Design, synthesis, characterization and property study of topological structures of DNA

Li, Dawei

2012

Li, D. (2012). Design, synthesis, characterization and property study of topological structures of DNA. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/51096>

<https://doi.org/10.32657/10356/51096>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

Design, synthesis, characterization and property study
of topological structures of DNA

LI DAWEI

**Design, synthesis, characterization and property study
of topological structures of DNA**

LI DAWEI

SCHOOL OF PHYSICAL AND MATHEMATICAL SCIENCES

2012

2012

**Design, synthesis, characterization and property study
of topological structures of DNA**

LI DAWEI

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

2012

Acknowledgements

It would not have been possible to write this doctoral thesis without the help and support of the kind people around me, to only some of whom it is possible to give particular mention here.

First and foremost, I would like to express my deep and sincere gratitude to my supervisor, Professor Li Tianhu, for his support, continuous guidance, meticulous suggestions and astute criticism during my graduate study in Nanyang Technological University. His unsurpassed knowledge and logical way of thinking have been of great value for me. The joy and enthusiasm he has for his research was contagious and motivational for me all the time. I am also thankful for the excellent example he has set as a successful chemist and professor.

I am most grateful to Dr. Yang Zhaoqi for helping me with important comments and suggestions and thanks to all other lab colleagues (Long Yi, Zhao Guanjia, Tan Hong Kee, Zhang Hao, Li Yiqin Jasmine, Hiew Shu Hui, Ng Tao Tao Magdeline, Li Cheng, Lei Qiong and Ba Sai) for their scientific help during my graduate study. I am also thankful to all my friends in Singapore, China and elsewhere for their support and encouragement throughout.

I would like to acknowledge all support staff in administrative office, teaching lab and the chemical store. I am also indebted to Dr. Wanxin Sun (Bruker, Singapore) for his assistance on the AFM sample determination.

In addition, the financial, academic and technical support of the Nanyang Technological University and Ministry of Education in Singapore is gratefully acknowledged.

Last but not least, I owe my loving thanks to my wife Lv Bei. Without her encouragement and understanding, it would have been impossible for me to finish my study. I also would like to give my deepest gratitude to my parents, who made me who I am. I will wish to dedicate this dissertation to them.

Table of Contents

Acknowledgements.....	I
Table of Contents.....	III
Abstract.....	VIII
List of Tables.....	XI
List of Figures.....	XII
List of Abbreviation.....	XVIII

Chapter 1 – Introduction

1.1 The Watson-Crick Model and B-form DNA.....	2
1.2 DNA Supercoiling and DNA Topological Conservation Law.....	8
1.2.1 Closed Circular DNA and DNA Supercoiling.....	8
1.2.2 Mathematical Expression of DNA Supercoiling: DNA Topological Conservation Law.....	11
1.3 DNA Bending.....	15
1.3.1 Intrinsic Curvature of DNA: Wedge Model or Junction Model?.....	16
1.3.2 DNA Flexibility: Forcible curvature of DNA as well as Protein-induced DNA bending.....	18
1.4 Alternative Conformations of DNA.....	20
1.4.1 Cruciform Structures in DNA.....	21
1.4.2 Four-Strands Nucleic Acids: G-quadruplexes.....	23

Chapter 2 – Manipulating DNA Writhe through Varying DNA

Sequences and General Topological Conservation Law of DNA

2.1 Introduction.....	27
-----------------------	----

2.2 Design of DNA Sequences.....	30
2.2.1 Design of Interwound Structures of DNA that Possesses Writhe Number of ± 1	30
2.2.2 Design of Toroidal Structures of DNA.....	33
2.2.3 Design of Double Interwound Structures of DNA.....	37
2.2.4 The DNA Sequence of Decatenated Kinetoplast DNA Minicircles.....	40
2.2.5 Design of Plasmid DNA Containing the DNA Sequence of kinetoplast DNA.....	42
2.2.6 Design of Plasmid DNA Containing the Replication Origins of Bacteriophage λ	44
2.3 Materials and Methods.....	49
2.3.1 Duplex DNA, Enzymes and Chemicals.....	49
2.3.2 Reactions of SacI with Duplex Linear DNA Precursors.....	50
2.3.3 Preparations of Circular DNA Using T4 Ligase.....	51
2.3.4 Degrade Linear DNA from Ligase Reaction Mixture Using Nuclease BAL-31 Exonuclease.....	51
2.3.5 Digest Circular DNA by SacI Endonuclease.....	52
2.3.6 Reactions of Human Topoisomerase II with Catenated Kinetoplast DNA.....	52
2.3.7 Reactions of Human Topoisomerase I with Circular DNA and Plasmid DNA.....	53
2.3.8 AFM Examination of Obtained Circular DNA.....	53
2.4 Results and Discussion.....	57
2.4.1 Synthesis and Confirmation of Interwound Structures of DNA that Possesses Writhe Number of ± 1	57
2.4.2 Synthesis and Confirmation of Toroidal Structures of DNA that Possesses Writhe Number of ± 1	61

2.4.3 Synthesis and Confirmation of Double Interwound Structures of DNA that Possesses Writhe Number of ± 2	64
2.4.4 Observation of Backbone Self-crossings of Kinetoplast DNA as well as Plasmid DNA Containing Kinetoplast DNA sequences.....	66
2.4.5 Observation of Backbone Self-crossings of Plasmid DNA Containing Repeats of Replication Origins of Bacteriophage λ Sequence.....	70
2.4.6 General DNA Topological Conservation Law of DNA.....	72
2.4.7 Significance of Our Studies.....	76
2.5 Conclusion.....	79

Chapter 3 – Precise Engineering and Visualization of Signs and Magnitudes of DNA Writhe on the Basis of PNA Invasion

3.1 Introduction.....	81
3.2 Design of DNA Sequences.....	89
3.2.1 Design of Linear DNA Precursors with One PNA Binding Site.....	89
3.2.2 Design of Linear DNA Precursors with Two PNA Binding Sites.....	92
3.3 Materials and Methods.....	94
3.3.1 Duplex DNA, Enzymes and Chemicals.....	94
3.3.2 Polymerase chain reactions for synthesis of Linear DNA 9 and Linear DNA 10.....	96
3.3.3 Reactions of SacI with Duplex Linear DNA Precursors.....	97
3.3.4 Preparations of Circular DNA Using T4 Ligase.....	97
3.3.5 Degrade Linear DNA from Ligase Reaction Mixture Using Nuclease BAL-31 Exonuclease.....	98
3.3.6 PNA Invasion.....	98
3.3.7 AFM Studies of Obtained Circular DNA.....	99
3.4 Results and Discussion.....	99

3.4.1 Engineering of DNA Supercoils with Writhe Number of -1 and +1.....	99
3.4.2 Engineering of DNA Supercoils with Writhe Number of -2 and +2.....	106
3.4.3 Significant of Our Studies.....	112
3.5 Conclusion.....	113

Chapter 4 –Positive Supercoiling Affiliated with Nucleosome Repairs

Non-B Structures of DNA

4.1 Introduction.....	115
4.2 Design of DNA Sequences.....	119
4.2.1 Design of Circular DNA with G-quadruplex Structures.....	119
4.2.2 Design of Circular DNA with Cruciform Structures.....	123
4.2.3 Design of Covalently Closed PNA-containing Circular DNA.....	125
4.3 Materials and Methods.....	127
4.3.1 Duplex DNA, Enzymes and Chemicals.....	127
4.3.2 Polymerase chain reactions for synthesis of Linear DNA 11 and Linear DNA 12.....	128
4.3.3 Reactions of SacI with Duplex Linear DNA Precursors.....	129
4.3.4 Preparations of Circular DNA Using T4 Ligase.....	130
4.3.5 Degrade Linear DNA from Ligase Reaction Mixture Using Nuclease BAL-31 Exonuclease.....	130
4.3.6 Reactions of Nt.BsmAI with Circular DNA.....	131
4.3.7 PNA Invasion.....	131
4.3.8 Nucleosome Assembly.....	131
4.3.9 AFM Studies of Obtained Circular DNA.....	132
4.4 Results and Discussion.....	133

