博士論文番号：981030
(Doctoral student number)

Molecular mechanisms underlying the genetic instability of inverted repeats in DNA

Lai Pey Jiun
Nara Institute of Science and Technology
Graduate School of Biological Sciences
Microbial Molecular Genetics
(Prof. Hisaji Maki)

Submitted on 2015/11/25

CONTENTS

Chapter 1

Introduction

1.1 Genome instability caused by unusual DNA structure

DNA structure is well known as the classical Watson and Crick B-DNA double helix. For many years it has been believed that DNA in general occurs in uniform structure. Not until year 1979 when the first alternative DNA structure- Z-DNA was discovered (Wang *et al*., 1979). DNA can exist with dynamic structure that can adopt a variety of alternative structure or can be known as non-B DNA structure. This includes hairpins/cruciform, Z-DNA, triplexes (H-DNA), tetraplexes (G-quadruplex DNA) and slipped DNA (Zhao *et al*., 2010) (Table 1).

One of the common features among these unusual DNA structure is the presence of specific defined sequences, for example mirror repeats, direct repeats, homopurine-homopyrimidine tracts or inverted repeats (Sinden 1994). These specific defined sequences have the tendency to adopt the alternative DNA structure. An inverted repeat is capable to form cruciform structure in double-stranded DNA or hairpin structure in single-stranded DNA (Smith 2008). DNA sequences with alternating pyrimidines and purines, for instance $(CG)_n$ or $(CA)_n$ favours the formation of Z-DNA (Rich *et al*., 1984). While, triples DNA (H-DNA) can be formed with DNA sequences containing the homopurine-homopyrimidine mirror repeats (Frank-Kamenetskii and Mirkin 1995). DNA sequence with repeating elements rich in guanosine residues can form the G-qudruplex DNA (Majumdar and Patel 2002). Slipped DNA is formed by direct repeats when the complementary repeats mispair and resulting in loop-out structure (Sinden *et al*., 2007).

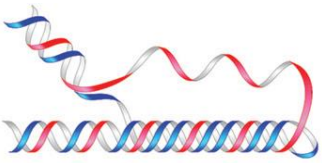
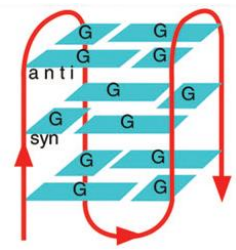
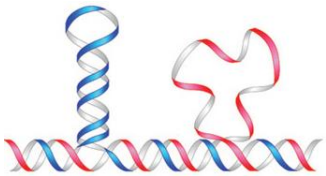
The presence of these unusual DNA structures has been suggested as a source of genome instability associated with chromosomal deletion, expansion, breakage or chromosomal rearrangement. It has been suggested that genome instability is dependent on DNA replication.

Of these repeats, tinucleotide repeat sequence (TRS) has been reported to be involved in human neurodegenerative diseases (Wells and Warren 1998; Oostra 1998). These diseases involved in the expansion of trinucleotide repeats for instance, (CTG)n/(CAG)n, (CGG)n/(CCG)n and (GAA)n/(TTC)n (Wells and Warren 1998). It is suggested that the instability of these trinucleotide repeats is dependent on the activity of DNA replication. The ability of secondary structure formation from these repeats happening during DNA replication may cause the expansion or contraction of this repeat depending on their orientation (Mirkin 2007).

H-DNA forms at DNA regions with homopurine-homopyrimidine mirror repeats capable to induce genetic instability in mammalian cells. Wang and Vasquez (2004) suggested that the H-DNA is mutagenic in mammalian cells, and double-stranded break (DSB) was also found near to the H-DNA. In addition H-DNA has also been shown to arrest DNA polymerase (Rao 1996; Krasilnikov *et al*., 1997). A Z-DNA-forming sequence is also known as a hotspot of genome instability. It was suggested that Z-DNA induces genetic instability differently in both bacterial and mammalian cells. In bacteria, Z-DNA-forming sequence induces small deletion within the CG-repeats which cause by the slippage events during DNA replication (Freund *et al*., 1989). DNA sequence rich in guanosine residue potentially forming the G-qudruplex structure can also interfere DNA replication. *In vitro* biochemical studies suggested that G-qudruplex capable in blocking DNA replication (Weitzmann *et al*., 1996).

DNA sequences with defined order sequence are found abundantly in both prokaryotes and eukaryotes genome. Studies are now suggesting that these defined order sequence DNA can cause genome instability which is dependent on DNA replication. However, there is still lacking the detail mechanism about how DNA replication is involved in causing the genome instability. Thus, in depth biochemical studies is needed for further confirmation.

Table 1 Types of non-B conformation and their characteristic

Name	**Conformation**	**General sequence requirements**	**Reference**
Cruciform		Inverted repeats ATGCAGTACTGCAT TACGTCATGACGTA	Smith 2008
Z-DNA		(YR.YR)n CGCGTGGCGTGTG GCGCACCGCACAC	Rich *et al*., 1984
Triplexes (H-DNA)		Mirror repeat AAGAGG GGAGAA T TCT CCCCTC T T	Frank-Kamenetskii and Mirkin 1995
Tetraplex (G-quadruplex DNA)		Oligo $(G)_n$ tracts $AG_3(T_2AG_3)_3$ Single strand	Majumdar and Patel 2002
Slipped DNA		Direct repeats TCGGTTCGGT AGCCAAGCCA	Sinden *et al*., 2007

1.2 Inverted repeats / palindrome sequences

Inverted repeats or can be known as palindrome is a DNA sequence with two copies of same nucleotide sequence oriented in opposite directions (Diagram 1A). Therefore this DNA sequence capable to base pair with the complement strands through intramolecular base pairing. As a consequent of this intrastrand basepairing, inverted repeats/palindrome sequences are capable to adopt two types of DNA secondary structure: (i) cruciform structure when in double-stranded DNA (dsDNA) (Diagram 1B) and (ii) hairpin structure when in single-stranded DNA (ssDNA) (Diagram 1C).

Formation of cruciform structure requires the melting of interstrand basepairing of the duplex DNA in order to form the intrastrand basepairing of the two DNA repeats oriented in opposite direction. Cruciform formation is energetically unfavorable in a relaxed DNA molecule. The formation of this cruciform requires the melting of the interstrand basepairing and stacking interactions in the loop as well as at the base of the stem. Transition from inverted repeats to cruciform structure is an energy driven process. The energy required for formation of cruciform structure is originated from the negative supercoiling (Lilley 1980; Panayotatos and Wells 1981). Mizuuchi *et al*. (1982) showed that cruciform structure is formed with the addition of DNA gyrase that introduce negative supercoiling.

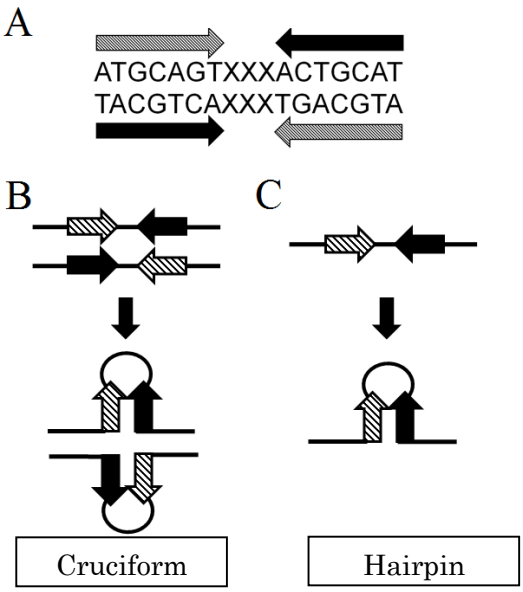


Diagram 1 **Schematic diagram showing the inverted repeats and the secondary structure.** (A) Inverted repeats are illustrated with the opposing arrows above the sequence. X represents the spacer, nucleotides between the inverted repeats. (B) Formation of cruciform structure when in double-stranded DNA. (C) Formation of hairpin structure when in single-stranded DNA.

1.2.1 Biological roles of inverted repeats and hairpin structure

Inverted repeats/palindrome sequences have been found abundantly in both prokaryotes and eukaryotes genome. Many of the inverted repeats lie within the regulatory sequences. Hence, the presence of these sequences capable of forming secondary structure may involve in several biological functions, for instance replication, transcription or translational control.

Inverted repeats have been shown to play a crucial role for the initiation of DNA replication in bacteria, eukaryotic, viruses and mammalian cells. Numerous studies have shown that secondary structure formed from the inverted repeats near or at the DNA replication origin as a recognition site for initiation of replication (Pearson *et al*., 1996). The 245 bp origin of replication in *E. coli* is known containing many inverted repeats. Within the *ori*C region of *Escherichia coli,* it contains three AT-rich 13 mer required for opening the duplex and initiation of replication (Bramhill and Kornberg 1988). Though, there is no direct evidence showed that these inverted repeats involved in alternative conformation (Sinden 1994). In the eukaryotic system, for example in the region of autonomously replicating sequences (ARS) of

Saccharomyces cerevisiae consist of 11 bp of conserved AT-rich elements (Stinchcomb *et al*., 1981; Broach *et al*., 1984). Cruciform structure formed from this region is required as the binding site for its origin recognition proteins (Callejo *et al*., 2002). Apart from that, the origin of phage G4 is shown to consist of three inverted repeats that capable to form hairpin structure, which are required for primase activity (Lambert *et al*., 1986). Hiasa *et al*. (1989) reported that the secondary structure of the stem-loops regardless of the DNA sequence is important for the priming.

Transcription is a process in which the binding of RNA polymerase (RNAP) to the promoter in DNA is involved. The overall DNA topology of the template DNA can influence the interaction of RNAP to the promoter as well as the regulator. One of the examples is the involvement of DNA hairpin in regulating the recognition of the N4 virion. N4 virion contains three early promoter sites which consist of short inverted repeats with 5 to 7 bp in length (Haynes and Rothman-Denes 1985). Glucksmann *et al*. (1992) showed that the presence of inverted repeats in the virion RNA polymerase (vRNAP) promoters that capable to form hairpin structure is required for vRNAP recognition. Later on, they also showed that extrusion of hairpin promoters as cruciform is mediated by the *E. coli* host DNA gyrase upon injection to *E. coli* (Dai *et al*., 1997, 1998). Apart from these, hairpin structure are also involved in the translational control mechanism. Studies have shown that secondary structure formed on mRNA determined the efficiency of translational initiation in prokaryotes (Gold 1988).

The presence of inverted repeats that potentially forming the secondary structure such as hairpin structure is important in the regulation of numerous essential biological functions. Either it is the inverted repeats or the formation of hairpin structure is required as the recognition site for those biological functions. Thus, the presence of inverted repeats is somehow essential and not avoidable in the genome.

1.2.2 Genome instability caused by inverted repeats is suggested dependent on DNA replication

The presence of inverted repeats has been shown to be functionally important in regulation of several biological processes. However, the occurrence of inverted repeats can also cause havoc in the genome by provoking genome instability. Two types of phenomenon have been reported with the presence of inverted repeats/palindrome sequences: (i) inviability which they can cause lethality to the carrier replicon, or (ii) instability whereby inverted repeats/ palindrome sequences is deleted at varying frequencies and the carrier is viable.

Attempts for cloning large inverted repeats/palindrome have been carried out back to the 70s but was not successful (Perricaudet *el al*., 1977). Studies showed that cloned long palindrome when transformed into *E.coli* underwent subsequent deletions and caused instability (Collins, 1981). Collins *et al*. (1982) later showed that higher stability of inverted repeats was observed in *recB sbcB* host. Lilley (1981) also demonstrated that the inability to clone a 260 bp palindrome is probably due to the instability of replicon of vector plasmid used. In addition short palindrome of 68 bp was also found unstable (Courey and Wang, 1983).

Difficulties in propagating the long palindrome in *E. coli* are suggested to be the cause of inviability of palindrome-containing plasmid. Hagan and Warren (1982, 1983) showed that a palindrome sequence with size greater than 800 bp is inviable when transformed into the Rec^+ *E. coli* strain. In addition, Leach and Stahl (1983) reported that λ-phage carrying a 3200 bp palindrome is inviable in Rec^+ strain. However, the inviability of palindrome-containing λ-phage is overcome in the *recBC sbcB* strains.

The instability and inviability of this palindrome sequence have been suggested to be dependent on DNA replication. Leach and Lindsey (1986) showed the loss of bacteriophage carrying a 530 bp palindromes in Rec^+ and *recA* strain but not *recBC sbcB* in the lytic condition. However, this phenomenon was not observed in Rec^+ host lysogenic for λ-phage. Hence, this suggested the correlation of DNA replication in causing the inviability of this palindrome

sequence. Shurvinton *et al*. (1987) also showed that λ-phage carrying palindrome of 8400∓100 bp is viable in *recBC sbcB* strain of *E. coli* but not rec^+ when DNA replication is active. Thus, they suggested that the activity of DNA replication leads to the formation of secondary structure from the palindrome sequence, which is the target of *recBC* and *sbcB* gene product. Lindsey and Leach (1989) found that λ-phage with a 571 bp of palindrome was able to be maintained in the $recBCD^+$ and $sbcB^+$ *E. coli* strain with subsequent DNA replication. They suggested that the ability for the maintenance of this palindrome-containing DNA is probably due to the reduced rate of replication. However, there are studies showed that *recBC sbcB* strain also carried the *sbcC* mutation (Llyod and Buckman, 1985) and suggested that *sbcC* may be important to determine the palindrome viability (Chalker *et al*., 1988).

1.3 The hypothesized model mechanism regarding the viability of palindrome sequences

Long palindrome has been shown to cause genome instability by inducing instability or inviability of palindrome sequence in a way dependent on the activity of DNA replication as suggested by the model mechanism shown in Diagram 2. It has been suggested that the deletion of palindrome sequence presumably due to the formation of hairpin structure could preferentially occur on the lagging strand (Trinh and Sinden 1991; Rosche *et al*., 1995; Pinder *et al*., 1998). In the concerted semi-discontinuous DNA replication, leading strand is synthesized continuously while lagging strand is synthesized discontinuously (Diagram 2B). It is suggested that during the discontinuous synthesis of lagging strand, the transiently exposed ssDNA on lagging strand template leads to the formation of hairpin structure (Diagram 2C).

Studies have shown that long palindromes are not tolerated in the wild type *E.coli* but capable to be propagated in the *sbcCD* mutants (Leach and Stahl 1983; Chalker *et al*., 1988; Gibson *et al*., 1992). Connelly *et al*. (1988) reported that the purified SbcCD protein of *E. coli* can cleave a hairpin structure. Hence, it is suggested that the inviability of long palindrome is dependent on the SbcCD complex, probably due to the cleavage of the hairpin structure (Leach

et al., 1997) (Diagram 2D).

Genetic evidence shown by Cromie *et al*. (2000) suggested that a 246 bp palindrome potentially forming a hairpin structure in ssDNA during DNA replication is a target for SbcCD nuclease. Cleavage of this hairpin structure could cause a DSB within this palindrome sequence which can be repaired by recombination. In addition, it was also suggested that hairpin cleavage occurs after replication through and beyond the palindrome. Later on, an *in vivo* physical analysis of the palindrome sequences on DNA revealed that a DSB was generated at the palindrome sequences in the *E.coli* and this process was dependent of DNA replication (Eykelenboom *et al*., 2008). Taken together with the genetic and physical evidences, they suggested a model in which a palindrome is the site for DSB formation and this cleavage occurs behind the replication fork (Diagram 2D).

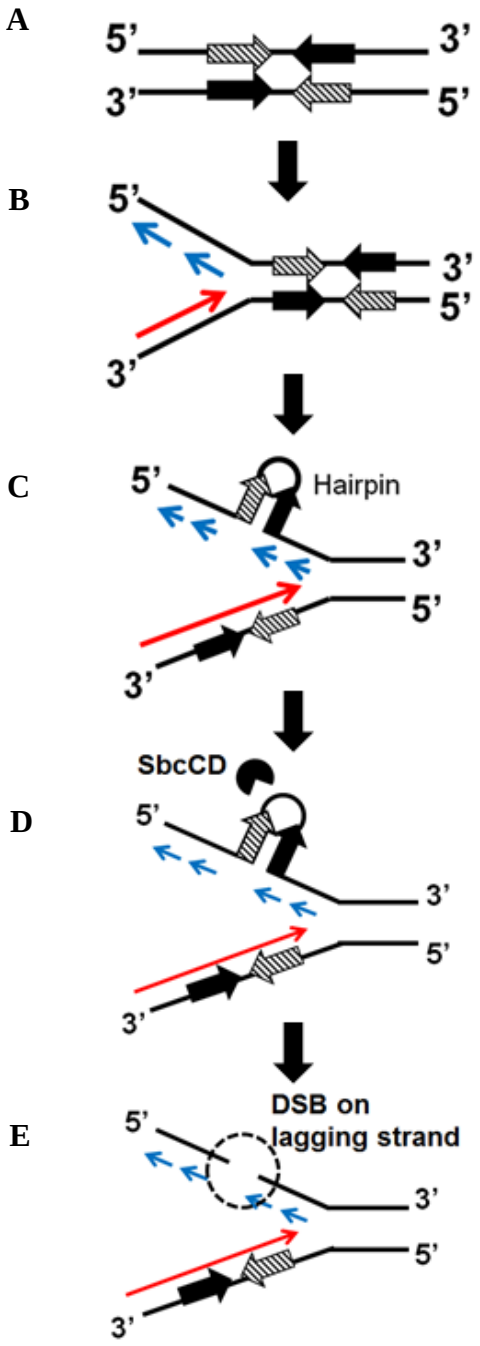


Diagram 2 **The hypothesized model mechanism of palindrome cleavage by SbcCD**. (A) Presence of inverted repeats on the DNA sequences. (B) Semi-discontinuous DNA replication mechanism involved in the continuous synthesis of leading strand (red line with arrowhead) and discontinuous synthesis of lagging strand (Blue line with arrowhead). (C) During the discontinuous synthesis of lagging strand, inverted repeats adopt the hairpin structure on the transiently exposed ssDNA on lagging strand template. This hairpin is formed after the replication fork and does not affect the replication fork. (D) Presence of this hairpin structure is cleaved by the SbcCD complex. (E) Thus, this cleavage will result a double-stranded break at the palindrome sequence. In the diagram, thin black line indicates the DNA template. Thick black line with arrowhead represents the inverted repeats.

1.4 Thesis rationale

As described above, the currently available model suggests that a hairpin structure is formed on the lagging strand template during DNA replication. Although the hairpin formation is hypothesized to be formed after the passage of replication fork it is predicted that polymerase on lagging strand must be affected. Numerous *in vitro* studies have shown that a variety of DNA polymerases, for instance DNA polymerase α (Weaver and DePamphilis 1982; Kaguni and Clayton 1982; Reckmann *et al*., 1985), Vaccinia DNA polymerase (Challberg and Englund 1979), bacteriophage T4 DNA polymerase (Huang *et al*., 1981) as well as *E. coli* polymerase III (LaDuca *et al*., 1983) were arrested at specific nucleotide sequence on the single-stranded DNA template that potentially forming a secondary structure, particularly hairpin loops leading to the pausing of DNA polymerase.

The resembles situation has been reported as previously shown in Higuchi *et al*. (2003), a DNA lesion located on lagging strand completely blocked the synthesis of Okazaki fragment extending toward the lesion site, but have not affected on the progression of the replication fork or leading strand synthesis. It is suggested that upon the lagging strand synthesis encounters the blocking lesion, the twin PolIII core catalysing the lagging strand DNA synthesis gets to recycle in the similar way as it is usually recycled during Okazaki Fragment synthesis. The unique characteristic of lagging strand type PolIII has been suggested by previous study (Maki & Kornberg 1988; Marians 1992; Li & Marians 2000) and thereby reduces the effect of lesion on lagging strand template to the concurrent DNA synthesis at the replication fork.

Mirkin's group (Voineage *et al*., 2008) has reported that replication forks are stalled at the inverted repeats in bacteria, yeast and mammalian cells. In contrast to that, the model suggested by Leach's group (Eykelenboom *et al*., 2008) as showed in diagram 2 suggested that an inverted repeat does not cause the collapse of replication; instead, the cleavage of hairpin structure formed from the inverted repeats by the SbcCD happened after passage of a replication fork.

With this observation, we couldn't explain how inverted repeats can caused the pausing of replication fork in the result observed in Mirkin's group. In addition the detail of biochemical mechanism regarding the progression of replication fork when encountering the inverted repeats is yet still unavailable.

In view of this controversy, I would like to know how DNA replication machinery will behave when the replication fork encounters the inverted repeats. Using the *in vitro oriC* DNA replication, I would like to find out (Diagram 3): (i) is hairpin structure formed on lagging strand template, and (ii) can DNA replication fork stall when it encounters the inverted repeats. With this in depth biochemical study, it will enable the understanding of the mechanism how inverted repeats can cause genetic instability dependent on DNA replication.

