April 1995

Primase structure and function

Mark A. Griep
University of Nebraska-Lincoln, mgriep1@unl.edu

Indian Journal of Biochemistry & Biophysics
Vol. 32, August 1995, pp.171–178

Minireview

Primase structure and function

Mark A Griep

Department of Chemistry and Center for Biotechnology, University of Nebraska, Lincoln, NE 68588-0304 USA

Abstract

Primase is the ssDNA-dependent RNA polymerase that synthesizes RNA primers during DNA replication. In common with all DNA and RNA polymerases, primase has structural and functional features involved in polymer elongation. As RNA polymerase, it has structural and functional features for initiating chain synthesis. As a primase, it has structural and functional features for initiating chain synthesis on ssDNA. Using amino acid sequence analysis the structure of *Escherichia coli* primase responsible for binding zinc, at least three magnesium, and DnaB helicase has been identified. One of the magnesium binding motifs resembles the "active magnesium" motif found in all DNA and RNA polymerases. This motif can be considered to be involved in phosphodiester bond formation. The region with the putative zinc binding motif is the most highly conserved portion, including more than 25% of identical residues among bacterial primases. The function of the zinc finger may be to bind ssDNA in a sequence-specific manner. Primase has "RNAP" motif, a sequence found in all RNA polymerases which may be involved in chain initiation. Many of the observations concerning primer synthesis initiation *in vivo* have been reproduced by several of the *in vitro* assay systems. Important among these is that Okazaki fragments are initiated *in vivo* from d(CTG) most of the time. This trinucleotide initiation specificity has been shown to be an intrinsic property of pure primase *in vitro*. Using artificial ssDNA templates, primase has been shown to be the slowest and most error-prone polymerase yet studied. The rate-determining step is the first phosphodiester bond formed. Any protein which can influence either the dinucleotide synthesis rate or primase–ssDNA binding affinity will also play a key role in the regulation of primer synthesis initiation.

Introduction

Primase synthesizes the RNA primers that are elongated by DNA polymerase. Primase from *Escherichia coli* has received attention pri marily as a result of its interaction with the other enzymes and proteins that participate in DNA replication[1–3]. Primer RNA is synthesized once to initiate leading strand DNA synthesis and repeatedly on the lagging strand ssDNA to initiate Okazaki fragment synthesis. It is on the lagging strand that primase plays its most important role. In fact, it has been stated that primase plays the key regulatory role in establishing the frequency of Okazaki initiation[4]. By controlling the initiation of Okazaki fragment synthesis, primase plays the key regulatory role during DNA replication. It is for these reasons that my lab has undertaken a study of the structure and function of primase. An understanding of how primase works will greatly enhance our understanding of how DNA replication is regulated.

Many insights into primase structure and function are a consequence of recognizing that it is an RNA polymerase. RNA polymerases are capable of starting chain synthesis while DNA polymerases are not. Even so, there are many differences between the action of the ssDNA-dependent primase and the dsDNA-dependent transcription RNA polymerases. Based on its function at a replication fork, primase must minimally bind to 2 magnesium, 2 NTP's, non-specific ssDNA, the initiating trinucleotide d(CTG) ssDNA, a RNA-DNA heteroduplex once primer synthesis has been initiated, DnaB helicase, and one or more of the ten subunits of DNA polymerase III holoenzyme. For two of the proteins with which primase interacts there is enough information that they have been the subject of specific reviews, DNA polymerase III holoenzyme[5,6] and SSB[7,8]. In contrast, another enzyme with which primase interacts, DnaB helicase has only recently been the subject study and has yet to be separately reviewed. Only those properties of SSB, DnaB helicase, and DNA polymerase III holoenzyme that are relevant to primase structure and function will be described.

Brief History of Primase Genetics

Temperature-sensitive phenotypes were used to identify the *dnaG* locus on the *E. coli* chromosome as being essential for lagging strand DNA replication initiation[9-11], Extracts from these temperature-sensitive cells lacked an enzyme essential for DNA replication of bacteriophage ssDNA. That enzyme "primed" DNA synthesis by synthesizing a RNA polymer which could be elongated by DNA polymerse III holoenzyme[12,13]. The DNA sequence and therefore protein primary sequence of this enzyme were determined[14]. At the same time, it was found that the genes encoding the 30S ribosomal factor S21, primase and the RNA polymerase sigma subunit are in an operon. This operon groups key proteins from the three processes of molecular biology. It has been recently demonstrated that the macromolecular synthesis operon is a conserved feature of a diverse group of gram-negative bacteria[1]. The genes for primase and the sigma subunit are also in an operon in the gram-positive bacteria Bacillus subtilis16 suggesting that it may be a conserved feature of many types of bacteria. By grouping primase and the sigma subunit in an operon, the cell has grouped proteins involved in the sequence-specific initiation of their respective systems. The sigma subunit of *E. coli* RNA polymerase holoenzyme is responsible for the sequence-specific binding to promoters, the first step in transcription. Expression of the proteins in this macromolecular synthesis operon are subject to many levels of control[17-19] such that primase is expressed at low levels. To overcome the natural low abundance of primase, several primase-overproducing strains have been produced. This eases the purification of large quantities of protein used in detailed functional and structural studies. Recently, the sequences of several primases was compared and it was predicted[20] and confirmed[21] that primase has a single zinc prosthetic group. The function of zinc is under active study in my lab.

