Relations Between Replication and Transcription

Daniel Castro-Roa and Nikolay Zenkin
Centre for Bacterial Cell Biology, Institute for Cell and Molecular Biosciences
Newcastle University, Newcastle upon Tyne
United Kingdom

1. Introduction

In the cell, RNA polymerase (RNAP) and the replisome share the same template DNA for their respective functions. The rate of replication is typically 20 times faster than transcription and six or more replication forks may be present at the same time on the chromosome (Gotta et al. 1991; Hirose et al. 1983). This implies that collisions between the two machineries are inevitable. In this chapter we will discuss the recent findings on the outcome of collisions between transcription and replication and their consequences, as well as cooperation between the two machineries.

2. Stability of RNA polymerase complexes with nucleic acids

In all living organisms transcription, the first step in gene expression, is accomplished by multisubunit RNAP. RNAP is highly evolutionary conserved, both structurally and functionally, in all three domains of life. Although there are differences in the mechanisms of initiation and regulation of transcription, mechanisms of catalysis are remarkably similar in all living organisms. Fully functional eukaryotic and archeal models involve 12-14 subunits (depending on the polymerase type and organism) with a total molecular weight greater than 500 kDa (Cramer et al. 2001). The simplified versions found in bacteria are composed of five subunits (subunits composition 2α, β', β and ω) with a molecular mass of approximately 400 kDa. Transcription is a cyclic pathway roughly composed of three steps: initiation, elongation and termination. RNAP is a flexible machine able to adopt different states required for various transcription stages and the mode and stability of binding of the enzyme to DNA at these stages are different. Various modes of RNAP binding to nucleic acids throughout the transcription cycle is one of the key factors which determine the fate of the replication fork progression along the DNA.

First, we briefly describe steps of transcription cycle and properties of complexes formed by RNAP with the nucleic acids at these steps.

The first stage of transcription is initiation. This phase begins with the search of the promoter by the core enzyme equipped with the transcription initiation σ factor through a scanning mechanism (Park et al. 1982; von Hippel 2007). The means by which RNAP is thought to find the promoter involves the tracking of a groove of the double helix throughout electrostatic interactions reinforced by the entropy that results from the

displacement of the counterion cloud that surrounds the DNA (Sakata-Sogawa et al. 2004). These initial interactions of the enzyme with DNA are weak and therefore unstable.

σ subunit recognizes a promoter, which usually is comprised of two hexameric sequences around -10 and -35 positions relative to transcription start site, to which it has high affinity. It has been demonstrated that promoter sequences are highly flexible making them more prone to RNAP binding (Ozoline et al. 1999; Travers 2004). The interactions of RNAP with promoter DNA are expanded by the wrapping of the DNA (demonstrated by DNA footprinting experiments) on the surface of the enzyme making the complex more stable (Ozoline et al. 1995). Wrapping also facilitates further rearrangements in both, RNAP and DNA. During this state, known as promoter closed complex (Li et al. 1998), the DNA helix remains double stranded and the complex, though being relatively stable, remains sensitive to high ionic strength and competitors such as heparin (Coulombe et al. 1999). The stability of this complex depends on the sequence of the promoter. Generally, the farther the sequence of a promoter (-10 and -35 elements) from the consensus, the less stable promoter complex will be formed on it (Fenton et al. 2001). Promoter sequence also determines the capacity of RNAP to compete with repressors and nucleoid proteins (Grainger et al. 2006).

After formation of the closed complex the double helix of DNA is destabilised by action of specific residues in the σ factor on a precise region of the promoter (Aiyar et al. 1994; deHaseth et al. 1995; Murakami et al. 2002). Then, the enzyme melts the double helix of the DNA and form a stretch ~17 nucleotides (nt) of unwound DNA known as the transcription bubble (for some particular σ factors, the energy required is obtained by ATP hydrolysis (Merrick 1993)). The melting generates further rearrangements of the DNA inside RNAP, placing the downstream DNA into the enzyme's DNA-binding clamp and positioning the template DNA in register with the active centre making the complex catalytically competent (Murakami et al. 2002; Vassylyev et al. 2002). This new configuration of RNAP is referred to as open promoter complex (Li et al. 1998; Mekler et al. 2002). At this point RNAP occupies a total of ~35 bp of the DNA and has undergone several structural rearrangements that provide higher stability compared to afore mentioned closed promoter complexes. Open promoter complex is capable to withstand higher ionic strength (200 mM KCl) and becomes resistant to competitors (Reppas et al. 2006; von Hippel et al. 1984).

Initiation of transcription starts with synthesis of short RNA transcripts (2-9 nucleotides long). Given that RNAP remains anchored to the promoter by σ factor, synthesis involves pulling of a stretch of downstream DNA of the same size inside the main channel of RNAP. This phenomenon is referred to as "scrunching". Scrunching results in increase of the size of the transcription bubble (given that its upstream edge is kept at the same position by σ subunit). Initiation ends when the energy accumulated during "scrunching" surpasses the energy that is anchoring RNAP to the promoter releasing the enzyme from it followed by the subsequent ejection of the σ factor leaving RNAP ready to elongate the RNA chain.

The stability of RNAP during elongation is greatly enhanced by multiple protein-nucleic acids contacts that are formed after synthesis of ~8 nucleotide long RNA and σ subunit release. In simple terms, if using a comparison of RNAP to a crab claw, the claw in initiation is more open, while in elongation it closes on the RNA-DNA hybrid almost fully surrounding it. These interactions of RNAP with nucleic acids make elongation complex highly stable and resistant to very high ionic strength (1 M KCl) and to competitors (Kuznedelov et al. 2002). The structure of the elongation complex is schematically shown in Figure 1. The characteristic feature of the elongation complex, that plays major role in its

stability, is the 9-10 base pair RNA-DNA hybrid. The length of the hybrid as well as the size of the transcription bubble remain the same throughout elongation.
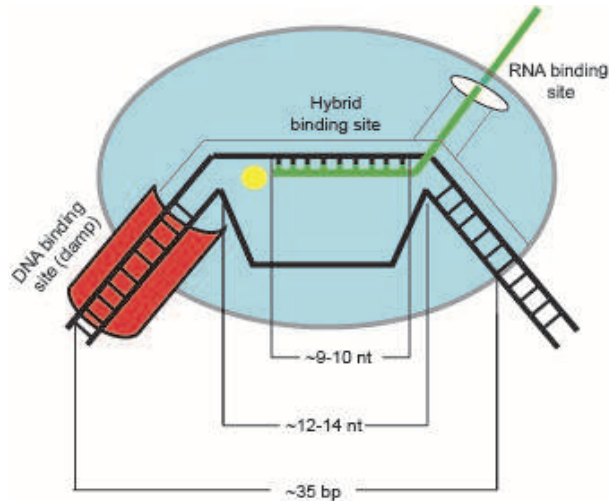

Fig. 1. Schematic representation of RNAP core enzyme (blue) during elongation Catalytic Mg^{2+} ions of active centre are shown as yellow circle. DNA is black, RNA is green. The three major interaction sites with nucleic acids that provide stability to the elongation complex and their length are shown.