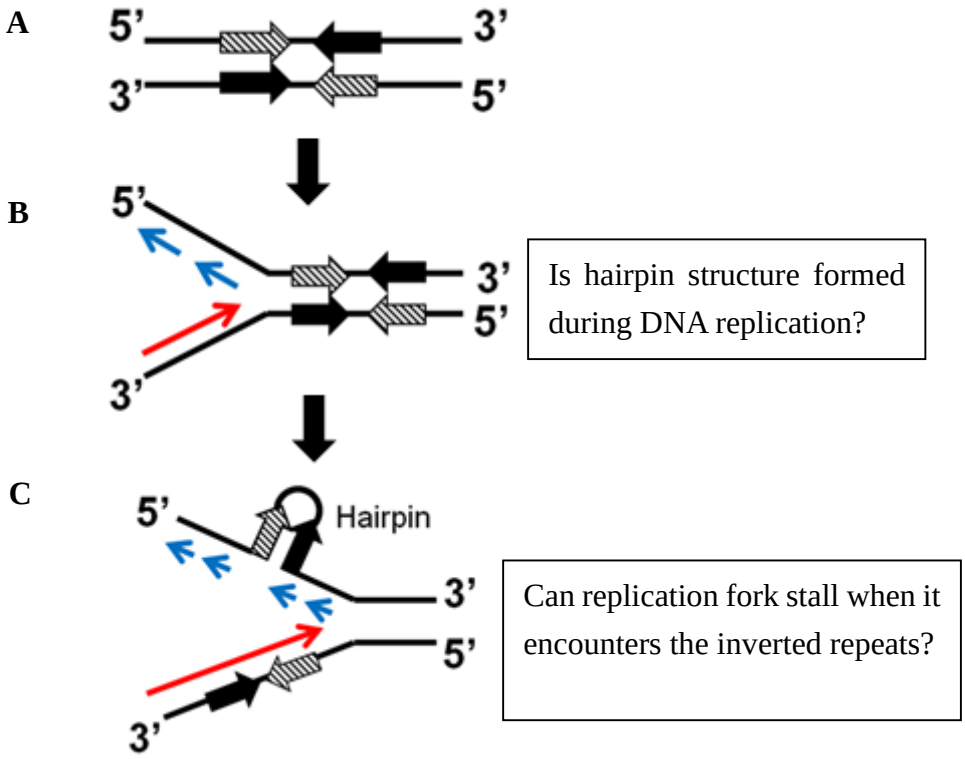


Diagram 3 **The objectives of this study.** (A) DNA sequence with the presence of inverted repeats. (B) During the semi-discontinuous DNA replication, leading strand is synthesized continuously (red line with arrowhead) and lagging strand is synthesized discontinuously producing small DNA fragment, Okazaki fragment (blue line with arrowhead). (C) Inverted repeats potentially forming hairpin structure is suggested to form on the transiently exposed ssDNA on lagging strand template and blocked Okazaki fragment synthesis. It is suggested that hairpin is formed on lagging strand after replication fork has passed through whole inverted repeats. Thin black line indicates the DNA template. Thick black line with arrowhead indicated the inverted repeats.

Chapter 2

Materials and Methods

2.1 Materials

2.1.1 Chemicals, reagents and enzyme

Chemical reagents used in this study were obtained from Sigma and Wako. Restriction enzymes and T7 endonuclease were from New England Biolabs. T4 polynucleotide kinase was from TOYOBO. In Fusion cloning kit was from Clontech.

2.1.2 *E. coli* strains (Table 2)

E. coli W3110 *tus*::Km used to maintain the *terB* containing *oriC* plasmid (pMOL7) was prepared as previously described (Hidaka *et al*., 1988). MK9200 with *sbcC* deletion was constructed with P1-mediated transduction of *tus*::Km from W3110 into DL 494 JM83 *sbcC201 phoR*::Tn10. This strain was used for the maintenance of palindrome-containing *oriC* plasmid. DL2774 JM83 *sbcC*::Km/ pLacD1-Pal246 is *E. coli* strain carrying pLacD1 which is the parental plasmid carrying the 246 bp palindrome sequence.

- **P1 Transduction**

Long DNA palindromes capable to form hairpin or cruciform structure are not tolerated in wild type *E. coli* (Leach 1994). Previous studies have shown that this problem can be overcome in *sbcC* or *sbcD* mutants (Chalker *et al*., 1988; Gibson *et al*., 1992). To stably maintain the 246 bp palindrome-containing *oriC* plasmid, *E. coli* strain with *sbcC* mutant was constructed using the P1 transduction.

In brief, an overnight culture of the donor strain (W3110 *tus*::Km) was sub-cultured in 5 ml LB broth-5 mM $CaCl_2$. P1 virus (MK811) was added to 1 ml of the donor strain subculture and incubated in 37 °C water bath for 20 min. Then 2.5 ml of R-top agar was added to this culture and plated on R-plates. These plates were incubated for 8 h at 37 °C. The layer of soft

agar was then scraped into a small centrifuge tube following by adding 5 drops of chloroform. The mixture was vortex vigorously and centrifuged with 9000 rpm at 4 °C for 20 min. The supernatant containing the P1 lysate carrying the *tus*::Km was transferred to new tube and 1-2 drops of chloroform was added before storage at 4 °C. This P1 lysate is ready for transduction.

For transduction, an overnight culture of recipient strain (DL494 JM83 *SbcC201 PhoR*::Tn10) was inoculated to fresh LB medium and grew with shaking at 37 °C until the OD 0.9-1.0. Then, 1 ml of this fresh culture was spun down with 3000 rpm at room temperature for 5 min. Next, the cell pellet was suspended with 1 ml of MC buffer. P1 lysate which prepared previously was added to this suspension cells and mix well. The cell mixture was incubated at 37 °C for 20 min. At this point, 200 µl of 1 M Sodium citrate was added to the cell mixture. Lastly this cell mixture was plate on LB plate with kanamycin selection.

2.1.3 Plasmid (Table 3)

The *oriC-terB* plasmid, pMOL7 was prepared as previously described (Higuchi *et al*., 2003). This plasmid was used as a vector for the construction of *oriC-terB* palindrome-containing plasmid. pLacD1 is the parental plasmid carrying the 111 bp inverted repeats separated by a 24 bp spacer. pMD-20T is the plasmid containing the 141 bp nucleotide sequence synthesized by the company which used to replace one of the palindrome sequence in the *oriC-terB* palindrome-containing plasmid for the preparation of palindrome-free control plasmid.

Table 2 List of *E. coli* strains

Strain	Genotype	Source or reference	Brief description
W3110 *tus*::Km	*LAM*$^{-}$, *in(rrnD-rrnE)1, rph-1,* $r^{+}m^{+}$, *Tus::km*^{r}		Host for the maintenance of *terB* containing plasmid
W3110 *tus*::Km/ pMOL7		Higuchi *et al*. (2003)	*E. coli* strain containing the *oriC* plasmid, pMOL7
DL494 JM83 *sbcC201 phoR*::Tn10	F^{-} *ara* Δ*(lac-proAB) rpsL* (Str^{r})*[φ80* d*lac*Δ*(lacZ)M15] thi*	Dr. David Leach	*E. coli* strain with the *SbcC* mutation
DL2774 JM83 *sbcC*::Km/ pLacD1-Pal246		Dr. David Leach	Host for the maintenance of pLacD1 plasmid containing the 246 bp palindrome sequence
MK9200 *sbcC201 phoR*::Tn10 *tus*::Km		This work	Host for the maintenance of 246 bp palindrome-containing *oriC-terB* plasmid

Table 3 List of plasmids

Plasmid	Brief description	Source or reference
pMOL7	*oriC-terB plasmid*	Higuchi *et al*. (2003)
pLacD1	Parental plasmid containing the 246 bp palindrome sequence	Dr. David Leach
pIR246	*oriC* plasmid containing the 246 bp palindrome sequence	This work
pCNS	Palindrome-free *oriC* plasmid	This work
pMD-20T	Plasmid containing 141 bp of the synthesized sequence	TAKARA

2.1.4 Oligonucleotide DNA

- **PCR primers (Table 4)**

PCR reaction was carried out using the Gene Amp PCR system 2400 (PE Applied Biosystems). Polymerases used for the PCR reaction were ExTaq Polymerase (Takara) and KOD-Plus Polymerase (Takara). PCR primers as listed in Table 4 were designed to amplify the vector plasmid, the palindrome sequence and the oligonucleotide sequence for preparation of control plasmid.

The PCR reaction conditions used for the amplification of 246 bp palindrome sequence were: 25 cycles of 30 sec at 96 °C, 30 sec at 60 °C, 30 sec at 72 °C with an initial denaturation step at 96 °C for 4 min and a final extension step at 72 °C for 5min. The PCR reaction conditions used for the amplification of vector plasmid, pMOL7 was the same as above except with longer elongation time: 4 min at 72 °C.

For the construction of palindrome-free control plasmid, pCNS, PCR reactions used to amplify the oligonucleotide sequence were: 30 cycles of 30 sec at 96 °C, 30 sec at 57.4 °C, 15 sec at 72 °C with an initial denaturation step at 96 °C for 4 min and a final extension step at 72 °C for 3min. While the PCR reaction conditions for the amplification of vector plasmid, pIR246, were: 30 cycles of 15 sec at 94 °C, 30 sec at 56.4 °C, 5 min at 68 °C with an initial denaturation step at 94 °C for 2 min and a final extension step at 68°C for 5min.

Table 4 List of primers

Name	Sequence
• To amplify pMOL7 Vector	
pMOL7_IFC2_For	5' AGCTTATGCATTTCTTTCCAGA 3'
pMOL7_IFC2_Rev	5' TCTCAATCGATAACCTTATTTTTG 3'
• To amplify 246 bp palindrome sequence from pLacD1	
Pal246_IFC_S_For	5' GGTTATCGATTGAGACACACAGGAAACAGCTATGA 3'
Pal246_IFC_S_Rev	5' AGAAATGCATAAGCTCAGTCACGACGTTGTAAAAC 3'
• To amplify pIR246 vector	
pIR246_V_For	5’ CGTCTTCGTCGTTTTACAACGT 3’
pIR246_V_Rev	5’ ATCAGCCTCATTTCTAAATATGAA 3’
• To amplify synthesized sequence from pMD-20T to prepare pCNS	
Reshuffled IR_Nsi_For	5’ AGAAATGAGGCTGATACTAGTAAC 3’
Reshuffled IR_Nsi_Rev	5’ AAAACGACGAAGACGATAAGAATG 3’
• To check for positive clone	
Pal246pMOL7_CS_For	5' TCACCGAGGCAGTTCCAT 3'
Pal246pMOL7_CS_Rev	5' GGTTTGGTTGATGCGAGTG 3'

- **Synthesized oligonucleotides sequence used for the construction of palindrome-free control plasmid, pCNS (Table 5A)**

A total of 141 bp double stranded DNA was synthesized by the company (TAKARA). This oligonucleotide was used to replace one arm of the palindrome sequence in pIR246 in order to construct the palindrome-free control plasmid. The sequence for this oligonucleotide is shown in Table 5A.

- **Probes used for the Southern blot hybridization (Table 5B)**

Southern blot hybridization probes used for the analysis of leading and lagging strand synthesis in alkaline agarose gel and sequencing gel were designed as listed in Table 5B. Probe Lead B and H were designed for the analysis of leading strand replication product. Probe Lead B was designed located approximately 437 bp upstream of the palindrome sequence and was used for the analysis of alkaline agarose gel Southern blotting. Probe Lead H which was used in the sequencing gel Southern blotting was located approximately 34 bp upstream of the palindrome sequence. These probes enable the determination of leading strand paused product within the palindrome sequence. Probe Lag A' was designed for the analysis of lagging strand replication product. It was located approximately 25 bp upstream of the palindrome sequence on lagging strand template and was used for the sequencing gel Southern blot analysis.

Table 5 List of oligonucleotide

A. Oligonucleotide DNA used in the construction of palindrome-free control plasmid, pCNS

5'AGAAATGAGGCTGATACTAGTAACGGTATCTCTCACGAATAATGCGTACTTA TACACTAGTATAATACTAGTCATTATCTATAGGTCTGTTCTATGATATAACTACT GTCTTCTTAGCATTCTTATCGTCTTCGTCGTTTT 3'

B. Probe used for Southern hybridization

Name	Sequence	Location
B	5' CGCATCGGGCTTCCCATACAATCGATCTTGGTCTG 3'	3387-3421
H	5' TATGCATTTCTTTCCAGACTTGTTCAAC 3'	2883-2911
A'	5' GTGTGTCTCAATCGATAACCTTATTTTTGACGAGG 3'	2643- 2678

2.1.5 Proteins

DnaA (300 ng/μl), GyrA (920 ng/μl), GyrB (460 ng/μl), HU (270 ng/μl) are gift from Prof. Katayama in Kyushu University. Tus (280 ng/μl) is given by Prof. Hidaka in National Institute for Basic Biology. SSB (570 ng/μl), β-clamp subunit (400 ng/μl), Pol III* Fraction IV (80 units/μl), DnaB (980 ng/μl), DnaC (333 ng/μl), DnaG (980 ng/μl) were previously prepared in our laboratory. In the case of SSB titration experiment and the analysis of the effect of DNA gyrase, Calmodulin-binding protein-tagged HolD pol III* (12.8 units/μl) was used. It was prepared by Dr. Furukohri as described previously (Butland *et al*., 2005).

2.2 Methods

2.2.1 Construction of plasmid using the In Fusion cloning method

- ***oriC-terB* palindrome containing plasmid, pIR246**

An *oriC-terB* plasmid containing the 111 bp inverted repeats separated by a 24 bp spacer was constructed by PCR and In-Fusion cloning method as illustrated in Diagram 4. pMOL7/W3110 *tus*::Km and pLacD1 PAL246/ DL2774 JM83 *sbcC*::Km were grown for 24 h at 30 °C in LB media containing 50 μg/ml Ampicilin and 20 μg/ml Chloramphenicol, respectively. These bacterial cultures were collected and plasmids were purified using the QIAGEN Midi prep kit (QIAGEN).

In-Fusion™ Advantage PCR cloning Kit (Clontech) was used for the construction of palindrome-containing plasmid. This method is designed to join DNA fragments that have 15 bases of homology at their ends. To do so, PCR amplification was carried out to amplify the vector plasmid, pMOL7. PCR primer used to amplify the 246 bp palindrome sequence from the pLacD1 was designed that share 15 bases of homology with the sequence at the ends of the amplified linearized cloning vector, pMOL7. In-Fusion cloning reaction was set up by mixing the amplified vector, pMOL7, amplified 246 bp palindrome sequence, 1× In-Fusion reaction buffer and In-Fusion Enzyme in a 10 μl reaction. This reaction mixture was incubated for 15 min at 37 °C, followed by 15 min incubation at 50 °C, and then placed on ice. This was transformed into MK9200 by heat-shock transformation and grown at 30 °C in LB plate containing 50 μg/ml Ampicilin for the selection of positive clone. The *oriC-terB* palindrome-containing plasmid is named pIR246.

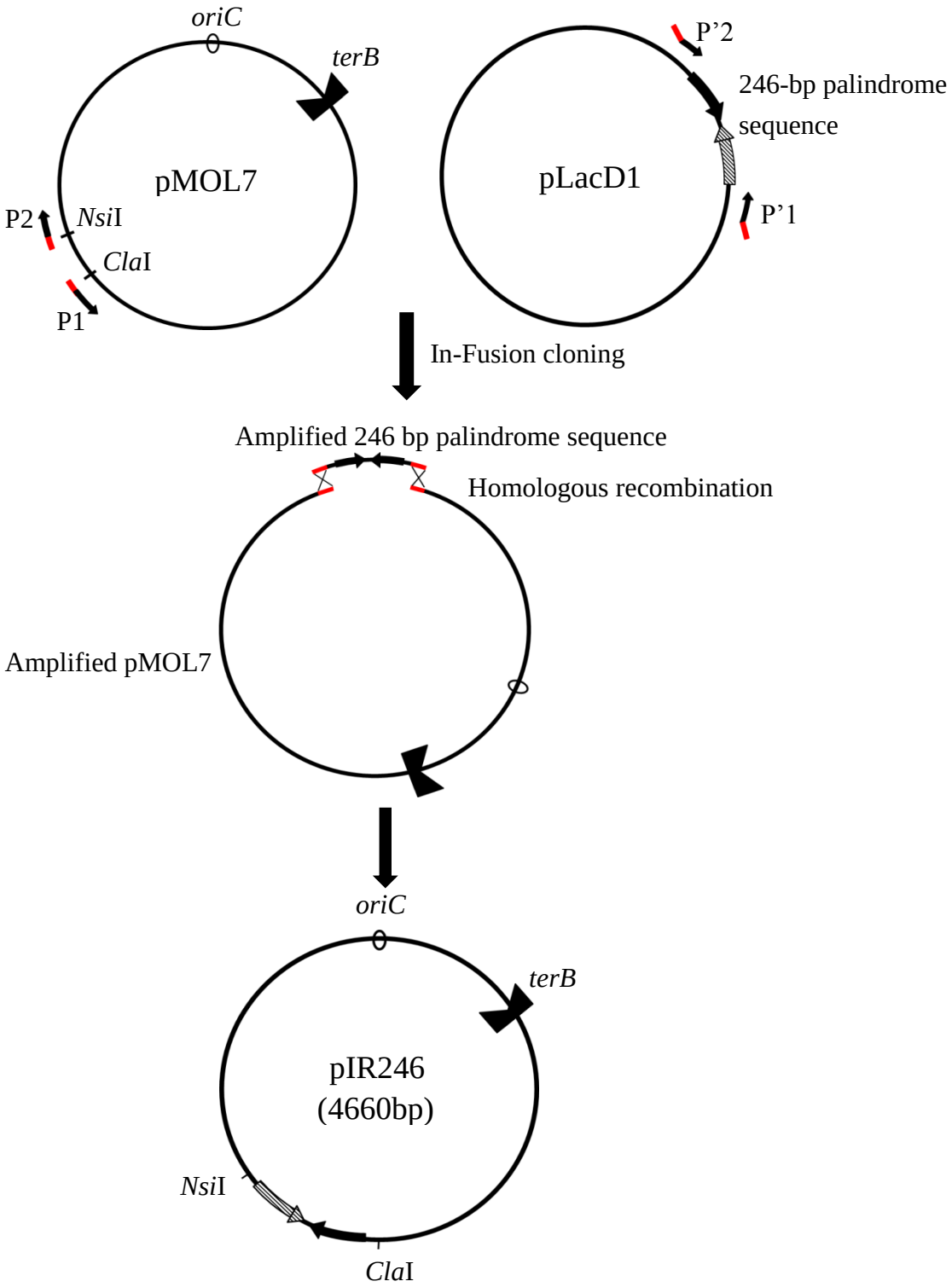


Diagram 4 **Schematic diagram showed the construction of pIR246**. A pair of PCR primer, P1 and P2 was used to amplify the vector plasmid, pMOL7. P'1 and P'2 with 15 bases of homology (as indicated in red) to the 5' end of the linearized vector plasmid, pMOL7 were used to amplify the 246 bp palindrome sequence from pLacD1. These two amplified products were combined by recognizing the 15 bp homology at their ends. The palindrome-containing *oriC-terB* plasmid is named, pIR246.

- ***oriC-terB* palindrome-free plasmid, pCNS**

An *oriC-terB* palindrome-free control plasmid was constructed by replacing one arm of the palindrome sequence in pIR246 with the synthesized oligonucleotide ordered from the company as illustrated in diagram 5. In-Fusion® HD Cloning kits (Clontech) was used for the construction of *oriC-terB* palindrome-free control plasmid due to the discontinuation production of In-Fusion™ Advantage PCR cloning Kit. The principle for both of these kits is similar except with slight modification in method which mention below. PCR amplification was carried out to amplify the vector plasmid, pIR246 as well as the synthesized oligonucleotide. In-Fusion cloning reaction was set up by mixing the amplified vector, pIR246, and amplified oligonucleotide sequence, 1× In-Fusion reaction buffer and In-Fusion Enzyme in a 10 μl reaction. This reaction mixture was incubated for 15 min incubation at 50 °C, and then placed on ice. Prior to transformation, this In-Fusion cloning reaction was diluted. Then, it was transformed into MK9200 using heat-shock transformation and grown at 30 °C in LB plate containing 50 μg/ml Ampicilin for the selection of positive clone. The *oriC-terB* palindrome-free control plasmid is named pCNS.

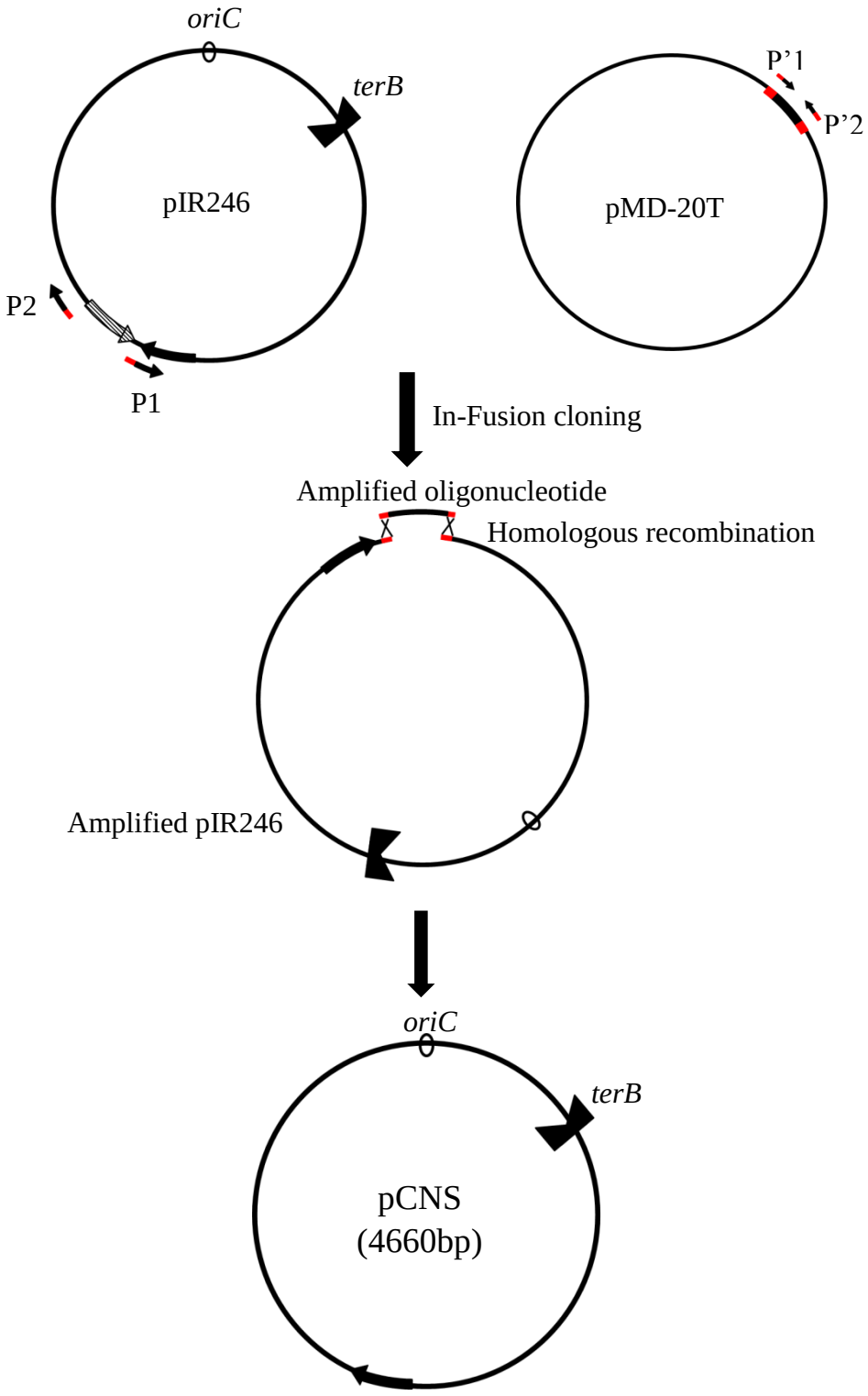


Diagram 5 **Schematic diagram showed the construction of pCNS**. PCR primer, P1 and P2 were used to amplify the vector plasmid, pIR246. P'1 and P'2 with 15 bases of homology (as indicated in red) to the 5' end of vector plasmid, pIR246 were used to amplify the synthesized oligonucleotide from pMD-20T. These two amplified products were combined by recognizing the 15 bp homology at their ends. The palindrome-free *oriC-terB* plasmid is named, pCNS.

