

Mechanisms of DNA Replication

Biochemistry 201
Advanced Molecular Biology
Doug Brutlag

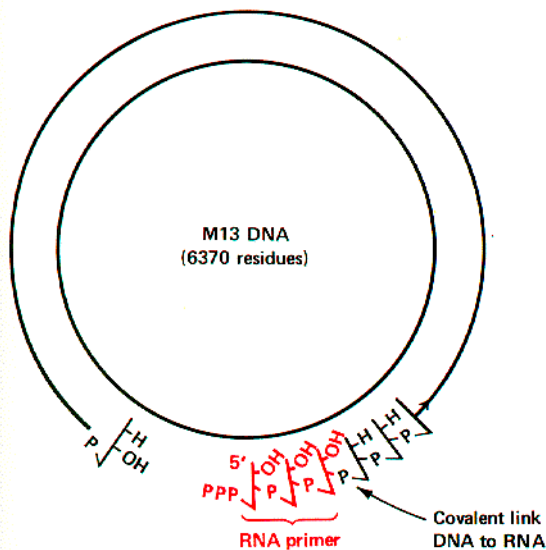
January 21, 2000

In 1971 it was known that DNA polymerases, unlike RNA polymerases, could not initiate new chains. In addition, it was known that there were three different DNA polymerases in *Escherichia coli* with different properties. The most abundant enzyme (200 molecules per cell) was DNA polymerase I which was not highly processive but which could mediate nick translation and strand displacement. DNA polymerase II was much less processive and, hence, much slower. It appeared to be primarily a repair DNA polymerase, limited to filling in short gaps resulting from repair processes. DNA polymerase III was induced 10 fold in response to SOS repair induction. Finally, the rarer DNA polymerase III (10 molecules per cell), while it could not mediate nick translation, could synthesize DNA in a highly processive fashion, and was the only enzyme that could keep up with the expected rates of DNA replication fork movement (1000 nt incorporation per second). So how could this mixture of DNA synthetic enzymes work to replicate the bacterial chromosome.

How are DNA Chains Initiated?

This problem was approached using the DNA from small circular single stranded phages M13, G4 and ØX174. It was known that the first step in the infection of these viral DNAs was the conversion of the viral DNA from a single stranded circular form to a duplex replicative form (RF) and this conversion was mediated by host cell

enzymes alone. Hence introducing the phage DNA into bacterial extracts and isolating the factors required for the conversion to RF would reveal the mechanism of initiation of DNA chains.



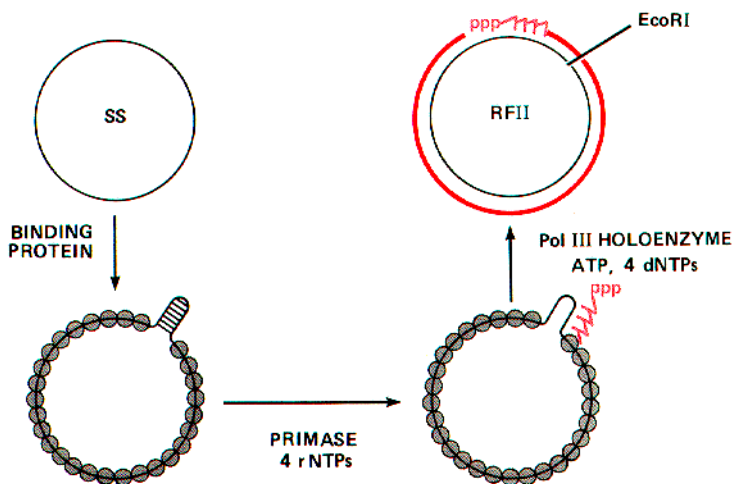
RNA Polymerase Primes M13

Based on *in vivo* results showing that M13 SS \rightarrow RF reaction was inhibited by rifampicin and other inhibitors of the host RNA polymerase, extracts of *Escherichia coli* were shown to convert M13 SS to a duplex form in a rifampicin sensitive reaction. This reaction required the four ribonucleotide triphosphates in addition to the four deoxynucleoside triphosphates. *In vitro*, a

short RNA primer was made from several locations on the M13 circle. These primers can server to initiate DNA synthesis by DNA polymerase III. Evidence favoring priming of DNA synthesis by RNA in M13 includes:

- Sensitivity to several inhibitors of the host cell RNA polymerase including rifampicin, streptolydigen and actinomycin D
- Requirement for all four rNTPs and all four dNTPs.
- Two stage conversion of SS to RF, the first of which requires rNTP and is rifampicin sensitive, the second of which is rifampicin resistant and requires dNTP in addition to ATP.
- The discovery of a phosphodiester linkage between RNA and DNA in the resulting DNA product.
- Persistence of an RNA segment in the RF in the absence of DNA polymerase I

Phages ØX174 and G4 Demonstrate Rifampicin Resistant RNA Priming



When DNA from either G4 or ØX174 is used in similar extracts, DNA synthesis is also observed, but in a rifampicin resistant reaction. Studies with the G4 template showed that it requires only three enzymes for its SS → RF conversion: single-strand DNA binding protein (SSB), dnaG protein (primase) and DNA polymerase III holoenzyme (see below). The synthesis also requires the presence of a specific origin

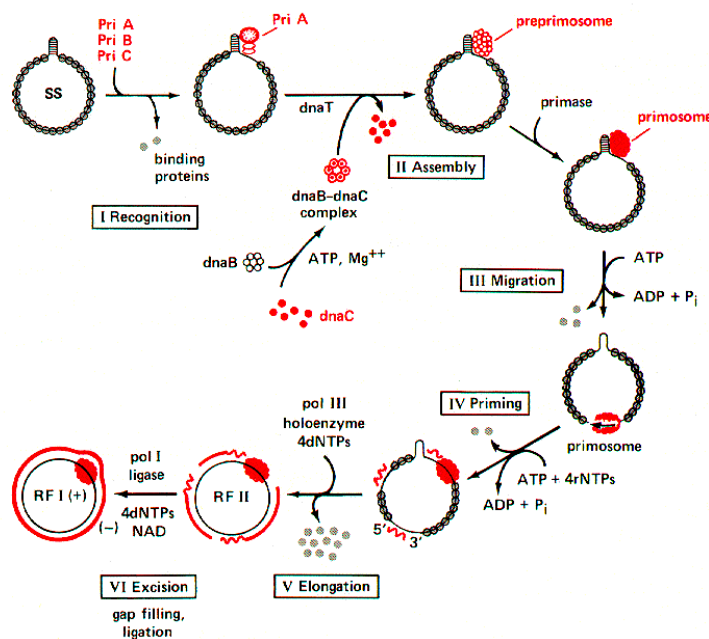
for SS DNA strand synthesis. The site of DNA priming is the same one that is used *in vivo* and it has a very particular structure including three hairpin loops. In the presence of dnaG protein (primase) a short 29 nucleotide RNA is made complementary to the first hairpin loop. In the presence of DNA polymerase III, the RNA made by dnaG protein is much shorter (9-11 bases). The primer always starts at a 5' CTG 3' template with the first two bases of the primer RNA being 5'AG complementary to the TC.

Phage ØX174 priming is much more complicated and appears to be more similar to the priming that occurs at the beginning of Okazaki fragments on the discontinuous strand of DNA synthesis at a replication fork. Initiation starts at a site known as a primer assembly site (pas) which bears no sequence similarity to the primase-binding

site in G4 or colE1 plasmid. There are extensive regions of DNA secondary structure however.

The steps in the formation of the ØX primosome involve:

- Coating of the single-stranded ØX174 DNA with *Escherichia coli* SSB DNA binding protein
- Binding of three proteins (PriA, priB and priC) to the primer assembly sequence.
- Formation of a complex of six subunits of dnaB protein coupled with six subunits of dnaC protein.