Primase Structure

High resolution structure of any primase is not yet available. Consequently most of the structural information about *E. coli* primase has been obtained from amino acid sequence analysis (Fig. 1). From the following analysis, an attempt is made to establish correlations between primase function and sequence motifs. The minimal cofactors to which primase binds are zinc, two magnesium, the initiating nucleotide ATP, the second nucleotide, nonspecific ssDNA, initiation template trinucleotide d(CTG), the primed template following initiation, DNA polymerase III holoenzyme, and DnaB helicase.

The *E. coli* protein has 582 amino acids, very close to the value from other bacteria. There are 67 identical residues found in the 6 bacterial primases for which the entire sequence is known. In addition, there are many highly conserved residues. The first motif common to bacterial primases consists of a cluster of as many as 5 lysines and arginines within 9 residues of one another (Fig. 1). This is the only region with such a high cluster of cationic residues. The sequence of this motif in *E. coli* primaseis 26-RVKLKKQGK, where the cationic residues are either identical or highly conserved among the bacterial primases. This cationic cluster has been proposed to be involved in the electrostatic binding of ssDNA by primase although there has been no experimental evidence put forward[21]. Electrostatic binding denotes the sequence non-specific binding mode of many DNA binding proteins. In contrast, sequence-specific binding is achieved through specific hydrogen bonding to the DNA via hydrophilic residues.