2.2.2 Preparation of cruciform-free pIR246 with Triton-lysis method

- **Triton-lysis method** (Kurahashi *et al*., 2004)

Plasmid DNA was isolated with a non-denaturing method, Triton-lysis method. In brief, *E. coli* cells from 500 ml of overnight culture were dissolved with 100 ml of lysis buffer containing 50 mM Tris-HCl (pH7.5), 5 % sucrose, 1.5 mg/ml lysozyme, 0.1 M EDTA, 25 mg/ml RNase and 0.75 % Triton X-100. The cell lysate was then purified with the ion-exchange column (QIAGEN) following the manufacture manual. Plasmid DNA was then precipitated with 2-propanol. The plasmid DNA pellet was dissolved with TE (pH8) and stored at -80 °C until used in an experiment. All of the procedures were performed at 4 °C in a cold room to avoid spontaneous cruciform formation during the procedure.

- **Cruciform extrusion assay**

To induce the formation of cruciform structure, Triton-extracted pIR246 and pCNS (100 ng/μl) were incubated in buffer containing 50 mM Tris-HCl (pH7.9), 10 mM $MgCl_2$ and 100 mM NaCl at different temperatures for 20 min. The range of temperatures used was 25, 30, 37, 50, 65, 80 and 95 °C. The heat-treated plasmids were quickly put on ice after the incubation. The heat-treated plasmids DNA were electrophoresis at 50 V for 1.5 h in a 1 % TBE agarose gel. The electrophoresis was carried out at 4 °C. pCNS was used as control for comparison.

- **T7 endonuclease digestion**

To confirm the Triton-extracted pIR246 is cruciform free, T7 endonuclease I that can recognize and cleave the four-way junction of cruciform structure was used. Both Triton-extracted and heat-treated plasmid DNA, pIR246 and pCNS respectively (100 ng/μl) were treated with T7 endonuclease I (1 units) in 20 μl of NEB buffer 3 for 5 min at room temperature. After digestion, the plasmid DNA was subjected to 1 % TBE agarose gel for 1.5 h. The electrophoresis was carried out at 4 °C.

2.2.3 *In vitro oriC* plasmid DNA replication assay

The *oriC* plasmid used in this study carry the *E. coli* origin of replication which mimic the chromosomal replication by formation of two replication forks. Using this *oriC* plasmid, an *in vitro oriC* replication system was reconstituted with the replicative proteins involved in DNA replication in *E. coli*.

The reaction mixture (15 μl) contained 40 mM HEPES-KOH (pH 7.6), 10 mM magnesium acetate, 0.1 mg/mL BSA, 10 mM DTT, 2 mM ATP, 1 mM GTP, 1 mM CTP, 1 mM UTP, 100 μM of each deoxynucleosidetriphosphate, 4 mM creatine phosphate, 40 μg/mL creatine kinase, 50 μM NAD, 120 ng of the template DNA, DnaA (60 ng), DnaB (90 ng), DnaC (55.2 ng), DnaG (97.2 ng), DNA gyrase A subunit (108 ng), DNA gyrase B subunit (216 ng), HU (5.1 ng), SSB (174 ng), Tus (67.2 ng), DNA polymerase III* (14.4 U), β-clamp (30 ng). The reaction mixtures were assembled at 4 °C in the absence of DnaA. The reaction was initiated by the addition of DnaA protein and incubated at 30 °C for 15 min. The reaction was stopped by adding 15 μl 50 mM EDTA and 0.15 % SDS. The replication products were incubated with 0.15 % SDS and 100 μg of Proteinase K for 30 min at 55 °C in order to remove the protein. Then, these replication products were purified with phenol/chloroform, and precipitated with ethanol. In the case of SSB titration experiment (result showed in 3.2.2), the amount of SSB added was 0, 17.4, 58, 174, and 696 ng. Whereas the amount of Gyrase A/B used in the gyrase titration experiment (result showed in 3.3.1) was 0/0, 54/108, 108/216, 216/432, 432/864 and 540/1080 ng, respectively. To observe the DNA replication, the replication product was radiolabelled by the addition of 1 μCi [α-^{32}P]-dATP (800 μCi/mmol, Perkin Elmer) to the reaction. Replication products analysed by Southern blot was carried out in the absence of radioisotopes. The nucleotide incorporation was measured by scintillation counter after precipitation with trichloroacetate and filtration on to Whatman GF/C glass-fibre filters

2.2.4 Analysis by Alkaline agarose gel electrophoresis

Alkaline agarose gel performs at high pH which causes the denaturation of double-stranded DNA by breaking the hydrogen bonds. This denatured DNA is maintained as single-stranded DNA and migrates through the gel according to its size. Replication products obtained after *in vitro oriC* replication assay were run on the alkaline agarose gel. This method allows the analysis of leading and lagging strand replication product. Replication products were treated with Proteinase K as described in 2.2.3 to remove the proteins before applying to the alkaline agarose gel. The replication products were suspended in a sample buffer containing 30 mM NaOH, 30 mM EDTA, 10 % sucrose and 0.1 % bromocresol green and resolved by electrophoresis in a 0.9 % agarose gel at 3.6 V/cm for 4 h in an electrophoresis buffer of 30 mM NaOH and 1 mM EDTA. *Eco*T14I digested bacteriophage λ DNA was 5′-end labelled with [γ-^{32}P]-ATP using the T4 polynucleotide kinase and was used as a size markers. At the end of the electrophoresis, the DNA was fixed in the gel with 7 % trichloroacetic acid and dried under vacuum on the DE81 paper (Whatman), and autoradiographed with a BAS2500 Bio Image Analyser (Fuji Film).

2.2.5 Analysis by sequencing gel electrophoresis

For the sequencing gel electrophoresis analysis, replication products obtained after *in vitro ori*C replication assay were digested with *Nsi*I (3 units) and *Ase*I (3 units) in 40 ml of NEB buffer 3 for 2 h at 37 ºC. The digested replication products were purified with phenol/chloroform and followed by ethanol precipitation. The DNA was resuspended in a sample buffer containing 98 % formamide, 10 mM EDTA, 0.05 % bromophenol blue, and 0.05 % xylene cianol and separated by electrophoresis in an 8 % polyacrylamide sequencing gel (8M urea, 6.3M Formamide, TBE buffer) (Ausubel *et al*., 1994) at 35 W for 1.5 h. The *Nsi*I/ *Ase*I digested template pIR246 and pCNS comprise of 348 and 357 mer oligo DNAs were 5′-end labelled with ^{32}P and used as size markers. BAS 2500 Bio Image Analyser was used to

visualize the resolved patterns of product DNA.

2.2.6 Southern blot analysis

DNA replication was carried out in the absence of radio-labelled nucleotides. For alkaline agarose gel Southern blotting, the replication products were digested with *Bgl*I (3 units) in 40 µl of NEB buffer 3 for 2 h at 37 ºC. The digested product DNAs were subjected to alkaline agarose gel electrophoresis as described above. Then this replication products were transferred onto a Hybond N+ membrane (Amersham Biosciences) by the capillary method with 10× SSC buffer as previously described (Ausubel *et al*. 1994). In case of the sequencing gel southern blotting, replication products were digested in 40 µl reaction volume containing with *Msc*I (3 units), *Ase*I (3 units) and 1× NEB buffer 3. These digested products were applied to the sequence gel electrophoresis as described above and transferred on to a Hybond N+ membrane using the TE 90 GeneSweep^{Tm} Sequencing Gel Transfer Unit (Hoefer Scientific Instruments).

After transferring, the membrane was exposed to 120 mJ UV. This was incubated at 55 °C for 2 h in a buffer containing 5× Denhalts' reagent, 6× SSC, 0.5 % SDS and 100 µg/mL of Salmon sperm testes DNA. After 2 h of incubation, ^{32}P-labelled oligomer DNA probe with approximately 2 x 10^6 CPM value were added and continue to incubate in the same buffer at 55 °C for 12 h. The membrane was washed four-times with a buffer (0.1× SSC, 0.1 % SDS) and autoradiographed with a BAS2500 Bio Image Analyser. Quantitative analysis of the autoradiogram was carried out using Image Gauge ver. 3.3 (Fuji Film). The probes used are listed in Table 5B.

Chapter 3

Results

3.1 Construction and preparation of cruciform free *oriC-terB* palindrome-containing plasmid, pIR246

In order to examine the effect of palindrome sequence during the concurrent synthesis of leading and lagging strand, an *oriC* plasmid carrying the palindrome sequence was constructed. pMOL7, an *oriC* plasmid carrying the chromosomal replication origin of *E.coli*, which was previously constructed by Higuchi *et al.* (2003) was used as a vector for the construction of palindrome-containing *oriC* plasmid. A 246 bp palindrome sequence containing the 111 bp inverted repeat separated by a 24 bp spacer was previously isolated from lambda SKK43 DNA (Kulkarni, 1990). This 246 bp palindrome sequence inserted in pLacD1 (Eykelenboom *et al*., 2008) was given by Dr. David Leach.

The palindrome-containing *oriC* plasmid was prepared by inserting the 246 bp palindrome sequence between the restriction sites, *Nsi*I and *Cla*I in the pMOL7. This palindrome-containing *oriC* plasmid is named pIR246 (Diagram 6A). In addition, a palindrome-free control plasmid was constructed by replacing one arm of the 246 bp palindrome in pIR246 with the synthesized oligonucleotide. This palindrome-free control plasmid is named pCNS (Diagram 6B).

Previous study suggested that long palindrome DNA was not able to maintain in wild type *E. coli* (Collins 1981; Hagen and Warren 1982; Leach and Stahl 1983; Sinden *et al.* 1991). However, it is stably maintained in *sbcCD* mutant *E. coli* strain (Leach 1994). Thus, I introduced pIR246 as well as pCNS into the *E. coli* strain with *sbcC* mutation for the maintenance of these plasmids. Sequencing the plasmid DNAs, confirmed that the full length 246 bp palindrome sequence was successfully inserted into the *oriC* plasmid and that the

palindrome-free control plasmid, pCNS was successfully constructed.

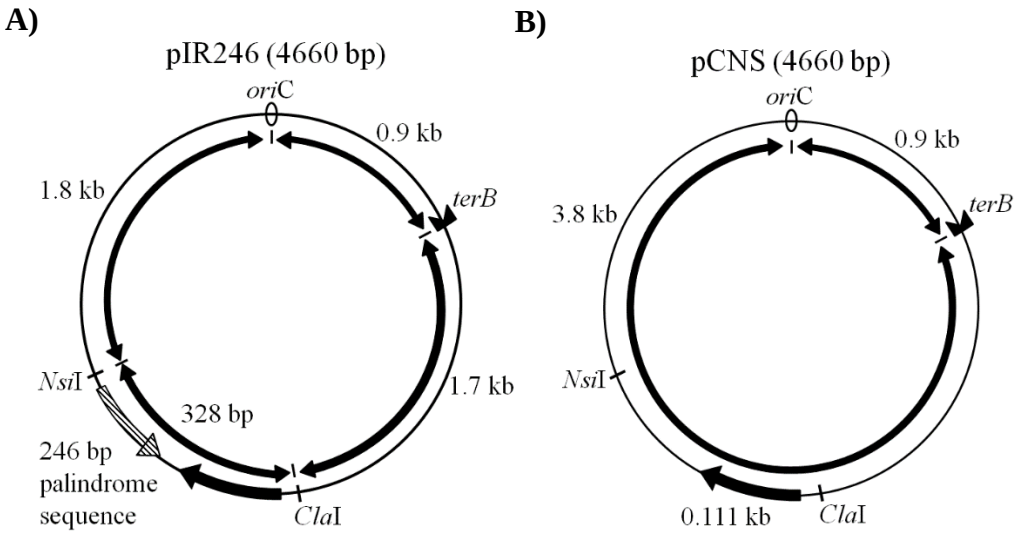


Diagram 6 **Schematic diagram of o*riC* plasmid used in this study.** (A) Palindrome-containing *oriC* plasmid named pIR246 was constructed in this study. This plasmid was constructed by cloning a 246 bp palindrome containing 111 bp inverted repeats separated by a 24 bp spacer into pMOL7, an *oriC* plasmid carrying the chromosomal replication origin of *E .coli*. (B) Palindrome-free control plasmid was constructed by replacing one arms of the palindrome sequence. This palindrome-free control plasmid is named pCNS.

3.1.1Triton-lysis method for the extraction of plasmid

Alkaline lysis method (denaturing method) is one of the most common methods used for the preparation of plasmid. However, this method involves the procedures of denaturation and reannealing of DNA strand. Therefore, this method has the possibility causing the formation of cruciform structure from the palindrome sequence. Kurahashi *et al.* (2004) showed that PATRR-containing plasmid (palindromic AT-rich repeats) exhibited cruciform extrusion when prepared with alkaline lysis method. However, PATRR-plasmids did not demonstrate cruciform extrusion when prepared by the Triton-lysis method which is a non-denaturing method.

In my study, I would like to investigate the behaviors of palindrome sequence during DNA replication. Therefore, it is very crucial to have a cruciform-free palindrome-containing plasmid in this study. Thus, Triton-lysis method was used for the preparation of cruciform-free pIR246 plasmid. In this method, cells were partially lysed to allow the plasmid molecules to escape, while the genomic DNA still trapped in the cell debris. The plasmid DNA was then recovered by isoprapanol precipitation. Figure 1 showed the image for the standard agarose gel electrophoresis of the Triton-extracted pCNS (lane 8) and pIR246 (lane 16), respectively. From the agarose gel result, it was observed that the Triton-extracted plasmids were separated into mainly two bands. The fastest migrating band represents the supercoiled DNA. While slow migrating bands may represent the open circular DNA.

3.1.2 Effect of heat on the formation of cruciform structure on pIR246

Palindrome sequence has the potential to transform from a lineform to a cruciform structure in double helix DNA. The extrusion of cruciform structure is initiated with the melting of DNA base pairs at the center of symmetry of a palindrome sequence forming the open regions on the loops. Subsequently, the unwinding of complementary strands allow the intrastrand base

pairing of the self-complementary sequence and make the existence of cruciform structure (Sinden 1994).

The extrusion of cruciform structure is a process driven by energy. This energy is needed to melt the center of palindrome sequence for further intrastrand base pairing to form the cruciform structure. The energy required to extrude the cruciform structure is originated from the negative DNA supercoiling. However the formation of cruciform structure is energetically unfavored in a relaxed DNA molecule (Lilley, 1980; Panayotatos and Wells, 1981; Mizuuchi *et al*., 1982).

Several factors are suggested to involve in the extrusion of cruciform structure, for instance, the base composition, temperature and salt (Sinden 1994). Since the preparation of palindrome-containing *oriC* pIR246 was succeeded, I then examined at what condition favor the extrusion of cruciform structure. In the current study, the base composition of the 246 bp palindrome sequences is fixed, thus I had examined the effect of temperature in affecting the extrusion of cruciform structure. Figure 1 showed the image for the agarose gel electrophoresis. Triton-extracted pCNS and pIR246 were mainly in a supercoiled form as shown in lane 8 and 16, respectively. These plasmids were incubated at different temperature for 20 min in the buffer containing 50 mM Tris-HCl (pH7.9), 10 mM $MgCl_2$ and 100 mM NaCl. It was observed that Triton-extracted pIR246 remained mostly as supercoiled structure after the incubation at room temperature till 50 °C (lane 9-12). However, mobility retardation of this supercoiled pIR246 was detected when incubated at higher temperature ranged from 65- 95 °C (lane 13-15). For negatively supercoiled DNA, a reduces in the twisting number would change the degree of supercoiling that consequently cause a reduction in electrophoretic mobility. Transition of palindrome sequence to cruciform structure involves unpairing of the complementary and consequently reduces the interstrand twist number. Thus, a character of laddering exhibited by this heat-treated pIR246 may reflect the various linking numbers of the plasmids. In contrast to

that, no mobility retardation was detected in the case of palindrome-free control plasmid, pCNS after incubation at various temperatures (Lane 1-7).

From this observation, it is suggested that the laddering bands observed in pIR246 is a result of a transition from a lineform palindrome sequence to a cruciform conformation. Apart from that, the formation of cruciform is found more prominent at higher temperature.

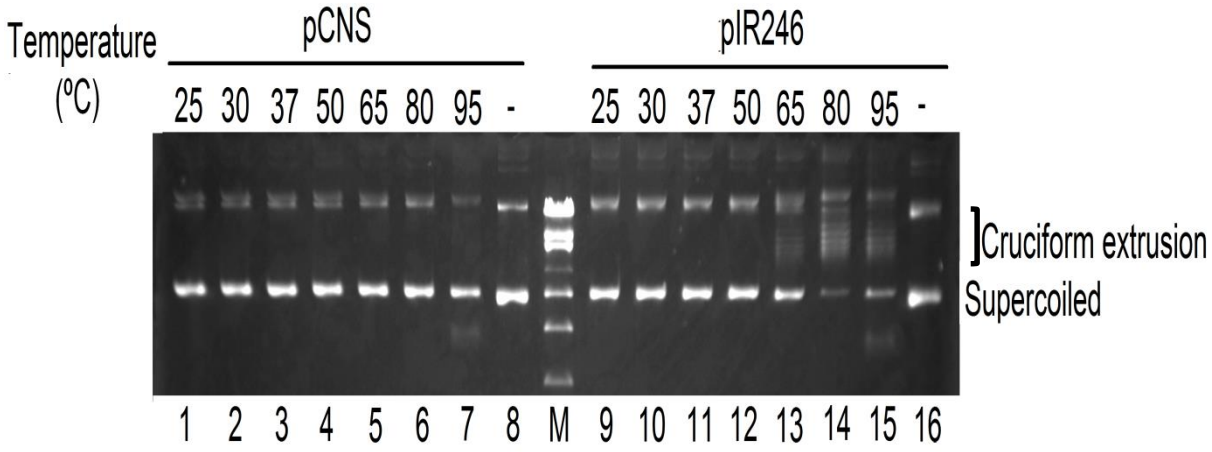


Figure 1 **Effect of temperature on the formation of cruciform in pIR246.** Triton-extracted pIR246 and pCNS were incubated in buffer (50 mM Tris-EDTA pH 7.9, 100 mM NaCl, 10 mM $MgCl_2$) at different temperature ranging from 25 to 95 °C. Gel images showed the effect of temperature on pCNS (lane 1-7) and pIR246 (lane 9-15). Lane 8 and 16 showed the Triton-extracted pCNS and pIR246 which occurred mainly as supercoiled plasmid, respectively. A character of laddering indicated with the black bracket was observed in pIR246 after incubation in buffer at higher temperature (lane 13-15). No character of laddering was observed for pCNS across the different temperature (lane 1-7).

3.1.3 Triton-extracted pIR246 is cruciform free

A cruciform structure consists of three components: the loops, the stems, and the four-way junction. This cruciform structure is known as putative substrates for several structure-specific nucleases. S1 or P1 nucleases are single-strand-specific nucleases that cleave the bases of loop regions (Lilley 1980; Haniford and Pulleyblank 1985). Whereas, the branched junction of cruciform DNA has been known as a substrate for the endonucleolytic activity of T4 endonuclease VI1 (Mizuuchi *et al*., 1982; Lilley and Kemper 1984) and T7 endonuclease I (Dickie *et al*., 1987; de Massey *et al*., 1987). To confirm the absence of cruciform structure in the Triton-extracted preparation of pIR246, their sensitivity to T7 endonuclease was examined. T7 endonuclease is capable to recognize the branched molecule and introduce a pair of nicks near to the base of the cruciform structure which subsequently cause the linearization of the template DNA (De´clais *et al*., 2003). If the cruciform structure formed by 246 bp palindrome sequence is digested by the nuclease, the pIR246 plasmid DNA will become linearized (Figure 2A).

Triton-extracted and heat-treated pIR246 were digested with T7 endonuclease I and analysed by agarose gel electrophoresis (Figure 2B). Lane 1 and 2 were Triton-extracted pIR246 and pCNS, respectively, which occurred mainly as supercoiled plasmid. Lane 3 and 4 showed the heat-treated pIR246 and pCNS, respectively. Heat-treated pIR246 showed a character of laddering while heat-treated pCNS remained as supercoiled plasmid. The size of the linearized pIR246 and pCNS were approximately 4.7 kb after the digestion with restriction enzyme, *Nsi*I, as shown in lane 9 and 10, respectively. When the heat-treated pIR246 (lane 3) was treated with T7 endonuclease I, a major product was the linearized DNA with approximately 4.7 kb (lane 5). In contrast, only a small amount of digested product was detected in heat-treated palindrome-free pCNS digested with T7 endonuclease I (lane 6). This result convincingly demonstrates that the mobility retardation for heat-treated pIR246 observed in Figure 1 correlates to the existence of a cruciform structure.