Besides unusual stability of the elongation complex, transcribing RNAP generates a considerable force. The calculated force of the actively transcribing RNAP is 20 pN per molecule of RNAP and is additive when more RNAP molecules collide with each other (Wang et al. 1998). This allows RNAP to overcome unwanted events and also dislodge proteins bound to DNA template (Epshtein et al. 2003). This powerful and stably bound to DNA machine is capable to proceed through millions (in eukaryotes) of base pairs without losing contact with the DNA template and nascent RNA.

Though RNAP is supposed to processively transcribe through long distances without interruption it is a subject to stringent regulation. RNAP recognizes different signals in the DNA that slow down its movement via various mechanisms. Some pause signals lead to structural modification of RNAP active centre, which slows down catalysis of the phosphodiester bond (Landick 2009). An unusual property of RNAP that also plays role in pausing is its ability to move backwards along the DNA template, a phenomenon called backtracking. During backtracking, RNAP shifts backwards in a manner of a zipper: the 3' end of RNA disengages from the template DNA strand and the active centre, while the rear end of RNA-DNA hybrid RNA anneals back to the template. This keeps the length of the RNA-DNA hybrid the same as in active elongation complex, which means that this arrested complex is as stable as the active one. These arrested complexes require separate factors (e.g. cleavage Gre factors, transcription-DNA repair coupling factor Mfd (Borukhov et al. 1993; Park et al. 2002)) for their resolution. Pausing and backtracking increase the probability of RNA polymerase encountering the replisome, and may be detrimental for genome integrity and cell viability if not resolved.

When RNAP encounters a termination sequence (typically a GC rich palindromic sequence followed by a stretch of uridines ~8nt long) it forms a weak RNA:DNA hybrid containing the poly U track of the terminator (Gusarov et al. 1999). This pauses (by backtracking mechanism) RNAP and gives enough time for a hairpin loop to form in the RNA exit channel followed by destabilization of the RNA:DNA hybrid leading to RNA release and aperture of the DNA binding clamp resulting in RNAP dissociation (Epshtein et al. 2007; Santangelo et al. 2004)

3. Interactions of transcription with replication

3.1 Collisions of RNAP and the replisome

The effect on collisions between RNAP and the replisome depends greatly in their directionality. Co-directional interactions occur in the leading strand whereas head-on collisions take place on the lagging strand (Figure. 2). Interestingly, analyses of genome organisation have shown that most of the essential genes, highly transcribed genes, and longer genes (Huvet et al. 2007; Omont et al. 2004; Price et al. 2005; Rocha et al. 2003) are oriented in the same direction as replication on the leading strand. This arrangement was also observed in *B. subtilis, Borrelia burgdorferi, Treponema pallidum, Haemophilus influenza, Helicobatcter pylori, Mycoplasma genitalium, mycoplasma pnumoniae*, and in some bacteriophages (McLean et al. 1998). The difference in gene arrangement is considered to be the outcome of natural selection on genome organisation (Mirkin et al. 2005). Although the evolution pressure that resulted in this organisation is still unclear, it is thought to be, at least in part, determined by differences of the interactions of RNAP with the replication fork during co-directional versus head-on collisions (Brewer 1988). This could be detrimental for RNAP completion of transcripts affecting the production of correct, full length proteins or most likely, because it could stall the replisome inhibiting the replication fork progression (Deng et al. 2005; Srivatsan et al. 2010). These mechanisms will be discussed later.

3.1.1 The effects of RNAP on replication and genome integrity

The manner in which replication forks compete with RNAP has been evaluated *in vivo* and *in vitro*. Unsurprisingly, most of the research has been performed in *B. subtilis* and *E. coli*. Early studies in *E. coli* analysed the outcome of the collision between the two machineries by inserting in the chromosome an IPTG inducible and unidirectional ColE1 replication origin upstream (for analysis of co-directional collisions with RNAP) or downstream (for analysis of head on collisions with RNAP) of the ribosomal operon *rrnB*.

This particular operon (5.4 kb) was chosen because of its length and high rate of transcription would make it easier to be visualised using electron microscopy (French 1992).

In the case of co-directional orientation, replication fork progression was the same in both non-transcribed and transcribed regions where no accumulation of the replisome was observed. Interestingly, during replication of the transcribed regions RNAP molecules were absent behind the replication fork but were still present in front of it. Additionally, after the replisome had completely passed through the transcribed region repopulation by RNAP took place. These observations suggested that the replisome was not neither slowed down nor displaced from the DNA by transcribing RNAP (Figure 3).

This observation is consistent with recent *in vitro* work which showed that after co-directional collision, RNAP is dislodged from the DNA and the replication fork resumes elongation by using the displaced RNA as a primer (Pomerantz et al. 2008). The mechanism in which RNAP acts as primase will be discussed later.
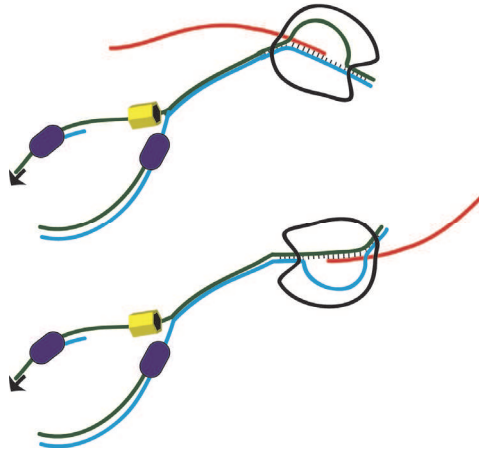
Fig. 2. Schematic representation of Co-directional and head-on collision between the replication fork and RNAP. RNAP with its characteristic claw-like shape is shown in black (right part) with nascent RNA (red line) and template (green) and non-template (blue) strands. The replisome (on the left) is represented by yellow 3D hexagon (DnaB) and DNA polymerases as purple ovals in lagging and leading strands respectively.