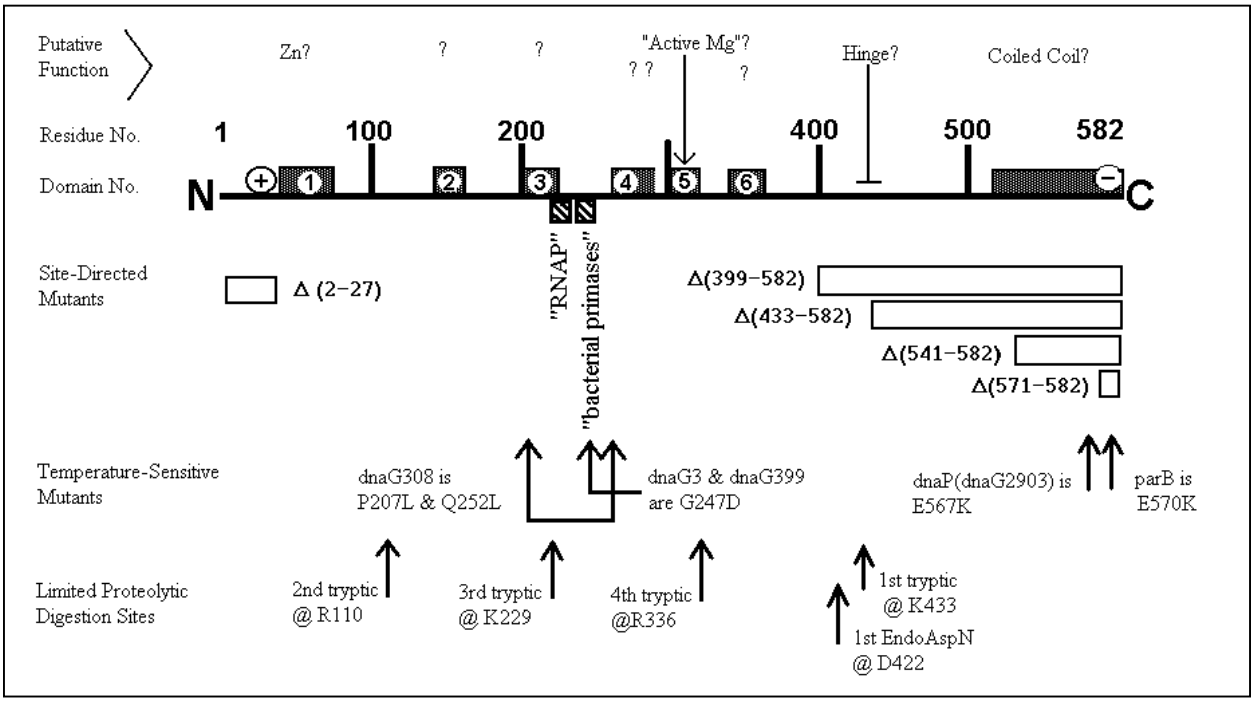

Fig. 1— Primase amino acid sequence analysis [The 582-amino acid sequence of *E. coli* primases is represented by the line. The locations of motifs, clusters of conserved or identical residues, are indicated by boxes. Possible functions for the motifs are indicated. The functions are mostly taken in analogy to similar motifs found in other enzymes. The shaded and numbered motifs were identified by Ilyina *et al.*[20], the striped motifs were identified by Versalovic and Lupski[26], and the speckled boxes were identified by Griep (unpublished results). Several site-directed deletion mutants have been synthesized and the deleted portions indicated by white boxes[21,31,32]. The carboxyl terminal deletion mutants have a few non-primase amino acids added to them. Temperature-sensitive mutations can reveal sites that are important for function but cannot determine whether these mutations alter protein folding or the function of adjacent catalytic domains. The sites of highest protease susceptibility may indicate highly exposed regions[31,32]. The motifs were identified by comparing the primases from eight bacteria and four bacteriophage. The full amino acid sequence from six bacteria were used: *Bacillus subtilis,* GenBank Accession number XO3897; *Clostridium acetobutylicum*, Z23080; *Escherichia coli*, J0[illegible]87; *Haemophilus influenzae*, L11044; *Rickettsia prowazekii*, M95860; and *Salmonella typhimurium*, M14427. Partial sequences were available for *Buchnera aphidicola*, P32000 and *Lactococcus lactis*, Dl0168. The sequences of four *E. coli* bacteriophage were used: P4, T4, T3 and T7. Even though the eukaryotic dimeric primases share some of these motifs, the motifs are arranged in a different order and were not included in the analysis.]

Motif I is a zinc finger motif and is involved in the chain initiation step (Fig. 1). This is the longest motif among the bacterial primases and involves about 25% of identical residues indicating that the zinc finger motif must play an important role in primase function. The residues in this motif were hypothesized to bind zinc because of their sequence similarity to other known zinc binding motifs[22]. The motif sequence in *E. coli* primase is 40—**C**PF**H**NEK TPSFTVNGE**K**QFYH**C**FG**C** where the putative zinc binding residues are in bold letters and the identical or conserved residues among the bacterial primases are underlined. The order of zinc ligating residues in all of the bacterial primases is CHCC, representing a unique class of zinc fingers. The bacteriophage primases fall into the CCCC class which has many members[23]. Some of the members of the CCCC class are involved in binding DNA, RNA or heteroduplexes and the tertiary structure of one of them was recently solved[23].

There are several possible functions for the zinc finger in primase. Zinc may play a catalytic role by acting as a Lewis acid to promote phosphodiester bond formation. If, instead, zinc plays a structural role, the amino acid residues in the zinc finger motif may participate in any of several functions: (i) The sequence may provide ssDNA template trinucleotide initiation specificity, which in the case of *E. coli* primase is the trinucleotide d(CTG). Template trinucleotide specificity was hypothesized as its role in bacteriophage T7 primase[22] and may be the role in *E. coli* primase as well. (ii) The sequence may bind to the initiating nucleotide, ATP, as does the B-site zinc sequence in *E. coli* RNA polymerase[24,25]. (iii) The sequence may be involved in protein-protein interactions as it seems to be in the case of bacteriophage T4 single-stranded DNA binding protein. (iv) Zinc may stabilize the folded conformation of primase in the same way that it does for many other zinc proteins.

Studies in my lab have shown that primase lacking zinc, apoprimase, is fully functional in site-specific primer synthesis (unpublished results). This negates a role for zinc in catalysis. Apoprimase also has the same conformation as primase as indicated by limited tryptic digestion. This negates a role in thermostability. The primary role for zinc appears to be to prevent the oxidation of the zinc-ligating cysteines. The evidence is that apoprimase becomes inactive upon storage at 4°C for only a few hours whereas primase is stable for months (Swart and Griep, unpublished results). Pretreatment of the inactive apoprimase with a reducing agent leads to full recovery of activity. While these results define the role of zinc, it does not establish the role of the amino acid residues. Studies are underway to test the remaining possibilities.

Primase motifs 2 and 3 have not been correlated with any function and will not be discussed here. However, overlapping motif 3 is the "RNAP" motif which is followed by the "bacterial primases" motifs[26]. The sequence in *E. coli* are 214-GRVIGFGGRVL and 229-PKYLN-SPETDIFHKG and nearly all of these residues are highly conserved or identical. It is interesting to note that the three temperature-sensitive mutations of primer synthesis are located in this general region (Fig. 1). *In vivo*, these mutant primases are fully active at low temperatures but inactive at 42°C. The mutated residues are not involved catalytically but somehow stabilize primer synthesis function. These residues may either be adjacent to critical residues or involved in protein folding. However, the sites of these mutations correlate with the amino acid analysis in establishing the importance of this region for function. The RNAP motif is common to all RNA polymerases. It may play a role in chain initiation since this is a function specific to all RNA polymerases and not DNA polymerases. The mechanism by which these residues carry out this role has not been postulated. However, I find that a secondary structure algorithm predicts that

the conserved glycines strongly correlate with a turn. Perhaps the conserved arginines at this turn coordinate the triphosphate moiety of the initiating ATP.

