerase accessory proteins. *J. Biol. Chem.* 283, 22838-22846.

64. Burgers PMJ, Bruford E, Blanco L, Christman MF, Woodgate R, Friederg EC, Hanaoka F, Hinkle D, Lawrence CW. (2001) Eukaryotic DNA polymerases: proposal for a revised nomenclature. *J. Biol. Chem.* 276, 43478-43490.

65. Steitz A. (1999) DNA polymerases: structural diversity and common mechanisms. *J. Biol. Chem.* 274, 17395-17398.

66. Hubscher U, and Spadari S. (2002) Eukaryotic DNA Polymerases. *Annu. Rev. Biochem.* 71, 133-163.

67. Brautigam CA, Steitz TA. (1998) Structural and functional insights provided by crystal structures of DNA polymerases and their substrate complexes. *Curr. Opin. Struct. Biol.* 8, 54-63.

68. Yamtich J, Sweasy JB. (2010) DNA polymerase Family X: Function, structure, and cellular roles. *Biochim. Biophy. Acta.* 1804, 1136-1150.

69. Moon AF, Batra VK, Beard WA, Bebenek K, Kunkel TA, Wilson SH, Pedersen LC. (2007) The X family portrait: structural insights into biological functions of X family polymerases. *DNA Repair.* 6, 1709-1725.

70. Tonegawa S. (1983) Somatic generation of antibody diversity. *Nature.* 302, 575-581.

71. Lieber M, Lu H, Gu J, and Schwarz K. (2008) Flexibility in the order of action and in the enzymology of the nuclease, polymerases, and ligase of vertebrate non-homologous DNA end joining: relevance to cancer, aging, and the immune system. *Cell Res.* 18, 125-133.

72. McElhinny N, Havener JM, Garcia-Diaz M, Juarez R, Bebenek K, Kee BL, Blanco L, Kunkel TA, and Ramsden DA. (2005) A gradient of template dependence defines distinct biological roles for family X polymerases in nonhomologous end joining. *Mol. Cell.* 19, 357-366.

73. Bertocci B, De SA, Berek C, Weill JC and Reynaud CA. (2003) Immunoglobulin kappa light chain gene rearrangement is impaired in mice deficient for DNA polymerase mu, *Immunity.* 19, 203-211.

74. Bebenek K, Blanco L, Kunkel TA. (2003) The frameshift infidelity of human DNA polymerase lambda: Implications for function. *J. Biol. Chem.* 278, 34685-34690.

75. Picher AJ, Bebenek K, Pedersen LC, Kunkel TA, Blanco L. (2006) Promiscuous mismatch extension by human DNA polymerase lambda. *Nucleic Acids Res.* 34, 3259-3266.

76. Zhang Y, Yuan F, Xie Z, Wang Z. (2001) Highly frequent frameshift DNA synthesis by human DNA polymerase mu. *Mol. Cell. Biol.* 21, 7995-8006.

77. Tippin B, Bertram JG, Goodman MF. (2004) To slip or skip, visualizing frameshift mutation dynamics for error-prone DNA polymerases. *J. Biol. Chem.* 279, 45360-45368.

78. Kirk BW, Kuchta RD. (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. *Biochem.* 38, 7727-7736.

79. Augustin MA, Huber R, Kaiser JT. (2001) Crystal structure of a DNA-dependent RNA polymerase (DNA primase). *Nat. Struct. Bio.* 8, 57-61.

80. Sawaya MR, Kumar A, Wilson SH, Kraut J. (1994) Crystal structure of rat DNA polymerase beta: evidence for a common polymerase mechanism. *Science.* 24, 1930-1935.

81. Woese CR, Fox GE. (1977) Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc. Natl. Acad. Sci. U S A.* 74, 5088-5090.

82. Kelman LM, Kelman Z. (2003) Archaea: an archetype for replication initiation studies? *Mol. Microbiol.* 48, 605-615.

83. Kelman Z. (2000) DNA replication in the third domain (of life). *Curr. Protein Peptide Sci.* 1, 139-154.

84. MacNeill SA. (2001) Understanding the enzymology of archaeal DNA replication: progress in form and function. *Mol. Microbiol.* 40, 520-529.

85. Grabowski B, Kelman Z. (2003) Archaeal dna replication: Eukaryal proteins in a bacterial context. *Annu. Rev. Microbiol.* 57, 487-517.

86. Böhlke K, Pisani FM, Rossi M, Antranikian G. (2002) Archaeal DNA replication: spotlight on a rapidly moving field. *Extremophiles*. 1-14.

87. Choi JY, Eoff RL, Pence MG, Wang J, Martin MV, Kim EJ, Folkmann LM, Guengerich FP. (2011) Roles of the four DNA polymerases of the crenarchaeon *Sulfolobus solfataricus* and accessory proteins in DNA replication. *J. of Bio. Chem.* 286, 31180-31193.

2.0 CHARACTERIZATION OF *SSO*DNAG: IMPLICATIONS FOR A DUAL PRIMASE SYSTEM[1]

The ability to initiate DNA synthesis by a DNA polymerase requires the generation of a primer stably bound to the template strand. Once the DNA replication origin is unwound by the DNA helicase to reveal the single strands, the DNA primase *de novo* synthesizes an RNA primer complementary to each unwound single strand; this primer can then be further extended by a DNA polymerase. This process is conserved in all three domains of life. In archaea, both classes of primases are found. The archaeal two-subunit Pri S&L primases have been well characterized in *Pab*, *Pho*, *Pfu* and *Sso* as discussed in the Introduction. In contrast, the bacterial-type DnaG primase in archaea has not been enzymatically or structurally studied since the gene encoding this primase was revealed(*1*). Previously, archaeal DnaG has been linked to the RNA exosome complex by association in a pull-down assay performed in *Sulfolobus solfataricus* (*Sso*), although no enzymatic role for *Sso*DnaG in the exosome has been shown (*2, 3*). Here we show that the DnaG primase from the archaeon *Sulfolobus solfataricus* can synthesize RNA primers of a similar but distinct size to that of *E. coli* DnaG. The activity of *Sso*DnaG is metal dependent and is optimal at high temperatures and roughly neutral pH. *Sso*DnaG exhibits moderate DNA binding consistent with bacterial DnaG that includes sequential and cooperative dimer formation.

[1] This chapter is primarily taken from *Journal of Molecular Biology,* 2010, 397, 664–676, by Zuo Z *et al*.

A conserved mutation in the active site shown previously to eliminate priming activity in *E. coli* reduces *Sso*DnaG priming activity to just above background, highlighting active site homology with bacterial primases. We also compared the priming rates between the two families of primases found in *Sso* and showed that *Sso*DnaG has a four-fold faster rate of primer synthesis. It was thought that these two primase families might have evolved independently, but inclusion of both primases in Archaea suggests that this might be an example of an evolutionary transformation in RNA priming. The consequences of the DNA priming roles for each primase during DNA replication in archaea are discussed.

2.1 MATERIALS AND METHODS

Materials

^{32}P labeled dNTP and NTP were purchased from PerkinElmer Life Sciences (Norton, Ohio). Unlabeled deoxynucleotides and ribonucleotides were purchased from Invitrogen (Carlsbad, CA). pET30a vector was obtained from Novagen (Madison, WI). Oligonucleotide primers were ordered from IDT (Coralville, IA) (Table 2-1). Single-stranded M13mp18 was purchased from USB (Cleveland, OH). All restriction enzymes are from New England Biolabs (Ipswich, MA). Both Pfx50 polymerase and T4 DNA ligase are from Invitrogen. Other chemicals were analytical grade or better.

Polymerization chain reaction (PCR) for DnaG

In a typical 25 µl reaction, about 75 ng of template DNA, 3 µM primers, 0.3 mM dNTPs, 1 mM Mg^{2+}, and 1.5 units of Pfx50 polymerase was used. The reaction started at 94 °C for 2 minutes, continued for 35 cycles of denaturation at 94 °C for 30 seconds, annealing at 50 °C for 40 seconds, extension at 70 °C for 60 seconds, and was ended with a final extension at 70 °C for 3 minutes.

Table 2-1. PCR primers used

DanG Forward primer	5'-CACC**CATATG**AGCTTCCCAATGAAATATGAT
DnaG Reverse primer	5'-ATTA**CTCGAG**AGAAGAAATAATATCGGTAAA
Reverse primer with stop codon	5'- ATTA**CTCGAGCTA**AGAAGAAATAATATCGGTAA
E175Q Forward primer	5'-CCTAATTTAATAATAGTA**C**AAGGAAGAGCTGATGTAA
E175Q Reverse primer	5'-TTACATCAGCTCTTCCTT**G**TACTATTATTAAATTAGG

* Red sequences are restriction sites. ** CTA in blue is the transcription stop codon *** The mutated sites were marked red in E175Q primers

Overlap extension PCR — DnaG-E175Q mutation

Two 50 µl reactions were run with one using a primer pair of the DnaG forward primer and E175Q reverse primer and another one using a primer pair of E175Q forward primer and DnaG reverse primer. All the reaction components had concentrations as described above and the same PCR temperature cycles were used. The two PCR products were purified and annealed using 10 cycles of above temperature program in the presence of all the reaction components but the DNA template and primers. The DnaG forward primer and the DnaG reverse primer were then added to the reaction at a concentration of 2 ng/µl. The reaction continued for 25 cycles.

Purification of DnaG gene after PCR

PCR solution was loaded on a 1% agarose gel and electrophoresed at a constant voltage of 110 volts for 40 minutes. The bands on the gel at the size of DnaG gene (46kD) were cut out under UV transilluminator. Then the gene product was isolated and purified using the Gel Purification Kit (Invitrogen).

Construction of expression vectors

Both the DnaG gene and pET30a plasmid vector were cut similarly using DNA restriction enzymes of *NdeI* and *XhoI*. Concentrations of digested DnaG gene and vector were estimated from a 1% agarose gel. The digested DnaG gene was ligated into the vector pET30a at a 5:1 ratio using T4 DNA ligase at 16 °C overnight. A hexa-histadine tag contained in the pET30a plasmid was included at the C-terminal end of recombinant DnaG . 2 µl of the ligation solution was incubated with 50 µl of competent XL10 Gold *E. coli* cells on the ice for 20 minutes followed by heat treatment in 42 °C water bath for 1 minute and then placed on the ice for 2 minutes. 600 µl of SOC medium was added to the cells solution and was then incubated in a 37 °C shaker for one hour. 50 to 150 µl of cells was spread on 2xYT plates containing 50 µg/ml kanamycin. The plate was incubated at 37 °C over night. Individual colonies were picked using sterile toothpicks and grew over night in 3 ml of the liquid 2xYT media containing 50 µg/ml kanamycin. The vector carrying DnaG gene (pET30a-DnaG-His) was extracted using FastPlasmid Mini prep Kit (5 Prime) and then digested to screen for the correct ligated product using Xba I which cuts at the site of number 384 nucleotide on the vector and the site of number 967 nucleotide on DnaG gene. The sizes of the two bands expected on the agarose gel for positively ligated expression vectors are 1005bp and 5324bp. The positive vector was amplified

by growing positive cells in 100ml of 2xYT media with same antibiotics and purified using a Midiprep Kit (Invitrogen). DNA sequencing was verified by the Genomics and Proteomics Core Laboratory (GPCL) at the University of Pittsburgh.

Protein Purification

pET30a-*Sso*DnaG or pET30a-*Sso*DnaG(E175Q) were transformed into BL21DE3 Rosetta 2 cells (Stratagene) and grown in 2xYT media supplemented with kanamycin and chloramphenicol and 0.1% glycerol at 37 °C and induced at room temperatures with 0.3 mM IPTG at OD_{600} ~ 0.6 for 16 hours. The cell pellets were resuspended in a solution containing 50 mM HEPES (pH 8.0), 100 mM NaCl, and 10 mM β-mercaptoethanol, 10% glycerol and a cocktail of protease inhibitors (Roche) and were lysed by sonication. After centrifugation, the supernatant was heat treated at 75 °C for 20 minutes. The supernatant was purified by HisTrap HP and HiTrap HP Q columns (GE Healthscience). Final cleanup and size selection was performed using a Superdex 200 26/60 gel filtration column. The extinction coefficient for *Sso*DnaG was calculated to be 22,900 $M^{-1}cm^{-1}$(*4*).

*Sso*DnaG protein was verified by mass spectrometry. The *Sso*DnaG band from SDS-PAGE gel was cut and digested by trypsin following the in-gel digestion protocol(*5*). The digested solution was analyzed by MALDI/TOF (Applied Biosystems, GPCL at the University of Pittsburgh) and data were searched against Mascot and confirmed with a significant threshold less than 0.05 (Matrix Science).

pETDuet-1-*Sso*PriS&L was provided by Steve Bell (Oxford University) and was transformed into BL21DE3 Rosetta 2 cells (Stratagene) and purified as previously described[4] with the following modifications. The Superdex 200 26/60 column was used as final step to

remove the minor impurities and monitor the correct size of monomeric Pri S&L. The plasmid containing *Ec*DnaG was provided by James Berger (E. California, Berkeley) and was purified as previously described (*6*).

Primase Activity Assay

The standard assay mixture contained *Sso*DnaG reaction buffer (50 mM Tris, pH 8.5), 1 mM $MnSO_4$, 0.1 mM NTPs, including 0.025 μCi/μL [α-^{32}P] GTP, and 4 nM M13 ssDNA unless indicated otherwise in the Figure Legends. Experiments were conducted at temperatures, primase and NTP concentrations as indicated in the Figure Legends. Aliquots of the reaction solution were quenched in an equal volume of stop solution (88% formamide, 10 mM EDTA, 1mg/ml bromophenol blue) at various time points to stop the reaction. Products were separated on a 20% (w/v) denaturing polyacrylamide gel (8 M urea in 1x TBE). The gel was exposed to phosphor screen and analyzed by autoradiography using a Storm PhosphorImager and included ImageQuant software (v. 5.0) (GE Healthscience). Various volumes of the reaction were spotted separately and phosphorimaged at the same time to calculate a standard curve of ^{32}P intensity versus the pmols of NTPs. Quantification of the primers in the priming reaction were calculated from the standard curve. The concentration of enzyme or NTPs was systematically varied as indicated. Priming reactions performed by *Sso*PriS&L were assayed as above except in *Sso*PriS&L buffer (50 mM Glycine pH 9.0, 10 mM $MnSO_4$, 5 mM $Mg(CH_3COO)_2$, 10 μM $ZnCl_2$).

The calculated rate of primer synthesis (pmol/min) was fit to a standard Michaelis-Menten equation using non-linear least squares analysis using Kaleidagraph (Synergy, Reading, PA) according to the following equation,

$$v = \frac{V_{max} \times [P]}{K_D + [P]} \quad (2\text{-}1)$$

where K_D is either the apparent catalytic dissociation constant for *Sso*DnaG or the K_M for NTPs, P is the concentration of enzyme, and V_{max} is maximal rate of synthesis.

Electrophoretic Mobility Shift Assay (EMSA)

EMSAs were performed essentially as previously described with a few modifications.[49] Assays were performed in a 10 μL reaction volume containing buffer 100 mM HEPES (pH 7.0), 100 mM NaCl, 25 % PEG with 4 nM 66mer ssDNA (5'-CGATGAGAGCGAGTCGCATGGTA TCGTCTAGCCGGTCGGGGTGGGTGGGAAGCGTAGGGAGAGGTG) labeled at the 5'-end using a standard polynucleotide kinase reaction and ^{32}P-γ-ATP, and the indicated amount of *Sso*DnaG. Binding reactions were allowed to equilibrate for 3 h at 60 °C followed by directly loading onto a gradient 4–15 % polyacrylamide/TBE ReadyGel (BioRad, Hercules, CA). Gels were run for 4 h at 13 volts cm^{-1} followed by drying and phosphorimaged as above. Quantification of the fraction of band shift was performed using the ImageQuant software (v5.0) and fit to Equation 3-2 using Kaleidagraph,

$$y = \frac{F_0 + (\Delta F \times [P]^n)}{K_D{}^n + [P]^n} \quad (2\text{-}2)$$

where F_0 is the background shift, ΔF is the fraction of DNA shifted, P is the *Sso*DnaG concentration, K_D is the dissociation constant, and n is the Hill coefficient which defines cooperativity.

Fluorescence Anisotropy Assay

Fluorescence anisotropy measurements were performed essentially as described previously (*7*). Briefly, the 66mer ssDNA substrate used in the EMSA experiments was

fluorescently labeled with fluorescein at the 5' end and HPLC purified (IDT Corp) and was used to monitor the anisotropy change upon binding *Sso*DnaG at 25 °C. Measurements were performed in primase reaction buffer in the absence of NTPs with 25 nM DNA using a FluoroMax-3 spectrofluorimeter (HORIBA Jobin Yvon). Fluorescence was excited at 485 nm, and the emission was monitored at 518 nm during 1 sec integration times and represents an average of six consecutive readings. The absolute fluorescence intensity at 518 nm did not change significantly with addition of a high concentration of *Sso*DnaG ruling out the possibility that *Sso*DnaG binds specifically to the fluorophore. The anisotropy as a function of the concentration of *Sso*DnaG was fit to Equation 2-2, where F_0 is the initial anisotropy value and ΔF is the change in anisotropy.

2.2 RESULTS

2.2.1 Purification of SsoDnaG

Annotation within the genomic sequence of *Sulfolobus solfataricus* (*8*) has identified a gene product, *Sso*0079 (*Sso*DnaG), with homology to the bacterial DnaG (*9*) responsible for initiating DNA synthesis on both the leading and lagging strands by the *de novo* synthesis of a RNA primer complementary to a DNA template. Interestingly, this gene is adjacent to a proposed B-family DNA polymerase typically involved in DNA replication (*10*). DnaG is conserved across

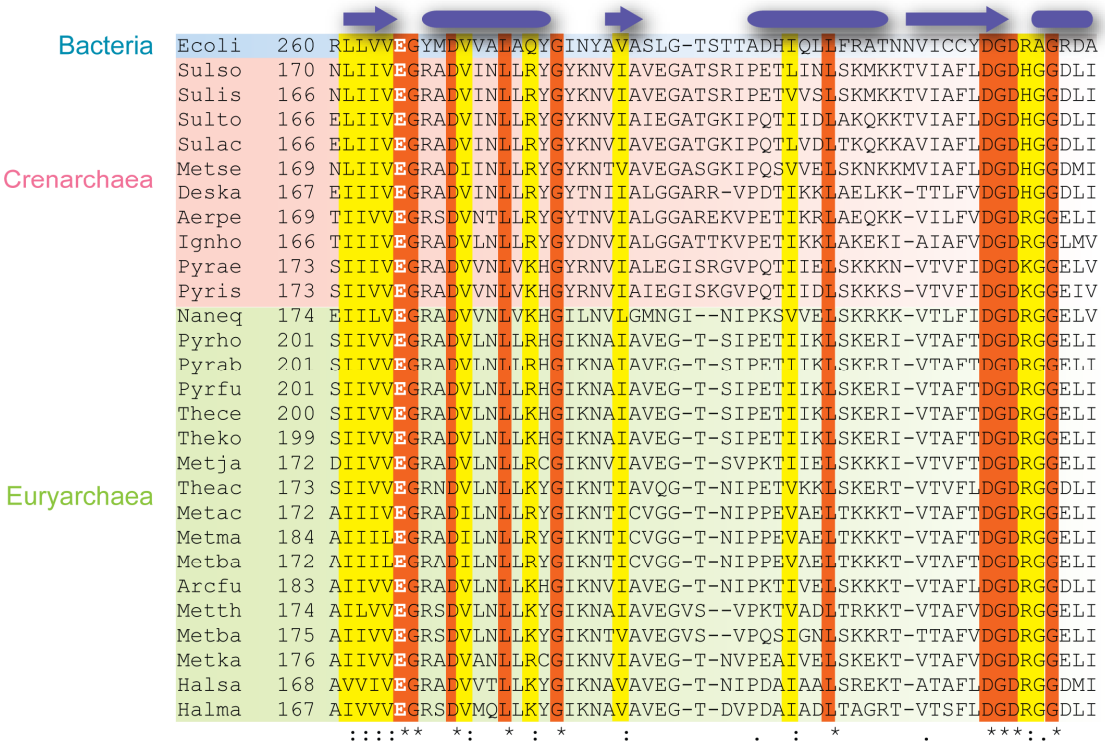

Figure 2-1. Multiple alignment of the TOPRIM core domain of DnaG primases

Grouped are representatives from bacteria (blue), and the two phyla from the arcahea domain; crenarchaea (red) and euryarchaea (green). The color-coding for absolutely conserved residues (orange), and highly conserved residues (yellow) from bacterial through archaeal species are shown. E175 in Sulso was highlighted (white) as the catalytic residue mutated in this manuscript. The secondary structure elements are derived from the crystal structure of *E. coli* DnaG (PDBID: 3B39). Species are listed as a five letter code for each species: Ecoli, *Escherichia coli*; Sulso, *Sulfolobus solfataricus*; Sulis, *Sulfolobus islandicus*; Sulto, *Sulfolobus tokodaii*; Sulac, *Sulfolobus acidicaldarius*; Metse, *Metallosphaera sedula*; Deska, *Desulfurococcus kamchatkensis*; Aerpe, *Aeropyrum pernix*; Ignho, *Ignicoccus hospitalis*; Pyrae, *Pyrobaculum aerophilum*; Pyris, *Pyrobaculum islandicum*; Naneq, *Nanoarchaeum equitans*; Pyrho, *Pyrococcus horikoshii*; Pyrab, *Pyrococcus abyssi*; Pyrfu, *Pyrococcus furiosus*; Thece, *Thermococcus celer*; Theko, *Thermococcus kodakarensis*; Metjan, *Methanocaldococcus jannaschii*; Theac, *Thermoplasma acidophilum*; Metac, *Methanosarcina acetivorans*; Metma, *Methanosarcina mazei*; Metba, *Methanosarcina barkeri*; Arcfu, *Archaeoglobus fulgidus*; Metth, *Methanothermobacter thermautotrophicum*; Metba, *Methanobrevi bactersmithii*; Metka *Methanopyrus kandleri*; Halsa, *Halobacterium salinarum*; Halmo, *Haloarcula marismortui* (http://www.ebi.ac.uk/Tools/clustalw2).

accepted archaeal phyla, crenarchaeota and euryarchaeota, and contains conserved residues within the TOPRIM domain responsible for catalysis in bacterial DnaG (Figure 2-1). *Sso*DnaG

lacks both the N-terminal Zn finger domain and the C-terminal DnaB helicase interacting domain (DnaB ID) found in bacterial DnaG. The Zn finger domain is thought to contribute to DNA binding, although a crystal structure of the core domain of *Ec*DnaG bound to DNA shows that a DNA binding and tracking groove exists adjacent to the active site (*11*). The DnaB ID is specific for interaction with the helicase during active DNA replication and stimulates primer production. Archaea lack a homologous DnaB replication helicase and instead utilize a eukaryotic-like MCM helicase (*12*).

Expression and purification of *Sso*DnaG was performed in *E. coli* as described in Materials and Methods and was shown to be > 95 % pure by coomassie stain (Figure 2-2A). After purification, gel filtration showed that *Sso*DnaG was present primarily in a soluble aggregate with a molecular weight greater than 750 kDa. The equilibrium between aggregation and individual species could be controlled somewhat by the addition of high concentrations of glycerol or polyethylene glycol (PEG) and incubation at 60 °C. We were not able to fully disrupt the aggregate even at high concentrations of PEG or glycerol. The degree of aggregation also varied slightly with different protein preparations. Expression of *Sso*DnaG with and without a His-tag or maltose binding protein tag did not alter the degree of aggregation or activity of the enzyme (data not shown). Activity and binding assays were performed using the unaggregated form of *Sso*DnaG separated by gel filtration, as the aggregate had minimal priming activity most likely due to a weak equilibrium between aggregate and monomer forms. We speculate that unknown auxiliary proteins and/or other factors are required to prevent aggregation and stabilize the active form of the enzyme or alternatively that the aggregate represents an inactive and improperly folded state of the enzyme.

2.2.2 Conditions for Maximal Priming Activity of SsoDnaG

Initially, we characterized the buffer conditions for optimal priming activity by *Sso*DnaG. Assays were performed using M13 as the DNA template and following the ^{32}P-α-NTP incorporation for different combinations of buffers, pH, metals, and temperatures. Initially, we examined the buffer and pH dependence on priming using 10 μM Zn^{2+}, 10 mM Mn^{2+}, and 5 mM Mg^{2+}. Background values in the absence of enzyme were subtracted from the rate of primer formation in the presence of *Sso*DnaG. Primer synthesis rates were calculated based on a standard curve of NTP concentrations after phosphorimaging. Assays performed with 50 mM Tris or HEPES buffer gave similar priming rates over a range of pH values, although 50 mM Tris at pH 8.5 resulted in maximal primer formation and was therefore used in all subsequent assays unless indicated otherwise (Figure 2-2B).