4.4.1 Construct Covalently Closed Circular DNA with G-quadruplex Structures.....	133
4.4.2 Disintegrate G-quadruplex Structures from Circular DNA through the Nucleosome Assembly Associated with Positive Supercoiling.....	137
4.4.3 Construct Covalently Closed Circular DNA with Cruciform Structures.....	141
4.4.4 Disintegrate Cruciform Structures from Covalently Closed Circular DNA through Introduction of Positive Supercoils Affiliated with Nucleosome Assembly.....	143
4.4.5 Construct and disintegrate Covalently Closed PNA-containing Circular DNA.....	146
4.4.6 Significance of Our Studies.....	150
4.5 Conclusion.....	151
References.....	153
Curriculum Vitae.....	161

Abstract

It has been recognized in the past that alteration of writhe of DNA in prokaryotic and eukaryotic cells is maneuvered solely by histone proteins and topological enzymes such as gyrase, reverse gyrase, topo I and topo II. It is demonstrated in the current studies for the first time that the shapes of DNA writher (toroidal and interwound forms) can be precisely created sheerly through maneuvering the sequence of DNA and with no involvement of topological enzymes. It is also shown unprecedentedly in our investigation that the size of DNA writhe could be accurately engineered by altering T-rich and A-rich segments in the target duplex DNAs. In addition, the results of our studies confirmed that the intrinsic curvatures of organismal DNAs alone could lead to the generation of duplex backbone self-crossings in their relaxed forms and the backbone self-crossings of those organismal DNAs could be readily confirmed through atomic force microscopic examination.

“DNA Topological Conservation Law ($Lk - Tw = Wr$)” was formulated by Professor F. B. Fuller in 1971 and further elucidated by Professor F. H. C. Crick in 1976 in order to describe the superhelical molecular architectures of DNA that had been discovered in nucleosomes in the eukaryotic cells at an earlier time. This law has since been widely cited in textbooks and taken as the fundamental principle that governs the topological behaviors of DNA in both prokaryotic and eukaryotic cells. It has been known nowadays, however, that the non-canonical B-form of DNA exists ubiquitously in both eukaryotic and prokaryotic genomes. In order to make the supercoiling behavior of non-canonical B-form-containing DNA describable as well,

we examined the non-B DNA issues experimentally and suggested an amended form (“General Topological Conservation Law of DNA”, $Lk - Tw + Nb = Wb + Wn = Wr$) of the original DNA Topological Conservation Law for more precisely describing the topological behaviors of DNA in prokaryotic and eukaryotic cells.

On the other hand, it is known that DNA is stored in organisms in either a right-handed or a left-handed form of supercoil with a fixed magnitude of writhe. Conversions of relaxed forms of DNA to their corresponding supercoiled conformations in the prokaryotic and eukaryotic cells are accomplished exclusively by gyrase, reverse gyrase and the combination of histones and topo I/II, courses of action that largely determines the function and activities of DNA in various cellular processes. It is shown in our studies for the first time, that the right and left handedness of DNA supercoils can be engineered precisely *in vitro* through utilization of the invading property of peptide nucleic acid. In addition, unlike the cellular process in which DNA can merely be converted into its supercoil with a fixed superhelical density, the PNA-invasion action can be utilized to engineer DNA supercoils with any desired magnitude of its writhes.

In addition, DNA damages refer commonly to chemical modifications of DNA structures in the prokaryotic and eukaryotic cells that make the DNA molecules incapable of resuming their original B conformations in a spontaneous manner. In response to the attack of cellular DNA by endogenous metabolites and exogenous causes, all organisms have evolved delicate DNA repairing mechanisms that are able to detect DNA damages, to activate productions of related enzymes and proteins, and

to further repair their damaged DNA. Besides these well-known chemical damages to DNA, physical alterations of canonical B-form of DNA such as formations of G-quadruplex, cruciform and sticky DNA routinely occur in organismal DNA that serve as signals for specified cellular actions. Similar to chemical damages of DNA, many of the non-B DNA structures, once formed, are incapable of resuming their original Watson-Crick base pairings in a spontaneous manner, which could cause damages to DNA in a physical fashion. It is conceivable that if stable non-B DNA structures cannot be repaired in time after their services as cellular signals in living organisms complete, these physically damaged DNA will obstruct the subsequent innate functions of cells in the same ways as chemically damaged DNA does. Unlike the repairing mechanisms of chemically damaged DNA, however, the driving forces and pathways for repairing physically damaged DNA in living organisms have not yet been well understood. In our studies, we demonstrated that positive supercoiling affiliated with nucleosome formation can act as the driving force to repair G-quadruplex, cruciform as well as a stable non-B DNA structure caused by peptide nucleic acid. Our discoveries of the new roles of DNA positive supercoiling affiliated with nucleosome formations may be relevant to the repairing mechanisms of physically damaged DNA in the living organisms.

List of Tables

Table 1.1	Helix Parameters of B-, A- and Z-form DNA.	7
Table 2.1	Nucleotide sequences of Circular DNA 1 and Circular DNA 2.	32
Table 2.2	Nucleotide sequences of Circular DNA 3 and Circular DNA 4.	35
Table 2.3	Nucleotide sequences of Circular DNA 5 and Circular DNA 6.	38
Table 2.4	Nucleotide sequences of kinetoplast DNA.	41
Table 2.5	Nucleotide sequences of Circular DNA 7.	43
Table 2.6	Nucleotide sequences of Circular DNA 8.	45
Table 2.7	Nucleotide sequences of vector pOK12.	47
Table 2.8	Nucleotide sequences of vector pSP73.	48
Table 3.1	Nucleotide sequences of Linear DNA 9.	90
Table 3.2	Nucleotide sequences of Linear DNA 10.	93
Table 3.3	Nucleotide sequences of primers used in polymerase chain reactions.	97
Table 3.4	Statistical data of DNA molecules examined using AFM and their measurement errors.	111
Table 4.1	Nucleotide sequences of Circular DNA 11.	120
Table 4.2	Nucleotide sequences of Circular DNA 12.	125
Table 4.3	Nucleotide sequences of primers used in polymerase chain reactions.	129

List of Figures

Figure 1.1	Schematic representation of the structure of nucleotide.	2
Figure 1.2	Schematic representation the chemical structures of purine and pyrimidine bases in DNA.	4
Figure 1.3	Schematic illustration of a polynucleotide chain, showing the phosphodiester bonds that connect adjacent nucleotide units.	4
Figure 1.4	Pictorial illustration of Watson–Crick base pairs.	5
Figure 1.5	The DNA double helix in solution: structural parameters.	6
Figure 1.6	Electron micrograph of two forms of DNA.	9
Figure 1.7	The definition of the node of negative and positive supercoils.	10
Figure 1.8	Pictorial illustration of generating supercoiled DNA governed by DNA Topological Conservation Law.	13
Figure 1.9	Different forms of supercoils: Interwound or Toroidal.	15
Figure 1.10	The wedge and junction models for DNA bending.	17
Figure 1.11	Schematic illustration of formation of forcible curvature.	18
Figure 1.12	Schematic illustration of formation of nucleosome.	19
Figure 1.13	Palindromes sequences in bacterial plasmid pBR322 and formation of cruciform structures.	21
Figure 1.14	Formation of cruciform structure in negative supercoiled circular DNA.	22
Figure 1.15	Pictorial illustration of the structure of G-quartet.	24