The Triton-extracted pIR246 was shown resistant to T7 endonuclease (lane 7). The T7 endonuclease treated Triton-extracted pIR246 remained mainly as supercoiled DNA as well as the Triton-extracted pCNS (lane 8). However, there was a small amount of linearized product was detected on Triton-extracted pIR246 and pCNS. This linearized product is probably due to the non-specific digestion of T7 endonuclease. Taken together with this, I concluded that pIR246 prepared with Triton-extraction method is free from cruciform structure.

In order to examine the behaviours of palindrome sequence during DNA replication, having a cruciform-free palindrome-containing plasmid is very crucial in my study. Using the Triton-lysis method, a supercoiled palindrome-containing pIR246 is obtained. This palindrome-containing pIR246 is capable to form cruciform structure and this formation is temperature sensitive. It is found that the formation of cruciform structure is most efficient with higher temperature. Most importantly, palindrome-containing pIR246 prepared using the Triton-lysis method is shown resistant to T7 endonuclease. Hence, this result convincingly confirms that Triton-extracted pIR246 is cruciform free.

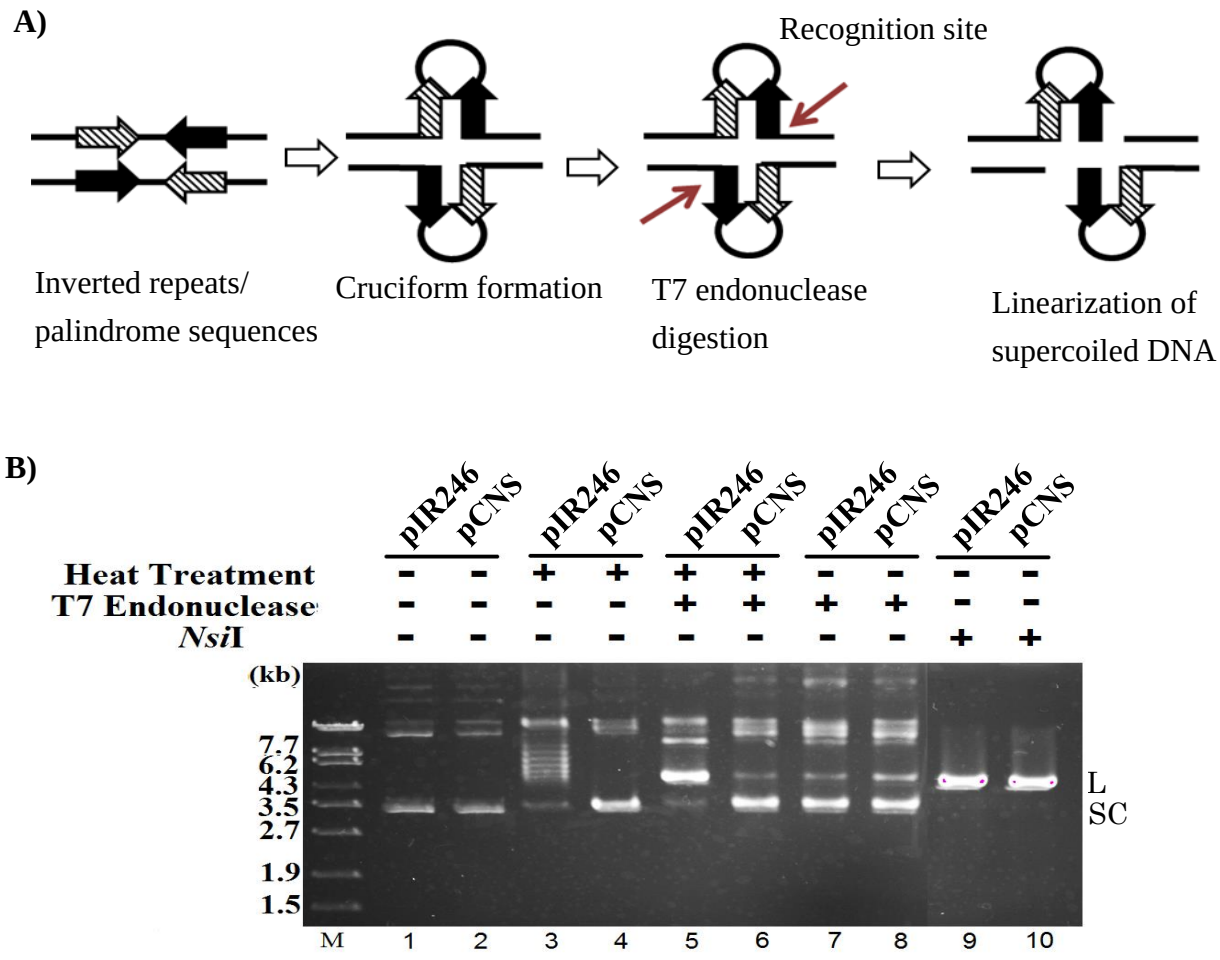


Figure 2 **Triton-extracted pIR246 is cruciform free.** (A) Schematic representation of T7 endonuclease digestion. If the 246 bp palindrome sequence adopts the cruciform structure, it will be cut into two fragments by a diagonal cleavage with T7 endonuclease. (B) Triton-lysis method was used to extract pIR246 and pCNS. Lane 1- 2 showed the supercoiled plasmid of Triton-extracted pIR246 and pCNS. Lane 3 showed the electrophoresis mobility shift caused by cruciform extrusion in heat-treated palindrome-containing plasmid, pIR246. Heat-treated pCNS remained as supercoiled plasmid (lane 4). Lane 5 showed that heat-treated pIR246 was linearized by T7 endonuculease I, while the heat-treated pCNS (lane 6) remained as supercoiled structure after T7 endonuclease I digestion. Lane 7 showed that Triton-extracted pIR246 and pCNS (lane 8) was resistant to T7 endonuclease I digestion. Lane 9-10 showed the single restriction digestion by *Nsi*I on pIR246 and pCNS, respectively. L and SC represent linearized and supercoiled plasmid, respectively.

3.2 Analysis of product DNA synthesized in *oriC* DNA replication *in vitro* with a standard condition.

In *E. coli*, chromosomal DNA replication initiates at the origin of replication, *oriC*. From the *oriC* region, DNA replication proceeds bidirectionally around the circular chromosome and ends in the terminus region nearly diametrically opposite to the *oriC* region (Prescott and Kuempel 1972; Bird *et al*., 1972). Reconstitution of *in vitro oriC* replication has become possible by having the *oriC*-containing plasmids (Yasuda and Hirota 1977; Oka *et al*., 1980) and the development of a soluble enzyme system (Fuller *et al*., 1981). Later on this crude soluble enzyme system was replaced by the purified replicative proteins for reconstitution of *in vitro oriC* replication (Kaguni and Kornberg 1984). This system is characterized by its ability for bidirectional replication from the *oriC* region as well as the concurrent synthesis of leading and lagging strand DNA (Kaguni *et al*., 1982; Hiasa and Marians 1994, 1999).

It is the first time that Higuchi *et al* (2003) converted the bidirectional *in vitro oriC* plasmid replication system into a semi-bidirectional replication system by introducing the *terB* sequence into the *oriC*-containing plasmid. When Tus protein binds to the *terB* sequence, replication fork moving clockwise is blocked. While the counter-clockwise replication fork continues on with replication until it is terminated at the *terB* sequence. This semi-bidirectional replication system is very useful to permit the tracing of replication fork that move in a particular direction from the *oriC* sequence. Having this established semi-bidirectional replication system, I examined the effects of 246 bp palindrome sequence on the progression of replication fork.

3.2.1 Alkaline agarose gel and sequencing gel analysis

Having confirmed that Triton-extracted pIR246 is cruciform-free, *in vitro oriC* replication assay was reconstituted to examine the effect of palindrome sequence on the progression of replication fork. Replication products obtained was applied onto the alkaline agarose gel electrophoresis, and the autoradiogram was shown in Figure 3B. In this semi-bidirectional replication system (Figure 3A), binding of Tus protein to the *terB* site blocks the clockwise replication and thus gave rise to the production of 0.9 kb DNA as the leading-strand replication product on both pIR246 and pCNS, respectively, as shown in Figure 3B. Whereas the counter-clockwise replication fork continued on with replication until it terminated at the *terB* sequence lead to the production of counter-clockwise leading-strand replication product with the size of approximately 3.8 kb as can be observed in pCNS and pIR246. Replication products range between 100-600 bp represents the Okazaki fragments as detected in both pIR246 and pCNS. From the result, it was observed that both leading-strand replication product as well as total Okazaki fragments product was not affected within the 246 bp palindrome sequence in pIR246.

Previous study suggested that hairpin structure can be formed on lagging strand and this structure can cause the pausing of polymerase. Thus, I would like to examine the lagging strand synthesis when it encounters the palindrome sequence. Using the alkaline agarose gel is not suitable for the detail analysis of lagging strand replication product due to the limitation of gel resolution. Therefore, sequencing gel which has better gel resolution was used for the in depth examination of lagging strand synthesis within the palindrome sequence. Replication products obtained from the *oriC* replication assay was digested with restriction enzyme, *Nsi*I and *Ase*I and applied onto the sequencing gel electrophoresis. The size of the replication products generated after restriction digestion is approximately 348 and 354 mer for leading-strand and lagging-strand replication product, respectively (Figure 4A). Assuming if palindrome sequence affects lagging or leading strand synthesis, the lagging-strand and leading-strand replication paused product is approximately 68 mer (Figure 4B) and 36 mer (Figure 4C), respectively.

Figure 4D showed the result for sequencing gel electrophoresis analysis. Lane 3 and 4 was the ^{32}P-labeled *Nsi*I/*Ase*I digested pIR246 and pCNS template as a marker, correspondingly. Lane 1 and 2 showed the digested replication products of pIR246 and pCNS with near identical size as shown in a single band whereby the size of leading-strand replication product is 348 mer and lagging-strand replication product is 354 mer. This result suggests that in pIR246, both leading-strand and lagging-strand replication can synthesized beyond the palindrome sequence. However, the estimated size of the replication paused product was not detected. In this observation, no significant pausing on leading and lagging strand replication was observed.

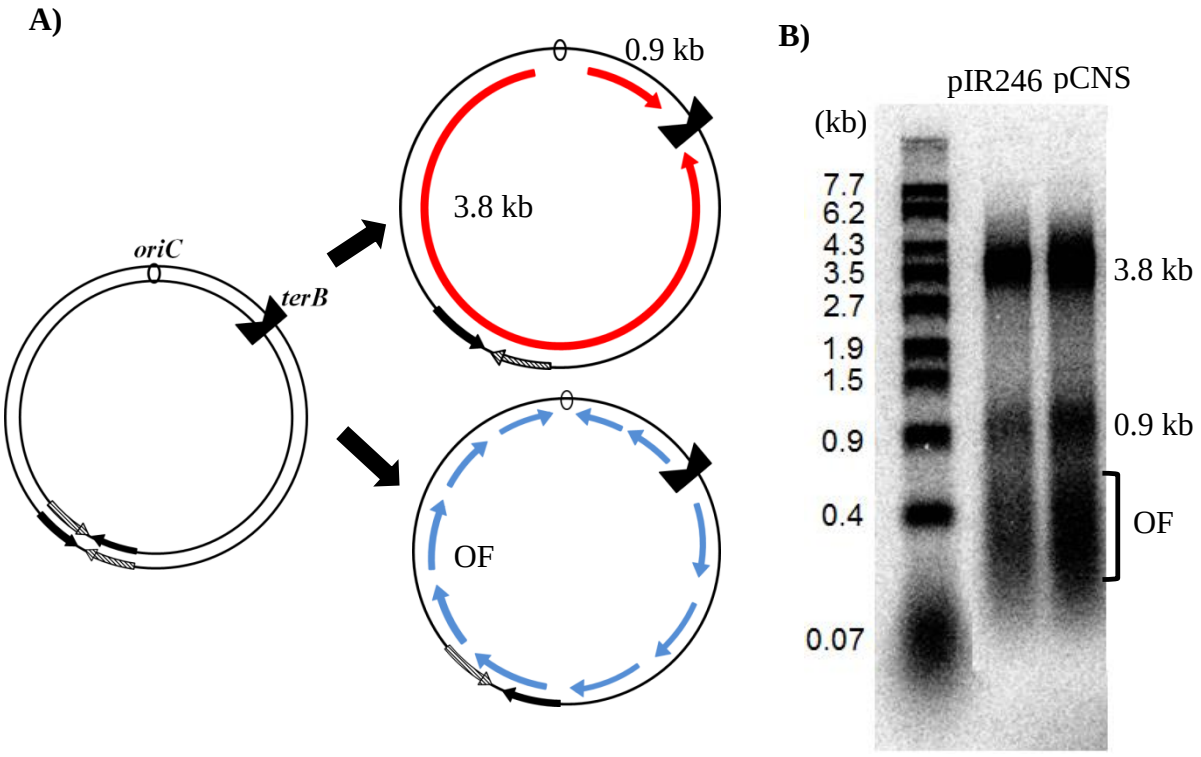


Figure 3 **Alkaline agarose gel result showed the effect of palindrome sequence on DNA chain elongation in leading- and lagging-strand DNA synthesis.** (A) Schematic diagram indicate the mode of semi-bidirectional DNA replication occurred in the majority of the template when it encountering the palindrome sequence. If DNA can synthesized beyond the palindrome sequence the size of the replication product is 3.8 kb while the clockwise replication product is 0.9 kb. OF indicates the Okazaki fragments, which ranged in size from 100 to 600 residues. Thin black lines indicate template strand DNA. Thick black lines with arrowheads indicate the 246 bp palindrome sequence. Red and blue lines with arrowheads indicate the leading-strand and the lagging-strand DNA synthesis, respectively. (B) The replication products were subjected to alkaline agarose gel analysis. The size of the counter-clockwise replication products was approximately 3.8 kb in both pIR246 and pCNS. Replication products with the size of approximately 0.9 kb represent the counter-clockwise replication product due to the binding of Tus to the *terB* site. The size of Okazaki fragments is ranged between 100-600 bp.

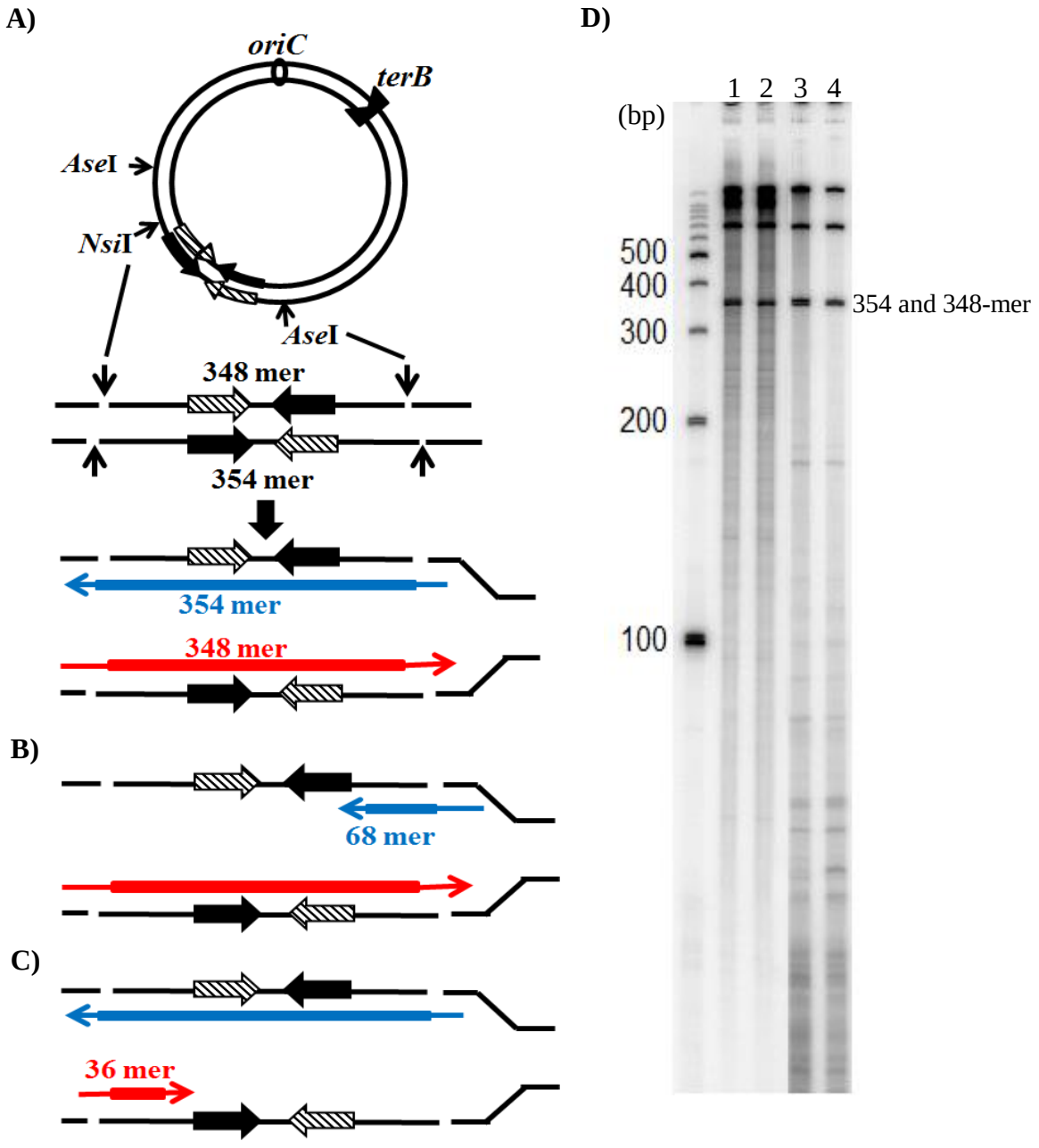


Figure 4 **Effect of inverted repeat on DNA chain elongation in leading- and lagging-strand DNA synthesis.** The *oriC* replication reaction was carried out for 15 min as described in the Experimental procedure. The replication products obtained were digested with *Nsi*I and *Ase*I and subjected to sequencing gel electrophoresis analysis. (A) The 246 bp palindrome sequence was integrated into the *oriC* plasmid, pMOL7 in between the *Nsi*I-*Ase*I site. The restriction-enzyme digestion of replication products prepared with intact template DNA causes hemireplicated *Nsi*I-*Ase*I cassette fragments to be excised from the products. In denaturing gel analysis, radio-labelled shorter DNAs (354 and 348 mer) are expected to be detected as a product of chain elongation on the lagging-strand template and one on the leading-strand template within the *Nsi*I-*Ase*I segment. (B) Assuming that the inverted repeat sequence on

lagging strand template blocks chain elongation on lagging strand, the restriction-enzyme digestion of lagging strand replication products would produce a 68 mer instead of the 354 mer. (C) If leading strands chain elongation is blocked by the inverted repeat on the leading strand template, the restriction-enzyme digestion of leading strand replication product is approximately 36 mer instead of 348 mer as illustrated. (D) An autoradiogram of the shorter DNAs excised from the replication products prepared with pIR246 (lane 1) and pCNS (lane 2). Lane 3 and 4 showed the ^{32}P-labelled *Nsi*I/*Ase*I digested pIR246 and pCNS template, respectively. Thin black line represents the template DNA strand. Thick black lines with arrowheads indicate the palindrome sequence. Blue and red lines with arrowheads represent the lagging and leading strand product respectively.

3.2.2 Effects of single-stranded binding protein (SSB)

The DNA polymerase III holoenzyme functions as an asymmetric dimer with two polymerase III cores that are responsible for the concurrent synthesis of leading and lagging strand (McHenry 1988). Owing to the anti-parallel structure of DNA and the 5'-3' direction of DNA chain elongation catalysed by DNA polymerase, DNA replication is characterized as the semi-discontinuous reaction. Therefore, leading strand is synthesized continuously towards which the DNA double helix is unwind, while lagging strand is synthesized discontinuously in the opposite direction resulting in numerous short DNA fragment knows as Okazaki fragment. The discontinuous synthesis of lagging strand transiently exposed the lagging strand template into the single-stranded loop with the size equivalent to the Okazaki fragment (Alberts *et al*., 1983)

Numerous *in vitro* DNA synthesis studies using the prokaryotic (Sherman and Gefter 1976, Challberg and Englund 1979; LaDuca *et al*., 1983; Huang *et al*., 1981) as well as the eukaryotic (Villani *et al*., 1981; Kaguni and Clayton 1982) DNA polymerases are suggested to be affected at specific sequence in the single-stranded DNA template. It is suggested that the arrest site corresponds to the palindromic sequence that possibly to form a stable hairpin structure. The addition of SSB has been shown to eliminate the barriers formed by the hairpin structure (LaDuca *et al*., 1983; Huang *et al*., 1981; Hacker and Alberts 1994) to facilitate replication through this barrier.

It has been hypothesized that hairpin structure is formed on lagging strand template caused the pausing of lagging strand synthesis. However result in 3.2.1 showed no significant pausing of lagging strand synthesis within the palindrome sequence. This observation may suggest that the presence of SSB in the reaction is sufficient to remove the hairpin structure as SSB is readily to binds on the ssDNA to prevent the formation of secondary structure. In view of this, I next examined the effect of SSB in relation to lagging strand synthesis when it encountering the 246 bp palindrome sequence.

In vitro oriC replication assay was carried out with different amount of SSB on both pIR246 and pCNS. Replication products obtained after 15 min incubation at 30 °C were digested with restriction enzymes, *Nsi*I and *Ase*I. The digested replication products were then applied onto the sequence gel electrophoresis. Figure 5 shows the result of the effect of SSB to lagging strand synthesis when it encounters the palindrome sequence. Lane 1 and 2 showed the ^{32}P-labelled *Nsi*I/*Ase*I digested pCNS and pIR246 template as a marker, respectively. Lane 3-7 showed the result for pCNS with different amount of SSB. Both leading and lagging strand replication product with the size of approximately 348 and 354 mer were detected on the sequencing gel. However, it was found that the overall amount of leading and lagging strand is less when in the absence of SSB (lane 3).