Motifs 4, 5 and 6 have been postulated to be magnesium binding sites[20,27]. The sequences of these motifs in *E. coli* primase are 261-LVVEGYMDVVAL, 304-VICCYDGDRAG, and 340-LPDGEDPDTLVRKEG. Even though each of these three motifs contains pairs of carboxylate side chains, it does necessarily indicate that three magnesium need to be bound for primase activity. It is quite possible that two of these motifs are required to coordinate a single magnesium. Based on sequence and structural similarity, motif 5 most closely resembles the "active" magnesium binding sites found in all other polymerases[28,29]. The active magnesium motif contains the sequence Asp-X-Asp which is preceded by a beta strand and followed by either a turn or another beta strand[30]. Secondary structure predictions confirm that the primase Asp-Gly Asp is preceded by a beta strand and followed by turn residues. This secondary structure corresponds to the transcription RNA polymerase type of active magnesium site.

Beginning near residue 400, bacterial primase sequence similarity becomes negligible. This is also the approximate position at which the bacteriophage primases either terminate or, in the case of bacteriophages P4 and T7, the sequence of a helicase is found[20]. These helicases show similarity to the *E. coli* DnaB helicase with which *E. coli* primase interacts. Perhaps as a result of bacteriophage information compactness, these two proteins were joined. This suggests that the carboxyl termini of bacterial primases bind to their respective helicases. This also suggests that the amino terminus of primase contains the structure necessary to carry out primer synthesis. Carboxyl terminal deletion mutants have been engineered which confirm both of these suggestions. The deletion mutants are named here by indicating the residues that have been deleted from native primase. The primer synthesis activities of the Δ(399-582) and Δ(433-582) terminal deletion mutants of *E. coli* primase are normal with only minor deficiencies[31,32]. In contrast, it was demonstrated that the Δ(433-582) mutant was no longer capable of interacting with DnaB helicase[32].

The residues near position 425 appear to be a hinge region (Fig. 1). Proteolytic susceptibility is commonly taken as a measure of surface exposure by the cleaved residue. The first digestion site by trypsin is lysine 433 and by endonuclease Asp-N is aspartate 422 (ref. 31,32). Since the first cut represents the single most accessible portion of the native protein conformation, this is considered to be a hinge region.

When the secondary structure probabilities of the carboxyl terminal residues of bacterial primases are compared (Griep, unpublished results), the final 70 residues show a strong probability for forming an alpha helix. Perhaps this helical "tail" interacts in the manner of a coiled coil interaction with the DnaB helicase.

The very carboxyl terminus of primase is rich in glutamate. This region has been postulated to be involved in membrane contact[33]. When either glutamate 567 or 570 are mutated to a lysine, the result is temperate-sensitivity. However, the result is not a primer synthesis deficiency at high temperature but difficulty in cellular partitioning followed DNA replication. The membrane duplication machinery continues to create daughter cell membrane but no partition is ever formed because the chromosomal DNA remains located in the parent cell. As a result, filamentous cells in which the DNA remains located in the central cell are observed under the microscope. However, interaction between primase and the membrane is not the only way to explain this type of mutation.. Any deficiency in DNA replication which does not lead to inhibition of the membrane synthesis machinery might achieve the same result.

From the sequence analysis it has been possible to identify regions for the binding of non-specific ssDNA, zinc, as many as three magnesium, and DnaB helicase. One of the magnesium is likely to be bound to the NTP which is incorporated into the growing strand. The zinc finger motif may be involved in trinucleotide d(CTG) ssDNA binding, initiating ATP binding, or both. It is less clear whether there is a particular motif associated with the binding to one of the subunits of DNA polymerase III holoenzyme. The "bacterial primases" motif may serve this function. No motif associated with RNA–DNA heteroduplex binding was expected because, even though crystal structures of DNA and RNA polymerases have established that they have a common "hand" structure with which they bind the template and primer, there is no sequence similarity that correlates with these structures.

Primase Function *in vivo* and *in vitro*

Many of the observations that have been made about Okazaki fragment synthesis *in vivo* have been accounted for or repeated using *in vitro* model systems. This correlation has held true even though many of the *in vitro* systems do not utilize all of the proteins thought to be present at the replication fork *in vivo*. The average RNA primer length *in vivo* is 11 ± 1 nucleotides[34,35]. The primers observed using purified proteins is nearly the same in many different replication systems[36]. *In vivo*, Okazaki fragments are an average 1500 nucleotides but have a distribution of between 500 and 3000 nucleotides[37]. Again, this result was confirmed with an artificial replication system[4,38]. *In vivo*, the primase initiation site consists of the trinucleotide sequence d(CTG) about 60% of the time, d(TTG) 20%, and other sequences the remaining percentage[34,39]. Unlike many enzymes, primase is actually more stringent *in vitro*; it will initiate from d(CTG) but not d(TTG)[40].

Brief History of Bacteriophage G4

Our most detailed understanding of primase function has come from studies of ssDNA bacteriophage initiation, such as bacteriophages G4 and ϕX174. These systems proved convenient because of the limited number of proteins required for primer synthesis initiation. Bacteriophage G4 is a ϕX-like phage that was originally isolated from four litres of raw sewage, "a gift of the people of Branford, CT"[41]. Shortly thereafter the minimal protein requirements for G4 complementary strand origin DNA synthesis were established. Only SSB need be present to cause primase to synthesize primers from a specific site in the origin[12]. The length of the RNA primer and the site of its initiation were then determined[42,43], DNA sequences with similarity to the G4 origin have been found in other bacteriophage and plasmids (now called G sites).