Figure 1.16	Different strand polarity arrangements of G-quadruplexes.	25
Figure 1.17	Schematic representation of structures of human telomere.	26
Figure 2.1	Pictorial illustration of an imaginary process for generating left-handed positive supercoiled DNA governed by the DNA Topological Conservation Law.	28
Figure 2.2	Schematic representation of our design of Circular DNA 1 with writhe number of ± 1 .	31
Figure 2.3	Schematic representation of our design of Circular DNA 2 with non-supertwisted (as controls) structures.	33
Figure 2.4	Schematic representation of our design of Circular DNA 3 with toroidal structures.	34
Figure 2.5	Schematic representation of our design of Circular DNA 4 in relaxed forms.	37
Figure 2.6	Schematic representation of our design of Circular DNA 5 with double interwound structures.	38
Figure 2.7	Schematic representation of our design of Circular DNA 6 in relaxed forms.	40
Figure 2.8	Schematic illustration of the formation of functionalized mica substrates with APS.	55
Figure 2.9	Electrophoretic analysis of synthesis of intrinsic curvature-containing Circular DNA 1 (676 bp in length) from Linear DNA 1.	57
Figure 2.10	AFM image of intrinsic curvature-containing Circular DNA 1.	58
Figure 2.11	Electrophoretic analysis of Circular DNA 1 with topo I.	59

Figure 2.12	AFM image of intrinsic curvature-containing Circular DNA 1 after reacting with Topo I.	60
Figure 2.13	Synthesis and examination of non-supertwisted structures of Circular DNA 2 that are in their relaxed forms.	61
Figure 2.14	Synthesis and examination of toroidal structures of Circular DNA 3 that are in their relaxed forms.	62
Figure 2.15	Synthesis and examination of non-supertwisted structures of Circular DNA 4 that are in their relaxed forms.	63
Figure 2.16	Synthesis and examination of double interwound structures of Circular DNA 5 that are in their relaxed forms.	65
Figure 2.17	Synthesis and examination of non-supertwisted structures of Circular DNA 6 that are in their relaxed forms.	66
Figure 2.18	Electrophoretic analysis of catenated kinetoplast DNA and decatenated kinetoplast DNA.	67
Figure 2.19	AFM image of decatenated kinetoplast DNA minicircles in their relaxed forms.	68
Figure 2.20	AFM images of Circular DNA 7 and pOK12 vector in their relaxed forms.	70
Figure 2.21	AFM images of Circular DNA 8 and pSP73 vector.	72
Figure 2.22	Schematic illustration of relationship among the parameters in canonical B-form DNA.	74
Figure 2.23	Schematic illustration of correlations among Lk, Tw, Nb, Wn, Wb and Wr in DNA that contain accumulable non-canonical B structures.	75

Figure 3.1	Illustration of supercoiled structure of DNA present in cells.	82
Figure 3.2	Chemical structures of DNA and PNA.	83
Figure 3.3	Pictorial illustration of P-loop structures.	83
Figure 3.4	Schematic illustrations of two possible routes for formation P-loop from bis-PNA.	84
Figure 3.5	Hoogsteen binding with protonated cytosine (I) and with pseudoisocytosine (II).	85
Figure 3.6	Schematic representation of reduction of linking number in linear DNA duplex by PNA.	86
Figure 3.7	Schematic representation of engineering of negatively supercoiled DNA by PNA invasion approach.	87
Figure 3.8	Schematic representation of engineering of positively supercoiled DNA on the base of PNA invasion.	88
Figure 3.9	Schematic illustrations of the routes for synthesis of Circular DNA 9 with writhe number of 0.	89
Figure 3.10	Schematic illustrations of engineering of Circular DNA N9 with writhe number of -1.	91
Figure 3.11	Schematic illustrations of engineering of Circular DNA P9 with writhe number of +1.	92
Figure 3.12	Schematic representation of molecular engineering of DNA supercoils with writhe number of ± 2 .	93
Figure 3.13	Synthesis and confirmation of Circular DNA 9 (530 bp in length) from Linear DNA 9 (558 bp in length).	100

Figure 3.14	Agarose gel electrophoretic analysis of the synthesis of Circular DNA N9.	101
Figure 3.15	AFM image of Circular DNA 9 with writhe numbers of -1.	102
Figure 3.16	Detail analysis of AFM images obtained from Circular DNA N9.	103
Figure 3.17	Synthesis and confirmation of Circular DNA P9.	104
Figure 3.18	Detail analysis of AFM images obtained from Circular DNA P9.	105
Figure 3.19	Synthesis and confirmation of Circular DNA 10 from Linear DNA 10.	107
Figure 3.20	Synthesis and confirmation of Circular DNA P10.	107
Figure 3.21	Detail analysis of AFM images obtained from Circular DNA N10.	109
Figure 3.22	Engineering of positive supercoiled Circular DNA P10 with writhe number of +2.	110
Figure 4.1	Pictorial illustration of topological relationship between circular DNA and nucleosome.	118
Figure 4.2	Pictorial illustration of generating G-quadruplex from duplex DNA using DNA gyrase.	120
Figure 4.3	Pictorial illustration of generating G-quadruplex from duplex DNA by alternative methods.	122
Figure 4.4	Pictorial illustration of our strategy for synthesis of cruciform-containing circular DNA.	124
Figure 4.5	Schematic illustrations of our strategy for synthesis of covalently closed PNA-containing circular DNA.	126

Figure 4.6	Examination of formation of G-quadruplex structures from duplex linear DNA with guanine-rich segment.	134
Figure 4.7	Gel electrophoresis analysis of formation of G-quadruplex structures in circular DNA.	135
Figure 4.8	AFM image of circular DNA with and without G-quadruplex structures.	136
Figure 4.9	Schematic illustrations of the disintegration of non-B structure (G-quadruplex) of DNA by nucleosome's positive-supercoil-introducing activity.	138
Figure 4.10	Examination of the disintegration of G-quadruplex structures from DNA circles.	139
Figure 4.11	Examination of the disintegration of G-quadruplex structures from DNA circles but in the absence of histone proteins.	140
Figure 4.12	Examination of synthesis of Circular DNA 12.	141
Figure 4.13	Examination of synthesis of Circular DNA C12.	142
Figure 4.14	Examination of the disintegration of cruciform structures from DNA circles.	144
Figure 4.15	Examination of the disintegration of cruciform structures from DNA circles but in the absence of histone proteins.	145
Figure 4.16	Examination of synthesis of PNA-containing circular DNA.	146
Figure 4.17	Examination of the disintegration of P-loop structures from DNA circles.	148
Figure 4.18	Examination of the disintegration of P-loop structures from DNA circles but in the absence of histone proteins.	149

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

DNA (deoxyribonucleic acid) is a nucleic acid that contains the genetic instructions used in the development and functioning of almost all known living organisms and some viruses.¹⁻⁵ The DNA segments carrying this genetic information are called genes which play a very important role in the dynamic biological processes such as replication, transcription and translocation.⁶⁻¹² The genetic information in DNA is stored as a code made up of four chemical bases: adenine (A), guanine (G), cytosine (C), and thymine (T). Similar to the way in which letters of the alphabet are arranged in a certain order to form words or make sentences, the sequence of those bases in DNA determine the information for building and maintaining an organism's cells as well as passing genetic traits to offspring. As the genetic material, the importance and significance of DNA can be appreciated from a deep understanding of DNA double helix, a structure that is apparently simple but also profound in terms of its implications for biological function.¹³⁻¹⁶ On the other hand, Because DNA is compressed into a crowded cellular environment, topological properties of DNA such as supercoiling (negative and positive), DNA curvature, cruciform structure, G-quadruplex structure and other non-canonical B-form structures of DNA influence virtually every major nucleic acid process.¹⁷ DNA, among all biological

macromolecule, has consequently attracted most attention and fired the imagination of scientists and non-scientists alike in the past.¹⁸⁻²⁰

1.1 The Watson-Crick Model and B-form DNA

The classical view of the DNA double helix was described by Watson and Crick in 1953, which was one of the most important scientific discoveries of the twentieth century.¹⁹⁻²⁰ Nine years later, they shared the Nobel Prize in Physiology or Medicine with Maurice Wilkins in 1962 for their discoveries concerning the molecular structure of nucleic acids and its significance for information transfer in living material. This model could perfectly reveal how DNA can fulfill its biological functions and satisfy the known chemical and physical properties of DNA. Half a century later, important new implications of this contribution to science are still coming to light.