It is hypothesized that when the amount of SSB is less the possibility of hairpin formation will be increased and vice versa. Hence, assuming if hairpin is formed and block lagging strand synthesis, lagging-strand replication-paused product is approximately 68 mer (Figure 4B). Lane 8-12 showed the result for pIR246 with different amounts of SSB. Both leading-strand and lagging-strand replication products with the size of 348 and 354 mer were detected when different amount of SSB were used. In the absence of SSB (lane 8), the overall leading and lagging strand replication product is less comparing to others (lane 9-12). This phenomenon is the same as observed in pCNS (lane 3). One of the explanations for this phenomenon is probably due to the less efficiency of initiation of DNA replication when in the absence of SSB.

From the result, no significant of paused product was detected even with the low amount of SSB. This result suggests that lagging strand is readily synthesized beyond the palindrome sequence.

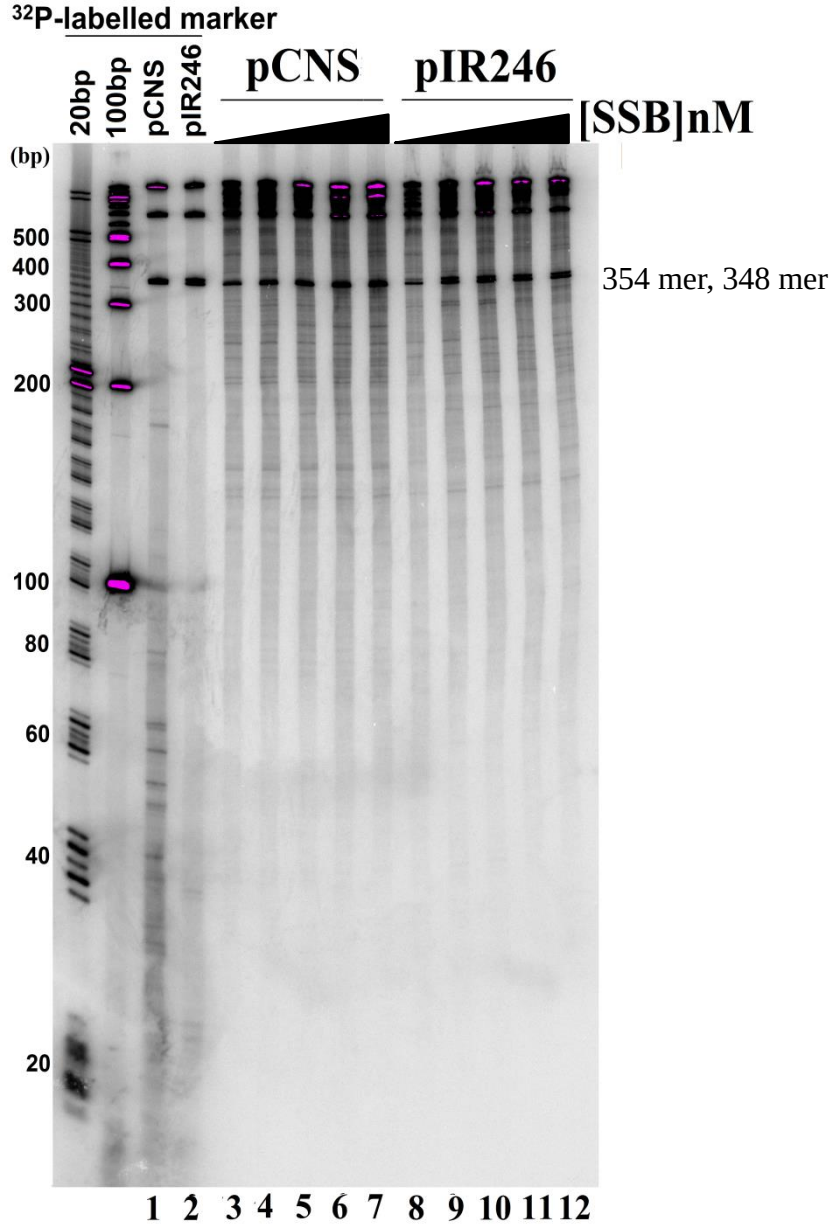


Figure 5 **Effect of SSB to lagging strand synthesis when it encountering the palindromes sequence.** The *oriC* replication reaction was carried out with various amount of SSB (0, 15.2, 50.7, 152 and 608 nM). The replication products obtained were digested with *Nsi*I and *Ase*I then subjected to sequencing gel analysis. Lane 1 and 2 showed the ^{32}P-labelled *Nsi*I/ *Ase*I digested pCNS and pIR246 template, respectively. Lanes 3-7 are the replication products for pCNS while lane 8-12 are the replication products for pIR246 that obtained when different amount of SSB were used, respectively. Replication products with the size of 354 and 348 mer were detected as a product of chain elongation on the lagging-strand template and one on the leading-strand template within the *Nsi*I- *Ase*I segment in both pCNS and pIR246. No significant paused products were detected in pIR246 even with low amount of SSB.

3.2.3 Southern blot analysis of leading strand replication product

Analysis using Alkaline agarose gel and sequencing gel showed no significant pausing of replication within the palindrome sequence. This suggests that both leading and lagging strands are readily synthesized beyond the palindrome. To further confirm this, Southern blotting analysis was carried out. Replication products obtained after *in vitro oriC* replication at 30 °C for 15 min were digested with restriction enzyme, *Bgl*I. This digested replication products were subjected to the alkaline agarose gel followed by Southern blot analysis (Figure 6B). Probe B specifically detecting the leading strand replication product was used (Figure 6A). A 4.7 kb DNA product was detected with Probe B on both pCNS and pIR246 which representing the template DNA indicated as T in Figure 6B. The full length leading strand replication product with the size of approximately 2.7 kb was detected in pCNS as well as pIR246. This suggested that the leading strand can be synthesized across the palindrome sequence.

However, it was surprising to observe a very faint but yet detectable short leading-strand replication product with the size of approximately 0.9 kb only in pIR246 but not pCNS. This detectable 0.9 kb replication product represent the leading-strand replication product from the *Bgl*I site to the far end of the palindrome sequence as illustrated in Figure 6A. The location of this leading-strand replication paused product is still uncertain and the detected band intensity is also very faint. From this observation, it is suggests that the presence of palindrome sequence affected the leading strand synthesis to some extent.

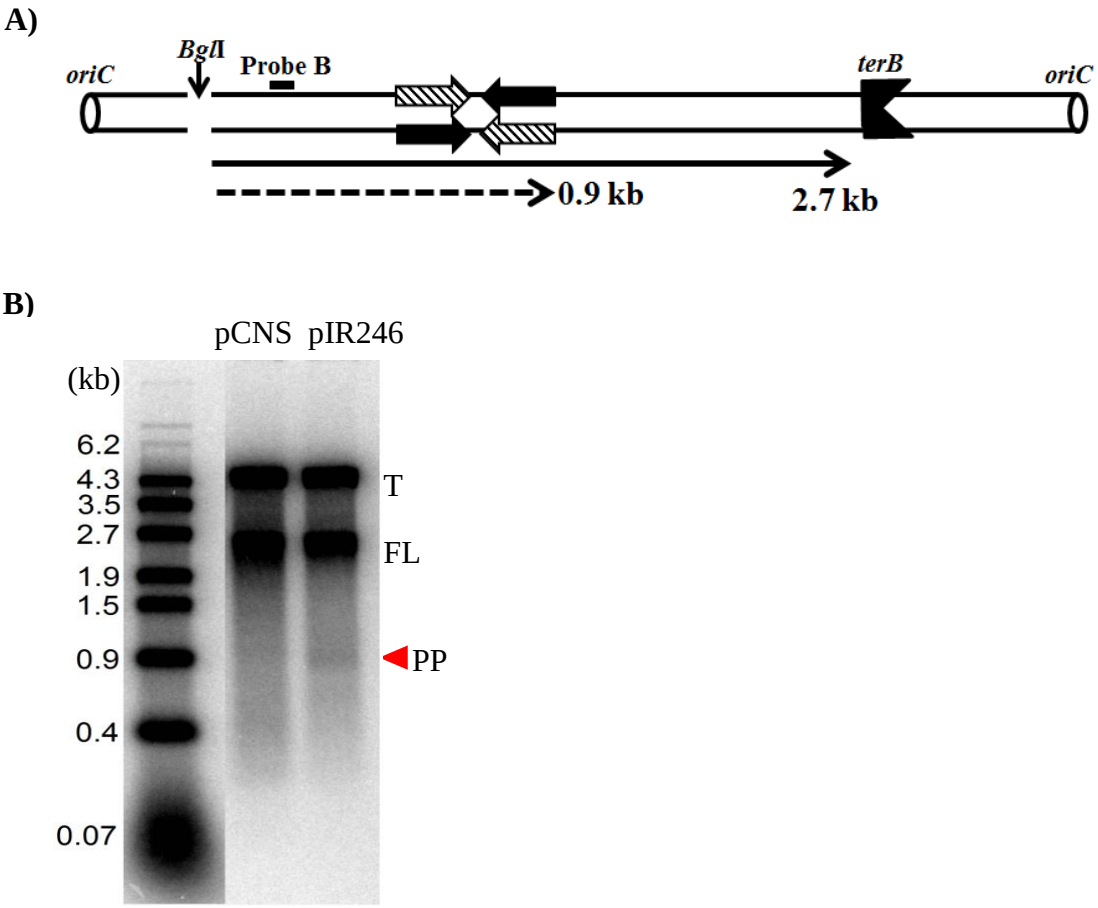


Figure 6 **Leading-strand replication-paused product was detected on pIR246**. Replication products obtained as described in experimental procedure were digested with *Bgl*I and subjected to alkaline agarose gel southern blot analysis. (A) Schematic diagram showed the location of probe used to detect the leading-strand replication products. Thin black lines represent the template DNA. Thick black lines with arrowheads indicate the 246 bp palindrome sequence. Black line with arrowhead indicates the full length leading strand product with the size of 2.7 kb. Black dashed line with arrowhead indicates the short leading strand replication product with the size of 0.9 kb. (B) Southern blot result showed that leading strand synthesis paused within the inverted repeat. T, FL and PP represent template, full length replication product and paused product, respectively.

3.3 The effects of DNA gyrase in affecting DNA synthesis on *oriC* plasmid containing the palindrome sequence

Result showed in 3.2 demonstrated that lagging strand synthesis is not significantly paused within the palindrome sequence. This result is contrast to the previous observation which suggested that lagging strand synthesis is paused at the hairpin structure. However the Southern blotting result (Figure 6) described in the previous section suggests that leading strand replication is affected by the palindrome sequence. It has been confirmed that Triton-extracted pIR246 plasmid used in this study is cruciform-free (Figure 2). Therefore from this observation, I hypothesized that the stalling of leading strand replication is probably due to a cruciform structure produced ahead of the replication fork after the initiation of DNA replication.

During the progression of replication fork through the double-stranded DNA, strand separation by helicase results in the accumulation of positive superhelical turns ahead of the replication fork (Alexandrov *et al*., 1999; Peter *et al*., 1998). The accumulation of this torsional stress needs to be released in order for further DNA replication. DNA gyrase, a type II DNA topoisomerase, is composed of two subunits encoded by *gyrA* and *gyrB* genes. This enzyme exists in solution as an A_2B_2 tetramer. Each of the subunits is presence with approximately 250 and 150 molecules per cells for *gyrA* and *gyrB,* respectively (McMacken *et al*., 1987). The physiological concentration of DNA gyrase is approximately 178 nM with the cell volume of *E. coli* is approximately 0.7 μm^3. It is the enzyme that responsible for introducing the negative supercoiling in DNA (Gellert *el al*., 1976). It is shown that DNA gyrase participates to release the positive supercoils formed ahead of the replication fork in order for the progression replication fork (Khodursky *et al*., 2000; Levine *et al*., 1998).

Thus, it seemed possible that the involvement of DNA gyrase in releasing the torsional stress ahead of the replication fork may contribute to the formation of cruciform structure within the palindrome sequence. Next, I examined the effect of DNA gyrase in causing the pausing on leading strand synthesis.

3.3.1 Titration of DNA gyrase

In vitro oriC replication assay was carried out with different amounts of DNA gyrase on both pCNS and pIR246. Replication products obtained were applied onto the alkaline agarose gel electrophoresis (Figure 7B). In the absence of gyrase, counter-clockwise replication was inhibited as can be observed in pCNS (Figure 7B, lane 1). However, only the 0.9 kb clockwise replication product and Okazaki fragments can be detected. The inhibition of counter-clockwise replication observed in pCNS is probably due to the accumulation of torsional stress ahead of the replication fork in the absence of gyrase. When gyrase was added to the reaction as shown in lane 2, a small amount of 3.8 kb counter-clockwise replication product can be detected. This suggested that torsional stress accumulated in front of the replication fork is released by the gyrase. However, this amount of gyrase may not be sufficient in completely releasing the torsional stress. When the amount of gyrase was increased (lane 3-6) there was an increment of the 3.8 kb counter-clockwise replication product. This suggests that when the amount of DNA gyrase is sufficient, the replication fork speed is increased and thereby enable the recycling of replication proteins to another unused template for DNA replication. In addition, it was observed that when there is excess amount of gyrase (lane 5-6), the size of the Okazaki fragment are slightly larger than with standard amount of gyrase was used (lane 3). This phenomenon is suggested relating to the replication fork speed. Higuchi *et al*. (2003) showed that the reduction of replication fork rate due to the lesion on leading strand template leads to the production of shorter Okazaki fragments. Therefore it is suggested that with the excess amount of gyrase the replication fork rate is increased and hence the size of the lagging strand replication is increased.

The same phenomena observed with pCNS were also found in the case of pIR246. When gyrase is absent from the reaction, the 3.8 kb counter-clockwise replication product was not detected in pIR246 (lane 7). Whereas a small amount of the 3.8 kb counter-clockwise replication product was detected when gyrase was added (lane 8). This counter-clockwise replication product was increased as the amount of the gyrase increased (lane 9-12). The 0.9 kb

clockwise replication product as well as Okazaki fragments were detected in a similar pattern as detected with pCNS. In addition to that, a leading-strand replication paused product was detected only with pIR246 but not pCNS. The size of the leading-strand replication paused product is different depending on the amount of DNA gyrase. Lane 7-8 showed that when there is less amount of gyrase, the size of this leading-strand replication-paused product is in between 1.9 to 2.7 kb. With the standard amount of DNA gyrase (lane 9) used in the *in vitro oriC* replication assay, two paused products with the size between 1.9 to 2.7 kb and 1.5 to 1.9kb were detected. However, the size of the paused product became in between 1.5 to 1.9 kb when the amount of gyrase is increased (lane 10-12). The concentration of DNA gyrase used in Lane 11 is approximately 148 nM which is nearly to the physiological concentration of DNA gyrase in the cell.

Assuming that palindrome sequence on leading strand template causes the pausing of leading strand synthesis; the pause product would be approximately 1.8 kb (Figure 7A). However the result shows that leading strand synthesis can be paused upstream, within or after the palindrome sequence dependent on the amount of DNA gyrase.

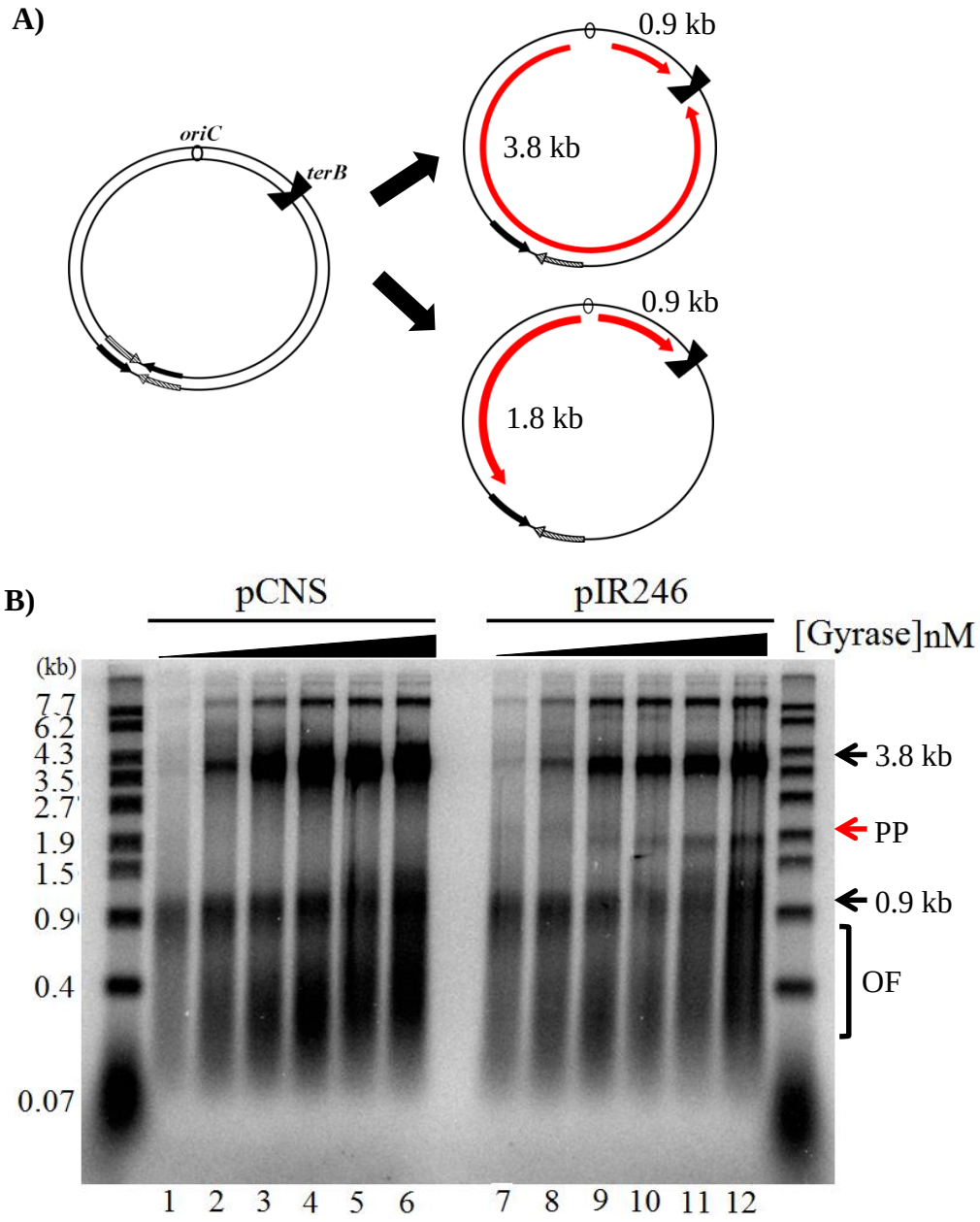


Figure 7 **Effect of gyrase to DNA synthesis within the palindrome sequence.** (A) Schematics diagram indicate the mode of DNA replication in majority of template when it encountering the palindrome sequence. If hairpin structure is formed on leading strand and block leading strand synthesis, the size of the paused product is approximately 1.8 kb. (B) *In vitro oriC* replicaiton assay was carried with titration of gyrase. The concentration of Gyr A/B is 0/0, 37/80, 74/160, 148/320, 296/640, 370/800 nM. Alkaline agarose gel result showed the replication products with the size of 3.8 kb and 0.9 kb which indicate the full length counter-clockwise and clock-wise replication products respectively in both pCNS and pIR246. OF indicating the Okazaki fragments ranged between 100-600 bp was observed in both pCNS and pIR246. PP indicates the paused product detected in pIR246 but not pCNS with different amount of gyrase. The size of the paused product is in between 1.5-1.9 kb.

3.3.2 Southern blot analysis for leading strand replication product

In the alkaline agarose gel (Figure 7B), it was observed that the paused products have different sizes. Thus, this makes the difficulty in concluding the exact location where the leading strand is paused. In order to confirm the size of the paused product Southern blotting was carried out using the leading strand specific probe B to estimate the size of the leading-strand replication-paused product separated on the alkaline agarose gel.

In vitro oriC assay was carried out with pCNS and pIR246 in two conditions: (i) the standard condition of gyrase (1×) containing 74 nM of Gyrase A and 160 nM of Gyrase B and (ii) the nearly to the physiological condition of gyrase (4×) containing 296 nM of Gyrase A and 640 nM of Gyrase B. Replication product obtained was digested with restriction enzyme, *Bgl*I before applied onto the alkaline gel. Southern blotting was then carried out by using Probe B (Figure 8A) that specifically detecting to the leading strand replication product. DNA product with the size of approximately 4.8 kb detected on both pCNS (lane 1 and 3) and pIR246 (lane 2 and 4) represent the template DNA as indicate as T in Figure 8B. A 2.7 kb full length leading strand replication product indicated as FL in Figure 8B was detected with pCNS as well as pIR246 when 1× and 4× gyrase were used. Leading-strand replication-paused products were detected only with pIR246 but not pCNS when 1× and 4× gyrase used. If leading strand replication is paused upstream of the palindrome sequence, the replication paused product is approximately 0.7 kb (Figure 8A). When 1× gyrase was used, two types of leading-strand replication-paused products were detected. The slow migrating product is approximately 0.9 kb while the fast migrating product is approximately 0.7 kb. The 0.9 kb paused product represent the leading-strand replication product from *Bgl*I to the far end of the palindrome sequence. Whereas the 0.7 kb represent the leading-strand replication product from *Bgl*I to the upstream of the palindrome sequence. When 4× gyrase was used only a 0.7 kb leading-strand replication-paused product was detected.

This result shows that leading strand synthesis can be paused before or after passing through

the palindrome sequence when 1× DNA gyrase was used. Moreover, when 4× DNA gyrase was used leading strand synthesis is paused before the palindrome sequence. This suggested that leading strand synthesis pausing within the 246 bp palindrome sequence is dependent on the amount of DNA gyrase.

To more precisely determine the location where leading strand replication is paused within the palindrome sequence, sequencing gel Southern blotting was carried out. This method enables for the identification of pausing site with single nucleotide resolution. Thus, replication products obtained after *in vitro oriC* replication assay in the condition with 1× and 4× gyrase were digested with restriction enzyme, *Msc*I and *Ase*I. The digested products were applied onto the sequencing gel followed by Southern blotting. Probe Lead H located 34 bp upstream of the palindrome sequence can detect the leading-strand replication product as well as the template DNA.