In the case of head-on collisions, RNAP was also dislodged but the replication fork progressed through the operon much more slowly than in the co-directional counterpart.
In other studies using DNA microarrays of wild type *B. subtilis* and of mutants carrying a long stretch of inverted DNA in the chromosome (to induce head-on collisions) the reduction of the speed of the replisome was also observed. In the wild type scenario, no interference of the fork progression by transcriptional machinery was detected. Surprisingly, in the case of the inverted mutants, replication was generally slowed down in both highly and low transcribed regions reinforcing the theory that the genome is organised to favour the co-directionality of both machineries (Wang et al. 2007).
The slowing down of the replication fork caused by head-on collisions could result in genomic instability due to inaccurate restart of the replisome by new collisions with transcribing RNAP molecules. Another possibility that could lead to instability is that, after collision, stalled replication fork could unwound generating a four stranded DNA structure that resembles a Holliday junction (a process known as replication fork regression) (Atkinson et al. 2009). This could lead to double strand breaks by cleavage of the DNA performed by Holiday junction resolvase affecting cell fitness and viability (McGlynn et al. 2000). In this study, performed in *E. coli* it was proposed that halted RNAP represents an impediment for replication fork progression and upon collisions it may require to restart. To reach this conclusion, cells were irradiated with UV light damaging the DNA. Transcribing RNAP cannot surpass the lesions on the DNA and therefore it stalls. Replication was observed to restart upon encountering the stalled RNAP which was assisted by endonucleolytic cleavage of the DNA by the Holiday junction resolvase RuvABC. Mutants lacking the RuvABC system, in the presence of high concentrations of the alarmone ppGpp were capable to survive UV irradiation demonstrating that few events of replication restart were occurring. This finding suggests that transcription factors are also involved in the resolution of conflicts between the two machineries.
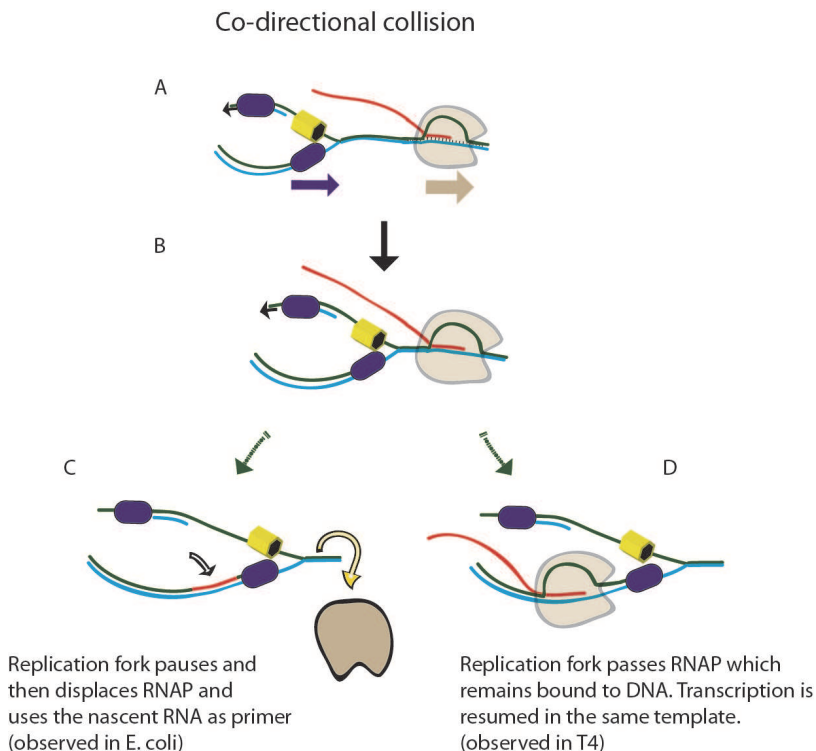
Co-directional collision

Fig. 3. Schematic representation of the observed outcome of the co-directional collision between the replication fork and RNAP. A) Machineries with arrows indicating their directionality. B) Replication fork colliding with RNAP. C) Replication fork displaces RNAP and utilises the newly synthesised RNA as primer. D) RNAP is temporarily displaced from the DNA allowing the replication fork to pass by. Then, transcription is resumed in the same template.

The mechanism in which the replisome is slowed down could be explained by an observation made *in vitro* with RNAP and replication system purified from *E. coli*. It was demonstrated that during head-on collisions RNAP was also displaced from the DNA, as in the case of co-directional collisions, but the replisome was stalled, though it did not fall apart and eventually was capable to resume elongation (Pomerantz et al. 2010). Earlier studies with plasmids suggested (Mirkin et al. 2005) that the outcome of head-on collision is due to the actual physical interactions between the two machineries (Figure 4). Another possible explanation for this phenomenon was that the topology of the DNA (positive supercoiling generated in front of the transcribing RNAP) could be the cause of the decrease in rate of replication (Brewer 1988; Deshpande et al. 1996).

Besides studying the interactions of actively elongating RNAP, different strategies were used to investigate the effects of initiating and terminating RNAP on the replication fork. In *E. coli*, an *in vivo* system was set up using two plasmids. One contained the strong T7A1 promoter where the equilibrium from promoter clearance towards abortive initiation was shifted by modifying the initial transcribed sequence. This would allow RNAP to be stably

bound to the promoter without entering into elongation. When challenged with replication in co-directional orientation, it was not detrimental for the replisome but during head-on collisions it turned out to be inhibitory. The second plasmid contained the same promoter but in this case RNAP was allowed to enter into elongation normally. Further downstream a transcription termination signal was placed. Interestingly, the replisome was stalled at co-oriented collisions with the terminating RNAP which, as mentioned before, could be in the backtracked state that occurs prior to termination. It is possible that the resulting effect upon encountering promoters and terminators might serve as polar "replication punctuation marks" that could facilitate clearing of mutations acquired in transcribed parts of the genome, by the mismatch repair or gene conversion machineries (Mirkin et al. 2006).
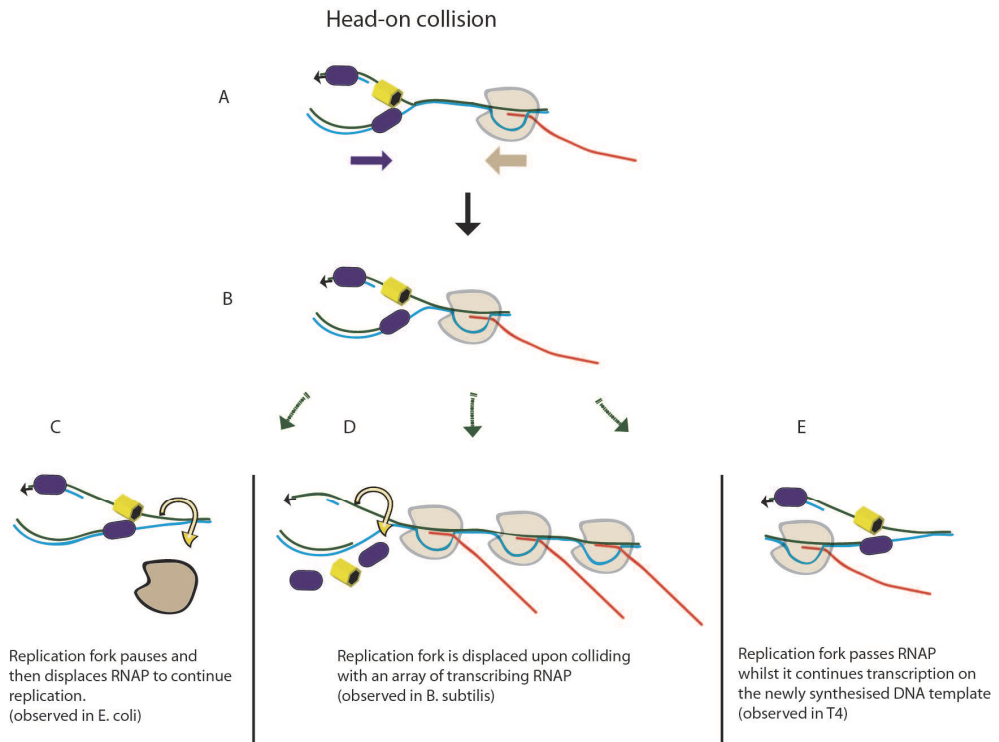

Fig. 4. Schematic representation of the observed outcomes of the head-on collision between the replication fork and RNAP. A) Machineries with arrows indicating their directionality. B) Replication fork colliding with RNAP. C) Replication fork briefly pauses and then displaces RNAP from the DNA D) Replication fork collapses after colliding with an array of RNAP. E) RNAP is temporarily dislodged from the DNA by the replisome but it is still bound to the DNA. After the replication fork has passed, transcription is resumed in the new DNA template

In contrast to the observations made with *E. coli*, studies conducted *in vitro* with reconstituted components using the bacteriophage T4 as a model showed that progression of the replication fork was paused regardless of the direction of the collision with longer stalling during head-on encounters (Liu et al. 1995). However, it was suggested that after co-

directional collisions RNAP was still bound to the DNA as the replication passed it by and was capable to resume transcription in the same strand. It was also postulated that during head-on collisions, RNAP would continue transcribing but both the enzyme and the transcript would switch strand, from the already existing (where the collision takes place) to the newly synthesised one (Liu et al. 1993).