Infectivity of mutant G4 has established that secondary structure is important[44] as is the spacing between and the lengths of the various secondary structures[45–48]. The minimal number of nucleotides that constitute the origin was limited to 274 by this method as well[49]. The respective roles of SSB, the hairpins and the initiating trinucleotide d(CTG) have been established from *in vitro* studies[40,50].

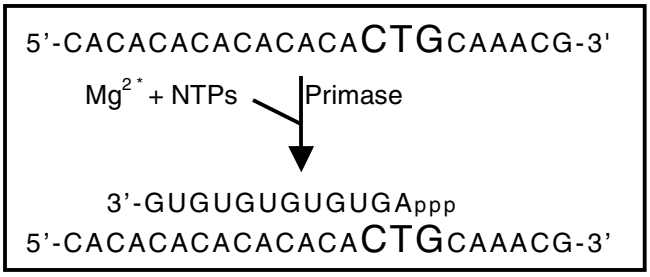

Fig. 2— Functional aspects of primer synthesis activity: On artificial ssDNA templates, primer synthesis is usually initiated with ATP at the thymine in the trinucleotide d(CTG)[40]. It is also possible to initiate opposite the d(CTG) cytosine by using high GTP concentrations (unpublished .results). In both cases, initiation is complementary to the d(CTG). At least six nucleotides are required on the 3'-flanking side for primer synthesis to be initiated from the d(CTG). The number of 3'-flanking nucleotides is more important than the actual sequence of those nucleotides. RNA polymers shorter than 12 nucleotides in length do not accumulate. This indicates that the rate limiting step precedes the formation of the first phosphodiester bond and that there is a mechanism to limit primers to a length of 12 nucleotides.

Site Specific Primer Synthesis Kinetics

Sequence-specificity may result from the need by all RNA polymerases to stabilize the first dinucleotide formed. For instance, both *E. coli* primase and RNA polymerase prefer to initiate with a purine[51]. If a special binding site on the enzyme is capable of binding ATP and some portion of the template, stabilization can be achieved. Primer synthesis initiates opposite the thymine in the trinucleotide d(CTG) with ATP (Fig. 2). Synthesis proceeds in the usual direction creating the dinucleotide pppApG. After this dinucleotide has been created, primer synthesis continues according to whatever sequence is present on the template. Primase will incorporate ddNTPs into an RNA primer, a feature that has allowed us to establish whether the RNA is a complementary to the ssDNA template[40].

The shortest observed primers are 12 nucleotides long. The ability to synthesize primers that are 12-nucleotides is an intrinsic property of primase that we have observed with several templates. This activity has implications regarding the initiation and termination of primer synthesis. Since no dinucleotides are observed, the rate-determining step of primer synthesis precedes dinucleotide bond formation. The next 10 phosphodiester bonds are formed so much more rapidly that the first primer observed is 12 nucleotides long. There are two structural possibilities for having such an ability: either the enzyme can recognize when it has made a 12-mer; or the enzyme maintains contact with the d(CTG) and stretches until it has made a 12-mer and can stretch no more. After primase has synthesized a 12-mer, it enters a distributive mode in which it becomes equally likely to elongate by one nucleotide or not.

Primer synthesis is slow *in vitro* with a rate of 0.00089 s^{-1} (Swart and Griep, communicated). Given the cellular primase concentration of 75 copies/cell[1], there is the potential to synthesize 0.067 primers/s. This is too slow to keep up with Okazaki fragment synthesis *in vivo* which is known to occur at the rate of about 1 primer/s. Other replication enzymes catalyze their reactions at an *in vitro* rate that is consistent with their estimated *in vivo* rates. For instance, DNA polymerse III holoenzyme synthesizes DNA *in vitro* at a rate of greater than 980 nucleotide/s (Griep and McHenry, unpublished results). If an average Okazaki fragment of 1500 nucleotides is synthesized once per second *in vivo*, then the *in vitro* rate for the DNA polymerase is almost fast enough. The solution to the primase conundrum may be an interaction with DnaB helicase. It

has been demonstrated that primase must bind to DnaB helicase to initiate primer synthesis at the replication fork[32]. DnaB helicase may greatly stimulate the rate of primer synthesis initiation as it does in the general priming system[52].

DNA replication has to have very high fidelity to conserve the genetic information. For instance, DNA polymerase III holoenzyme misincorporates less than once in 100,000 nucleotides and there are many enzymes to correct misincorporated nucleotides. In contrast, we find that primase misincorporates once in 20 nucleotides (Swart and Griep, communicated). This is even higher than that in eukaryotic primase which misincorporates once in 30 (ref. 53). Perhaps the error-prone nature of primases is needed to insure that initiation does take place. It is also possible that an error-fi1led RNA primer is a better target for the cellular machinery that removes the primer, fills in the gap and then ligates the lagging strands together .

