The study of G-quadruplex in supercoiled DNA

Lv, Bei

2016

Lv, B. (2016). The study of G-quadruplex in supercoiled DNA. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/68928>

<https://doi.org/10.32657/10356/68928>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

The Study of G-quadruplex in Supercoiled DNA

LV BEI

SCHOOL OF PHYSICAL AND MATHEMATICAL SCIENCES

2016

The Study of G-quadruplex in Supercoiled DNA

LV BEI

School of Physical and Mathematical Sciences

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy

2016

Acknowledgements

First and foremost, I would like to express my sincere and deep thank to my supervisor. Professor Li Tianhu has encouraged and inspired me during my graduate study in Nanyang Technological University. I have benefited greatly from the comprehensive knowledge and logical way of thinking he have taught me those years.

I appreciate so much the help and advice on DNA sequence design from Dr. Li Dawei. He encouraged me and supported me throughout my studies. I will always be truly grateful for this.

I would like to thank all my laboratory mates (Dr. Zhang Hao, Li Yiqin Jasmine, Hiew Shu Hui, Dr. Ng Tao Tao Magdeline, Dr. Li Cheng, Dr. Lei Qiong and Dr. Ba Sai) in Prof Li Tianhu's group, it was fun working with you. I am also grateful to all my friends in Singapore, China and elsewhere for the support and encouragement all the time.

Furthermore, I would like to acknowledge all support staff in CBC general office, and the chemical store. The financial support of the Ministry of Education and Nanyang Technological University in Singapore is gratefully acknowledged.

Most importantly, I owe my loving thanks to my parents and lovely daughter. Without their encouragement and understanding, it would have been impossible for me to finish my graduate study in NTU. Thank you!

Table of Contents

Acknowledgements.....	I
Table of Contents.....	II
Abstract.....	VI
List of Tables.....	IX
List of Figures.....	X
List of Abbreviation.....	XV
Chapter 1 – Introduction	
1.1 Background.....	1
1.2 The Conformation of B-form DNA.....	3
1.2.1 The Primary Structure of DNA.....	3
1.2.2 Secondary Structure of DNA and Watson-Crick Model.....	5
1.2.3 B-form DNA.....	7
1.3 DNA Supercoiling.....	9
1.3.1 Measurement of DNA Supercoiling.....	11
1.3.2 Nicked and Relaxed Circular DNA.....	12
1.3.3 Negative Supercoiling.....	13
1.3.4 Positive Supercoiling.....	14
1.4 G-quadruplex and Non-B DNA structures.....	15
1.4.1 The importance of Non-B DNA Structures.....	15
1.4.2 Basic Characterization of the G-quadruplex Structures.....	15
1.4.3 Variations on the Structures of G-quadruplexes.....	17

1.4.4 G-quadruplex functions----Telomeres and Telomerase.....	18
1.4.5 G-quadruplex functions---- Transcription Regulation.....	20
1.4.6 G-rich Sequences in Genome: Duplex-Quadruplex Competition.....	21

Chapter 2 –DNA gyrase-driven generation of a G-quadruplex from plasmid DNA

2.1 Introduction.....	23
2.2 Sequence Design of the Circular DNA with G-quadruplex.....	26
2.2.1 The Strategy to Construct a Mini-plasmid DNA with Circular Backbone.....	26
2.2.2 The Strategy to Construct a Circular DNA with guanine-rich segment.....	27
2.3 Materials and Methods.....	30
2.3.1 Vectors, Oligonucleotide, Enzymes and Chemicals.....	30
2.3.2 Synthesis of Linear DNA 1 and Linear DNA 3 through Polymerase Chain Reaction.....	31
2.3.3 Synthesis of Linear DNA 2 and Linear DNA 4 with Two Identical Cohesive Ends through SacI Digestion.....	33
2.3.4 Circularization Reaction Catalyzed by T4 Ligase.....	33
2.3.5 Removing Linear DNA Products from Ligase Reaction Mixture Using Nuclease BAL-31 Exonuclease.....	34
2.3.6 Introducing Negative Supercoils into DNA Circles by DNA Gyrase under Physiological-like Conditions.....	35
2.3.7 Removing DNA supercoils by Nicking Endonucleases Nt.BsmAI.....	35
2.3.8 AFM Examination of Obtained Circular DNA.....	36
2.3.9 Reaction of T7 endonuclease with Non-B DNA.....	38

2.4 Results and Discussion.....	38
2.4.1 Synthesis and Structure Examination of G-rich-containing Circular DNA 1.....	38
2.4.2 Formations of G-quadruplexes Facilitated by DNA Gyrase under Physiological Conditions of K ⁺ and their Conformation by Electrophoresis.....	42
2.4.3 Examining the Formations of G-quadruplexes Facilitated by DNA Gyrase under Physiological Conditions of K ⁺ using AFM.....	44
2.4.4 Analysis of Reaction between DNA Gyrase and Circular DNA 1 under the Concentration of K ⁺ of Non-physiological Condition.....	49
2.4.5 Confirmation of Absence of G-quadruplex Structure in Non- guanine-rich Circular DNA 5.....	52
2.4.6 Confirmation of the Existence of G-quadruplex in Circular DNA by Endonuclease.....	55
2.5 Conclusion.....	60

Chapter 3 –Disintegration of cruciform and G-quadruplex structures during the course of helicase-dependent amplification (HDA)

3.1 Introduction.....	61
3.2 Sequence Design of the Circular Template DNA with Non-B Structures.....	89
3.2.1 The General Strategy to Construct a Template DNA with Circular Backbone.....	69
3.2.2 The Strategy to Construct a Circular DNA with Cruciform and the Following Structural Confirmation.....	70
3.2.3 The Strategy to Construct DNA 3 and DNA 5 and the Following Structural Confirmation.....	79
3.3 Materials and Methods.....	83
3.3.1 Vectors, Oligonucleotide, Enzymes and Chemicals.....	83

3.3.2 Experimental Procedures for Helicase-Dependent Isothermal DNA Amplification and AFM Examination.....	84
3.3.3 Experimental Procedures for Synthesis and Structural Confirmation of DNA 1.....	86
3.3.4 Experimental Procedures for Synthesis and Structural Confirmation of DNA 3 and DNA 5.....	88
3.4 Results and Discussion.....	90
3.4.1 Breaking Down Cruciform in the Course of Isothermal DNA Replication (HDA).....	90
3.4.2 Confirmation of Breaking Down Cruciform Structures by Topo I Relaxation.....	93
3.4.3 Examination of the Stability of Cruciform Structures in HDA Buffers.....	95
3.4.4 The Control Experiment to Examine Breaking down Cruciform is affiliated with Positive Supercoils.....	96
3.4.5 Breaking Down G-quadruplex in the Course of Isothermal DNA Replication (HDA).....	98
3.4.6 The Control Experiment to Examine Breaking down G-quadruplex is affiliated with Positive Supercoils.....	101
3.5 Conclusion.....	103
References.....	105
List of Publications.....	117

Abstract

G-quadruplex is a kind of non-B structure, which is formed in guanine-rich nucleotide sequence with various conformation and stabilized by positive ion. Motifs for generating of G-quadruplex structures are widespread in genomic sequences of prokaryote and eukaryote. Its unique spatial arrangement as well as its great biological significance have received considerable attention in the past few years. Except 3' overhang in telomere, G-rich sequences widely exist in the duplex regions of genomic DNA, where the formation of G-quadruplex was blocked by not only its complementary strands but also the adjacent duplex regions. It therefore becomes more and more important to study the competition mechanism between duplex and G-quadruplex. DNA supercoiling, on the other hand, can change the structures of double helix by unwinding or overwinding DNA double strands. The mechanisms of formation and disintegration of G-quadruplex structures in the supercoiled circular DNA are studied and discussed in this thesis.

In the first project, we demonstrated that the DNA gyrase, an essential bacterial enzyme which possesses the unique ability to introduce negative supercoils into a DNA circle, can drive the G-quadruplex generation from plasmid DNA under the intracellular concentration of potassium in prokaryotes. It is showed in our studies that

the formed G-quadruplex in circular duplex DNA is evidently verifiable through electrophoretic analyses, atomic force microscopic examinations as well as enzymatic assays. Since the formation of G-quadruplex structure from G-rich segments of a long duplex DNA is not a spontaneous process, a driving force must be attributed to initiate G-quadruplex formation in genomic DNA. We speculate that once those negative supercoils induced by DNA gyrase are formed, the double helix structures between two single strands are unwound and intrastrand base-pairing could occur if there are guanine-rich segments in the DNA sequences. Since DNA gyrase is a prokaryote-exclusively owned enzyme that is absent in eukaryotes, the outcomes of our investigations could suggest that prokaryotic cells might utilize this topological enzyme to regulate the generation of G-quadruplex to comply with their subsequent cellular functions.

In the second project, the relationship between positive supercoiling and thermodynamically stable non-B structures was studied. It is known that physical alterations of B-form of DNA such as G-quadruplex and cruciform structures occur commonly in organisms that serve as signals for specified cellular events. Although the modes of action for repairing of chemically damaged DNA have been well studied nowadays, the repairing mechanisms for physically altered DNA structures have not yet been understood. Our *in vitro* studies show that both breakdown of thermodynamically stable G-quadruplex and cruciform structures and resumption of canonical B-conformation of DNA can take place during the courses of isothermal helicase-dependent amplification (HDA). Since positive DNA supercoils is overwound, the DNA structures with positive supercoils are anticipated to hold more

backbone constraints than its negative counterpart does. We speculated that the pathway that makes the non-B structures repairable is the relieving of the accumulated torsional stress that was caused by the positive supercoiling. Our new findings suggest that living organisms might have evolved this distinct and economical pathway for repairing their physically altered DNA structures.

List of Tables

Table 2.1	Double-stranded sequence of Linear DNA 1.	28
Table 2.2	Double-stranded sequence of Linear DNA 2.	29
Table 2.3	Double-stranded sequence of Circular DNA 1.	29
Table 2.4	Vectors, Oligonucleotide, Enzymes and Chemicals used in this research.	30
Table 2.5	Nucleotide sequences of primers used in polymerase chain reactions.	31
Table 2.6	Analysis of yield (%) of circularization in different concentration of substrate and reaction time.	40
Table 2.7	Double-stranded sequence of Circular DNA 4.	52
Table 3.1	Sequence of plasmid DNA X2420 (3593 bp) and X4511E (4220 bp).	71
Table 3.2	Double-stranded sequence of DNA 1.	77
Table 3.3	Double-stranded sequence of DNA 3	82
Table 3.4	Vectors, Oligonucleotide, Enzymes and Chemicals used in this research.	83
Table 3.5	Nucleotide sequences of primers used in isothermal helicase-dependent amplification (HDA).	85
Table 3.6	Nucleotide sequences of primers used in our polymerase chain reactions.	87

List of Figures

Figure 1.1	Pictorial illustration of B-form DNA and some typical non-B DNA structures.	2
Figure 1.2	Molecular structures of bases in DNA.	3
Figure 1.3	Schematic illustration of primary structure of DNA, a linear sequence of nucleotides.	4
Figure 1.4	Pictorial illustration of duplex structure of DNA and base-pairing stacking.	6
Figure 1.5	Pictorial illustration of B-form DNA and right-handed conformation.	7
Figure 1.6	Molecular model of A-form, B-form and Z-form DNA.	8
Figure 1.7	Pictorial illustration of DNA supercoiling.	10
Figure 1.8	Pictorial illustration of negative (A), relaxed (B) and positive supercoil (C).	14
Figure 1.9	Structures of G-quartet and G-quadruplex.	16
Figure 1.10	Strand stoichiometry variation of G-quadruplexes.	17
Figure 1.11	Strand arrangements of G-quadruplexes.	18
Figure 1.12	Schematic representation of structures of human telomere and formation of G-quadruplex in the single strand region of human telomere.	19
Figure 1.13	Formation of a G-quadruplex in promoter of a gene can affect the level and nature of transcription from that gene.	20

Figure 2.1	Schematic illustrations of competitive mechanism between G-quadruplex and duplex.	24
Figure 2.2	Schematic illustrations of our general strategy to exam the possibility of formation G-quadruplex in negatively supercoiled DNA caused by DNA gyrase.	25
Figure 2.3	Diagrammatic illustration of synthetic route toward Circular DNA 1.	27
Figure 2.4	Electrophoretic analysis of intermediate DNA molecules generated during the synthesis of Circular DNA 1.	39
Figure 2.5	AFM images of Circular DNA 1 with its scale bar of 200 nm.	41
Figure 2.6	Diagrammatic illustration of reactions of Circular DNA 1 upon the action of DNA gyrase and other enzymes.	42
Figure 2.7	Electrophoretic analysis of products of enzymatic reactions on Circular DNA 1.	43
Figure 2.8	AFM examination of DNA with G-quadruplex-containing negative supercoils	44
Figure 2.9	AFM images of Circular DNA 2.	45
Figure 2.10	Section analyses of AFM images of Circular DNA 2.	46
Figure 2.11	Frequency distributions of the lengths (nm) of Circular DNA 1 (A) and Circular DNA 2 (B).	47
Figure 2.12	AFM examination of G-quadruplex-containing linear DNA.	49
Figure 2.13	Examination of action of DNA gyrase on Circular DNA 1 under a non-physiological concentration of potassium ions.	50

Figure 2.14	AFM examination of the topological structures of Circular DNA 1 and Circular DNA 3.	51
Figure 2.15	Examination of action of DNA gyrase on Circular DNA 4 under physiological concentrations of potassium ions.	53
Figure 2.16	AFM images of the obtained relaxed form of closed circular DNA.	55
Figure 2.17	Diagrammatic illustration of the non-matched sites in G-quadruplex- containing DNA.	55
Figure 2.18	Diagrammatic illustration of our enzymatic confirmation of presence of G-quadruplex structures in Circular DNA.	56
Figure 2.19	Enzymatic confirmation of presence of G-quadruplex structures in Circular DNA 2.	57
Figure 2.20	AFM examination of DNA products obtained after T7 Endonuclease I cleavage.	58
Figure 2.21	Enzymatic confirmation of absence of G-quadruplex structures in Circular DNA 5.	59
Figure 3.1	Diagrammatic illustration of the DNA reparation mechanisms by photoreactivation and alkyltransferase.	62
Figure 3.2	Diagrammatic illustration of positive and negative supercoiling.	64
Figure 3.3	Diagrammatic illustration of HDA.	67
Figure 3.4	Schematic illustration of the topological relationships during the course of DNA replication <i>in vitro</i> within a circular DNA.	68

Figure 3.5	Schematic illustration of the broken down of non-B structures during HDA within a circular DNA.	69
Figure 3.6	Schematic illustration of our synthetic route towards DNA 1.	70
Figure 3.7	Synthesis and structural confirmation of DNA S3.	76
Figure 3.8	Synthesis and structural confirmation of DNA 1.	78
Figure 3.9	Schematic illustration of our synthetic route towards DNA 3 and DNA 5.	79
Figure 3.10	Synthesis and structural confirmation of DNA S7.	80
Figure 3.11	Synthesis and structural confirmation of DNA 3 and DNA 5 according to our previously reported method.	81
Figure 3.12	Pictorial diagram of an envisioned disintegration of DNA cruciform structures by positive DNA supercoiling.	91
Figure 3.13	Disintegration of DNA cruciform structures during the course of isothermal HDA.	92
Figure 3.14	Examination of the absence of cruciform structures by removing the supercoils in DNA 2 with Topo I.	94
Figure 3.15	Examination of effects of buffer and salts on the stability of the cruciform residing in DNA 1.	95
Figure 3.16	Examination of the interaction between helicase and cruciform in DNA 1.	96
Figure 3.17	Pictorial diagram of an envisioned disintegration of G-quadruplex structures by positive DNA supercoiling.	98

Figure 3.18	Disintegration of G-quadruplex structures during the course of isothermal HDA.	100
Figure 3.19	Pictorial diagram of an envisioned reaction pathway of G-quadruplex-containing DNA 5 in HDA reaction.	101
Figure 3.20	Examination of the presence of G-quadruplex in DNA 6 after the action of isothermal HDA.	102

Table of Abbreviations

AFM	Atomic Force Microscopy
APS	1-(3-aminopropyl)silatrane
A-tract	Adenine-tract
Bp	Base pairs
BSA	Bovine Serum Albumin
°C	degree Celsius
DNA	Deoxyribonucleic acid
dsDNA	double stranded DNA
ssDNA	single stranded DNA
Lk	Linking number
Tw	Twist number
Wr	Writhe number
EB	Ethidium Bromide
PNA	Peptide Nucleic Acid
PCR	Polymerase Chain Reaction
Topo I	Human topoisomerase I
Topo II	Human topoisomerase II
TAE	Tris, Ammonium acetate, EDTA buffer
TBE	Tris, Boric acid, EDTA buffer
EDTA	Ethylenediaminetetraacetic acid
TRIS	Tris(hydroxymethyl)aminomethane

Chapter 1

Introduction

1.1 Background

DNA is a biological macromolecule that contains the genetic information used by living organisms.(1-3) In 1953, Watson and Crick deduced a model for the structure of DNA, which is called double-helical model. In this model, pyrimidine bases can only form hydrogen bonds to purine bases (adenine (A) pairing only to thymine (T) and cytosine (C) bonding only to guanine (G)) by forming two or three hydrogen bonds.(4-6) Structurally, a negative charged sugar-phosphate backbone is on the outside, which can make DNA molecules hydrophilic. On the other hand, all the important atoms of bases are protected from the environmental chemical damage by forming a large hydrophobic interior.(1) B-form is a kind of structure which has a right-handed double helix with a major groove and a minor groove. It has been well studied that Watson and Crick base pair and B-form conformation are the most common structures adopted by DNA *in vivo*.(5-7)

However, particular DNA sequences under certain conditions can also form alternative conformations of DNA due to the different base pair arrangement, which

are categorized as non-B DNA structures such as cruciform structures(8), bulges(9), G-quadruplexes(10), triplexes(11), slipped structures(12) as shown in Figure 1. Although double helix models which is most adopted by DNA, non-B DNA conformations also exist in both eukaryotic and prokaryotic genomes. Among all the non-B DNA structures, G-quadruplex, a thermodynamically stable structural entity, has been deemed to possess great biological significance.(13-17) Since the sequences of formation of G-quadruplex structures widely exist in both prokaryotic and eukaryotic genomes, competition mechanisms between DNA duplex and G-quadruplex have attracted more and more attentions in the past few years. This research will focus on the formation and disintegration of G-quadruplex DNA structures within genomic DNA under physiological conditions.

Figure 1.1 Pictorial illustration of B-form DNA and some typical non-B DNA structures.

1.2 The Conformation of B-form DNA

1.2.1 The Primary Structure of DNA

Figure 1.2 Molecular structures of bases in DNA.

DNA is the carrier of genetic information. The basic component of DNA is a nucleotide, which is consisted of a pentose carbon sugar (2'-deoxyribose), a nitrogen containing base and a phosphate group. Purine and pyrimidine are two kinds of bases in DNA. Structurally, a purine is composed of carbon and nitrogen, which is known as heterocyclic aromatic base. There are two common purine bases are found in DNA: adenine and guanine. A pyrimidine base is the six-membered rings and the two common pyrimidine bases in DNA are thymine and cytosine (Figure 1.2). To form a nucleotide, the nitrogen bases need to be covalently attached to the pentose carbon sugar by a glycosidic bond. The purine bases can form the glycosidic bonds using the nitrogen at 9 position while the pyrimidine bases attach to 2'-deoxyribose with nitrogen at 1 position. The purine and pyrimidine bases can only linked to the carbon at 1' position of the deoxyribose sugar. Finally, the phosphate group forms an ester

bond with 5' hydroxy group of the pentose sugar by one of the negatively charged oxygen groups.

Figure 1.3 Schematic illustration of primary structure of DNA, a linear sequence of nucleotides.(3)

The primary structure of DNA refers commonly to a linear sequence of nucleotides (Figure 1.3). In DNA, monomer nucleotides are linked by phosphodiester bonds between the 5' and 3' carbon atoms, which are known as polynucleotides. Each polynucleotide chain has two distinct ends which are named as 5' and 3'. Since the chemical and biological properties of the two ends are quite different, polarity can be found in a DNA strand. Generally, a hydroxy group appears at 3' end of a DNA strand while a single phosphate group exists in the 5' end. By convention, sequence of a DNA strand is usually described from its 5' end to the 3' end. Since no branch can be observed in a nucleic acid strand normally, covalent structure of the DNA molecule

can be described through specifying the sequence of DNA strand, which is why the sequence of the polynucleotides is defined as the primary structure of DNA.

1.2.2 Secondary Structure of DNA and Watson-Crick Model

The secondary structure of DNA refers commonly to the base-pairing interactions within a single DNA molecule. It can be described that a group of bases which are paired in a DNA molecule. The most famous theory to represent the secondary structure of DNA is called Watson-Crick Model.(5)

In the early 1950s, Chargaff reported that the amount of adenine bases always equaled to the amount of thymine bases in DNA molecules while the amount of guanine bases equaled to the amount of cytosine bases, which was known as “Chargaff’s Rules”(18,19). It came for the fact that all the DNA samples used by Chargaff were purified from numerous different organisms and the GC (or AT) content in a DNA molecule is unrelated to the rule. In addition, the experimental data of X-ray diffraction of DNA fibers was also reported at that time.(20,21) Based on the two important pieces of information, Watson and Crick deduced a model to describe the structure of DNA double helix, which was known as Watson-Crick Model, one of the most important scientific discoveries in twentieth century. In this model, duplex DNA possesses a right-handed helical structure, which is constructed by two complementary single-stranded DNA arranges in an antiparallel fashion. According to the model, one DNA strand aligns in the direction of 5' to 3' and the other strand is close to it in the direction of 3' to 5'. With such a configuration, the pentose carbon sugar and phosphate group construct the backbone of the double helix structures. This

stereo arrangement of sugar and phosphate group can form a hydrophilic sugar-phosphate backbone, which is outside of the duplex DNA. On the other hand, the large hydrophobic interior containing bases are surrounded inside. Because the genetic information mainly stores in the nitrogen bases, the double helix structure can protect all the important function groups and atoms of the base from hydrolysis or chemical damage in the "tough" condition in the cell.(1)

Figure 1.4 Pictorial illustration of duplex structure of DNA and base-pairing stacking.

This picture was contributed by *Madeleine Price Ball* in Wikipedia.

The hydrogen bonds between paired bases can hold two strands together, which are known as “Watson-Crick base pairing”. An adenine base can form two hydrogen bonds with thymine base while three hydrogen bonds link guanine base and cytosine base. The paired bases are stacked each other and it is believed that the effects of base-pairing stacking contribution into thermal stability of the DNA double helix

significantly as shown in Figure 1.4.(1,22) It should be pointed out that “Watson-Crick base pairing” is not the only conformation for base-pairing. Due to the effects of tautomerization and ionization in solution, the properties of bases can be changed and alternative ways to form hydrogen bonds between two bases are observed.(1)

1.2.3 B-form DNA

Figure 1.5 Pictorial illustration of B-form DNA and right-handed conformation.

The double helix conformation of B-form is believed to represent the most common form adopted by DNA in cells.(1-3,23,24) The structure of B-form was first derived from X-ray diffraction analysis of sodium salt of DNA fibers at 92% relative humidity.(25,26) Several years ago, more detail parameters reported to described the helical structure of double helix B-form structures.(27-30) As shown in Figure 1.5, the helical tune in B-form DNA is measured to be every 10.4 to 10.5 bp. It means that with

helical axis rising one 360° rotation needs to be finished in every 10.4 to 10.5 bp. The axial rise in B-form DNA is about $3.4 \text{ \AA}$, which means that the distance between two neighboring bases is measured to be $3.4 \text{ \AA}$ in B-form DNA. The conformation of the ribose sugar in B-form DNA is C2'-endo. What is particularly intriguing is that there are two distinct grooves in B-form DNA as shown in Figure 1.5. During the dynamic cellular processes, different proteins can recognize and bind to the major groove or minor groove for different biological purposes.(31-33) In addition, particular drugs can be designed to fit and target the major groove or minor groove of the B-form DNA for some therapeutic aims.(34-36)

Figure 1.6 Molecular models of A-form, B-form and Z-form DNA.