Activities of other primases from bacteria, archaea, and eukaryotes have shown that the presence of certain divalent metals can influence primer length(*13-15*).[5, 10, 27] The quantity of primer product was somewhat dependent on divalent metals but the distribution of primer product lengths was unchanged for *Sso*DnaG. Assays including either Mg^{2+} or Mn^{2+} stimulated primer formation over that without any metals or Zn^{2+} (Figure 2-2C) A combination of both Mg^{2+} and Mn^{2+} did not stimulate further primer synthesis over that of the single metal Mn^{2+} (data not shown). We also noted that the presence of a high concentration of Zn^{2+} (> 1 mM) created a high background of apparent primer products in the absence of enzyme on a denaturing

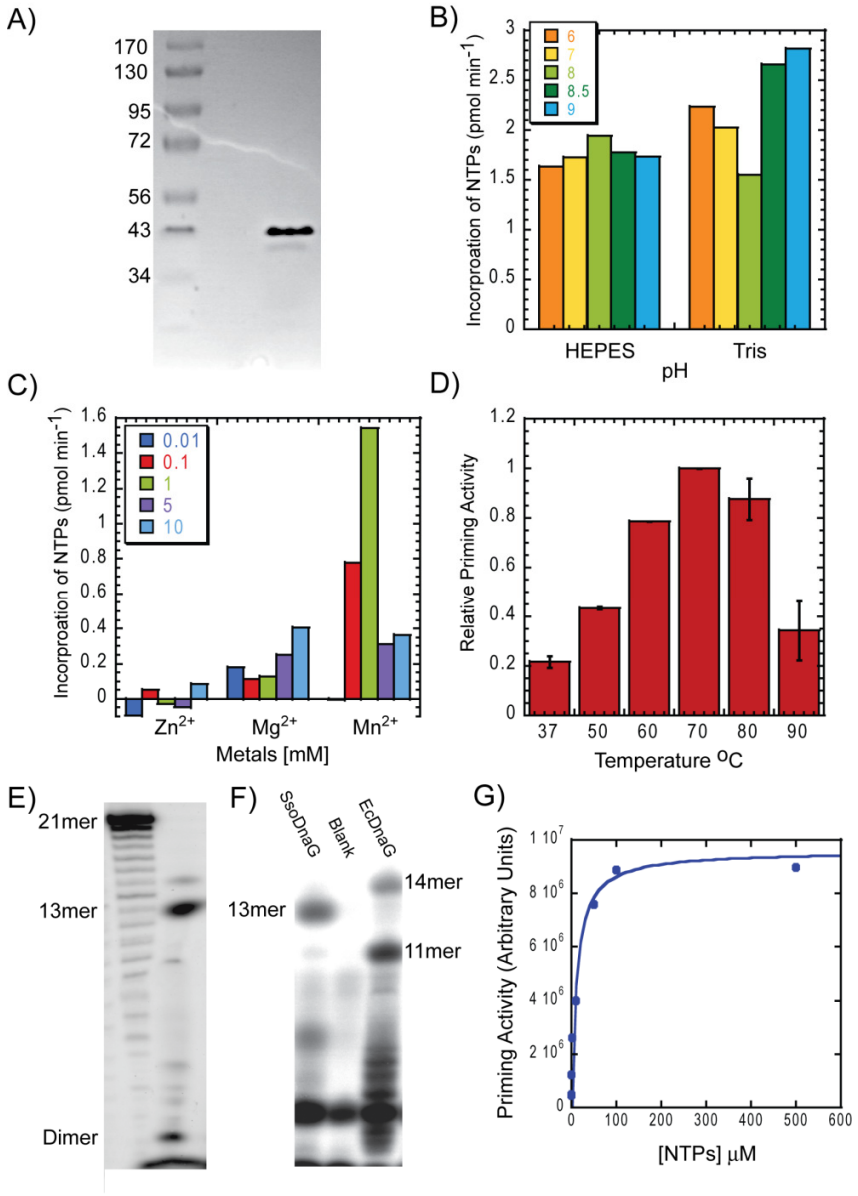

Figure 2-2. Priming Activity of *Sso*DnaG

A) Shown is purified *Sso*DnaG on a coomassie stained SDS-PAGE gel. The relative *Sso*DnaG priming activity at 1.5 µM of enzyme with different buffer conditions were tested. Assays were performed as described in Materials and Methods and measure the B) effect of pH in 50 mM HEPES or Tris Buffer, C) effect of divalent metals in 50 mM Tris buffer (pH 8.5) with 1 mM Mn^{2+} and D) Temperature dependence with 50 mM Tris (pH 8.5) with 1 mM Mn^{2+} on the relative priming activity normalized to 70 °C. E) Shows a representative image of the distribution of primers synthesized by *Sso*DnaG next to a 21 base ^{32}P 5' labeled oligo. F) Shows the distribution of primers synthesized by either Sulfolobus (*Sso*DnaG) or E.coli (*Ec*DnaG) DnaG. G) The binding affinity of NTPs or K_M was also determined by measuring the priming rate at different concentrations of [NTPs] at 60 °C, in 50 mM HEPES (pH = 8.0) and 1 mM Mn^{2+} and 5 mM Mg^{2+}, and 10 µM Zn^{2+} buffer.

acrylamide gel. Due to the lack of stimulation and the potential interfering background signal from Zn^{2+}, we only included 1 mM Mn^{2+} in all further enzymatic assays.

The temperature dependence on priming activity showed a maximum at 70 °C (Figure 2-2D). Again, the distribution of primer sizes did not significantly change over this temperature range. The measured ^{32}P intensity of total primer formation was used to quantify a relative priming activity at different temperatures. Optimal priming activity occurring at 70 °C rules out the possibility that contaminating *Ec*DnaG from the purification was providing the priming activity because that enzyme has limited priming ability above 50 °C. The growth temperature of *Sulfolobus* also lies within this temperature range, supporting a physiological role for this enzyme *in vivo*.

2.2.3 Primer Products Synthesized by SsoDnaG

The heterodimeric eukaryotic DNA primases from archaea are able to incorporate either NTP or dNTP substrates to synthesize both RNA, DNA, or hybrid primer products (*15, 16*). *Sso*DnaG, on the other hand, is only able to utilize NTPs and not dNTPs (data not shown) under all buffer conditions, metal concentrations, and temperature ranges tested. This is similar to what has been shown for the bacterial primase, where dNTPs have actually been shown to inhibit priming activity (*17, 18*). Instead, incorporation of NTPs led to the formation of primer products that ranged from 2-15 nucleotides (Figure 2-2E). The most common RNA primers synthesized were dimer, 4mer, or 13mer products. Once again, this is similar to a variety of reports of primer size for the bacterial primase that ranges from 2-16 nucleotides with maximum products of either 11 or 12 nucleotides, depending on the study (*14, 18*). Kinetic experiments performed as a function

of NTP concentration were used to determine the affinity of *Sso*DnaG (K_M) for NTPs (Figure 2-2F). *Sso*DnaG exhibits Michaelis-Menten type kinetics in the presence of both 5 mM Mg^{2+} and 10 mM Mn^{2+} and the K_m for NTPs was fit to 11 ± 1 μM. Therefore all further assays were performed with 100 μM NTPs.

2.2.4 Priming Activity of SsoDnaG

Primer synthesis as a function of primase concentration was monitored by individually quantifying both the larger 13mer and the smaller dimer primer products (Figure 2-3A). A fit to Equation 2-1 identified similar apparent catalytic dissociation constants (K_D) for both the 13mer and dimer of 1.2 ± 0.3 and 1.4 ± 0.3 μM, respectively, although the V_{max} values were very different (Figure 2-3).

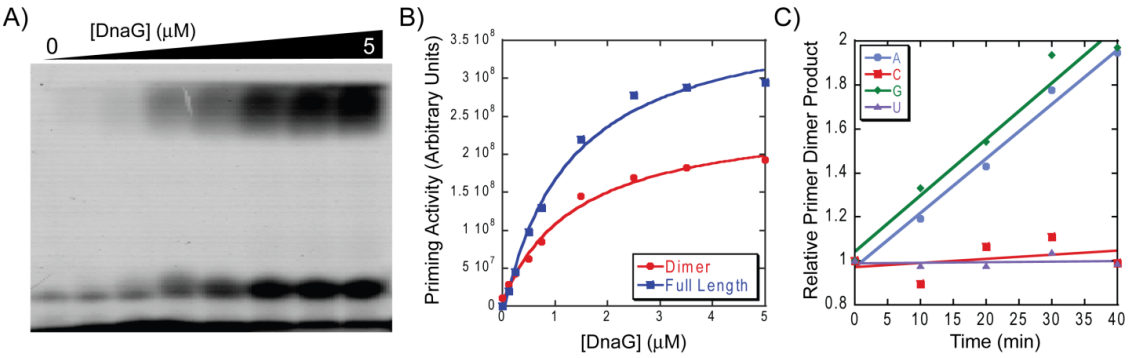

Figure 2-3. Priming Kinetics of *Sso*DnaG

A) *Sso*DnaG at different concentrations (0.01-5 μM) were incubated with 4 nM M13mp18, a mixture of all four NTPs (100 μM), and 0.25 μCi α-^{32}P-GTP for 50 minutes at 60 °C. B) The intensity of either dimer or 13mers primer products were quantified, plotted against increasing enzyme concentration, and the apparent catalytic K_D's were calculated for each. C) The effect of four separate ^{32}P labeled nucleotides on the relative steady-state rate of dimer primer products using 1 μM *Sso*DnaG.

In order to gain insight into the start site of primer synthesis, we performed priming reactions with each of the four ^{32}P nucleotides and analyzed the composition of the suspected dimer as a function of time. Figure 2-3C shows that there is a linear correlation in the rate of dimer product formation using either α-^{32}P-ATP or α-^{32}P-GTP. Neither α-^{32}P-CTP nor α-^{32}P-UTP was significantly incorporated into the dimer product over the time course of the reaction. The identical steady-state rates of incorporation for both ATP and GTP suggest that these nucleotides are equally incorporated into the first two positions of the primer product and must therefore besynthesized opposite either 5'-TC or 5'-CT in the M13 template.

2.2.5 DNA Binding of SsoDnaG

DNA binding of DnaG primases in bacteria has been shown to be weak and nonspecific at best, allowing for only a transient association with the DNA template (*19, 20*). We have detected binding of *Sso*DnaG to DNA by both electrophoretic mobility shift assay (EMSA) and fluorescence anisotropy. To visualize *Sso*DnaG binding to DNA, we monitored the shift of ^{32}P labelled DNA with increasing concentration of enzyme. Shown in Figure 2-4A and quantified in Figure 2-4B is the gel shift of *Sso*DnaG binding to a 5'-^{32}P labeled 66mer DNA. We do note that the shifted EMSA band has only migrated slightly into the gel, and we can not fully eliminate the possibility of the aggregation of DnaG with DNA. Therefore, we have performed multiple EMSA experiments with the error bars reported in Figure 2-4B. The EMSA data was initially fit to a standard single site binding curve, but the inclusion of a cooperativity parameter (n) significantly increased the quality of the fit. The calculated binding constant (K_D) is (2.3 ± 0.1) x 10^2 nM and the cooperativity is 2.6 ± 0.3. The positive cooperativity parameter can also identify

the stoichiometry of protein binding to DNA. This result is within the range of binding measured previously for bacterial DnaG (*11, 20*). *Sso*DnaG lacks the Zn finger motif of bacterial DnaG known to be important in DNA binding. Even without this known binding motif, *Sso*DnaG is able to bind DNA similarly to bacterial DnaG.

In order to further quantify and verify DNA binding using a second independent assay, we performed fluorescence anisotropy binding experiments to a 5' fluorescein labeled 66mer ssDNA substrate. Fluorescence anisotropy monitors the relative rotational diffusion rates of molecules so that an increase in molecular mass upon complex formation produces an increase in anisotropy. A fit of the change in anisotropy upon binding *Sso*DnaG identifies a moderate binding constant (K_d = 2.5 ± 0.1 x 10^2 nM) and Hill coefficient of 1.7 ± 0.1 (Figure 2-4C). The binding of *Sso*DnaG to smaller DNA substrates (≤ 25 bases) was not reproducibly detected in EMSA or anisotropy experiments. The binding affinity measured between these two techniques is almost identical and has also identified a possible positive cooperativity of binding to DNA. Positive cooperativity suggests that two molecules of *Sso*DnaG bind sequentially, and the second molecule's binding affinity is increased after binding of the first molecule. The positive cooperativity identified in the fit of our data could either be enzymatic dimer formation necessary for optimal catalysis or lateral binding of individually competent molecules along the length of DNA.

2.2.6 Active Site Mutant of *Sso*DnaG Reduces Priming Activity

The homology within the primary sequence of *Sso*DnaG to the core TOPRIM domain of *Ec*DnaG allowed for the initial identification of this gene and alignment of other archaeal DnaG

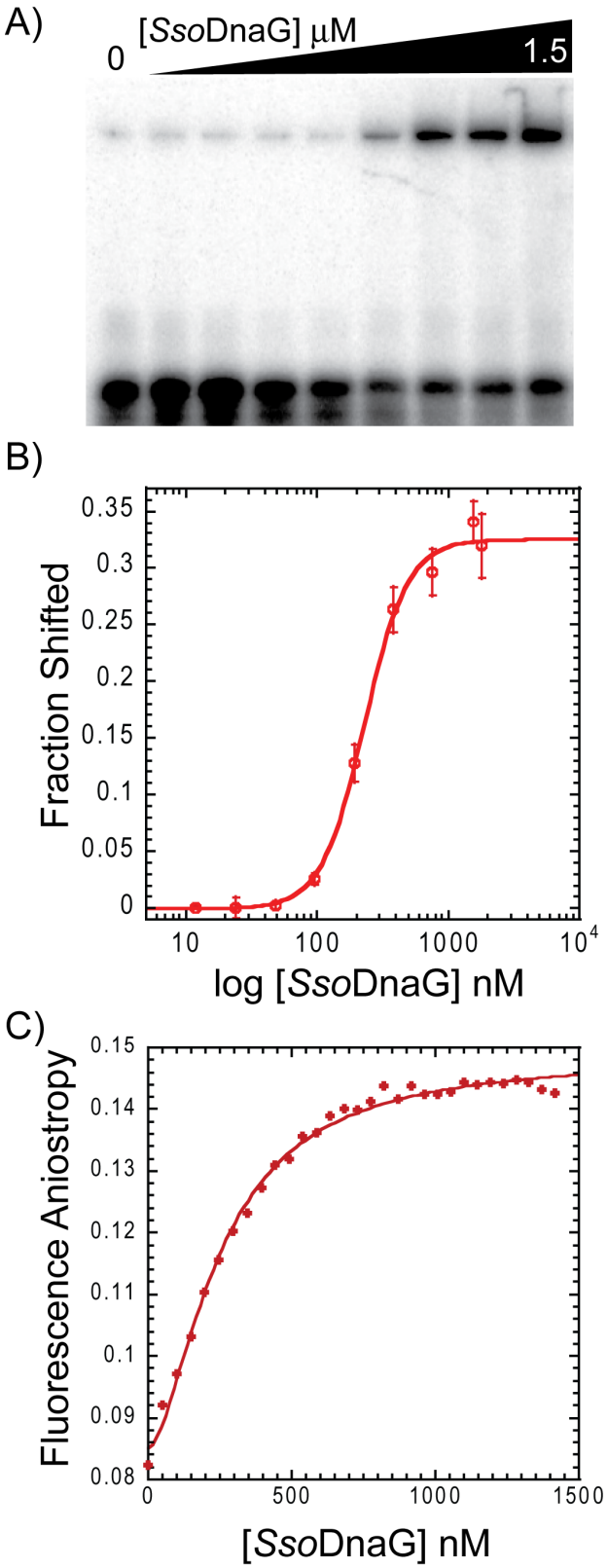

Figure 2-4. DNA Binding of *Sso*DnaG

A) Electrophoretic mobility shift assay (EMSA) of the interaction of *Sso*DnaG with a 66mer ssDNA substrate labeled at the 5' end with ^{32}P. The concentration of *Sso*DnaG was increased as shown above the gel. B) A plot of the fraction of the total DNA shifted as a function of the log of the concentration of *Sso*DnaG. Data was fit to Equation 2-2 and yielded a K_d of 228 ± 11 nM and a Hill coefficient equal to 2.6 ± 0.3. C) Fluorescence anisotropy was used to quantify binding *Sso*DnaG to a 66-nt single strand DNA substrate labeled at the 5' end with fluorescein. Data points were fit to Equation 2-2 and yielded a K_d equal to 253 ± 9 nM and a Hill coefficient equal to 1.7 ± 0.1.

proteins (Figure 2-1) (*9*). Within the core domain is a highly conserved and essential catalytic glutamate (E) and aspartate dyad (DGD). Mutation of the conserved glutamate to glutamine has been shown to abolish priming activity but retain DNA binding activity in bacteriophage T4 and *E. coli* (*21, 22*). To test whether *Sso*DnaG utilizes the conserved glutamate in the TOPRIM domain as a catalytic residue, we made an E175Q mutation (*Sso*DnaG Cat$^-$) and tested the DNA priming ability of this mutant. Under optimal experimental conditions described above for WT *Sso*DnaG, the priming activity of *Sso*DnaG Cat$^-$ at identical concentrations of enzyme (1 μM) was shown to be drastically reduced but not eliminated (Figure 2-5). A low level of DNA priming activity was detected for this mutant, whereas the homologous mutation in bacterial DnaG abolishes activity (*21, 22*). This may be due to the higher temperatures used for *Sso*DnaG as well as the two remaining and conserved aspartates at the active site contributing to some level of reduced catalysis. Nevertheless, the priming ability of *Sso*DnaG Cat$^-$ was severely inhibited and supports the role for E175 as a catalytically important residue.

Quantification of the priming activity of WT *Sso*DnaG versus *Sso*DnaG Cat$^-$ shows a 13-fold reduction in the rate of primer formation (Figure 2-5B). WT *Sso*DnaG synthesized primers at a rate of 2.3 pmol min^{-1} versus 0.1 pmol min^{-1} for *Sso*DnaG Cat$^-$. The rate of primer formation for *Sso*DnaG Cat$^-$ is only slightly higher than background levels and is only significant due to the trace production of 13mer primers at times barely visible on the image. Quantification of the rate of primer formation as a function of the primase concentration shows no significant increase in product within increasing concentration of E175Q (Figure 2-5C). This lack of activity for the catalytic mutant also provides evidence that *Ec*DnaG is not contaminating our purification of *Sso*DnaG. On the other hand, WT *Sso*DnaG priming ability is stimulated with increasing enzyme concentration. A fit of this data to Equation 2-1 reveals a V_{max} of 6.8 ± 0.6 and 0.26 ± 0.02 pmols

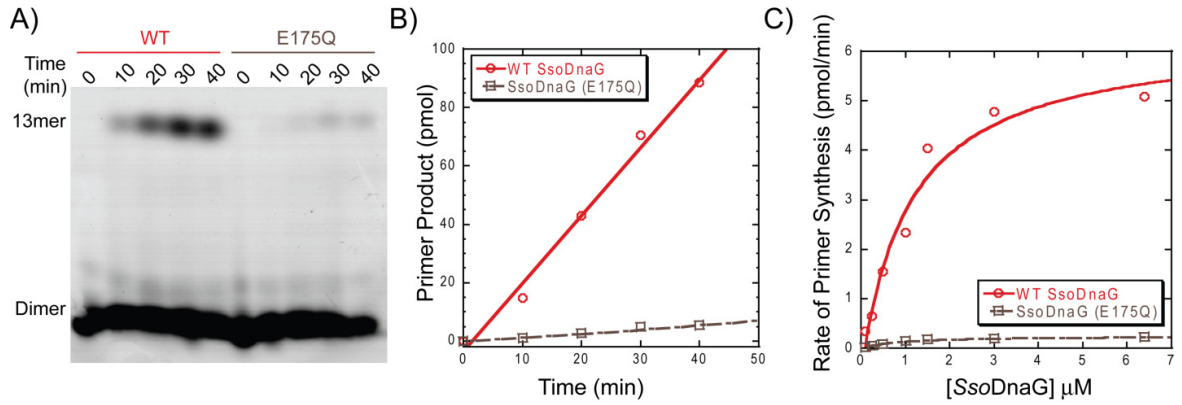

Figure 2-5. Priming Activity of Active site Mutant of *Sso*DnaG

A) Reactions containing 1 μM enzyme (either wt *Sso*DnaG or E175Q *Sso*DnaG), 0.25 μCi ^{32}P-GTP, 100 μM NTPs, and 4 nM M13 DNA template, were incubated at 60 °C and processed as described in Materials and Methods. B) Quantification of the band intensities were used to plot the steady state rate of primer formation (pmol/min) calculated from a standard curve of pmol NTPs. C) Concentration dependent study comparing primer formation of either *Sso*DnaG WT or E175Q. Equation 2-1 was used to fit the data for WT or E175Q *Sso*DnaG with a catalytic K_D equal to 1.1 ± 0.4 or 0.8 ± 0.3 μM, respectively. The V_{max} for WT or E175Q is 6.8 ± 0.6 and 0.26 ± 0.02 pmols/min, respectively.

min^{-1} for WT and E175Q *Sso*DnaG, respectively (Table 2-1). The catalytic K_D of total primer formation for WT *Sso*DnaG equals 1.1 ± 0.4 μM. This K_D value is equivalent to that calculated for either dimer or 13mer primer formation in Figure 2-3B.

2.2.7 Comparison of Priming Activities of SsoDnaG to SsoPriS&L

In order to better understand the priming abilities of each primase family contained within Archaea, we have directly compared the priming kinetics of the bacterial-like *Sso*DnaG and the eukaryotic heterodimer *Sso*PriS&L. Previously, *Sso*PriS&L was shown to synthesize both RNA and DNA primers with a wide distribution in size (*16*). Major RNA primer products synthesized

by *Sso*PriS&L were primarily 7 and 14 nucleotides long. Because *Sso*DnaG is only capable of synthesizing RNA primers, we have directly compared only the RNA primer synthesis ability for each primase using M13 as the substrate. Optimal buffer conditions reported for *Sso*PriS&L(*16*) were used for the priming assays.

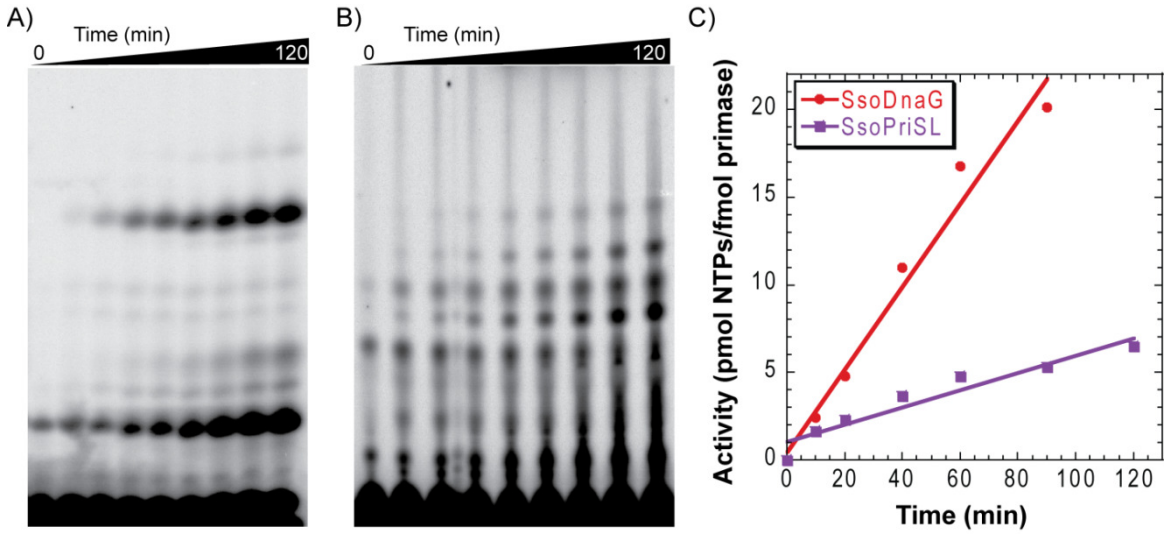

Figure 2-6. Comparison of Priming Activity for *Sso*DnaG and *Sso*PriS&L

Timecourse for primer formation of both A) *Sso*DnaG and B) *Sso*PriS&L. 2.5 μM of enzyme was incubated at 60 °C with 4 nM M13 and a mixture of 100 μM NTPs and 0.25 μCi α-^{32}P-GTP. C) Rates of priming activity of both *Sso*DnaG (red circles) and *Sso*PriS&L (purple squares) per mole of enzyme was calculated to be 0.2 and 0.05 pmol NTPs fmol primase^{-1} min^{-1}. The rate of primer formation by *Sso*DnaG is about 4 times greater than that of *Sso*PriS&L.