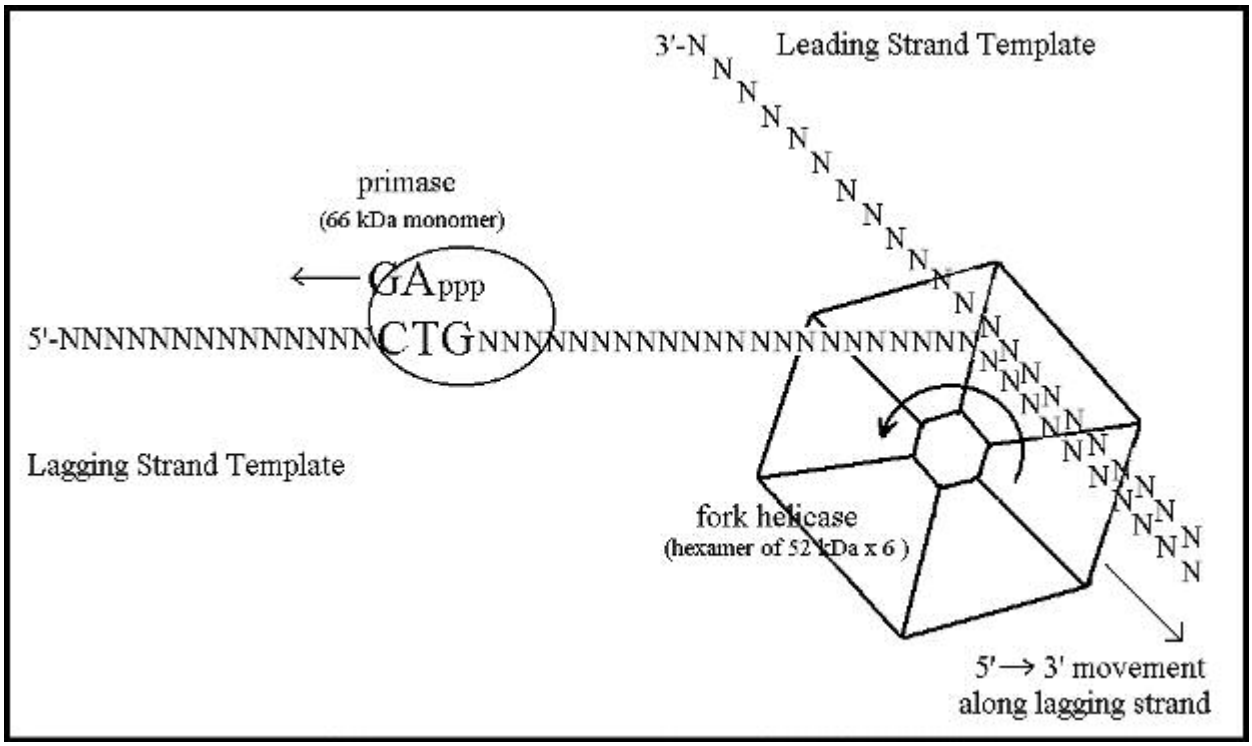

Fig. 3— Action of primase at the replication fork: The relative positions or leading and lagging strand ssDNA, primase, and DnaB helicase demonstrate a possible mechanism for the interaction between primase and helicase at the replication fork. DnaB helicase is a lagging strand helicase which means that it maintains contact with the lagging strand ssDNA after duplex unwinding[63]. As it is a hexameric helicase, it may have a mechanism considerably different from that of the dimeric helicases. It is certainly true that many dimeric helicases are leading strand helicases. Nevertheless, an analogy is made here with the “rolling” mechanism of dimeric helicase which couples ATP hydrolysis to sequential duplex unwinding[64]. If DnaB helicase has its identical subunits arranged adjacent to one another, then duplex unwinding may involve a simple “rolling' along the template. As the ssDNA is created, primase can initiate primer synthesis on the exposed lagging strand. At some point in the primer synthesis cycle, primase and its growing primer are released from the helicase. In this model, the “rolling” of the DnaB helicase places primase in the correct direction for primer synthesis antiparallel to the lagging strand once it is released.

Possible Order in Which the Primase Motifs Function

Having identified possible functions for many of the ten primase motifs, I will conclude by hypothesizing the order in which those motifs function. The initiating nucleotide ATP binds to at least two of the motifs. Primase from *E. coli* can bind ATP in the absence of magnesium or ssDNA (Urlacher and Griep, unpublished results). If primase is like *E. coli* RNA polymerase[24], the adenine moiety will be bound near the zinc finger motif. In addition, the triphosphate moiety of ATP may be bound at the RNAP motif as discussed above. Tertiary folding of primase must bring the zinc finger motif adjacent to the RNAP motif.

Our preliminary results also indicate that free primase can bind the second nucleotide but requires magnesium to do so (Urlacher and Griep, unpublished results). Consistent with our result is that T7 primase will synthesize random diribonucleotides in the absence of ssDNA template but in the presence of manganese[22,54]. In the case of *E. coli* primase, the magnesium, and therefore triphosphate moiety of GTP would be bound by motif 5 located in the middle portion of primase. The T7 primase results indicate that by doing so, the triphosphate of the second nucleotide is brought into proximity to the 3'-hyroxyl of the initiating nucleotide ribose so that a phosphodiester bond can be formed.