Apart from B-form structures, duplex DNA can adopt other helical structures such as A-form and Z-form as shown in Figure 1.6. A-form DNA holds the same right-handed helical structures as B-form DNA does. What makes these two structure different is that the conformation of the ribose sugar in A-form is C3'-endo while B-form DNA holds a C2'-endo ribose sugar as mentioned above. RNA molecules can also form double stranded structures in some certain conditions, which was believed to

be in a A-form conformation.(37) Z-form DNA is in a distinct left-handed helical structures, in which sugar-phosphate backbone arranges in a zig-zag pattern.(38) The formation of Z-form DNA was reported in some DNA with alternating pyrimidine-purine sequences (such as d(GC)₂ in high salt condition or negative supercoiled conformation.(39-42) However, in real cellular condition double helical B-form structure is the most common conformation that is adopted by DNA molecules.

Living organisms storing most of their genetic information into B-form DNA may have great biological significance. First of all, as described above, the stereo arrangement of B-form DNA can protect the central of the helix and make purine and pyrimidine bases chemically inert, where genetic instructions stored in it. Secondly, during the course of semi-conservative replication, two complementary strands in B-form DNA can serve as the templates for DNA polymerase to produce two exact same copies of daughter DNA.(43-45) Thirdly, when DNA damage caused by UV irradiation or chemicals, the undamaged strand can provide template for the DNA repair machine, which is known as Double-Strand Break (DSB) repair pathway in DNA recombination.(46-50)

1.3 DNA Supercoiling

DNA supercoiling is the tertiary structure of nucleic acids. It means that molecular architecture of DNA that exists in space in a self-twisted fashion as shown in Figure 1.7. It is believed that DNA supercoiling (the global alteration of DNA structure) is caused directly from the double helical property of the DNA molecule.(39,51-53) Shortly after the establishment of Watson-Crick Model in 1953,

two forms of DNA (Form I and Form II) were found when scientists studied the DNA molecules of tumour virus with sedimentation analysis that separates components according to the size and compactness of DNA molecules.(1,54,55) Although all the two components with same molecular weight exhibited double helix structures, different properties were observed. DNA with Form I DNA was tested to be more compact (with higher sedimentation coefficient) than Form II DNA.

Figure 1.7 Pictorial illustration of DNA supercoiling.

It was therefore suggested that Form I DNA possessed a circular structure which was constructed by formation of two phosphodiester bonds between 5' and 3' ends of the single linear DNA molecule. At the same time, electron micrograph analysis showed that Form I DNA was observed to have more self-crossings within its molecule. It was why Form I DNA showed more compact conformation. The Form I DNA is now called to be supercoiled DNA, which exhibits in a tangled and twisted structure. On the other hand, Form II DNA appeared in the relaxed circular structures and one broken phosphodiester bond was often found in one of the two complementary helical strands, which is known as the nicked or relaxed DNA. If the breakage of DNA phosphodiester backbone occurred in both strands of duplex circular DNA at the same point or the two broken points are very near, the circular structure

could be changed into a linear form, which is Form III DNA and it has been named as linear DNA now.

As discussed above, DNA supercoiling should only be discussed in a circular system, because the broken strands can rotate around the intact strand to dissipate the torsional constraints.(3) Duplex DNA with a circular conformation is known as a cccDNA (covalently closed circular DNA). It has been studied that the molecular structures of DNA stored in prokaryotic cell such as bacteria, archaea and mitochondria existed almost all in a circular structures.(56,57) For example, plasmid DNA purified from bacterial cell is a cccDNA and it is often classified into Form I DNA. On the other hand, although eukaryotic chromosomal DNA has two open ends and appears in a linear conformation, the supercoiling behaviors can also be observed within its molecular structure. This happens because any backbone of inside duplex DNA segments in eukaryotic chromosome cannot be freely rotated to remove the torsional stress if the two ends of chromosomal DNA are too far away to be reachable.(58) In this case, any segments in eukaryotic chromosome should be considered as virtually circular DNA.

1.3.1 Measurement of DNA Supercoiling

The measurement of DNA supercoiling can be described through mathematical methods, which has been reported in many research articles.(59-65) To address the properties of DNA supercoiling, three important key mathematical concepts were used: Tw, Wr and Lk. Tw is the twist number, which represents the total number of helical turns within a DNA molecule or a given segment of duplex DNA. The value of Tw

can be calculated by dividing the total numbers of DNA base pairs by 10.5 (One turn in B-form DNA need 10.5 bp in standard solution).(3) W_r is writhe number, which describes the spatial coiling of DNA duplex backbone in a three-dimensional space. L_k is the linking number of DNA double helix, which means the total number of crossing points of the two DNA strands if the whole DNA molecule is flattened on a geometric plane. Theoretically, linking number of a DNA must be an integer since two strands must always be coiled each other in an integral number of times. The value of L_k is equal to the sum of T_w plus W_r as shown in Equation 1.1. It should be pointed out that the value of linking number of any cccDNA is fixed and the value of twist number is changeable according to the type and concentration of the salts, nucleotide sequence, temperature, pH value, pressure and other factors in solution.(66,67)

$$L_k = W_r + T_w \quad (\text{Equation 1.1})$$

1.3.2 Nicked and Relaxed Circular DNA

DNA with a linear structure possesses two open ends which can rotate freely in solution. Such conformation of DNA double helix represents a preferred structure, in which the helical repeats is around 10.4 to 10.5 bp per helical turn.(6) One or several single-stranded breaks can be found in the circular backbone of a nicked DNA, where the same arrangement of double helix appears because of the swivel around the nicked site(s) between the two DNA strands. The lowest energy can be achieved in both linear and nicked DNA molecules.(4,68,69)

The relaxed DNA is a kind of DNA molecule which is also in the same helical structural arrangement as those in linear or nicked DNA but the backbone of this DNA is covalently closed. Theoretically, no supercoiling can be found in relaxed DNA and energy in relaxed DNA is also lower than the supercoiling counterpart. The linking number in relaxed DNA is defined as Lk_0 , which can be calculated with the equation as shown in Equation 1.2. N in this equation is the total number of base pairs in DNA molecules. As shown in Figure 1.8B, the linking number (Lk) in the relaxed DNA equals to Lk_0 ($210/10.5 = 20$). The twist number (Tw) is also 20 ($210/10.5$). Therefore, the linking number (Lk) equals to the twist number (Tw) in this DNA and the writhe number (Wr) should be zero based on Equation 1.1, which means that there is no self-crossing point in the DNA backbone and no supercoil appears in the relaxed DNA. Similar with Tw , the value of Lk_0 is also changeable according to the type and concentration of the salts, nucleotide sequence, temperature, pH value, pressure and other factors in solution.

$$Lk_0 = N/10.5 \quad (\text{Equation 1.2})$$

1.3.3 Negative Supercoiling

In living organism, DNA is mainly stored in negatively supercoiling structures.(4,68,70,71) The helical conformation in negative supercoiling is underwound, which means that fewer helical turns can be found in the molecule than those in a relax DNA. It has been established that negative supercoiling may weaken the interactions between paired bases and such conformation facilitates the formation of denaturation bubbles along the circular DNA backbone during the course of

replication and transcription.(39,68) The linking number (Lk) of a negative supercoiling DNA is less than Lk_0 . As shown in Figure 1.8A, the linking number (Lk) in this DNA is 19 and the decrease of linking number can only be achieved by topoisomerase. The twist number (Tw) equals to Lk_0 (210/10.5). As a result, the writhe number in this DNA is -1 according to Equation 1.2 ($19 - 20 = -1$). It means that DNA is in a right-handed negative supercoil structure and the self-crossing point is one.

Figure 1.8 Pictorial illustration of negative (A), relaxed (B) and positive supercoil (C).

1.3.4 Positive supercoiling

Although DNA isolated from most living cells is in a negative supercoiling structure, positive supercoiling can also be adopted by genomic DNA in some organism such as hyperthermophiles.(72-74) In addition, during the course of DNA replication positive supercoils can be generated ahead of the replication forks, which should be solved by topoisomerases *in vivo*.(68,75) Positive supercoil is overwound in the helical structure of DNA, where more helical turns appear in the given DNA double helix than those in its relaxed counterpart. In Figure 1.8C, 21 of linking

number (Lk) can be achieved by topoisomerase and the twist number (Tw) in this DNA also equals to 20 ($210/10.5 = 20$). This DNA is in a left-handed positive supercoil structure and the self-crossing point is one. It should be pointed out that higher energy exists in both negatively and positively supercoiled DNA than the relaxed form.

1.4 G-quadruplex and Non-B DNA structures

1.4.1 The importance of Non-B DNA Structures

Besides the well-recognized canonical B-form conformation, many other structural forms of DNA are known to exist under physiological conditions such as cruciforms, G-quadruplex, i-motif, triplexes, slipped structures, folded slipped structure, and left-handed Z-DNA, which are often named “non-B DNA structures” (Figure 1.1).(76-80) It has been demonstrated in the past years that these non-B DNA structures are present *in vivo* and play vital roles in various cellular processes.(76,81,82) The cruciform structures of DNA, for example, are believed to form at or near replication origins of some eukaryotic cells and serve as recognition signals for DNA replication.(83-85) In addition, it has been demonstrated that G-quadruplex generated in the promoter region of *c-myc* gene acts as a transcriptional repressor element for the expression of the gene.(14,86) Moreover, some previous studies suggested that H-DNA could play important roles in transcription, replication as well as genetic recombination.(87-89)

1.4.2 Basic Characterization of the G-quadruplex Structures

In 1965, Gellert and co-workers reported a structure called G-quadruplex, which is composed of guanine bases only.(90) Thus in 1989, Williamson and colleagues suggested a model, in which the formation of a cyclic arrangement of eight of Hoogsteen hydrogen bonds gives a planar structure which is called G-quartet. Each guanine base serves as both the acceptor and donor in a G-G Hoogsteen hydrogen bond and there is a cavity formed in the core to serve as the binding site for a monovalent (or divalent) cation which can help to maintain the stability of the structure. More generally, the radius of cation is a crucial parameter to determine how the G-quartets are stabilized (Figure 1.9). In the monovalent cation series, potassium ion is the best choice as a stabilizer and the order is generally $K^+ \gg Na^+ > Rb^+ > NH_4^+ > Cs^+ \gg Li^+$. For the earth alkali series (divalent cation), strontium ion has the strongest capacity to stabilize the G-quadruplex and the order is $Sr^{2+} \gg Ba^{2+} > Ca^{2+} > Mg^{2+}$.(10) Since these G-quartets have large π -surfaces, they can overlap each other due to the π - π interaction and a series of nucleic acid secondary structures can be formed, which are so called G-quadruplexes.(91)

Figure 1.9 Structures of G-quartet and G-quadruplex. Left: a G-quartet. Right: an intramolecular G-quadruplex.(91)

G-quadruplex is a type of non-B structure of nucleic acids that is composed of stacked G-tetrads in its columnar congregation (Figure 1.9).(14,77,82,86,91-93) G-quadruplexes, among the non-B DNA structures, could possess their melting points as high as 90 °C under physiological conditions.(10,91,94-96) In addition, it was demonstrated in the past that G-quadruplexes served as a transcriptional repressor element for the expression of promoter region of *c-myc* gene.(97,98) It is estimated that there are a million guanine-rich sites in eukaryotic cells that have the potential for forming G-quadruplexes, which exhibits possible prevalence of this non-B structure *in vivo*.(95,99,100)

1.4.3 Variations on the Structures of G-quadruplexes

Figure 1.10. Strand stoichiometry variation of G-quadruplexes.(10)

G-quadruplexes can be formed by one strand (Figure 1.10A), two strands (Figure 1.10B) or four strands of DNA (Figure 1.10C).(10) Theoretically, there is a possibility to form a G-quadruplex with the arrangement of three strands but have yet to be proved. It has been reported that G-quadruplexes with variation of strand stoichiometry depend on the concentration of each DNA strands. On the other hand, orientations of the DNA strands lead to different G-quadruplex conformations. Since there is a strand polarity which is customarily described as from the 5' end to the 3'

end, the polymorphism of the four strands arrangement should be observed. Figure 4 depicts the possible structures formed with intermolecular or intramolecular interaction due to different strand stoichiometries and folding patterns. A is formed by different four parallel strands of DNA; B is constructed with single molecular with the parallel strands adjacent to each other, which is also described as antiparallel structure; C is unimolecular antiparallel structure with alternating parallel strands; D - F are unimolecular structures with parallel or antiparallel strands, which have been observed in the regions of human telomeric repeat.(91)

Figure 1.11. Strand arrangements of G-quadruplexes.(91)

1.4.4 G-quadruplex functions----Telomeres and Telomerase

Since DNA polymerase *in vivo* can exclusively amplify DNA in a direction from 5' end to 3' end, only one daughter strand is synthesized with a continuous manner which is called leading strand. On the other hand, another daughter strand (lagging strand) is synthesized in discontinuous pieces called “Okazaki fragments” using RNA primers and fragments of DNA are joined together by DNA ligase.(101-103) At the very end of chromosomes, lagging strand cannot be made because there is no binding site for RNA primers to attach. The result is the daughter DNA become smaller and smaller, which is known as “end replication problem”.(104-106)

To solve this problem, chromosome DNA is put a cap on its end which is called telomere.(90,107-111) A telomere does not contain any genetic information and then it will not lead to severe consequences for the cells. It has been demonstrated in the past that telomere can help to protect the end of the chromosome from deterioration or fusion with neighboring chromosomes.(112,113) There are repeated sequences located at the ends of telomere regions in eukaryotic genomes and the character of those sequences is that short guanine tracts spaced periodically along the DNA backbones. In all vertebrates, such repetitive sequence is $d(GGGTTA)_n$ and other organisms generally have very similar sequences.(110,114,115) As shown in Figure 1.12, there are two parts in the region of human DNA telomere, one is double stranded human telomeric repeats, the length of which ranges from 5 to 15 kilo-base pairs depending on the tissue type and several other factors. On the other hand, another part of human telomeric DNA is 75 to 300 nucleotides of single stranded 3' overhang. It has been studied that eukaryotes maintains their intracellular concentration of potassium ion from 0.1 to 0.6 M under physiological condition, which makes the single strands G-rich DNA fold into a stable G-quadruplex structure as shown in Figure 1.12.(115)

Figure 1.12 Schematic representation of structures of human telomere and formation of G-quadruplex in the single strand region of human telomere.

1.4.5 G-quadruplex functions---- Transcription Regulation

Gene expression is precisely regulated by various methods. It has been studied that G-quadruplex formation sequences are widely spread in the promoter region of the genes.(116) A novel method has been established to investigate the potential G-quadruplex formation sites in the promoter regions of human genes.(95) By this analysis, it is found that there are almost half of all genes (about 43%) containing G-quadruplex formation motifs in their promoter regions.(117) More interestingly, G-quadruplexes are more likely to appear near the transcription start site (TSS) of the cancer genes. It is shown that more than half of those cancer genes (67%) possesses G-quadruplex promoter.

Figure 1.13 Formation of a G-quadruplex in promoter of a gene can affect the level and nature of transcription from that gene.(91)

It is known that the expression of *c-myc gene* can give an important transcription factor which is involved in regulation of 15% of all human genes.(118) On the other hand, overexpression of *c-myc gene in vivo* can lead to a wide range of cancers. Therefore, *c-myc gene* is believed to be an oncogene. It is found that the sequence d(GGGGAGGGTGGGGAGGGTGGGGAAGG) appears in the 115 to 142 base-pairs upstream of the TSS, where some secondary structures can be formed without wrapping around histone proteins because this region is highly sensitive to nucleases. *In vitro* studies showed that a family of polymorphic G-quadruplexes can be folded through that G-rich strand. Hurley and co-workers demonstrated a model to describe the mechanism, by which the formation of G-quadruplex in promoter region regulating the expression of oncogene *c-myc* is explained.(118) As shown in Figure 1.13, the formation of the G-quadruplex structure in the promoter region of oncogene *c-myc* can affect the level and nature of transcription from this gene. Generally, formation of the G-quadruplex may block the transcription machinery. Although it is also possible that the formation of G-quadruplex could be an activating domain because the accessibility of the other strand leads to increased transcriptional activity, only inhibition effect was found in other genes such as *Bcl-2*, *c-kit*, *VEGF* and *Ret*.(95)

1.4.6 G-rich Sequences in Genome: Duplex-Quadruplex Competition

Except 3' overhang in telomere, G-rich sequences wildly exist in the duplex regions of genomic DNA, where the formation of G-quadruplex was blocked by not only its complementary strands but also the adjacent duplex regions. It therefore becomes more and more important to study the competition mechanism between

duplex and G-quadruplex. It has been reported that molecular crowded conditions caused by PEG facilitate the formation of G-quadruplex because of its significant G-quadruplex stabilization and duplex destabilization.(119-123) Unfortunately, those investigations were only performed using oligonucleotides as substrates and the experiments were conducted by heating the DNA to denaturation temperature and cooling down. Such processes cannot occur *in vivo*.

DNA Supercoiling, on the other hand, can change the structures of double helix by unwinding or overwinding DNA double strands. In our studies, some supercoiled DNA molecules were engineered. The formation pattern and structure properties of G-quadruplexes were studied in the negative supercoiled DNA while the possibility of disintegration of G-quadruplex and cruciform structures within positive supercoiled DNA is discussed as well. It is our hope that competition mechanism between duplex and G-quadruplex was preliminarily established from the point of DNA supercoiling, which could benefit our understanding of the roles of non-B DNA structures in some important biological processes.

Chapter 2

DNA gyrase-driven generation of a G-quadruplex from plasmid DNA

2.1 Introduction

As we discussed in chapter 1, double-helical model and B-form DNA can describe the most common structure adopted by DNA in various vital cellular processes. In double-helical model, pyrimidine bases can only form hydrogen bonds with purine bases (adenine-thymine, guanine-cytosine) by forming two or three hydrogen bonds.(5,6) Structurally, a negative charged sugar-phosphate backbone is on the outside, which can make DNA hydrophilic. On the other hand, all the important atoms of bases are protected from chemical damage by the environment by forming a large hydrophobic interior.(124-127) B-form DNA with double-helical conformation is a kind of structure which has a right-handed double helix with a major groove and a minor groove. However, particular DNA sequences (such as G-rich regions) under certain conditions can also form alternative conformations of DNA due to the different base pair arrangements, which are categorized as non-B DNA structures.(8,128-131)

Among the non-B DNA structures, G-quadruplex is a four-stranded DNA structure composed of two or more stacks of G-quartets which are stabilized by the formation of planar arrays of four hydrogen-bonded guanines.(10,14,82,91,132) Motifs for the formation of G-quadruplex DNA structures are widely dispersed in

prokaryotic and eukaryotic cells and its unique spatial arrangement as well as its great biological significance have received considerable attention in the past few years. For example, it has been estimated that there are a million guanine-rich sites in eukaryotic cells that have the potential for forming G-quadruplexes, which exhibits possible prevalence of this non-B structure *in vivo*.(95,99,100) In addition, it has been reported in the past that within the promoter region of *c-myc* gene, G-quadruplexes can serve as a transcriptional repressor element for the gene expression.(97,98)

The spatial organization of G-quadruplex structure can be readily constructed from a single-stranded segment of DNA such as G-quadruplexes in telomere regions, which has been proved through fluorescence labelling *in vivo*.(13,80,133) On the other hand, G-quadruplex formation from a long perfectly matched duplex DNA cannot proceed in a spontaneous manner and it need to compete with the duplex that is normally generated with the complementary cytosine-rich strand under physiological conditions.(85,95,134,135) However, G-quadruplex structures could possess their melting points as high as 90 °C under physiological conditions,(136) which indicated that those table structural entity have the potential to compete with duplex structures *in vivo* as shown in Figure 2.1.

Figure 2.1 Schematic illustrations of competitive mechanism between G-quadruplex and duplex.

As discussed in section 1.2.3, the B-form double helical conformation is also believed to be a thermodynamically stable structural entity. We therefore speculate that formation of G-quadruplex will require the weakening of Watson-Crick interactions in its duplex DNA precursor in the first place (96,97,119,121,137) and an additional driving force is needed for lessening guanine-rich duplex DNA structures in order to facilitate the generation of G-quadruplex from duplex DNA both *in vitro* and *in vivo*. (3,97)

Figure 2.2 Schematic illustrations of our general strategy to exam the possibility of formation G-quadruplex in negatively supercoiled DNA caused by DNA gyrase.

Besides its likely emergence in the genomic and telomeric regions of linear DNA in the eukaryotic cells (90,138-140), the sequences of G-quadruplex formation (G-rich sequences) are known to be prevalent in the circular structures of DNA in the prokaryotes(141,142). Different from the way that eukaryotic cells mainly use topoisomerases and histone proteins to control their topological features of DNA (143), prokaryotes, on the other hand, utilize their exclusively owned DNA gyrase to

manipulate the formation of supercoiled structures in their circular chromosomal and plasmid DNA (144,145). DNA gyrase, on the other hand, is an essential bacterial enzyme which possesses the unique ability to introduce negative supercoils into a DNA circle using the free energy from ATP hydrolysis.(145-147) Since negative supercoiling weakens Watson-Crick base pairings when it is introduced by DNA gyrase (1,146), we have examined the possible correlation between the action of DNA gyrase and generation of G-quadruplex as shown in Figure 2.2. It has been demonstrated that the action of DNA gyrase can readily drive the generation of G-quadruplex structures from perfectly matched guanine-rich circular duplex DNAs under physiological-like conditions. In addition, our studies showed that the G-quadruplexes generated in circular DNAs can be readily verifiable and analyzable through using atomic force microscopy and its associated software (148-151), as well as through electrophoretic analyses and enzymatic assays.

2.2 Sequence Design of the Circular DNA with G-quadruplex

2.2.1 The Strategy to Construct a Mini-plasmid DNA with Circular Backbone

According to our previous reports,(149) duplex circular DNA can be prepared from linear DNA precursor which have two identical cohesive ends. Those linear DNA can be obtained by the reaction of endonuclease SacI digestion. Polymerase Chain Reaction (PCR) can produce the linear DNA with two identical restriction enzyme cutting sites of SacI based on particular design. To achieve acceptable final

circularization yield, the length of linear DNA precursor was designed to be around 500 bp, which is the optimal length of linear DNA for circularization and has been proved by us in the past.(152-154)

2.2.2 The Strategy to Construct a Circular DNA with guanine-rich segment

Figure 2.3 Diagrammatic illustration of synthetic route toward Circular DNA 1.

It has been known that both chromosomal and plasmid DNAs in the prokaryotes are circular in their backbones. More importantly, all the topological behavioural must be discussed in a circular system. A Circular DNA with one G-rich segment was designed and synthesized in our studies. Specifically, a 573 base-paired circular DNA (Circular DNA 1) that contains a single potential G-quadruplex-forming site was

designed at first. To introduce one guanine-rich segment d(TTAGGG)₄ into Circular DNA 1, a particularly designed oligonucleotide (ssODN-1, see section 2.3.1 for detail sequence information) was used as one of the two primers in the initial stage of Polymerase Chain Reactions (PCR) on a DNA template of X2420G. As shown in Figure 2.3, the first step of PCR amplification (Step 1 in Figure 2.3) can generate a 601 base-paired linear duplex DNA that upheld a single guanine-rich segment adjacent to one terminus of its duplex sequence, which was named as Linear DNA 1 (see Table 2.1).