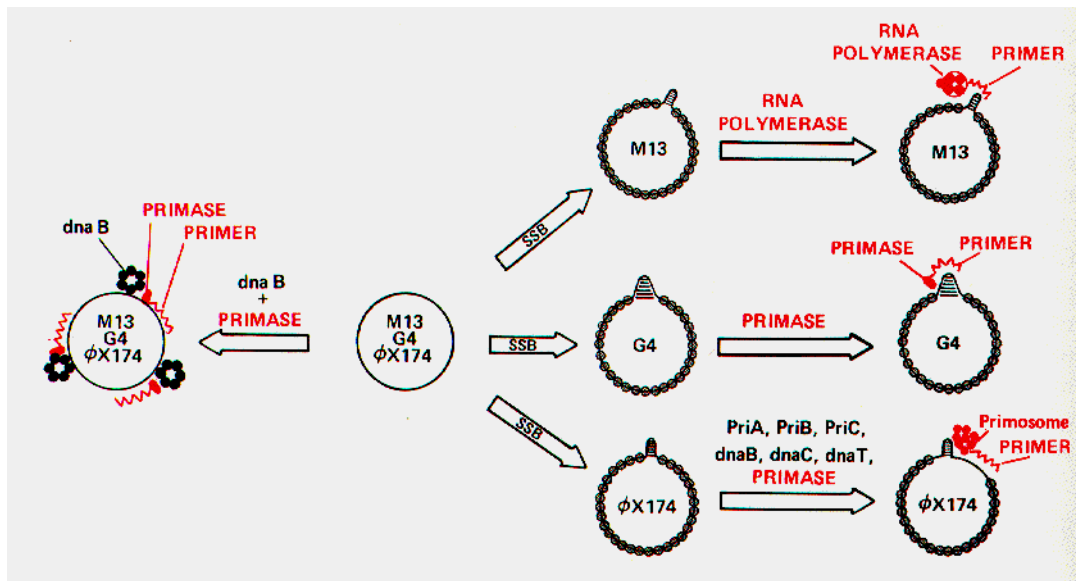


- Transfer of the complex of dnaB-dnaC to the priA-B-C complex at the primer assembly site via the dnaT gene product. dnaC dissociates at this step and the resulting complex is known as the preprimosome.

- Binding of primase (dnaG) to the preprimosome complex to form the primosome.

The mature primosome can then proceed in an ATP dependent fashion to traverse the DNA. The primosome can apparently be driven by either the dnaB protein in a 3'-5' direction or by the priA protein in the 5'-> 3' direction. Both the dnaB protein and the priA protein in the primosome can serve as DNA helicase activities. The priA protein can also displace SSB from in front of the moving primosome. dnaB cannot displace SSB and can only move on a naked DNA template. During either of these motions, the primase activity can synthesize primers 11 ± 1 nucleotides in length at various sites along the template in a reaction requiring the four rNTPs. Once DNA polymerase III extends these primers, the SSB protein is permanently displaced from the single-stranded DNA template. Removal of the RNA primers and ultimate sealing of the nicks in the DNA require the combined action of 5' exonuclease of DNA polymerase I and DNA ligase.

It is believed that similar primosome complexes are present on the lagging-strand of DNA synthesis. This is suggested by the sensitivity of DNA synthesis to disruption by temperature-sensitive mutations in *dnaB*, *dnaC*, *dnaT* and *dnaG*.



Primer sites can be generated by nuclease in some plasmid and viral DNA templates.

The continued replication of ϕ X174 RF DNA requires a viral encoded protein gpA (gene product A) which serves to nick the covalently closed circular DNA at a specific site, becoming covalently attached to the 5' end of the nicked DNA. The gpA protein then associates with a cellular DNA helicase (rep) and remains bound to it as rep unwinds the two strands of ϕ X RF. The 3-prime end of the original nick then serves as a primer terminus for extension by DNA polymerase III. At the end of a round of replication, the gpA protein mediates a strand transfer reaction releasing a single-stranded DNA circle and continuing the rolling circle DNA synthetic step. The ϕ X gpA protein contains two tyrosines that can become covalently attached to DNA during these processes. It is believed that the protein uses the two tyrosines alternately in each round of single strand synthesis.

Other viruses also use specific nicks as well as specific proteins to prime DNA synthesis. Adenovirus encodes a protein with an adenylylated serine residue that can serve as a primer beginning at the end of adenoviral DNA molecules for example