Although the mechanism in which this strand swapping might occur is still unclear it was proposed that RNAP and transcript transiently disengage from the template but remained bound to the DNA and an active elongation complex reforms after the replisome has passed through.

Most recent studies monitoring the association of the replicative helicase and replication start proteins by ChIP, ChIP-qPCR and ChIP-chip also observed that collision leads to replication restart *in vivo* in *B. subtilis* in highly transcribed genes both co-directional and head-on collisions (Merrikh et al. 2011).

Analysis of the rolling circle M13 replication of plasmids in head-on orientation towards RNAP revealed a significant loss of plasmid stability along with deletions in and downstream the transcribed region. Notably, it was also observed that replication of the parental plasmid decreased (due to interference of the replisome by RNAP) leading to an enrichment of the plasmids that had undergone deletions (Vilette et al. 1996).

Mutations as result of head-on collisions between the replisome and transcription have also been reported for bacterial chromosomal DNA. In *B. subtilis*, analysis of the *rpoB* gene (which encodes for the RNAP β subunit) positioned in the lagging strand, head-on towards replication fork progression, showed an increased mutation rate (~3 fold) compared to the isogenic control with no inversion (Wang et al. 2007). Since the observation was done in one of the longest transcriptional units, the study also suggested that replication fork progression is affected more dramatically when not just a single RNAP but an array of RNAP molecules are transcribing the gene where the collision takes place. This supports the hypothesis that states that long genes are arranged in a co-directional manner to avoid this sort of events (Omont et al. 2004). In the same study major disruption of the replisome was detected on ribosomal genes orientated head-on to the replication fork by following RecA filaments which indicate generation of single stranded DNA or double stranded DNA ends.

In eukaryotic cells, even though the rate of replication is similar to the rate of transcription and there are specific barriers that block the replisome from entering ribosomal genes (Brewer et al. 1988), interactions of the two machineries have also been characterised. In *Saccharomyces cereviseae* by genome wide analysis it was shown that replication fork is paused by RNAP II in highly transcribed genes (Azvolinsky et al. 2009).

Other studies on yeast showed that genome integrity is compromised by impairment of the replisome during head-on collisions with RNAP II which might promote transcription associated recombination (Kim et al. 2007; Prado et al. 2005).

All these findings relate directly to the mechanisms of overcoming obstacles by the two machineries with major emphasis on the outcome of the replisome in terms of its integrity and stability that clearly is affected depending on the directionality of the collisions. However there are other direct or indirect roles played by RNAP in replication.

4. Effects of transcription on replication initiation and completion

DNA replication is normally initiated from a defined position on the chromosome, the origin, when assembly of the replication complex takes place. There is evidence that

transcription might be involved in the regulation of the initiation of replication. After unwinding of the DNA duplex, DNA polymerases (DNAP) cannot initiate the synthesis of a new DNA chain, without pre-existing primers. Due to the 5'→ 3' directionality of the DNAP one of the strands is polymerised continuously (leading strand) and the other one in a discontinuous manner (lagging strand). The leading strand needs to be primed at the beginning, whereas the lagging strand requires being primed frequently to generate short chains (Okazaki fragments) until its completion (Masai et al. 1996). This process is achieved by a special primase, which is part of the replication machinery, but action of RNAP has also been described in the priming of replication (Hassan et al. 1994).

The mechanisms in which the transcriptional machinery is involved in replication initiation and priming have been studied over the past 30 years and will be discussed below.

4.1 Transcription of mioC and its effect on replication initiation

The replication origin oriC is a highly conserved region in enteric bacteria. In its ~245 bp it contains five *dnaA* binding sites (Kornberg 1991; Messer et al. 2001) and the mioC gene which contains no apparent transcriptional terminator which allows most transcripts to read through the oriC region (Nozaki et al. 1988; Schauzu et al. 1987). First indication of RNAP involvement in initiation was obtained in early experiments in *E. coli* which demonstrated that replication was sensitive to rifampicin (which inhibits RNAP) independently from the effects on protein synthesis (Lark 1972; Messer 1972). Temperature sensitive mutants of DnaA are lethal at the non-permissive temperature. Interestingly, this effect is suppressed by specific mutations on *rpoB* (Atlung 1984; Bagdasarian et al. 1977). Also, an increased copy number of oriC containing chromosomes was observed in *rpoB* or *rpoC* mutants (Rasmussen et al. 1983). This suggested that modified RNAP could assist replication initiation but the actual possible mechanism remained obscure.

In vitro experiments of the replication from oriC showed that RNAP is required under specific conditions that affect DNA unwinding such as reduced negative superhelicity or reduced temperature (Baker et al. 1988). Similar to the case of phage λ replication where transcription from the P$_R$ promoter induces opening of the DNA under the same conditions (Keppel 1988). It was also suggested that RNAP involvement was not due to priming activity because transcription terminated by incorporation of 3'-dATP had no effect on replication initiation, meaning that there is no need for the transcript to cross the oriC sequence. (Baker et al. 1988). Contrary to these observations, when primase was omitted from the reactions in the presence of RNAP, the replication efficiency was reduced tenfold but not inhibited contemplating the possibility that the transcripts could have been used as primers (Ogawa et al. 1985).

Further analysis showed that transcription originated from the mioC promoter enters the oriC region and seems to be involved in control of initiation of replication (deWind 1987). A study revealed that the copy number of minichromosomes in the presence of mioC transcription, was 2 fold higher compared to the mioC deletion control. An apparently contradicting observation was made when the wild type mioC promoter was replaced by stronger lac or Cm^r promoters. Stronger transcription that entered oriC resulted in inhibition of minichromosome replication (Tanaka et al. 1983). Other experiments also performed in minichromosomes demonstrated that DnaA inhibits transcription from mioC promoter by binding to the dnaA box which is located right upstream of it (Lother et al. 1985; Stuitje et al. 1986).