Shown in Figure 2-6 A and B are representative kinetic time courses of primer formation for identical concentrations of either *Sso*DnaG or *Sso*PriS&L. An initial comparison shows a difference in the distribution of primer products between the two primases. Whereas *Sso*DnaG again primarily produces dimer, 4mer, and 13mer products, *Sso*PriS&L produces a range of primer products similar to that reported previously (*16*). Quantification of the total primer

products (pmol) from Figure 2-6A and B was calculated from a standard curve of ^{32}P intensity versus concentration of NTPs. The rate of primer product formation per fmol of either primase enzyme is represented in Figure 2-6C. As a comparison, the rate of *Sso*PriS&L at 100 μM NTPs was measured previously to be ~100 fmol NTPs min^{-1} μg $primase^{-1}$ or 0.007 pmol min^{-1} fmol $primase^{-1}$ (*16*). The average rates of total primer production measured here was calculated from three independent measurements and found to be 0.17 ± 0.06 and 0.039 ± 0.008 pmol NTPs incorporated fmol $primase^{-1}$ min^{-1} for *Sso*DnaG and *Sso*PriS&L, respectively. Our priming rates for *Sso*PriS&L under identical conditions are slightly higher than those previously published. This is most likely due to different ^{32}P labeled NTPs used in the experiments (^{32}P-α-GTP versus ^{32}P-α-ATP) which depending on the primer sequence synthesized for each enzyme may affect our calculation of primer products. Using this upper priming limit for *Sso*PriS&L measured in this manuscript, there is a 4-fold greater *in vitro* NTP incorporation rate for *Sso*DnaG over that of *Sso*PriS&L under optimal conditions for each enzyme.

2.3 DISCUSSIONS

Until now, it was widely accepted that the DNA priming activity necessary for DNA replication initiation and elongation in Archaea was performed by the heterodimeric PriS&L protein complex which shares homology and is in the same family as its eukaryotic counterpart, polα-primosome. Curiously, this archaeal PriS&L primase complex has the ability to synthesize both RNA and DNA products over a large size range as well as possesses an alternative terminal transferase activity. The prevailing view was that this primase complex in archaea can initially

synthesize RNA products that are then extended into DNA to initiate DNA replication in a process similar to that of the polα complex in eukaryotes, which contains a DNA polymerase subunit. This RNA/DNA hybrid can then be further extended by the DNA replication polymerase holoenzyme complex. The dual activities of PriS&L in archaea allow for a simpler yet homologous alternative to the four subunit polα complex in eukaryotes.

Our discovery of *Sso*DnaG RNA priming activity that is similar to known enzymatic properties of bacterial primases suggests that discreet roles may exist for each primase family found in archaea. Most of the DNA replication machinery from Archaea is highly homologous to their respective eukaryotic counterparts. So, the presence of both bacterial and eukaryotic-like DNA primases in Archaea has important implications for our understanding of the evolutionary development of DNA replication in all domains of life. Separate priming activities from two primases in Archaea also question the role of each in nucleic acid processing *in vivo*.

2.3.1 Comparison of the DnaG Primase from Archaea and Bacterial Domains

The initial identification of a homologous TOPRIM domain in proposed archaeal DnaG primases (*9*) led us to investigate the enzymatic ability of this enzyme in *Sulfolobus solfataricus*. We had difficulty purifying the active form of the enzyme as it tends to aggregate in solution. It is unclear if the aggregation is in a slow or nonexistent equilibrium, but the isolated monomer species by gel filtration can be concentrated without loss of activity. The aggregate fraction can be controlled somewhat by the addition of PEG or glycerol or expression at lower temperatures. This suggests that the high molecular weight aggregate may be a misfolded species that

maintains solubility, and a small fraction can refold into its native state under appropriate conditions.

Bacteriophage and bacterial primases generally recognize a canonical three base sequence on the template strand and begin synthesizing primers complementary to the second base in the sequence (*23*). Based on the identical incorporation rate for both ATP and GTP into the dimeric primer product, we propose that the recognition site for *Sso*DnaG is either 5'-d(CTX) or 5'-d(TCX), where X is currently unknown. In *E. coli*, RNA primers are most often initiated *in vivo* by pppApG(*24*), which is complementary to the second two bases in the 5'-d(CTG) DnaG recognition site determined *in vitro* (*25*). Our results determining the composition of the RNA dimer product (G & A) made from *Sso*DnaG are identical to the first two nucleotides in bacterial RNA primer synthesis and consistent with the bacterial primase recognition site.

The length of primer products synthesized by *Sso*DnaG is also very reminiscent for what is found for bacterial DnaG, although the absolute distribution is different. The primary products for *Sso*DnaG are dimer, 4mer and 13mer. The primer products for *Ec*DnaG range from 10-14 nucleotides with the majority being 12 ± 1 mer products(*25*). The distribution of primer sizes created by *Sso*DnaG has a much smaller range. The overwhelming primer product is a discrete 13mer and only small fractions of primers of other lengths are detected. This implies that the processivity of *Sso*DnaG and absolute primer size is more tightly controlled than in bacteria. The similarity in primer length between bacterial and archaeal primases suggests a conserved mode of synthesis in which a single extra base may be thermodynamically important. The problem of primer handoff is still an important issue in these thermophiles because a 13mer RNA primer will not be stably associated with the DNA template at 75 °C. It is likely that another replication protein interacts with *Sso*DnaG or the primer after synthesis, stabilizing the primer on DNA

while awaiting recruitment of the DNA polymerase. This proposed process is similar to what has been suggested for *E. coli* (*26*).

*Sso*DnaG lacks the canonical zinc finger binding domain found in *Ec*DnaG, but was still able to bind DNA with modest affinity. DNA primases are notorious for transient and nonspecific binding to DNA templates (*23*). The inability of *Sso*DnaG to fully shift the DNA in our EMSA experiments indicates a similar binding profile in archaea. Interestingly, we have detected a slight positive cooperativity for *Sso*DnaG binding to DNA consistent with sequential and cooperative dimer formation. Dimer formation has also been observed for *Ec*DnaG, and primases from other homologous phages are thought to form hexamers (*20,27,28*). Cooperativity can either be due to dimer formation required for catalysis or sequential binding events along the length of the DNA template. The DNA template used in our binding experiments is 66 bases long. Based on the site size proposed for the *Ec*DnaG core domain (7-8 bases) (*11*) and the homology between enzymes, this could potentially allow binding of 7-8 molecules of *Sso*DnaG along the length of the template. We were only able to detect a positive cooperativity with a Hill coefficient equal to 2 in two independent assays (EMSA and anisotropy) suggesting that dimer formation is more likely than polymer formation for *Sso*DnaG binding to DNA.

Finally and maybe most importantly for comparison to bacterial primases, a mutation in the conserved glutamate contained in the TOPRIM active site significantly impaired the priming activity of *Sso*DnaG. This glutamate is conserved across all currently sequenced species of archaea including *Nanoarchaeum equitans* which has the smallest known archaeal genome (*29*). This level of conservation throughout the archaeal domain which also includes two other proposed catalytic aspartates roughly 45 residues upstream is indicative of an essential enzyme

for archaeal life. Interestingly, none of these archaeal family members include a zinc binding motif or DnaB interacting domain found in *E. coli*. These domains must have evolved separately in bacteria and are not required for optimal function in archaea.

2.3.2 Potential Roles of Each Primase in Archaea

The *Sso*DnaG gene and protein product has been detected *in vivo* at significant levels using both global gene expression (*30, 31*) and proteomics (*32*). Therefore, individual roles for each primase probably exist in *Sulfolobus*. Analysis of the data from a global gene response to UV damage (*30*) shows that *Sso*DnaG is immediately down regulated followed by up-regulation after 2 hours of recovery. If *Sso*DnaG is involved in DNA replication initiation of Okazaki fragments, then it makes sense that the down regulation of this gene would halt DNA replication to allow an SOS response to repair DNA by reversing UV damage. After repair, *Sso*DnaG is transcribed to continue DNA replication. In contrast, *Sso*PriS and *Sso*PriL do not follow the same pattern or even act synergistically in response to UV damage. *Sso*PriS is initially increased and then returns to basal levels while *Sso*PriL is essentially unchanged and then down regulated. Genetic knockout of the homologous PriS&L in halobacterium is lethal (*33*), although to our knowledge, no knockouts of bacterial primases have been performed in archaea.

Protein interactions of these primases suggest a different story. The only proteins that *Sso*DnaG has been found to interact with are those within the exosome (*2*). Interestingly, the function of the exosome is to degrade RNA products for transcriptional turnovers. This is explicitly opposite to the RNA synthesis function we detected for *Sso*DnaG. Addition of poly (A)-tails to RNA molecules is common to all organisms examined so far. Ironically, RNA

polyadenylation is part of RNA degradation process. In certain bacteria cells, the addition of poly (A)-rich tails is carried out by polynucleotide phosphorylase, PNPase, a homologue of archaea exosome (*34-36*). Polyadenylated RNA molecules are then degraded rapidly by hydrophilic exoribonucleases of RNase I/R family or by phosphorolytic exoribonuclease PNPase. In archaea, polyadenylation of RNA was found to be mediated by exosome complex (*37*). The biological role of *Sso*DnaG in exosome has yet to be explored. However, it's attractive to speculate that *Sso*DnaG might play a role in the polymerization of poly (A)-rich tails before the RNA molecules can be degraded by archaea exosomes. Our finding of incorporation of A into the dimeric primer (Figure 2-3C) is supportive to this hypothesis. On the other hand, *Sso*PriS&L has been found to interact specifically with the clamp-loader (*Sso*RFC) responsible for catalyzing the formation of the DNA polymerase holoenzyme necessary for DNA replication (*38*). In fact, this interaction is proposed to be necessary for efficient primer hand-off to the DNA polymerase in *Sulfolobus*.

There are also significant enzymatic differences between the bacterial DnaG and the archaeal-eukaryotic primases. Our results show that *Sso*DnaG is only able to incorporate NTPs similar to those of bacterial DnaG. Early reports on the priming ability of *Ec*DnaG suggested that it could incorporate dNTPs (*17*), but subsequent studies have shown that dNTPs actually inhibit priming(*39*). We were unable to detect any primer products using dNTPs. On the other hand, PriS&L in a variety of archaeal species has been shown to incorporate either NTPs or dNTPs into primers. PriS&L has also been shown to possess a terminal transferase activity to extend ssDNA substrates (*15,16,40*). The multiple enzymatic activities of the PriS&L family of primases indicate potential involvement in alternative processes within DNA replication or

repair, whereas the singular priming activity detected for *Sso*DnaG is focused on replication initiation or Okazaki fragment maturation.

Further studies will be required to untangle the *in vivo* enzymatic roles of each primase family within archaea. Evolutionarily speaking, phylogenetic analysis has revealed that the archaeal eukaryotic primase family may have evolved from the same common ancestor of the family B DNA polymerase (*41*). The combination of RNA and DNA synthesis abilities would be extremely important in the transition from an RNA to a DNA genetic world. The hybrid activity of the PriSL primase may have been retained in the form of replication initiation from the origin on the leading strand. We speculate that the specialized role of DNA priming by sole incorporation of NTPs by archaeal DnaG would have evolved to initiate DNA synthesis on the lagging strand. To our knowledge, no species of life within the bacterial or eukaryotic domains includes both classes of DNA primases. Just about all of the other conserved DNA replication proteins in archaea are eukaryotic-like and provide a convenient model for understanding DNA replication in higher organisms. The presence of dual primase families in archaea is conflicting, but this anomaly may provide a model for the examination of the evolutionary split and development of DNA replication in bacteria and eukaryotes.

REFERENCES

1. Koonin EV, Mushegian AR, Galperin MY, Walker DR. (1997) Comparison of archaeal and bacterial genomes: computer analysis of protein sequences predicts novel functions and suggests a chimeric origin for the archaea. *Mol. Microbiol. 25*, 619-637.

2. Walter P, Klein F, Lorentzen E, Ilchmann A, Klug G, Evguenieva-Hackenberg E. (2006) Characterization of native and reconstituted exosome complexes from the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Mol. Microbiol.* 62(4):1076-89

3. Evguenieva-Hackenberg E, Lassek C and Klug G. (2011) Subcellular localization of RNA degrading proteins and protein complexes in prokaryotes. *RNA Biology 81*, 49-54.

4. Gill SC, von Hippel PH. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* 182, 319-326.

5. Shevchenko A, Tomas H, Havlis J, Olsen JV and Mann. (2006) In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* 1, 2856-2860.

6. Corn JE, Pease PJ, Hura GL, Berger JM. (2005) Crosstalk between primase subunits can act to regulate primer synthesis in trans. *Mol Cell.* 20, 391-401.

7. Mikheikin AL, Lin HK, Mehta P, Jen-Jacobson L and Trakselis MA. (2009) A trimeric DNA polymerase complex increases the native replication processivity. *Nucleic Acids Res.* 37, 7194-7205.

8. She Q, Singh RK, Confalonieri F, Zivanovic Y, Allard G, Awayez MJ, Chan-Weiher CC, Clausen IG, Curtis BA, De MA, Erauso G, Fletcher C, Gordon PM, Heikamp-de J, Jeffries AC, Kozera CJ, Medina N, Peng X, Thi-Ngoc HP, Redder P, Schenk ME, Theriault C, Tolstrup N, Charlebois RL, Doolittle WF, Duguet M, Gaasterland T, Garrett RA, Ragan MA, Sensen CW and Van der OJ. (2001) The complete genome of the crenarchaeon Sulfolobus solfataricus P2. *Proc. Natl. Acad. Sci. U S A.* 98, 7835-7840.

9. Aravind L, Leipe DD & Koonin EV. (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, 4205-4213.

10. Edgell DR, Klenk HP and Doolittle WF. (1997) Gene duplications in evolution of archaeal family B DNA polymerases, *J. Bacteriol.* 179, 2632-2640.

11. Corn JE, Pelton JG and Berger JM. (2008) Identification of a DNA primase template tracking site redefines the geometry of primer synthesis. *Nat. Struct. Mol. Biol.* 15, 163-169.

12. McGeoch AT, Trakselis MA, Laskey RA and Bell SD. (2005) Organization of the archaeal MCM complex on DNA and implications for the helicase mechanism. *Nat. Struct. Mol. Biol.* 12, 756-762.

13. Schneider A, Smith RW, Kautz AR, Weisshart K, Grosse F and Nasheuer HP. (1998) Primase activity of human DNA polymerase alpha-primase. Divalent cations stabilize the enzyme activity of the p48 subunit, *J. Biol. Chem.* 273, 21608-21615.

14. Zechner EL, Wu CA and Marians KJ. (1992) Coordinated leading- and lagging-strand synthesis at the *Escherichia coli* DNA replication fork. III. A polymerase-primase interaction governs primer size, *J. Biol. Chem.* 267, 4054-4063.

15. Le Breton M, Henneke G, Norais C, Flament D, Myllykallio H, Querellou J and Raffin JP. (2007) The heterodimeric primase from the euryarchaeon *Pyrococcus abyssi*: a multifunctional enzyme for initiation and repair. *J. Mol. Biol.* 374, 1172-1185.

16. Lao-Sirieix SH, 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, 1251-1263.

17. Rowen L, Kornberg A. (1978) A ribo-deoxyribonucleotide primer synthesized by primase. *J. Biol. Chem.* 253, 770-774.

18. Johnson SK, Bhattacharyya S and Griep MA. (2000) DnaB helicase stimulates primer synthesis activity on short oligonucleotide templates. *Biochemistry.* 39, 736-744.

19. Swart JR, Griep MA. (1995) Primer synthesis kinetics by *Escherichia coli* primase on single-stranded DNA templates. *Biochemistry.* 34, 16097-16106.

20. Khopde S, Biswas EE and Biswas SB. (2002) Affinity and sequence specificity of DNA binding and site selection for primer synthesis by *Escherichia coli* primase. *Biochemistry* 41, 14820-14830.

21. Strack B, Less M, Calendar R and Lanka E. (1992) A common sequence motif, -E-G-Y-A-T-A-, identified within the primase domains of plasmid-encoded I- and P-type DNA primases and the alpha protein of the *Escherichia coli* satellite phage P4. *J. Biol. Chem.* 267, 13062-13072.

22. Trakselis MA, Roccasecca RM, Yang J, Valentine AM and Benkovic SJ. (2003) Dissociative properties of the proteins within the bacteriophage T4 replisome. *J. Biol. Chem.* 278, 49839-49849.

23. Frick DN and Richardson CC. (2001) DNA primases. *Annu. Rev. Biochem.* 70, 39-80.

24. Kitani T, Yoda K, Ogawa T and Okazaki T. (1985) Evidence that discontinuous DNA replication in *Escherichia coli* is primed by approximately 10 to 12 residues of RNA starting with a purine. *J. Mol. Biol.* 184, 45-52.

25. Bhattacharyya S, Griep MA. (2000) DnaB helicase affects the initiation specificity of *Escherichia coli* primase on single-stranded DNA templates. *Biochemistry.* 39, 745-752.

26. Yuzhakov A, Kelman Z, O'Donnell M. (1999) Trading places on DNA--a three-point switch underlies primer handoff from primase to the replicative DNA polymerase. *Cell.* 96, 153-163.

27. Frick DN and Richardson CC. (1999) Interaction of bacteriophage T7 gene 4 primase with its template recognition site. *J. Biol. Chem.* 274, 35889-35898.

28. Valentine AM, Ishmael FT, Shier VK and Benkovic SJ. (2001) A zinc ribbon protein in DNA replication: primer synthesis and macromolecular interactions by the bacteriophage T4 primase. *Biochemistry.* 40, 15074-15085.

29. Huber H, Hohn MJ, Rachel R, Fuchs T, Wimmer VC, Stetter KO. (2002) A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. *Nature.* 417, 63-67.

30. Götz D, Paytubi S, Munro S, Lundgren M, Bernander R, White MF. (2007) Responses of hyperthermophilic crenarchaea to UV irradiation. *Genome Biol.* 8, R220.

31. Andersson AF, Lundgren M, Eriksson S, Rosenlund M, Bernander R, Nilsson P. (2006) Global analysis of mRNA stability in the archaeon *Sulfolobus*. *Genome Biol.* 7, R99.

32. Assiddiq BF, Snijders AP, Chong PK, Wright PC, Dickman MJ. (2008) Identification and characterization of *sulfolobus solfataricus* P2 proteome using multidimensional liquid phase protein separations. *J. Proteome. Res.* 7, 2253-2261.

33. Berquist BR, DasSarma P and DasSarma S. (2007) Essential and non-essential DNA replication genes in the model halophilic Archaeon, Halobacterium sp NRC-1. *BMC. Genet. 8, 31*.

34. Mohanty BK, Kushner SR. (2000) Polynucleotide phosphorylase functions both as a 30 to 50 exonuclease and a poly(A) polymerase in *Escherichia coli*. *Proc. Natl. Acad. Sci. U S A.* 97, 11966-11971.

35. Bollenbach TJ, Schuster G, Stern DB. (2004) Cooperation of endo- and exoribonucleases in chloroplast mRNA turnover. *Prog. Nucleic. Acid. Res. Mol. Biol.* 78, 305-337.

36. Slomovic S, Portnoy V, Yehudai-Resheff S, Bronshtein E, Schuster G. (2008) Polynucleotide phosphorylase and the archaeal exosome as poly(A)-polymerases, *Biochimica et Biophysica Acta.* 1779, 247-255.

37. Portnoy V, Evguenieva-Hackenberg E, Klein F, Walter P, Lorentzen E, Klug G, Schuster G. (2005) RNA polyadenylation in Archaea: not observed in Haloferax while the exosome polyadenylates RNA in *Sulfolobus*. *EMBO Rep.* 6, 1188-1193.

38. Wu K, Lai X, Guo X, Hu J, Xiang X, Huang L. (2007) Interplay between primase and replication factor C in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Mol. Microbiol.* 63, 826-837.

39. Koepsell S, Bastola D, Hinrichs SH and Griep MA. (2004) Thermally denaturing high-performance liquid chromatography analysis of primase activity. *Anal. Biochem.* 332, 330-336.

40. Beck K, Lipps G. (2007) Properties of an unusual DNA primase from an archaeal plasmid, *Nucleic Acids Res.* 35, 5635-5645.

41. 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, 3875-3896.

3.0 STRAND ANNEALING AND TERMINAL TRANSFERASE ACTIVITIES OF *SSO*DPO1[1]

The six families of DNA polymerases play individual roles in DNA replication and repair in all domains of life in an effort to maintain the genomic integrity of the cell. In most respects, B-family DNA replication polymerases require a template to direct synthesis and incorporation of nucleotides into a growing strand. The stability of a primer on a template substrate allows for the faithful copying of the complementary strand. When DNA damage is encountered on the template strand, specialized Y-family lesion bypass DNA polymerases are recruited to transverse the damage (*1, 2*). Even these lesion bypass polymerases absolutely require the presence of a template strand to direct synthesis in a potentially mutagenic manner (*1*). Some DNA polymerases, most notably Taq, can incorporate a single adenine base on the end of a double strand DNA (dsDNA) product in a single template independent terminal transferase step that creates an A-tail (*3*).

Within the X-family of DNA polymerases, there exists terminal deoxynucleotidyl transferase (TdT) that can extend single strand DNA (ssDNA) in a template independent manner (*4*). Essentially, TdT can incorporate dNTPs onto the 3' end of DNA indiscriminately. TdT was first identified in mammals (*5*) and is thought to play a role in the diversification of the

[1] This chapter is primarily taken from *Biochemistry* 2011, 50, 5379–5390, by Zuo Z. *et al.*

vertebrate immune system through V(D)J recombination (*6*) or in specific aspects of DNA double strand break repair (*7*). V(D)J recombination is the rearrangement of antigen receptors to create unique antibody specificities to thwart biological attacks on the immune system (*8, 9*). Rearrangement requires the creation of double strand breaks within immunoglobin (Ig) genes, TdT activity to extend ssDNA6, followed by reannealing of random gene regions to propagate immune diversity (*10, 11*). Eukaryotic X-family DNA polymerases (including TdT, pol μ, and pol λ) have also been shown to participate in nonhomologous end joining (NHEJ) by extending ssDNA at double strand breaks as a template for annealing (*12*).

Archaea are single cell prokaryotic organisms and not thought to need a V(D)J recombination system for diversification of the immune system, but growth at high temperatures would likely lead to accelerated DNA damage including double strand breaks (DSB) (*13*). Interestingly, it has been noted that the mutational frequencies in the genome are equal to or lower than those found in mesophilc organisms suggesting a highly active DNA repair system (*14*). Homologous recombination (HR), not NHEJ, is thought to be the preferred mode for the repair of double strand breaks in archaea (*15, 16*). The need for an X-family polymerase with TdT activity may not be necessary in archaea in light of the discovery that the eukaryotic-like primase (PriSL) has terminal transferase activity (*17*). Interestingly, a crystal structure of the *Pyrococcus* PriS subunit has revealed unexpected structural homology with X-family DNA polymerases prompting the authors to predict an unrecognized role for archaeal DNA primases in DNA repair (*17-19*). Although archaea lack a putative X-family DNA polymerase, they have both a B-family DNA replication polymerase, and a repair Y-family lesion bypass polymerase (*20, 21*). The genetic role, if any, for terminal transferase activity in archaea remains uncharacterized.

DNA replication and repair at high temperatures creates significant thermodynamic problems for maintaining annealed double strand DNA templates. In fact, the small RNA primer required for initiation of DNA replication on both the leading and lagging strand can not be stably associated with the template at those high temperatures through thermodynamic interactions alone and would require help from other proteins to remain annealed (*22*). It has been proposed previously that single-strand binding protein (SSB) or the processivity factor, PCNA, might stabilize short RNA primers before being captured by the DNA polymerase (*23, 24*). Enzymes such as topoisomerase (*25*) and the HEL112 helicase (*26*) in *Sso* have been shown to have strand annealing activities necessary for their respective functions in genomic maintenance, but they are not thought to stabilize RNA primers during DNA replication.

Classical B-family polymerases involved in DNA replication have been well characterized kinetically and structurally in *Sso* (*27-29*). Recently, we have reported that the primary DNA polymerase in *Sso* (Dpo1) can form a trimer that activates the template directed polymerization ability of the enzyme (*30*). We now show that Dpo1 has a remarkable ability to stabilize weak base pairing interactions to replicate a template strand at high temperatures. This annealing activity is not recognized for either *Thermus aquaticus* pol I (Taq) or *Pyrococcus furiosus* B-family DNA polymerase (*Pfu*-Pol). In addition to stabilizing DNA templates, Dpo1 has a robust and unique terminal deoxynucleotidyl transferase activity that can extend ssDNA into extremely long products at high temperatures. We have confirmed that Dpo1's TdT activity can proceed by two mechanisms with differing kinetics: a traditional template independent or a more unique template dependent looping. The implications of strand annealing and terminal transferase activities for this multifaceted B-family DNA polymerase are discussed with regards to initiating DNA replication and repair for the maintenance of the genome in archaea.

3.1 MATERIALS AND METHODS

Materials

^{32}P-ATP was purchased from MP Biomedicals (Solon, Ohio). Unlabeled deoxyribonucleotides and ribonucleotides were purchased from Invitrogen (Carlsbad, CA). All single stranded DNA (ssDNA) were ordered from IDT (Coralville, IA) and gel purified as described previously (*30*). DNA substrates are listed in Table 2-1. Optikinase was purchased from USB (Cleveland, OH). Taq DNA polymerase is from Bioline (Taunton, MA) and *Pfu*Turbo DNA polymerase is from Stratagene (Santa Clara, CA). Wild-type (WT) Dpo1 and exonuclease deficient (exo$^-$) Dpo1 (D231A/D318A) were purified as described previously (*30*). Restriction enzymes were purchased from New England Biolabs (Ipswich, MA). All other reagents are analytical grade or better.