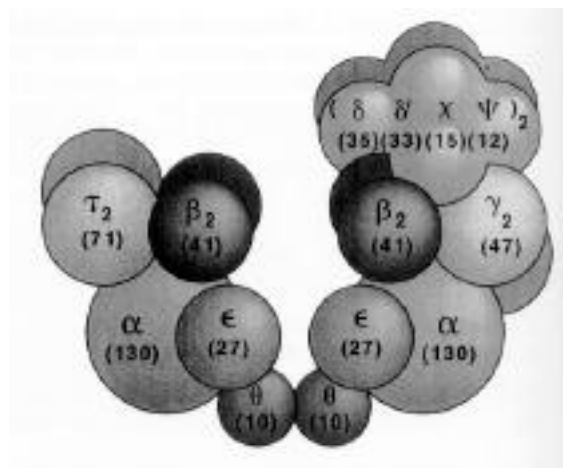
DNA Polymerase III Holoenzyme

Table 5-2
Major subunits and subassemblies of pol III holoenzyme

| Subunit | Mass (kDa) | Gene | Subassembly | | |
|------------|------------|-------------|--------------------|------------|--------------|
| α | 130* | dnaE | } pol III (core) | } pol III* | } holoenzyme |
| ϵ | 27.5* | dnaQ (mutD) | | | |
| θ | 10 | | | | |
| τ | 71* | dnaX | } γ complex | | |
| γ | 47.5* | dnaX | | | |
| δ | 35 | | | | |
| δ' | 33 | | | | |
| χ | 15 | | | | |
| ψ | 12 | | | | |
| β | 40.6* | dnaN | | | |

* Based on DNA sequence; others are based on electrophoresis.

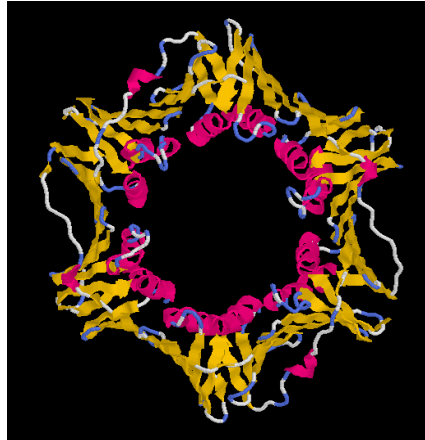
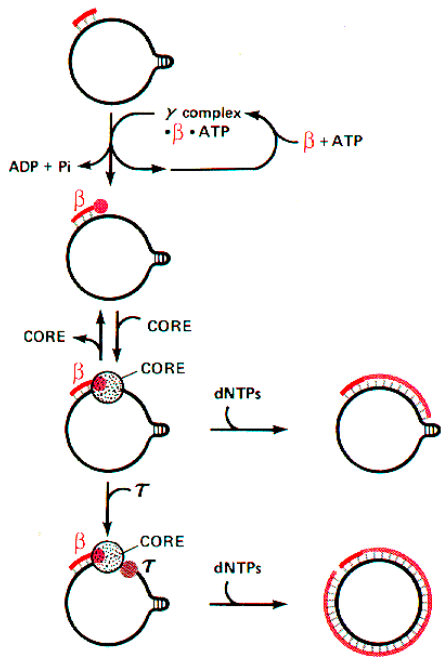
DNA polymerase III is a very complex replication protein consisting of many subunits and subassemblies. The core enzyme consists of three subunits, the α subunit (encoded by dnaE) is the primary synthetic unit, the ϵ subunit encodes the 3' exonuclease and the function of the θ subunit is unknown. The τ subunit by itself expresses no overt exonuclease activity. Only when the τ subunit is associated with the δ subunit does one see activity. Mutations in the δ subunit are of two forms. A temperature sensitive allele is known in dnaQ suggesting the subunit is essential. Another allele is known as mutD, one of the original mutator strains of *Escherichia coli*. This phenotype suggests that the δ subunit may play a role in fidelity of replication.



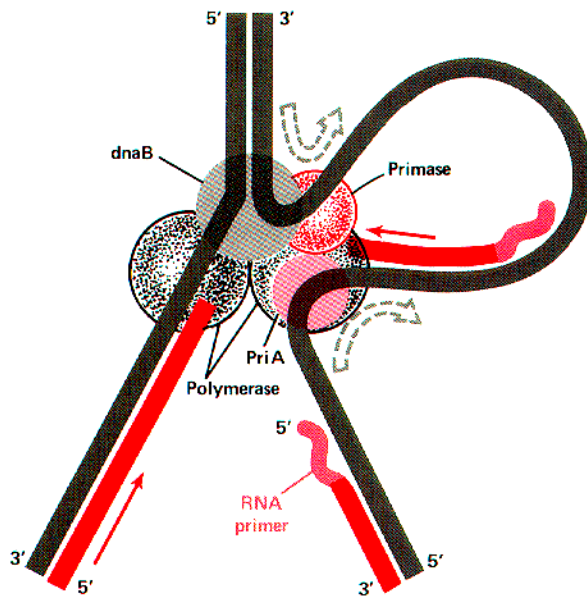
A holoenzyme form of DNA polymerase III appears to contain a dimer of the core enzyme whose formation requires the participation of a τ subunit and ATP. In addition, δ subunits are present in each half of the core dimer that give rise to extreme processivity and speed of polymerase action. In addition to α and ϵ subunits, the enzyme contains a very high molecular weight complex known as the γ complex for the main subunit that holds the complex together. The τ complex, like the δ subunit lends processivity