Table 2.1 Double-stranded sequence of Linear DNA 1

5'	CCGAGCTCAGGATCCGGATGATCCCTAACCCTAACCCTAACCCTAACCAGTCCGTAATACGACTCACTTAAGGCCTTGACTAGAGGGTAC	
3'	GGCTCGAGTCCTAGGCCTACTAGGGATTGGGATTGGGATTGGGATTGGTCAGGCATTATGCTGAGTGAATCCGGAAGTATCTCCCATG	
	CAACCTAGGTATCTAGAACCAGTCTCGAGCCATAACTTCGTATAGCATACATTATACGAAGTTATATAAGCTGTCAAACATGAGAATTCCTTG	
	GTTGGATCCATAGATCTTGCCAGAGCTCGGTATTGAAGCATATCGTATGTAATATGCTTCAATATATTCGACAGTTTGTACTCTTAAGAAC	
	TTATAGGTTAATGTCATGATAAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTATTTTT	
	AATATCCAATTACAGTACTATTATTACCAAAGAATCTGCAGTCCACCGTGAAAAGCCCTTTACACGCGCCTTGGGGATAAACAAATAAAAA	
	CTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAAC	
	GATTTATGTAAGTTTATACATAGGCGAGTACTCTGTTATTGGGACTATTACGAAGTTATTATAACTTTTTCCTTCTCATACTCATAAGTTG	
	ATTTCCGTGTCGCCCTTATTCCCTTTTTTGGCGCATTTTGCCTTCTGTTTTTGGCTCACCCAGAAACGCTGGTGAAAGTAAAGATGCTGAA	
	TAAAGGCACAGCGGGAATAAGGGAAAAACGCCGTAAACGGAAGGACAAAAACGAGTGGGTCTTTGCGACCACTTTTCATTTTCTACGACTT	
	GATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGTTAAGCTTTTTGCACAACATGGGGGATCATGTAACCTCG	
	CTAGTCAACCCACGTGCTACCCAATGTAGCTTGACCTAGAGTTGTCGCCATTCAATTCGAAAAACGTGTTGTACCCCTAGTACATTGAGC	
	CCTTGATCGAAGGAGAGAAGAGCTGGAGCTCAATGAAGCCATACCAAACGA	3'
	GGAAGTACGTTCTCTCTTCTCGACCTCGAGTTACTTCGGTATGGTTTGCT	5'

In addition, two restriction enzyme cutting sites were designed in the ends of its linear structure. The subsequent endonuclease SacI cutting (Step 2 in Figure 2.3) can create two cohesive ends on Linear DNA 1 to give another linear DNA named as Linear DNA 2 (see Table 2.2). The two cohesive ends can paired each other within one DNA molecule to form a intermediate randomly (Step 3 in Figure 2.3). The following circularization reaction (Step 4 in Figure 2.3) was catalyzed by T4 DNA

ligase, which can catalyze the formation of a phosphodiester bond between 5' phosphate group and 3' hydroxyl termini in duplex DNA. After those four steps of reaction, a circular DNA (Circular DNA 1, see Table 2.3) can be obtained.

Table 2.2 Double-stranded sequence of Linear DNA 2.

```

5' CAGGATCCGGATGATCCCTAACCCCTAACCCCTAACCCCTAACCCAGTCCGTAATACGACTCACTTAAGGCCTTGACTAGAGGGTAC
3' TCGAGTCCTAGGCCTACTAGGGATTGGGATTGGGATTGGGATTGGTCAGGCATTATGCTGAGTGAATCCGGAACCTGATCTCCCATG

CAACCTAGGTATCTAGAACCGGTCTCGAGCCATAACTTCGTATAGCATACATTATACGAAGTTATATAAGCTGTCAAACATGAGAATTCTTG
GTTGGATCCATAGATCTTGGCCAGAGCTCGGTATTGAAGCATATCGTATGTAATATGCTTCAATATATTCGACAGTTTGTACTCTTAAGAAC

TTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTT
AATATCCAATTACAGTACTATTATTACCAAAGAATCTGCAGTCCACCGTGAAAAGCCCCTTTACACGCGCCTTGGGGATAAACAAATAAAAA

CTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAAC
GATTTATGTAAGTTTATACATAGGCGAGTACTCTGTTATTGGGACTATTTACGAAGTTATTATAACTTTTTCTCTCATACTCATAAGTTG

ATTTCCGTGTCGCCCTTATTCCTTTTTTTCGGGCATTTTGCCTTCCTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAA
TAAAGGCACAGCGGGAATAAGGGAAAAACGCCGTAAAACGGAAGGACAAAAACGAGTGGGTCTTTGCGACCACCTTCATTTTCTACGACTT

GATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGTTAAGCTTTTTGCACAACATGGGGGATCATGTAACCTG
CTAGTCAACCCACGTGCTCACCCAATGTAGCTTGACCTAGAGTTGTGCGCCATTCAATTCGAAAAACGTGTTGTACCCCCCTAGTACATTGAGC

CCTTGATCGAAGGAGAGAAGAGCTGGAGCT 3'
GGAAC TAGCTTCCTCTCTTCTCGACC 5'

```

Table 2.3 Double-stranded sequence of Circular DNA 1.

```

< 5' GAGCTCAGGATCCGGATGATCCCTAACCCCTAACCCCTAACCCCTAACCCAGTCCGTAATACGACTCACTTAAGGCCTTGACTAGAGGGTAC
3' CTGAGTCCTAGGCCTACTAGGGATTGGGATTGGGATTGGGATTGGTCAGGCATTATGCTGAGTGAATCCGGAACCTGATCTCCCATG

CAACCTAGGTATCTAGAACCGGTCTCGAGCCATAACTTCGTATAGCATACATTATACGAAGTTATATAAGCTGTCAAACATGAGAATTCTTG
GTTGGATCCATAGATCTTGGCCAGAGCTCGGTATTGAAGCATATCGTATGTAATATGCTTCAATATATTCGACAGTTTGTACTCTTAAGAAC

TTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTT
AATATCCAATTACAGTACTATTATTACCAAAGAATCTGCAGTCCACCGTGAAAAGCCCCTTTACACGCGCCTTGGGGATAAACAAATAAAAA

CTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAAC
GATTTATGTAAGTTTATACATAGGCGAGTACTCTGTTATTGGGACTATTTACGAAGTTATTATAACTTTTTCTCTCATACTCATAAGTTG

ATTTCCGTGTCGCCCTTATTCCTTTTTTTCGGGCATTTTGCCTTCCTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAA
TAAAGGCACAGCGGGAATAAGGGAAAAACGCCGTAAAACGGAAGGACAAAAACGAGTGGGTCTTTGCGACCACCTTCATTTTCTACGACTT

GATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGTTAAGCTTTTTGCACAACATGGGGGATCATGTAACCTG
CTAGTCAACCCACGTGCTCACCCAATGTAGCTTGACCTAGAGTTGTGCGCCATTCAATTCGAAAAACGTGTTGTACCCCCCTAGTACATTGAGC

CCTTGATCGAAGGAGAGAAGAGCTG 3'
GGAAC TAGCTTCCTCTCTTCTCGAC 5' >

```

Notes: (1) This DNA is a circular DNA; and

(2) < and > stand for the termini that are covalently connected.

2.3 Materials and Methods

2.3.1 Vectors, Oligonucleotide, Enzymes and Chemicals

Most of the Vectors, Oligonucleotide, Enzymes and Chemicals used in this research were listed as shown as follows (see Table 2.4). Items that are not in the list were obtained from Sigma-Aldrich with analytical grade or molecular biology grade.

Table 2.4 Vectors, Oligonucleotide, Enzymes and Chemicals used in this research.

Item(s)	Supplier(s)	Item(s)	Supplier(s)
Vector DNA (X2420G)	Generay Biotech (Shanghai, China)	Oligodeoxyribonucleotides (primers)	Sigma-Proligo (Singapore)
DNA ladder (100 bp)	Fermentas (Singapore)	DNA ladder (1 Kb)	New England Biolabs (Ipswich, MA, US)
QIAquick PCR purification kit	Qiagen (Singapore)	QIAquick Gel Extraction Kit	Qiagen (Singapore)
<i>Taq</i> Polymerase	New England Biolabs (Ipswich, MA, US)	SacI endonuclease	New England Biolabs (Ipswich, MA, US)
T4 DNA ligase	New England Biolabs (Ipswich, MA, US)	BAL 31 exonuclease	New England Biolabs (Ipswich, MA, US)

DNA gyrase	New England Biolabs (Ipswich, MA, US)	T7 Endonuclease I	New England Biolabs (Ipswich, MA, US)
Endonucleases Nt.BsmAI	New England Biolabs (Ipswich, MA, US)	Ethidium bromide	Research Biolabs (Singapore)
Mini Prep Cell	Bio-Rad (Hercules, CA, US)	Biological purity water	1st Base Pte. Ltd (Singapore)
TAE, TBE, TRIS	1st Base Pte. Ltd (Singapore)	Agarose	Invitrogen (Carlsbad, CA, US)

2.3.2 Synthesis of Linear DNA 1 and Linear DNA 3 through Polymerase Chain Reaction

Polymerase Chain Reaction (PCR) is a common and indispensable technique used in medical and biological research labs for a variety of applications. In most cases, PCR is a method of *in vitro* DNA synthesis relying on thermal cycling, which is powerful and sensitive. Almost all PCR applications employ a heat-stable DNA polymerase such as *Taq* polymerase.

Table 2.5 Nucleotide sequences of primers used in polymerase chain reactions.

Name of DNA	Nucleotide sequence
ssODN-1	5'- CCGAGCTCAGGATCCGGATGATCCCTAACCC

	TAACCCTAACCCTAACCAGTCCGTAATACGACTCAC-3'
ssODN-2	5'- TCGTTTGGTATGGCTTCATT -3'
ssODN-3	5'- CCGAGCTCAGGATCCGGATGATGGATGTGGA GTTGATGGTGGATGTCCAGTCCGTAATACGACTCAC-3'

In order to synthesize linear precursor for the following DNA circularization, Linear DNA 1 and Linear DNA 3 were amplified by PCR reactions. The sequences of the forward and reverse primers were shown in Table 2.5. The ssODN-1 and ssODN-2 are forward and reverse primers for Linear DNA 1 while the ssODN-3 and primer 2 are forward and reverse primers for Linear DNA 3. PCR amplification was conducted based on the standard procedures reported in literature(155) using a vector DNA of X2420G as the template.

A reaction mixture containing 5 ng vector DNA (X2420G), 0.25 μ M forward primer (ssODN-1 or ssODN-3), 0.25 μ M reverse primer (ssODN-2), 200 μ M dNTP, 1.5 U *Taq* polymerase in a total volume of 50 μ l reaction buffer (20 mM Tris-HCl, 10 mM $(\text{NH}_4)_2\text{SO}_4$, 10 mM KCl, 2 mM MgSO_4 , 0.1% Triton X-100, pH 8.8 @ 25 $^{\circ}\text{C}$) was put into a thermal cycling machine. The program for the PCR can be described as follows:

Step 1: Denaturation of the template at 95 $^{\circ}\text{C}$ for 120 sec;

Step 2: Primers annealing with target sites at 56 $^{\circ}\text{C}$ for 30 sec;

Step 3: Elongating target strands at 72 $^{\circ}\text{C}$ for 30 sec;

Step 4: Turning back to Step 2 for 30 times;

Step 5: Keeping the reaction temperature at 72 °C for 15 min to allow complete elongation of all DNA products;

Step 6: Cooling to 12 °C and finishing the reaction.

The products of polymerase chain reaction were further analyzed using agarose electrophoresis (1.5%) and purified using QIAquick PCR purification kit before the next steps.

2.3.3 Synthesis of Linear DNA 2 and Linear DNA 4 with Two Identical Cohesive Ends through SacI Digestion

As discussed in section 2.2, circularization of the duplex DNA from linear DNA needs two identical cohesive ends. Since two same restriction enzyme cutting sites were designed in the ends of Linear DNA 1 and Linear DNA 3, SacI endonuclease was used. The reaction was then conducted by incubation of a solution that contained 10 mM Bis-Tris-propane-HCl, 10 mM MgCl₂, 1 mM dithiothreitol, Linear DNA 1 (500 ng) and 10 U SacI at 37 °C for 1 hr. The products of SacI digestion were further analyzed using agarose electrophoresis (1.5%) and purified using QIAquick PCR purification kit before the next steps.

2.3.4 Circularization Reaction Catalyzed by T4 Ligase

Since two identical cohesive ends were obtained by digestion of endonuclease Sac I in Linear DNA 2 and Linear DNA 4, Circular DNA 1 and Circular DNA 4 were prepared by circularization reaction catalyzed by T4 DNA Ligase as described as

follows: A 50 μ l solution containing 10 mM $MgCl_2$, 50 mM Tris-HCl, 10 mM dithiothreitol, 1 mM ATP, 500 to 1000 ng of Linear DNA 2 or Linear DNA 4 and 5 U T4 DNA ligase was incubated at 16 $^{\circ}C$ for 12 hrs. The obtained reaction mixtures were further analyzed using electrophoretic analysis (1.5% agarose) and purified using QIAquick PCR purification kit before the next steps.

2.3.5 Removing Linear DNA Products from Ligase Reaction Mixture Using Nuclease BAL-31 Exonuclease

Since the above mentioned ligase reaction can also produce some by-products such as linear dimers, trimers and some circular DNA with nicked sites, removing those by-products from reaction mixture is necessary. Nuclease BAL-31 is a exonuclease, which can degrade both 5' and 3' ends of duplex DNA without generating internal scissions. The enzyme is also a highly specific single-stranded endonuclease which cleaves at nicks, gaps and single-stranded regions of duplex DNA. The ligase reaction mixture was then treated by nuclease BAL-31 in order to remove the linear dimers or trimer DNA from the reaction mixture. In addition, the treatment by nuclease BAL-31 can also further confirm that no nicks, gaps and single-stranded regions is in the backbone of obtained circular DNA. The nuclease BAL-31 degradation reaction can be conducted as follows: 500 ng reaction products of ligase reactoin, 2 U exonuclease BAL-31 and 1X Nuclease BAL-31 Reaction Buffer (20 mM Tris-HCl, 600 mM NaCl, 12 mM $CaCl_2$, 12 mM $MgCl_2$, 1 mM EDTA) was incubated in a total volume of 50 μ l at 30 $^{\circ}C$ for 10 hrs. The obtained circular DNA products

were further analyzed using electrophoretic analysis (1.5% agarose), purified using QIAquick PCR purification kit and immobilized on mica for AFM examination.

2.3.6 Introducing Negative Supercoils into DNA Circles by DNA Gyrase under Physiological-like Conditions

DNA gyrase is an essential bacterial enzyme and it can introduce negative supercoils into a DNA circle using the free energy from ATP hydrolysis. G-quadruplex-containing negative supercoil of DNA formed from Circular DNA 1 by the action of DNA gyrase under a physiological concentration of potassium ions (150 mM KCl). This DNA was obtained through incubation of a solution (50 μ l) that contained 35 mM Tris-HCl, 150 mM KCl (a physiological concentration of potassium ions), 4 mM NaCl, 4 mM MgCl₂, 2 mM DTT, 1.75 mM ATP, 5 mM spermidine, 0.1 mg/ml BSA, 500 ng Circular DNA 1 and 2 U DNA gyrase at 37 °C for 1 hr. The obtained products were further analyzed using electrophoretic analysis (1.5% agarose) and immobilized on mica for AFM examination.

2.3.7 Removing DNA supercoils by Nicking Endonucleases Nt.BsmAI

Because the spatial compactness of the resultant supercoiled DNA could make it difficult to verify the G-quadruplex with duplex along the backbone of DNA molecules using AFM, nicking endonucleases Nt.BsmAI was used to remove the supercoils from Circular DNA 2. The reaction is performed as described as follows: 200 ng negative supercoiled circular DNA, 1 U nicking endonucleases Nt.BsmAI and 1X Nt.BsmAI Reaction Buffer (20 mM Tris-acetate, 50 mM potassium acetate, 10

mM Magnesium Acetate, 1 mM Dithiothreitol) was incubated in a total volume of 50 μ l at 37 °C for 2 hrs. The obtained products were further analyzed using electrophoretic analysis (1.5% agarose) and immobilized on mica for AFM examination.

2.3.8 AFM Studies of Obtained Circular DNA

Atomic Force Microscope (AFM) is a powerful tool to determine some certain subtle alternations in DNA topological features.(156-159) Not only can the two-dimensional but also 3-D topological information be obtained through the AFM examination of DNA. Before scanning, DNA molecules need to be attached on a substrate with atomic level smooth surface. Crystal mica possesses such character and is commonly used as the substrate for AFM examination. Since both mica and DNA backbone are negatively charged, the modification of mica surface to be positively charged is necessary. The mica modification, sample preparation, AFM examination and final data processing are shown as follows:

Procedures of Mica modification and sample preparation: All micas were modified on their surfaces with (3-aminopropyl)triethoxysilane(150) before use. The resulted modified mica is the APS-mica and DNA samples can be attached to its surface without the help of divalent cation. DNA samples for AFM examination were prepared into solutions initially that contained 20 mM Tris-HCl (pH = 7) and 0.1 to 0.01 μ g/ml DNA. 5 μ l to 10 μ l of the obtained DNA solutions were placed next in the middle of the newly prepared APS-mica plates ($\sim 1 \times 1$ cm²), which were further kept

at room temperature for 5 minutes. The surfaces of the APS-mica plates bound by DNA were then rinsed for 3 times using distilled water.

Procedures of AFM examination: AFM images of DNA molecules on the APS-mica plates were obtained in Tapping ModeTM on a MultimodeTM AFM (Veeco, Santa Barbara, CA) in connection with a Nanoscope VTM controller. Antimony (n) doped Si cantilevers with nominal spring constants between 20 and 80 N/m were selected. Scan frequency was 1.9 Hz per line and the modulation amplitude was in a nanometer range.

Procedures of final data processing: All DNA sample determinations were carried out in air at room temperature. The observed shapes were significantly different from anything seen on pure duplex DNA. As a result, all of these structures were included in the dataset. Since variations in the imaging surface and/or kinks in the circular DNA, small raised structures (blobs) were occasionally seen on pure duplex DNA. To distinguish the newly formed non-B structures from the features occasionally found on the pure duplex DNA, a criterion was set according to the previous studies.⁽¹⁶⁰⁾ The normal height and the peak height were determined for 20 duplex DNA molecules. The mean of normal height was 0.51 ± 0.01 nm, and the mean of peak height was 0.67 ± 0.02 nm, with a highest absolute value of 0.83 nm. Consequently, any blob < 0.9 nm in height was excluded from the dataset and any blob > 1 nm was included. The height measurements were taken across the middle of each blob. Frequency distributions of lengths (in nm) of DNA were obtained by detecting the circumference along the

backbone of circular DNA, which were measured by drawing a series of very short lines along the DNA contour and summing the lengths.(160)

2.3.9 Reaction of T7 endonuclease with Non-B DNA

T7 endonuclease I is a type of endonuclease that recognizes and cleaves non-perfectly matched DNA, cruciform DNA structures, Holliday structures or junctions, heteroduplex DNA and more slowly, nicked double-stranded DNA. With the purpose to exam that the G-quadruplex structures are present in the backbone of circular DNA molecules, T7 endonuclease I was used in our studies. The reaction is performed as decribed as follows: a solution that contained 50 mM NaCl, 10 mM Tris-HCl, 10 mM MgCl₂, 1 mM DTT, 200 ng circular DNA and 5 U T7 Endonuclease I was incubated at 37 °C for 1 hr.

2.4 Results and Discussion

2.4.1 Synthesis and Structure Examination of G-rich-containing Circular DNA 1

Since the topological properties need to be studied in a circular DNA which is the common structure of both chromosomal and plasmid DNAs in the prokaryotes, a 573 base-paired mini-plasmid DNA (Circular DNA 1) that contains a single potential G-quadruplex-forming site was designed and synthesized at first in our studies (Figure 2.3). Human telomere consists of repetitive stretches of TTAGGG at the end of

chromosomes of human cells, which can form tetraplex assembly in the presence of K^+ or Na^+ .(90,139) With the aim to examine whether a duplex circular DNA containing human telomeric sequence $d(TTAGGG)_4$ can form G-quadruplex with assistance of DNA gyrase under physiological conditions of K^+ , Circular DNA 1 was designed and synthesized in our studies at first.

Figure 2.4 Electrophoretic analysis of intermediate DNA molecules generated during the synthesis of Circular DNA 1. Lane M: molecular weight markers; Lane 1: Linear DNA 1 generated through PCR amplification (Step 1); Lane 2: Linear DNA 2 with its cohesive ends created through using *SacI* (Step 2); Lane 3: crude product of Circular DNA 1 produced through reaction of Linear DNA 2 and T4 DNA ligase (Step 3 and Step 4); Lane 4: pure Circular DNA 2 obtained through hydrolysis of crude product of Circular DNA 1 by Nuclease BAL-31.

Polymerase Chain Reaction was carried out at first during our investigations in which a plasmid vector (X2420G) served as the template and specific primers (ssODN-1 and ssODN-2, see Table 2.5) were used as the forward primer and reverse primer to generate a duplex linear DNA (Linear DNA 1, Step 1 in Figure 2.3) that contains 601 base pairs in length (Lane 1 in Figure 2.4). Linear DNA 1 was accordingly examined by the electrophoresis as shown in Figure 2.4. The ssODN-1

was designed to contain cytosine-rich segment from which guanine-rich region of DNA duplex could be formed in the later stages of the synthetic process. In addition, there were two restriction endonuclease SacI cutting sites in the sequence Linear DNA 1 as shown as section 2.3. Subsequently, two cohesive ends on Linear DNA 1 were created using SacI endonuclease and Linear DNA 2 was formed (Lane 2 in Figure 2.4). The DNA in Lane 1 in Figure 2.4 was obtained from PCR amplification using particular pair of primers (Table 2.5), which possessed two blunt ends. The DNA in Lane 2 in Figure 2.4, on the other hand, was hydrolysis product of the DNA in Lane 1 by SacI (restriction endonucleases), which was shorter than the DNA in Lane 1 by 26 base pairs and contained two cohesive ends.

The following ligase reaction brought about desired 573 base-paired Circular DNA 1 in its crude form (Lane 3 in Figure 2.4). In this ligase reaction, higher yield of the circularization can be achieved by decreasing the concentration of linear precursor. This happens because the circularization is a intermolecular reaction, in which each linear DNA need more "space" to pair each ends in the same molecule. In addition, there is no significant effect with increasing the reaction time as shown in Table 2.6. The yield (%) of circularization is calculated based on comparison of the band density data in Lane 3 in Fig. 3A. The measurement of the densities of these bands was conducted using “Gel Documentation System (G-Box HR, Syngene, and Cambridge, UK)”, which were further analyzed using “Gene Tools Software”.

Table 2.6 Analysis of yield (%) of circularization in different concentration of substrate and reaction time.

	1	2	3	4	5	6	7	8
T4 DNA ligase (U)	5	5	5	5	5	5	5	5
Linear DNA (ng/ul)	500	125	25	5	500	125	25	5
Temperature (°C)	16	16	16	16	16	16	16	16
Reaction time (h)	2	2	2	2	12	12	12	12
Yield (%)	7.5	10.2	15.7	30.5	10.2	13.1	17.3	39.2

Formation of a circular backbone in Circular DNA 1 was further purified and confirmed using Nuclease BAL-31 (exonuclease, Lane 4 in Figure 2.4) which can remove the single-stranded by-products and nick- or gap-containing circular DNA. AFM can give the direct evidence of the topological properties of DNA molecules. The obtained Circular DNA 1 was also examined by AFM. As shown in Figure 2.5, a circular structure can be clearly observed along the backbone of each DNA molecules.

Figure 2.5 AFM images of Circular DNA 1 with its scale bar of 200 nm. (the sample used for the AFM examination was the same batch of sample as the one loaded in Lane 4 in Figure 2.4).

2.4.2 Formations of G-quadruplexes Facilitated by DNA Gyrase under Physiological Conditions of K^+ and their Conformation by Electrophoresis

Figure 2.6 Diagrammatic illustration of reactions of Circular DNA 1 upon the action of DNA gyrase and other enzymes.

It has been well established in the past that negative superhelicity could be a crucial factor for inducing the formation of G-quadruplex structures due to the intrastrand base-pairing in DNA duplex.(161,162) DNA gyrase, on the other hand, is an essential bacterial enzyme that catalyzes the ATP-dependent negative supercoiling of double-stranded closed-circular DNA. Circular DNA 1 was accordingly treated with DNA gyrase and negative supercoiled DNA circles were obtained. Figure 2.6 depicts the anticipated course of generation of G-quadruplex from Circular DNA 1 driven by the action of DNA gyrase under the physiological concentrations of cations.

Figure 2.7 Electrophoretic analysis of products of enzymatic reactions on Circular DNA 1. Lane M: molecular weight markers; Lane 1: Circular DNA 1; Lane 2: negatively supercoiled DNA formed through incubation of Circular DNA 1 with DNA gyrase under a non-physiological concentration of potassium ion (24 mM KCl, this lane serves as a control for Lane 3 of Figure 2.7); Lane 3: G-quadruplex-containing DNA (Structure 2) generated by the action of DNA gyrase under physiological concentrations of potassium ions (150 mM KCl and 4 mM NaCl at pH 7.5); Lane 4: nicked form of G-quadruplex-containing circular DNA (Structure 3); Lane 5: Circular DNA 2.