Determination of Melting Temperatures

DNA melting temperature measurements were conducted on Varian Cary 100 Bio UV-Visible Spectrophotometer. UV absorbance at 260 nm was measured for 1.5 μM DNA with or without 1.5 μM Dpo1exo$^-$ in assay buffer [50 mM Glycine (pH8), 5 mM $Mg(CH_3COO)_2$] at every integral temperature point from 4 °C to 90 °C programmed at a rate of 1 °C min^{-1}. The melting temperature was determined from a plot of the UV absorbance versus temperature for multiple replicates and analyzed using the included software.

Strand Annealing and Terminal Transferase Assays

Unless specified otherwise, 1.5 μM Dpo1 (WT or exo$^-$) was incubated with approximately 80 nM ^{32}P 5'- end labeled DNA and 100 μM of specified nucleotides in glycine buffer (50 mM Glycine (pH8.0), and 5 mM $Mg(CH_3COO)_2$) in 10 μl reaction volumes at 70 °C for 60 minutes. ^{32}P 5'-end labeling was performed using Optikinase according to manufacturer's directions. The reaction was stopped by adding equal volume of the quench solution (88% formamide, 10 mM EDTA, 1 mg/ml bromophenol blue, 0.1% SDS). The quenched reaction samples were separated on a 20% denaturing polyacrylamide gel or a 0.8% alkaline agarose gel, imaged overnight, and scanned using a phosphorimager (Storm 20, GE Healthsciences). Quantification of the band intensities and lengths was performed using ImageQuant software (v5.0) compared with a ^{32}P end labeled 1kb ladder (Invitrogen).

Restriction Digest Assays

1.0 μM Dpo1exo$^-$ was incubated with 100 μM dNTPs and approximate 280 nM ^{32}P 5'-end labeled DNA with a restriction site close to the 5'-end in assay buffer for 60 minutes at 70 °C. 3 μl of the reaction was used in a 10 μl restriction digest reaction containing 0 or 25 units of the restriction enzyme and the specific buffer. Digestion reactions were run at optimal temperatures as specified by the manufacturer for two hours. The reactions were stopped and analyzed by denaturing polyacrylamide gel as described above.

DNA Polymerase and Transferase Kinetics

Time courses for transferase and polymerase assays were analyzed by calculating the fraction of product (+1, +2, +3, or total) to substrate. Data from reactions with $(dT)_{20}AAA$ were fit to a single exponential equation using Kaleidagraph (Synergy Software) according to Equation 3-1:

$$[Product] = A[1 - \exp(-k_{obs}t)] \quad (3\text{-}1)$$

where A represents the amplitude and k_{obs} is the observed rate constant. Data from reactions with $(dT)_{20}$ were fit and modeled according to the minimal kinetic scheme outlined in Figure 3-7C using a kinetic simulation with Berkeley Madonna (University of California, Berkeley)

Simulation of Dpo1 Terminal Transferase/Polymerase Reactions

The simplified kinetic scheme in Figure 3-9E was used to develop a model to fit the kinetic data for extension of either $(dT)_{20}$ or $(dT)_{20}AAA$ and simulate the individual kinetic parameters. The following reaction scheme was simulated using Berkeley Madonna (University of California, Berkeley, CA) using our kinetic data according to the following differential equations:

A = [ssDNA-Dpo1] (3-2)

B = [ssDNA(+n)-Dpo1] (3-3)

C = [ssDNA(+3)-Dpo1] (3-4)

D = [ssDNA(+3)-Dpo1*] (3-5)

E = [ssDNA(+n)-Dpo1*] (3-6)

$$\frac{d[A]}{dt} = -k_1A - k_2A \quad (3\text{-}7)$$

$$\frac{d[B]}{dt} = k_1 A \quad (3\text{-}8)$$

$$\frac{d[C]}{dt} = k_2 A - k_3 C + k_{-3} D \quad (3\text{-}9)$$

$$\frac{d[D]}{dt} = k_3 C - k_{-3} D + k_4 D \quad (3\text{-}10)$$

$$\frac{d[E]}{dt} = k_4 D \quad (3\text{-}11)$$

k_1 is calculated from the observed rate constant for template independent transferase activity modeled from the data for dTTP on $(dT)_{20}$ or dATP $(dA)_{20}$. k_2 was simulated in our kinetic scheme and represents the upper estimate of the rate for the formation of the +1 to +3 products in a quasi template dependent fashion. k_3 and k_{-3} represent the forward and revered rate constants for annealing simulated from the lag in the data from $(dT)_{20}$ and dATP/dTTP (Figure 3-8C). k_4 was obtained from the exponential fit of the data from $(dT)_{20}$AAA and dATP/dTTP (Figure 3-8D). k_a and k_b represent annealing and denaturation, respectively, for products greater than +3 and were not simulated directly in our model. These values would change depending on the thermostability of the DNA base pairs at any particular step.

3.2 RESULTS

3.2.1 Thermal Stabilization and Annealing Activity of *Sso*Dpo1

We designed four hairpin DNA substrates (named according to their melting temperatures, T_M) with different T_Ms that place the annealed bases at the 3' end leaving 10 thymine bases in the template strand. A stretch of thymidines in the template region was included to limit any hairpin

formation to the designed site. The T_Ms of the DNA substrates without and with Dpo1 were experimentally measured according to Materials and Methods (Table 3-1). In all cases, addition of Dpo1 increased the T_M and stabilized weak base-pairing interactions for the DNA hairpins. There was an average increase of 5.5 °C in the hairpin T_M when Dpo1 was included.

Table 3-1. DNA Sequence and Melting Temperatures

DNA NAME	Sequence	T_M (°C)[2] DNA	DNA/Dpo1	ΔT_M (°C)
DNA32	T T T ATG T ||| T T TACTTTTTTTTT-5′	31.6 ± 1.3	37.0 ± 2.3	5.4
DNA37	T T T TACG T |||| T T T ATGCTTTTTTTTT-5′	37.0 ± 1.0	42.5 ± 1.4	5.8
DNA39	T T T TACG T |||| T T ATGCTTTTTTTTT-5′	38.6 ± 1.9	44.5 ± 1.4	5.5
DNA88	T CGCCGGCCCGGG T |||||||||||| T GCGGCCGGGCCCTTTTTTTTT-5?	88	n/d[3]	n/d[3]

1. Melting temperatures were only determined for hairpin DNA
2. Experimentally determined. 3. n/d – not determined

Remarkably, we have also found that Dpo1 is able to rapidly incorporate 10 bases to the 3'-end of DNA39 at a temperature (60 °C) 25 to 30 degrees higher than the T_M (Figure 3-1A). Although we primarily used trimeric concentrations of Dpo1 (1 μM) (*30*), lower monomeric concentrations (150 nM) had very similar activity (data not shown). We compared this associated annealing and polymerase activity of Dpo1 to other thermophilic polymerases, Taq and *Pfu*-Pol and found that this DNA stabilization and annealing activity is unique to Dpo1. At 40 °C, Dpo1 and Taq can polymerize DNA39, but only Dpo1 can function at 60 °C (Figure 3-1B). At 60 °C,

there is no polymerase product seen for Taq, while Pfu has only DNA exonuclease activity. As a positive control, all three polymerases can rapidly polymerize DNA88 at 60 °C (Figure 3-1C).

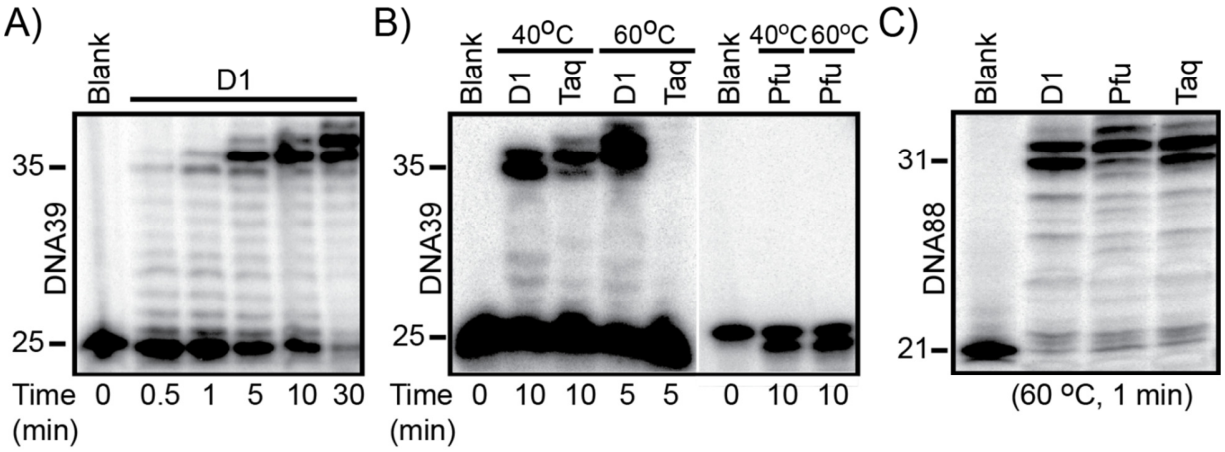

Figure 3-1. Strand Annealing Activity of *Sso*Dpo1

A) Dpo1 and 5'-end labeled DNA39 were incubated in a standard polymerase reaction as described in Materials and Methods at 60 °C for indicated times. B) 1 μM of Dpo1, 2.5 units of *Taq* DNA polymerase, or 1.25 units of *Pfu*Turbo DNA polymerase was incubated with ^{32}P 5'-end labeled DNA39 at 40 °C or 60 °C for indicated times. C) A thermally stable hairpin, DNA88, was used as a positive control for each DNA polymerase under standard reaction conditions for 1 minute at 60 °C.

To further explore the DNA annealing activity, we monitored a time course for polymerase activity of Dpo1 on DNA32, DNA37 and DNA39 at 50 °C (Figure 3-2). Dpo1 has polymerization activity on all three DNA hairpins, although the rate is much faster on DNA37 and DNA39 (Figure 3-2D). At longer times, there are polymerase products when using DNA32, but the low T_M of this substrate limits Dpo1's activity. We also examined the annealing activity for Dpo1 at two other temperatures (30 °C and 60 °C) for each DNA substrate (Figure 3-3). At

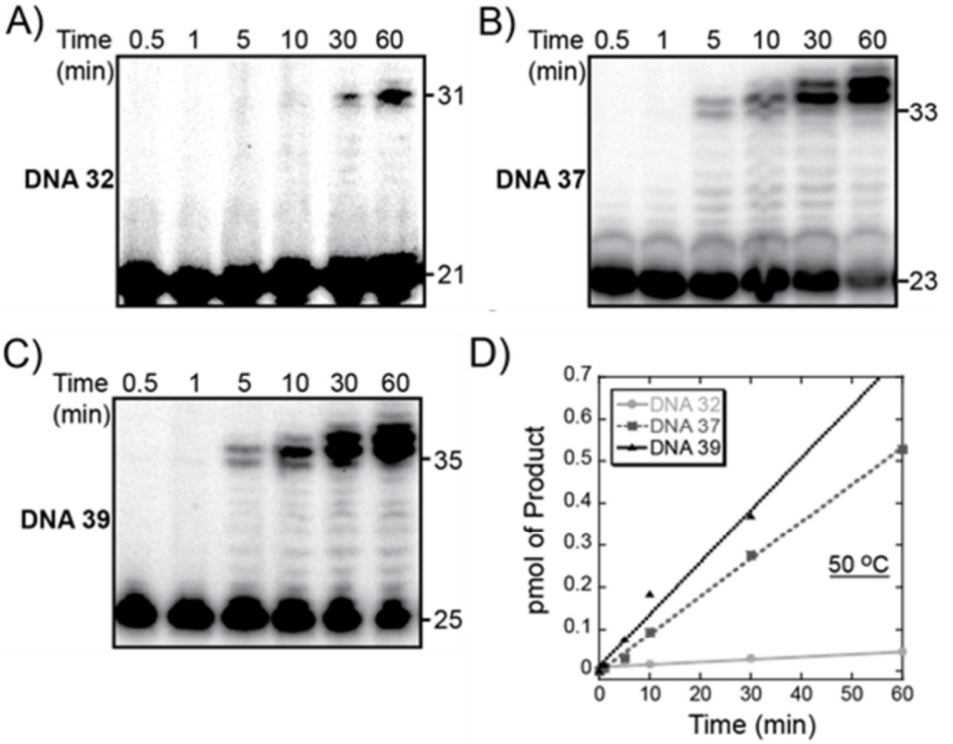

Figure 3-2. Time Course for Polymerase Activity at 50 °C

The time course for polymerase activity on DNA hairpin substrates A) DNA32, B) DNA37, and C) DNA39 at 50 °C. Concentrations of Dpo1 and DNA were held constant at 1 μM and 80 nM, respectively. The size of the polymerase products is shown to the right of the gel figures. D) The rate of polymerase activity is quantified and compared between each DNA hairpin, DNA32 (●), DNA37 (■), and DNA39 (▲) at 50 °C. The rates are 0.6 ± 0.1, 8.9 ± 0.1, and 12.4 ± 0.1 fmol min^{-1} for DNA32, DNA37, and DNA39, respectively.

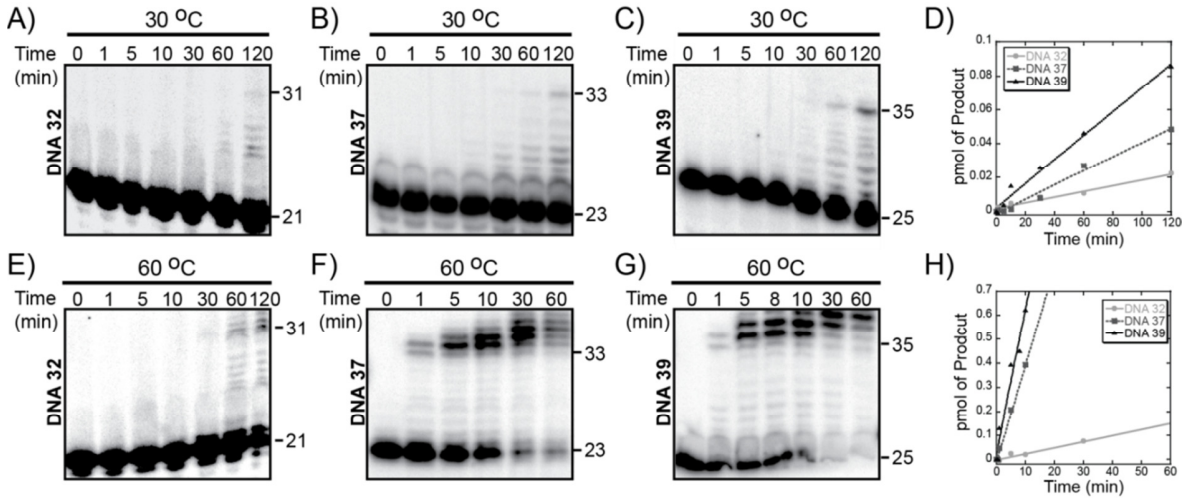

Figure 3-3. Time Course for Polymerase Activity at 30 and 60°C

The time course for polymerase activity on hairpin substrates A) or E) DNA32, B) or F) DNA37, and C) or G) DNA39 at 30 or 60 °C, respectively. Concentrations of Dpo1 were held constant at 1 M as described in Materials and Methods. The size of the polymerase products is shown to the right of the gel figures. D) The rate of polymerase activity at 30 °C is quantified and compared between each DNA hairpin, DNA32 (●), DNA37 (■), and DNA39 (▲). The rates at 30 °C are 0.2 ± 0.1, 0.4 ± 0.2, and 0.8 ± 0.1 fmol min^{-1} for DNA32, DNA37, and DNA39, respectively. H) The rate of polymerase activity at 60 °C is quantified and compared between each DNA hairpin, DNA32 (●), DNA37 (■), and DNA39 (▲). The rates are 2.6 ± 0.3, 40 ± 2, and 60 ± 6 fmol min^{-1} for DNA32, DNA37, and DNA39, respectively.

30 °C, the rate of synthesis is more equivalent for the three DNA substrates. The results at 60 °C are more similar to 50 °C where polymerization on DNA37 and DNA39 are significantly faster than on DNA32. Therefore, the efficiency of extension by Dpo1 is dependent on the reaction temperature and the T_M of the DNA substrate. Surprising still is Dpo1's ability to polymerize any product at a temperature that is 30 °C above the melting temperature of the DNA primer/template.

3.2.2 Template Specific Terminal Transferase Activities

Noticeably, Dpo1 begins to incorporate additional bases to the 3' end of the fully polymerized hairpin at higher temperatures (Figure 3-1B & C and Figure 3-3 F & G). All of the DNA hairpins used included ten thymidines in the template to limit hairpin formation to a single site. It is possible that template fluidity in these regions and resulting polymerase slippage could account for longer products. To test this, we monitored polymerase products using either a DNA template with a homopolymeric 5'-overhang (DNA39) or a template that contained three interspaced cytosines in the 5' overhang (DNA39C) (Figure 3-4, Table 3-2). Reactions with DNA39 resulted in fast extension to a 36 or 37 base product. This is consistent with either the addition of 1 or 2 extra bases to the end of the template in a terminal transferase reaction or the result of template dependent polymerase slippage. When DNA39C was used instead at 60 °C, the major product was 35 bases with a significant decrease of longer products. At 70 °C, though, a limited amount of longer products were detected with DNA39C. Therefore, the composition of the template sequence and the reaction temperature may result in the insertion of additional bases during DNA replication by Dpo1 in a polymerase slipping mechanism.

Table 3-2. ssDNA templates

ssDNA name	ssDNA sequence
d(A)20	5′-AAAAAAAAAAAAAAAAAAAA
d(T)20	5′-TTTTTTTTTTTTTTTTTTTT
d(G)20	5′-GGGGGGGGGGGGGGGGGGGG
d(C)20	5′-CCCCCCCCCCCCCCCCCCCC
d(T)20AAA	5′-TTTTTTTTTTTTTTTTTTTTAAA
EcoRV	5′-TTTTTGATATCTTTTTTTTTTTTTTTTT
Hind III	5′-TTTTAAGCTTTTTTTTTTTTTTTTTT
Apa I	5′-TTTTTGGGCCCTTTTTTTTTTTTTTTT
Swa I	5′-TTTTTATTTAAATTTTTTTTTTTTTTTT
DNA39C	5′-TTCTTCTTCTCGTATTTTTTTACG

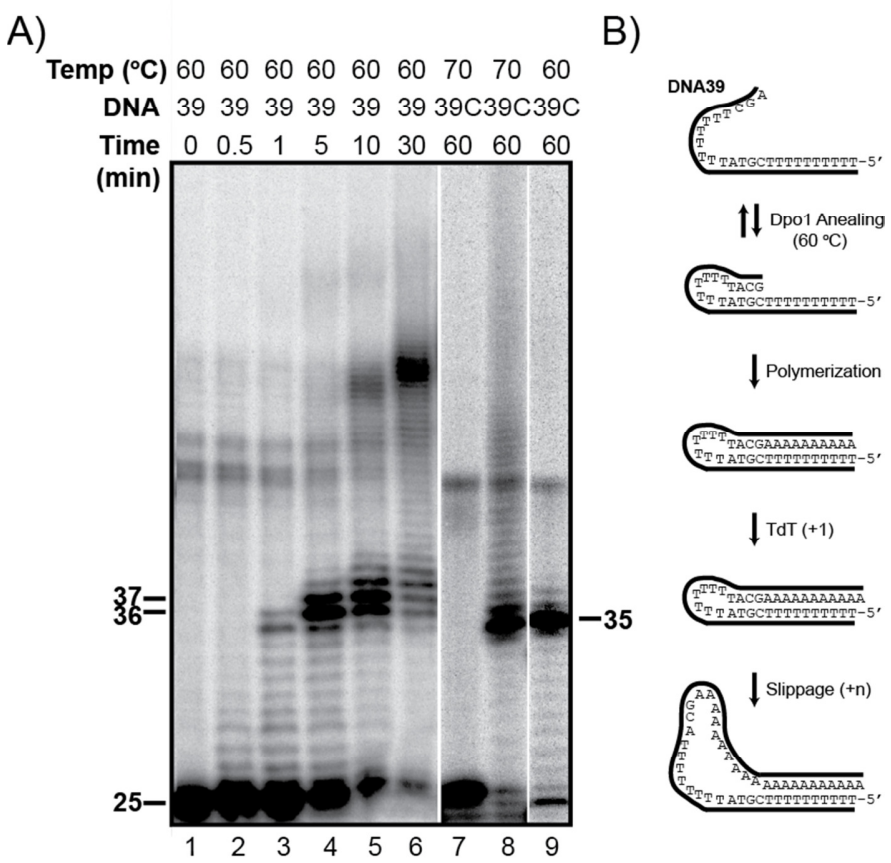

Figure 3-4. Polymerase Slippage Mechanism on Hairpin DNA39

A) The reactions were carried out as described in Materials and Methods with temperature, DNA substrates and time periods as indicated. Lanes 8 and 9 show that insertion of three cytosines on the 5' overhang/template (DNA39C) of the hairpin reduces the production of long products (>36 bases) observed for the DNA39 hairpin that only includes 10 thymidines in the 5' overhang (lanes 2-7). Lanes 1 and 7 show DNA only for DNA39 and DNA39C, respectively. B) Proposed model for the slippage of a DNA template during polymerization by Dpo1.

In order to verify whether direct terminal transferase activity is also possible for Dpo1, we switched to linear homopolymeric DNA substrates [$(dT)_{20}$, $(dC)_{20}$, $(dA)_{20}$, and $(dG)_{20}$] to remove any possible interfering Dpo1 directed annealing or hairpin formation. Homopolymeric DNA substrates greater than 10 bases were required for detectable TdT activity (data not shown). Similar to the DNA hairpins above, the terminal transferase activity on $(dT)_{20}$ not only increased with temperature but seemed to be activated at physiological temperatures greater than 60 °C (Figure 3-5A). Dpo1 was incubated at 70 °C with each homopolymer labeled at the 5'-end with ^{32}P for the indicated times (Figure 3-5C). Dpo1 shows strong TdT activity on both $(dT)_{20}$ and $(dA)_{20}$ but very weak activity on $(dG)_{20}$ and $(dC)_{20}$. We have noticed that at times greater than 60 minutes, there is a decrease in the amount of product that remains in the wells. We attribute this loss to precipitation or the inability of extremely large DNA molecules to enter an acrylamide gel. TdT activity is detected for both wild-type and exonuclease mutant (Dpo1exo$^-$) forms of Dpo1 on $(dT)_{20}$ although the rate is accelerated roughly two-fold for Dpo1exo$^-$ (Figure 3-5E). In order to eliminate any competing nuclease proofreading activity and to more accurately measure transferase activity, Dpo1exo$^-$ is used in all subsequent experiments.

In order to better visualize longer TdT products, we analyzed the reactions on denaturing agarose gels. After a significant lag phase of greater than 10 minutes, abundant TdT products greater than several thousand bases are seen (Figure 3-6). There is a wide distribution of product length during the reaction times. The average rate of transferase activity over 20-40 minutes on $(dT)_{20}$ was 115 ± 4 bases/min while the maximal rate was 310 ± 27 bases/min.