to the pol III holoenzyme. The τ subunit binds to the core enzyme while the δ subunit binds the γ complex to the core enzyme. The presence of two molecules of core DNA polymerase III in the holoenzyme may allow the holoenzyme to catalyze the synthesis of both leading-strand and lagging-strand DNA simultaneously.

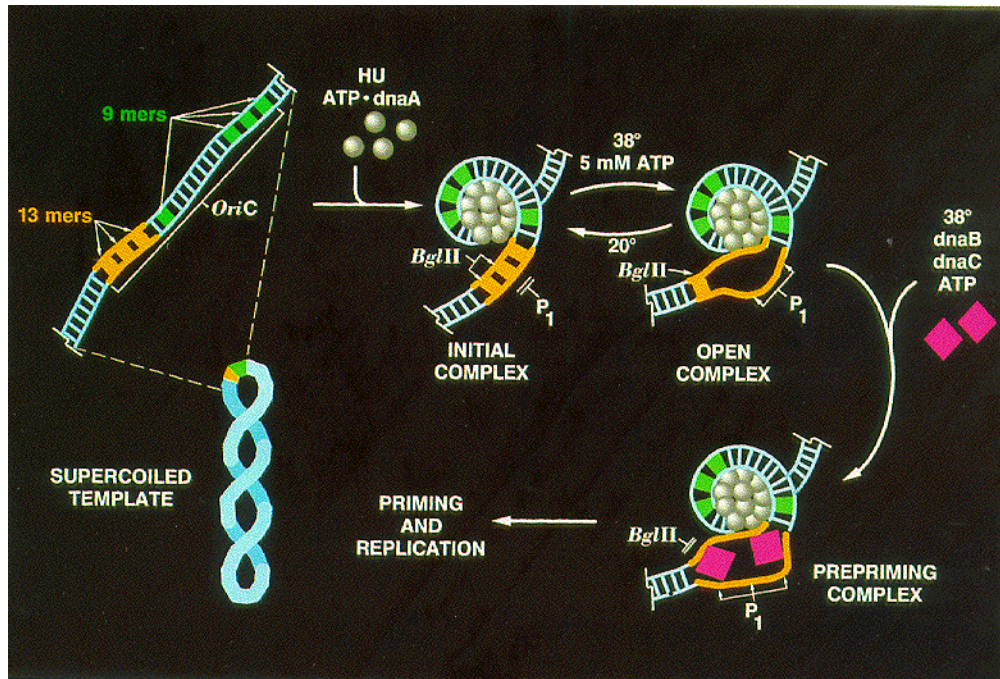
The formation of a highly processive form of the DNA polymerase III requires the action of the γ subunit to link a complex of the β subunits to DNA to which a core will bind. The β subunits form a ring about the DNA (sliding clamp) to which the core DNA polymerase III associates. This sliding clamp ensures that the polymerase stays locked onto the DNA and proceeds in a highly processive fashion as it synthesizes DNA. It may also permit one-dimensional diffusion of the DNA polymerase III along duplex DNA once it has completed filling in a gap on the lagging-strand.



The Replication Fork



An overall model of the replication fork includes a dimer DNA polymerase III holoenzyme bound on the leading-strand with a molecule of the rep helicase associated with the parental DNA strand serving as template. The half of the DNA polymerase III holoenzyme with the complex attached serves to fill in the most proximal gap between Okazaki fragments on the lagging-strand. In order for the dimeric DNA polymerase III holoenzyme to synthesize DNA on both strands simultaneously, the lagging-strand template must bend around 180° to present itself to the enzyme in the same physical direction as the leading-strand. This would give the DNA synthetic activity the appearance of a sewing machine as it acted upon the lagging-strand. The primers for the lagging-strand synthesis are synthesized by a primosome located at the replication fork, perhaps in association with the DNA polymerase III holoenzyme. This primosome may also utilize the helicase activity of dnaB to help unwind the parental DNA strands. The other primers on the discontinuous strand would be synthesized by other primosomes propelled by the PriA activity. SSB would be bound at all single strand gaps in front of both the leading-strand polymerase III and in front of the lagging-strand synthesis.