For the comparison purpose, Circular DNA 1 was incubated firstly with DNA gyrase in a buffer solution (pH 7.5) that contained 24 mM KCl, a concentration that was much lower than the physiological level of the cation (0.1-0.6 M) (163-166). As shown in Lane 2 in Figure 2.7, a band that migrated faster than the one of Circular DNA 1 (Lane 1) was observed, which signified that more compact negative supercoil of DNA was formed by the action of DNA gyrase (Structure 1 in Figure 2.6) under a non-physiological concentration of potassium ion. It is known, on the other hand, that prokaryotes commonly maintains their intracellular concentration of (1) potassium ion from 0.1 – 0.6 M (the overwhelmingly predominant cation insides prokaryotic cells),

(2) sodium ion ≤ 10 mM as well as keeps up their pH from 6.0 to 8.0 (163-166). Circular DNA 1 was consequently incubated next with DNA gyrase under certain physiological concentrations of cations (150 mM KCl and 4 mM NaCl at pH 7.5) during our examinations. As shown in Lane 3 in Figure 2.7, a band was generated that migrated slower than the one generated under a non-physiological concentrations of cations (Lane 2, Structure 1 in Figure 2.6). Since formation of G-quadruplex in DNA negative supercoil should in theory alter the compactness of the DNA structure, the observation of slower-moving band in Lane 3 (Figure 2.7) is consistent with the suggestion that non-B DNA structures were produced in the negatively supercoiled circular DNA (Structure 2 in Figure 2.6) driven by the action of DNA gyrase under physiological concentrations of cations.

2.4.3 Examining the Formations of G-quadruplexes Facilitated by DNA Gyrase under Physiological Conditions of K^+ using AFM

Figure 2.8 AFM examination of DNA with G-quadruplex-containing negative supercoils. (The DNA sample used for this AFM examination was the same batch of sample as the one loaded in Lane 3 in Figure 2.7). The structures of DNA in the AFM images were G-quadruplex-containing negative supercoils that were generated from

relaxed forms of guanine-rich segment-containing circular DNA (Circular DNA 1 in Fig. 2A) under a physiological concentration of potassium ion (150 mM KCl). G-quadruplex structures in these negatively supercoils are not identifiable using AFM owing to their spatial compactness.

To exam the topological difference between Circular DNA 1 and G-quadruplex-containing circular DNA, atomic force microscope (AFM) was used. Originally, we decided exam whether G-quadruplexes appeared in DNA samples of Lane 3 in Figure 2.7 by AFM. As shown in Figure 2.8, the spatial compactness of the DNA backbone makes it difficult to verify the presence of G-quadruplex using AFM. This happens because the formed G-quadruplex co-exists with DNA negative supercoil (Structure 2 in Figure 2.6) and it is very difficult to identify between the formed secondary structures and self-crossings caused by supercoiling. The supercoils in Structure 2 (DNA samples of Lane 3 in Figure 2.7) were accordingly relaxed next using nicking endonucleases Nt.BsmAI and further re-ligated using DNA ligase (Lane 4 and Lane 5 in Figure 2.7). The obtained DNA was accordingly named as Circular DNA 2.

Figure 2.9 AFM images of Circular DNA 2. (A) AFM image with its scale bar of 200 nm. (B) 3D AFM image of Circular DNA 2.

With aim to verify the presence of G-quadruplexes in Circular DNA 2, the same batch of sample as the one loaded in Lane 5 in Figure 2.7 was tested by AFM. As shown in Figure 2.9, sharp turns associated with brighter dots along the circular DNA backbones appeared in the AFM images of the obtained DNA molecules (Circular DNA 2), which is the indication that some non-B DNA structures were present along their circular DNA backbones.

Figure 2.10 Section analyses of AFM images of Circular DNA 2.

To further confirm the observed "sharp turns" is different from DNA duplex, subsequent section analysis by AFM associated software was conducted. The section analysis along the bright dot-crossing line (Line 1 in Figure 2.10) revealed that the

height of the section at Site 1 (1.1~1.3 nm) is ~1.8 times greater than the one on Site 2 (0.6~0.8 nm). In addition, the section at Site 1 (38 nm) is twice as wide as the one at Site 2 (19 nm). As a comparison study, additional section analysis was carried out along a line (Line 2 in Figure 2.10 that is perpendicular to Line 1, which unveiled that the widths and heights on both sides of the DNA backbones (Site 3 and Site 4) were nearly equal.

Since variations in the imaging surface and/or kinks in the circular DNA, small raised structures (blobs) were occasionally seen on pure duplex DNA. To distinguish the newly formed non-B structures from the features occasionally found on the pure duplex DNA, a criterion was set according to the previous studies as described in section 2.3. According to this method, frequency distributions of lengths of Circular DNA 1 and Circular DNA 2 were counted and showed in Figure 2.11. Because the length of plasmid DNA backbone in Circular DNA 2 should theoretically be shorter than those in Circular DNA 1 after the formation of non-B structures, these distributions clearly show that some secondary structures are formed in Circular DNA 2 (160).

Figure 2.11 Frequency distributions of the lengths (nm) of Circular DNA 1 (A) and Circular DNA 2 (B).

The observed "sharp turns" is most likely G-quadruplex because guanine-rich segment contains in Circular DNA 2 and the formation of those structures is potassium dependent. With purpose to exam the non-B structures shown above is indeed G-quadruplex, positive control in the AFM should be performed to prove that the character described above is G-quadruplex. Therefore, we constructed a G-quadruplex-containing linear DNA using reported method(149) and tested by AFM. Since the guanine-rich segment d(TTAGGG)₄ was introduced through using Polymerase Chain Reactions (PCR), Linear DNA 1 should in theory uphold a single guanine-rich segment adjacent to one terminus of its duplex sequence. On the other hand, it has been established that G-quadruplex could preferentially form and dominate over duplex structure under molecular crowding condition created by PEG 200 as a result of significant G-quadruplex stabilization and duplex destabilization.(167,168) Linear DNA 1 was accordingly incubated with 150 mM KCl and 40% PEG 200 and G-quadruplex-containing linear DNA was obtained. The obtained G-quadruplex-containing linear DNA was then analyzed by AFM. As anticipated, bright dots appeared at the end of those G-quadruplex-containing in the AFM images (Figure 2.12). The following section analysis of G-quadruplex-containing linear DNA showed that height of G-quadruplex part (1.1~1.4 nm) is ~1.8 times greater than the duplex part (0.6~0.8 nm), which is consistent with the character described in Figure 2.10.

Figure 2.12 AFM examination of G-quadruplex-containing linear DNA. [the examinations in the current section serve as a positive control to prove that the character described in Figure 2.10 is G-quadruplex]. (A) AFM examination of G-quadruplex-containing linear DNA. The structures of DNA in the AFM images were G-quadruplex-containing linear DNA that were generated from Linear DNA 1 (the same batch of DNA sample as the one loaded into Lane 1 in Figure 2.7) incubated with 150 mM KCl and 40% PEG 200 at 95°C for 5 min and cooled down to room temperature.(168) (B) Section analysis of non-B structure of an AFM image in Figure 2.12. (C) Section analysis of duplex part of an AFM image in Figure 2.12.

2.4.4 Analysis of Reaction between DNA Gyrase and Circular DNA 1 under the Concentration of K^+ of Non-physiological Condition

In order to exclude the possibility that reaction buffers and other factors caused the formation of G-quadruplex coincidentally, we carried out proper control experiments as well. The control experiments were performed in the same way as those shown in Figure 2.7 except that the concentration of K^+ was reduced from a

physiological concentration of 150 mM to a non-physiological concentration of 24 mM. The designed reaction route and following electrophoretic analysis are shown in Figure 2.13. The final DNA product was accordingly names as Circular DNA 3.

Figure 2.13 Examination of action of DNA gyrase on Circular DNA 1 under a non-physiological concentration of potassium ions. (A) Diagrammatic illustration of reactions of Circular DNA 1 upon the actions of DNA gyrase and other enzymes under non-physiological conditions. (B) Electrophoretic analysis of reaction products of Circular DNA 1 upon the actions of DNA gyrase and other enzymes under non-physiological conditions. Lane M: molecular weight markers; Lane 1: Circular DNA 1 alone; Lane 2: negative supercoil of circular DNA (Structure 1) generated under a non-physiological concentration of potassium ion (24 mM KCl); Lane 3: nicked form of circular DNA; Lane 4: relaxed Circular DNA 3.

In order to verify whether a G-quadruplex existed in the final DNA products, AFM examination was also conducted as shown in Figure 2.14. As anticipated, no "sharp turns" as shown in Figure 2.9 can be observed (Figure 2.14 A). The section

analyses on the AFM images of DNA obtained from these control studies also revealed no evidence of formation of non-B structure along the backbones of the circular DNA molecules (Figure 2.14 B). In addition, frequency distributions of lengths of Circular DNA 1 and Circular DNA 3 were measured, in which no apparent shortening of the plasmid occurred after Circular DNA 1 was treated with DNA gyrase under a non-physiological concentration of potassium ions (Figure 2.14 C). Based on the analysis shown above, the topological structures of Circular DNA 3 are the same as those of Circular DNA 1.

Figure 2.14 AFM examinations of the topological structures of Circular DNA 1 and Circular DNA 3. (A) AFM images of relaxed closed Circular DNA 3 (the DNA sample used for this AFM examination was the same batch of sample as the one

loaded in Lane 4 in Figure 2.13). (B) Section analysis of AFM image of a DNA molecule taken from Figure 2.14A. (C) – (D) Frequency distributions of the lengths (nm) of Circular DNA 1 and Circular DNA 3.

2.4.5 Confirmation of Absence of G-quadruplex Structure in Non-guanine-rich Circular DNA 5

Table 2.7 Double-stranded sequence of Circular DNA 4.

```

< 5' GAGCTCAGGATCCGGATGATCCTACACCTCAACTACCACCTACACCGTCCGTAATACGACTCACTTAAGGCCTTGACTAGAGGGTAC
3' CTCGAGTCCTAGGCCCTACTAGGATGTGGAGTTGATGGTGGATGTGGTCAGGCATTATGCTGAGTGAATCCGGAACCTGATCTCCCATG

CAACCTAGGTATCTAGAACC GGTCGAGCCATAACTTCGTATAGCATA CATTATACGAAGTTATATAAGCTGTCAAACATGAGAATTCTTG
GTTGGATCCATAGATCTTGGCCAGAGCTCGGTATTGAAGCATATCGTATGTAATATGCTTCAATATATTGACAGTTTGTACTCTTAAGAAC

TTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTT
AATATCCAATTACAGTACTATTATTACCAAGAATCTGCAGTCCACCGTAAAAGCCCCTTTACACGCGCCTTGGGGATAAACAATAAAAA

CTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAAC
GATTTATGTAAGTTTATACATAGGCGAGTACTCTGTTATTTGGGACTATTTACGAAGTTATTATAACTTTTTCCTTCTCATACTCATAAGTTG

ATTTCCGTGTCGCCCTTATTCCCTTTTTTGGGGCATTTCCTGCTGTTTTCCTCAGCAAGAACGCTGGTGAAAGTAAAGATGCTGAA
TAAAGGCACAGCGGGAATAAGGGAAAAAACGCCGTAAAACGGAAGGACAAAAACGAGTGGGTCTTTGCGACCACTTTCATTTTCTACGACTT

GATCAGTTGGGTGCACGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAGTTAAGCTTTTGCACAACATGGGGGATCATGTAACCTG
CTAGTCAACCCACGTGCTCACC AATGTAGCTTGACCTAGAGTTGTGCGCATTC AATTCGAAAAACGTGTTGTACCCCTAGTACATTGAGC

CCTTGATCGAAGGAGAGAAGAGCTG 3'
GGAAGTACGCTTCTCTCTCGAC 5'>

```

Notes: (1) This DNA is a circular DNA; and

(2) < and > stand for the termini that are covalently connected.

With the aim of confirming that the observed non-B structures shown in Figure 2.9 are indeed associated with the guanine-rich sequence of d(TTAGGG)₄ in Circular DNA 1, a different 573 base-paired circular DNA (Circular DNA 4) was designed and synthesized next during our investigations, which possesses the same length and the same nucleotide composition as Circular DNA 1. Unlike Circular DNA 1, Circular DNA 4 possesses no apparent guanine-rich segment because the nucleotide sequences of Circular DNA 4 are identical to those of Circular DNA 1 except that (TTAGGG)₄ was replaced with GGATGTGGAGTTGATGGTGGATGT (see Table 2.5 for primers

information and Table 2.7 for entire nucleotide sequences of Circular DNA 4). As a result, G-quadruplex structure is in theory not capable of being formed in Circular DNA 4 owing to the absence of guanine nucleotides aligned in a row along its sequence. The same procedures as the ones shown in Figure 2.7 were subsequently carried out in our studies except that Circular DNA 1 was replaced with Circular DNA 4. The designed reaction route and following electrophoretic analysis are shown in Figure 2.15. The final DNA product was accordingly named as Circular DNA 5. It is shown that no apparent mobility shift difference can be observed between Lane 2 and Lane 3, which indicated that the supercoiling structures generated from Circular DNA 4 may remain unchanged in spite of the variation of the concentration of potassium cation. In addition, the same mobility shift occurred between Lane 1 and Lane 5, which suggested that Circular DNA 4 may hold a same topological conformation with Circular DNA 5.

Figure 2.15 Examination of action of DNA gyrase on Circular DNA 4 under physiological concentrations of potassium ions. [The examinations in the current section (non-guanine-rich-segment-containing circular DNA) serve as control experiments for the studies shown in Figure 2.7 and 2.9 (guanine-rich-segment-containing circular DNA)]. A. Illustration of reactions of Circular DNA 4 upon the actions of DNA gyrase, nicking enzyme and DNA ligase. (B) Electrophoretic analysis of products of enzymatic reactions on Circular DNA 4. Lane M: molecular weight markers; Lane 1: Circular DNA 4 alone; Lane 2: negative supercoils of circular DNA generated under a non-physiological concentration of potassium ions (24 mM). Lane 3: negative supercoiled circular DNA (Structure 1 in Figure 2.15A) generated under a physiological concentration of potassium ions (150 mM KCl). Lane 4: nicked form of circular DNA (Structure 2 in Figure 2.15A). Lane 5: relaxed form of closed circular DNA (Circular DNA 5 in Figure 2.15A). This relaxed form of DNA was obtained through incubation of a solution (50 μ l) that contained 50 mM Tris-HCl, 10 mM MgCl₂, 1 mM ATP, 10 mM dithiothreitol, 500 ng nicked form of circular DNA (the same batch of DNA sample as the one loaded into Lane 4 in Figure 2.15B) and 20 U T4 DNA ligase at 16 °C for 8 hrs.

Similarly, AFM was used to exam whether G-quadruplex structures existed in the backbone of Circular DNA 5. Our AFM examination and further section analysis of AFM image revealed that no non-B structure was formed from the obtained final DNA products (Figure 2.16), which is consistent with the suggestion that the non-B structure observed in Circular DNA 2 (Figure 2.9) is G-quadruplex.

Figure 2.16 AFM images of the obtained relaxed form of closed circular DNA. (A) AFM images with its scale bar of 200 nm. The DNA sample for this AFM examination was the same batch of sample as the one loaded in Lane 5 in Figure 2.15B. (D) Section analysis of an AFM image in Figure 2.16A

2.4.6 Confirmation of the Existence of G-quadruplex in Circular DNA by Endonuclease

Figure 2.17 Diagrammatic illustration of the non-matched sites in G-quadruplex-containing DNA.

Because G-quadruplex structure belongs to non-B DNA structures, non-matched sites and C-rich single-stranded segments (at physiological conditions of pH values) should be contained in its molecular backbone as shown in Figure 2.17. T7 endonuclease I is a type of endonuclease that recognizes and cleaves non-perfectly matched DNA. We therefore decided to exam the generation of non-B structures driven by the action of DNA gyrase using enzymatic methods as well in our studies. The diagrammatic illustration of anticipated reactions of G-quadruplex-containing or non-G-quadruplex-containing circular DNA with T7 Endonuclease I are shown in Figure 2.18. Because of the presence of non-matched sites in G-quadruplex-containing circular DNA, this circular DNA will be linearized after the cleavage of T7 Endonuclease I as shown in Figure 2.18A. However, the perfect matched non-G-quadruplex-containing circular DNA could be keep its circular conformation after it is treated with T7 Endonuclease I as shown in Figure 2.18B.

Figure 2.18 Diagrammatic illustration of our enzymatic confirmation of presence of G-quadruplex structures in Circular DNA. (A) Illustration of anticipated reactions of

G-quadruplex-containing circular DNA with T7 Endonuclease I. (B) Illustration of anticipated reactions of perfect matched with T7 Endonuclease I.

Circular DNA 2 (G-quadruplex-containing) and Circular DNA 3 (non-G-quadruplex-containing) were accordingly treated with T7 Endonuclease I to exam the existence of non-B structure in the circular DNA and the products were tested by electrophoretic analysis. As shown in Figure 2.19A, a new band with a faster rate of mobility shift was observed, which is the product of Circular DNA 2 treated with T7 Endonuclease I. The result given in above electrophoretic analysis suggested that a DNA with linear conformation could be formed (Lane 2 in Figure 2.19A). On the other hand, no mobility shift difference can be observed between Circular DNA 3 and its products treated with T7 Endonuclease I (Lane 2 in Figure 2.19B).

Figure 2.19 Enzymatic confirmation of presence of G-quadruplex structures in Circular DNA 2. (A) Electrophoretic analysis of reactions of Circular DNA 2 and T7 Endonuclease I. Lane 1: Circular DNA 2; Lane 2: linear DNA as the cleavage product of Circular DNA 2 by T7 Endonuclease I. (B) Electrophoretic analysis of reactions of

Circular 3 and T7 Endonuclease I. Lane 1: Circular DNA 3; Lane 2: a mixture obtained from incubation of Circular DNA 3 with T7 Endonuclease I.

Further AFM examination on the DNA isolated from the band in Lane 2 in Figure 2.19 unveiled that the DNA product was indeed linear (Figure 2.20A), which indicated that Circular DNA 2 was cut by T7 Endonuclease I due to the presence of non-B structures. As a control study, the circular DNA (Circular DNA 3) obtained under a non-physiological condition was incubated with T7 Endonuclease I as well. Our further AFM examination (Figure 2.20B) revealed that Circular DNA 3 remained intact after it was incubated with T7 Endonuclease I.

Figure 2.20 AFM examination of DNA products obtained after T7 Endonuclease I cleavage. (A) AFM images of linear DNA obtained from the reactions of Circular DNA 2 with T7 Endonuclease I. (B) AFM images of Circular DNA 3 upon its incubation with T7 Endonuclease I. Scale bar indicates 200 nm

In addition, Circular DNA 5 (no guanine-rich segments) that was prepared from circular DNA 4 (see Figure 2.15) was also tested by the enzymatic method to confirm no G-quadruplex in it. As shown in Figure 2.21, a no mobility shift difference

can be found in its electrophoretic analysis and the final products showed a circular conformation in the AFM examination.

Figure 2.21 Enzymatic confirmation of absence of G-quadruplex structures in Circular DNA 5. (A) Diagrammatic illustration of anticipated reactions of non-G-quadruplex-containing Circular DNA 5 with T7 Endonuclease I. (B) Electrophoretic analysis of reactions of non-G-quadruplex-containing Circular 5 (Figure 2.15) and T7 Endonuclease I. Lane 1: Circular DNA 5 (the same batch of DNA sample as the one loaded into Lane 5 in Fig. 2.15B); Lane 2: a mixture obtained from incubation of Circular DNA 5 with T7 Endonuclease I. (C) AFM images of Circular DNA 5 upon its incubation with T7 Endonuclease I with its scale bar of 200 nm. The DNA sample for this AFM examination was the same batch of sample as the one loaded in Lane 2 in Figure 2.21B.

2.5 Conclusion

Formation of G-quadruplex in circular duplex DNA demonstrated in the current studies could have certain implications in our understanding of the structural interconversion between G-quadruplex and duplex in prokaryotic cells. Among non-B DNA structures, G-quadruplex structures have been extensively investigated because of its attractively biological effects. However, it is not clear whether G-quadruplex could form from conventional Watson-Crick duplex prokaryotic genomes to date. Since DNA gyrase as essential topoisomerase is widely dispersed in prokaryotic bacteria, our current studies implied that G-quadruplex could be formed from duplex DNA circles under near physiological conditions and the crucial factor of G-quadruplex formation could be negative supercoiling affiliated with DNA gyrase.

In conclusion, the results of our (1) gel mobility shift analysis, (2) AFM examination, (3) endonuclease assays as well as (4) our control studies signify that the action of DNA gyrase can readily drive the generation of G-quadruplex from guanine-rich segment-containing plasmid DNA under the intracellular ion concentrations of prokaryotic cells. Since DNA gyrase is a prokaryote-exclusively owned enzyme that is absent in eukaryotes, the outcomes of our current investigations could suggest that prokaryotic cells might utilize this topological enzyme to regulate the generation of G-quadruplex to comply with their subsequent cellular functions.

Chapter 3

Disintegration of cruciform and G-quadruplex structures during the course of helicase-dependent amplification (HDA)

3.1 Introduction

DNA damages refer commonly to chemical changes of DNA structures in the prokaryotic and eukaryotic genome. DNA *in vivo* undergoes damage spontaneously from hydrolysis and deamination of DNA molecules. In addition, DNA is often damaged by alkylation, oxidation and radiation (Gamma and X-rays) that can cause double-strand breaks and are particularly hazardous. Some damages, such as thymine dimer, nick or breaks in the DNA backbone can create impediments to replication or transcription. Other damages caused by the structural changes of bases, however, have no effect on replication, but bring about mispairing which in turn can be converted to mutation.(169-171)

The damaged DNA molecules will lose the ability to resume their original double helical B-forms.(172-174) To detect DNA damages and repair the lesions by activating DNA reparation machines, all organisms have evolved delicate DNA repairing mechanisms. For example, direct reversal of DNA damage by photoreactivation is an error-free repair mechanism, which can form monomer from thymine dimers by DNA photolyases in the presence of visible light. Alkyltransferase

removes the methyl group from the methylated O⁶-methylguanine and the methyl group is transferred to the protein itself as shown in Figure 3.1.

Figure 3.1 Diagrammatic illustration of the DNA repair mechanisms by photoreactivation and alkyltransferase.

Besides these well-recognized chemical damages to DNA, the observations from the abovementioned investigations in chapter 2 as well as from other *in vitro* studies illustrate that many non-B DNA conformations are factually stable structural entities (e.g. formations of G-quadruplex and cruciform) that are not readily disintegrated to form a original B-form double helical structure under physiological conditions owing to their high thermodynamically stabilities. Those physical alterations of canonical B-form of DNA occur prevalently in organisms that serve as signals for specified cellular events. It has been demonstrated in the past years that G-quadruplex and cruciform structures are present *in vivo* and play important roles in various cellular processes such as replication, transcription and recombination. G-quadruplex, for

example, has been proved to exist *in vivo* and it can be formed in the regions of promoter of some cancer genes and act as the "molecular switch" for the expression of those genes.(118) In addition, The cruciform structures of DNA are believed to form at or near replication origins of some eukaryotic cells and serve as recognition signals for DNA replication.

In theory, once its service as a cellular signal in a living organism is completed, a non-B DNA structure should be broken down instantaneously into their canonical B forms in order to continue their subsequent cellular functions. However, the pathways and driving forces that lead to the disintegration of non-B DNA structures in cells have not yet been well investigated thus far. Comparing with the chemical damages on DNA in which covalent bonds are either broken down or newly formed (e.g. UV-mediated dimerization of pyrimidines)(175), no covalent bond is broken or newly formed in the course of the formation of non-B structures from canonical B-form of DNA. As a result, living organisms may not use the same mechanisms for repairing chemically damaged DNA. Particularly, for the cases of DNA cruciform structures in which both opposite strands take shape of non-B DNA structures(176,177), there is no intact single strand left to serve as a template if these physically altered DNA structures had presumably undergone "single-stranded DNA repairing mechanisms"(178,179). Therefore, an economical pathway for repairing their physically altered DNA structures may be chosen by some living organisms.

The DNA stored in mesophilic prokaryotic cells and all eukaryotic cells exists in the forms of negative supercoils, which are produced either by DNA gyrase or by the

combining actions of histone proteins and topoisomerases. Contrast to these negative supercoil-storing mesophiles, on the other hand, all forms of hyperthermophiles possess reverse gyrase inside their cellular structures that introduces positive supercoils to their DNA. Structurally, right-handed negative supercoiling unwinds DNA double helix, which facilitates the denaturation bubbles during the replication and transcription. However, positive supercoiling is right-handed and it overwinds the helical turns of DNA as shown in Figure 3.2.

Figure 3.2 Diagrammatic illustration of positive and negative supercoiling.