Notably, in a recent study on *E. coli* chromosome, a mioC gene mutant in which TN5 transposon was inserted in the promoter region was isolated. This mutant allowed mioC transcription constitutively on the chromosome throughout the cell cycle. It was shown that the read-through of mutant mioC through oriC impaired the initiation of chromosomal replication and also suggested that the transcriptional fluctuation during the cell cycle might be essential for the control of initiation of replication (Su'etsugu et al. 2003). This was in agreement with observations made in minichromosomes where transcription of mioC gene appeared to be repressed prior to the initiation of replication and derepressed right after initiation of replication (Ogawa et al. 1994; Theisen et al. 1993). The proposed mechanisms by which transcription from mioC could inhibit replication were that transcription could impair formation of initiation complex at the origin via physical interactions which would displace DnaA from DNA or through the generation of local changes in the DNA topology destabilizing the initiation complex. However, while the replication time in cells lacking mioC promoter was unaffected (Bates et al. 1997; Lobner-Olesen et al. 1992), transcription from this promoter became essential for cell viability when the dnaA box R4 was deleted suggesting a positive role for mioC transcription in replication initiation (Bates et al. 1997).

Despite of all the data obtained over 3 decades, no clear consensus on the role of transcription from the mioC gene on replication has been achieved. The lack of consensus in the observations among these studies could be due to the differences in the plasmid structures that could result in different transcription modes which affect the effect on replication (Masai et al. 1988). Further investigation is required to elucidate the mechanisms of regulation of the two machineries.

4.2 The role of RNAP as primase

As mentioned in the previous section, replication initiation and elongation of the lagging strand needs to be primed. DNAP is incapable of initiating DNA synthesis *de novo* and require an accessible and properly positioned hydroxyl group to attach a deoxyribonucleotide. Such insufficiency might have emerged through evolution to increase processivity of these enzymes by impairing the sequence specific interactions with template, which are needed for specific initiation of *de novo* synthesis (Zenkin et al. 2008). In the majority of cases DNA replication is primed by enzymes called primases. Primases are DNA-dependent RNA polymerases that synthesise oligoribonucleotides that remain annealed to template DNA and are used by DNAP as primers. Some other strategies have been acquired by different replicons to fulfil these requirements. For example, preformed tRNA (by retroviral reverse transcriptases), DNA primers generated by endonucleolytic cleavage (by gpA of øX174), and serine OH group (adenoviral 55 kDa terminal protein). Strikingly, some replicons rely on RNAP and not on primases for initiation of replication of their genomes. This event was first observed 30 years ago in reconstituted replication system by A. Kornberg (Brutlag et al. 1971; Wickner et al. 1972). Before the discovery of DnaG (Bouche et al. 1975; Rowen et al. 1978; Schekman et al. 1972) RNAP was thought to be the only primase in the cell.

RNAP priming activity is directly utilised in the replication of the leading strand of filamentous and T-odd phages. The filamentous phages, such as M13, fd and f1, contain a simple, relatively small (6407 nt long), single stranded circular genome which contains 10 genes (Herrmann et al. 1978; Ohsumi et al. 1978). An important feature of the genome is an intergenic region that takes up to 8% of the genome. This region does not encode proteins but serves as the origin for the + and - strand synthesis (Chen et al. 1978). This region contains five

palindromes (Zinder et al. 1985). After infection, adsorption and penetration, the replication of the single stranded genome starts. RNAP equipped with σ^{70} binds to one of the hairpins (which has no homology to any *E. coli* promoter) in the (+) strand of the intergenic region. Though RNAP bound to this region protects ~125 nt, only ~20 bp long hairpin is enough for specific recognition. RNAP synthesises a short (18 nt long) RNA chain (Zenkin et al. 2006). This RNA serves as a primer and is further extended by DNAP (Kaguni et al. 1982).

The mechanism in which RNAP synthesises the M13 replication primer has been recently elucidated (Figure 5). It was observed that the minimal M13 origin occupies the downstream DNA binding channel and positions the single-stranded template in the active centre of RNAP leaving the enzyme ready to initiate synthesis of RNA (this is very unusual mode of initiation as compared with "normal" transcription where, upon RNAP binding, the double stranded DNA is melted and abortive synthesis of RNA occurs until the enzyme has stabilised both the transcript and the DNA template in the active centre. This series of events allows further synthesis of RNA). RNAP lacking the σ subunit (core enzyme), was observed to be capable to specifically initiate RNA synthesis on the M13 *ori* hairpin. The synthesis of a dinucleotide suggested that the stabilisation of the initiation complex was achieved through "structure-specific" interactions of the downstream DNA binding channel of RNAP with the *ori* hairpin. However, no synthesis beyond the first phosphodiester bond was observed. When RNAP containing σ subunit was used, stabilisation of short transcripts was achieved, allowing further extension of the RNA. σ factor was shown to be needed at least until a trinucleotide was formed. When the growing RNA:DNA hybrid reached the size of 10-12 bases, steric constrains lead to the disengagement of the rear part of the hybrid from the active cleft and leads to formation of overextended hybrid. Upon synthesis of an 18-20 bases long RNA irreversible structural rearrangement is induced that leads to accommodation of the hybrid in the downstream DNA binding channel. After the rearrangement, the exposed 3′ end of the RNA becomes available to be used by DNAP (Zenkin et al. 2008). It is still unclear if DNAP recognises the 3′ end of the RNA in complex with RNAP or if RNAP is displaced before initiation of DNA synthesis.
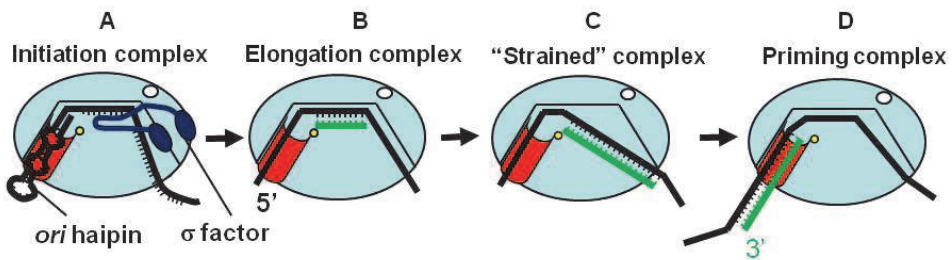

Fig. 5. Schematic representation of the mechanism of primer synthesis by RNAP. A) RNAP (blue) accommodates the *ori* hairpin (black) in the downstream DNA binding channel (red). Synthesis of RNA starts and the transcript is stabilised by σ factor (purple). B) Synthesis of an RNA (green) 8-9 nt long, proceeds in a normal way. C) Upon reaching the RNA:DNA hybrid of the size of 10-12 nt long, the rear end of the hybrid disengages from the active cleft D) Upon reaching critical length of 18-20 nt, RNAP backslides to reposition the hybrid in the downstream DNA binding channel making the 3′ end of RNA available to serve as primer for replication.