Figure 1.1 Schematic representation of the structure of nucleotide.²

DNA is a polymer and the monomer units of DNA are nucleotides (Figure 1.1). Each nucleotide consists of a 5-carbon sugar (2'-deoxyribose), a phosphate group, and a nitrogen containing base attached to the sugar. The bases in DNA can be classified into two types: (1) the purine bases, five-membered and six-membered heterocyclic compounds (adenine and guanine) and (2) the pyrimidine bases, the six-membered rings (thymine and cytosine) (Figure 1.2). The backbone of the DNA strand comprises alternating phosphate and sugar residues while the DNA chains (single strand DNA) have two distinct ends which was known as 5' (five prime) and 3' (three prime), with the 5' end having a terminal phosphate group and the 3' end containing a terminal hydroxyl group. With forming phosphodiester bonds which connects the 3'-hydroxyl group of one sugar to the 5'-hydroxyl of the next, the sugars are joined together by phosphodiester linkages and a polynucleotide chain is made by those joining the sugars and bases, which constitutes the primary structure of DNA (Figure 1.3).

Due to the experimental data of X-ray diffraction of DNA fibres as well as chemical data, Watson and Crick deduced a model for the structure of DNA.²⁰ In the double-helical model, purine bases can form hydrogen bonds to pyrimidine bases, with adenine (A) pairing only to thymine (T) and cytosine (C) bonding only to guanine (G). The linking between two nitrogenous bases on opposite complementary DNA strands that are connected via hydrogen bonds is called a base pair. As shown in Figure 1.4, there are two types of base pairs which form different numbers of hydrogen bonds: (I) Two hydrogen bonds can be formed between adenine (A) bonding and thymine (T); (II) Three hydrogen bonds can be formed between cytosine (C) bonding and guanine (G). Besides the “Watson-Crick” base pairs, other base pairs

schemes are possible. For example, “Hoogsteen” base pairs can also be found in DNA triplex structures and apparent mismatches can occur in some particular DNA and RNA structures.²¹

Figure 1.2 Schematic representation the chemical structures of purine and pyrimidine bases in DNA.

Figure 1.3 Schematic illustration of a polynucleotide chain, showing the phosphodiester bonds that connect adjacent nucleotide units.²

Figure 1.4 Pictorial illustration of Watson-Crick base pairs. The hydrogen bond distances and distances between the C1' positions of the ribose sugars are indicated.¹

The double helix structure of DNA is made of two strands which are coiled around each other in an antiparallel and right-handed fashion. This spatial arrangement of two DNA strands facilitates forming a structure with a largely hydrophobic interior consisting of the planar DNA bases stacked on each other with the hydrophilic sugar-phosphate backbone on the outside, which physically protects all the important atoms of bases from chemical damage by the environment. A number of parameters are used to define the double helix structure of DNA. As shown in Figure 1.5, there are 10.5 base pairs (bp) for every turn of the helix. Since 360° constitutes one helical turn, there would be a 34.3° ($360/10.5$) twist angle or rotation per residue between adjacent base pairs. The helix pitch (the length of one complete helical turn) is 35.7 Å. The helix diameter (the width in Å across the helix) is about 20 Å. Axial rise (the distance between adjacent planar bases in DNA double helix) is 3.4 Å. The position of major groove and minor groove are also indicated in Figure 1.5.¹

Figure 1.5 The DNA double helix in solution: structural parameters.¹

This form of the DNA double helix shown above is known as the canonical B-form, a structure which is thought to represent the conformation of most DNA found in cells. The basic structural parameters of canonical B-form DNA were originally derived from X-ray diffraction analysis of sodium salt of DNA fibers at 92% relative humidity.²¹⁻²² The dominant feature that distinguishes B-form DNA from other forms is the major and minor grooves which provide very distinct surfaces with which proteins can interact. There are 10.5 bp per right-handed helical turn and the sugar pucker (the form of the ribose sugar) is C2'-endo. A number of parameters are defined to describe the conformation of the double helix shown in Figure 1.5 and Table 1.1. Commonly, term B-form DNA will be used to refer to the right-handed helical form found for DNA in solution.

Apart from canonical B-form DNA, A-form DNA was first identified by Fuller from X-ray fibers diffraction analysis at 75% relative humidity.²³ The significant difference between B-form DNA and A-form DNA is that the conformation of the ribose sugar: A-DNA is normally C3'-endo and B-DNA is C2'-endo. Moreover, the helix conformation of A-form DNA is border and untwisted²⁴ (data shown in Table 1.1). Besides A- and B-form DNA, Z-DNA which can be formed in the particular sequences under certain condition²⁵⁻²⁸ (the presence of some certain divalent cations, high salt concentration or DNA supercoiling) is a left-handed helix conformation that is distinct from right-handed DNA forms.²⁹⁻³⁰ Another distinguishing feature of Z-form DNA is the zigzag path of the sugar-phosphate backbone, which is why it was named Z-DNA.³¹ There is some evidence that Z-DNA occurs in both prokaryotes and eukaryotes during the course of cellular processes such as transcription and gene activation.³²

Table 1.1 Helix Parameters of B-, A- and Z-form DNA.¹

<i>Parameters</i>	<i>B-form DNA</i>	<i>A-form DNA</i>	<i>Z-form DNA</i>
Helix handedness	Right	Right	Left
Residue per turn	10.4~10.5	11.0	12
Sugar pucker	C2'-endo	C3'-endo	C2'-endo (pyrimidine) C3'-endo (purine)
Helix diameter	~2.0 nm (20 Å)	~2.6 nm (26 Å)	~1.8 nm (18 Å)
Major groove	Wide and deep	Narrow and deep	Flat
Minor groove	Narrow and deep	Wide and deep	Narrow and deep

1.2 DNA Supercoiling and DNA Topological Conservation Law

The structure of DNA does not only exist as secondary structures such as double helix we stated above, but it can fold up on itself to form tertiary structures by supercoiling.³³ DNA that is stored in both prokaryotic and eukaryotic cells exists almost all in either negatively or positively supercoiled forms.³⁴⁻³⁷ During the dynamic processes of replication, transcription and translocation, these supercoiled structures of DNA are transformed into their relaxed matching conformations and will further resume their supercoiled states after these genetic actions complete.³⁸⁻³⁹

1.2.1 Closed Circular DNA and DNA Supercoiling

The macromolecule of DNA in bacteria, archaea and mitochondria of eukaryotes is circular in its backbone while eukaryotic chromosomal DNA possesses open ends. Even though it is linear, a segment of any genetic DNA in eukaryotes must be considered as a circle when its topological property is evaluated.⁴⁰⁻⁴¹ This happens because the termini of chromosomal DNA are too far away to be reachable, which makes the backbones of any inside duplex DNA segments virtually not freely rotatable. The circular DNA possesses a covalently closed backbone, which means the two phosphodiester backbones are intact and covalently continuous. A simple plasmid DNA molecule which was purified from bacterial cell is a circular DNA which originally called Form I DNA. This supercoiled plasmid DNA appears as a tangled and twisted structure while the nicked circular plasmid DNA molecules are relaxed

and lose the twists (Figure 1.6). Nicked DNA contains a single break in one strand and has been called Form II DNA. The circular DNA with one or more nicked site contains no super twists and will appear as a relaxed and untwisted circular ring because the broken strand can rotate about the intact strand to dissipate the torsional stress. If the circular DNA was broken in both phosphate backbone of the two strands at the same point or very near point, a linear DNA which was named Form III DNA can be formed. The terms of Form I, II as well as III DNA were used to describe the different conformations of DNA as shown above, but it should be replaced by some new concepts such as “supercoiled DNA”, “relaxed DNA” as well as “linear DNA”.