Since the discovery of reverse gyrase and its catalysis for the formation of positive DNA supercoils in 1972, various studies have been carried out in order to understand the innate roles of the positive-supercoil-introducing enzyme in hyperthermophiles. From these earlier studies, it is known now that reverse gyrase is composed of the structural domains of both helicase and type I topoisomerase that enable this hyperthermophilic enzyme to introduce positive supercoiling to its substrate DNA. In addition, it has become apparent that presence of reverse gyrase is a prerequisite for hyperthermophiles to flourish even though hyperthermophilic

organisms can “barely sustain life” in the absence of reverse gyrase gene. Furthermore, reverse gyrase has been shown to be capable of coating the nicked sites arising in DNA strands and to contribute to the stabilization of DNA at high temperatures. Even though a great deal of information about reverse gyrase and its cellular functions have been obtained in the past thirty years, it has not yet been clearly understood up to now as to what the fundamental principles and benefits are that have driven mesophiles and hyperthermophiles to evolve and adopt opposite signed (positive and negative) supercoils of DNA for the safekeeping of their genomic DNA. Moreover, the current available information about reverse gyrase has not been sufficient for experts to draw a conclusion about why reverse gyrase can be found in all hyperthermophiles and what exactly the innate roles of reverse gyrase are in supporting hyperthermophilic life. Unfortunately, examination of the properties of positive DNA supercoils under the conditions that mimic those in the hot springs and hydrothermal vents as well as comparison of them with the ones of negative DNA supercoils and relaxed forms of DNA have not yet been carried out systematically thus far. In addition, the information about whether thermodynamically stable non-B structures of DNA (e.g. extraordinarily stable G-quadruplex) could indeed be disintegrable by reverse gyrase-based positive DNA supercoiling has not yet been known.

The common biotopes for hyperthermophiles are known to be the hot springs on the land and hydrothermal vents in the ocean, which are evidently different from those for accommodating mesophiles. Some characteristics of these two types of hot fields include high temperature, extremely low pH (pH from 0 to 9), and low salt concentration, which could in theory destabilize the unwound structures of negative

DNA supercoils in either physical or chemical manners. Consequently, use of positively supercoiled forms of DNA for the storage of their genetic information could possibly be a “brilliant” strategy adopted by hyperthermophiles for the purpose of avoiding formation of thermodynamically stable non-B structures of DNA (e.g. G-quadruplex with its melting point of 90 °C) at hyperthermophilic temperatures, preventing protonation of nucleobases and phosphodiester backbones under extremely acidic conditions (pH = ~0) and compensating the stability loss caused by low salinity. In addition, it could be possible that reverse gyrase’s positive-supercoil-introducing action has been evolved to repair the physical damages (e.g. formation of extraordinarily stable G-quadruplex) of DNA produced at hyperthermophilic temperatures as well.

The activity of DNA helicase, on the other hand, are often associated with DNA replication. One of the common characteristic of helicase and reverse gyrase is known to be that they are capable of generating positive supercoil in the DNA that they are acting on. In both prokaryotic and eukaryotic cells, helicases are known as motor proteins which can move directionally along a nucleic acid phosphodiester backbone and separate two annealed nucleic acid strands.(102,180) With the special abilities of unwinding DNA double strands, helicases in the replication process unwinds the DNA duplex where it binds and further moves forward at the replication fork that it has created. During this course of action, a positive supercoiling can be formed within the sequence ahead of the DNA replication fork.(3,71) As a result, this part of DNA backbone is forced to rotate, which will lead to the formation of a torsional stress in the whole DNA circle.

Figure 3.3 Diagrammatic illustration of HDA. 1: UvrD helicase unwind DNA template in the presence of single-stranded binding proteins (SSB) and other accessory proteins at 37 °C. 2: Primers anneals to single-stranded parts of DNA. 3: Exo-Klenow polymerase extends the primers. 4: Double-stranded DNA is separated by UvrD helicase and two semiconservative replication products is obtained. The picture is taken from the report by *M. Vincent et al* (2004). (181)

Helicase-dependent amplification (HDA), on the other hand, is an *in vitro* isothermal DNA replication technology, which utilizes a DNA helicase to generate single-stranded templates for primer hybridization and subsequent primer extension by a DNA polymerase as shown in Figure 3.3.(181,182) Because there is no topoisomerase involved in HDA, the topological problem ahead the replication fork

cannot be solved in time. Positive supercoils must be generated within the circular structure of template DNA as shown in Figure 3.4.

Figure 3.4 Schematic illustration of the topological relationships during the course of DNA replication *in vitro* within a circular DNA. Formation of positive supercoiling caused by Helicase-dependent amplification (HDA).

We accordingly choose HDA as a modeling to investigate whether the transformation of non-B structures occurs in this course of DNA replication *in vitro*. It has been reported that positive supercoiling is overwound (144,183) and some non-B structures can be repaired by positive supercoiling affiliated with nucleosome formation (79). We therefore speculate that breakdown of non-B structure and resumption of the original canonical B conformation of DNA should be able to relieve the torsional stress accumulated in the backbone of a positively supercoiled DNA, thus leading to the non-B DNA disintegration. Figure 3.5 depicts our anticipated actions of positive supercoiling on a non-B DNA structure when these two types of structures co-exist in the same duplex DNA strands.

Figure 3.5 Schematic illustration of the broken down of non-B structures during HDA within a circular DNA. Our envisioned "repairing" mechanisms of non-B DNA structures by positive supercoiling affiliated with DNA replication.

3.2 Sequence Design of the Circular Template DNA with Non-B Structures

3.2.1 The General Strategy to Construct a Template DNA with Circular Backbone

The strategy for constructing a circular DNA is as same as the way mentioned above. Polymerase Chain Reaction (PCR) can give the linear DNA with two identical restriction enzyme cutting sites of SacI based on particular design. Linear DNA with two cohesive ends can be produced by the reaction of endonuclease SacI digestion. Circular DNA can be synthesized from linear DNA precursor which has two identical cohesive ends. To achieve acceptable final circularization yield, the length of linear

DNA precursor was designed to be around 500 or 1000 bp, which are the optimal length of linear DNA for circularization and has been proved by us in the past.

3.2.2 The Strategy to Construct a Circular DNA with Cruciform and the Following Structural Confirmation

Figure 3.6 Schematic illustration of our synthetic route towards DNA 1.

Among many non-B DNA structures, cruciforms play an important role in various biological processes including replication, regulation of gene expression, nucleosome structure and recombination.(1,176,184,185) It is believed that formations of cruciform structures occur at or near replication origins of some

eukaryotic cells and serve as recognition signals for DNA replication.(186) We accordingly chose cruciforms as the first example of non-B DNA structures in the beginning of our investigation. It has been well established that a segment of any genetic DNA must be considered as a circle when its topological property is evaluated.(162) Therefore, a cruciform-containing circular DNA (DNA 1 in Figure 3.6) was subsequently designed as the initial DNA substrate to examine whether such a widespread non-B structure could be transformed into its original B-form conformation. The cruciform in DNA 1 possesses 56 base pairs and 3 bases in each of its stems and loop regions respectively. The size of the cruciform in DNA 1 was designed to be large enough for us to verify its presence using AFM examination.(153) The synthetic route toward DNA 1 was shown in Figure 3.6.

Table 3.1 Sequence of plasmid DNA X2420 (3593 bp) and X4511E (4220 bp). Sequence only shows one strand from 5' to 3'. Gray shadow indicates the sequence of linear DNA which employed to prepare DNA 3 and DNA 1 respectively. The sequence in red color can be produced by HDA.

Name of DNA	Nucleotide sequence
Vector: X2420G	5 'CTTGCAAACAAAAAACCACCGCTACCAGCGGTGGTTTGTGTTG CCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAAGTGGCTTC AGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAG TTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTC GCTCTGCTAATCCTGTTACCAAGTGGCTGCTGCCAGTGGCGATAAG TCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAG GCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGC TTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAG CATTGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGG TATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAG CTTCCAGGGGAAACGCCTGGTATCTTTATAGTCCTGTGCGGGTTT CGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGG

	GGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTTACGG TTCCTGGCCTTTTGTCTGGCCTTTTGTCTACATGTGAAATTGTAAA CGTTAATATTTTGTAAAATTCGCGTTAAATTTTTGTAAATCAG CTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAA ATCAAAAGAATAGACCGAGATAGGGTTGAGTGTGTGTTCCAGTTTG GAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGG GCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATC ACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAA TCGGAACCCTAAAGGGATGCCCCGATTTAGAGCTTGACGGGGAAA GCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGC GGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAAC CACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCA TACATGTGCTGAGGATCGAGTCTTAATTACTGCCGGCCTTGTAGA AACGCAAAAAGGCCATCCGTCAGGATGGCCTTCTGCTTAGTTTGA TGCTTGGCAGTTTATGGCGGGCGTCCTGCCCGCCACCCTCCGGGC CGTTGCTTCACAACGTTCAAATCCGCTCCCGGCGGATTTGTCTTA CTCAGGAGAGCGTTCACCGACAAACAACAGATAAAACGAAAGGCC CAGTCTACCGACTGAGCCTTTCGTTTTATTTGATGCCTGGCAGTT CCCTACTCTCGCGTTAACGCTAGCATGGATGTTTTCCAGTCACG ACGTTGTAAACGACGGCCAGTCTTAAGCTCGGGCCCCAAATAAT GATTTTATTTTGAAGTATAGTGACCTGTTGTTGCAACAAATTGA TGAGCAATGCTTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGG CTTGAAGGAATTCGGCAAGTCTTCCCAC TTAGTGGAATCCTCGTCG CAAAACCGATTAAGTTGGGTAACGCCAGGGTTTTCCAGTCACGA CGTTGTAAACGACGGCCAGTCCGTAATACGACTCA CTTAAGGCC TTGACTAGAGGGTACCAACCTAGGTATCTAGAACCGGTCTCGAGC CATAACTTCGTATAGCATAACATTATACGAAGTTATATAAGCTGTC AAACATGAGAATTCTTGTTATAGGTTAATGTCATGATAATAATGG TTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAA CCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGC TCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAA GGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCT TTTTTGCGGCATTTTGCCTTCCTGTTTTTGTCTACCCAGAAACGC TGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGG GTTACATCGAACTGGATCTCAACAGCGGTAAGTTAAGCTTTTTGC ACAACATGGGGGATCATGTAACCTCGCCTTGATCGAAGGAGAGAAG AGCTGGAGCTCAATGAAGCCATACCAAACGACGAGCGTGACACCA CGATGCCTGCAGGAATTCTCGAGCCATAACTTCGTATAGCATAAC ATTATACGAAGTTATCCATGGACTAGTGAGTCGTATTACGTAGCT TGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATC CGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTA
--	--

	AAGCCTGGGGGGGGTTAACCATGGATCCGGTAAGTGGGATATCGA AGACTTGCCGCTAGAATTCGATCCCCCTATAGTGAGTCGTATTACA TGGTCATAGCTGTTTCCTGGCAGCTCTGGCCCGTGTCTCAAAATC TCTGATGTTACATTGCACAAGATAAAAAATATATCATCATGAACAA TAAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGA GCCATATTCAACGGGAAACGTCGAGGCCGCGATTAAATTCCAACA TGGATGCTGATTTATATGGGTATAAATGGGCCCCGCGATAATGTCTG GGCAATCAGGTGCGACAATCTATCGCTTGTATGGGAAGCCCGATG CGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATG ATGTTACAGATGAGATGGTCAGACTAACTGGCTGACGGAATTTA TGCCTCTCCCGACCATCAAGCATTTTATCCGTACTCCTGATGATG CATGGTTACTCACCCTGCGATCCCCGGAAAAACAGCATTCCAGG TATTAGAAGAATATCCTGATTCAGGTGAAAAATATTGTTGATGCGC TGGCAGTGTTTCTGCGCCGGTTGCATTTCGATTCCTGTTTGTAATT GTCCTTTTAAACAGCGATCGCGTATTTTCGTCTCGCTCAGGCGCAAT CACGAATGAATAACGGTTTGGTTGATGCGAGTGATTTTGATGACG AGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCATA AACTTTTGCCATTCTCACCGGATTCAGTCGTCACTCATGGTGATT TCTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTT GTATTGATGTTGGACGAGTCGGAATCGCAGACCGATAACCAGGATC TTGCCATCCTATGGAACGCTCGGTGAGTTTCTCCTTCATTAC AGAAACGGCTTTTTTCAAAAATATGGTATTGATAATCCTGATATGA ATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAATCAG AATTGGTTAATTGGTTGTAACTGGCAGAGCATTACGCTGACTT GACGGGACGGCGCAAGCTCATGACCAAATCCCTTAACGTGAGTT ACGCGTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAG GATCTTCTTGAGATCCTTTTTTTTCTGCGCGTAATCTGCTG3 '
Vector: X4511E	5 'CTAAATTGTAAGCGTTAATATTTTTGTTAAATTTCGCGTTAAAT TTTTGTAAATCAGCTCATTTTTTAAACCAATAGGCCGAAATCGGC AAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGT GTTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAACGTGGAC TCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCA CTACGTGAACCATCACCTAATCAAGTTTTTTGGGGTTCGAGGTGC CGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTAGA GCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAG AAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTC ACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTA CAGGGCGCGTCCCATTGCGCATTCAGGCTGCGCAACTGTTGGGAA GGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAA GGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTT TCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAGCGCGCGTAA

TACGACTCACTATAGGGCGAATTGGGTACGGCCGTCAAGGCCAAG CTTCCCAGTCAGAGGTGGATCCTCGTCGCAAAACGAGCTCCTCGA TGAAAGATCCTTTCCGGAGATCCTTTTGGCGAGCGGTGGTTTGAT AAGCTCCGGCAGTCCGCCTTGACTAGAGGGTACCAACCTAGGTAT CTAGAACGAATTCCGGAGCCTGAATCGGCCAACGCGCGGGGAGAG GCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGATT CGCTGCGCTCGGTTCGTTTCGGCTGCGGCGAGCGGTATCAGCTCACT CAAAGGCGGTAAATACGGTTATCCACAGAATCAGGGGATAACGCAG GAAAGAACATGTGAGCAATCAAGGCCAGCAAAAGGCCAGGAACCG TAAACAAGGCCGCGTTGCTGGCGTGACGAGCATCACAAACAATCG ACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATA CCAGGCGTTTCCGACTAGTGCCCTGGAAGCTCCCTCGTGCGCTCA TAAGAAGGAGAGAAGCTAAGAGAGGAACTGGACTCTCAAACATGA AACGTTTTGTTATAGGTTAATGTCATGATAATAATGGTTTTCTTAG ACGTCAGGTGGCACTTTTTCGGGGAAATGTGCGCGGAACCCCTATT TGTAATAACATTCAAATATGTATCCGCTCATGATACAATAAGTCT CCCCTGATAAATGCTTCAATGAAGGAAGAGTATGAGTATTCAACA TTTCCGTGTCGCCCTTATTCCCTTTTGCACAACATGGGGGATCAT GTAACTCGCCTTGATCGGAGCTGAATGAAGCCATACCAAACGACG AGCGTGACACCACGATGCCTGCAGCTCGAGCCCTGAATGTATTTA GCGCCAGGGTTTTCCCAGTCACGACCGCACATTTCCCCGAAAAGT GCCACCTGACGTCTAAGAAACCATTATTATCATGACTCCTGTGTG AAATTGTTATCCGCTCACGAGGCCCTTTCGCCTCGCGCGTTTCGG TGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGGCGGT CACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCA GGGCGCGTCAGCGGGTGTGGCGGGTGTGCGGGCTGGCTTAACTA TGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGGACAT ATTGTCGTTACCGAATTCATGGACTAGTGAATCGTATTACGTCTG TGTGATTGTTATCCGAGCTTATCAAACCACCGCTCGCCAAAAGGA TCTCCGGAAGGATCTTTCATCGAGCTCGGGGTAAACCATGGATC CGGAGATCTTAAGTGGGATATCACGTGAAGCTTGCAAGCTCCAGC TTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCA TGGTATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATT CCACACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGT GCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTG CCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGA ATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCT TCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTTCGTTTCGGCTG CGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAAATACGGTTATCC ACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGC CAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTT

	<p> TTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGC TCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAG GCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACC CTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGC GTGGCGCTTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTG TAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTT CAGCCCGACCGCTGCGCCTTATCCGGTAACCTATCGTCTTGAGTCC AACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGT AACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTC TTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTT GGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTT GGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGT TTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCT CAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGG AACGAAAACCTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAA AGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAA TCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCA TGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGT TCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATA CGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGA GATCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCA GCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCC TCCATCCAGTCTATTAATTGTTGCCGGAAGCTAGAGTAAGTAGT TCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGC ATCGTGGTGTACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCC GGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGC AAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGT AAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCAT AATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACT GGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGA CCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCA CATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCG GGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTTCG ATGTAACCCACTCGTGCACCCAAGTATCTTCAGCATCTTTTACT TTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCC GCAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATA CTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGT CTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAA ATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCAC 3' </p>
--	---

The inverted repeated sequences were introduced into target DNA molecules through using two primers (Primer 1 and 2 see Table 3.6 in Materials and Methods section) in the initial stage of Polymerase Chain Reactions (PCR) on a particular designed DNA template of X4511E (see Table 3.1 for the sequence).

Figure 3.7 Synthesis and structural confirmation of DNA S3. (A) Electrophoretic analysis of DNA products involved in synthesizing DNA S3. (B) Structural confirmation of DNA S3 using AFM. Scale bar indicates 200 nm.

This PCR amplification gave a linear DNA (DNA S1 in Figure 3.6) which contains 1294 base pairs in length and holds two separate inverted segments at its two terminuses. Through using our previous reported methods (187,188), a circular DNA (DNA S3 in Figure 3.6) was constructed. Formation of a circular backbone in DNA 1 was further purified and confirmed using Nuclease BAL-31, which can remove the single-stranded by-products and nick- or gap-containing circular DNA. The DNA products generated in the related reactions are anglicized using gel electrophoresis as shown in Figure 3.7A. To separate the linear and circular products, the agrose gel and buffer solutions containing ~20 ng/μl of ethidium bromide were used when running

the gel electrophoresis. Lane 1 is DNA S1 obtained by PCR. Lane 2 is DNA S2 obtained by SacI cleavage. Lane 3 is reaction mixture of ligase reaction and Lane 4 is DNA S3 obtained from the reaction mixture of ligase reaction followed by Nuclease BAL-31 hydrolysis. The formation of a circular backbone in DNA S3 was also confirmed using AFM examination (150) (Figure 3.7B).

Table 3.2 Double-stranded sequence of DNA 1.

<p>< 5' GCTCACTGATTTCGCTGCGCTCGGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGG 3' CGAGTGACTAAGCGACGCGAGCCAGCAAGCCGACGCCCTCGCCATAGTCGAGTGAGTTCCGCCATTATGCCAATAGGTGTCTTACTCCCC</p>	
<p>ATAACGCGAGAAAGAACATGTGAGCAATCAAGGCCAGCAAAAGGCCAGGAACCGTAACAAGGCCGCGTTGCTGGCGTGACGAGCATCACAA TATTGCGTCTCTTCTGTACACTCGTTAGTTCGGTTCGTTTCCGGTCTTGGCATTGTTCCGGCGCAACGACCGCCTGCTCGTAGTGT</p>	
<p>ACAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCGACTAGTGCCCTGGAAGCTCCCTCGTG TGTTAGCTGCGAGTTTCAGTCTCCACCGCTTGGGCTGTCTGATATTTCTATGGTCCGCAAGGCTGATCACGGGACCTTCGAGGGAGCACG</p>	
<p>GCTCATAAGAAGGAGAGAAGCTAAGAGAGGAACCTGGACTCTCAACATGAAACGTTTTGTTATAGGTTAATGTCATGATAAATAATGGTTTTCT CGAGTATTCCTCTCTCTCGATTCTCTCTTGTGACCTGAGAGTTGTACTTTGCAAAACAATATCCAATTACAGTACTATTATTACCAAAGA</p>	
<p>TAGACGTCAGGTGGCACTTTTCGGGAAATGTGCGCGGAACCCCTATTTGTAAATACATTCAAATATGTATCCGCTCATGATACAATAAGTC ATCTGCAGTCCACCGTGAAAGCCCTTTACACGCGCCTTGGGGATAAACATTTATGTAAGTTTATACATAGGCGAGTACTATGTTATTACG</p>	
<p>TCCCTTGATAAATGCTTCAATGAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCTTTTGACAACATGGGGGATCATG AGGGGACTATTTACGAAGTTACTTCTTCTCATACTCATAAGTTGTAAGGCACAGCGGGAATAAGGGAAAACGTGTTGACCCCTAGTAC</p>	
<p>TAACTCGCCTTGATCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGCAGCTCGAGCCCTGAATGTATTAGCG ATTGAGCGGAACTAGCCTCGACTTACTTCGGTATGGTTTGTGCTCGCACTGTGGTGCTACGGACGTCGAGCTCGGGACTTACATAAATCGC</p>	
<p>CCAGGGTTTTCCAGTCACGACCGCACATTTCCCGGAAAAGTGCCACCTGACGCTCTAAGAAACCATTTATTATCATGACTCCTGTGTGAAATT GGTCCCAAAGGGTCAGTGTGCGGTGTAAGGGGCTTTTACGGTGGACTGCAGATTCTTTGGTAATAATAGTACTGAGGACACACTTTAA</p>	
<p>GTTATCCGCTCACGAGGCCCTTTTCGCTCGCGCGTTTCGGTGATGACGCTGAAAACCTCTGACACATGCAGCTCCCGGAGGCGGTACAGCT CAATAGCGAGTGTCTCCGGGAAAGCGGAGCGCGCAAGCCACTACTGCCACTTTTGGAGACTGTGTACGTCGAGGGCCTCCGCCAGTGTGCA</p>	
<p>TGCTGTAAAGCGGATGCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTGGCGGGTGTGCGGGTGGCTTAAGTATGCGGCATC ACAGACATTGCGCTACGGCCCTCGTCTGTTTCGGGAGTCCCGCGAGTCGCCCAACCGCCACAGCCCGACCGCAATTGATACGCCGTAG</p>	
<p>AGAGCAGATTGTACTGAGAGTGACCATATGGACATATTGTCGTTACCGAATTCATGGACTAGTGAATCGTATTACGCTCTGTGTATTGTTA TCTCGTCTAACATGACTCTCAGCTGGTATACCTGTATAACAGCAATGGCTTAAGTACCTGATCACTTAGCATAATGCAGACACACTAACAAAT</p>	
<p style="text-align: center;">cruciform region</p> <div style="display: flex; justify-content: space-between;"> <div style="text-align: left;"> <p>TCC AGG</p> </div> <div style="border: 1px solid black; padding: 5px; flex-grow: 1;"> <p>GAGCTTATCAAACACCGCTCGCCAAAAGGATCTCCGAAAGGATCTTTCATCGAGCTCCTCGATGAAAGATCCTTCCGGAGATCCTT</p> <p>CTCGAATAGTTTGGTGGCGAGCGGTTTTCTTAGAGGCCCTTCTAGAAAAGTAGCTCGAGGAGCTACTTTCTAGGAAAGGCCCTTAGGAA</p> </div> </div> <p style="text-align: center;">cruciform region</p> <div style="display: flex; justify-content: space-between;"> <div style="text-align: left;"> <p>cruciform region</p> <div style="border: 1px solid black; padding: 5px; width: 150px;"> <p>TTGGCGAGCGGTGGTTTGATAAGCTC</p> </div> </div> <div style="text-align: right;"> <p>cruciform region</p> <div style="border: 1px solid black; padding: 5px; width: 150px;"> <p>CGGCAGTCCGCCTTGACTAGAGGGTACCAACCTAGGTATCTAGAACAATTCGGAGCCTGAATCG</p> </div> </div> </div> <div style="display: flex; justify-content: space-between;"> <div style="text-align: left;"> <p>cruciform region</p> <div style="border: 1px solid black; padding: 5px; width: 150px;"> <p>AACCGCTCGCCACCAAACTATTCGAG</p> </div> </div> <div style="text-align: right;"> <p>GCCGTCAGGCGGAAGTATCTCCCATGGTTGGATCCATAGATCTTGCTTAAGGCCTCGGACTTAGC</p> </div> </div> <p style="text-align: center;">cruciform region</p> <p>GCCAACGCGCGGGGAGAGGCGGTTTGCCTATTGGGCGCTCTTCGCTTCCTC 3'</p> <p>CGGTTGCGGCCCCCTCTCCGCCAAACGCATAACCCGCGAGAAGCGCAAGGAG 5' ></p>	

Notes: (1) DNA 1 is a circular DNA; and
(2) < and > stand for the termini that are covalently connected.