The size of the full length leading strand product is approximately 377 mer when detected with probe Lead H (Figure 9A). As illustrated in Figure 9B, if formation of hairpin structure from the palindrome sequence is solely responsible for arresting the leading strand polymerase, the size of the leading-strand replication-paused product is approximately 65 mer. Whereas if leading strand synthesis stop after passing through the first arm of the palindrome sequence, the size of the paused product is approximately 176 mer (Figure 9B).

Lane 5- 6 are the *Msc*I/*Ase*I digested pIR246 and pCNS template DNA detected by probe Lead H as a marker which the size is approximately 377 mer. The 377 mer full length leading-strand replication product was also found with both pIR246 (lane 1 and 3) and pCNS (lane2 and 4), respectively. However, a 65 mer leading-strand replication paused product was detected only with pIR246 but not pCNS. This result suggests the presence of hairpin structure on leading strand template which consequently caused the pausing of leading strand polymerase at the base of the hairpin stem. Besides that, it was noticed that at 4× gyrase the amount of leading-strand replication-paused product at the base of hairpin stem seems higher than at the

condition of 1× gyrase. Leading-strand replication product with a range of size (105-156 mer) was also detected with pIR246 suggesting that leading strand synthesis is paused within the first arm of the palindrome sequence. The observation of this character of laddering replication paused product suggested that leading strand polymerase slowly overcome the hairpin structure probably through strand displacement activity.

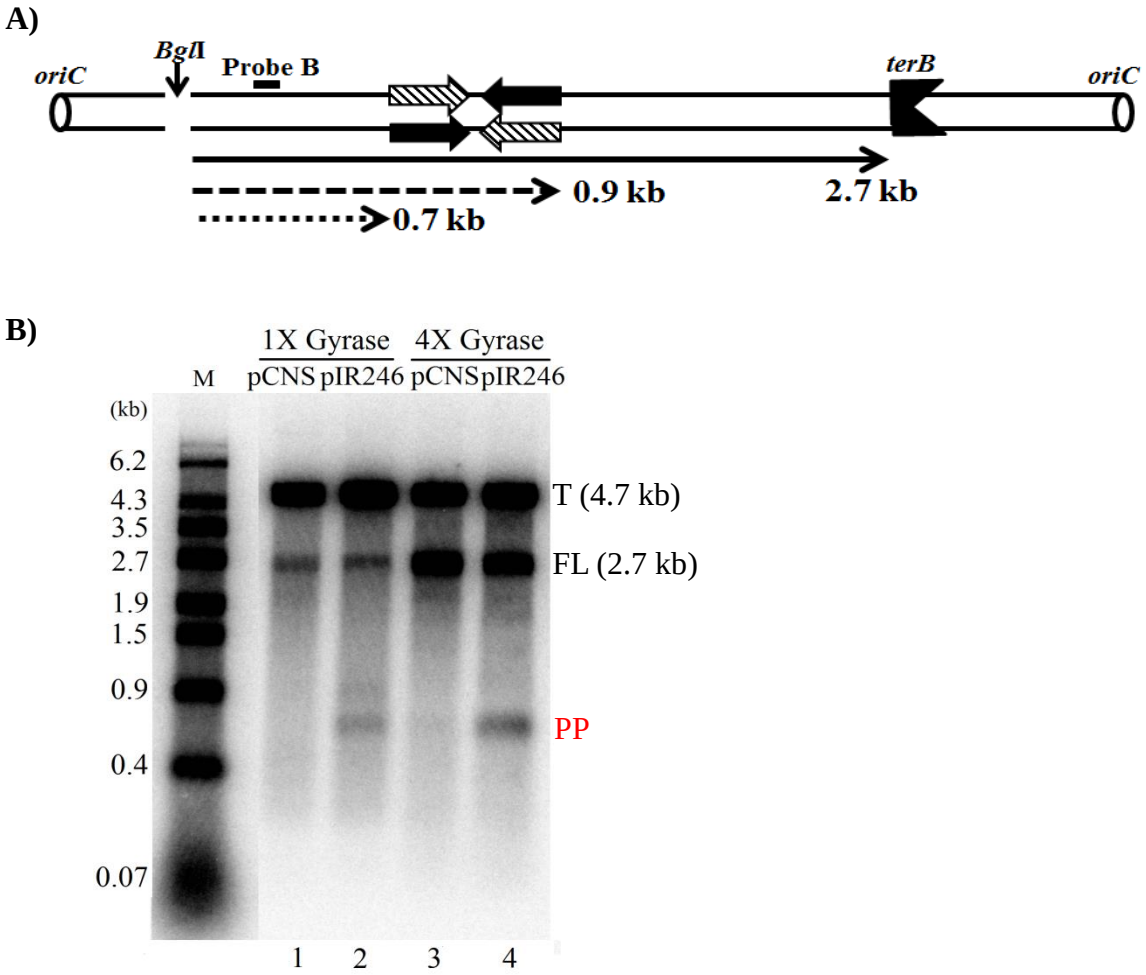


Figure 8 **Alkaline agarose gel Southern blotting analysis of leading strand products when it encounters the palindrome sequence.** Replication products obtained as described in experimental procedure were digested with *Bgl*I and subjected to alkaline agarose gel southern blot analysis. (A) Schematic diagram showed the location of probe used to detect the leading strand replication products. Thin black lines represent the template DNA. Thick black lines with arrowheads indicate the 246 bp palindrome sequence. Black line with arrowhead indicates the full length leading strand product with the size of 2.7 kb. Black dashed line with arrowhead indicates the 0.9 kb leading strand replication paused product that paused after passing through the palindrome sequence. Black dotted line with arrowhead indicate the 0.7 kb leading strand replication paused product that paused upstream of the palindrome sequence (B) Replication products were prepared and subjected to alkaline agarose gel Southern blotting analysis. An autoradiogram of the replication products detected with probes B. Lane 1 and 3 are replication products prepared with the pCNS, while lane 2 and 4 are those prepared with pIR246 with different amount of gyrase respectively. Paused product with the size in between 0.4-0.9 kb was detected in pIR246 when 1× and 4× gyrase were used. T and FL denote the leading-strand template DNA and full length leading strand replication product respectively. PP indicates the paused leading strand replication product.

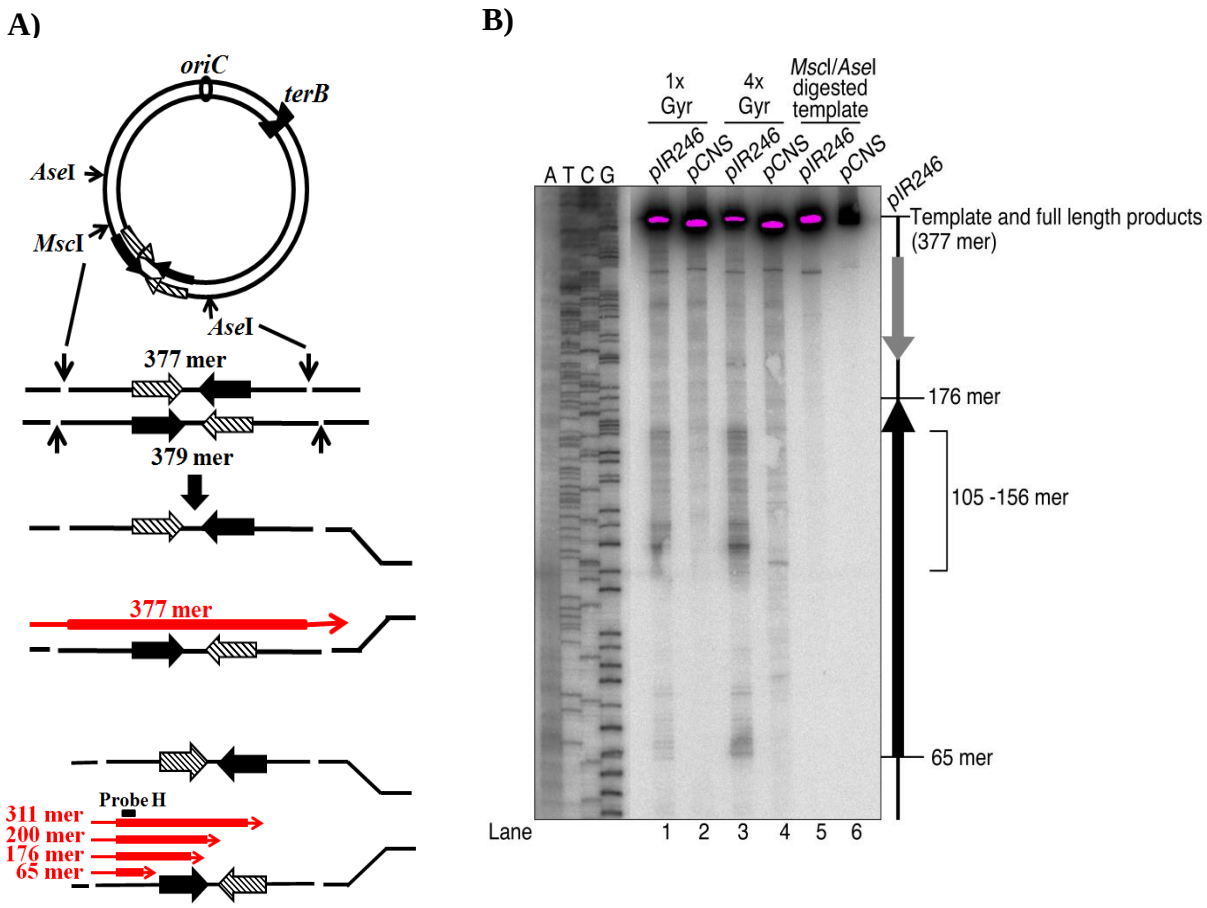


Figure 9 **Leading strand synthesis paused at the base of the hairpin stem.** (A) Schematic diagram showing the predicted leading-strand products detected by probe Lead H after the restriction enzyme treatment. Thin black line represents the template DNA. Thick black line with arrowheads represents the palindrome sequence which located between *Msc*I and *Ase*I. Red thick line with arrowhead denote the leading-strand replication product detect by probe Lead H. (B) Replication products were digested with restriction enzyme, *Msc*I and *Ase*I and applied to the sequencing gel followed by Southern blot analysis as described in the Experimental procedure. Lane 1 and 3 are replication products prepared with the pIR246, while lane 2 and 4 are those prepared with pCNS with different amount of gyrase, respectively. Lane 5 and 6 showed the *Msc*I/*Ase*I digested pIR246 and pCNS template detected by probe Lead H, respectively. The locations of inverted repeats in pIR246 are shown as black and grey arrows in the image. Full length leading strand replication product and template DNA with the size of approximately 377 mer was detected in pIR246 and pCNS. Paused products with the size of approximately 65 mer and 105-156 mer (black bracket) were observed in pIR246 with different amount of gyrase but not pCNS.

3.3.3 Southern blot analysis for lagging strand replication product

It seems that the emerge of cruciform structure ahead of the replication fork caused the pausing of leading strand replication as shown in the alkaline agarose gel (Figure 8B) as well as the sequencing gel Southern blotting results (Figure 9C). Hence, hairpin structure is also suggested to be present on the lagging strand. To examine the effect of palindrome sequence on lagging strand replication, *in vitro oriC* replication assay in the condition of 1× and 4× gyrase was carried out, and the replication product was subjected to a sequence gel Southern blotting using a lagging strand specific Probe Lag A' located 25 bp upstream of the palindrome sequence. Lagging strand replication product as well as the template detected by probe Lag A' is approximately 379 mer (Figure 10A). Lane 5-6 were the *Msc*I/*Ase*I digested pIR246 and pCNS DNA template as a marker detected by Probe Lag A', the size of which is 379 mer (Figure 10B). Additionally, the 379 mer full length lagging strand replication product as well as the template was found with pIR246 (lane 1 and 3) and pCNS (lane 2 and 4), respectively (Figure 10C).

In can be observed that multiple bands were detected by probe lag A' with pCNS (lane 2 and 4). This multiple bands reflect the priming site for the lagging strand synthesis. Lagging strand is synthesized in short fragments known as Okazaki fragments and the synthesis of each Okazaki fragment is initiated with the synthesis of RNA primer (Figure 10B). Therefore the band pattern observed on pCNS may reflect the different priming site for the synthesis of Okazaki fragments.

In the case of pIR246, if lagging strand synthesis is paused upstream of the palindrome sequence, a 68 mer paused product will be detected (Figure 10A). However result showed that there is no particular strong pausing at 68 mer when 1× and 4× of DNA gyrase were used (Figure 10C). Instead, multiple bands were also detected on pIR246. It is observed that lagging-strand product detected at the upstream of the palindrome sequence with the size below 68 mer has similar band pattern in both pCNS and pIR246. The lagging-strand product detected by probe Lag A' on pCNS (Figure 10C) with the size within 68 to 179 mer has similar DNA sequence

with the first arm of the palindrome sequence in pIR246 (Figure 10C). Therefore, it is predicted that similar band pattern of lagging-strand product will be observed on both pIR246 and pCNS. However, the lagging-strand product detected within the size of 68 to 179 mer in pIR246 has different band pattern as compared to pCNS.

From this observation, it suggests that lagging-strand polymerase was not stopped at the palindrome sequence as there is no particular strong lagging-strand replication paused-product before the palindrome was detected. However it is found that the priming location for lagging strand synthesis within the palindrome sequence is different between pIR246 and pCNS. Hence this suggests the possibility of the presence of secondary structure, particularly hairpin structure on lagging strand and it is probably unwind by DNA helicase and hence it delay the priming site for Okazaki fragments synthesis.

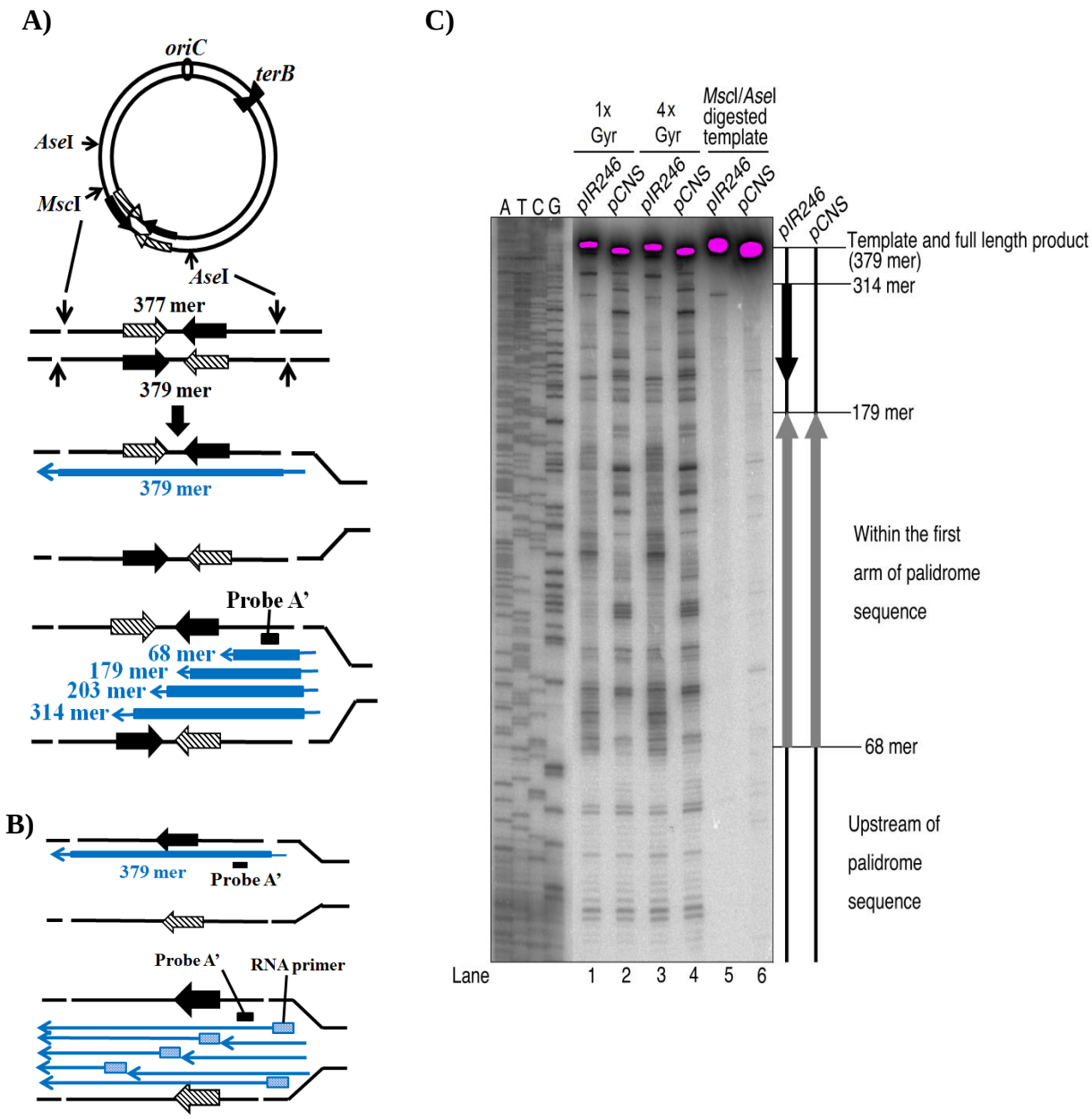


Figure 10 **Lagging strand synthesis was affected within the palindrome sequences.** (A) Schematic diagram showed the location of lagging strand product within the 246 bp palindrome sequence detected by probe Lag A' in pIR246. Thin black line represents the template DNA. Thick black line with arrowheads represents the palindrome sequence which located between *Msc*I and *Ase*I. Blue thick line with arrowhead denote the lagging strand replication product detect by probe Lag A'. (B) Schematic diagram showed the location of lagging strand product detected by probe Lag A' in pCNS and the different priming site during Okazaki fragments synthesis. (C) Replication products obtained were digested with restriction enzyme, *Msc*I and *Ase*I. The digested replication products were applied to the sequencing gel followed by Southern blot analysis. Lane 1 and 3 are replication products prepared with the pIR246, while lane 2 and 4 are those prepared with pCNS with different amount of gyrase, respectively. Lane

5 and 6 showed the *Msc*I/*Ase*I digested pIR246 and pCNS template detected by probe Lag A', respectively. The locations of inverted repeats in pIR246 are shown as black and grey arrows in the image. Full length lagging-strand replication-product and template DNA with the size of approximately 379 mer were detected in pIR246 and pCNS. No significant pausing of lagging strand synthesis within the palindrome sequences in pIR246.

Chapter 4

Discussions

4.1 Triton-extracted pIR246 is cruciform free and can extrude cruciform structure when exposed to heat.

To begin the study, an *oriC-terB* palindrome-containing plasmid, pIR246 and *oriC-terB* palindrome-free control plasmid, pCNS were constructed. In order to examine how the palindrome sequence behaves during DNA replication, it is very important to obtain a DNA preparation of cruciform-free palindrome-containing plasmid. For this purpose, Triton-lysis method was used to prepare the pIR246 plasmid DNA as well as the pCNS. This Triton-lysis method is a non-denaturing method that can avoid the formation of cruciform structure during plasmid preparation.

The palindrome-containing pIR246 and palindrome-free pCNS obtained from Triton-lysis method were DNA mainly in a supercoiled form. In this report, I found that cruciform extrusion is temperature sensitive. The formation of cruciform structure is significantly stimulated at higher temperature in the presence of salt. The heat-treated palindrome-containing pIR246 plasmid exhibited a mobility retardation characterized by the laddering bands observed in agarose gel, suggesting the transition of linear-form palindrome sequence into the cruciform structure. When cruciform is formed from the palindromic sequence, the palindrome-containing supercoiled DNA is known to be relaxed. Therefore, the laddering bands observed in heat-treated pIR246 reflect the various different linking numbers of the plasmid due to the extrusion of cruciform structure.

To further confirm that Triton-extracted pIR246 is cruciform-free, T7 endonuclease which can cleave the cruciform structure was used. The heat-treated pIR246 was readily digested by T7 endonuclease and become linearized. On the other hand, Triton-extracted pIR246 was resistant to the digestion by T7 endonuclease. This observation convincingly

showed that laddering pattern observed with heat-treated pIR246 is the cruciform structure. Most importantly, pIR246 plasmid DNA prepared by Triton-lysis method appeared to contain little or no cruciform structure. Therefore, this result confirms that Triton-extracted pIR246 DNA preparation is cruciform free.

4.2 No significant replication pausing was observed in the standard condition of *in vitro oriC* DNA replication reaction even without SSB.

Having confirmed that Triton-extracted pIR246 is cruciform-free, the *in vitro oriC* DNA replication assay was reconstituted to examine the progression of replication fork when encountering the palindrome sequence. Previous model (Eykelenboom *et al*., 2008) has suggested that hairpin structure formed on lagging strand after the passage of the replication fork. Higuchi *et al.* (2003) showed that the lesion on lagging strand completely blocked the synthesis of Okazaki fragment extending toward the lesion site. Thus, I expected that Okazaki fragment synthesis might be blocked if hairpin structure is formed on lagging strand template after the replication fork passes through the palindrome sequence.

From the result shown in section 3.2.1 (Figure 3 and 4), both leading and lagging strands can be synthesized beyond the 246 bp palindrome sequence. In addition, no significant pausing within the palindrome sequence was detected on lagging strand. From this result, it appeared that hairpin structure could not be formed on the lagging strand template. One of the possibilities for this observation was probably due to the presence of SSB in the *in vitro oriC* replication assay. Single-stranded DNA binding protein (SSB) is known by its function of binding tightly on ssDNA to prevent the formation of any secondary structure. Pausing of DNA polymerase at hairpin structure on ssDNA has been shown to be strongly inhibited by the addition of SSB. Therefore, the possibility of hairpin formation would be reduced if an excess amount of SSB is present.