RNAP synthesised RNA that promote replication have also been characterised in the *E. coli* double stranded plasmid ColE1. As observed *in vitro*, initiation of replication of the continuous strand depends on three host enzymes: RNAP, RNase H and DNAP I. It was shown that if RNAP was inhibited by rifampicin replication of the plasmid would not occur (Itoh et al. 1980; Itoh et al. 1982). RNAP was shown to synthesise a 555 nt transcript which enters a region where the transition between RNA and DNA occurs (Selzer et al. 1982). The transcript, called RNA II contains a series of stems and loops which are necessary for the formation of an extended RNA:DNA hybrid at the end of the transcript (Masai et al. 1996). RNA II is tightly regulated by the antisense RNA I which binds to the 5′ of RNA II affecting its secondary structure inhibiting the formation of the RNA:DNA hybrid (Tamm et al. 1983). Cleavage of RNA:DNA hybrid mediated by RNase H exposes the 3′-OH ends needed for DNA synthesis by DNAP I (Naito et al. 1984) which after polymerisation of 400 nt is replaced by the highly processive DNAP III which will continue the replication of the strand. 5′ end secondary structure is required to abolish reformation of upstream DNA duplex behind RNAP, thus forcing RNAP to transcribe single stranded template. Therefore, the mechanism of primer formation may be the same as for single stranded M13 phage.

In a recent study conducted *in vitro* with a reconstituted *E. coli* replisome lacking primase, the transcriptional machineries from both *E. coli* and T7 RNA polymerases were challenged with the replication fork progressing in a co-directional manner (Pomerantz et al. 2008). Replication was initiated on a forked DNA template. The progression of replication was followed by the analysis of the radioactively labeled nascent DNA strand. RNAP, (either *E. coli* or T7) elongation complex was stalled in front of replisome by the omission of one of the 4 NTPs. Strikingly, after displacement of RNAP from the template, the nascent transcript apparently remained annealed to DNA template and was used by replisome as a primer. This can explain the outcomes of co-directional collisions observed *in vivo* (see above). The mechanism of primer formation in this case may also be similar to M13, given that the replisome approaching RNAP from behind restricts the upstream DNA duplex formation resulting in RNAP synthesising on single stranded DNA.

In eukaryotic cells evidence of RNAP acting as a specific primase which initiates the synthesis of the heavy-strand and the light-strand of the double stranded 16.6 kb long mitochondrial DNA has also been observed (Fuste et al. 2010; Wanrooij et al. 2008).

5. Conclusions

Altogether, through the data generated in four decades, it has become evident that transcription has played a major role in the evolution of genome organisation, architecture and integrity. Even though the mechanisms and the outcomes of the collisions of the replisome with RNAP, the regulatory function of transcription on replication initiation and its contribution in elongation by serving as primase have been demonstrated, many other questions remain unanswered. Further studies on the mechanisms in which mioC acts on oriC replication are needed. Also, the mechanism(s) of action of transcription factors such as DksA and the alarmone ppGpp modulate not only the activity of RNAP but also its interactions with replication and other machineries require to be properly elucidated. The phenomenon of template swapping by RNAP upon head-on collisions and the by-passing of the replication fork in co-directional collisions with the T4 replication machinery remains obscure and needs to be clarified.

6. References

Aiyar, S. E., et al. (1994). "Mutations in sigma factor that affect the temperature dependence of transcription from a promoter, but not from a mismatch bubble in double-stranded DNA." *Biochemistry* 33(38): 11501-11506.

Atkinson, J., et al. (2009). "Replication fork reversal and the maintenance of genome stability." *Nucleic acids research* 37(11): 3475-3492.

Atlung, T. (1984). "Allele-specific suppression of dnaA(Ts) mutations by rpoB mutations in Escherichia coli." *Molecular & general genetics : MGG* 197(1): 125-128.

Azvolinsky, A., et al. (2009). "Highly transcribed RNA polymerase II genes are impediments to replication fork progression in Saccharomyces cerevisiae." *Molecular cell* 34(6): 722-734.

Bagdasarian, M. M., et al. (1977). "Suppression of the DnaA phenotype by mutations in the rpoB cistron of ribonucleic acid polymerase in Salmonella typhimurium and Escherichia coli." *Journal of bacteriology* 130(2): 577-582.

Baker, T. A., et al. (1988). "Transcriptional activation of initiation of replication from the E. coli chromosomal origin: an RNA-DNA hybrid near oriC." *Cell* 55(1): 113-123.

Bates, D. B., et al. (1997). "The absence of effect of gid or mioC transcription on the initiation of chromosomal replication in Escherichia coli." *Proceedings of the National Academy of Sciences of the United States of America* 94(23): 12497-12502.

Borukhov, S., et al. (1993). "Transcript cleavage factors from E. coli." *Cell* 72(3): 459-466.

Bouche, J. P., et al. (1975). "dnaG gene product, a rifampicin-resistant RNA polymerase, initiates the conversion of a single-stranded coliphage DNA to its duplex replicative form." *The Journal of biological chemistry* 250(15): 5995-6001.

Brewer, B. J. (1988). "When polymerases collide: replication and the transcriptional organization of the E. coli chromosome." *Cell* 53(5): 679-686.

Brewer, B. J., et al. (1988). "A replication fork barrier at the 3' end of yeast ribosomal RNA genes." *Cell* 55(4): 637-643.

Brutlag, D., et al. (1971). "A possible role for RNA polymerase in the initiation of M13 DNA synthesis." *Proceedings of the National Academy of Sciences of the United States of America* 68(11): 2826-2829.

Chen, T. C., et al. (1978). "Replication of bacteriophage M13. XIII. Structure and replication of cloned M13 miniphage." *Journal of molecular biology* 125(2): 107-121.

Coulombe, B., et al. (1999). "DNA bending and wrapping around RNA polymerase: a "revolutionary" model describing transcriptional mechanisms." *Microbiol Mol Biol Rev*: 457-478.

Cramer, P., et al. (2001). "Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution." *Science* 292(5523): 1863-1876.

deHaseth, P. L., et al. (1995). "Open complex formation by Escherichia coli RNA polymerase: the mechanism of polymerase-induced strand separation of double helical DNA." *Mol Microbiol* 16(5): 817-824.

Deng, S., et al. (2005). "Organization of supercoil domains and their reorganization by transcription." *Molecular microbiology* 57(6): 1511-1521.

Deshpande, A. M., et al. (1996). "DNA replication fork pause sites dependent on transcription." *Science* 272(5264): 1030-1033.

deWind, N., Parren, P., Stuitje, A. R. and Meijer, < (1987). "Evidence for the involvement of the 16kD gene promoter in initiation of chromosomal replication of Escherichia coli

strains carrying a B/r-derived replication origin." *Nucleic acids research* 15: 4907-4914.

Epshtein, V., et al. (2007). "An allosteric path to transcription termination." *Molecular cell* 28(6): 991-1001.

Epshtein, V., et al. (2003). "Transcription through the roadblocks: the role of RNA polymerase cooperation." *EMBO J* 22(18): 4719-4727.

Fenton, M. S., et al. (2001). "Function of the bacterial TATAAT -10 element as single-stranded DNA during RNA polymerase isomerization." *Proceedings of the National Academy of Sciences of the United States of America* 98(16): 9020-9025.

French, S. (1992). "Consequences of replication fork movement through transcription units in vivo." *Science* 258(5086): 1362-1365.

Fuste, J. M., et al. (2010). "Mitochondrial RNA polymerase is needed for activation of the origin of light-strand DNA replication." *Molecular cell* 37(1): 67-78.