It is clear in the past that the cruciform structures can be formed in a negative supercoiled circular DNA which promotes breathing effect in the double helix.(189,190) We accordingly introduced negative supercoils into DNA circles through using DNA gyrase (DNA S4). The resulting DNA samples were incubated in 60 mM NaCl, a condition that benefits the formation of cruciform structures.(191,192) The nucleotide sequences of DNA 1 is given in Table 3.2 and the formation of cruciform was authenticated using gel electrophoresis (DNA 1, Lan 2 in Figure 3.8A) and AFM examination (Figure 3.8B).

Figure 3.8 Synthesis and structural confirmation of DNA 1. (A) Electrophoretic analysis of DNA products involved in synthesizing DNA 1. No ethidium bromide is used during electrophoresis. The gel was incubated in a solution containing 10 ng/ul EtBr and 1x TAE buffer for 30 minutes. The gel was visualized using Gel Documentation System. Lane M: molecular weight makers; Lane 1: negatively supercoiled DNA S4 obtained through incubation of DNA S3 with DNA gyrase; Lane 2: cruciform-containing DNA 1 obtained by incubation of DNA S4 in 60 mM NaCl. (B) Structural confirmation of DNA 1 using AFM. Scale bar indicates 200 nm.

3.2.3 The Strategy to Construct DNA 3 and DNA 5 and the Following Structural Confirmation

Figure 3.9 Schematic illustration of our synthetic route towards DNA 3 and DNA 5.

Besides cruciform, G-quadruplexes are four-stranded DNA structures composed of two or more stacks of G-quartets in which four guanines are arranged in a square planar array. Previous *in vitro* studies demonstrated that most of the G-quadruplex structures exhibit their melting points ranging from 50 to 80 °C, and are incapable of resuming their original B-conformation under the physiological conditions once they are formed. Because of its high thermodynamic stability and

likely pervasiveness in the eukaryotic cells, G-quadruplex was chosen as well in our studies for examining the possibility of disintegration of non-B DNA structure during the course of DNA replication. A G-quadruplex-containing circular DNA with relaxed conformation (DNA 3) was accordingly designed and synthesized (Figure 3.9) as DNA substrate according to our previously reported method. The formation of DNA S7 (Figure 3.10) and G-quadruplex-containing circular DNA 3 was examined by gel electrophoresis and AFM (Figure 3.11)

Figure 3.10 Synthesis and structural confirmation of DNA S7. (A) Electrophoretic analysis of DNA products involved in synthesizing DNA S7 (see Figure 3.9). To separate the linear and circular products, the agarose gel and buffer solutions containing ~20 ng/ μ l of ethidium bromide were used when running the gel electrophoresis. Lane M: molecular weight markers; Lane 1: DNA S5 generated through PCR amplification reactions (Step 1 in Figure 3.9); Lane 2: DNA S6 with its cohesive ends created through using SacI (Step 2 in Figure 3.9); Lane 3: crude product of DNA S7 produced through reaction of DNA S6 and T4 DNA ligase; Lane 4: pure DNA S7 obtained through hydrolysis of crude product of DNA S7 by Nuclease BAL-31 (Step 3 in Figure 3.9). (B) Structural confirmation of DNA S7 using AFM. The DNA sample used for this AFM examination was the same batch of sample as the one loaded into

Lane 4 in Figure 3.10A. The backbone circularity of DNA S7 was verifiable by the naked eyes.

Figure 3.11 Synthesis and structural confirmation of DNA 3 and DNA 5 according to our previously reported method(188). (A) Electrophoretic analysis of DNA products involved in synthesizing DNA 3 and DNA 5. No ethidium bromide is used here. Lane M: molecular weight markers; Lane 1: negatively supercoiled DNA S8 obtained through incubation of DNA S7 with DNA gyrase (Step 4 in Figure 3.9); Lane 2: G-quadruplex-containing circular DNA (DNA 3 in Figure 3.9) obtained by incubation of DNA S8 in 150 mM KCl (Step 5 in Figure 3.9); Lane 3: both nicked site- and G-quadruplex-containing circular DNA (DNA 5 in Figure 3.9) obtained by incubation of DNA 3 with Nt.BsmAI. (B) Structural confirmation of DNA 3 using AFM. Left: AFM

With the aim to confirm whether the disintegration of G-quadruplex is indeed associated with positive supercoiling, a both nicked site- and G-quadruplex-containing circular DNA was synthesized (DNA 5 in Figure 3.9). The structure of DNA 5 was also tested by gel electrophoresis and AFM (Figure 3.11). The double-stranded sequence of DNA 5 are the same as DNA 3 as shown in Table 3.3.

3.3 Materials and Methods

3.3.1 Vectors, Oligonucleotide, Enzymes and Chemicals

Most of the Vectors, Oligonucleotide, Enzymes and Chemicals used in this research were listed as shown as follows (see Table 3.4). Items that are not in the list were obtained from Sigma-Aldrich with analytical grade or molecular biology grade.

Table 3.4 Vectors, Oligonucleotide, Enzymes and Chemicals used in this research.

Item(s)	Supplier(s)	Item(s)	Supplier(s)
Vector DNA (X2420G and X4511E)	Generay Biotech (Shanghai, China)	Oligodeoxyribonucleotides (primers)	Sigma-Proligo (Singapore)
DNA ladder (100 bp)	Fermentas (Singapore)	DNA ladder (1 Kb and 100 bp)	New England Biolabs (Ipswich, MA, US)

QIAquick PCR purification kit	Qiagen (Singapore)	QIAquick Gel Extraction Kit	Qiagen (Singapore)
<i>Taq</i> Polymerase	New England Biolabs (Ipswich, MA, US)	SacI endonuclease	New England Biolabs (Ipswich, MA, US)
T4 DNA ligase	New England Biolabs (Ipswich, MA, US)	BAL 31 exonuclease	New England Biolabs (Ipswich, MA, US)
DNA gyrase	New England Biolabs (Ipswich, MA, US)	Topoisomerase I	New England Biolabs (Ipswich, MA, US)
Endonucleases Nt.BsmAI	New England Biolabs (Ipswich, MA, US)	Ethidium bromide	Research Biolabs (Singapore)
Mini Prep Cell	Bio-Rad (Hercules, CA, US)	Biological purity water	1st Base Pte. Ltd (Singapore)
TAE, TBE, TRIS	1st Base Pte. Ltd (Singapore)	Agarose	Invitrogen (Carlsbad, CA, US)

3.3.2 Experimental Procedures for Helicase-Dependent Isothermal DNA Amplification and AFM Examination

Experimental procedures for helicase-dependent isothermal DNA amplification: A mixture containing plasmid DNA (0.05 pmol), forward primer (10 pmol) and reverse primer (10 pmol), UvrD helicase (100 ng), MutL (400 ng), T4 gene 32 protein (4.5 mg), ATP (0.15 mmol) and exo-Klenow polymerase (5 U) was incubated at 37 °C for 2 hours.(181) The nucleotide sequences of forward primer and reverse primer are shown in Table 3.5.

Table 3.5 Nucleotide sequences of primers used in isothermal helicase-dependent amplification (HDA). Primer-f-1 and Primer-r-1 are forward primer and reverse primer for DNA 1. Primer-f-2 and Primer-r-2 are forward primer and reverse primer for DNA 3.

Name of DNA	Nucleotide sequence
Primer-f-1	5' CGCCAGGGTTTCCAGTCACGAC 3'
Primer-r-1	5' AGCGGATAACAATTTACACAGGA 3'
Primer-f-2	5' TTAGTGGATCCTCGTCGCAA 3'
Primer-r-2	5' TGAGTCGTATTACGGACTGG 3'

Experimental procedures for DNA sample preparations and AFM examination: All micas were modified on their surfaces with (3-aminopropyl)triethoxysilane(150) before use. DNA samples for AFM examination were prepared into solutions initially that contained 20 mM Tris-HCl (pH = 7) and 0.1 to 0.01 µg/ml DNA. 5 µl to 10 µl of the obtained DNA solutions were placed next in the middle of the newly prepared APS-mica plates (~1 x 1 cm²), which were further kept at room temperature for 5 minutes. The surfaces of the APS-mica plates bound by DNA were then rinsed for 3 times using distilled water. AFM images of DNA molecules on the APS-mica plates

were obtained in Tapping ModeTM on a MultimodeTM AFM (Veeco, Santa Barbara, CA) in connection with a Nanoscope VTM controller. Antimony (n) doped Si cantilevers with nominal spring constants between 20 and 80 N/m were selected. Scan frequency was 1.9 Hz per line and the modulation amplitude was in a nanometer range. All DNA sample determinations were carried out in air at room temperature. In the case of Holliday junction (cruciform) and spurs, the observed shapes were significantly different from anything seen on pure duplex DNA. As a result, all of these structures were included in the dataset. Since variations in the imaging surface and/or kinks in the circular DNA, small raised structures (blobs) were occasionally seen on pure duplex DNA. To distinguish the newly formed non-B structures from the features occasionally found on the pure duplex DNA, a criterion was set according to previous the studies. The normal height and the peak height were determined for 20 duplex DNA molecules. The mean of normal height was 0.51 ± 0.01 nm, and the mean of peak height was 0.67 ± 0.02 nm, with a highest absolute value of 0.83 nm. Consequently, any blob < 0.9 nm in height was excluded from the dataset and any blob ≥ 1 nm was included. The height measurements were taken across the middle of each blob. Frequency distributions of lengths (in nm) of DNA were obtained by detecting the circumference along the backbone of circular DNA, which were measured by drawing a series of very short lines along the DNA contour and summing the lengths.

3.3.3 Experimental Procedures for Synthesis and Structural Confirmation of DNA 1

Table 3.6. Nucleotide sequences of primers used in our polymerase chain reactions.

Name of DNA	Nucleotide sequence
Primer 1	5' GTGGATCCTCGTCGCAAAAC 3'
Primer 2	5' CCGGATCCATGGTTAACCCC 3'
Primer 3	5' CCGAGCTCAGGATCCGGATGATCCCTAACCCCTAACCCCTAACCCCTAATACATGTGCTGAGGATCGAG 3'
Primer 4	5' TCGTTTGGTATGGCTTCATT 3'

Step 1 in Figure 3.6: X4511E (a plasmid DNA, see Table 3.1 for the sequence) was purchased from Generay Biotech (Shanghai, China) and used as the template for PCR. Nucleotide sequences of both Forward Primer (Primer 1) and Reverse Primer (Primer 2) are shown in Table 3.6. The PCR amplification reactions were carried out following reported procedures with an annealing temperature of 58 °C (193,194) and the amplification products were verified through electrophoresis (Lane 1 in Figure 3.7A).

Step 2 in Figure 3.6: A mixture containing 10 mM Bis-Tris-Propane-HCl, 10 mM MgCl₂, 1 mM Dithiothreitol, 10 units SacI and ~2 µg DNA S1 was incubated at 37 °C for 1 hour to generate a cohesive end-containing linear DNA (DNA S2, see Lane 2 in Figure 3.7A).

Step 3 in Figure 3.6: A mixture containing 50 mM Tris-HCl, 10 mM MgCl₂, 1 mM ATP, 10 mM dithiothreitol, 20 units T4 DNA ligase and ~500 ng DNA S2 was incubated at 16 °C for 8 hours (Lane 3 in Figure 3.7A). The resultant reaction mixture was allowed next to react with BAL-31 (an exonuclease that hydrolyzes open end-

containing DNA) in order to acquire pure closed circular DNA products (DNA S3, see Lane 4 in Figure 3.7A).

Step 4 in Figure 3.6: A mixture containing 5 units of DNA gyrase, 1 x DNA gyrase buffer (35 mM Tris-HCl, 24 mM KCl, 4 mM MgCl₂, 2 mM DTT, 1.75 mM ATP, 5 mM spermidine, 0.1 mg/ml BSA and 6.5% Glycerol) and ~500 ng DNA S3 was incubated at 37 °C for 1 hour to generate negatively supercoiled DNA (DNA S4; also see Lane 1 in Figure 3.8A).

Step 5 in Figure 3.6: A mixture containing 20 mM Tris-HCl (PH = 7), 60 mM NaCl and ~500 ng DNA S4 was kept at room temperature overnight to produce cruciform-containing circular DNA (DNA 1 whose sequence is shown in Table S3, also see Lane 2 in Figure 3.8A).

3.3.4 Experimental Procedures for Synthesis and Structural Confirmation of DNA 3 and DNA 5

Step 1 in Figure 3.9: X2420G (plasmid DNA) was purchased from Generay Biotech (Shanghai, China). The forward primer (Primer 3 in Table 3.6) contained the cytosine-rich segment. The detailed nucleotide sequences of Forward Primer (Primer 3) and reverse primer (Primer 4) used in the studies are shown in Table 3.6. The PCR amplification reactions were carried out following reported procedures with a annealing temperature of 61 °C and the amplification product (DNA S5) was verified through electrophoresis (DNA S5, Lane 1 in Figure 3.10A).

Step 2 in Figure 3.9: A mixture containing 10 mM Bis-Tris-Propane-HCl, 10 mM MgCl₂, 1 mM Dithiothreitol, 10 units SacI and ~2 µg DNA S5 was incubated at 37 °C for 1 hour, which gave rise to a cohesive end-containing linear DNA (DNA S6; Lane 2 in Figure 3.10A).

Step 3 in Figure 3.9: A mixture containing 50 mM Tris-HCl, 10 mM MgCl₂, 1 mM ATP, 10 mM dithiothreitol, 20 units T4 DNA ligase and ~500 ng DNA S6 was incubated at 16 °C for 8 hours (Lane 3 in Figure 3.10A). The resultant reaction mixture was allowed next to react with BAL-31 (an exonuclease that hydrolyzes opening end-containing DNA) in order to acquire pure closed circular DNA products (DNA S7, Lane 4 in Figure 3.10A).

Step 4 in Figure 3.9: A mixture containing 5 units of DNA gyrase, 1 x DNA gyrase buffer (35 mM Tris-HCl, 24 mM KCl, 4 mM MgCl₂, 2 mM DTT, 1.75 mM ATP, 5 mM spermidine, 0.1 mg/ml BSA and 6.5% Glycerol) and ~500 ng DNA S7 was incubated at 37 °C for 1 hour to generate negatively supercoiled DNA (DNA S8; also see Lane 1 in Figure 3.11A).

Step 5 in Figure 3.9: A mixture containing 20 mM Tris-HCl (PH = 7), 150 mM KCl and ~500 ng DNA S8 was kept at room temperature overnight to produce G-quadruplex-containing circular DNA (DNA 3 whose sequence is shown in Table S4, also see Lane 2 in Figure 3.11A).

Step 6 in Figure 3.9: A mixture containing 5 units of Nt.BsmAI, 1 x Nt.BsmAI buffer (20 mM Tris-acetate, 50 mM potassium acetate, 10 mM Magnesium Acetate, 1

mM Dithiothreitol) and ~500 ng DNA 3 was incubated at 37 °C for 1 hour to generate a both nicked site- and G-quadruplex-containing circular DNA (DNA 5; also see Lane 3 in Figure 3.11A).

3.4 Results and Discussion

3.4.1 Breaking down Cruciform in the Course of Isothermal DNA Replication (HDA)

Figure 3.12 depicts our designed route for examining the possible breakdown of cruciform structures during the course of helicase-dependent amplification (HDA). Two strands of duplex DNA are separated by DNA helicases and coated by single-stranded DNA (ssDNA)-binding proteins, which facilitate the hybridization of two sequence-specific primers with each border of the target DNA. At the same time, DNA polymerases extend the primers annealed to the templates to produce a dsDNA. Because the partial separation of duplex segments (replication bubble) occurred and no topoisomerase is involved in this course of action, positive supercoiling can be generated within the closed DNA circles (Structure 1 in Figure 3.12) in order to satisfy the “DNA Topological Conservation Law”(63,195). It has been well investigated that the helical turns of the positive supercoiling is highly overwound. We therefore speculated that disintegration of cruciform structures and restoration of its original B conformation could reduce the torsional constraint (Structure 2 in Figure 3.12). In addition, it is well known that the disintegration of a non-B structure within a

covalently closed circular DNA will lead to a negative supercoiling. This happens because the part of cruciform base pairing will contribute to the increase of twist number for the whole DNA circle while the linking number still remain during this course. According to the Equation 1.1 in Chapter 1 ($Lk = Wr + Tw$), the decrease of writhe number will occurred and the negative supercoiling should be generated.(1,187) In order to visualize and confirm the disappearance of cruciform structures, we decided next to use exonuclease and proteinase K to remove the linear DNA amplified by HDA and proteins within the reaction system.

Figure 3.12 Pictorial diagram of an envisioned disintegration of DNA cruciform structures by positive DNA supercoiling.

Figure 3.13 summarize the experimental evidences confirming that such a process takes place. As helicases are able to unwind duplex DNA enzymatically, we tested whether the entire HDA reaction could be carried out at 37 °C using a plasmid DNA (X4511) as template (see Experimental Section in Section 3.3 for detail experimental procedures(181)). As shown in Figure 3.13A (Lane 1), a band with

Figure 3.13 Disintegration of DNA cruciform structures during the course of isothermal HDA. (A) Electrophoretic analyses on DNA products amplified by HDA using B-form (lane 1) or non-B containing (lane 2) plasmid DNA as template. The DNA products in above two lanes were designed with same sequences. (B) Electrophoretic analyses on the reaction mixtures involved in the production of DNA 2 from DNA 1. Lane M: Molecular weight markers; Lane 1: DNA 1 alone; Lane 2: A product obtained after the HDA reaction mixture were digested by lambda exonuclease and proteinase K. (C) AFM images of DNA 1 (scale bar: 150 nm). The sample used for this AFM examination was the same batch of sample as the one loaded into Lane 1 in Figure 3.13B. (D) AFM images of DNA 2 (scale bar: 150 nm). The sample used for this AFM examination was the same batch of sample as the one loaded into Lane 2 in Figure 3.13B.

almost the same mobility shift as 100 base pairs marker can be observed, which indicated that the HDA reaction occurred and a linear DNA with 104 base pairs was

amplified as our initially designed (Lane 1 in Figure 3.13A, Table 3.1). At this experimental stage, we decided to use our newly synthesized DNA 1 as the template to investigate whether the isothermal HDA reaction can be conducted within this non-B structure containing mini-plasmid DNA (DNA 1). Our results show that a DNA product with the length of around 100 base pairs can also be detected, which indicated that the isothermal HDA reaction occurred when using the non-B-containing substrate as DNA template (Lane 2 in Figure 3.13A). The reaction mixtures involved in the production of DNA 2 from DNA 1 were accordingly tested by electrophoreses and AFM. As anticipated, DNA circles with negative supercoiling (DNA 2 in Figure 3.12) were observed during the AFM examination while mobility shift difference between DNA 1 and DNA 2 can also be detected by electrophoretic analysis as shown in Figure 3.13B to D. The observations above clearly showed that the conformation change occurred from DNA 1 to DNA 2 and the cruciform structures in DNA 1 were disintegrated.

3.4.2 Confirmation of Breaking Down Cruciform Structures by Topo I Relaxation

In order to further confirm the disappearance of cruciform structure along the backbone of DNA 2, we decided to treat DNA 2 with Topo I, an enzyme can remove negative supercoiling from DNA circles. As shown in Figure 3.14A, the slower migration was observed when DNA 2 was treated with Topo I. By AFM examination, DNA samples from Lane 2 in Figure 3.14A was proved to be in a circular structures and there was indeed no cruciform structure left in the final DNA molecules (Figure

3.14B). To further confirm the structure difference between DNA 1 (images in Figure 3.13C) and Topo I relaxation products (images in Figure 3.14B), frequency distributions of the lengths of DNA molecules was measured. In Figure 3.14C and D, the curves indicate the fitted Gaussian functions. The mean length of DNA 1 in Figure 3.13C is 393.14 ± 0.11 nm while the mean length of DNA in Figure 3.14B is 412.60 ± 0.41 nm, which indicate that the perimeter differences between the circular backbones of DNA 1 in Figure 3.13C and DNA molecules in Figure 3.14B are detectable in the overwhelming number of their AFM images. Because the length of circular backbone of Topo I relaxation products (images in Figure 3.14B) should theoretically be longer than those in DNA 1 after breaking down cruciform structures, these distributions clearly show that those secondary structures are disintegrated from DNA 1.

Figure 3.14 Examination of the absence of cruciform structures by removing the supercoils in DNA 2 with Topo I. (A) Electrophoretic analysis of DNA products involved in Topo I relaxation reaction. Lane 1: DNA 2 alone; Lane 2: mixtures

obtained by the Topo I relaxation with DNA 2. (B) AFM images of the DNA products obtained by the Topo I relaxation with DNA 2 (scale bar: 150 nm). The sample used for this AFM examination was the same batch of sample as the one loaded into Lane 2 in Figure 3.14A. (C) – (D) Frequency distributions of the lengths (nm) of DNA 1 (images in Figure 3.13C) and Topo I relaxation products (images in Figure 3.14B).

3.4.3 Examination of the Stability of Cruciform Structures in HDA Buffers

Figure 3.15 Examination of effects of buffer and salts on the stability of the cruciform residing in DNA 1. (the tests conducted in the current section served as the control experiments for those shown in Figure 3.13). (A) Agarose gel electrophoretic analysis of our control studies. (B) AFM images of the final products of our current control studies.

With the aim of finding out whether buffers and salts used in our studies could possibly interfere the stability of cruciform structures, control experiments were conducted. In the agarose gel electrophoretic analysis (Figure 3.15A), the samples

loaded into Lane 1 and Lane 2 were prepared in the same ways as for those loaded into Lane 1 and Lane 2 in Figure 3.13 except that helicase and its associated proteins were not used in the control experiments. The DNA sample used for this AFM examination was the same batch of sample as the one loaded into Lane 2 in Figure 3.15. The presence of cruciform in the final products obtained from our control studies was verifiable by the naked eyes. Our AFM examination confirms that the DNA obtained from our control studies still contained cruciform structures in its sequence (Figure 3.15).

3.4.4 The Control Experiment to Examine Breaking down Cruciform is affiliated with Positive Supercoils

According to the previous reports,(176,190,192,196) if a cruciform-containing circular DNA is linearized or nicked, the cruciform structure is destabilized and eventually disappears. Therefore, it is very hard to construct a both nicked site- and cruciform-containing circular DNA and use it as the substrate for the experiment (to directly confirm whether the breaking down of cruciform is affiliated with positive supercoils) as described in Figure 3.19 and Figure 3.20.

Figure 3.16 Examination of the interaction between helicase and cruciform in DNA 1.

(A) Agarose gel electrophoretic analysis. Lane M: Molecular weight markers; Lane 1: DNA 1; Lane 2: Reaction mixture obtained by incubation of DNA 1 with helicase only in the HDA buffer with ATP; (B) AFM images of the final products of our current studies.

To rule out the possibility of the interaction between helicase and cruciform leading to the disintegration of cruciform structure, we treated cruciform-containing DNA 1 with helicase only in the HDA buffer with ATP as shown in Figure 3.16. The reaction is conducted in the absence of primer and associated proteins (e.g., T4 gene 32 protein, exo-Klenow polymerase). The DNA sample used for the following AFM examination was the same batch of sample as the one loaded into Lane 2 in Figure 3.16A. The presence of cruciform in the final products obtained from the control studies can be verifiable by the naked eyes. Our results show that the cruciform structures still remain after DNA 1 was treated with helicase and buffers.

Based on above results, we speculated that (1) there is no direct interaction between helicase and cruciform or the interaction is too weak to disintegrate the cruciform structures; (2) even helicase has the special abilities of unwinding DNA double strands, only very short duplex regions can be separated randomly because the separated two single stranded DNA will re-anneal if there is no stabilization factors (e.g., Single-Strand Binding (SSB) Proteins). Those randomly opened short region of DNA duplex can only lead to a low level of positive supercoiling, which cannot supply the enough energy needed to disintegrate the cruciform structures within DNA

1. On the other hand, with the assistance of its associated proteins and primers used in HDA, helicase can open up a long part of duplex for the binding of primers and DNA polymerase, which lead to the formation of relatively high level of positive supercoiling and reach the energy needed to disintegrate the cruciform structures as shown in Figure 3.13.

3.4.5 Breaking Down G-quadruplex in the Course of Isothermal DNA Replication (HDA)

Figure 3.17 Pictorial diagram of an envisioned disintegration of G-quadruplex structures by positive DNA supercoiling.