Thus, I next examined the effect of SSB on the lagging strand synthesis when it

encounters the 246 bp palindrome sequence on the lagging strand template. In order to exclude the possibility of SSB contamination in all the purified replicative proteins, Western blotting has been carried out and showed that Fraction IV Pol III* contained a large amount of SSB contamination (data not shown). Therefore, Calmodulin-binding protein (CBP)-tagged HolD Pol III * with a minute amount of SSB contamination comparing to the Fraction IV Pol III* was used in the SSB titration experiment. *In vitro oriC* replication assay was reconstituted using the CBP-tagged HolD Pol III* with various amount of SSB. It was speculated that hairpin structure may be formed more easily when the amount of SSB is less and vice versa.

From the result shown in 3.2.2 (Figure 5), it was surprising to find that lagging strand can be synthesized beyond the palindrome sequence in the absence of SSB as well as in the excess amount of SSB condition. In addition, there was no detectable lagging-strand replication-paused product in the absence of SSB. Hence this phenomenon cannot be explained by the possibility of SSB contamination for inhibiting the formation of hairpin structure on lagging strand template. The efficiency and the stability of hairpin formation on lagging strand template are apparently lower than the detectable level of our measurements. In the current experiments, with a standard condition of *oriC* reaction it is clear that a hairpin structure is not efficiently formed on the lagging strand template even in the absence of SSB.

Owing to this observation I next used a more sensitive method; Southern blotting that allows the detection of much lower amount of DNA sample. Using the Southern blotting method, it was surprising to observe that a very faint but yet detectable leading-strand replication-paused product on the alkaline gel (3.2.3 Figure 6) although paused-product was not detectable on the lagging strand. No replication-paused product was detected in the palindrome-free pCNS plasmid DNA. This result distinctively suggests that the leading strand synthesis is affected within the palindrome sequence. The pausing of leading strand synthesis within the palindrome sequence suggested a presence of hairpin structure on leading strand template during DNA replication. The low but detectable amount of leading-strand replication-paused

product might be dependent on the stability and efficiency of the hairpin structure on the leading strand template.

In the concerted semi-discontinuous DNA replication process, leading strand is synthesized continuously and thereby it is unlikely for the hairpin structure to be formed on the leading strand template. However, there is a dispute regarding the semi-discontinuous DNA replication model. There are several *in vivo* studies have suggested the discontinuous replication model in *E. coli* (Sternglanz *et al*., 1976; Gottesman *et al*., 1973; Konrad *et al*., 1974; Okazaki *et al*., 1971). If leading strand is synthesized discontinuously, hence there is a possibility for hairpin formation on the leading strand. In addition to that, if DNA helicase is uncouple with leading strand polymerase, DNA helicase may move ahead of the leading strand polymerase and thereby an ssDNA may occur on the leading strand template which may lead to the formation of hairpin structure. Although at this stage these two explanations may not be sufficient to explain the pausing of leading strand synthesis within the palindrome sequence, nonetheless it shouldn't be rule out these two possibilities.

Despite these two explanations, perhaps a more feasible explanation related to the observation of leading-strand replication-paused product within the 246 bp palindrome is the pre-formation of cruciform structure in front of the replication fork. During the progression of replication fork through the double-stranded DNA, strand separation by DNA helicase results in the accumulation of positive superhelical turns ahead of the replication fork (Alexandrov *et al*., 1999; Peter *et al*., 1998). The accumulation of this torsional stress needs to be released in order for further DNA replication. DNA gyrase responsible for introducing the negative supercoiling in DNA (Gellert *el al*., 1976) is shown to participate for releasing the positive supercoils formed ahead of the replication fork in order for the progression of replication fork (Khodursky *et al*., 2000; Levine *et al*., 1998). As a consequent of the releasing torsional stress by DNA gyrase, cruciform structure may be extruded during DNA replication. It is known that cruciform formation is driven by the energy from negative supercoiling.

4.3 Leading strand synthesis is paused at the base of the hairpin stem on the leading strand template.

To examine the possibility of cruciform formation by DNA gyrase, which leads to the pausing of leading strand synthesis, *in vitro oriC* replication assay was reconstituted with various amount of DNA gyrase. As shown in 3.3.1 (Figure 7), a replication-paused product was detected in pIR246 but not pCNS. It is noticed that there is a difference with the pattern of paused product detected in pIR246. When the amount of DNA gyrase is less, the size of the detected paused-product is approximately 2.0 kb suggesting that this paused-product stopped after passing through the palindrome sequence. Whereas a 1.8 kb paused product was detected when higher amount of DNA gyrase was used, and this suggests that leading strand synthesis is stopped upstream of the palindrome sequence.

To further confirm the size of the leading-strand replication-paused product, Southern blotting was carried. In the Southern blotting result (Figure 8), it is clearly shown that there are two leading-strand replication-paused products with the size of approximately 0.7 and 0.9 kb were detected in the standard condition of DNA gyrase. Nonetheless, when excess amounts of DNA gyrase, only a 0.7 kb leading-strand replication-paused product was detected. The fast migrating leading-strand replication-paused product with the size of approximately 0.7 kb suggests that leading strand synthesis is paused upstream of the palindrome sequence. Whereas the slow migrating leading-strand replication-paused product with the size of approximately 0.9 kb suggests that leading strand synthesis stopped after passing through the palindrome sequence. This observation suggests that leading strand synthesis can pause upstream of the palindrome sequence presumably due to the formation of hairpin structure. However, it is surprising that leading strand synthesis is also found stopped after passing through the palindrome sequence when *in vitro oriC* replication assay was carried out with standard amount of DNA gyrase. From this observation, I suggest that pausing of leading strand synthesis is dependent on the amount of DNA gyrase.

The sequencing gel Southern blotting that enables the identification of pausing site with single nucleotide resolution was carried out to precisely estimate the location where leading strand synthesis is paused within the palindrome sequence. Result in 3.3.2 (Figure 9) clearly showed that leading strand synthesis is paused at the base of the hairpin stem when in the standard and excess amount of DNA gyrase used. Whereas, this leading-strand replication-paused product was not detected in the palindrome-free pCNS. Apart from that, a laddering of leading-strand replication-paused product with the size ranging from 105-156 mer was also detected in both conditions of DNA gyrase. The sizes of leading-strand replication-paused product suggest that leading strand can be synthesized within the palindrome sequence.

The above results further confirm the presence of hairpin structure on leading strand template which subsequently caused the pausing of the leading strand synthesis. In addition to that, the observation of laddering pattern of leading-strand replication-paused product suggests a possibility of strand displacement activity of Polymerase III to overcome the hairpin structure formed on the leading strand template. In accordance with this observation, it is suggested that the presence of hairpin structure does not cause an absolute pausing on the leading strand template. Indeed the pausing on the leading strand synthesis is suggested to be temporal due to the strand displacement activity to overcome the hairpin structure.

4.4 The different pausing pattern of leading strand synthesis in accordance with the amount of DNA gyrase

Apart from the observation mentioned above, the pattern of semi-bidirectional replication product with control plasmid, pCNS is slightly different when different amount of DNA gyrase was used (Figure 7). When DNA gyrase is insufficient (absence or less DNA gyrase), the amount of counter-clockwise leading strand replication product is either absence or less. This phenomenon is due to the accumulation of torsional stress in front of the replication fork that inhibits further DNA chain elongation. This was not observed when the amount of DNA gyrase

is sufficient whereby the accumulation of torsional stress is released and hence it enables for further progression of replication fork. In addition to the different occurrence of counter-clockwise leading strand replication product, there is also a difference in the size of lagging strand replication product. When the amount of DNA gyrase is inadequate, the size of lagging strand replication product is slightly smaller comparing when there is adequate amount of DNA gyrase. This observation is probably correlated with the rate of the replication fork. As suggested by Higuchi *et al*. (2003) that a lesion on leading strand template resulted in a reduced rate of replication fork movement which leads to the production of shorter Okazaki fragments. Therefore when DNA gyrase is inadequate, the reduced rate of replication fork can cause the smaller size of Okazaki fragments production. Moreover, the larger Okazaki fragments are produced when the rate of replication fork is increased due to the adequate amount of DNA gyrase.

The results shown in alkaline agrose gel (Figure 7) as well as the Southern blotting (Figure 8 and 9) demonstrated that the pattern of the leading-strand replication-paused product with pIR246 is different dependent on the amount of DNA gyrase. One of the possibilities of this difference may be due to the size of the extruded hairpin structure. The complete extruded hairpin structure may cause the pausing of leading strand synthesis upstream of the palindrome sequence. On the other hand, if the hairpin structure is not completely extruded it may cause the pausing of leading strand synthesis within the palindrome sequence. In fact a perhaps more possible explanation regarding the different pattern of pausing in the standard and excess amount of DNA gyrase condition is correlated to the torsional stress in front of the replication fork. Within the region of palindrome sequence on the leading strand template there is a locally accumulating negative supercoiling due to the hairpin structure. Therefore, leading strand synthesis is first stopped at the base of the hairpin stem. Then by the strand displacement activity of the leading strand polymerase it may slowly overcome the hairpin structure. The strand displacement activity thus unwinds the hairpin structure on the leading strand template

may induce some extra torsional stress in front of the replication fork. During the standard amount of DNA gyrase condition, the insufficient of DNA gyrase activity may cause the pausing of DNA polymerase after passing through the hairpin structure due to the accumulation of torsional stress ahead of the replication fork. This phenomenon is not observed when excess amount of DNA gyrase is used, suggesting that there is sufficient amount of DNA gyrase activity to completely release the torsional stress in front of the replication fork. Therefore, only the paused product stopped at the base of the hairpin stem is detected when there is excess amount of DNA gyrase.

4.5 Lagging strand synthesis is suggested to be affected within the palindrome sequence

The observation of leading strand replication pausing is probably due to the presence of cruciform structure ahead of the replication fork. Hence, hairpin structure is also expected to be present on the lagging strand template. The sequencing gel Southern blotting was carried out to examine the effect of palindrome sequence on DNA synthesis on the lagging strand template.

The result shown in Figure 10 demonstrated that there were multiple bands detected with pCNS. These bands may reflect the priming site for the DNA synthesis on lagging strand template. Lagging strand is synthesized in short fragments known as Okazaki fragments. The synthesis of each Okazaki fragments is initiated by the synthesis of RNA primer which is catalyzed by DNA primase at specific recognition sites. Hence the observation of this multiple bands may represent the priming site for lagging strand synthesis.

From the result (Figure 10), multiple bands were also detected with pIR246 but the band pattern was different from that observed with pCNS. The DNA sequence detected by probe Lag A' on lagging strand replication product with pCNS has the same DNA sequence with the first arm of the palindrome sequence on the lagging-strand replication product with pIR246. Hence it is assumed that same band pattern of lagging strand replication product will be observed in both plasmids. Instead, a different priming site for the synthesis of Okazaki fragments is

observed with pIR246 comparing to pCNS. This result is surprising and hence brings into the suggestion that possibly a hairpin structure is also presence on the lagging strand template and subsequently causes a delay of priming for lagging strand synthesis in pIR246.

In fact, the occurrence of hairpin structure on lagging strand template does not show significant pausing at the base of the hairpin stem. One of the possibilities of this observation is probably due to the involvement of DNA helicase. In the replisome, DNA helicase is located on the lagging strand template. Thus, DNA helicase will first encounter the hairpin structure before the lagging strand DNA polymerase. Consequently, DNA helicase would be involved in unwinding the hairpin structure as it is to unwind the duplex DNA. After unwinding the first arm of inverted repeats until the spacer, the second arm of the inverted repeats will be covered by SSB. Therefore this observation also suggests that lagging strand synthesis is affected within the palindrome sequence probably due to the hairpin structure. However, this does not cause any pausing possibly due to the effect of DNA helicase located on the lagging strand template in unwinding the hairpin structure.

4.6 Conclusions

From all the observations obtained in this study, a new model explaining how DNA replication machinery behaves when it encounters the palindrome sequence is proposed (Figure 11). Mirkin's group found that the progression of replication fork is stalled when it encounters the palindrome sequence. However this phenomenon cannot be explained by the model proposed by Leach's group, in which the presence of palindrome sequence does not cause the pausing of replication fork.

My results indicated that the progression of replication fork is affected by the palindrome sequence. Both leading- and lagging-strand synthesis were affected within the 246 bp palindrome sequence. The leading strand synthesis was paused before the palindrome sequence (Figure 7 and 8) and further study revealed leading-strand synthesis to be paused at the base of the hairpin stem (Figure 9). These suggest that hairpin structure was probably formed on the leading strand template and caused the pausing of leading-strand synthesis. Nonetheless, the pausing of leading-strand synthesis within the palindrome sequence seems not to be absolute but likely temporal. This is because a laddering pattern was observed with leading-strand replication-paused product analyzed by the sequencing-gel Southern-blotting, which is probably due to the strand displacement activity of DNA polymerase to overcome the hairpin structure. In addition, full length of leading-strand replication-product was also.

On the other hand, lagging-strand synthesis was also affected within the 246 bp palindrome sequence. Although no significant pausing of lagging-strand replication-product at the base of the hairpin stem, it was found the distribution of priming sites for the Okazaki fragments synthesis was different in replication of pIR246 and pCNS (Figure 10). The difference in priming sites brings into the suggestion that hairpin structure is also present on the lagging strand template and DNA helicase may unwind the hairpin structure. As a consequence, the action of DNA helicase in hairpin structure would change the timing of the priming of Okazaki fragments synthesis on pIR246.

Previous study has suggested the cleavage of hairpin structure by SbcCD is on the lagging strand. However, there is no direct evidence for this notion. Instead of hairpin structure formation on the lagging strand template, I newly propose that cruciform structure is formed from the palindrome sequence in front of the replication fork (Diagram 11).

During DNA replication, the cruciform structure could be formed ahead of the replication fork (Figure 11B). Therefore, it is possible both leading and lagging strand synthesis are interrupted within the 246 bp palindrome sequence (Figure 11C and D). However, the cruciform structure formed from the palindrome sequence would not necessarily cause an absolute block of DNA replication fork, instead it could slow down the DNA replication fork. As the dimeric Polymerase III approaching the cruciform structure, leading strand Polymerase III would be paused by the hairpin structure formed on the leading strand template (Figure 11C). Subsequently by the strand displacement activity of leading-strand Polymerase, the hairpin structure could be replicated (Figure 11D). On the other hand, hairpin structure formed on the lagging strand template is probably unwound by the DNA helicase and thereby lagging strand Polymerase III could continue to synthesize the lagging strand (Figure 11C and D). As a result, both Polymerase IIIs would be able to continue the chain elongation after passing through the palindrome sequence (Figure 11E).

I suggest that cruciform structure is formed during DNA replication and affect both the leading and lagging strand synthesis. The possibility of cruciform-extruded template used in the *in vitro oriC* replication assay is precluded as it has already been confirmed that the palindrome-containing pIR246 plasmid DNA is cruciform free. Instead a cruciform structure seems to be formed during DNA replication presumably due to the action of DNA gyrase. During DNA replication, strand separation of the DNA duplex for chain elongation will generates some torsional stress (positive supercoiling) in front of the replication fork. This torsional stress needs to be released in order for further chain elongation in front of the replication fork. DNA gyrase is the enzyme involved in introduction of negative supercoiling.

Hence, DNA gyrase functions in releasing the torsional stress accumulate in front of the replication fork. Under this circumstance the reaction of DNA gyrase to introduce the negative supercoiling may facilitate the formation of cruciform structure as the formation of cruciform structure is favored in the presence of negative supercoiling. (Figure 11A and B)

It has been known that negative superhelical density is an essential factor for stabilizing most of the non-B structures, including cruciform structure. Napierala *et al*. (2005) has suggested that increased negative superhelical density *in vivo* enhances the genetic instability of triplet repeat sequences due to the conformation of non-B DNA structure-forming repeat. Thus, the involvement of DNA gyrase in releasing the torsional stress in front of the replication fork by introducing the negative supercoiling would cause the extrusion of cruciform structure. As a consequence of this phenomenon, hairpin structure should be present not only on lagging strand template but also on the leading strand template. However, further experiments need to be carried out to confirm the involvement of DNA gyrase in inducing the cruciform structure from the palindrome sequence.

Results obtained in this study are now providing a new insight how palindrome sequence affects DNA replication. I found that both leading- and lagging-strand synthesis are affected within the 246 bp palindrome sequence probably due to the formation of cruciform structure dependent on DNA replication. However the DNA replication machinery itself can overcome the cruciform structure probably involving the DNA helicase to unwind the hairpin structure on lagging-strand template and the strand displacement activity of Polymerase III to overcome the hairpin structure on leading-strand template. Previous studies have suggested that palindrome sequence is not able to be maintained in the presence of SbcCD due to the cleavage of hairpin structure by SbcCD protein. However this inviability of palindrome sequence is overcome when in the absence of SbcCD nuclease. The *in vivo* study by Cromie *et al*. (2000) suggested that when in the absence of SbcCD, RecQ helicase is involved in the palindrome viability in a *sbcC* mutant. The involvement of RecQ with the helicase activity is suggested to unwind the

secondary structure and allow the replication to pass through the palindrome sequence. Thus, it is suggested that in the *sbcC* mutant, a helicase-dependent pathway exists to allow replication bypass of secondary structures.

As a conclusion, I propose that genome instability stimulated by the palindrome sequence is dependent on DNA replication. This is probably due to the formation of cruciform structure by the action of DNA gyrase during DNA replication.

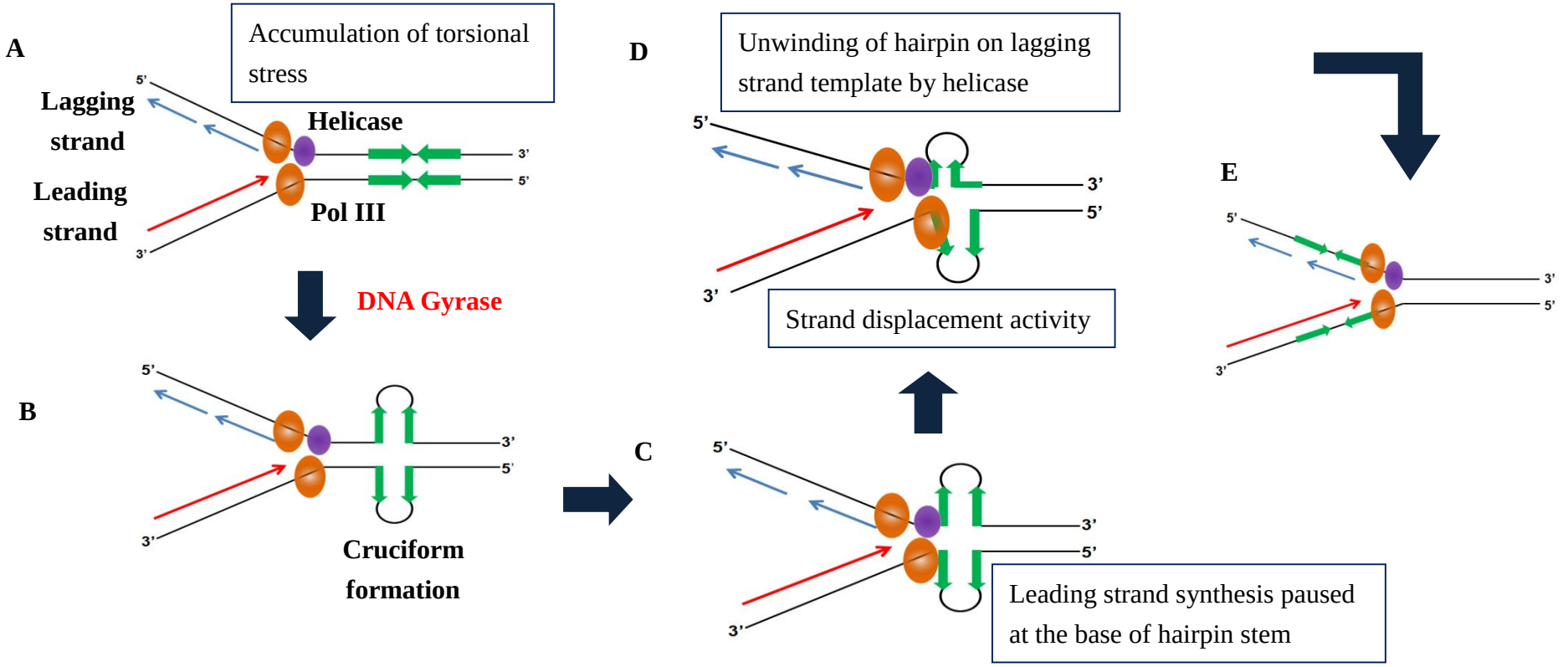


Figure 11 **Model mechanism of DNA replication machinery when there is presence of palindrome sequence.** (A) During DNA replication, torsional stress is accumulated ahead of replication forks due to the strand separation. (B) DNA gyrase involved in releasing the torsional stress by introducing negative supercoiling causes the extrusion of cruciform structure ahead of replication fork. (C) This cruciform structure affects both the leading and lagging strand synthesis. (D) Strand displacement activity involved to overcome the hairpin on the leading strand template. On the lagging strand synthesis, DNA helicase may involve to unwind the hairpin structure on the lagging strand template. (E) Consequently, both leading and lagging strand polymerases are now able to continue the chain elongation after passing through the palindrome sequence. Blue and red lines with arrowhead represent the leading and lagging strand synthesis, respectively. Thin black line represents the DNA template. The thick green line with arrowhead indicates the palindrome sequence. Orange and purple circles represent the DNA polymerase and DNA helicase.

Acknowledgements

First and foremost, I would like to express my heartfelt gratitude to my supervisor, Prof. Hisaji Maki, for giving me an opportunity to study in his lab. Throughout these four years, he has provided constant encouragement, patient, helps and guidance to me. In addition I would also like to thank him for providing me a very good environment and ample facilities in order to carry out my study. Without a suitable environment and the facilities, this project would not be successful.

I would also like to take this opportunity to express my appreciation to Dr. Asako Furukohri who is the assistant prof. in Maki's laboratory. Throughout my study, she has provided me with endless of technical as well as scientific advices. I am indebted to her for all the time and effort she has dedicated to my work. Apart from that, she has been helping me a lot in my daily life especially when I was sicked.