The initiation of Okazaki fragment synthesis is under strict control[36,38]. For example, 90% of the primers synthesized are utilized. This indicates that there is coordination between the completion of the Okazaki fragment and the initiation of the next. As indicated above, primer synthesis initiation is the rate-determining step for primase. This is the step that must be regulated during Okazaki fragment synthesis. Perhaps as DNA polymerase III holoenzyme completes one Okazaki fragment, it slows down and this is communicated to the DnaB helicase at the fork. As a dimeric DNA polymerase, the leading and lagging strand halves would be able to communicate with one another and the other enzyme at the replication fork[55,56]. As SSB does not play a role on the leading strand ssDNA[57], the leading strand DNA polymerase half probably abuts the DnaB helicase as it progresses. Thus, when the lagging strand DNA polymerase stalls at the end of Okazaki fragment synthesis, this could be communicated to the helicase. At some point in the cycle, the carboxyl terminal "tail" of primase binds to the helicase. When the lagging strand DNA polymerase is ready or near ready, primase nonspecifically binds to the lagging strand ssDNA bound to DnaB helicase. It may bind nonspecifically via its positive cluster near the zinc finger motif. To do this it does not release from the helicase but probably bends at its "hinge".

DnaB helicase binds about 60 nucleotides of ssDNA[58]. A given trinucleotide will appear once every 64 nucleotides if the sequence consists of roughly 25% of each nucleotide and occurs "randomly". Thus, primase will encounter each trinucleotide during each complete cycle of the hexameric helicase (Fig. 3). CTG may be available at higher than average frequency because it is one of the leucine codons and leucine is the most frequently coded for amino acid in *E. coli* proteins[59]. At the next suitable d(CTG) in the ssDNA, the primase zinc finger may be involved in arranging the d(CTG), ATP, GTP, and magnesium sites for phosphodiester bond formation. The classical type of zinc finger CCHH binds to the major groove of duplex DNA[60,61]. Perhaps the zinc finger in primase (and all other zinc-containing RNA polymerases) is responsible for creating their own major groove. The creation of this major groove from d(CTG), ATP and GTP would then stimulate dinucleotide synthesis and the initiation of Okazaki fragment synthesis. DnaB helicase may stimulate dinucleotide bond formation by making the templates less flexible.

The next 10 phosphodiesters bonds may or may not be formed while primase remains bound to the DnaB helicase. This primer elongation cycle involves sequential pyrophosphate–magnesium release, template translocation, and correct NTP–magnesium binding. Primase does not readily release the RNA primer upon its completion and DNA polymerase III holoenzyme probably first binds to primase rather than to the RNA polymer[62]. Ultimately, primase is displaced and can re-enter the Okazaki fragment synthesis cycle.

Acknowledgment

I thank my students for their diligence in research and their critical reading of this review during its preparation. The work described was funded by the National Institutes of Health (Grant GM 47490) and the UNL Center for Biotechnology.

References

1 McMacken R, Silver L & Georgopoulos C, (1987) in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, (Neidhardt, F C, ed.), pp. 564–612, American Society for Microbiology, Washington, DC.