Form I DNA (Supercoiled DNA)

Form II DNA (Nicked DNA)

Figure 1.6 Electron micrograph of two forms of DNA. The twisted and tangled structure is supercoiled DNA (Form I DNA) while the relaxed and untwisted structure is nicked DNA. The plasmid DNA molecule shown above is about 9000 bp in length.¹

Figure 1.7 The definition of the node of negative and positive supercoils.⁴²

It is well known that superhelical turns may be of either right-handedness or left-handedness: the right-handed DNA helix rotates in a right-handed (clockwise) direction while the left-handed DNA helix rotates in opposite motion. On the other hand, DNA supercoiling can also be divided into negative and positive. The underwinding of DNA helix corresponds to a negative supercoiling (right-handed super-twisted) while the overwinding of DNA helix leads to a positive supercoiling (left-handed super-twisted). If the front segment of DNA is aligned with the back segment using a rotation of $0^\circ \sim 180^\circ$, the sign of the node can be identified by tracking the movement and direction of DNA (Figure 1.7 A). Once the front segments of DNA rotate in a clockwise manner to align it with the back segment, the crossovers of DNA can be defined as negative supercoiling, whereas positive supercoiling require a counterclockwise rotation (Figure 1.7 B). Polyoma DNA was

identified as the negative supercoils by Vinograd,⁴³ although the positive supercoiling occurs when building up ahead of a replication fork positive.⁴⁴

1.2.2 Mathematical Expression of DNA Supercoiling: DNA Topological Conservation Law

The conversion of relaxed forms of DNA into their supercoiled conformations is maneuvered *in vivo* exclusively by topoisomerases which can catalyze the interconversion between different topological forms of DNA⁴⁴⁻⁴⁷ (e.g. DNA gyrase and reverse gyrase) and histone proteins³⁵, a delicate action that has been hardly imitable by any other non-enzymatic means. Shortly after the molecular assembly of nucleosomes and chromatins had been characterized in the late 1960s and early 1970s, “DNA Topological Conservation Law”⁴⁸⁻⁵⁴ was formulated for describing the topological features of DNA formed in these constituents of chromosomes:

$$Lk - Tw = Wr \quad (\text{Equation 1.1})$$

This law has since been widely cited in textbooks and applied in nucleic acid research as the fundamental principle that rules the emergence of DNA supercoiling as well as transformation between the supercoiled and relaxed conformations of DNA *in vivo* and *in vitro*.⁵⁵⁻⁶⁰

To quantitatively measure the supercoiling of a DNA molecule, a series of mathematical concepts must be introduced: linking number (Lk), twist number (Tw)

and writhe number (Wr). Linking number (Lk) is the number of intertwinings between two complementary DNA strands which means one DNA strand crosses the other in the geometric plane surface; twist number is the total number of turns of double stranded DNA around its helical axis; and writhe number is a measure of coiling of the axis of double helix.⁵² When a circular DNA or a virtually circular DNA of different sizes is acted on by topoisomerases, alteration of topological features of the DNA could occur. A 177 base pair circular DNA (Figure 1.8), for example, could be transformed into a positive supercoil when reverse gyrase (a type of topoisomerases) is allowed to interact with it. Since this small circular DNA (Structure 3 in Figure 1.8) is obtained through a ligation reaction from its linear precursor catalyzed by DNA ligase, the DNA circle must exist in its relaxed form because there is absence of gyrase or reverse gyrase activity in DNA ligase. Consequently, both linking number (Lk) and the twist number (Tw) of this DNA circle should be 17 ($177/10.4$). According to DNA Topological Conservation Law the writhe (Wr) number in this case should be equal to zero ($Wr = Lk - Tw = 17 - 17 = 0$), which indicates that there is structurally no self-crossing in the backbones of circular duplex DNA. After reverse gyrase acts on the relaxed form of the 177 base pair DNA circle in its initial stage, on the other hand, the linking number of this circular DNA will be altered from 17 to 18. The average rotation per residue will accordingly change from 34.3° in its original relaxed form to a higher degree in the new state. Because the new state (Structure 4 in Figure 1.8) represents an unfavorable conformation of the DNA double helix, the winding tension generated in the DNA circle will be relieved by forming a left-handed supercoil (Structure 5 in Figure 1.8). According to DNA

Topological Conservation Law, the writhe (Wr) of this DNA is equal to one ($Wr = Lk - Tw = 18 - 17 = 1$).⁶¹⁻⁶²

Figure 1.8 Pictorial illustration of generating supercoiled DNA governed by DNA Topological Conservation Law. Structure 1: linear DNA with two cohesive ends; Structure 2: two cohesive ends pair each other; Structure 3: covalently closed circular DNA; Structure 4: high energy intermediate; Structure 5: positive supercoiled DNA.

It is well known that the ends of linear DNA molecule are free to rotate, which is the preferred helical repeat and represents lowest energy form (Structure 1 in Figure 1.8). When the state of helical twist exists in a covalently closed circular DNA, the DNA molecule is relaxed and contains no supercoils (Structure 3 in Figure 1.8). The linking number of the DNA in relaxed form is defined as Lk_o , which can be obtained from the following equation:

$$Lk_o = N / 10.4 \quad (\text{Equation 1.2})$$

where N is the number of base pairs of the whole DNA sequence and 10.4 refers to the helical repeat. In this case, the linking number (Lk) equals the twist number (Tw) and writhe number should be zero (Structure 3 in Figure 1.8). It should be pointed out that the linking number will be same whether the DNA molecule existed in a linear, nicked or covalently closed form.^{1, 63}

The torsional stress can be introduced when the relaxed circular DNA is transformed into its supercoiled conformation. Different degree of torsional stress can lead to different free energy of supercoiling. To give a measure of the levels of supercoiling that can be used for comparisons between DNA molecules, the term σ (superhelical density)⁶⁴⁻⁶⁶ is utilized, which is independent of DNA length and can be calculated using the following equation:

$$\sigma = (Lk - Lk_o) / Lk_o \quad (\text{Equation 1.3})$$

The σ value for underwound DNA is always negative and for overwound DNA is always positive. For example, the length of bacterial plasmid DNA or the *Escherichia*

coli chromosome⁶⁷ varies greatly, a superhelical density of around - 0.06 has been identified.

Figure 1.9 Different forms of supercoils: Interwound or Toroidal.

The conformation of supercoiled DNA can exist not only in an interwound structure as shown in Figure 1.8 (Structure 5) but also in a toroidal form. The difference between interwound and toroidal structures is shown in Figure 1.9.^{40, 60-61, 68} Similar to the interwound supercoils, DNA with toroidal structures also plays a very important role in living cells. It is well known the negative supercoils can be introduced by DNA wrapping around octamer histones which was known as nucleosomes in eukaryotes. DNA supercoils in this case are in a toroidal coils conformation and the free energy of supercoiling is restrained in the stable writhing around proteins.