Gotta, S. L., et al. (1991). "rRNA transcription rate in Escherichia coli." *Journal of bacteriology* 173(20): 6647-6649.

Grainger, D. C., et al. (2006). "Association of nucleoid proteins with coding and non-coding segments of the Escherichia coli genome." *Nucleic acids research* 34(16): 4642-4652.

Gusarov, I., et al. (1999). "The mechanism of intrinsic transcription termination." *Molecular cell* 3: 495-504.

Hassan, A. B., et al. (1994). "Does transcription by RNA polymerase play a direct role in the initiation of replication?" *Journal of cell science* 107 (Pt 6): 1381-1387.

Herrmann, R., et al. (1978). "Transposition of a DNA sequence determining kanamycin resistance into the single-stranded genome of bacteriophage fd." *Molecular & general genetics : MGG* 159(2): 171-178.

Hirose, S., et al. (1983). "Initiation site of deoxyribonucleotide polymerization at the replication origin of the Escherichia coli chromosome." *Molecular & general genetics : MGG* 189(3): 422-431.

Huvet, M., et al. (2007). "Human gene organization driven by the coordination of replication and transcription." *Genome research* 17(9): 1278-1285.

Itoh, T., et al. (1980). "Formation of an RNA primer for initiation of replication of ColE1 DNA by ribonuclease H." *Proceedings of the National Academy of Sciences of the United States of America* 77(5): 2450-2454.

Itoh, T., et al. (1982). "Purification of ribonuclease H as a factor required for initiation of in vitro Co1E1 DNA replication." *Nucleic acids research* 10(19): 5949-5965.

Kaguni, J. M., et al. (1982). "Enzymatic replication of E. coli chromosomal origin is bidirectional." *Nature* 296(5858): 623-627.

Keppel, F., Fayet, O., Georgopoulos, C., Ed. (1988). The bacteriophages, Plenum Press, New York.

Kim, N., et al. (2007). "Transcription-associated mutagenesis in yeast is directly proportional to the level of gene expression and influenced by the direction of DNA replication." *DNA repair* 6(9): 1285-1296.

Kornberg, A. (1991). "Control of initiation of the Escherichia coli chromosome." *Cold Spring Harbor symposia on quantitative biology* 56: 275-278.

Kuznedelov, K., et al. (2002). "Structure-based analysis of RNA polymerase function: the largest subunit's rudder contributes critically to elongation complex stability and is

not involved in the maintenance of RNA-DNA hybrid length." *EMBO J* 21(6): 1369-1378.

Landick, R. (2009). "Transcriptional pausing without backtracking." *Proceedings of the National Academy of Sciences of the United States of America* 106(22): 8797-8798.

Lark, K. G. (1972). "Evidence for the direct involvement of RNA in the initiation of DNA replication in Escherichia coli 15T." *Journal of molecular biology* 64(1): 47-60.

Li, X. Y., et al. (1998). "Characterization of the closed complex intermediate formed during transcription initiation by Escherichia coli RNA polymerase." *The Journal of biological chemistry* 273(36): 23549-23557.

Liu, B., et al. (1995). "Head-on collision between a DNA replication apparatus and RNA polymerase transcription complex." *Science* 267(5201): 1131-1137.

Liu, B., et al. (1993). "The DNA replication fork can pass RNA polymerase without displacing the nascent transcript." *Nature* 366(6450): 33-39.

Lobner-Olesen, A., et al. (1992). "Different effects of mioC transcription on initiation of chromosomal and minichromosomal replication in Escherichia coli." *Nucleic acids research* 20(12): 3029-3036.

Lother, H., et al. (1985). "dnaA protein-regulated transcription: effects on the in vitro replication of Escherichia coli minichromosomes." *The EMBO journal* 4(2): 555-560.

Masai, H., et al. (1988). "Operon structure of dnaT and dnaC genes essential for normal and stable DNA replication of Escherichia coli chromosome." *The Journal of biological chemistry* 263(29): 15083-15093.

Masai, H., et al. (1996). "Mechanisms of primer RNA synthesis and D-loop/R-loop-dependent DNA replication in Escherichia coli." *Biochimie* 78(11-12): 1109-1117.

McGlynn, P., et al. (2000). "Modulation of RNA polymerase by (p)ppGpp reveals a RecG-dependent mechanism for replication fork progression." *Cell* 101(1): 35-45.

McLean, M. J., et al. (1998). "Base composition skews, replication orientation, and gene orientation in 12 prokaryote genomes." *Journal of molecular evolution* 47(6): 691-696.

Mekler, V., et al. (2002). "Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex." *Cell* 108: 599-614.

Merrick, M. J. (1993). "In a class of its own--the RNA polymerase sigma factor sigma 54 (sigma N)." *Mol Microbiol* 10(5): 903-909.

Merrikh, H., et al. (2011). "Co-directional replication-transcription conflicts lead to replication restart." *Nature* 470(7335): 554-557.

Messer, W. (1972). "Initiation of deoxyribonucleic acid replication in Escherichia coli B-r: chronology of events and transcriptional control of initiation." *Journal of bacteriology* 112(1): 7-12.

Messer, W., et al. (2001). "Bacterial replication initiator DnaA. Rules for DnaA binding and roles of DnaA in origin unwinding and helicase loading." *Biochimie* 83(1): 5-12.

Mirkin, E. V., et al. (2006). "Transcription regulatory elements are punctuation marks for DNA replication." *Proceedings of the National Academy of Sciences of the United States of America* 103(19): 7276-7281.

Mirkin, E. V., et al. (2005). "Mechanisms of transcription-replication collisions in bacteria." *Molecular and cellular biology* 25(3): 888-895.

Murakami, K. S., et al. (2002). "Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex." *Science* 296(5571): 1285-1290.

Murakami, K. S., et al. (2002). "Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 A resolution." *Science* 296(5571): 1280-1284.

Naito, S., et al. (1984). "Escherichia coli mutants suppressing replication-defective mutations of the ColE1 plasmid." *Proceedings of the National Academy of Sciences of the United States of America* 81(2): 550-554.

Nozaki, N., et al. (1988). "In vitro transcription of the origin region of replication of the Escherichia coli chromosome." *The Journal of biological chemistry* 263(28): 14176-14183.

Ogawa, T., et al. (1985). "Initiation of enzymatic replication at the origin of the Escherichia coli chromosome: contributions of RNA polymerase and primase." *Proceedings of the National Academy of Sciences of the United States of America* 82(11): 3562-3566.

Ogawa, T., et al. (1994). "Cell cycle-dependent transcription from the gid and mioC promoters of Escherichia coli." *Journal of bacteriology* 176(6): 1609-1615.

Ohsumi, M., et al. (1978). "The isolation and characterization of an in vivo recombinant between the filamentous bacteriophage f1 and the plasmid pSC101." *Virology* 89(2): 438-449.

Omont, N., et al. (2004). "Transcription/replication collisions cause bacterial transcription units to be longer on the leading strand of replication." *Bioinformatics* 20(16): 2719-2725.

Ozoline, O. N., et al. (1999). "DNA bendability--a novel feature in E. coli promoter recognition." *Journal of biomolecular structure & dynamics* 16(4): 825-831.