Figure 3.17 shows the schematic illustration of our examination on the breakdown of a G-quadruplex structure that resides in DNA 3 during the course of HDA. With the purpose of investigating whether the G-quadruplex-containing DNA circle can serve as the template and produce desired linear DNA through HDA reaction, DNA 3 and primers were mixed with enzyme mixture (helicase, DNA polymerase and other proteins) and incubated at 37 °C. Similar with cruciform, linear DNA products and proteins were digested by exonuclease and proteinase K, by which

Structure 3 in Figure 3.17 can be obtained. In order to visualize and confirm the disappearance of G-quadruplex structures in a more accurate manner using AFM, Topo I was used to remove the remaining supercoiling left in the target circular DNA.(188)

Our experimental evidences verifying the disappearance of G-quadruplex structure during the course of HDA are shown in Figure 3.18. In the agarose gel electrophoretic analysis, a band (Lane 2 in Figure 3.18A) with anticipated mobility shift was observed, which confirmed that HDA reaction occurred. Our further AFM analysis (Figure 3.18D) confirmed that there was indeed no G-quadruplex structure left in the final DNA molecules (DNA 4 in Figure 3.17). In addition, statistical examination on the lengths of backbones of DNA 3 (initial DNA substrate) and DNA 4 (final disintegration products) in their AFM images were carried out in our studies (Figure 3.18E – F). Generally, frequency distributions of lengths (in nm) of DNA were obtained by detecting the circumference along the backbone of circular DNA, which were measured by drawing a series of very short lines along the DNA contour and summing the lengths. Our results showed that the mean length of DNA 3 is 376.78 ± 0.37 nm while the mean length of DNA 4 is 386.58 ± 0.17 nm, which indicate that the perimeter differences between the circular backbones of DNA 3 and DNA 4 are detectable in the overwhelming number of their AFM images. Since the length of circular backbone of DNA 4 (images in Figure 3.18D and F) should theoretical be longer than those in DNA 3 (images in Figure 3.18C and E) after the disintegration of non-B structures, these distributions clearly show that the G-quadruplex structures are disintegrated from DNA 3.

Figure 3.18 Disintegration of G-quadruplex structures during the course of isothermal HDA. (A) Electrophoretic analyses on DNA products amplified by HDA using B-form (lane 1) or non-B containing (lane 2) plasmid DNA as template. The DNA products in above two lanes were designed with same sequences. (B) Electrophoretic analyses on the reaction mixtures involved in the production of DNA 4 from DNA 3. Lane M: Molecular weight markers; Lane 1: DNA 3 (G-quadruplex-containing circular DNA) alone; Lane 2: A product obtained after the HDA reaction mixture were digested by lambda exonuclease and proteinase K.; Lane 3: DNA 4 obtained by the Topo I relaxation with the products in lane 2. (C) Left: AFM images of DNA 3 (scale bar: 150 nm). The sample used for this AFM examination was the same batch of sample as the one loaded into Lane 1 in Figure 3.18B. Right: section analyses of an AFM image of DNA 3. (D) Left: AFM images of DNA 4 (scale bar 150 nm). The

sample used for this AFM examination was the same batch of sample as the one loaded in Lane 3 in Figure 3.18B. Right: Section analyses of an AFM images DNA 4. (E) – (F) Frequency distributions of the lengths (nm) of DNA 3 and DNA 4 in their AFM images. The curves indicate the fitted Gaussian functions.

3.4.6 The Control Experiment to Examine Breaking down G-quadruplex is affiliated with Positive Supercoils

With the aim to confirm whether the disintegration of G-quadruplex is indeed associated with positive supercoiling, both nicked site- and G-quadruplex-containing DNA 5 (Figure 3.11) was used as the substrate in HDA reaction. It has been well known that the circular DNA with a nicked site is a "most" relaxed conformation and no supercoiling can be introduced into the DNA circle unless the nicked site can be sealed by DNA ligase or topoisomerases.(66) We accordingly used DNA 5 as the template and performed the HDA reaction in the same way as aforementioned procedures as shown in Figure 3.19.

Figure 3.19 Pictorial diagram of an envisioned reaction pathway of G-quadruplex-containing DNA 5 in HDA reaction.

As shown in Figure 3.20A, AFM examination showed that G-quadruplex structures still remained in the final circular DNA products after HDA reaction (DNA 6 in Figure 3.19). In addition, frequency distributions of lengths (in nm) and statistical analysis of DNA 5 and DNA 6 are measured and shown in Figure 3.20C to D. These distributions show that no apparent increasing of the plasmid can be detected, which indicated that disintegration of G-quadruplex is indeed associated with positive supercoiling formed by HDA reaction.

Figure 3.20 Examination of the presence of G-quadruplex in DNA 6 after the action of isothermal HDA. (A) Left: AFM images of DNA 5 (scale bar: 150 nm). Right: section analyses of an AFM image of DNA 5. (B) Left: AFM images of DNA 6 (scale bar 150 nm). Right: Section analyses of an AFM images DNA 6. (C) – (D) Frequency distributions of the lengths (nm) of DNA 5 and DNA 6 in their AFM images. The curves indicate the fitted Gaussian functions.

In addition, it has been reported that vigorous G-quadruplex unwinding is a conserved feature of Pif1 helicases, whereas three RecQ helicases had much lower G-

quadruplex unwinding activity than the Pif1 helicases.(197,198) In the current studies, UvrD helicase was used and no direct G-quadruplex unwinding activity was observed (Figure 3.20). It is our speculation that living organisms may employ different or multiple G-quadruplex disintegration mechanisms during the various cellular processes.

3.5 Conclusion

In conclusion, HDA reaction was chosen as a modeling to investigate whether the transformation of non-B structures into its normal duplex structure occurs in this course of DNA replication *in vitro*. Circular DNAs containing some typical non-B DNA structures such as cruciform and G-quadruplex were designed and prepared during our investigations, whose structures were subsequently verified using gel electrophoresis and AFM. The topological conformation changes of cruciform- and G-quadruplex-containing template DNAs after HDA reaction were examined through electrophoretic analyses and AFM examination. The outcomes of our AFM studies show that those non-B DNA structures that had resided in the original circular DNA sequences were disintegrated and the original canonical B forms of DNA resumed. Since positive supercoiling is constantly affiliated with DNA replication, it is likely that the torsional stress associated with the overwound structures of positive DNA supercoils acts as the driving forces to break down non-B DNA structures.

On one hand, it is showed that negative supercoiling of DNA facilitate the folding of some non-B structures such as G-quadruplex, cruciform and H-DNA according to accumulating evidences in our studies and literature reports. These evidences imply that living organisms will create those non-B structures as cellular signals through maneuvering the formation of their DNA negative supercoiling when the particular cellular processes (such as replication, transcription and recombination) need to be regulated. On the other hand, the studies in Chapter 3 demonstrated that positive supercoiling of DNA can break down the structural entities of non-B DNA and those segments were replaced by newly constructed B-form double helical structures. These new observations could suggest that when the services of non-B DNA structures as cellular signals complete, living organisms would rely on the positive DNA supercoiling to disintegrate the stable physically altered DNA structures. Consequently, utilization of the complementary roles of negative supercoiling and positive supercoiling in a sequential order could likely be an economical and efficient way for living organisms to manipulate non-B structures to act as cellular signals for their DNA transactions.

References

1. Sinden, R.R. (1994) *DNA structure and function*. Academic Press, San Diego.
2. Zull, J.E. and Smith, S.K. (1990) Is genetic code redundancy related to retention of structural information in both DNA strands? *Trends Biochem Sci*, **15**, 257-261.
3. Bates, A.D. and Maxwell, A. (2005) *DNA topology*. 2nd ed. Oxford University Press, Oxford ; New York.
4. Champoux, J.J. (2001) DNA topoisomerases: structure, function, and mechanism. *Annual review of biochemistry*, **70**, 369-413.
5. Watson, J.D. and Crick, F.H. (1953) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*, **171**, 737-738.
6. Watson, J.D. and Crick, F.H. (2003) A structure for deoxyribose nucleic acid. 1953. *Nature*, **421**, 397-398; discussion 396.
7. Cozzarelli, N.R. and Wang, J.C. (1990) *DNA topology and its biological effects*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
8. Brazda, V., Laister, R.C., Jagelska, E.B. and Arrowsmith, C. (2011) Cruciform structures are a common DNA feature important for regulating biological processes. *BMC Mol Biol*, **12**, 33.
9. Schreck, J.S., Ouldrige, T.E., Romano, F., Louis, A.A. and Doye, J.P. (2015) Characterizing the bending and flexibility induced by bulges in DNA duplexes. *J Chem Phys*, **142**, 165101.
10. Simonsson, T. (2001) G-quadruplex DNA structures--variations on a theme. *Biological chemistry*, **382**, 621-628.
11. Vollmer, S. and Richert, C. (2015) DNA Triplexes That Bind Several Cofactor Molecules. *Chemistry*, **21**, 18613-18622.
12. Sinden, R.R., Pytlos-Sinden, M.J. and Potaman, V.N. (2007) Slipped strand DNA structures. *Frontiers in bioscience : a journal and virtual library*, **12**, 4788-4799.
13. Lipps, H.J. and Rhodes, D. (2009) G-quadruplex structures: in vivo evidence and function. *Trends in cell biology*, **19**, 414-422.
14. Lipps, H.J. and Rhodes, D. (2009) G-quadruplex structures: in vivo evidence and function. *Trends in cell biology*, **19**, 414-422.
15. Zhao, J.H., Bacolla, A., Wang, G.L. and Vasquez, K.M. (2010) Non-B DNA structure-induced genetic instability and evolution. *Cell Mol Life Sci*, **67**, 43-62.
16. Lv, L., Guo, Z., Wang, J. and Wang, E. (2012) G-quadruplex as signal transducer for biorecognition events. *Current pharmaceutical design*, **18**, 2076-2095.
17. Yu, H.Q., Miyoshi, D. and Sugimoto, N. (2006) Characterization of structure and stability of long telomeric DNA G-quadruplexes. *J Am Chem Soc*, **128**, 15461-15468.

18. Chargaff, E. (1951) Structure and function of nucleic acids as cell constituents. *Federation proceedings*, **10**, 654-659.
19. Zamenhof, S., Brawerman, G. and Chargaff, E. (1952) On the desoxypentose nucleic acids from several microorganisms. *Biochimica et biophysica acta*, **9**, 402-405.
20. Wilkins, M.H. and Randall, J.T. (1953) Crystallinity in sperm heads: molecular structure of nucleoprotein in vivo. *Biochimica et biophysica acta*, **10**, 192-193.
21. Wilkins, M.H., Stokes, A.R. and Wilson, H.R. (1953) Molecular structure of deoxypentose nucleic acids. *Nature*, **171**, 738-740.
22. Gorodetsky, A.A., Green, O., Yavin, E. and Barton, J.K. (2007) Coupling into the base pair stack is necessary for DNA-mediated electrochemistry. *Bioconjugate chemistry*, **18**, 1434-1441.
23. McCarty, M. (1994) A retrospective look: how we identified the pneumococcal transforming substance as DNA. *J Exp Med*, **179**, 385-394.
24. Case-Green, S.C., Mir, K.U., Pritchard, C.E. and Southern, E.M. (1998) Analysing genetic information with DNA arrays. *Curr Opin Chem Biol*, **2**, 404-410.
25. Wing, R., Drew, H., Takano, T., Broka, C., Tanaka, S., Itakura, K. and Dickerson, R.E. (1980) Crystal structure analysis of a complete turn of B-DNA. *Nature*, **287**, 755-758.
26. Gowers, D.M. and Fox, K.R. (1999) Towards mixed sequence recognition by triple helix formation. *Nucleic Acids Res*, **27**, 1569-1577.
27. Skuratovskii, I.Y., Volkova, L.I., Kapitonova, K.A. and Bartenev, V.N. (1979) Position of caesium ions in crystalline B-form DNA determined by synchrotron radiation diffraction. Preliminary results. *J Mol Biol*, **134**, 369-374.
28. Zavriev, S.K., Minchenkova, L.E., Frank-Kamenetskii, M.D. and Ivanov, V.I. (1978) On the flexibility of the boundaries between the A-form and B-form sections in DNA molecule. *Nucleic acids research*, **5**, 2657-2663.
29. Pabo, C.O., Jordan, S.R. and Frankel, A.D. (1983) Systematic analysis of possible hydrogen bonds between amino acid side chains and B-form DNA. *J Biomol Struct Dyn*, **1**, 1039-1049.
30. Dutta, S., Parrack, P.K. and Sasisekharan, V. (1984) Fibre diffraction of lithium DNA shows structural variability and deviation from a regular helical structure for the B-form. *Febs Lett*, **176**, 110-114.
31. Oakley, M.G., Mrksich, M. and Dervan, P.B. (1992) Evidence that a minor groove-binding peptide and a major groove-binding protein can simultaneously occupy a common site on DNA. *Biochemistry*, **31**, 10969-10975.
32. McBryant, S.J., Gedulin, B., Clemens, K.R., Wright, P.E. and Gottesfeld, J.M. (1996) Assessment of major and minor groove DNA interactions by the zinc fingers of *Xenopus* transcription factor IIIA. *Nucleic acids research*, **24**, 2567-2574.
33. Chen, L. and Frankel, A.D. (1995) A peptide interaction in the major groove of RNA resembles protein interactions in the minor groove of DNA. *Proceedings of the National Academy of Sciences of the United States of America*, **92**, 5077-5081.

34. Tuite, E., Sehlstedt, U., Hagmar, P., Norden, B. and Takahashi, M. (1997) Effects of minor and major groove-binding drugs and intercalators on the DNA association of minor groove-binding proteins RecA and deoxyribonuclease I detected by flow linear dichroism. *European journal of biochemistry / FEBS*, **243**, 482-492.
35. Nguyen-Hackley, D.H., Ramm, E., Taylor, C.M., Joung, J.K., Dervan, P.B. and Pabo, C.O. (2004) Allosteric inhibition of zinc-finger binding in the major groove of DNA by minor-groove binding ligands. *Biochemistry*, **43**, 3880-3890.
36. Arimondo, P.B., Bailly, C., Bourtoune, A.S., Ryabinin, V.A., Syniakov, A.N., Sun, J.S., Garestier, T. and Helene, C. (2001) Directing topoisomerase I mediated DNA cleavage to specific sites by camptothecin tethered to minor- and major-groove ligands. *Angewandte Chemie*, **40**, 3045-3048.
37. Shakked, Z., Rabinovich, D., Cruse, W.B., Egert, E., Kennard, O., Sala, G., Salisbury, S.A. and Viswamitra, M.A. (1981) Crystalline A-dna: the X-ray analysis of the fragment d(G-G-T-A-T-A-C-C). *Proc R Soc Lond B Biol Sci*, **213**, 479-487.
38. Pohl, F.M. and Jovin, T.M. (1972) Salt-induced co-operative conformational change of a synthetic DNA: equilibrium and kinetic studies with poly (dG-dC). *J Mol Biol*, **67**, 375-396.
39. Wang, J.C., Peck, L.J. and Becherer, K. (1983) DNA supercoiling and its effects on DNA structure and function. *Cold Spring Harb Symp Quant Biol*, **47 Pt 1**, 85-91.
40. Drew, H., Takano, T., Tanaka, S., Itakura, K. and Dickerson, R.E. (1980) High-salt d(CpGpCpG), a left-handed Z' DNA double helix. *Nature*, **286**, 567-573.
41. Klysik, J., Stirdivant, S.M., Larson, J.E., Hart, P.A. and Wells, R.D. (1981) Left-handed DNA in restriction fragments and a recombinant plasmid. *Nature*, **290**, 672-677.
42. Peck, L.J., Nordheim, A., Rich, A. and Wang, J.C. (1982) Flipping of cloned d(pCpG)n.d(pCpG)n DNA sequences from right- to left-handed helical structure by salt, Co(III), or negative supercoiling. *Proc Natl Acad Sci U S A*, **79**, 4560-4564.
43. Gerbi, S.A. and Bielinsky, A.K. (2002) DNA replication and chromatin. *Current opinion in genetics & development*, **12**, 243-248.
44. Jaunin, F. and Fakan, S. (2002) DNA replication and nuclear architecture. *J Cell Biochem*, **85**, 1-9.
45. Boyer, A.S., Walter, D. and Sorensen, C.S. (2016) DNA replication and cancer: From dysfunctional replication origin activities to therapeutic opportunities. *Seminars in cancer biology*.
46. Gorbunova, V. and Seluanov, A. (2016) DNA double strand break repair, aging and the chromatin connection. *Mutation research*.
47. Doksani, Y. and de Lange, T. (2014) The role of double-strand break repair pathways at functional and dysfunctional telomeres. *Cold Spring Harbor perspectives in biology*, **6**, a016576.

48. Srivastava, M. and Raghavan, S.C. (2015) DNA double-strand break repair inhibitors as cancer therapeutics. *Chemistry & biology*, **22**, 17-29.
49. Ceccaldi, R., Rondinelli, B. and D'Andrea, A.D. (2016) Repair Pathway Choices and Consequences at the Double-Strand Break. *Trends in cell biology*, **26**, 52-64.
50. Glickman, M.S. (2014) Double-Strand DNA Break Repair in Mycobacteria. *Microbiology spectrum*, **2**.
51. Swigon, D., Lim, S. and Kim, Y. (2013) Dynamical simulations of DNA supercoiling and compression. *Biochem Soc Trans*, **41**, 554-558.
52. Gilbert, N. and Allan, J. (2014) Supercoiling in DNA and chromatin. *Current opinion in genetics & development*, **25**, 15-21.
53. Akbuga, J., Aral, C., Ozbas-Turan, S., Kabasakal, L. and Keyer-Uysal, M. (2003) Transfection efficiency of chitosan microspheres: effect of DNA topology. *Stp Pharma Sci*, **13**, 99-103.
54. Dulbecco, R. and Vogt, M. (1963) Evidence for a Ring Structure of Polyoma Virus DNA. *Proc Natl Acad Sci U S A*, **50**, 236-243.
55. Lesko, S.A., Jr. and Emery, A.J., Jr. (1966) Subunit form of calf thymus DNA. *Biochem Biophys Res Commun*, **23**, 707-712.
56. Dorman, C.J. and Corcoran, C.P. (2009) Bacterial DNA topology and infectious disease. *Nucleic acids research*, **37**, 672-678.
57. Huang, W.M. (1996) Bacterial diversity based on type II DNA topoisomerase genes. *Annu Rev Genet*, **30**, 79-107.
58. Li, D., Lv, B., Zhang, H., Lee, J.Y. and Li, T. (2014) Positive supercoiling affiliated with nucleosome formation repairs non-B DNA structures. *Chemical communications*, **50**, 10641-10644.
59. Pohl, W.F. (1968) Self-Linking Number of a Closed Space Curve. *J Math Mech*, **17**, 975-&.
60. White, J.H. (1969) Self-Linking and Gauss-Integral in Higher Dimensions. *Am J Math*, **91**, 693-&.
61. Fuller, F.B. (1971) The writhing number of a space curve. *Proc Natl Acad Sci U S A*, **68**, 815-819.
62. Crick, F.H.C. (1976) Linking Numbers and Nucleosomes. *P Natl Acad Sci USA*, **73**, 2639-2643.
63. Fuller, F.B. (1978) Decomposition of the linking number of a closed ribbon: A problem from molecular biology. *Proceedings of the National Academy of Sciences of the United States of America*, **75**, 3557-3561.
64. Miller, D. and Benham, C. (1996) Fixed-writhe isotopies and the topological conservation law for closed, circular DNA. *J Knot Theor Ramif*, **5**, 859-866.
65. Benham, C. and Miller, D. (2000) Writhing and linking densities for closed, circular DNA. *J Knot Theor Ramif*, **9**, 577-585.
66. Cozzarelli, N.R. and Wang, J.C. (1990) *DNA topology and its biological effects*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
67. Levene, S.D. (2007) Topology in biology: From DNA mechanics to enzymology. *Biol Med Phys Biomed*, 3-21.

68. Schwartzman, J.B., Martinez-Robles, M.L., Hernandez, P. and Krimer, D.B. (2013) The benefit of DNA supercoiling during replication. *Biochem Soc Trans*, **41**, 646-651.
69. Vos, S.M., Tretter, E.M., Schmidt, B.H. and Berger, J.M. (2011) All tangled up: how cells direct, manage and exploit topoisomerase function. *Nature reviews. Molecular cell biology*, **12**, 827-841.
70. Thumm, W., Seidl, A. and Hinz, H.J. (1988) Energy-structure correlations of plasmid DNA in different topological forms. *Nucleic acids research*, **16**, 11737-11757.
71. Schwartzman, J.B., Martinez-Robles, M.L., Hernandez, P. and Krimer, D.B. (2010) Plasmid DNA replication and topology as visualized by two-dimensional agarose gel electrophoresis. *Plasmid*, **63**, 1-10.
72. Delatour, C.B., Portemer, C., Huber, R., Forterre, P. and Duguet, M. (1991) Reverse Gyrase in Thermophilic Eubacteria. *J Bacteriol*, **173**, 3921-3923.
73. Kovalsky, O.I., Kozyavkin, S.A. and Slesarev, A.I. (1990) Archaeobacterial Reverse Gyrase Cleavage-Site Specificity Is Similar to That of Eubacterial DNA Topoisomerases-I. *Nucleic acids research*, **18**, 2801-2805.
74. Forterre, P., Bouthier De La Tour, C., Philippe, H. and Duguet, M. (2000) Reverse gyrase from hyperthermophiles: probable transfer of a thermoadaptation trait from archaea to bacteria. *Trends in genetics : TIG*, **16**, 152-154.
75. Hayashi, M., Hayashi, M.N. and Spiegelmans. (1963) Replicating form of a single-stranded DNA virus: isolation and properties. *Science*, **140**, 1313-1316.
76. Wells, R.D. (2007) Non-B DNA conformations, mutagenesis and disease. *Trends Biochem Sci*, **32**, 271-278.
77. Wells, R.D. (2009) Discovery of the role of non-B DNA structures in mutagenesis and human genomic disorders. *J Biol Chem*, **284**, 8997-9009.
78. Du, Y. and Zhou, X. (2013) Targeting non-B-form DNA in living cells. *Chem Rec*, **13**, 371-384.
79. Li, D., Lv, B., Zhang, H., Lee, J.Y. and Li, T. (2014) Positive supercoiling affiliated with nucleosome formation repairs non-B DNA structures. *Chemical communications*, **50**, 10641-10644.
80. Du, Y. and Zhou, X. (2013) Targeting non-B-form DNA in living cells. *Chem Rec*, **13**, 371-384.
81. van Holde, K. and Zlatanova, J. (1994) Unusual DNA structures, chromatin and transcription. *BioEssays : news and reviews in molecular, cellular and developmental biology*, **16**, 59-68.
82. Bochman, M.L., Paeschke, K. and Zakian, V.A. (2012) DNA secondary structures: stability and function of G-quadruplex structures. *Nat Rev Genet*, **13**, 770-780.
83. McFadden, G. and Morgan, A.R. (1982) DNA cruciform structures: implications for telomere replication in eukaryotes and instability of long palindromic DNA sequences in prokaryotes. *J Theor Biol*, **97**, 343-349.
84. Tenen, D.G., Haines, L.L., Hansen, U.M., Martin, R.G. and Livingston, D.M. (1985) Formation of a cruciform structure at the simian virus 40 replication origin abolishes T-antigen binding to the origin in vitro. *J Virol*, **56**, 293-297.