1.3 DNA Bending

More than twenty years ago, the concept of sequence-dependent DNA structure was proposed,⁶⁹ which indicated that the DNA sequence not only encodes genetic information but also greatly affect the spatial structure of DNA in dynamic processes of replication, transcription, translocation as well as DNA packaging within the nucleus. There are two related but distinct phenomena associated with DNA bending: intrinsic curvature of DNA and DNA flexibility. Intrinsic curvature, which can be observed in some specific sequence such as short adenine tracts, can lead to a bending of the DNA helix axis. On the other hand, if certain DNA sequences can be bent, for example, DNA wrapping around octamer histones, it can be called DNA flexibility.

1.3.1 Intrinsic Curvature of DNA: Wedge Model or Junction Model?

Although the X-ray diffraction data showed that a small degree of bending occurs in many DNA sequence, the polyacrylamide gel electrophoresis experiments give us more convincing results.⁶⁹ The DNA molecules with higher degree of curvature show less migration distance than those straight molecules of the same size. This happens because the curved DNA molecules pass through the pores of polyacrylamide gel less easily than their straight counterparts. On the other hand, one of the common characteristics of these intrinsically bending DNA sequences is its possession of short adenine tracts spaced periodically along their DNA backbones (e.g. $\{AAAAACGGC\}_n$). When those short adenine tracts are periodically placed with a spacing of 10.4 ~ 10.5 bp, a distinct structure from random sequence can be

observed.⁷⁰⁻⁷¹ Those spaced adenine tracts in phase with the helical repeat of DNA are prone to form the stable DNA curvatures.

Wedge Model

Junction Model

Figure 1.10 The wedge and junction models for DNA bending.¹

There are two controversial theories to explain the intrinsic curvature in DNA: Wedge Model and Junction model. In Wedge Model, it assumes that there is a “Wedge” angle existing between the AA dinucleotide, which causes a deflection in the axis of the DNA double helix.⁷²⁻⁷³ The bending arises from the “Wedge” at various positions in DNA and total bend will be the sum of all individual wedges. Ulanovsky *et al.* synthesized small DNA circles using double-stranded DNA possessing a strong 10.5 bp periodicity of the runs of adenines and calculated the wedge angle of an AA dinucleotide.⁷³ On the other hand, Junction Model suggests that the adenine tracts form a slight different double helix than the normal B-form conformation and the bending is derived from the abrupt change between adenine tracts and B-form

structures in the direction of the helix axis as shown in Figure 1.10.⁷⁴⁻⁷⁵ DNA bending sequences occur naturally, for example, kinetoplast DNA is the mitochondrial DNA of trypanosomes which is composed of thousands of relaxed interlocked minicircles. One of the most intriguing characteristics of the kinetoplast DNA is its possession of various spaced adenine tracts in its circular structure.⁷⁶⁻⁷⁷ Furthermore, it is possible to synthesize artificial DNA sequences with phased adenine tracts and make DNA circles in the test tube as small as 105 bp.⁷³

1.3.2 DNA Flexibility: Forcible curvature of DNA as well as Protein-induced DNA bending

Figure 1.11 Schematic illustration of formation of forcible curvature.

There are two types of DNA flexibility: isotropic and anisotropic. Isotropic flexibility means that the DNA molecule can bend equally in all directions, on the other hand, anisotropic flexibility means that the DNA can only bend in a preferred direction.² Apart from the intrinsic curvature in DNA, there is another kind of curvature named forcible curvature, which is related to DNA flexibility as well as DNA cyclization. Forcible curvature refers to the phenomena that a linear DNA with certain sequence can be bended through cyclization reaction to form a circular structure⁷⁸⁻⁸⁰ (Figure 1.11). Regardless of adenine tracts existed in DNA sequence or not, the newly formed circular DNA contains a bending structure which is introduced from ligase reaction. It has been well established the covalently closed circular DNA exhibited more thermodynamically stability than its linear counterpart through some thermal and alkaline melting experiments.³³

Figure 1.12 Schematic illustration of formation of nucleosome.

Besides the forcible bending as discussed above, DNA-protein interactions can also lead to bending conformations which is associated with DNA flexibility. Many DNA-binding proteins can bend DNA and introduce the curvature structures, for example, DNA with 146 bp in length are wrapped in 1.8 turns around the histone protein octamer, which is known as nucleosome (Figure 1.12). There are some experimental evidences that certain sequence motifs within the 146 bp region is preferred when forming nucleosome, even though it can be regarded as non-specific in terms of DNA sequence.⁸¹⁻⁸³

1.4 Alternative Conformations of DNA

Although the X-ray diffraction analysis of DNA gave us a sophisticated view of DNA double helix,⁸⁴ two-dimensional and three-dimensional nuclear magnetic resonance (NMR) provided us more elaborate data of DNA in solution,⁸⁵⁻⁸⁷ which allow us to know that the double helix structure is not the extremely uniform and monotonous conformation that DNA can adopt. Besides the well-recognized canonical B-forms 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”. 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.⁸⁸⁻

1.4.1 Cruciform Structures in DNA

The cruciform structures of DNA is important for the critical biological processes of DNA recombination and repair that occur in the cell,⁹⁰⁻⁹² and it is believed to form at or near replication origins of some eukaryotic cells and serve as recognition signals for DNA replication.⁹³⁻⁹⁴ The consequence of intrastrand base-pairing in double-stranded DNA leads to cruciforms which consist of a pair of stem and loop structures. The presence of palindromes sequences (inverted repeats) in DNA is necessary when forming cruciform structures. Bacterial plasmid pBR322, for example, consists of palindromes sequence and a cruciform structure with stems of 11 base pair can be found in its supercoiled structures⁹⁵⁻⁹⁶ (Figure 1.13).

Figure 1.13 Palindromes sequences in bacterial plasmid pBR322 and formation of cruciform structures. The 11 bp inverted repeat sequence is shown in bold.²

Figure 1.14 Formation of cruciform structure in negative supercoiled circular DNA.

It is clear that formation of a cruciform structure in B-form DNA with linear conformation will be thermodynamically unfavourable even though the inverted repeat sequence occurs in it. This happens because the energy is required to melt the center of the palindromes sequences to allow the intrastrand base-pairing. On the other hand, the cruciform structures can be formed in negative supercoiled circular DNA which promotes breathing effect in the double helix and it has been experimentally confirmed in the past. To form a cruciform structure, the double helix structures between two single strands are unwound and intrastrand base-pairing occurs to form the stems of cruciform. The formation of cruciform structures in negative supercoiled DNA results in relaxation of supercoils, which arises from the unfavourable free energy associated with negative supercoiling⁹⁷⁻⁹⁹ (Figure 1.14). When circular DNA is at or below a certain superhelical density, no DNA molecule with cruciform structure can be observed. The addition of one or more negative supercoil (negative writhe number) can lead to almost all molecules with cruciform structures, which has been

identified by two-dimensional gel¹⁰⁰ and the preference of single-strand-specific reagents (such as T7 endonuclease).