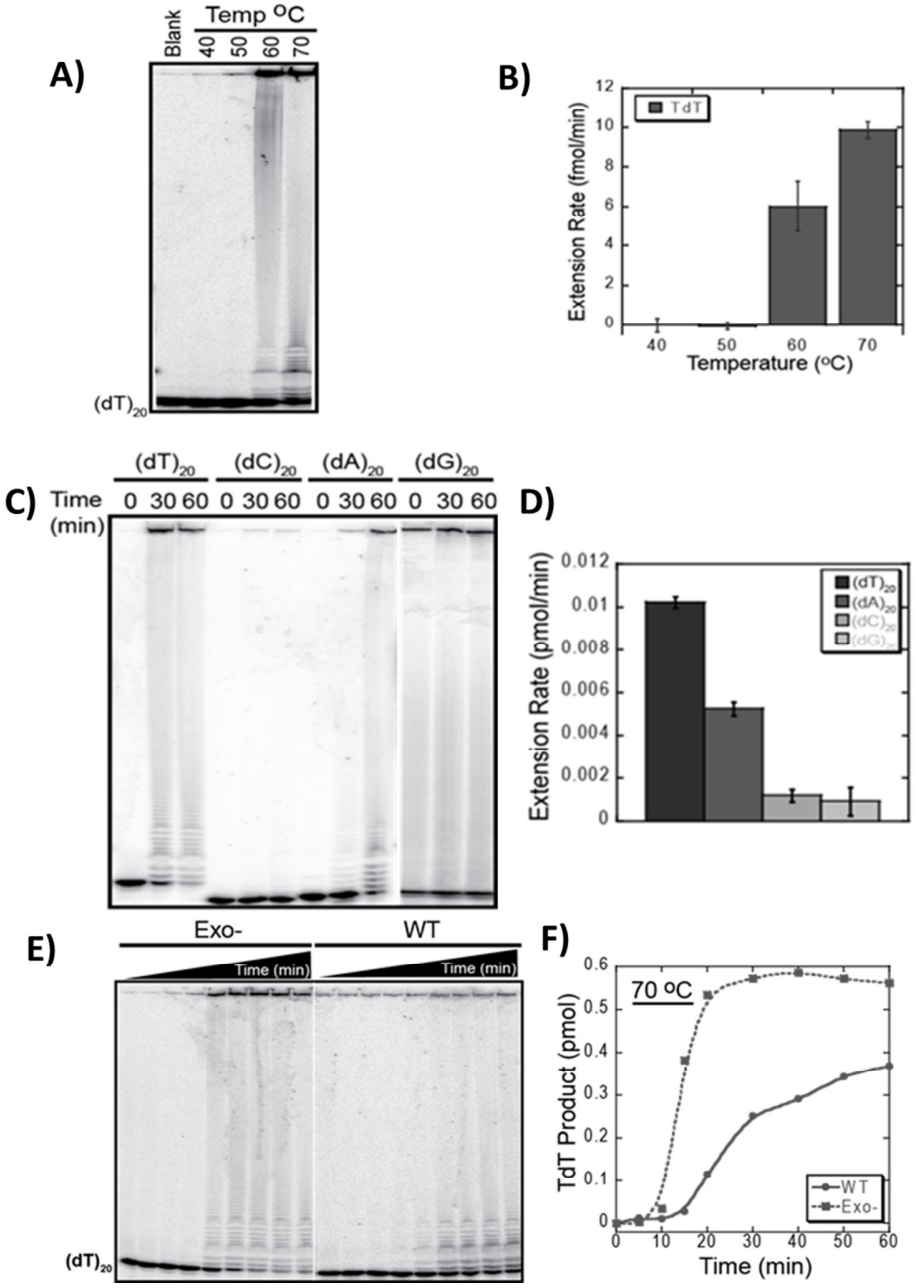

Figure 3-5. DNA Template Dependence of Transferase Activity

A) Transferase activity of Dpo1 on $(dT)_{20}$ were tested at 40, 50, 60 and 70 °C for 60 minutes as described in Materials and Methods. B) Products from three independent experiments were quantified as rates and plotted as a function of temperature. The rate of transferase activity was essentially zero for 40 °C and 50 °C, 6.0 ± 1.3 fmol min^{-1} at 60 °C, and 9.8 ± 0.4 fmol min^{-1} at 70 °C. C) Terminal transferase activity of Dpo1 measured using homopolymeric oligos of $(dT)_{20}$, $(dC)_{20}$, $(dA)_{20}$ and $(dG)_{20}$ tested in standard assay reactions for 30 or 60 minutes. D) Quantification of the transferase rates are 10.2 ± 0.2 fmol min^{-1} for $(dT)_{20}$, 5.3 ± 0.3 fmol min^{-1} for $(dA)_{20}$, 1.2 ± 0.3 fmol min^{-1} for $(dC)_{20}$, and 0.9 ± 0.7 fmol min^{-1} for $(dG)_{20}$. E) Shows a time

(Figure 3-5, continued)
course of terminal transferase for both Dpo1exo$^-$ and WT Dpo1 on $(dT)_{20}$. Time points are 0, 5, 10, 15, 20, 30, 40, 50, and 60 minutes. F) The elongation products were quantified and plotted as a function of time for Dpo1exo$^-$ (■) and WT Dpo1 (●). There is a roughly 35% reduction in transferase activity for WT Dpo1 compared to Dpo1 exo$^-$.

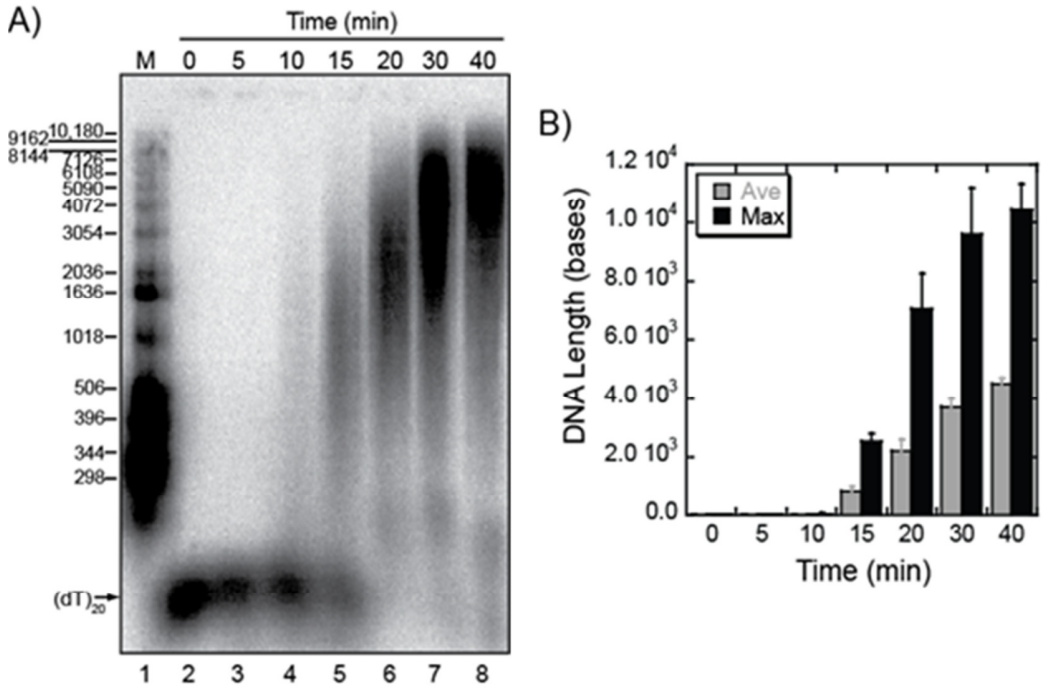

Figure 3-6. Product Length of *Sso*Dpo1 TdT Activity

Product size distribution of Dpo1 TdT activity was determined by 0.8% agarose gel. A) A 40-minute time course for standard TdT assay as described in the Materials & Methods. The marker size is labeled on the left to the gel, time points on the top, and lane number at the bottom. B) The maximum size (●) and average size (○) of the product at each time point were calculated using the ImageQuant.

3.2.3 Nucleotide Specificity for TdT Activity

We then examined the incorporation of individual nucleotides on $(dT)_{20}$, $(dA)_{20}$, $(dC)_{20}$, and $(dG)_{20}$. Dpo1 transferase reactions were incubated with individual nucleotides and quenched after 60 minutes. Dpo1 was unable to incorporate ribonucleotides (rNTPs) into any DNA substrate we tested (data not shown). No visible elongation of $(dT)_{20}$ was observed with dGTP or dCTP (Figure 3-7A). dAMP was clearly incorporated repetitively on the 3'-end of $(dT)_{20}$ to create long terminal transferase products. Reduced TdT activity was also noted with dTTP on a $(dT)_{20}$ substrate. A time course of this reaction showed a slow increase in product length with time (data not shown). When only dATP is used as the substrate, there is a faster addition of three elongation of $(dT)_{20}$ was observed with dGTP or dCTP (Figure 3-7A). dAMP was clearly incorporated repetitively on the 3'-end of $(dT)_{20}$ to create long terminal transferase products. Reduced TdT activity was also noted with dTTP on a $(dT)_{20}$ substrate. A time course of this reaction showed a slow increase in product length with time (data not shown). When only dATP is used as the substrate, there is a faster addition of three bases followed by a slower extension of the template. This is indicated in lanes 2, 6, 8, and 9 in Figure 3-7A, where there is a buildup of the +1, +2, and +3 products (also seen in Figure 3-9A). Experiments with dual combinations of each dNTP and constant dATP indicate that inclusion of dTTP stimulates the transferase rate equivalent to when all dNTPs are utilized (Figure 3-7A). There is a disappearance in the +1 to +3 products indicating that subsequent steps are no longer rate limiting when dTTP is included with dATP. Inclusion of dGTP or dCTP with dATP did not increase transferase activity over dATP alone.In contrast, dTTP is required to efficiently elongate $(dA)_{20}$ by Dpo1, while dATP alone resulted in a slow extension of TdT products. Neither dGTP nor dCTP resulted in any detectable products over the time course. Again, there is a build up of the +1, +2, and +3 products with

dTTP indicating an initial fast addition followed by slower extension to generate products greater than +4. Inclusion of dATP with dTTP or all dNTPs elevates the terminal transferase activity similarly to a rate seen with $(dT)_{20}$ above (Figure 3-7B). Interestingly, although incorporation of dNTPs onto $(dC)_{20}$ or $(dG)_{20}$ is greatly reduced, the preference is for the complementary base to the template to be added to the 3' end (Figure 3-7C) similar to that detected for $(dT)_{20}$ and $(dA)_{20}$.

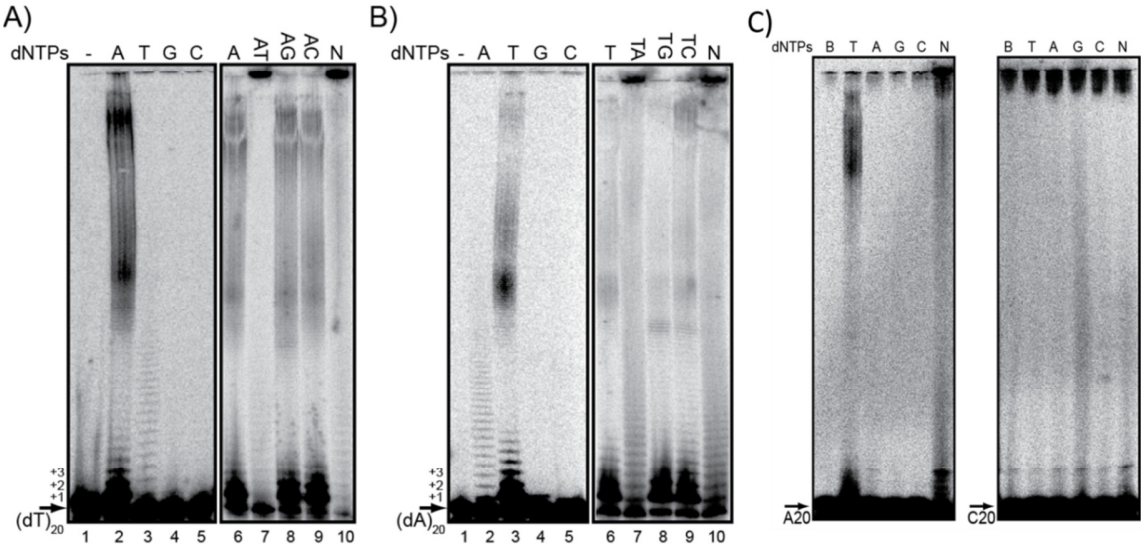

Figure 3-7. Effect of of Deoxyribonucleotides on *Sso*Dpo1 TdT Activity

Test of individual or combination of deoxyribonucleotides on the terminal transferase activity on A) $(dT)_{20}$ and B) $(dA)_{20}$ in standard assay reactions for 60 minutes at 70 °C as described in Materials and Methods. Various combinations of dNTPs were used in the reaction as indicated. N represents the inclusion of all dNTPs. Concentrations of dNTPs were held constant at 0.1 µM. +1, +2, and +3 note the location of individual TdT products. C) Shows the nucleotide specificity T, A, G, or C on the terminal transferase activity of Dpo1 on either $(dA)_{20}$ or $(dC)_{20}$. Complementary bases, dTTP and dGTP, are utilized specifically for $(dA)_{20}$ or $(dC)_{20}$, respectively. Inclusion of all dNTPs (N) show increased TdT activity. B represents the DNA only.

3.2.4 Loop Back Annealing Model for Transferase Activity

The efficient incorporation of complementary bases to the homopolymeric templates suggests that Dpo1 is promoting annealing as a mechanism to extend an ssDNA template. To test this, we designed four ssDNA substrates that included restriction sequences near the 5'-end. The templates had at least 15 thymidines at the 3' end determined to be an optimal template for Dpo1's TdT activity. If after addition of a few complementary bases to the 3'-end, Dpo1 utilized the ssDNA as a template for a loop back polymerization reaction, then dsDNA would be synthesized across the restriction site, and subsequent restriction digestion would liberate a small ^{32}P-labeled ssDNA product (Figure 3-8A). Terminal transferase reactions were performed as above at 70 °C with each of the four DNA templates containing distinct restriction sites. After the TdT reactions, each of the four restriction enzymes were added and incubated in standard conditions. The appearance of a 7-9 base product after restriction digest (Figure 3-8B, lanes 4, 8, 12, and 16) is indicative of looping back and synthesizing DNA in a template dependent manner across the restriction site. Templates with restriction sites, *EcoRV* and *SwaI*, resulted in 8 bases products, while the *HindIII* and *ApaI* templates resulted in 7 and 9 base products, respectively. Control reactions with no restriction enzyme added did not show any degradation (Figure 3-8B, lanes 3, 7, 11, and 15). Interestingly, not all of the transferase products were digested even after two hours of reaction. These products can be attributed to either a template independent transferase activity or severe template dependent slippage precluding replication across the restriction site.

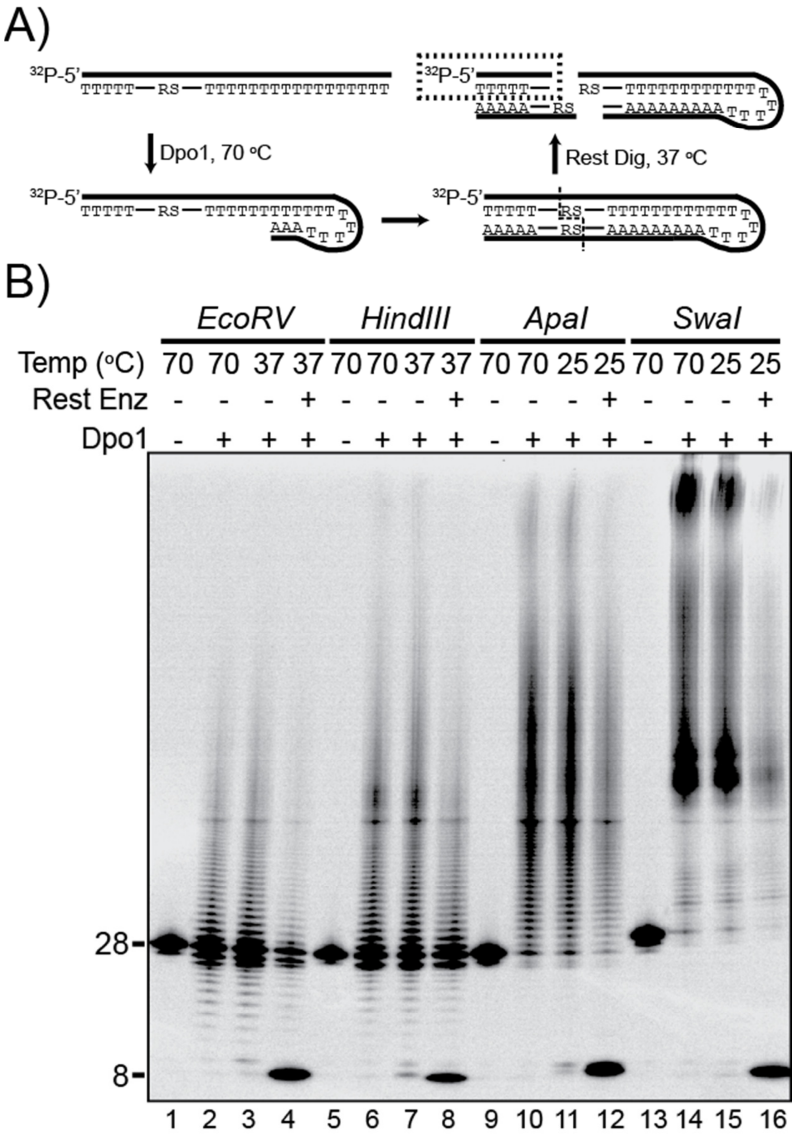

Figure 3-8. Loop Back Annealing Mechanism

A) Model of hairpin DNA with a restriction site close to the 5' end. A loop back mechanism that involves template dependent polymerization would create a competent restriction site. Cleavage of the resulting product with a specific restriction enzyme would liberate a short (7-9 base) ^{32}P-labeled ssDNA at the 5' end. B) Restriction digestion analysis of Dpo1 transferase products. Lanes 1, 5, 9 and 13 are blanks consisting of DNA only for transferase reactions of lanes 2, 6, 10 and 14 that include Dpo1. Lanes 3, 7, 11 and 15 are control reactions for restriction digestion that don't include restriction enzymes. Lanes 4, 8, 12 and 16 are the restriction digested products from the Dpo1 transferase assays incubated with specific restriction enzymes resulting in 8, 7, 9, and 8 base products, respectively.

3.2.5 Kinetic Mechanism for TdT Activity

Kinetic TdT assays were performed at 70 °C on $(dT)_{20}$ or $(dT)_{20}AAA$ with either dATP or dATP/dTTP as a function of time at 70 °C. The kinetic time course with $(dT)_{20}$ involved a lag phase followed by robust TdT activity with either dATP or dATP/dTTP (Figure 3-9A, B, C & E). During this lag, there is an obvious buildup of +1 to +3 products with $(dT)_{20}$ and dATP. The lag phase is reproducible at different temperatures and combinations of nucleotides. Preincubation of enzyme with $(dT)_{20}$ template at 70 °C and initiation with dNTPs shows the same lag as when the reaction is initiated by the addition of enzyme (data not shown). Although the lag persists, the +2 and +3 products disappear when dTTP is included (Figure 3-9E) indicating that formation of +2 or +3 products are no longer rate limiting. Therefore, the lag is most likely due to a slow step in the rearrangement of Dpo1 and/or the DNA template into the correct conformation for rapid template dependent extension at the 3' OH end.

Because the incorporation of dNTPs on the 3' end seemed to be most active when the complementary base was used (Figure 3-7) and we noted the buildup of a +3 product, we also tested TdT activity on $(dT)_{20}AAA$ with either dATP or dATP/dTTP (Figure 3-9D & F). Interestingly, the lag phase and any +2 or +3 products disappeared, and the results fit easily to a single exponential equation consistent with a single observed rate constant with dATP (0.027 ± 0.003 min^{-1}) or dATP/dTTP (0.090 ± 0.005 min^{-1}). The observed rate with dATP/dTTP is three fold faster than for dATP alone most likely due to the ability to perform template directed synthesis equally across from template thymidines and adenines or newly incorporated bases complementary to the template $(dT)_{20}AAA$. The addition of three adenines at the 3'-end of the ssDNA template also allows for the possible preorganization of the loop back annealing

eliminating the lag and resulting in faster template dependent transferase activity than with $(dT)_{20}$ alone. Interestingly, these reactions are performed at 70 $^{\circ}$C, which is well above the melting temperature for formation of a hairpin with $(dT)_{20}$AAA, and so, Dpo1 again is able to stabilize weak base pairing interactions for rapid extension in a template dependent manner. Kinetic modeling of the data from Figure 3-9E & F to the minimal reaction scheme shown in Figure 9G resulted in the fit of rate constants for each of the competing steps. The observed rate constant for template independent base additions (k_1) was modeled from the average of the data from Figure 3-7A, lane 3 and Figure 3-7B, lane 2 and a kinetic time course (data not shown) following the incorporation of dT or dA only on a $(dT)_{20}$ or $(dA)_{20}$ template, respectively, where no base pairing is possible. In the presence of all nucleotides, terminal transferase activity for the first few base additions is slightly faster (k_2). On the basis of the results from Figure 3-7, this initial transferase activity is dependent on the formation of a quasi-stable state that incorporates nucleotides complementary to the template but not stable enough to iteratively incorporate nucleotides opposite the ssDNA template without melting.

The rate constants for annealing (k_3 and k_{-3}) were derived from a fit of the data in Figure 3-9E for $(dT)_{20}$ and dATP/dTTP while independently modeling each rate constant (k_1, k_2, and k_4) based on initial values described above (see Materials and Methods in this chapter for details). Inclusion of the slower quasi stable step for +1 to +3 additions (k_2) was necessary to improve the fit of the simulation. Based on the simulation results, a reversible step of annealing was fit with a forward rate constant (k_3) of 0.16 min^{-1} and a reverse rate constant (k_{-3}) of 0.03 min^{-1} for annealing of the +3 product. The rate constant for this equilibrium step is modeled from data using $(dT)_{20}$ and will likely increase if more thermodynamically stable annealing intermediates are available (either k_a or k_b). The rate constant for rapid extension (k_4) was fit to 0.11 min^{-1} and

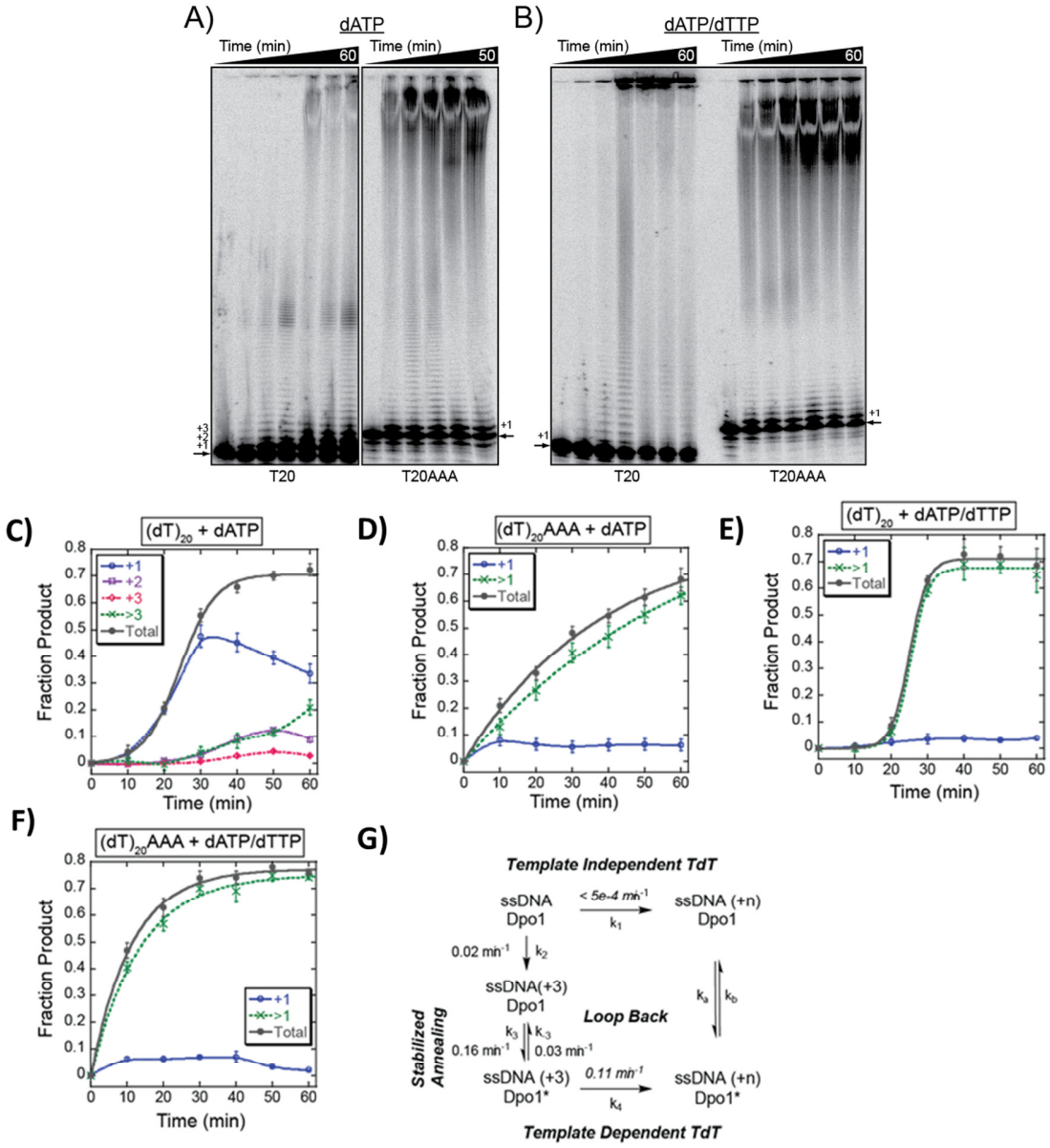

Figure 3-9. Kinetic Mechanism for TdT Activity

A) and B) Show a time course for templates $(dT)_{20}$ and $(dT)_{20}AAA$ with dATP (A) and (B) dATP/dTTP as substrates. Reactions were conducted as described in Materials and Methods in this chapter. Time points are 0, 10, 20, 30, 40, 50 and 60 minutes. Arrow indicates the position of either $(dT)_{20}$ or $(dT)_{20}AAA$ template. +1, +2, or +3 indicate the position of initial extension products. C-F show quantification of the time course for $(dT)_{20}$ and $(dT)_{20}AAA$ using (C and D) dATP and (E and F) dATP/dTTP as nucleotide substrates. All values of data points for +1, +2, +3, >1, >3 and total were averaged from at least three independent experiments. The data points (>1 and total) for $(dT)_{20}AAA$

(Figure 3-9, continued)
were fit to a single exponential Equation 3-1 described in Materials and Methods in this chapter to extract observed rate constants for dATP (0.031 ± 0.001 min^{-1}) and dATP/dTTP (0.089 ± 0.006 min^{-1}). G) The kinetic reaction pathway was simulated and modeled based on the experimental data from C) and F) as well as Figure 3-5 as described in Materials and Methods. k_1 is an upper estimate of the template independent transferase rate. k_2 represents the initial formation of +1to +3 products required for stable loop back annealing. k_2 may or may not be necessary and will depend on the sequence of the DNA and available probability of hairpin formation [$(dT)_{20}$ vs $(dT)_{20}AAA$]. k_3 and k_{-3} represent the loop back step required for the template dependent transferase mechanism resulting in a starred (*) ssDNA(+3)/Dpo1* species. k_4 represents the rate constant for template dependent terminal transferase activity calculated from the fit of the data for $(T_{20})AAA$ and in F). k_a and k_b represent subsequent annealing equilibria steps for longer terminal transferase products dependent of DNA sequence and temperature.

compares well with the observed rate constant for extension of $(dT)_{20}AAA$ (Figure 3-9F). Based on our results above for a template of this type, Dpo1 is most likely extending ssDNA in a template dependent manner that involves slippage of the synthesized polyA strand on a polyT template. Therefore, k_4 includes individual components of template dependent DNA synthesis and DNA template slippage. Inclusion and simulation of reverse rate constants for k_1, k_2, and k_4 did not improve the fits and were removed for simplicity. As the rate constant for template dependent synthesis for Dpo1 was measured previously to be 189 s^{-1} at 56 °C (*27*), k_4 most likely represents the slower polymerization slippage rate.