Further appreciation goes to Dr. Satoko Maki and Assoc. Prof. Akiyama who has been of great help to me throughout my study. Discussion session that we had had during progress report is very useful to me. Special thanks are dedicated to lab members in Maki's laboratory especially Ikeda-san and Murao-san for their helps when I first arrived in Japan. Not to forget my neighbor who sit next to me, Hang-san who has been a very good friend to me. We shared not only laughter but also we encouraged each other when we are feeling down. Her good and optimistic attitude has also inspired me and I should learn from her.

I wish to extend my gratitude to all my friends from Malaysia who are currently studying in NAIST as they had helped me to walk through every difficulty that I faced. Their support, advices and encouragement have helped me to make this study a success. In addition, I would also like to offer my special thanks to my friends from China, Jacky Yang, Lay-san and Yu-san for their helps with my daily life in Japan.

Last but not least, I am indebted to my family for their encouragement, support and most importantly their understanding. My special thanks are extended to my husband, Mr. Teoh Seddon for his companion ship. Most importantly is his support to me for coming to Japan to pursue my PhD study. He has been a very good listener to me when I am feeling down and also a very good helper to ease my burden by helping me with the housework. Their care and love to me have given me the strength to complete my study.

References

Anusubel, F. M., Brent, R., Kingston, R. E., Moore, D. D. and Seidman, J. G. (1994) *Current protocols in Molecular Biology.* John Wiley and Sons, New York.

Alberts, B. M., Barry, J., Bedinger, P., Formosa, T., Jongeneel, C,V. and Kreuzer, K. N. (1983) Studies on DNA replication in the bacteriophage T4 *in vitro* system. Cold Spring Harb. Symp. Quant. Biol. 47, 655–668.

Alexandrov, A. I., Cozzarelli, N. R., Holmes, V. F., Khodursky, A. B., Peter, B. J., Postow, L., Rybenkov, V. and Vologodskii, A. V.(1999). Mechanisms of separation of the complementary strands of DNA during replication. Genetica 106, 131–140.

Bird, R. E., Louarn, J., Martuscelli, J., and Caro, L. (1972) Origin and Sequence of Chromosome Replication in *Escherichia coli.* J. Mol. Biol. 70**,** 549–566.

Bramhill, D. and Kornberg, A. (1988) A model for initiation at origins of DNA replication. Cell 23, 915-8.

Broach, J. R., Li, Y.Y., Feldman, J., Jayaram, M., Abraham, J., Nasmyth, K. A. and Hicks, J.B. (1984) Localization and sequence analysis of yeast origins of DNA replication. Cold Spring Harbor Symp. Quant. Biol. 47, 1165-1173.

Callejo, M., Alvarez, D., Price, G. B. and Zannis-Hadjopoulos, M. (2002) The 14-3-3 protein homologues from *Saccharomyces cerevisiae*, Bmh1p and Bmh2p, have cruciform DNAbinding activity and associate *in vivo* with ARS307. J. Biol. Chem. 277, 38416–38423.

Chalker, A. F., Leach, D. R. and Lloyd, R. G. (1988) *Escherichia coli sbcC* mutants permit stable propagation of DNA replicons containing a long palindrome. Gene 71, 201-205.

Challberg, M. D. and Englund, P. T. (1979) The effect of template secondary structure on Vaccinia DNA polymerase*. J. Biol. Chem. 254, 16 7820-7826.

Collins, J. (1981) Instability of palindromic DNA in *Escherichia coli*. Cold Spring Harbor Symp. Quant. Biol. 45, 409-416.

Collins, J., Volckaert, G. and Nevers, P. (1982) Precise and nearly-precise excision of the symmetrical inverted repeats of Tn5; common features of *recA*-independent deletion events in *Escherichia coli*. Gene 19, 139-146.

Connelly, J. C., Kirkham, L. A. and Leach, D. R. F. (1998) The SbcCD nuclease of *Escherichia coli* a structural maintenance of chromosomes (SMC) family protein that cleaves hairpin DNA. Proc. Natl. Acad. Sci. USA 95, 7969-7974.

Courey, A.J. and Wang, J.C. (1983) Cruciform formation in a negatively supercoiled DNA may kinetically forbidden under physiological conditions. Cell 33, 817-829.

Cromie, G. A., Millar, C. B., Schmidt, K. H. and Leach, D. R. F. (2000) Palindromes as substrates for multiple pathways of recombination in *Escherichia coli*. Genetics 154, 513-522.

Dai, X., Greizerstein, M., Nadas-Chinni, K. and Rothman-Denes, L. B. (1997) Supercoil-induced extrusion of a regulatory DNA hairpin. Proc. Natl. Acad. Sci. USA 94, 2174-2179.

Dai, X. and Rothman-Denes, L. B. (1998) Sequence and DNA structural determinants of N4 virion RNA polymerase-promoter recognition. Genes Dev. 12, 2782-2790.

De´clais, A., Fogg, J. M., Freman, A. D. J., Coste, F., Hadden, J. M., Philips, S. E.V. and Lilley, D. M. J. (2003) The complex between a four-way junction DNA and T7 endonuclease I. EMBO J. 22,1398–1409.

de Massey, B., Weisberg, R. A. and Studier, F. W. (1987) Gene 3 endonuclease of bacteriophage T7 resolves conformationally branched structures in double-stranded DNA. J. Mol. Biol. 193, 359-376.

Dickie, P., McFadden, G. and Morgan, A. R. (1987) The site-specific cleavage of synthetic Holliday junction analogs and related branched structures by bacteriophage T7 endonuclease I. J. Biol. Chem. 262, 14826-14836.

Eykelenboom, J. K., Blackwood, J. K., Okely, E. and Leach, D. R. F. (2008) SbcCD causes a double-strand break at a DNA palindrome in the *Escherichia coli* chromosome. Mol. Cell 29, 644-651.

Frank-Kamenetskii, M.D. and Mirkin, S.M. (1995) Triplex DNA structures. Annu. Rev. Bio. Chem. 64, 65-69.

Freund, A. M., Bichara, M., and Fuchs, R. P. (1989) Z-DNA-forming sequences are spontaneous deletion hot spots. Proc. Natl. Acad. Sci. USA 86, 7465–7469.

Fuller, R. S., Kaguni, J. M. and Kornberg, A. (1981) Enzymatic replication of the origin of the *Escherichia coli* chromosome. Proc. Natl. Acad. Sci. USA. 78**,** 7370–7374.

Butland, G., Perigrin-Alvarez, J. M., Li, J., Yang, W. H., Yang, X. C., Canadien, V., Starostine, A., Richards, D., Beattie, B., Krogan, N., Davey, M., Parkison, J., Greenblatt, J. and Emili, A. (2005) Interaction network containing conserved and essential protein complexes in *Escherichia coli*. Nature 433, 531-537.

Gellert, M., Mizuuchi, K., O'Dea, M. H. and Nash, H. (1976) DNA gyrase: an enzyme that introduces superhelical turns into DNA. Proc. Natl. Acad. Sci. USA 73, 3872-3876.

Gibson, F. P ., Leach, D. R. F. and Lloyd, R. G. (1992) Identification of *sbcD* mutations as co-suppressers of *recBC* that allow propagation of DNA palindromes in *Escherichia coli* K-12. J. Bacteriol. 174, 1222-1228.

Glucksmann, M. A., Markiewicz, P., Malone, C. and Rothman-Denes, L. B. (1992) Specific sequences and a hairpin structure in the template strand are required for N4 virion RNA polymerase promoter recognition. Cell 70, 491-500.

Gold, L. (1988) Postranscriptional regulatory mechanisms in Escherichia coli. Annu. Rev. Biochem. 57, 199-233.

Gottesman, M. M., Hicks, M. L. and Gellert, M. (1973) Genetics and function of DNA ligase in *Escherichia coli*. J. Mol. Biol. 77, 531–547.

Hagan, C.E. and Warren, G.J. (1982) Lethality of palindromic DNA and its use in the selection of recombinant plasmids. Gene 19, 147-151.

Hagan, C. E. and Warren, G. J. (1983) Viability of palindromic DNA is restored by deletions occurring at low but variable frequency in plasmids of *E. coli*. Gene 24, 317-326.

Hacker, K. J. and Albert, B. M. (1994) The rapid dissociation of the T4 DNA Polymerase holoenzyme when stopped by a DNA hairpin helix. J. Biol. Chem. 269, 24221-24228.

Haniford, D. B. and Pulleyblank, D. E. (1985) Transition of a cloned d(AT)n-d(AT)n tract to a cruciform *in vivo*. Nucleic Acids Res. 13, 4343-4363.

Haynes, L. L. and Rothman-Denes, L. B. (1985) N4 virion RNA polymerase sites of transcription initiation. Cell 41, 597- 605.

Hiasa, H. and Marians, K. J. (1994) Primase couples leading- and lagging-strand DNA synthesis from *oriC*. J. Biol. Chem. 269, 6058–6063.

Hiasa, H. and Marians, K. J. (1999) Initiation of bidirectional replication at the chromosomal origin is directed by the interaction between helicase and primase. J. Biol. Chem. 274, 27244–27248.

Hiasa, H., Tanaka, K., Sakai, H., Yoshida, K., Honda, Y., Komano, T. and Godson, N. (1989) Distinct functional contributions of three potential secondary structures in the phage G4 origin of complementary DNA strand synthesis. Gene 84, 17-22.

Hidaka, M., Akiyama, M. and Horiuchi, T. (1988) A consensus sequence of three DNA replication terminus sites on the *E. coli* chromosome is highly homologous to the *terR* sites of the R6K plasmid. Cell 55, 467–4

Higuchi, K., Katayama, T., Iwai, S., Hidaka, M., Horiuchi, T. and Maki, H. (2003) Fate of DNA replication fork encountering a single DNA lesion during *oriC* plasmid DNA replication *in vitro.* Genes to cell **8,** 437–449.

Huang, C.C., Hearst, J. E. and Alberts, B. M. (1981) Two types of replication proteins increase the rate at which T4 DNA polymerase traverses the helical regions in a single-stranded DNA template. J. Biol. Chem. 256, 4087-4094.

Kaguni, J.M. and Kornberg, A. (1984) Replication initiated at the origin (*oriC*) of the *E. coli* chromosome reconstituted with purified enzymes. Cell 138, 183-190.

Kaguni, J.M., Fuller, R.S. and Kornberg, A. (1982) Enzymatic replication of *E. coli* chromosomal origin is bidirectional. Nature 296, 623-627.

Kaguni, L. S. and Clayton, D.A. (1982) Template-directed pausing in *in vitro* DNA synthesis by DNA polymerase a from *Drosophila melanogaster* embryos. Proc .Nat. Acad. Sci. USA 79, 983-987.

Khodursky, A. B., Peter, B. J., Schmidt, M. B., DeRisi, J., Botstein, D., Brown, P. O. and Cozzarelli, N. R. (2000) Analysis of topoisomerase function in bacterial replication fork movement: Use of DNA microarrays. Proc. Natl. Acad. Sci. USA 97, 9419–9424.

Konrad, E. B., Modrich, P. and Lehman, I. R. (1974) DNA synthesis in strains of *Escherichia coli* K12 with temperature-sensitive DNA ligase and DNA polymerase I. J. Mol. Biol. 90, 115–126.

Kulkarni, S. (1990) Genetic studies on DNA palindromes in phage lambda and *E. coli*. Ph.D. Thesis, University of Oregon, pp. 38–62.

Kurahashi, H., Inagaki, H., Yamada, K., Ohye, T., Taniguchi, M., Beverly S. Emanuel and Toda, T. (2004) Cruciform DNA structure underlies the etiology for palindrome-mediated human chromosomal translocations. J. Biol. Chem. 279, 35377–35383.

Krasilnikov, A. S., Panyutin, I. G., Samadashwily, G. M., Cox, R., Lazurkin, Y. S. and Mirkin, S. M. (1997) Mechanisms of triplex-caused polymerization arrest. Nucleic Acids Res. 25, 1339–1346.

LaDuca, R. J., Fay, P. J., Chuang, C., McHenry, C. S. and Bambara, R. A. (1983) Site-specific pausing of deoxyribonucleic Acid synthesis catalyzed by four forms of *Escherichia coli* DNA polymerase III. Biochemistry 22, 5177-5188.

Lambert, P.F., Waring, D. A., Wells, R. D. and Reznikoff, W.S. (1986) DNA Requirements at the Bacteriophage G4 Origin of Complementary-Strand DNA Synthesis. J. Virol. 58, 450-458.

Leach, D. R. F. (1994) Long DNA palindromes, cruciform structures, genetic instability and secondary structure repair. BioAssays 16, 893-900.

Leach, D. R. F. and Lindsey, J. (1986) In vivo loss of supercoiled DNA carrying a palindromic sequence. Mol. Gen. Genet. 204, 322-327.

Leach, D. F. R., Okely, E. A. and Pinder, D. J. (1997) Repair by recombination of DNA containing a palindromic sequence. Mol. Microbiol. 26, 597-606.

Leach, D. R. F. and Stahl, F. W. (1983) Viability of λ phages carrying a perfect palindrome in the absence of recombination nucleases. Nature 305, 448-451.

Levine, C., Hiasa, H. and Marians, K. J. (1998) DNA gyrase and topoisomerase IV: Biochemical activities, physiological roles during chromosome replication, and drug sensitivities. Biochim. Biohys. Acta. 1400, 29–43.

Li, X. and Marians, K. J. (2000) Two distinct triggers for cycling of the lagging strand polymerase at the replication fork. J. Biol.Chem. 275, 34757–34765.

Lilley, D. M. J. (1980) The inverted repeat as a recognizable structural feature in supercoiled DNA molecules. Proc. Nat. Acad. Sci. USA 77, 6468-6472.

Lilley, D. M. J. (1981) *In vivo* consequences of plasmid topology. Nature 292, 380-382.

Lilley, D. M. and Kemper, B. (1984) Cruciform-resolvase interactions in supercoiled DNA. Cell 36**,** 413–422.

Lindsey, J. C. and Leach, D. R. F. (1989) Slow replication of palindrome-containing DNA. J. Mol. Biol. 206, 779-782

Lloyd, R. G. and Buckman, C. (1985) Identification and genetic analysis of *sbcC* mutations in commonly used *recBC sbcB* strains of *Escherichia coli* K-12. J. Bacteriol. 164, 836–844.

Majumdar, A. and Patel, D. J. (2002) Identifying hydrogen bond alignments in multistrand DNA architectures by NMR. Acc. Chem. Res. 35,1-11.

Maki, S. and Kornberg, A. (1988) DNA polymerase III holoenzyme of *Escherichia coli*. III. Distinctive processive polymerases reconstituted from purified subunits. J. Biol. Chem. 263, 6561–6569.

Marians, K. J. (1992) Prokaryotic DNA replication. Annu. Rev.Biochem. 61, 673–719.

McHenry, C. S. (1988) DNA polymerase III holoenzyme of *Escherichia coli*. Annu. Rev. Biochem. 57, 519–550.

McMacken, R., Silver, L., and Georgopoulos, C. (1987) Section C. Class III Reactions: Formation and processing of polymers. 39. DNA Replication. In Neidhardt, F.C., Ingraham, J. L., Low, K. B., Magasanik, B., Schaechter, M., and Umbarger, H. E. (eds) *Escherichia Coli* and *Salmonella Typhimurium*. Cellular and Molecular Biology. Vol I. American Society for Microbiology.

Mirkin, S. M. (2007) Expandable DNA repeats and human disease. Nature 447, 932-940.

Mizuuchi, K., Kemper, B., Hays, J., Weisberg, R.A. (1982) T4 endonuclease VII cleaves Holliday structures. Cell **29,** 357–365.

Mizuuchi, K., Mizuuchi, M. and Gellert, M. (1982) Cruciform structures in palindromic DNA are favoured by DNA supercoiling. J. Mol. Biol. 156, 229-243.

Napierala, M., Bacolla, A., and Wells, R.D. (2005) Increased negative superhelical density *in vivo* enhances the genetic instability of triplet repeat sequences. J. Biol. Chem. 280, 37366-37376.

Oka, A., Sugimoto, K., Takanami, M., and Hirota, Y. (1980) Replication origin of the *Escherichia coli* K-12 chromosome: the size and structure of the minimum DNA segment carrying the information for autonomous replication. Mol. Gen. Genet. 178, 9-20.

Okazaki, R., Arisawa, M. and Sugino, A. (1971) Slow joining of newly replicated DNA chains in DNA polymerase I-deficient *Escherichia coli* mutants. Proc. Natl. Acad. Sci. USA 68, 2954–2957.

Oostra, B. A. (1998) Trinucleotide diseases and instability Heildelberg: Springer-Verlag.

Panayotatos, N. and Wells, R. D. (1981) Cruciform structures in supercoiled DNA. Nature 289, 466−470.

Pearson, C.E., Zorbas, H., Price, G. B. and Zannis-Hadjopoulos, M. (1996) Inverted repeats, stem-loops and cruciform: significance for initiation of DNA replication. J. Cell. Biochem. 63, 1-22.

Perricaudet, M., Fritsch, A., Pettersson, U., Philipson, L. and Tiollas, P. (1977) Excision and recombination of adenovirus DNA fragments in *Escherichia coli*. Science 196, 208-210.

Peter, B. J., Ullsperger, C., Hiasa, H., Marians, K. J. and Cozzarelli, N. R. (1998). The structure of supercoiled intermediates in DNA replication. Cell 94, 819–827.

Pinder, D. J., Blake, C. E., Lindsey, J. C. and Leach, D. F. R. (1998) Replication strand preference for deletions associated with DNA palindromes. Mol. Microbiol. 28, 719-727.

Prescott, D. M. and Kuempel, P. L. (1972) Bidirectional Replication of the chromosome in *Escherichia coli.* Proc. Natl. Acad. Sci. USA 69**,** 2842–2845.

Rao, B. S. (1996) Regulation of DNA replication by homopurine/ homopyrimidine sequences Mol. Cell. Biochem. 156, 163–168.

Reckmann, B., Grosse, F., Urbanke, C., Frank, R., BlÖcker, H. and Krauss, G. (1985) Analysis of secondary structures in M13mp8 (+) single-stranded DNA by the pausing of DNA polymerase α. Eur. J. Biochem. 152, 633-643.

Rich, A., Nordheim, A. and Wang, H. J. (1984) The chemistry and biology of left-handed Z-DNA. Annu. Rev. Biochem. 53, 377-378.

Rosche, W. A., Trinh, T. Q. and Sinden, R. R. (1995) Differential DNA secondary structure-mediated deletion mutation on the leading and lagging strands. J. Bacteriol. 177, 4385-4391.

Sherman, L. A. and Gefter, M. L. (1976) Studies on the mechanism of enzymatic DNA elongation by *Escherichia coli* DNA polymerase II. J. Mol. Biol. **103,** 61–76.

Shurvinton, C. E., Stahl, M. M. and Stahl, F. W. (1987) Large palindromes in the λ phage genome are preserved in a *rec*$^+$ host by inhibiting λ DNA replication. Proc. Natl. Acad. Sci. USA 84, 1624-1628.

Sinden, R. R. (1994) *DNA structure and function*. Academic Press, San Diego. Academic Press, Inc., San Diego,CA.

Sinden, R. R., Pytlos-Sinden, M. J., and Potaman, V. N. (2007) Slipped strand DNA structures. Front. Biosci. 12, 4788-4799.

Sinder, R. R., Zheng, G. X., Brankamp, R. G. and Allen, K. N. (1991) On the deletion of inverted repeated DNA in *Escherichia coli:* effects of length, thermal stability, and cruciform formation *in vivo*. Genetics 129, 991-1005.

Smith, G.R. (2008) Meeting DNA palindrome head-to-head. Genes Dev. 22, 2612-2620.

Sternglanz, R., Wang, H. F. and Donegan, J. J. (1976). Evidence that both growing DNA chains at a replication fork are synthesized discontinuously. Biochem. 15, 1838–1843.

Stinchcomb, D. T., Mann, C., Selker, E. and Davis, R.W. (1981) DNA sequences that allow the replication and segregation of yeast chromosomes. ICN-UCLA Symp. Mol. Cell. Biol. 22, 473-488.

Trinh, T. Q. and Sinden, R. P. (1991) Preferential DNA secondary structure mutagenesis in the lagging strand of replication in *E. coli*. Nature 352, 544-547.

Villani, G., Fay, P. J., Bambara, R. A. and Lehman, I. R. (1981) Elongation of RNA-primed DNA templates by DNA polymerase alpha from *Drosophila melanogaster* embryos. J. Biol. Chem. 256, 8202-8209.

Voineagu, I., Narayanan, V., Lobachev, K. S. and Mirkin, S. M. (2008) Replication stalling at unstable inverted repeats: interplay between DNA hairpins and fork stabilizing proteins. PNAS 105, 9936-9941.

Wang, A. H., Quigley, G. J., Kolpak, F. J., Crawford, J. L., van Boom, J. H., van der Marel, G. and Rich, A. (1979) Molecular structure of a left-handed double helical DNA fragment at atomic resolution. Nature 13, 680-686.

Wang, G. and Vasquez, K. M. (2004) Naturally occurring H-DNA-forming sequences are mutagenic in mammalian cells. Proc. Natl. Acad. Sci. USA 101, 13448–13453.

Weaver, D.T. and DePamphilis, M. L. (1982) Specific Sequences in native DNA that arrest synthesis by DNA polymerase α*. J. Biol. Chem. 257, 2075-2086.

Weitzmann, M. N., Woodford, K. J. and Usdin, K. (1996) The development and use of a DNA polymerase arrest assay for the evaluation of parameters affecting intrastrand tetraplex formation. J. Biol. Chem. 271, 20958–20964.

Wells, R. D. and Warren, S. T. (1998) Genetic instabilities and hereritary Neurological Diseases. San Diego: Academic Press

Yasuda, S. and Hirota, Y. (1977) Cloning and mapping of the replication origin of *Escherichia coli*. Proc. Natl. Acad. Sci. USA 74, 5458-5462.

Zhao, J., Bacolla, A., Wang, G., Vasquez, K. M. (2010) Non-B DNA structure-induced genetic instability and evolution. Cell. Mol. Life Sci. 67, 43-62.