1.4.2 Four-Strands Nucleic Acids: G-quadruplexes

The famous Watson-Crick duplex form is the regular conformation adopted by DNA as discussed above. The four-stranded DNA structures, on the other hand, known as G-quadruplexes have attracted many attentions during the years and the motifs for the formation of G-quadruplex DNA structures are widely used in eukaryotic genomes, and many significant biological process.¹⁰¹⁻¹⁰⁵

The most usual structure of DNA is a double-stranded helix with stacked base pairs (adenine-thymine, guanine-cytosine) on the inside and with a negatively charged backbone (deoxyribose phosphate) on the outside. The two antiparallel strands are held together between the bases by complementary basepairing. Three hydrogen bonds exist in GC and two hydrogen bonds occur between AT, which was named as Watson-Crick base pairing as shown in Figure 1.4. Unlike double helical structures, G-quadruplexes have a core that is made up of guanine bases only, which is held together by a cyclic arrangement of eight Hoogsteen hydrogen bonds around the edges of the resulting square. These planar structures are called G-quartets (Figure 1.15). Apart from the classical Watson-Crick G-C base paring for duplex DNA, G-quartets can overlap and a series of nucleic acid secondary structures can be formed, which were named G-quadruplexes. The presence of a central cation (typically potassium)

helps to maintain the stability of the structure, which may be very stable under physiological conditions.¹⁰⁶⁻¹⁰⁷ More generally, the ionic radius is a parameter that aptly describes how well guanine tetrads are stabilized by various cations. In the alkali series the order is generally $K^+ > Na^+ > Rb^+ > NH_4^+ > Cs^+ > Li^+$ and for the earth alkali series the order is $Sr^{2+} > Ba^{2+} > Ca^{2+} > Mg^{2+}$.¹⁰³

Figure 1.15 Pictorial illustration of the structure of G-quartet. Four guanines can hydrogen bond in a square arrangement to form a G-quartet. M^+ , a central cation (typically potassium), helps to maintain the stability of the structure.

G-quadruplexes can be formed by one, two, or four strands of DNA due to stoichiometry variation of strands.¹⁰⁸⁻¹¹³ In principle, three strand arrangements are possible but have yet to be substantiated. Moreover, if one strand could form a unimolecular structure it could also form bimolecular or quadrimolecular structures.¹¹⁴⁻¹¹⁶ The concentration of DNA strands determined which structures can be adopted.¹¹⁷ Since the strands were customarily described as from the 5' end to the 3' end, the possibility of topological variants for these four strands appears.

Irrespective of whether they are part of the same molecule or not, the strand or strands that constitute a G-quadruplex can come together in four different ways, which can be described as all parallel, three parallel and one antiparallel, adjacent parallel or alternating antiparallel as shown in Figure 1.16.

Figure 1.16 Different strand polarity arrangements of G-quadruplexes. (A) Four strands structure with all strands parallel; (B) antiparallel structure formed by two strands with adjacent parallel strands; (C) antiparallel structure formed by single strand with alternating parallel strands; (D) single strand parallel structure with three double chain reversal loops; (E) single strand antiparallel structure with adjacent parallel strands and a diagonal loop; (F) single strand mixed structure with three parallel and one antiparallel strands. All three structures (D), (E) and (F) have been observed for the human telomeric repeat.¹⁰¹

There are repeated sequences at the ends of the linear chromosomes in eukaryotic cells, which facilitate distinguishing between chromosome ends and unexpected

breaks in the DNA.¹¹⁸⁻¹¹⁹ In all vertebrates this repeated sequence is d(GGGTTA)_n and other organisms generally have very similar sequences, characterized by runs of GGG with intervening bases.¹²⁰ Human telomeres consist of repetitive stretches of TTAGGG at the end of chromosomes of human cells which prevent the ends of chromosomal DNA strands from destruction during the course of replication.¹²¹⁻¹²² The average length of human telomere varies between 5 and 15 kilo-bases depending on the tissue type and several other factors. At the very end of human telomeric DNA, there is a single-stranded 3' overhang of 75 to 300 nucleotides¹²³⁻¹²⁴ (Figure 1.17). It is well known that the length of the double-stranded region of the telomere becomes shorter with every cell division and it is so-called end replication problem. Accordingly the telomere becomes too short, as a result, chromosome fusion, senescence and apoptosis occurred. If something would elongate the telomeres, there can be lifetime on cells. There is an enzyme called telomerase which can prolong the telomeres using an internal RNA template. The cancerous cells can also overcome the limit on cell divisions by expressing telomerase and it has attracted much interest in developing methods to reduce the activity of telomerase for therapeutic purposes.¹²⁵⁻

128

Figure 1.17 Schematic representation of structures of human telomere.

Chapter 2

Manipulating DNA Writhe through Varying DNA Sequences and General Topological Conservation Law of DNA

2.1 Introduction

During the dynamic processes of replication, transcription and genetic recombination in prokaryotic and eukaryotic cells, supercoiled structures of DNA^{1, 35} are transformed into their relaxed matching conformations and will further resume their supercoiled states after these genetic actions complete^{38-39, 47}. The conversion of these relaxed forms of DNA into their supercoiled conformations *in vivo* is maneuvered exclusively by topological enzymes (e.g. DNA gyrase, reverse gyrase and topoisomerase I) and histone proteins.^{45, 81-83, 129} Shortly after the molecular assemblies of nucleosomes and chromatin in the eukaryotic cells had been characterized in the late 1960s, a mathematical equation of $Lk - Tw = Wr$ was introduced for describing the topological features of DNA formed in these constituents of chromosomes.⁵³ This mathematical equation was further elucidated by Crick from the perspectives of molecular biology and was later named “DNA Topological Conservation Law”⁴⁸⁻⁵⁴. According to this law, the magnitude of DNA writhe could be maneuvered by the

action of gyrase and reverse gyrase through their alteration of the linking number (Lk) of the target DNA^{28, 46, 56-57, 130} as illustrated in Figure 1.

Figure 2.1 Pictorial illustration of an imaginary process for generating left-handed positive supercoiled DNA governed by the DNA Topological Conservation Law. When a reverse gyrase adds one linking number to a circular DNA, for instance new writhe will be formed according to the mathematical equation of $Wr = Lk - Tw$.⁶¹⁻⁶²

Various specific sequences of DNA, on the other hand, are known to be capable of existing as intrinsically bent structures^{74-75, 131-132}, a type of non-canonical B conformations that occurs periodically in eukaryotic genomes as well as in prokaryotic DNA.^{18, 133-134} One of the common characteristics of these intrinsically bent DNA sequences is its possession of short adenine tracts spaced periodically along their DNA backbones (e.g. (AAAAACGGC)_n).⁷⁵ In addition, it has been well established that the degree of bending in a sequence of DNA is dependent on the lengths of short adenine tracts as well as the nature and lengths of the nucleotides spaced between adenine tracts. Moreover, when a 105 base-paired linear DNA sequence was designed to

contain properly edited sequence of adenine and non-adenine tracts, the two termini of the linear DNA sequence could position themselves in close proximity as the consequence of its possession of intrinsic curvature.⁷³ On the basis of our recent analysis on the currently available information about DNA curvatures, we speculate that besides the known alteration of writhe by the actions of topological enzymes (e.g. gyrase and reverse gyrase), the shape (interwound and toroidal forms) and magnitude of DNA writhes could be maneuvered by varying directions and degrees of intrinsic DNA curvatures. Our speculation on the DNA writhe issues has subsequently been tested experientially in our lab recently. Here we report our design, synthesis and confirmation of DNA that possess desired the shapes (either interwound or toroidal forms) and magnitude of writhes without the assistance of gyrase, reverse gyrase and histone proteins.