The rate of transferase activity in base pairs per minutes is difficult to quantify, as there is a broad distribution of product sizes. As can be seen in Figure 3-9B for $(dT)_{20}AAA$ at ten minutes, there are smaller products of less than 20 bases, but there is also significant longer products that remain in the well. The rate of transferase activity is fastest when using ssDNA as the initiator as long products are seen within ten or twenty minutes depending on the template. Transferase activity is observed on duplex and pt/DNA (Figure 3-2) but the length of products are much less within that time. Terminal transferase activity on a stable primer-template hairpin

(DNA39C) is estimated to be < 0.5 bases/min at 70 °C. The maximal template dependent terminal transferase rate measured on $(dT)_{20}AAA$ is greater than 20 bases/min at 60 °C.

3.3 DISCUSSION

DNA dependent DNA polymerases coordinate the positioning of the incoming nucleotide with the template strand for optimal catalysis. Upon completion of polymerization opposite a template, further nucleotide addition is disfavored when no templating base is available. Some DNA polymerases that lack a proofreading exonuclease domain can inefficiently incorporate a single (+1) nucleotide addition (*31-33*). Other polymerases from the X-family such as calf thymus TdT, Pol λ, and Pol μ, have the surprising ability to incorporate successive nucleotide additions in a template independent manner. Only X-family DNA polymerases have been shown to have consecutive nucleotide additions on single strand DNA (*7, 12, 34*), and Pol λ and Pol μ are the only DNA polymerase known to possess both template dependent and independent activities (*35, 36*). Neither of these X-family DNA polymerases have a proofreading exonuclease domain that can limit the terminal transferase activity.

Sulfolobus solfataricus is a hyperthermophilic aerobic archaeon which optimally grows at 80 °C. How this organism survives and propagates at such extreme conditions is not fully understood. Severe destabilizing thermodynamic forces that persist at high temperatures would make maintaining annealed dsDNA templates difficult. It is thought dsDNA is stabilized by specific binding proteins, which is necessary to protect the integrity of the genome at high temperatures. *Sso*Dpo1 is classified as a model B-family DNA polymerase due to sequence and functional homologies (*37*). It has rapid polymerization kinetics (*27*), efficient exonuclease

proofreading (*38*), high processivity when bound to PCNA (*39*), and structural homology to other B-family members (*28*). In this study, we provide evidence that Dpo1 has the inherent ability to stabilize thermodynamically weak base pairing interactions to facilitate template dependent DNA polymerization. Surprisingly, this B-family DNA polymerase (Dpo1) is also found to have robust terminal transferase activity that proceeds by two independent mechanisms: a loop back annealing template dependent polymerase activity as well as a slower template independent TdT activity. These annealing and transferase activities have not been noted for any other DNA replication polymerase within this family and may provide a mechanism for efficient replication and repair at high temperatures.

3.3.1 Role of DNA Polymerase Strand Annealing/Stabilization in DNA Replication

During initiation of DNA replication, a small RNA primer is synthesized by a DNA primase as a template for the DNA polymerase. The length of this RNA primer ranges from 4 - 14 nucleotides depending on the species (*22, 40, 41*).The stability of the primer template in hyperthermophilic organisms such as archaea is a significant thermodynamic issue that is thought to be overcome primarily through protein binding and stabilization. Even mesophilic organisms require the stabilization of this short RNA primer before a DNA polymerase can be loaded at the priming site. In *E. coli* and bacteriophage T4, this is accomplished through interactions and stabilization by SSB and PCNA (*23, 24*). In this report, we can show that the B-family DNA polymerase in *Sso* (Dpo1) is able to stabilize weak base pairing interactions of a primer template in order to extend the primer in a template dependent manner. This activity may be required to efficiently extend a short RNA primer during DNA replication at high temperatures.

We have been able to directly measure this increase in stabilization in DNA melting temperature assays that include Dpo1. Although there is a significant increase in the T_m when Dpo1 is included, the value does not approach the reaction temperature. Therefore, there is still a dynamic equilibrium between annealed hairpin and free DNA at 70 °C. Dpo1 must capture the annealed conformation and in the presence of dNTPs extend the hairpin in a template dependent manner. Unlike Dpo1, neither Taq nor *Pfu*-Pol could extend primers at 20 °C above the T_M. The polymerization of DNA with this template (to 35 bases) does not result in a thermodynamically stable duplex at 70 °C even with Dpo1 bound and an equilibrium between melted and annealed hairpin would persist. Therefore, there is a balance between the stability of the primer template (T_M) and the reaction temperature that determines the efficiency of extension for Dpo1.

3.3.2 DNA Polymerase Template Slipping

Homopolymeric stretches or repeats of DNA sequence are typically difficult to replicate accurately without template slipping that results in errors including insertions or deletions (*42-44*) and are even associated with a variety of diseases (*45*). The ability of one strand to slide relative to the other in these stretches of DNA is increased at higher temperatures. In order to specifically control annealing and hairpin formation by Dpo1, we utilized a stretch of thymidines in the template region that could not participate in base pairing. As a consequence of using these templates, we detected the formation of extremely long DNA products that are significantly greater than would be expected if synthesis only occurred opposite the template thymidines. This activity was more robust at higher temperatures. Therefore, either polymerase slippage or template independent terminal transferase activity is occurring. As a means to reduce template slippage, three regularly spaced cytosines were included in the template region (DNA39C). With

this template, slippage was reduced but did not eliminate the formation of longer DNA products. Therefore, longer products synthesized from DNA39C are most likely the result of limited DNA slippage and/or template independent terminal transferase activity.

3.3.3 Both Template Independent and Dependent Terminal Transferase Activities

Terminal transferase activity is characterized by the incorporation of nucleotides onto the 3' end of a single-stranded DNA in a template-independent manner. Biochemical studies have shown that calf thymus TdT requires ssDNA as an initiator that is at least three nucleotides long (*46*). The elongation of a homopolymer by calf thymus TdT requires at least six nucleotides for poly(dA) and five nucleotides for poly(dT) (*46*). Dpo1 requires a slightly linger homopolymeric template (10-15 bases) for efficient TdT activity. This is most likely due to the larger site size of Dpo1 for binding DNA(*30*) compared with calf thymus TdT.

Dpo1 is able to preferentially extend $(dT)_{20}$ and $(dA)_{20}$ over that of $(dG)_{20}$ or $(dC)_{20}$. From a molecular point of view, this TdT efficiency variation on different ssDNA initiators may be the result of optimal DNA conformations or favorable interactions between the template and the active site amino acids that direct the ssDNA elongation. This mechanism may resemble that shown by Pol μ where a conserved His orientates an incoming nucleotide towards a productive complex in the absence of a templating base for TdT activity (*12*). Additionally, both $(dG)_{20}$ and $(dC)_{20}$ will most likely have significant secondary structures, including quartets and quadraplexes (*47*), that would preclude Dpo1 binding and terminal transferase activity.

The studies on nucleotide selections of TdT enzymes have shown diverse rules. Other DNA polymerases have been shown to place an adenine opposite an abasic site during DNA

translesion repair (*48*) or at the end of a polymerized product in the +1 position (*3*). The short isoform of murine TdT was reported to indiscriminately incorporate both dNTPs and rNTPs to the 3' end of a $(dA)_{10}$ initiator with a higher efficiency for dNTPs resulting in longer products (~18 bases added to the initiator) (*49*). The preference to incorporate complementary bases to the end of the homopolymer DNA initiators by Dpo1 has similarities to the selection by human polymerase λ. Pol λ was found to incorporate nucleotides on the 3' end either independently resulting in +1 or +2 products or following a loop back model that utilized the template strand to initiate elongation resulting in longer +15-17 products (*36*). Interestingly, Pol λ is classified as a terminal transferase enzyme even though the primary mechanism for extension is template dependent.

Even though, complementary base addition is favorable, Dpo1 can incorporate nucleotides identical to the homopolymeric template at a slower rate in a template independent manner. For example, incorporation of dA on $(dA)_{20}$ and dT on $(dT)_{20}$ is clearly seen. This template independent terminal transferase activity has not been detected for any other B-family DNA polymerase. Therefore, Dpo1 has the unique ability to extend ssDNA with at least two terminal transferase mechanisms.

3.3.4 Looping Back as a Mechanism for Terminal Transferase Activity

The transferase activity of Dpo1 is optimal at higher temperatures (> 50 °C) resulting in extremely long products. Using homoligomeric DNA substrates, we detect a lag phase where Dpo1 extends and rearranges the ssDNA template into a productive complex before template dependent terminal transferase activity of Dpo1 is activated. This lag phase can be reduced with DNA substrates that include three residues on the end ($(dT)_{20}$AAA) that can participate in the

formation of a weakly stable hairpin. Because of this result, we speculate that Dpo1 can facilitate the looping back and stabilization of weak base pairing interactions to extend ssDNA in a template dependent manner.

Upon modeling and simulation of the kinetics of terminal transferase activity for Dpo1, it is clear that there are multiple kinetic routes. The preferred pathway involves the stabilization of weak base pairing by Dpo1 to form a hairpin where extension occurs in a template dependent manner. The rate limiting step for this pathway involves the template independent extension of ssDNA to produce favorable loop back annealing by Dpo1. The annealing step is reversible and will be dependent on the thermodynamics of the resultant DNA hairpin. If a short complementary sequence can be readily stabilized by Dpo1, then template dependent terminal transferase is the preferred mechanism. If no complementary region can be stabilized, then Dpo1 can extend DNA in a template independent manner at a much reduced rate before possible switching to a template dependent mechanism.

When dATP/dTTP was included in a reaction with either $(dT)_{20}$ or $(dA)_{20}$, the rate of transferase activity was significantly increased over that with the single nucleotide. This extension rate with dATP/dTTP is equivalent to when all nucleotides are included. This provides evidence for an increasingly complicated loop-back mechanism for transferase activity. When only dATP is available, looping back occurs and slippage is restricted to the $(dT)_{20}$ region because dATP is the only available substrate complementary to the T template. However, when both dATP and dTTP are available, the slippage reaction is expanded from the $(dT)_{20}$ region to the newly synthesized poly-dA region. The availability of both complementary nucleotides would eliminate any restrictions on slippage and increase the transferase rate and product length. For $(dT)_{20}$ when dATP is used, initial complementary extension is preferred resulting in a

buildup of a +1 to +3 products. Our results suggest that initial +3 extension exists in a quasi stable state that fluctuates between annealed and melted states. Once three bases of complementarity are available, rapid template dependent extension occurs as indicated by the results with $(dT)_{20}AAA$.

The length of terminal transferase products synthesized by Dpo1 routinely exceeds 5000 bases. These extreme TdT product lengths synthesized by Dpo1 are much greater than the 10-20 base products produced by any other DNA polymerase (Pol μ, TdT, or Pol λ). Interestingly, analysis of the crystal structure identified a small extra alpha helical region within Dpo1 that is not homologous to other polymerase members within the B-family (*28*). This region makes contact with the fingers domain which undergoes large conformational changes during nucleotide selection. The uniqueness of this region may provide for the conformational transition required for the synthesis of TdT products by Dpo1.

3.3.5 Role of Terminal Transferase Activity in Archaea

In addition to the suggested DNA priming role of archaeal PriSL, this enzyme also has a promiscuous terminal nucleotidyl transferase activity (*17*). In contrast to Dpo1, the terminal transferase activity of PriSL was greater for incorporation of NTPs versus dNTPs on a DNA template. PriSL also displayed some complex template and nucleotide specificities that gave products ranging from 2 to 7000 bases. Interestingly, the optimal temperature for PriSL transferase activity (~50 °C) was less than the physiological temperature for *Sso*, while the RNA primer products (2-14 bases) were more abundant at 70 °C. Dpo1, on the other hand, has an absolute preference for dNTPs and is active at physiological temperatures. It is conceivable that

the terminal transferase activity of PriSL is vestigial and therefore active only at lower temperatures and its primary role is DNA priming in *Sso*.

Terminal transferase activity has been shown to be required for both V(D)J recombination and NHEJ repair in eukaryotes. As there is no obvious need for V(D)J recombination in archaea, the terminal transferase activity of Dpo1 would most likely participate in the repair of double strand breaks. There are two pathways of double strand break (DSB) repair in eukaryotes. Homologous recombination (HR) is the more conserved and precise method for repairing DSBs. Within *Pyrococcus furiosus*, four HR genes have been identified and reconstituted biochemically to efficiently process the DNA for homologous strand exchange *in vitro* (*50*). Further experiments have examined the transcriptional response to double strand breaks in archaea and have identified that the homologous recombination genes, RadA and associated HR genes, were up regulated (*13*). Although HR seems to be the preferred pathway for double strand break repair in archaea, a 3' phosphoesterase module within the NHEJ DNA ligase that performs 3' healing reactions at DSBs has been shown to be conserved in some archaeal species giving some credence to this mode repair as well (*51*). Further genomic and biochemical characterizations will be needed to determine if NHEJ is a functional repair pathway in archaea, and if the terminal transferase activity of Dpo1 shown here can participate.

REFERENCES

1. Prakash S, Johnson RE, and Prakash L. (2005) Eukaryotic translesion synthesis DNA polymerases: Specificity of structure and function. *Annu. Rev. Biochem.* 74, 317-353.

2. Goodman, M. F. (2002) Error-prone repair DNA polymerases in prokaryotes and eukaryotes. *Annu. Rev. Biochem.* 71, 17-50.

3. Forrest, R., Zhou, H., and Hickford, J. G. (2000) 3'overhangs influence PCR-SSCP patterns. *Biotechniques.* 29, 958-960, 962.

4. Motea, E. A., and Berdis, A. J. (2010) Terminal deoxynucleotidyl transferase: The story of a misguided DNA polymerase. *Biochim. Biophys. Acta.* 1804, 1151-1166.

5. Bollum, F. J. (1959) Thermal conversion of nonpriming deoxyribonucleic acid to primer. *J. Biol. Chem.* 234, 2733-2734.

6. Desiderio, S. V., Yancopoulos, G. D., Paskind, M., Thomas, E., Boss, M. A., Landau, N., Alt, F. W., and Baltimore, D. (1984) Insertion of N regions into heavy-chain genes is correlated with expression of terminal deoxytransferase in B cells. *Nature.* 311, 752-755.

7. Moon, A. F., Garcia-Diaz, M., Batra, V. K., Beard, W. A., Bebenek, K., Kunkel, T. A., Wilson, S. H., and Pedersen, L. C. (2007) The X family portrait: Structural insights into biological functions of X family polymerases. *DNA Repair (Amst).* 6, 1709-1725.

8. Tonegawa, S. (1983) Somatic generation of antibody diversity. *Nature.* 302, 575-581.

9. Lieber, M. R., Lu, H., Gu, J., and Schwarz, K. (2008) Flexibility in the order of action and in the enzymology of the nuclease, polymerases, and ligase of vertebrate non-homologous DNA end joining: Relevance to cancer, aging, and the immune system. *Cell Res.* 18, 125-133.

10. Nick McElhinny, S. A., Havener, J. M., Garcia-Diaz, M., Juarez, R., Bebenek, K., Kee, B. L., Blanco, L., Kunkel, T. A., and Ramsden, D. A. (2005) A gradient of template dependence defines distinct biological roles for family X polymerases in nonhomologous end joining. *Mol. Cell.* 19, 357-366.

11. Bertocci, B., De, S. A., Berek, C., Weill, J. C., and Reynaud, C. A. (2003) Immunoglobulin kappa light chain gene rearrangement is impaired in mice deficient for DNA polymerase mu. *Immunity.* 19, 203-211.

12. Andrade, P., Martin, M. J., Juarez, R., Lopez de, S. F., and Blanco, L. (2009) Limited terminal transferase in human DNA polymerase mu defines the required balance between accuracy and efficiency in NHEJ. *Proc. Natl. Acad. Sci. U S A.* 106, 16203-16208.

13. Rolfsmeier, M. L., Laughery, M. F., and Haseltine, C. A. (2010) Repair of DNA double-strand breaks following UV damage in three *Sulfolobus solfataricus* strains. *J. Bacteriol.* 192, 4954-4962.

14. Zhang, L., Lou, H., Guo, L., Zhan, Z., Duan, Z., Guo, X., and Huang, L. (2010) Accurate DNA synthesis by *Sulfolobus solfataricus* DNA polymerase B1 at high temperature. Extremophiles. 14, 107-117.

15. Grogan, D. W. (2000) The question of DNA repair in hyperthermophilic archaea. *Trends Microbiol.* 8, 180-185.

16. White, M. F. (2011) Homologous recombination in the archaea: the means justify the ends. *Biochem. Soc. Trans.* 39, 15-19.

17. Lao-Sirieix, S. H., and Bell, S. D. (2004) The heterodimeric primase of the hyperthermophilic archaeon *Sulfolobus solfataricus* possesses DNA and RNA primase, polymerase and 3'-terminal nucleotidyl transferase activities. *J. Mol. Biol.* 344, 1251-1263.

18. Augustin, M. A., Huber, R., and Kaiser, J. T. (2001) Crystal structure of a DNA-dependent RNA polymerase (DNA primase). *Nat. Struct. Biol.* 8, 57-61.

19. Ito, N., Nureki, O., Shirouzu, M., Yokoyama, S., and Hanaoka, F. (2003) Crystal structure of the Pyrococcus horikoshii DNA primase-UTP complex: implications for the mechanism of primer synthesis. *Genes Cells.* 8, 913-923.

20. Grabowski, B., and Kelman, Z. (2003) Archeal DNA replication: Eukaryal proteins in a bacterial context. *Annu. Rev. Microbiol.* 57, 487-516.

21. Barry, E. R., and Bell, S. D. (2006) DNA replication in the archaea. *Microbiol. Mol. Biol. Rev.* 70, 876-887.

22. Zuo, Z., Rodgers, C. J., Mikheikin, A. L., and Trakselis, M. A. (2010) Characterization of a functional DnaG-type primase in archaea: Implications for a dual-primase system. *J. Mol. Biol.* 397, 664-676.

23. Yuzhakov, A., Kelman, Z., and O'Donnell, M. (1999) Trading places on DNA--a three-point switch underlies primer handoff from primase to the replicative DNA polymerase. *Cell.* 96, 153-163.

24. Nelson, S. W., Kumar, R., and Benkovic, S. J. (2008) RNA primer handoff in bacteriophage T4 DNA replication: The role of single-stranded DNA binding protein and polymerase accessory proteins. *J. Biol. Chem.* 283, 33, 22838-46

25. Chen, L., and Huang, L. (2006) Oligonucleotide cleavage and rejoining by topoisomerase III from the hyperthermophilic archaeon *Sulfolobus solfataricus*: Temperature

dependence and strand annealing-promoted DNA religation. *Mol. Microbiol.* 60, 783-794.

26. De Felice, M., Aria, V., Esposito, L., De Falco, M., Pucci, B., Rossi, M., and Pisani, F. M. (2007) A novel DNA helicase with strand-annealing activity from the crenarchaeon *Sulfolobus solfataricus*. *Biochem J.* 408, 87-95.

27. Brown, J. A., and Suo, Z. (2009) Elucidating the kinetic mechanism of DNA polymerization catalyzed by *Sulfolobus solfataricus* P2 DNA polymerase B1. *Biochemistry.* 48, 7502-7511.

28. Savino, C., Federici, L., Johnson, K. A., Vallone, B., Nastopoulos, V., Rossi, M., Pisani, F. M., and Tsernoglou, D. (2004) Insights into DNA replication: the crystal structure of DNA polymerase B1 from the archaeon *Sulfolobus solfataricus*. *Structure.* 12, 2001-2008.

29. Pisani, F. M., De, F. M., Manco, G., and Rossi, M. (1998) Domain organization and biochemical features of Sulfolobus solfataricus DNA polymerase. *Extremophiles.* 2, 171-177.

30. Mikheikin, A. L., Lin, H. K., Mehta, P., Jen-Jacobson, L., and Trakselis, M. A. (2009) A trimeric DNA polymerase complex increases the native replication processivity. *Nucleic Acids Res.* 37, 7194-7205.

31. Clark, J. M., Joyce, C. M., and Beardsley, G. P. (1987) Novel blunt-end addition reactions catalyzed by DNA polymerase I of *Escherichia coli*. *J. Mol. Biol.* 198, 123-127.

32. Clark, J. M. (1988) Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. *Nucleic Acids Res.* 16, 9677-9686.

33. Golinelli, M. P., and Hughes, S. H. (2002) Nontemplated base addition by HIV-1 RT can induce nonspecific strand transfer in vitro. *Virology.* 294, 122-134.

34. Delarue, M., Boule, J. B., Lescar, J., Expert-Bezancon, N., Jourdan, N., Sukumar, N., Rougeon, F., and Papanicolaou, C. (2002) Crystal structures of a template-independent DNA polymerase: murine terminal deoxynucleotidyltransferase. *EMBO.* 21, 427-439.

35. Dominguez, O., Ruiz, J. F., Lain de, L. T., Garcia-Diaz, M., Gonzalez, M. A., Kirchhoff, T., Martinez, A., Bernad, A., and Blanco, L. (2000) DNA polymerase mu (Pol mu), homologous to TdT, could act as a DNA mutator in eukaryotic cells. *EMBO.* 19, 1731-1742.

36. Maga, G., Ramadan, K., Locatelli, G. A., Shevelev, I., Spadari, S., and Hubscher, U. (2005) DNA elongation by the human DNA polymerase lambda polymerase and terminal transferase activities are differentially coordinated by proliferating cell nuclear antigen and replication protein A. *J. Biol. Chem.* 280, 1971-1981.

37. Pisani, F. M., Manco, G., Carratore, V., and Rossi, M. (1996) Domain organization and DNA-induced conformational changes of an archaeal family B DNA polymerase. *Biochemistry.* 35, 9158-9166.

38. Pisani, F. M., De, F. M., and Rossi, M. (1998) Amino acid residues involved in determining the processivity of the 3'-5' exonuclease activity in a family B DNA polymerase from the thermoacidophilic archaeon *Sulfolobus solfataricus*. *Biochemistry.* 37, 15005-15012.

39. Dionne, I., Nookala, R. K., Jackson, S. P., Doherty, A. J., and Bell, S. D. (2003) A heterotrimeric PCNA in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Mol. Cell.* 11, 275-282.

40. Frick, D. N., and Richardson, C. C. (2001) DNA primases. *Annu. Rev. Biochem.* 70, 39-80.

41. Kuchta, R. D., and Stengel, G. (2009) Mechanism and evolution of DNA primases. *Biochim. Biophys. Acta*. 1804, 5, 1180-9

42. Schlotterer, C., and Tautz, D. (1992) Slippage synthesis of simple sequence DNA, *Nucleic Acids Res.* 20, 211-215.

43. Shinde, D., Lai, Y., Sun, F., and Arnheim, N. (2003) Taq DNA polymerase slippage mutation rates measured by PCR and quasi-likelihood analysis: (CA/GT)n and (A/T)n microsatellites. *Nucleic Acids Res.* 31, 974-980.

44. Kuhner, F., Morfill, J., Neher, R. A., Blank, K., and Gaub, H. E. (2007) Force-induced DNA slippage, *Biophys. J.* 92, 2491-2497.

45. Richard, G. F., Kerrest, A., and Dujon, B. (2008) Comparative genomics and molecular dynamics of DNA repeats in eukaryotes. *Microbiol. Mol. Biol. Rev.* 72, 686-727.

46. Kato, K. I., Goncalves, J. M., Houts, G. E., and Bollum, F. J. (1967) Deoxynucleotide-polymerizing enzymes of calf thymus gland. II. Properties of the terminal deoxynucleotidyltransferase, *J. Biol. Chem.* 242, 2780-2789.

47. Neidle, S., and Parkinson, G. N. . (2003) The structure of telomeric DNA. *Curr. Opin. in Structl. Bio.* 13, 275-283.

48. Zhang, L., Brown, J. A., Newmister, S. A., and Suo, Z. (2009) Polymerization Fidelity of a Replicative DNA Polymerase from the Hyperthermophilic Archaeon *Sulfolobus solfataricus* P2. *Biochemistry*. 48, 31, 7492-501

49. Boule, J. B., Rougeon, F., and Papanicolaou, C. (2001) Terminal deoxynucleotidyl transferase indiscriminately incorporates ribonucleotides and deoxyribonucleotides, *J. Biol. Chem.* 276, 31388-31393.