Initiation of DNA Replication

Analysis of the initiation of DNA replication at an origin was studied in much the same way as the initiation of DNA chains. A specific small molecular substrate containing an origin was designed and then used in an *in vitro* assay for components that would initiate replication on a duplex DNA circle.

A plasmid containing the *Escherichia coli* origin of replication (*oriC*) was selected by ligating fragments of *Escherichia coli* DNA with a drug resistance marker (ampicillin resistance gene) and selecting ampicillin resistance due to autonomously replicating plasmid circles in ampicillin sensitive host cells. The size of the resulting *Escherichia coli* origin fragment was reduced in size by a variety of techniques until the smallest fragment that would support autonomous replication was isolated. This strategy yielded a 245 base pair minimal fragment known as the *oriC* fragment on the first figure.

Comparison of origins from a variety of widely divergent bacterial species revealed:

- 1) The presence of 14 *dam* methylase sites supporting post-replication mismatch repair and guarantying accurate replication of this region. These sites also induce a high state of methylation of the origin.
- 2) The presence of four repeats of a 9 base sequence that is the binding site for the *dnaA* protein, required for chromosomal initiation.

- 3) The presence of a 60 base pair very AT rich region that is highly conserved in evolution.
- 4) The presence of three 13 base repeats in this AT rich region that bind dnaA in the prepriming reaction and melt allowing dnaB helicase to bind.
- 5) Other regions whose sequence is less conserved but which are of fixed length in evolution

Using plasmids with the oriC sequence described above, a reaction permitting initiation and bi-directional DNA replication was purified with the following properties:

- Initiation depends on the plasmid being supercoiled
- Replication starts in the oriC sequence and spreads bidirectionally.
- Initiation depends on transcription of sequences nearby the oriC but the transcripts do NOT serve as primers.
- Because of the need for supercoiled template and RNA transcription, the reaction is sensitive to inhibitors of gyrase and RNA polymerase
- High concentrations of proteins or macromolecules such as polyethylene glycol are required for the reaction to occur.
- High concentrations of ATP and an ATP generating system are required.
- The reaction can only be observed in partially purified fractions due to the presence of numerous inhibitors in the crude lysate.

The steps and components involved in the initiation reaction include:

- Binding of dnaA to the four highly conserved 9-mer sequences. About 20-40 molecules of dnaA (a 52-kilodalton monomer) bind to the right end of the 245-bp oriC sequence forming a large “nucleosome-like” complex. Nearly 200 base pairs are protected from DNase I in this complex.
- ATP is bound in two forms to the dnaA complex, both a high and a low affinity bound form.
- In a second step, an open complex is formed in which dnaA now associates with the remaining 60 base pairs of AT rich DNA including the three 13 base pair repeats. Conditions required for the open complex are the presence of superhelical density >

-0.04, HU protein (or possibly the related IHF - integrative host factor protein), temperature above 21° C

- This second reaction can be inhibited by the presence of another origin specific DNA binding protein IciA that is a helix-turn-helix protein that binds to the 13mers inhibiting their melting.
- The third step is the formation of the prepriming complex. A complex of dnaB and dnaC protein interacts with the open complex (much like it does with the PriA, B, C complex in X SS-->RF).
- dnaB helicase then continues to open the region begun by dnaA using its helicase activity. This opening requires SSB, DNA gyrase activity as well as ATP for the dnaB helicase. Two dnaB complexes are bound at each end of the opening on what will become the lagging-strand of synthesis.
- dnaG (primase) then binds dnaB and primes DNA synthesis by DNA polymerase III holoenzyme at each fork. Although initiated by dnaB-dnaG primase on the “lagging” strand, the first DNA polymerase II holoenzyme quickly proceeds to the opposite fork, becoming the leading-strand enzyme for this opposite fork.