2 Marians K J, (1992) *Annu Rev Biochem*, 61, 673–719.

3 Kornberg A & Baker T A, (1992) *DNA Replication*, W H Freeman and Company, New York.

4 Wu C A, Zechner E L & Marians K J, (1992) *J Biol Chem*, 267, 4030–4044.

5 McHenry C S, (1988) *Annu Rev Biochem*, 57, 519–550.

6 O'Donnell M, (1992) *BioEssays*, 14, 105–111.

7 Chase J W & Williams K R, (1986) *Annu Rev Biochem*, 55, 103–136.

8 Lohman T M & Ferrari M E, (1994) *Annu Rev Biochem*, 63, 527–570.

9 Wechsler J A & Gross J D, (1971) *Mol Gen Genet*, 113, 273–284.

10 Lark K G, (1972) *Nature New Biology*, 240, 237–240.

11 Louarn J-M, (1974) *Mol Gen Genet*, 133, 193–200.

12 Bouche J-P, Zechel K & Kornberg A, (1975) *J Biol Chem*, 250, 5995–6001.

13 Wickner S, (1977) *Proc Natl Acad Sci USA*, 74, 2815–2819.

14 Burton Z F, Gross C A, Watanabe K K & Burgess R R, (1983) *Cell*, 32, 335–349.

15 Versalovic J, Koeuth T, Britton R, Geszvain K & Lupski J R, (1993) *Mol Microbiol*, 8, 343–355.

16 Wang L-F, Price C W & Doi R H, (1985) *J Biol Chem*, 260, 3368–3372.

17 Konigsberg W & Godson G N, (1983) *Proc Natl Acad Sci USA*, 80, 687–691.

18 Lupski J R, Ruiz A A & Godson G N, (1984.) *Mol Gen Genet*, 195, 391–401.

19 Nesin M, Lupski J R & Godson G N, (1988) *J Bacteriol*, 170, 5759–5764.

20 Ilyina T V, Gorbalenya A E & Koonin E V, (1992) *J Mol Evol*, 34, 351–357.

21 Stamford N P J, Lilley P E & Dixon N E, (1992) *Biochim Biophys Acta*, 1132, 17–25.

22 Bernstein J A & Richardson C C, (1988) *Proc Natl Acad Sci USA*, 85, 396–400.

23 Qian X, Gozani S N, Yoon H, Jeon C, Agarwal K & Weiss M A, (1993) *Biochemistry*, 32, 9944–9959.

24 Chatterji D & Wu F Y-H, (1982) *Biochemistry,* 21, 4657–4664.

25 Chatterji D, Wu C-W & Wu F Y-H, (1984*) J Bioi Chem,* 259, 284–289.

26 Versalovic J & Lupski J R, (1993)*Gene ,* 136, 281–286.
27 Pansegrau W & Lanka E, (1992) *Nucleic Acids Res*, 20, 4931.
28 Argos P, (1988) *Nucleic Acids Res*, 16, 9909–9916.
29 Copeland W C & Wang T S-F, (1993) *J Biol Chem*, 268, 11028–11040.
30 Delarue M, Poch O, Tordo N, Moras D & Argos P, (1990) *Protein Eng*, 3, 461–467.
31 Sun W, Tornlo J, Steitz T A & Godson G N, (1994) *Proc Natl Acad Sci USA*, 91, 11462–11466.
32 Tougu K, Peng H & Marians K J, (1994) *J Biol Chem*, 269, 4675–4682.
33 Grompe M, Versa1ovic J, Koeuth T & Lupski J R, (1991) *J Bacteriol*, 173, 1268–1278.
34 Kitani T, Yoda K-y, Ogawa T & Okazaki T, (1985) *J Mol Biol*, 184, 45–52.
35 Yoda K-y, Yasuda H, Jiang X-W & Okazaki T, (1988) *Nucleic Acids Res*, 16, 6531–6546.
36 Zechner EL, Wu C A & Marians K J, (1992) *J Biol Chem*, 267 , 4054–4063.
37 Ogawa T & Okazaki T, (1980) *Annu Rev Biochem*, 49, 421–457.
38 Zechner EL, Wu C A & Marians K J, (1992) *J Biol Chem*, 267, 4045–4053.
39 Yoda K-y & Okazaki T, (1991) *Mol Gen Genet*, 227, 1–8.
40 Swart J R & Griep M A, (1993) *J Biol Chem*, 268, 12970–12976.
41 Godson G N, (1974) *Virology,* 58, 272–289.
42 Bouche J-P, Rowen L & Komberg A, (1978) *J Biol Chem*, 253, 765–769.
43 Rowen L & Komberg A, (1978) *J Biol Chem*, 253, 770–774.
44 Lambert P F, Kawashima E & Reznikoff W S, (1987) *Gene*, 53, 257–264.
45 Sakai H, Hiasa H, Iwamoto K, Horimoto T, Komano T & Godson G N, (1988) *Gene*, 71, 323–330.
46 Hiasa H, Tanaka K, Sakai H, Yoshida K, Komano T & Godson G N, (1989) *Gene*, 84, 17–22.
47 Hiasa H, Sakai H, Tanaka K, Honda Y, Komano T & Godson G N, (1989) *Gene*, 84, 9–16.
48 Hiasa H, Sakai H, Komano T & Godson G N, (1990) *Nucleic Acids Res*, 18, 4825–4831.
49 Lambert P F, Waring D A, Wells R D & Reznikoff W S, (1986) *J Virol*, 58, 450–458.
50 Sun W & Godson G N, (1993) *J Biol Chem*, 268, 8026–8039.
51 Maitra U & Hurwitz J, (1965) *Proc Natl Acad Sci USA*, 54, 815–819.
52 Arai K-i & Kornberg A, (1981) *J Biol Chem*, 256, 5260–5266.
53 Sheaff R J & Kuchta R D, (1994) *J Biol Chem*, 269, 19225–19231.
54 Mendelman L V, Beauchamp B B & Richardson C C, (1994) *EMBO J*, 13, 3909–3916.
55 McHenry C S, (1988) *Biochim Biophys Acta*, 951, 240–248.
56 McHenry C S, (1991) *J Biol Chem*, 266, 19127–19130.
57 Mok M & Marians K J, (1987) *J Biol Chem*, 262, 16644–16654.
58 Arai K-i & Km berg A, (1981) *J Biol Chem*, 256, 5253–5259.
59 Karlin S, Blaisdell B E & Bucher P, (1992) *Protein Eng*, 5, 729–738.
60 Pavletich N P & Pabo C O, (1991) *Science*, 252, 809–817.
61 Pavletich N P & Pabo C O, (1993) *Science*, 261, 1701–1707.

62 Griep M A & McHenry C S, (1989) *J Biol Chem*, 264, 11294–11301.
63 LeBowitz J R & McMacken R, (1986) *J Biol Chem*, 261, 4738–4748.
64 Lohman T M, (1993) *J Biol Chem*, 268, 2269–2272.