Ozoline, O. N., et al. (1995). "Structure of open promoter complexes with Escherichia coli RNA polymerase as revealed by the DNase I footprinting technique: compilation analysis." *Nucleic Acids Res.* 23: 4533-4541.

Park, C. S., et al. (1982). "Molecular mechanism of promoter selection in gene transcription. II. Kinetic evidence for promoter search by a one-dimensional diffusion of RNA polymerase molecule along the DNA template." *The Journal of biological chemistry* 257(12): 6950-6956.

Park, J. S., et al. (2002). "E. coli Transcription repair coupling factor (Mfd protein) rescues arrested complexes by promoting forward translocation." *Cell* 109(6): 757-767.

Pomerantz, R. T., et al. (2008). "The replisome uses mRNA as a primer after colliding with RNA polymerase." *Nature* 456(7223): 762-766.

Pomerantz, R. T., et al. (2010). "Direct restart of a replication fork stalled by a head-on RNA polymerase." *Science* 327(5965): 590-592.

Prado, F., et al. (2005). "Impairment of replication fork progression mediates RNA polII transcription-associated recombination." *The EMBO journal* 24(6): 1267-1276.

Price, M. N., et al. (2005). "Interruptions in gene expression drive highly expressed operons to the leading strand of DNA replication." *Nucleic acids research* 33(10): 3224-3234.

Rasmussen, K. V., et al. (1983). "Conditional change of DNA replication control in an RNA polymerase mutant of Escherichia coli." *Journal of bacteriology* 154(1): 443-451.

Reppas, N. B., et al. (2006). "The transition between transcriptional initiation and elongation in E. coli is highly variable and often rate limiting." *Molecular cell* 24(5): 747-757.

Rocha, E. P., et al. (2003). "Essentiality, not expressiveness, drives gene-strand bias in bacteria." *Nature genetics* 34(4): 377-378.

Rowen, L., et al. (1978). "Primase, the dnaG protein of Escherichia coli. An enzyme which starts DNA chains." *The Journal of biological chemistry* 253(3): 758-764.

Sakata-Sogawa, K., et al. (2004). "RNA polymerase can track a DNA groove during promoter search." *Proceedings of the National Academy of Sciences of the United States of America* 101(41): 14731-14735.

Santangelo, T. J., et al. (2004). "Forward translocation is the natural pathway of RNA release at an intrinsic terminator." *Molecular cell* 14(1): 117-126.

Schauzu, M. A., et al. (1987). "Transcripts within the replication origin, oriC, of Escherichia coli." *Nucleic acids research* 15(6): 2479-2497.

Schekman, R., et al. (1972). "Initiation of DNA synthesis: synthesis of phiX174 replicative form requires RNA synthesis resistant to rifampicin." *Proceedings of the National Academy of Sciences of the United States of America* 69(9): 2691-2695.

Selzer, G., et al. (1982). "Specific cleavage of the p15A primer precursor by ribonuclease H at the origin of DNA replication." *Proceedings of the National Academy of Sciences of the United States of America* 79(23): 7082-7086.

Srivatsan, A., et al. (2010). "Co-orientation of replication and transcription preserves genome integrity." *PLoS genetics* 6(1): e1000810.

Stuitje, A. R., et al. (1986). "Dissection of promoter sequences involved in transcriptional activation of the Escherichia coli replication origin." *Nucleic acids research* 14(5): 2333-2344.

Su'etsugu, M., et al. (2003). "Transcriptional control for initiation of chromosomal replication in Escherichia coli: fluctuation of the level of origin transcription ensures timely initiation." *Genes to cells : devoted to molecular & cellular mechanisms* 8(9): 731-745.

Tamm, J., et al. (1983). "Structural analysis of RNA molecules involved in plasmid copy number control." *Nucleic acids research* 11(18): 6381-6397.

Tanaka, M., et al. (1983). "A novel type of E. coli mutants with increased chromosomal copy number." *Molecular & general genetics : MGG* 192(1-2): 51-60.

Theisen, P. W., et al. (1993). "Correlation of gene transcription with the time of initiation of chromosome replication in Escherichia coli." *Molecular microbiology* 10(3): 575-584.

Travers, A. A. (2004). "The structural basis of DNA flexibility." *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences* 362(1820): 1423-1438.

Vassylyev, D. G., et al. (2002). "Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 A resolution." *Nature* 417(6890): 712-719.

Vilette, D., et al. (1996). "Transcription-induced deletions in plasmid vectors: M13 DNA replication as a source of instability." *Molecular & general genetics : MGG* 252(4): 398-403.

von Hippel, P. H. (2007). "From "simple" DNA-protein interactions to the macromolecular machines of gene expression." *Annual review of biophysics and biomolecular structure* 36: 79-105.

von Hippel, P. H., et al. (1984). "Protein-nucleic acid interactions in transcription: A molecular analysis." *Ann. Rev. Biochem.* 53: 389-446.

Wang, J. D., et al. (2007). "Genome-wide coorientation of replication and transcription reduces adverse effects on replication in Bacillus subtilis." *Proceedings of the National Academy of Sciences of the United States of America* 104(13): 5608-5613.

Wang, M. D., et al. (1998). "Force and velocity measured for single molecules of RNA polymerase." *Science* 282: 902-906.

Wanrooij, S., et al. (2008). "Human mitochondrial RNA polymerase primes lagging-strand DNA synthesis in vitro." *Proceedings of the National Academy of Sciences of the United States of America* 105(32): 11122-11127.

Wickner, W., et al. (1972). "RNA synthesis initiates in vitro conversion of M13 DNA to its replicative form." *Proceedings of the National Academy of Sciences of the United States of America* 69(4): 965-969.

Zenkin, N., et al. (2006). "The mechanism of DNA replication primer synthesis by RNA polymerase." *Nature* 439(7076): 617-620.

Zenkin, N., et al. (2008). "RNA polymerase -the third class of primases." *Cell Mol Life Sci* 65(15): 2280-2288.

Zinder, N. D., et al. (1985). "Multiregulatory element of filamentous bacteriophages." *Microbiological reviews* 49(2): 101-106.
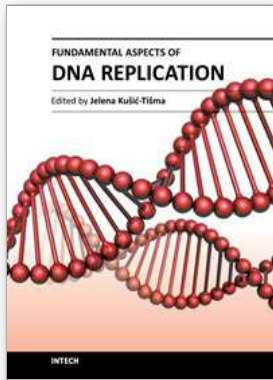
Fundamental Aspects of DNA Replication

Edited by Dr. Jelena Kusic-Tisma

DNA replication, the process of copying one double stranded DNA molecule to produce two identical copies, is at the heart of cell proliferation. This book highlights new insights into the replication process in eukaryotes, from the assembly of pre-replication complex and features of DNA replication origins, through polymerization mechanisms, to propagation of epigenetic states. It also covers cell cycle control of replication initiation and includes the latest on mechanisms of replication in prokaryotes. The association between genome replication and transcription is also addressed. We hope that readers will find this book interesting, helpful and inspiring.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

INTECH
open science | open minds