Other factors that influence the rate and specificity of the oriC replication include HU protein, RNA polymerase transcription of nearby regions, ribonuclease H degradation of short non-specific RNA synthesis

Termination of DNA Replication

Termination of DNA replication occurs at specific sites. These terminator sites and proteins that bound to them were first demonstrated in *B. subtilis*, but the reactions in *Escherichia coli* will be discussed here.

The terminator sequences in *Escherichia coli* consist of four repeats of a specific 20 base pair terminator sequence which, when bound with a terminator binding protein impeded the progress of a replication fork primarily by inhibiting the passage of the DNA helicases present at these forks. Two of these four terminator sequences (the most clockwise) are oriented so as to block the replication fork arriving in the clockwise direction. The other two terminator sequences located most counter clockwise are oriented to block the replication fork arriving in the counterclockwise direction. This disposition of sequences permits the short region between the two pairs of terminators to be replicated by either fork, which ever arrives first.

This mechanism for termination has been tested by building plasmids with single copies of the terminator sequences or with two copies in opposite orientation and examining their effects on replication forks progressing from a nearby oriC, either in the presence or the absence of the ter binding protein (TBP). Not only can one demonstrate

that the *ter* sequence, together with the TBP can block a replication fork in an orientation specific manner, they can also be demonstrated to block all three major DNA helicases present in *Escherichia coli* including *dnaB*, *rep* and *uvrD* helicase.

References

- Asai, T. and Kogoma, T. (1994). D-loops and R-loops: alternative mechanisms for the initiation of chromosome replication in *Escherichia coli*. *J Bacteriol* **176** (7), 1807-12.
- Baker, T. A. (1995). Replication arrest. *Cell* **80** (4), 521-4.
- Beckman, R. A. and Loeb, L. A. (1993). Multi-stage proofreading in DNA replication. *Q Rev Biophys* **26** (3), 225-331.
- Bierne, H. and Michel, B. (1994). When replication forks stop. *Mol Microbiol* **13** (1), 17-23.
- Bramhill, D. and Kornberg, A. (1988). A model for Initiation at Origins of DNA Replication. *Cell* **54**, 915-918.
- Crooke, E., Hwang, D. S., Skarstad, K., Thony, B. and Kornberg, A. (1991). *E. coli* minichromosome replication: regulation of initiation at *oriC*. *Res Microbiol* **142** (2-3), 127-30.
- Donachie, W. D. and Begg, K. (1990). Genes and the replication cycle of *Escherichia coli*. *Res Microbiol* **141** (1), 64-75.
- Echols, H. and Goodman, M. F. (1991). Fidelity mechanisms in DNA replication. *Annu Rev Biochem* **60** (477), 477-511.
- Goodman, M. F., Creighton, S., Bloom, L. B. and Petruska, J. (1993). Biochemical basis of DNA replication fidelity. *Crit Rev Biochem Mol Biol* **28** (2), 83-126.
- Kelman, Z. and O'Donnell, M. (1994). DNA replication: enzymology and mechanisms. *Curr Opin Genet Dev* **4** (2), 185-95.**
- Kornberg, A. (1988). DNA Replication. *J. Biol. Chem.* **263** (1), 1-4.
- Kornberg, A. (1991). Control of initiation of the *Escherichia coli* chromosome. *Cold Spring Harb Symp Quant Biol* **56**, 275-8.
- Kuriyan, J. and O'Donnell, M. (1993). Sliding clamps of DNA polymerases. *J Mol Biol* **234** (4), 915-25.

Losick, Richard and Shapiro, Lucy (1998) DNA Replication: Bringing the Mountain to Mohammed. *Science* 282, pp. 1430 - 1431.

McHenry, C. S. (1991). DNA polymerase III holoenzyme. Components, structure, and mechanism of a true replicative complex. *J Biol Chem* **266** (29), 19127-30.

O'Donnell, M. (1992). Accessory protein function in the DNA polymerase III holoenzyme from *E. coli*. *Bioessays* 14 (2), 105-11.

O'Donnell, M. (1994). Beta sliding clamp dynamics within *E. coli* DNA polymerase III holoenzyme. *Ann N Y Acad Sci* **726** , 144-53; discussion 153-5.

O'Donnell, M., Kuriyan, J., Kong, X. P., Stukenberg, P. T. and Onrust, R. (1992). The sliding clamp of DNA polymerase III holoenzyme encircles DNA. *Mol Biol Cell* **3** (9), 953-7.