50. Hopkins, B. B., and Paull, T. T. (2008) The P. furiosus mre11/rad50 complex promotes 5' strand resection at a DNA double-strand break. *Cell.* 135, 250-260.

51. Nair, P. A., Smith, P., and Shuman, S. (2010) Structure of bacterial LigD 3'-phosphoesterase unveils a DNA repair superfamily. *Proc. Natl. Acad. Sci. U S A.* 107, 12822-12827.

4.0 PRIMER HANDOFF IN *SULFOLOBUS SOLFATARICUS* TO INITIATE DNA REPLICATION

The primary biological function of the primase is to synthesize RNA primers for the DNA polymerase to elongate into ssDNA. Primases are error-prone RNA polymerases. In humans, the polα/primase complex misincorporates NTPs at a frequency of 0.01 errors per nucleotide polymerized(*1*). Therefore, the RNA primers in prokaryotes or RNA-DNA hybrid products synthesized by polα complex in eukaryotes must be transferred to a replicative DNA polymerase with high fidelity in order to maintain the accurate DNA replication. Primer transfer at high temperatures, like that in hyperthermophilic archaea, has significant thermodynamic challenges as maintaining small RNA primers annealed with the DNA template is difficult. Efficient primer transfer often relies on help from other proteins to remain annealed (*2*). Therefore, it is critical to reveal other replicative factors in the primer transfer process in archaea in order to understand how the accurate DNA replication initiates in the hyperthomophiles.

Primer transfer mechanisms have been well studied in both the prokaryotic domain and the eukaryotic domain(*3*). For example, the T7 primase associates with the polymerase to directly transfer the 4-nt RNA primers between the enzymes for elongation(*4*). In *E. coli*, the primer transfer mechanism is much more complicated. O'Donnel *et al* revealed a three-point switch mechanism for primer transfer that requires two other cofactors: Pol III's processivity

factor β (β sliding clamp) and γ clamp-loading complex in addition to the primase and the replicative DNA polymerase(*5*). After the RNA primer is synthesized, the *E. coli* DnaG primase remains attached on the primed template by interacting with single strand binding protein (SSB). Then the γ clamp-loading complex makes contacts with SSB and disrupts the primase-SSB interaction, resulting in primase dissociation from the primed template. Finally, the γ complex loads β sliding clamp onto the primed template and then DNA Pol III associates with β sliding clamp to form a processive DNA polymerase.

In eukaryotes, primer transfer is very unique in that two handoff events are involved. First, the eukaryotic primases directly transfer the synthesized RNA primers to polymerase α subunit without dissociating from the template because these subunits directly interact to form the pri α-primosome complex (*6*). In humans, Pol α elongates the RNA primer up to 3000 nucleotides long to form RNA/DNA pre-Okazaki fragments (*7*). Second, these hybrid fragments are transferred from the polα/primase complex to the replicative DNA polymerases, Pol δ or Pol ε, for elongation into the daughter strands (*8*). The second handoff process of the RNA/DNA hybrid products is also mediated by single stranded binding protein (RPA in eukaryotes), the clamp-loading complex and the clamp complex like that in *E coli*. The clamp loader in eukaryotes, RFC, interacts with RPA to release Pol α complex and recruits the clamp complex, PCNA, to site (4, *9*).

In contrast, the primer transfer mechanism in archaea is not well studied. Successful primer transfer requires that the RNA primers stay associated with the template before the DNA polymerase is recruited. It is apparent that because archaea grow at high temperatures the thermodynamic processes hinder direct primer transfer and will require stabilization through other DNA replication accessory proteins. We hypothesize that these thermophilic organisms

have evolved efficient primer transfer mechanisms that require the action of other replication proteins to stabilize this process. It has been speculated that the DNA replication in *P. furiosus* is initiated from RNA primers like the organisms in other domains, because *Pfu*PCNA and *Pfu*RFC can stimulate both *Pfu*Dpo1 and *Pfu*DpoD to elongate the RNA products synthesized by *Pfu*PriS&L on the single-stranded M13 template (*10*). There are still many unanswered questions about the molecular mechanism of primer transfer in archaea, such as how and when the primase dissociates from the primed template, what is the role of single stranded binding protein as a mediator in this process, and if the stimulation by PCNA and RFC applies to other species of archaea.

In this chapter, preliminary results for primer transfer from both PriS&L and DnaG to Dpo1 are described in *Sso*. We have found that *Sso*Dpo1 can elongate synthetic DNA primers on both circular and linear templates. The elongation of *de novo* DNA primers synthesized by *Sso*PriS&L on the circular template was also detected. However, *Sso*Dpo1 was found only to elongate synthetic RNA primers on a linear template but not on a circular template. Extension of *de novo* RNA primers synthesized by either *Sso*PriS&L or *Sso*DnaG on the circular template was not detected in any case. *Sso*PCNA, *Sso*RFC, and *Sso*SSB were found to be able to stimulate the elongation DNA primers. These results indicate that the mechanism of primer transfer in organisms in the third domain of life might be more complex than that as suggested above for other domains. We hypothesize that *Sso*RFC, *Sso*PCNA, and *Sso*SSB are required for efficient primer handoff to *Sso*Dpo1 for RNA primers synthesized by *Sso*DnaG or RNA-DNA hybrid primers synthesized by *Sso*PriS&L.

4.1 MATERIALS AND METHODS

Materials

^{32}P-ATP, ^{32}P-GTP and ^{32}P-dATP were purchased from MP Biomedicals (Solon, Ohio). Unlabeled deoxyribonucleotides and ribonucleotides were purchased from Invitrogen (Carlsbad, CA). All single stranded DNA (ssDNA) and RNA were synthesized by IDT (Coralville, IA, Table 4-1) and gel purified as described previously (*11*). Optikinase was purchased from USB (Cleveland, OH). TdT is from NEB (Boston, MA). Exonuclease deficient (exo$^-$) Dpo1 (D231A/D318A), DnaG and Pri S&L were purified as described previously (*2, 11, 12*). RFC and PCNA are generously provided by Hsiang-kai Lin. All other reagents are analytical grade or better.

Table 4-1. DNA template and synthetic primers

Oligonucleotides	Nucleotide Sequence	T_M (°C)
MAT15	5'-CACCTCTCCCTACGCTTCCCACCCACCCCGACCG GCATCTGCTATGGTACGCTGAGCGAGAGTAGC	NC[1]
MAT15 DNA primer	5'-GCTACTCTCGCTCAGCGTACCATAGCAG	64
MAT15 RNA primer	5'-GCTACTCTCGCTCAGrCrGrUrArCrCrArUrArGrCrArG	62
M13 DNA primer	5'- CCGGAAACCAGGCAAAGCGCCATTCG	66
M13 RNA primer	5'-CCGGAAACCAGGrCrArArArGrCrGrCrCrArUrUrCrG	61

[1] Not Calculated.

Synthetic primer elongation by Dpo1

In 10 µl reactions, 1.0 µM Dpo1 (exo$^-$) was incubated with approximately 0.2 µM ^{32}P 5'-end labeled synthetic primers, 150 µM of dNTPs, and 60 nM M13 template or 200 nM MAT23 template in elongation buffer (100 mM HEPES (pH7.5), 150 mM K(CH_3COO) and 50 mM $Mg(CH_3COO)_2$) in 10 µl reaction volumes at 60 °C for 5 minutes for M13 template or 30 seconds for MAT15 template. 1.0 µM of RFC and PCNA were included in reactions as specified in figures. ^{32}P 5'-end labeling was performed using Optikinase according to manufacturer's directions. The template DNA and primers was heated at 95 °C for 5 min and then cooled down to room temperature for primer annealing in a 10- µl reaction in elongation buffers prior addition of the rest substrates. The reaction was stopped by adding equal volume of the quench solution (88% formamide, 10 mM EDTA, 1 mg/ml bromophenol blue, 0.1% SDS). The quenched reaction samples were separated on a 20% denaturing polyacrylamide gel or a 0.8% alkaline agarose gel, imaged overnight, and scanned using a phosphorimager (Storm 820, GE Healthsciences).

Two-step primase priming and Dpo1 elongating assay

In the first step of priming reaction, PriS&L priming reactions were conducted as described previously (*13*) and DnaG priming reaction were performed as in Chapter 2 except that these cold reactions didn't include the radioactive nucleotides. After the priming reaction was completed, the reaction solution was heated up to 95 °C, kept at 95 °C for 5 min, and then cooled down to room temperature. In the second elongation step, Dpo1, RFC, PCNA and dNTPs were mixed with 3.7 µl priming reaction solution at the same concentrations as for the synthetic primer elongation assays in a 10 µl reaction volume under the same conditions. In hot priming reaction, 200 ng of α-^{32}P-GTP was included for both PriS&L and DnaG to label the primer

products, while 200 ng of α-^{32}P-dATP was used in elongation step to label the elongation products if the cold priming reaction solution was used. The elongation reactions were then stopped and analyzed as above.

Coupled priming/elongating assays

In a typical 25 µl reaction, 1.5 µM of the primase (Pri S&L or DnaG) was incubated with 500 ng of M13, 1 µM of dNTPs and/or NTPs, and different combinations of 1 µM of Dpo1, 1 µM of RFC and 1 µM of PCNA for 2 hours at 60 °C for Pri S&L or at 70 °C for DnaG. The reaction was stopped and analyzed as above.

5' and 3' end labeling for the priming reaction products

^{32}P 5'-end labeling was performed using Optikinase (USB, Cleveland, OH). according to manufacturer's directions. Briefly, 6 µl of cold priming solution was mixed with 1 µl of γ-^{32}P-ATP, 1 µl of 10x Opti buffer and 1µl of Opti kinase. ^{32}P 3' end labeling was performed using TdT. In a typical 25 µl reaction, 10 µl of cold priming solution was incubated with 2.5 µl of terminal deoxytransferase (TdT) (NEB, Ipswitch, MA), 2.5 µl of 10x TdT buffer and 2.5 µl of $CoCl_2$ for 20 min at 37 °C for labeling reaction and then 70 °C for 30 min to deactivate the enzyme.

4.2 RESULTS

4.2.1 Elongation of synthetic primers

In order to test the primer elongating activity of Dpo1, we designed two synthetic DNA primers and two synthetic RNA primers of the same length as listed in Materials and Methods. The clamp loading complex (RFC in *Sso*) and the clamp complex (PCNA in *Sso*) was shown to stimulate the primer elongation in both prokaryotic and eukaryotic domains (*10, 11, 12*). The stimulating effect of both enzymes was also observed in Pfu (*11*). Therefore, the elongation reaction was tested in a combination of Dpo1, RFC, and PCNA on a circular single stranded DNA template, M13, and a linear 66 base DNA template, MAT15. Dpo1 alone can efficiently extend synthetic DNA primers on the M13 template (Figure 4-1A). The stimulating effect of RFC and PCNA was not clear due to the large size of the elongation products that stuck in the wells of the 20% acrylamide gel used (Figure 4-1A), however an experiment using an 0.8% agarose gel shows that PCNA can stimulate the DNA elongation of DNA primers by Dpo1 on M13 template (Figure 4-2A).

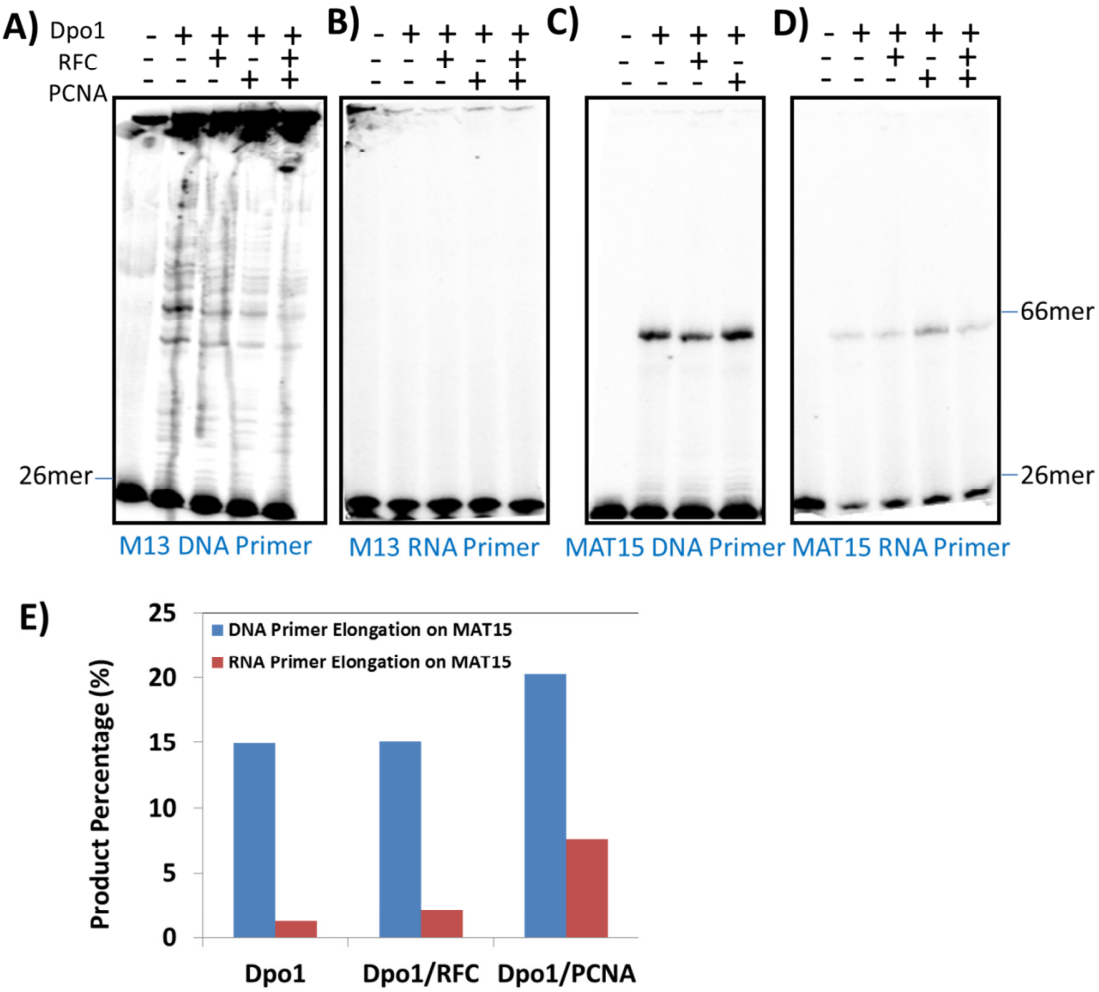

Figure 4-1. Elongation of synthetic primers by Dpo1

5'end ^{32}P labeled synthetic primers were studied on circular single stranded M13 DNA and linear synthetic 66mer MAT15 DNA. Different synthetic primers as labeled and M13 (A & B) or MAT15 (C &D) were incubated with different combinations of Dpo1, RFC and PCNA as described in Materials and Methods. The primer size is 26mer as labeled. 1 mM of ATP was included when RFC was used. The bands of 66mer products in C) (■) and D) (■) were quantified and normalized to the total radiation in each lane (E).

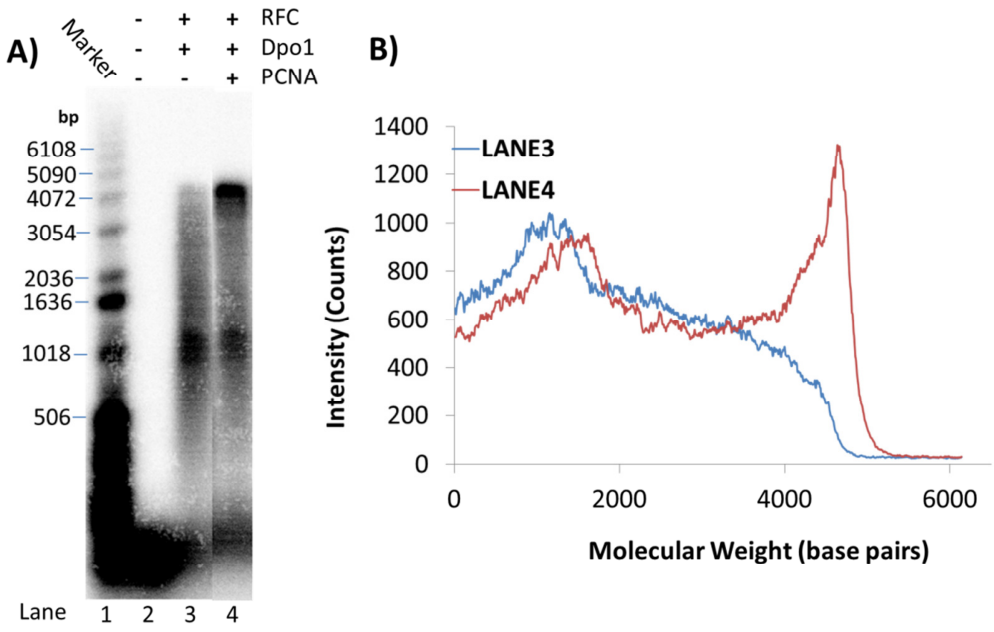

Figure 4-2. Stimulating effect of PCNA on DNA elongation on the M13 template

A) Shows the 0.8% agarose gel image of the stimulatory effect on DNA elongation for RFC and PCNA. The reactions were conducted and analyzed on a 0.8% alkaline agarose gel as described in Method and Materials. 1 kb oligomer ladder were labeled at 5' end with ^{32}P and used as standards (lane1). The protein combinations are indicated on the top the gel. The product intensity for lane 3 and lane 4 were quantified and calibrated against the standards in lane 1 using ImageQuant software. The intensity was plotted against the molecular weight (B).

The quantification of the product distribution as a function of DNA length for lanes 3 and 4 in Figure 4-2A show that the intensity of the longest elongation product dramatically increases when PCNA is included (Figure 4-2B). Therefore the addition of PCNA increases the elongation rate of Dpo1 on an M13 template. Extension on linear DNA templates also shows that PCNA increases the elongation of DNA primers by 30% while inclusion of RFC alone doesn't stimulate the elongation (Figure 4-1C & E). On the other hand, elongation of RNA primers shows conflicting results on the two templates. When the M13 template was used, extension of the RNA primer was not visibly or quantitatively detected (Figure 4-1B). On the other hand, using a

linear MAT15 template (66 bases), the stimulating effect of both RFC and PCNA is obviously seen (Figure 4-1D & E). Quantification of the results shows a one-fold and six-fold stimulation for RFC or PCNA, respectively (Figure 4-1 E). Based upon these results, we speculate that both RFC and PCNA can stimulate Dpo1 to elongate both DNA and RNA primers, though further experiments are needed to clarify this effect on RNA primer elongation on the circular M13. Possibly with longer templates inclusion of SSB is required to organize the DNA for efficient RNA primer extension.

4.2.2 De novo primer transfer from PriS&L to Dpo1

PriS&L was shown to synthesize both RNA products (up to 1 kb) and DNA products (up to 7 kb) *in vitro* (*13*). Similarly, the hot priming reaction by PriS&L (when ^{32}P NTP is included) in our experiments generated long RNA products that are stuck in the gel well (Figure 4-3A lane 3). The RNA primer products synthesized by PriS&L in cold priming reactions were successfully identified by ^{32}P 5' end labeling (Figure 4-3A lane 1). The 5' end labeling for the cold priming reaction shows similar size of primers but different intensity distribution compared with products of the hot priming reaction. This variation in primer distribution is not clearly known but can be attributed to preferential labeling of shorter RNA molecules. Alternatively, short RNA primers may not have ^{32}P-ATP within the sequence precluding labeling in the hot reaction. For successful primer elongation by the DNA polymerase, a free 3' hydroxyl group at the end of the primer is required. Therefore, the 3'-end labeling of the primer was conducted to verify the availability of the 3' hydroxyl group by a terminal transferase (TdT) reaction. TdT is known to transfer 1-3 nucleotides under optimal conditions to the 3' end of single stranded DNA

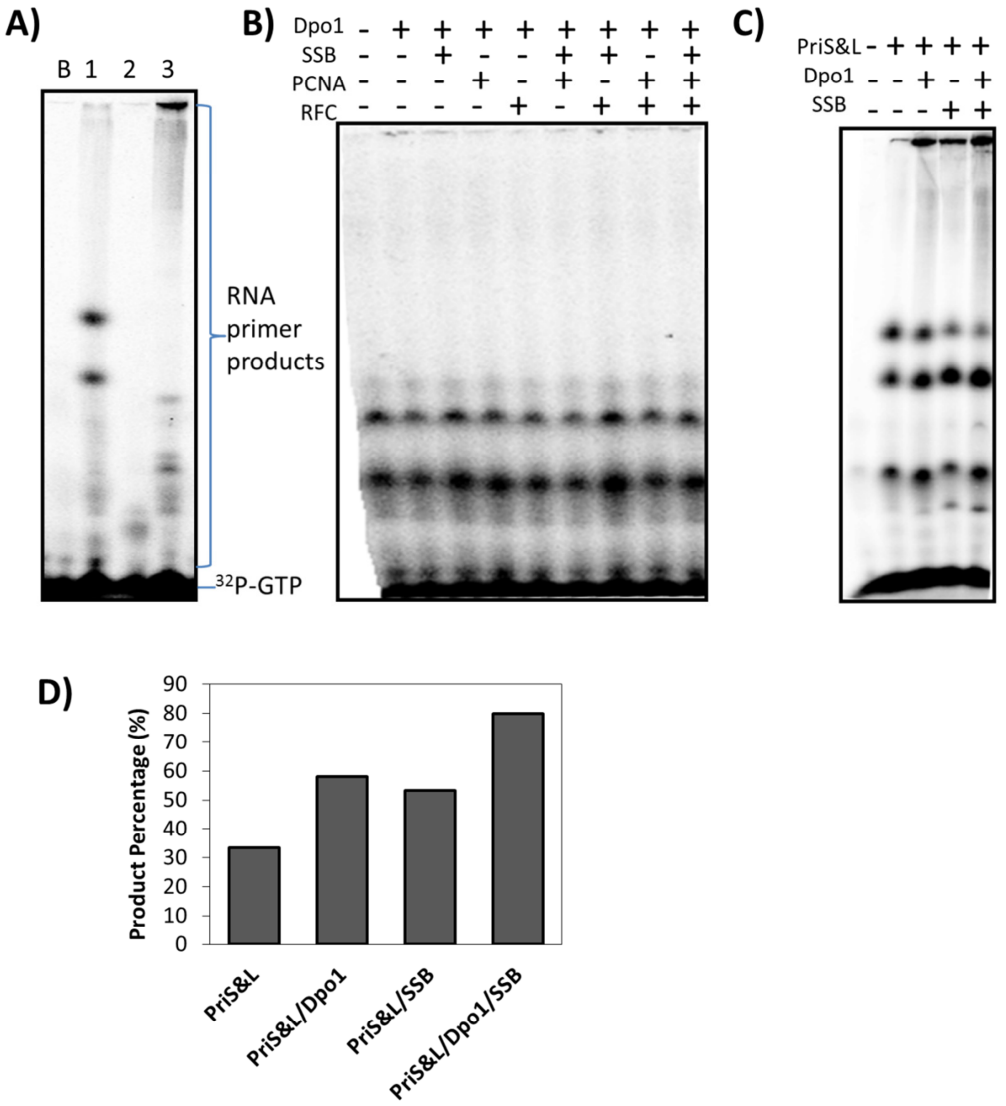

Figure 4-3. *De novo* primer transfer from Pri S&L to Dpo1

A) Shows the RNA products synthesized by PriS&L using the M13 template as described in Materials and Methods. Cold priming reaction solutions were radioactively labeled at both 5' end (lane 1) and 3' end (lane 2) in comparison with hot priming reaction (lane 3). Primer products were identified by compared with blank reaction (lane B). B) Two-step method as described in Materials and Methods was utilized to study the elongation of Pri S&L RNA primer by Dpo1 in combination with SSB, PCNA, RFC as labeled. Cold priming reaction was performed using M13 template and NTPs without radioactive labels. The elongation reaction included dNTPs and α-^{32}P-dATP. C) Coupled priming/elongating method was used to study the elongation of Pri S&L DNA primer in combination with Dpo1 and SSB. dNTPs and α-^{32}P-dATP were used as substrates. D) Products in each lane in C) were quantified and normalized to total radiation in each lane as labeled. The background percentage was subtracted from product percentage of the other four lanes.

molecules which must have a free hydroxyl group there (*12*). Unfortunately, the 3' end labeling of the RNA primer didn't efficiently label the RNA primer products (Figure 4-3A lane 2).However, this deficiency of 3' labeling does not necessarily mean unavailability of the 3' hydroxyl group of the RNA primers but could be the result of inefficiency of transferase activity on RNA substrates. Previously, it was shown that two ribonucleotide residues at the 3' end of the DNA initiator inhibit the elongation by TdT (*15*). Thus, further experiments are required to verify the availability of the hydroxyl group at the 3' end as discussed later.