85. Belmont, P., Constant, J.F. and Demeunynck, M. (2001) Nucleic acid conformation diversity: from structure to function and regulation. *Chem Soc Rev*, **30**, 70-81.
86. Franceschin, M. (2009) G-Quadruplex DNA Structures and Organic Chemistry: More Than One Connection. *Eur J Org Chem*, 2225-2238.
87. Wang, G.L. and Vasquez, K.M. (2004) Naturally occurring H-DNA-forming sequences are mutagenic in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 13448-13453.
88. Dey, I. and Rath, P.C. (2005) A novel rat genomic simple repeat DNA with RNA-homology shows triplex (H-DNA)-like structure and tissue-specific RNA expression. *Biochem Bioph Res Co*, **327**, 276-286.
89. Bacolla, A., Wojciechowska, M., Kosmider, B., Larson, J.E. and Wells, R.D. (2006) The involvement of non-B DNA structures in gross chromosomal rearrangements. *DNA repair*, **5**, 1161-1170.
90. Neidle, S. (2010) Human telomeric G-quadruplex: the current status of telomeric G-quadruplexes as therapeutic targets in human cancer. *FEBS J*, **277**, 1118-1125.
91. Huppert, J.L. (2008) Four-stranded nucleic acids: structure, function and targeting of G-quadruplexes. *Chem Soc Rev*, **37**, 1375-1384.
92. Bacolla, A. and Wells, R.D. (2004) Non-B DNA conformations, genomic rearrangements, and human disease. *J Biol Chem*, **279**, 47411-47414.
93. Haider, S.M., Neidle, S. and Parkinson, G.N. (2011) A structural analysis of G-quadruplex/ligand interactions. *Biochimie*, **93**, 1239-1251.
94. Tran, P.L., Mergny, J.L. and Alberti, P. (2011) Stability of telomeric G-quadruplexes. *Nucleic acids research*, **39**, 3282-3294.
95. Huppert, J.L. and Balasubramanian, S. (2005) Prevalence of quadruplexes in the human genome. *Nucleic acids research*, **33**, 2908-2916.
96. Hurley, L.H., Von Hoff, D.D., Siddiqui-Jain, A. and Yang, D. (2006) Drug targeting of the c-MYC promoter to repress gene expression via a G-quadruplex silencer element. *Seminars in oncology*, **33**, 498-512.
97. Sun, D. and Hurley, L.H. (2009) The importance of negative superhelicity in inducing the formation of G-quadruplex and i-motif structures in the c-Myc promoter: implications for drug targeting and control of gene expression. *Journal of medicinal chemistry*, **52**, 2863-2874.
98. Evans, T., Schon, E., Gora-Maslak, G., Patterson, J. and Efstratiadis, A. (1984) S1-hypersensitive sites in eukaryotic promoter regions. *Nucleic Acids Res*, **12**, 8043-8058.
99. Lam, E.Y., Beraldi, D., Tannahill, D. and Balasubramanian, S. (2013) G-quadruplex structures are stable and detectable in human genomic DNA. *Nat Commun*, **4**, 1796.
100. Capra, J.A., Paeschke, K., Singh, M. and Zakian, V.A. (2010) G-quadruplex DNA sequences are evolutionarily conserved and associated with distinct genomic features in *Saccharomyces cerevisiae*. *PLoS computational biology*, **6**, e1000861.
101. Holt, I.J. and Reyes, A. (2012) Human mitochondrial DNA replication. *Cold Spring Harbor perspectives in biology*, **4**.

102. Li, Y. and Araki, H. (2013) Loading and activation of DNA replicative helicases: the key step of initiation of DNA replication. *Genes Cells*, **18**, 266-277.
103. Balakrishnan, L. and Bambara, R.A. (2013) Okazaki fragment metabolism. *Cold Spring Harbor perspectives in biology*, **5**.
104. Levy, M.Z., Allsopp, R.C., Futcher, A.B., Greider, C.W. and Harley, C.B. (1992) Telomere end-replication problem and cell aging. *J Mol Biol*, **225**, 951-960.
105. Lingner, J., Cooper, J.P. and Cech, T.R. (1995) Telomerase and DNA end replication: no longer a lagging strand problem? *Science*, **269**, 1533-1534.
106. Lundblad, V. (1997) The end replication problem: more than one solution. *Nat Med*, **3**, 1198-1199.
107. Chaires, J.B. (2010) Human telomeric G-quadruplex: thermodynamic and kinetic studies of telomeric quadruplex stability. *Febs J*, **277**, 1098-1106.
108. Granotier, C., Pennarun, G., Riou, L., Hoffschir, F., Gauthier, L.R., De Cian, A., Gomez, D., Mandine, E., Riou, J.F., Mergny, J.L. *et al.* (2005) Preferential binding of a G-quadruplex ligand to human chromosome ends. *Nucleic acids research*, **33**, 4182-4190.
109. Phan, A.T. (2010) Human telomeric G-quadruplex: structures of DNA and RNA sequences. *Febs J*, **277**, 1107-1117.
110. Kong, C.M., Lee, X.W. and Wang, X. (2013) Telomere shortening in human diseases. *Febs J*, **280**, 3180-3193.
111. Mikhelson, V.M. and Gamaley, I.A. (2012) Telomere shortening is a sole mechanism of aging in mammals. *Curr Aging Sci*, **5**, 203-208.
112. Edwards, P.A. (2010) Fusion genes and chromosome translocations in the common epithelial cancers. *J Pathol*, **220**, 244-254.
113. Rauscher, F.J., 3rd. (1997) Chromosome translocation-mediated conversion of a tumor suppressor gene into a dominant oncogene: fusion of EWS1 to WT1 in desmoplastic small round cell tumors. *Curr Top Microbiol Immunol*, **220**, 151-162.
114. Cesare, A.J. and Karlseder, J. (2012) A three-state model of telomere control over human proliferative boundaries. *Curr Opin Cell Biol*, **24**, 731-738.
115. Eisenberg, D.T. (2011) An evolutionary review of human telomere biology: the thrifty telomere hypothesis and notes on potential adaptive paternal effects. *Am J Hum Biol*, **23**, 149-167.
116. Patel, D.J., Phan, A.T. and Kuryavyi, V. (2007) Human telomere, oncogenic promoter and 5'-UTR G-quadruplexes: diverse higher order DNA and RNA targets for cancer therapeutics. *Nucleic acids research*, **35**, 7429-7455.
117. Huppert, J.L. and Balasubramanian, S. (2007) G-quadruplexes in promoters throughout the human genome. *Nucleic acids research*, **35**, 406-413.
118. Siddiqui-Jain, A., Grand, C.L., Bearss, D.J. and Hurley, L.H. (2002) Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 11593-11598.

119. Miyoshi, D., Matsumura, S., Nakano, S. and Sugimoto, N. (2004) Duplex dissociation of telomere DNAs induced by molecular crowding. *J Am Chem Soc*, **126**, 165-169.
120. Karimata, H., Miyoshi, D. and Sugimoto, N. (2005) Structure and stability of DNA quadruplexes under molecular crowding conditions. *Nucleic Acids Symp Ser (Oxf)*, 239-240.
121. Kan, Z.Y., Lin, Y., Wang, F., Zhuang, X.Y., Zhao, Y., Pang, D.W., Hao, Y.H. and Tan, Z. (2007) G-quadruplex formation in human telomeric (TTAGGG)₄ sequence with complementary strand in close vicinity under molecularly crowded condition. *Nucleic acids research*, **35**, 3646-3653.
122. Sugimoto, N. (2014) Noncanonical structures and their thermodynamics of DNA and RNA under molecular crowding: beyond the Watson-Crick double helix. *International review of cell and molecular biology*, **307**, 205-273.
123. Miyoshi, D., Fujimoto, T. and Sugimoto, N. (2013) Molecular crowding and hydration regulating of G-quadruplex formation. *Top Curr Chem*, **330**, 87-110.
124. Moens, P.B. (2003) The double-stranded DNA helix in recombination at meiosis. *Genome*, **46**, 936-937.
125. Chin, J.Y., Schleifman, E.B. and Glazer, P.M. (2007) Repair and recombination induced by triple helix DNA. *Front Biosci*, **12**, 4288-4297.
126. Murakami, Y. (2003) Studies on the structure of double helix DNA in the last half century: Changes in the nature of scientific research. *Tanpakushitsu Kakusan Koso*, **48**, 698-699.
127. Ogawa, H. (2003) Studies on the structure of double-helix DNA in the last half century: DNA tail. *Tanpakushitsu Kakusan Koso*, **48**, 678-680.
128. Voineagu, I., Freudenreich, C.H. and Mirkin, S.M. (2009) Checkpoint Responses to Unusual Structures Formed by DNA Repeats. *Molecular Carcinogenesis*, **48**, 309-318.
129. Mikheikin, A.L., Lushnikov, A.Y. and Lyubchenko, Y.L. (2006) Effect of DNA supercoiling on the geometry of holliday junctions. *Biochemistry*, **45**, 12998-13006.
130. Blackburn, G.M. (2006) *Nucleic acids in chemistry and biology*. 3rd ed. RSC Pub., Cambridge, UK.
131. Bacolla, A. and Wells, R.D. (2009) Non-B DNA Conformations as Determinants of Mutagenesis and Human Disease. *Molecular Carcinogenesis*, **48**, 273-285.
132. Wang, X.L., Huang, J., Zhou, Y.Y., Yan, S.Y., Weng, X.C., Wu, X.J., Deng, M.G. and Zhou, X.A. (2010) Conformational Switching of G-Quadruplex DNA by Photoregulation. *Angew Chem Int Edit*, **49**, 5305-5309.
133. Vummidi, B.R., Alzeer, J. and Luedtke, N.W. (2013) Fluorescent probes for G-quadruplex structures. *Chembiochem*, **14**, 540-558.
134. Zhou, J., Wei, C., Jia, G., Wang, X., Feng, Z. and Li, C. (2009) Human telomeric G-quadruplex formed from duplex under near physiological conditions: spectroscopic evidence and kinetics. *Biochimie*, **91**, 1104-1111.
135. Li, T., Liu, D., Chen, J., Lee, A.H., Qi, J. and Chan, A.S. (2001) Construction of circular oligodeoxyribonucleotides on the new structural basis of i-motif. *J Am Chem Soc*, **123**, 12901-12902.

136. Lane, A.N., Chaires, J.B., Gray, R.D. and Trent, J.O. (2008) Stability and kinetics of G-quadruplex structures. *Nucleic acids research*, **36**, 5482-5515.
137. Todd, A.K., Johnston, M. and Neidle, S. (2005) Highly prevalent putative quadruplex sequence motifs in human DNA. *Nucleic Acids Res*, **33**, 2901-2907.
138. Maizels, N. (2006) Dynamic roles for G4 DNA in the biology of eukaryotic cells. *Nat Struct Mol Biol*, **13**, 1055-1059.
139. Neidle, S. and Parkinson, G.N. (2003) The structure of telomeric DNA. *Current opinion in structural biology*, **13**, 275-283.
140. Morris, M.J., Wingate, K.L., Silwal, J., Leeper, T.C. and Basu, S. (2012) The porphyrin TmPyP4 unfolds the extremely stable G-quadruplex in MT3-MMP mRNA and alleviates its repressive effect to enhance translation in eukaryotic cells. *Nucleic Acids Res*, **40**, 4137-4145.
141. Hatfield, G.W. and Benham, C.J. (2002) DNA topology-mediated control of global gene expression in Escherichia coli. *Annu Rev Genet*, **36**, 175-203.
142. Rawal, P., Kummaraleti, V.B.R., Ravindran, J., Kumar, N., Halder, K., Sharma, R., Mukerji, M., Das, S.K. and Chowdhury, S. (2006) Genome-wide prediction of G4 DNA as regulatory motifs: Role in Escherichia coli global regulation. *Genome Res*, **16**, 644-655.
143. Wang, J.C. and Kirkegaard, K. (1981) DNA topoisomerases. *Gene Amplif Anal*, **2**, 455-473.
144. Furuyama, T. and Henikoff, S. (2009) Centromeric nucleosomes induce positive DNA supercoils. *Cell*, **138**, 104-113.
145. Reece, R.J. and Maxwell, A. (1991) DNA gyrase: structure and function. *Crit Rev Biochem Mol Biol*, **26**, 335-375.
146. Cozzarelli, N.R. (1980) DNA gyrase and the supercoiling of DNA. *Science*, **207**, 953-960.
147. Wang, J.C. and Kirkegaard, K. (1981) DNA topoisomerases. *Gene Amplif Anal*, **2**, 455-473.
148. Li, D., Yang, Z., Zhao, G., Long, Y., Lv, B., Li, C., Hiew, S., Ng, M.T.T., Guo, J., Tan, H. *et al.* (2011) Manipulating DNA writhe through varying DNA sequences. *Chemical communications*, **47**, 7479-7481.
149. Li, D.W., Yang, Z.Q., Long, Y., Zhao, G., Lv, B., Hiew, S., Magdeline, T.T.N., Guo, J.J., Tan, H., Zhang, H. *et al.* (2011) Precise engineering and visualization of signs and magnitudes of DNA writhe on the basis of PNA invasion. *Chemical communications*, **47**, 10695-10697.
150. Lyubchenko, Y.L. and Shlyakhtenko, L.S. (2009) AFM for analysis of structure and dynamics of DNA and protein-DNA complexes. *Methods*, **47**, 206-213.
151. Lyubchenko, Y.L., Shlyakhtenko, L.S. and Gall, A.A. (2009) Atomic force microscopy imaging and probing of DNA, proteins, and protein DNA complexes: silatrane surface chemistry. *Methods Mol Biol*, **543**, 337-351.
152. Bates, A.D., Noy, A., Piperakis, M.M., Harris, S.A. and Maxwell, A. (2013) Small DNA circles as probes of DNA topology. *Biochem Soc T*, **41**, 565-570.
153. Li, D.W., Yang, Z.Q., Zhao, G.J., Long, Y., Lv, B., Li, C., Hiew, S., Ng, M.T.T., Guo, J.J., Tan, H. *et al.* (2011) Manipulating DNA writhe through varying DNA sequences. *Chemical communications*, **47**, 7479-7481.

154. Yang, Z.Q., Li, D.W. and Li, T.H. (2011) Design and synthesis of catenated rings based on toroidal DNA structures. *Chemical communications*, **47**, 11930-11932.
155. Erlich, H.A. (1989) Polymerase chain reaction. *J Clin Immunol*, **9**, 437-447.
156. Lyubchenko, Y.L., Shlyakhtenko, L.S., Aki, T. and Adhya, S. (1997) Atomic force microscopic demonstration of DNA looping by GalR and HU. *Nucleic Acids Res*, **25**, 873-876.
157. Lyubchenko, Y.L., Gall, A.A., Shlyakhtenko, L.S., Harrington, R.E., Jacobs, B.L., Oden, P.I. and Lindsay, S.M. (1992) Atomic force microscopy imaging of double stranded DNA and RNA. *J Biomol Struct Dyn*, **10**, 589-606.
158. Shlyakhtenko, L.S., Miloskeska, L., Potaman, V.N., Sinden, R.R. and Lyubchenko, Y.L. (2003) Intersegmental interactions in supercoiled DNA: atomic force microscope study. *Ultramicroscopy*, **97**, 263-270.
159. Lyubchenko, Y.L., Oden, P.I., Lampner, D., Lindsay, S.M. and Dunker, K.A. (1993) Atomic force microscopy of DNA and bacteriophage in air, water and propanol: the role of adhesion forces. *Nucleic Acids Res*, **21**, 1117-1123.
160. Neaves, K.J., Huppert, J.L., Henderson, R.M. and Edwardson, J.M. (2009) Direct visualization of G-quadruplexes in DNA using atomic force microscopy. *Nucleic acids research*, **37**, 6269-6275.
161. Kahn, W.J.D. (2002) DNA topology: Applications to gene expression. *Curr Org Chem*, **6**, 815-826.
162. Monastyrskii, M.I.i. (2007) *Topology in molecular biology*. Springer, Berlin ; New York.
163. McLaggan, D., Naprstek, J., Buurman, E.T. and Epstein, W. (1994) Interdependence of K⁺ and glutamate accumulation during osmotic adaptation of Escherichia coli. *J Biol Chem*, **269**, 1911-1917.
164. Douglas, R.M., Ritchie, G.Y., Munro, A.W., McLaggan, D. and Booth, I.R. (1994) The K(+) -efflux system, KefC, in Escherichia coli: genetic evidence for oligomeric structure. *Mol Membr Biol*, **11**, 55-61.
165. Luoto, H.H., Baykov, A.A., Lahti, R. and Malinen, A.M. (2013) Membrane-integral pyrophosphatase subfamily capable of translocating both Na⁺ and H⁺. *Proc Natl Acad Sci U S A*, **110**, 1255-1260.
166. Yuan, Z., Cai, T., Tian, J., Ivanov, A.V., Giovannucci, D.R. and Xie, Z. (2005) Na/K-ATPase tethers phospholipase C and IP₃ receptor into a calcium-regulatory complex. *Mol Biol Cell*, **16**, 4034-4045.
167. Miyoshi, D. and Sugimoto, N. (2008) Molecular crowding effects on structure and stability of DNA. *Biochimie*, **90**, 1040-1051.
168. Zheng, K.W., Chen, Z., Hao, Y.H. and Tan, Z. (2010) Molecular crowding creates an essential environment for the formation of stable G-quadruplexes in long double-stranded DNA. *Nucleic acids research*, **38**, 327-338.
169. Lord, C.J. and Ashworth, A. (2012) The DNA damage response and cancer therapy. *Nature*, **481**, 287-294.
170. Soutoglou, E. and Misteli, T. (2008) Activation of the cellular DNA damage response in the absence of DNA lesions. *Science*, **320**, 1507-1510.
171. Kottemann, M.C. and Smogorzewska, A. (2013) Fanconi anaemia and the repair of Watson and Crick DNA crosslinks. *Nature*, **493**, 356-363.

172. Maugeri-Sacca, M., Bartucci, M. and De Maria, R. (2012) DNA damage repair pathways in cancer stem cells. *Molecular cancer therapeutics*, **11**, 1627-1636.
173. Melis, J.P., van Steeg, H. and Luijten, M. (2013) Oxidative DNA damage and nucleotide excision repair. *Antioxidants & redox signaling*, **18**, 2409-2419.
174. Surova, O. and Zhivotovsky, B. (2013) Various modes of cell death induced by DNA damage. *Oncogene*, **32**, 3789-3797.
175. Heil, K., Pearson, D. and Carell, T. (2011) Chemical investigation of light induced DNA bipyrimidine damage and repair. *Chem Soc Rev*, **40**, 4271-4278.
176. Murchie, A.I. and Lilley, D.M. (1992) Supercoiled DNA and cruciform structures. *Methods in enzymology*, **211**, 158-180.
177. Son, L.S., Bacolla, A. and Wells, R.D. (2006) Sticky DNA: in vivo formation in E-coli and in vitro association of long GAA center dot TTC tracts to generate two independent supercoiled domains. *J Mol Biol*, **360**, 267-284.
178. Wood, R.D. (1996) DNA repair in eukaryotes. *Annual review of biochemistry*, **65**, 135-167.
179. Sancar, A., Lindsey-Boltz, L.A., Unsal-Kacmaz, K. and Linn, S. (2004) Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annual review of biochemistry*, **73**, 39-85.
180. Singleton, M.R., Dillingham, M.S. and Wigley, D.B. (2007) Structure and mechanism of helicases and nucleic acid translocases. *Annual Review of Biochemistry*, **76**, 23-50.
181. Vincent, M., Xu, Y. and Kong, H.M. (2004) Helicase-dependent isothermal DNA amplification. *Embo Rep*, **5**, 795-800.
182. Jeong, Y.J., Park, K. and Kim, D.E. (2009) Isothermal DNA amplification in vitro: the helicase-dependent amplification system. *Cell Mol Life Sci*, **66**, 3325-3336.
183. Valenti, A., Perugino, G., Rossi, M. and Ciaramella, M. (2011) Positive supercoiling in thermophiles and mesophiles: of the good and evil. *Biochem Soc Trans*, **39**, 58-63.
184. Leach, D.R. (1994) Long DNA palindromes, cruciform structures, genetic instability and secondary structure repair. *BioEssays : news and reviews in molecular, cellular and developmental biology*, **16**, 893-900.
185. Dayn, A., Malkhosyan, S. and Mirkin, S.M. (1992) Transcriptionally driven cruciform formation in vivo. *Nucleic acids research*, **20**, 5991-5997.
186. Brazda, V., Laister, R.C., Jagelska, E.B. and Arrowsmith, C. (2011) Cruciform structures are a common DNA feature important for regulating biological processes. *Bmc Mol Biol*, **12**, 33.
187. Li, D., Yang, Z., Long, Y., Zhao, G., Lv, B., Hiew, S., Ng, M.T., Guo, J., Tan, H., Zhang, H. *et al.* (2011) Precise engineering and visualization of signs and magnitudes of DNA writhe on the basis of PNA invasion. *Chemical communications*, **47**, 10695-10697.
188. Lv, B., Li, D., Zhang, H., Lee, J.Y. and Li, T. (2013) DNA gyrase-driven generation of a G-quadruplex from plasmid DNA. *Chemical communications*, **49**, 8317-8319.
189. Panayotatos, N. and Wells, R.D. (1981) Cruciform structures in supercoiled DNA. *Nature*, **289**, 466-470.

190. Lilley, D.M., Gough, G.W., Hallam, L.R. and Sullivan, K.M. (1985) The physical chemistry of cruciform structures in supercoiled DNA molecules. *Biochimie*, **67**, 697-706.
191. Sullivan, K.M. and Lilley, D.M. (1987) Influence of cation size and charge on the extrusion of a salt-dependent cruciform. *J Mol Biol*, **193**, 397-404.
192. Singleton, C.K. (1983) Effects of salts, temperature, and stem length on supercoil-induced formation of cruciforms. *The Journal of biological chemistry*, **258**, 7661-7668.
193. Pavlov, A.R., Pavlova, N.V., Kozyavkin, S.A. and Slesarev, A.I. (2004) Recent developments in the optimization of thermostable DNA polymerases for efficient applications. *Trends Biotechnol*, **22**, 253-260.
194. Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, **239**, 487-491.
195. Li, D., Yang, Z., Lv, B. and Li, T. (2012) Observation of backbone self-crossings of organismal DNAs through atomic force microscopy. *Bioorg Med Chem Lett*, **22**, 833-836.
196. Vologodskaya, M.Y. and Vologodskii, A.V. (1999) Effect of magnesium on cruciform extrusion in supercoiled DNA. *J Mol Biol*, **289**, 851-859.
197. Paeschke, K., Bochman, M.L., Garcia, P.D., Cejka, P., Friedman, K.L., Kowalczykowski, S.C. and Zakian, V.A. (2013) Pif1 family helicases suppress genome instability at G-quadruplex motifs. *Nature*, **497**, 458-462.
198. Bochman, M.L., Judge, C.P. and Zakian, V.A. (2011) The Pif1 family in prokaryotes: what are our helicases doing in your bacteria? *Molecular biology of the cell*, **22**, 1955-1959.

List of Publications

1. **Lv, B.;** Dai, Y; Liu, J; Zhuge, Q; Li, D.; The Effect of Dimethyl Sulfoxide on Supercoiled DNA Relaxation Catalyzed by Type I Topoisomerases, *BioMed Research International*, vol. 2015, Article ID 320490.
2. Li, D.; **Lv, B. (co-first author);** Zhang, H.; Lee, J. Y.; Li, T., Disintegration of cruciform and G-quadruplex structures during the course of helicase-dependent amplification (HDA). *Bioorg Med Chem Lett* 2015, 25, 1709-1714.
3. Li, D.; **Lv, B.;** Zhang, H.; Lee, J. Y.; Li, T., Positive supercoiling affiliated with nucleosome formation repairs non-B DNA structures. *Chem Commun*, 2014, 50, 10641-10644.
4. **Lv, B.;** Li, D. W.; Zhang, H.; Li, Y. Q. J; Li, T. H., DNA gyrase-driven generation of a G-quadruplex from plasmid DNA. *Chem Commun*. 2013, 49, 8317-8319. (Highlighted on cover)
5. Zhang, H; Guo, J. J. Li, D. W.; Magdeline, T. T. N; Lee, J. Y.; **Lv, B.;** Ng, C. W.; Selvi Lee; Shao, F. W.; Li, T. H. Confirmation of quinolone-induced formation of gyrase–DNA conjugates using AFM. *Bioorg Med Chem Lett* 2013, 23, 4622-4626.
6. Li, D. W.; Yang, Z. Q.; **Lv, B.;** Li, T. H., Observation of backbone self-crossings of organismal DNAs through atomic force microscopy. *Bioorg Med Chem Lett* 2012, 22, 833-836.
7. Li, D. W.; Yang, Z. Q.; Long, Y.; Zhao, G.; **Lv, B.;** Hiew, S.; Magdeline, T. T. N.; Guo, J. J.; Tan, H.; Zhang, H.; Yuan, W. X.; Su, H. B.; Li, T. H. , Precise engineering and visualization of signs and magnitudes of DNA writhe on the basis of PNA invasion. *Chem Commun* 2011, 47, 10695-10697.
8. Li, D. W.; Yang, Z. Q.; Zhao, G. J.; Long, Y.; **Lv, B.;** Li, C.; Hiew, S.; Ng, M. T. T.; Guo, J. J.; Tan, H.; Zhang, H.; Li, T. H., Manipulating DNA writhe through varying DNA sequences. *Chem Commun* 2011, 47, 7479-7481.