The *de novo* RNA primer transfer from PriS&L to Dpo1 was conducted in a two-step handoff experiment. In the first step, the RNA primers synthesized by PriS&L in a cold priming reaction were used as the primer substrates for the elongation reaction by Dpo1 as the second step. In the elongating reaction, the α-^{32}P-dATP was utilized to radioactively label elongation products by Dpo1. Surprisingly, the RNA products synthesized by PriS&L as identified by the 5' end labeling in Figure 4-2A were not elongated by Dpo1 even in the presence of RFC, PCNA and SSB, as only the background bands were observed (Figure 4-3B). For the *de novo* DNA primer transfer, a coupled priming/elongating strategy was employed. Briefly, dNTPs and α-^{32}P-dATP were incubated in one reaction with all enzymes as labeled in Figure 4-3C. Visibly, there are more products stuck in the gel wells when Dpo1 was added. The quantification of the product bands shows that the reactions including Dpo1 generate ~20% more products than the reactions that have PriS&L but not Dpo1 (Figure 4-2D). This indicates that Dpo1 can either elongate the DNA primers synthesized by PriS&L or increase the elongation activity of PriSL directly. The addition of SSB into the reaction can stimulate both the priming reaction and the elongation reaction (Figure 4-3D). Stimulation by SSB indicates that SSB may also play a role in the primer transfer in archaea similar to that in both prokaryotes and eukaryotes. Further experiments

including PCNA and RFC in the reactions are necessary to reveal the stimulating effect of both enzymes on the primer transfer and elongation.

4.2.3 De novo Primer handoff from DnaG to Dpo1

DnaG only synthesizes RNA primers as shown in Chapter 2 and Figure 4-4A. Both the hot priming reaction solution and the 5' labeling of the cold priming reaction solution identified the RNA primer products (Figure 4-4A). Again, the 3' end labeling by TdT was not successful in labeling the RNA. Both two-step and one-step strategies were also utilized to study the RNA primer transfer from DnaG to Dpo1. No primer elongation by Dpo1 in combination with RFC, PCNA, and SSB was observed by the two-step strategy (Figure 4-4B). Interestingly, when both DnaG and Dpo1 were included in the one-step reaction, there is a smear of bands on the gel which may suggest Dpo1 elongated the primer products (Figure 4-4C). Unfortunately, these longer products were later proved to be the consequence of the TdT activity of Dpo1 (Chapter 3).

After altering the protein purification protocol for DnaG, contaminating DNA fragments from *E coli* cells were introduced as impurities into the experiments shown in Figure 4-4C. Control experiments in the absence of NTPs showed that Dpo1 could extend these contaminating DNA fragments and not the RNA primers. Actually, this is how we discovered the TdT activity of Dpo1 in Chapter 3. Future experiments will require more pure DnaG to assess the ability of Dpo1 to extend these de novo synthesized RNA primers.

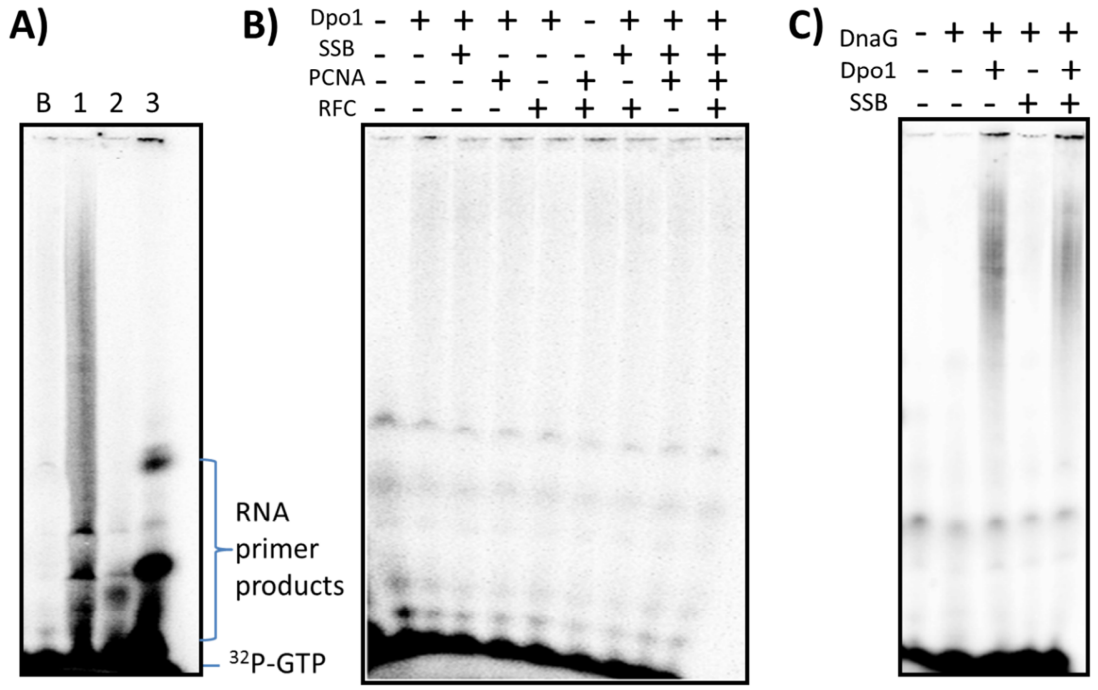

Figure 4-4. *De novo* RNA primer transfer from DnaG to Dpo1

A) Shows the RNA products synthesized by DnaG as described in Materials and Methods. Cold priming reaction solutions were labeled at both 5' end (lane 1) and 3' end (lane 2) in comparison with hot priming reaction (lane 3). Primer products were identified by compared with blank reaction (lane B). B) Two-step method as described in Materials and Methods was utilized to study the elongation of DnaG RNA primer by Dpo1 in combination with SSB, PCNA, RFC as labeled. C) Coupling priming/elongating method was used to study the elongation of DnaG RNA primer in combination with Dpo1 and SSB. Both NTPs and dNTPs were included as substrates with α-^{32}P-dATP as the probe.

4.3 DISCUSSION

The archaeal PriS&L primase is able to synthesize both DNA and RNA primers up to several thousand nucleotides long *in vitro*. Recently, we have shown *Sso*DnaG is active *in vitro* as well to synthesize RNA primers similar to *E coli* DnaG (Chapter 2). Questions about which DNA primase is functional in archaeal cells arise for this dual functional cellular primase system. DNA primases are required for DNA replication initiation on the leading strand and for each Okazaki fragment on the lagging strand. Characterization of the protein interactions with each primase as well as the primer handoff mechanism will be required to identify their individual roles. Therefore, primer transfer in archaea may be more complex than that in eukaryotic domain or prokaryotic domain.

In addition, Archaeal organisms contain three B-family DNA polymerases which all have exonuclease proof-reading domains and can act as replication polymerases. The B-family *Sso*Dpo1 is widely believed to be the replicative DNA polymerase in archaea due to its greater rate, fidelity, and processivity (16, *17*), however the molecular roles of the other two B-family polymerases (*Sso*Dpo2 and *Sso*Dpo3) are not clear currently. These multiple DNA polymerases also with high fidelity will also contribute to the complexity of the primer transfer in archaea.

In prokaryotes, the replicative polymerase can extend the RNA primers as shown in *E. coli*, T4 and T7 systems (*4, 18*). In contrast, the replicative DNA polymerase in eukaryotes elongates hybrid RNA-DNA pre-Okazaki fragments synthesized by the polα complex from the 3' DNA end. Our experiments show that *Sso*Dpo1 was not able to elongate RNA primers on the single stranded circular M13 template in contrast that DNA primers can be extended efficiently on M13 (Figure 4-1A & B). Even though *Sso*Dpo1 can extend the RNA primers on a small linear synthetic DNA template, the extension of DNA primers is more efficient than that of RNA

primers. Therefore and maybe not surprisingly, *Sso*Dpo1 prefers to extend DNA primers. The *de novo* primer transfer experiments also support this conclusion. *Sso*Dpo1 can elongate DNA primers synthesized by *Sso*PriS&L (Figure 4-3C), but can't elongate RNA primers synthesized by either primase (Figure 4-3B & 4-4B). This substrate discrimination by the replicative polymerase in *Sso* is indicative that the primer transfer mechanism in archaea is possibly homologous with that in eukaryotes. The primers that can be elongated by *Sso*Dpo1 can be RNA/DNA hybrids with the DNA fragment at the 3' end like those synthesized by Pol α complex in eukaryotes. Even though Archaea lack a homolog of the Pol α subunit which incorporates dNTPs to the 3' end of the RNA primer, archaeal PriS&L has the capabilities to synthesize both RNA and DNA products (*12*). This dual function of this primase can play the roles of Pol α complex in eukaryotes to synthesize RNA/DNA hybrid products. The evidence to support this mechanism homology with eukaryotes also comes from the studies of protein-protein interaction in archaea. *Sso*RFC was shown to interact with *Sso*PriS&L and inhibit the priming activity of the primase (*19*). The interaction between the clamp loader and the primase was believed to important for the primase dissociation from the primed site in eukaryotes *(4,19-21*). Further supportive evidence includes the stimulation of the processivity of *Sso*Dpo1 by *Sso*PCNA (*11*) and stimulation of *Sso*SSB on elongation as observed in Figure 4-3C. Therefore, we hypothesize that *Sso*PriS&L creates a hybrid RNA/DNA primer that is transferred to *Sso*Dpo1 through the concerted action of *Sso*RFC and *Sso*PCNA. Future experiments should be examined for the effect of *Sso*SSB on *Sso*RFC inhibition of the primase activity and stimulation of *Sso*PCNA binding will be helpful to elucidate the exact primase/polymerase switch.

An available 3' hydroxyl group of the primer is required for successful primer transfer. But our experiments cannot dictate the availability of the 3' hydroxyl group of the RNA primers

synthesized by *Sso*PriS&L or *Sso*DnaG (Figure 4-2A and 4-3A) due to the possible low TdT efficiency. To identify this necessary hydroxyl group, TdT labeling of the synthetic RNA primer as used in Figure 4-1 can be conducted as a positive control experiment. If TdT can't label the synthetic RNA primers, T4 RNA ligase can be used instead. T4 RNA ligase incorporates a phosphate group at the 3' end of RNA molecules when the hydroxyl group is available (*22*).

The RNA primers synthesized by *Sso*DnaG were also not successfully elongated by *Sso*Dpo1 in our experiments. The priming activity of *Sso*DnaG that we characterized (Chapter 2) and the genomic conservation of the active residues strongly suggest that DnaG primase may be essential for DNA replication in archaeal cells. It's been shown that Okazaki fragments in archaea are similar to those in eukaryotes with a short RNA primer at the 5' end (*23*). The length of the RNA primer ≥10nt) detected in the Okazaki fragments is consistent with the RNA primers synthesized by *Sso*DnaG. Therefore, it may need other unknown factors to facilitate the RNA primer transfer in archaea. Further experiments are needed to elucidate the primer transfer mechanism in archaea and dissect the individual roles of each primase.

REFERENCES

1. Zhang SS, G. F. (1990) Accuracy of DNA primase. *J. Mol. Biol.* 216, 475-479.

2. Zuo Z., Rodgers CJ, Mikheikin AL, and Trakselis MA. (2010) Characterization of a functional DnaG-type primase in archaea: implications for a dual-primase system. *J. Mol. Biol.* 397, 664-676.

3. Frick DN, and Richardson CC. (2001) DNA primases. *Annu. Rev. Biochem.* 70, 39-80.

4. Kato MIT, Wagner G, and Ellenberger T. (2004) A Molecular Handoff between Bacteriophage T7 DNA Primase and T7 DNA Polymerase Initiates DNA Synthesis. *J Biol Chem.* 279, 30554-30562.

5. Yuzhakov AKZ, O'Donnell M. (1999) Trading places on DNA — a three-point switch underlies primer handoff from primase to the replicative DNA polymerase. *Cell.* 96, 153-163.

6. Kuchta RD, and Stengel G. (2010) Mechanism and evolution of DNA primases, *Biochim. Biophys. Acta*. 1804, 5, 1180-1189

7. Wang TS, Korn D. (1984) DNA primase from KB cells. Characterization of a primase activity tightly associated with immunoaffinity-purified DNA polymerase-alpha. *J Biol Chem.* 259, 1854-1865.

8. Baker SP, Bell TA. (1998) Polymerases and the Replisome: Machines within Machines. *Cell.* 92, 295-305.

9. Tsurimoto T, Stillman B. (1991) Replication factors required for SV40 DNA replication in vitro. II. Switching of DNA polymerase a and d during initiation of leading and lagging strand synthesis. *J. Biol. Chem.* 266, 3, 1961-1968.

10. Liu L, Ishino S, Bocquier AA, Cann IK, Kohda D & Ishino Y. (2001) The archaeal DNA primase: Biochemical characterization of the p41-p46 complex from *Pyrococcus furiosus*. *J. Biol. Chem.* 276, 45484-45490.

11. Mikheikin AL, Lin HK, Mehta P, Jen-Jacobson L, and Trakselis MA. (2009) A trimeric DNA polymerase complex increases the native replication processivity. *Nucleic Acids Res.* 37, 7194-7205.

12. Lao-Sirieix SH, and 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, 1251-1263.

13. Yuzhakov A, Kelman Z, and O'Donnell M. (1999) Trading places on DNA--a three-point switch underlies primer handoff from primase to the replicative DNA polymerase. *Cell.* 96, 153-163.

14. Motea EA, and Berdis AJ. (2010) Terminal deoxynucleotidyl transferase: the story of a misguided DNA polymerase. *Biochim. Biophys. Acta.* 1804, 1151-1166.

15. Roychoudhury R. (1972) Enzymic synthesis of polynucleotides. Oligodeoxynucleotides with one 3′-terminal ribonucleotide as primers for polydeoxynucleotide synthesis, *J. Biol. Chem. 247*, 3910-3917.

16. Choi JY, Pence MG, Wang J, Martin MV, Kim EJ, Folkmann LM, and Guengerich FP. (2011) Roles of the four DNA polymerases of the crenarchaeon *Sulfolobus solfataricus* and Accessory proteins in DNA replication. *J. of Biol. Chem.* 286, 31180-31193.

17. Zhang L, Brown JA, Suo Z. (2009) Polymerization fidelity of a replicative DNA polymerase from the hyperthermophilic archaeon *Sulfolobus solfataricus* P2. *Biochemistry*. 48, 31, 7492-501.

18. Nelson SW, Benkovic SJ. (2008) RNA primer handoff in bacteriophage T4 DNA replication: The role of single-stranded DNA binding protein and polymerase accessory proteins, *J. Biol. Chem.* 283, 22838-22846.

19. Wu KY, Guo X, Hu JC, Xiang XY and Huang L. (2007) Interplay between primase and replication factor C in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Mol. Microbio.* 63, 826-837.

20. Maga GS. M, Spadari S, and Hubscher U. (2000) DNA polymerase switching. I. Replication factor C displaces DNA polymerase alpha prior to PCNA loading. *J. Mol. Biol*. 791-801.

21. Mossi R, Ferrari E, and Hubscher U. (2000) DNA polymerase switching. II. Replication factor C abrogates primer synthesis by DNA polymerase alpha at a critical length. *J. Mol. Biol*. 803-814.

22. Viollet S, Munafo DB, Zhuang FL, and Robb GB. (2011) T4 RNA Ligase 2 truncated active site mutants: Improved tools for RNA analysis. *BMC Biotech.* 11, 14.

23. Matsunaga F, Forterre P, and Myllykallio H. (2003) Identification of short 'eukaryotic' Okazaki fragments synthesized from a prokaryotic replication origin. *EMBO.* 4, 154-158.

5.0 FUTURE WORK

5.1 MOLECULAR ROLE OF *SSO*DNAG IN DNA REPLICATION IN ARCHAEA

The DNA replication machinery in archaea is more closely related to eukaryotes than prokaryotes (*1-3*). Nearly all other identified DNA replication proteins in archaea except for the DnaG primase are eukaryotic-like (*4*). It is widely believed that PriS&L is the functional archaeal primase for DNA replication initiation and elongation in archaea, however, the priming activity of *Sso*DnaG characterized in Chapter 2 has generated questions about the enzymatic roles of each primase *in vivo*. Genetic knockout of the homologous PriS&L in halobacterium is lethal (*5*). We would expect that knockouts of *Sso*DnaG will also be lethal to archaeal cells. This hypothesis is well supported by some studies. The catalytic glutamate contained in the TOPRIM domain and two other proposed catalytic aspartates are well conserved though out the archaeal domain (Figure 2-1). This high level conservation is indicative that the archaeal DnaG primase is essential for the life in this domain. DnaG was identified in cells by proteomics analysis (*6*) showing that expression of this protein occurs in growing cells. *Sso*DnaG was also reported to be a component of the exosome, whose function is to degrade RNA molecules. This finding along with the priming activity we characterized suggests that *Sso*DnaG might be required in multiple essential roles in archaeal cells.

The DNA replication proteins conduct their activities in a way of collaboration and cooperation with multiple proteins to form the replication machinery. Interactions of *Sso*DnaG with other replication proteins will further reveal the molecular role of *Sso*DnaG. It has been recently detected that *Sso*DnaG interacts with cognate *Sso*MCM helicase in the Trakselis lab (data not shown). The direct interaction between the primase and the helicase indicates that *Sso*DnaG can be recruited to the replication fork after the helicase opens the double-stranded genome in archaea. Therefore, *Sso*DnaG is most likely participating at the replication fork similarly to the phage or bacterial helicase-primase complex (*7*) but different than in eukaryotes. To test this hypothesis, further experiments of protein interactions between *Sso*DnaG and other replication proteins and their effects on enzyme activities are needed. The clamp loader, clamp, and single-strand binding proteins play central roles in priming and primer transfer as shown in both eukaryotes and prokaryotes. Therefore, experiments to study interactions of *Sso*DnaG with *Sso*RFC, *Sso*SSB, and *Sso*MCM will definitely be critical to show the enzymatic role of *Sso*DnaG in DNA replication machinery in archaea.

5.2 TDT ACTIVITY OF DPO1 IN ARCHAEAL DNA REPAIR

In contrast to DNA replication, DNA repair in Archaea is poorly understood. This is perhaps surprising given that the challenges faced by hyperthemophilic organisms in preservation of the genome integrity as the growth at high temperatures would increase the DNA damages including double-strand breaks (DSB) and oxidative damage (*8*). Interestingly, it has been shown that the mutational frequencies in the archaeal genome are equal to or lower than those found in mesophilc organisms suggesting a highly efficient DNA repair system in archaea (*9*). Homology

searches suggest that archaeal organisms have similar repair proteins as eukaryotes (*1*), however, archaea may have simpler repair pathways as many of the repair proteins are absent comparing to those in eukaryotes (*1*). Homologous recombination (HR), not non-homologous end joining (NHEJ), was originally thought to be the preferred pathway to repair double-strand breaks (DSB) in archaea as most of the HR proteins were identified in archaea (*10, 11*). Recently, a core component of NHEJ repair pathways, LigD 3' phosphoesterase (PE), was identified and characterized in archaeal organisms (*12, 13*). Thus, NHEJ may be also conserved in archaea as a pathway to repair double strand breaks. Eukaryotic X-family DNA polymerases (including TdT, Pol μ, and Pol λ) have been shown to participate in non-homologous end joining (NHEJ) by extending ssDNA at double strand breaks as a template for annealing (*14*). The discovery of TdT activities for *Sso*PriS&L and *Sso*Dpo1 supports that NHEJ is an active DSB repair pathway in archaea. The role of TdT activities of these two enzymes in NHEJ repair of double strand breaks should be explored further.

5.3 PRIMER TRANSFER IN ARCHAEA

The puzzle of primer transfer in archaea still needs to be deciphered. Single-strand binding protein (SSB), clamp loader, and clamp were shown to be critical in the primer transfer in both prokaryotic and eukaryotic domains. The prevailing view of the primer transfer mechanism in archaea is that it is similar to that in eukaryotes where the RNA-DNA hybrid fragments synthesized by the polα complex are transferred to the replicative DNA polymerases. During this process, PriS&L in archaea can imitate the polα complex in eukaryotes and the single strand binding protein can act as the mediator that interacts with both primase and clamp loading

complex. This mechanism predicts that, in the presence of SSB in archaea, the RFC should inhibit the priming activity of the primase and stimulate the DNA synthesis of Dpo1 or other replicative DNA polymerases. *Sso*PriS&L was shown to synthesize both RNA and DNA primers *in vitro* but no RNA/DNA hybrid products were reported (*15*). It was shown that *Pfu*PriS&L can synthesize RNA/DNA fragments in dose dependence of dNTPs in the substrates (*16*). When equal concentrations of NTPs and dNTPs were present in the reaction, *Pfu*PriS&L synthesizes no RNA primers. The primer elongation assays indicate that the hybrid primer with ribonucleotides at the 5' end can be more efficiently elongated by *Pfu*Dpo1 than the DNA primers. Therefore, the synthesizing activity of RNA/DNA hybrid by *Sso*PriS&L can be studied first, followed by experiments of the effect of SSB on the primase activity. Finally, the DNA polymerase activity after priming can unravel the mystery of the primer transfer between the primase and the polymerase.

Our results show that *Sso*Dpo1 was not able to elongate RNA primers *de novo* but can elongate the synthetic RNA primers with a lower efficiency than the DNA primer. In this context, the sole RNA synthesizing ability of *Sso*DnaG appears to be less favored by *Sso*Dpo1 in primer transfer in *Sso*. However, the RNA primers synthesized by *Sso*DnaG are possibly transferred in an unknown mechanism or to a separate DNA polymerase through the action of accessory proteins. Due to the robust and accurate DNA synthesis ability, *Sso*Dpo1 is the most likely candidate for the replicative DNA polymerase. However, it is not known if Dpo1 is the only replicative polymerase, like Pol III in *E coli,* or if multiple replicative polymerases exist in archaea like those in eukaryotes. Therefore, the other two B-family DNA polymerases in archaea, Dpo2 and Dpo3, should also be examined in the RNA primer transfer mechanism. A bold hypothesis is that two or more DNA polymerases are involved in the RNA primer transfer

mechanism possibly even different on the leading and lagging strands. The RNA primer synthesized by *Sso*DnaG can be elongated by a separate polymerase into RNA/DNA hybrid fragments. Then the *Sso*Dpo1 replaces that polymerase for further elongation of the hybrid fragments. Therefore, knowledge of RNA elongating ability of Dpo2, Dpo3, and Dpo4 would be useful to elucidate this mechanism.

REFERENCES

1. Kelman Z, W. M. (2005) Archaeal DNA replication and repair. *Curr. Opinion in Microb.* 8, 669-676.

2. MacNeill SA. (2001) Understanding of the enzymology of archaeal DNA replication: progress in form and function. *Mol. Microbiol.* 40, 520-529.

3. Kelman Z. (2000) DNA replication in the third domain (of life). *Curr. Protein Peptide Sci.* 1, 139-154.

4. Zuo Z, Rodgers CJ, Mikheikin AL, and Trakselis MA. (2010) Characterization of a functional DnaG-type primase in archaea: Implications for a dual-primase system. *J. Mol. Biol.* 397, 664-676.

5. Berquist B R., D. P. D. S. (2007) Essential and non-essential DNA replication genes in the model halophilic Archaeon, halobacterium sp NRC-1, *BMC. Genet. 8*, 31-42.

6. Assiddiq BF, Snijders AP, Chong PK, Wright PC, and Dickman M J. (2008) Identification and characterization of *sulfolobus solfataricus* P2 proteome using multidimensional liquid phase protein separations. *J. Proteome. Res.* 7, 2253-2261.

7. Kuchta RD. (2010) Mechanism and evolution of DNA primases. *Biochim. Biophys. Acta.* 1804, 1180-1189.

8. Rolfsmeier ML, Laughery MF, and Haseltine CA. (2010) Repair of DNA double-strand breaks following UV damage in three Sulfolobus solfataricus strains. *J. Bacteriol.* 192, 4954-4962.

9. Zhang L, Lou H, Guo L, Zhan Z, Duan Z, Guo X, and Huang L. (2010) Accurate DNA synthesis by *Sulfolobus solfataricus* DNA polymerase B1 at high temperature. *Extremophiles.* 14, 107-117.

10. Grogan DW. (2000) The question of DNA repair in hyperthermophilic archaea. *Trends Microbiol.* 8, 180-185.

11. White MF. (2011) Homologous recombination in the archaea: the means justify the ends, *Biochem. Soc. Trans.* 39, 15-19.

12. Smith P, Das U, Zhu H, Shuman S. (2011) Structures and activities of archaeal members of the LigD 3'-phosphoesterase DNA repair enzyme superfamily. *Nucleic Acids Res.* 39, 3310-3320.

13. Nair PA, Shuman S. (2010) Structure of bacterial LigD 3'-phosphoesterase unveils a DNA repair superfamily. *Proc. Natl. Acad. Sci. U S A*. 107, 128222-112827.

14. Andrade P, Juarez R, Saro F, Blanco L. (2009) Limited terminal transferase in human DNA polymerase mu defines the required balance between accuracy and efficiency in NHEJ. *Proc. Natl. Acad. Sci. U.S.A*. 106, 16203-16208.

15. Lao-Sirieix SH, and 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, 1251-1263.

16. Liu L, Ishino S, Bocquier AA, Cann IK, Kohda D & Ishino Y. (2001) The archaeal DNA primase: Biochemical characterization of the p41-p46 complex from Pyrococcus furiosus. *J. Biol. Chem*. 276, 45484-45490.