Besides those artificially designed DNAs that were capable of forming interwound and toroidal structures in their relaxed forms,⁶¹ we speculated that certain DNA sequences that occur in organisms could display non-zero writhe number as well, and have consequently carried out some new studies. Here we report our atomic force microscopic examination on kinetoplast DNA⁷⁶ as well as the DNA that contain sequences at the replication origin of Bacteriophage λ ¹³⁵. Our new results illustrate that these naturally occurring relaxed form of organismal DNA could indeed exhibit backbone self-crossings in their AFM images as well.⁷⁷ Consequently, it was our further speculation that the mathematical correlation between Lk, Tw and Wr in DNA Topological Conservation Law might not be held validly any longer to describe the topological features of DNA that contain intrinsically bend or other non-canonical B-

DNA segments. Here we present the evidences showing non-conservation of “DNA Topological Conservation Law” in both designed and organismal DNA. In addition, since the non-canonical B-forms of DNA are nowadays known to exist ubiquitously in both eukaryotic and prokaryotic genomes and to play crucial roles in various cellular processes, an amended form of original conservation law is proposed and named as “General Topological Conservation Law of DNA” in our studies, in which the contributions of both canonical B-forms and non-canonical B-forms of DNA are taken into consideration.^{61, 77}

2.2 Design of DNA Sequences

2.2.1 Design of Interwound Structures of DNA that Possesses Writhe Number of ± 1

With the aim of demonstrating that fabrication of an interwound DNA structure with writhe number of one in the absence of gyrase or reverse gyrase is feasible, a linear duplex DNA sequence (Linear DNA 1 in Figure 2.2) was designed during our investigation that contains 676 base pairs in length. One of the uniqueness of the 676 base pair linear DNA is that this sequence contains two consecutive spaced tracts of adenines with nearly equal lengths that spread in its two opposite strands respectively. Upon the action of T4 DNA ligase on this linear DNA precursor (Figure 2.2), a circular DNA (Circular DNA 1) was obtained in our studies. Since there was neither gyrase nor reverse gyrase involved in the ligation reaction, the newly formed DNA

circle must exist in its relaxed form, which implies that the linking number and twist number of the newly formed relaxed DNA circle are equal to each other ($Lk = Tw = 676/10.4 = 65$). According to DNA Topological Conservation Law, the writhe number of the relaxed form of 676 base pair DNA circle should be equal to zero ($Wr = Lk - Tw = 65 - 65 = 0$).

Figure 2.2 Schematic representation of our design of Circular DNA 1 with writhe number of ± 1 . See Table 2.1 for detailed information about the nucleotide sequences of Circular DNA 1.

Table 2.1 Nucleotide sequences of Circular DNA 1 and Circular DNA 2. Only one of the two strands of DNA from 5' end to 3' end is shown in the table. * (1) Junctions between the segments that are highlighted in red and in blue represent the points at which adenine-rich and thymine-rich sequence alternate between the two opposite strands. (2) This DNA contains 2 segments of continuous spaces adenine tracts.

Name of DNA	Nucleotide sequence
Circular DNA 1*	CGAAAAGTGCAAAAAGTCGGAAAAATCCGTGCAAAAATCGTCAAAA GGCCCGAAAAAATAGCTAAAAATCGTCGAAAAACTGCGTTGAAAAA GCTTAAAAACGATGCAAAAAGTCATTCAAAAATGGGCAAAAAGTG GCCAAAAGCTATAAAAAACGCGCAAAAATCGCACTTTTTGGAGT TTTTTCGGGCTTTTTTTGGATCATTTTTTAGTCGTTTTTGGCCATT TTCGGCACTTTTTTGCATATATTTTGCCCGTTTTTGCCAATTTTTTC GTATTTTCGCTATTTTTTGGCATTTTTTTGGCCATTTTTTAGGCTTTT GGGTGGTTTTTCGGCCGTTTTTTGGAGGTTTTCCAGATTTTGCATTT TCAGTGCGTTTTTGGCCATTTTCGGCTGTTTTTTGCCATATATTTT GCCCGTTTTTGCCAATTTTTTTGCCAATTTTCGGGGTATTTTCGCT ATTTTTTGGCATTTTTTTGGCCATTTTGCAATGTTTTTTTCAGTTTTT TTGTGCAAAAGCAGTGAAAAGTGGCCAAAAATGCCGAAAAATCAGC GAAAAGCCCGAAAAATGGCCGTAAAAATGGTAAAAGAAATCAAAAA TTCATAAAACCCAAAATGGCTGAGAAAAGGAGTGAAAAATGGTCC TAAAACCGCAAAAATCTCTCAAAAATGAGCT
Circular DNA 2	CGCTACATAATACGACTCACTATTATATGTATAACTTCGTATAATG TATGCTATACGAAGTTATTGCTCGCAGTGTTACTGCAATCATCGTG GTGATTAATCCTTGTTGTGTAATTCGTTACTCAACGAAGGTTAATTC ACTATAGTTGTCTGGTACTCTCTAGTGAATTCCTTAAGTGAGTAG TATTAAGAAGTAAGTGTAAGATGCTTCGAGTTATGTGACTGATAAG TATTCAATCAAGTCATTCTGAGAATAATGTATGTTACTATAATCAT GATTAGAACTCGAGTTGCTCTTGTCATGGTGTCAACGTTGGATAATA CTGACATAGCAGAACTTTAAAAGTGTTCAATTATTGGAAGATCTGCG AACATGCTCAACGTTCTTACCTCTGTTGAGATCCAGTTTCGATATAA TTCACCTGTGCACCCAACTTATCTTCAGCATTACTTTACCAGAGT TTCTGTGTGAGTAATCGGGAAACAGGAAGGCAAGATTCAAATCTTA AAGTGAATAAGTTCGACACAGAAATGTTGAATGCTCATACTCTTCC TTTAGCGACTTCAATACGCTTATTGAAGCATTTATAAGGGTTATTG TTACGTGAGTATAACTTCGTATAATGTATGCTATACGAAGTTATGG CTCGAAGTCAGTTATAGATATCTAAGTGAGCT

In order to confirm that the observed self-crossings of DNA in Figure 2.2 are indeed associated with the intrinsic curvatures of DNA, a different 676 base pair circular DNA (Circular DNA 2) as control was designed from Linear DNA 2 next during our investigations, which possesses the same length and the same nucleotide composition as Circular DNA 1, however it possesses a flexible structure in its whole sequence as shown in Figure 2.3.

Figure 2.3 Schematic representation of our design of Circular DNA 2 with non-supertwisted (as controls) structures. See Table 2.1 for detailed information about the nucleotide sequences of Circular DNA 2.

2.2.2 Design of Toroidal Structures of DNA

Supercoiled DNA is known to be capable of existing in either interwound or toroidal forms *in vivo* and *in vitro*.^{35, 40, 136} With the aim of examining whether intrinsic curvature-containing DNA could adopt the structural features beyond the interwound forms shown in Figure 2.2, a new 1154 base pair circular DNA (Circular DNA 3) was designed next during our investigations, in which all spaced tracts of adenines occur in the same strand of its duplex structure as shown in Figure 2.4.

Figure 2.4 Schematic representation of our design of Circular DNA 3 with toroidal structures. See Table 2.2 for detailed information about the nucleotide sequences of Circular DNA 3.

Table 2.2 Nucleotide sequences of Circular DNA 3 and Circular DNA 4. Only one of the two strands of DNA from 5' end to 3' end is shown in the table. * (1) Continuous spaced adenine tracts occur exclusively in one of the double strands. (2) The duplex segments highlighted in blue and green are those that possess high and low degrees of curvatures separately.

Name of DNA	Nucleotide sequence
Circular DNA 3*	CAGTTGGGTAATTTTTAGGGTTTTCCCAGTTTTGACGTTGTTTTT CGACGGAATTCCCTTTTTACGACTCACTTTTTGCCTTGACTAGAG GGTTTTTACCAAGCTTCTATTTTTGGTCTTTTGCCATAACTTTT TATAGCATACATTTTACGAGTTTTATAAGCTGTTTTTCATGAGGC TTCTTTTTTATAGGTTTTTGTTCATGATTTTAATGGTATCTTTTTTCG TCGGTGGCATTTTTTCGGGGTTTTGCGCGGATCCCCTTTTTGTTTA TGGGCCTTTTTACATCAGGTTTTTCCGCTCAGCAATGATTTTTTG CCCTTTTAGATTTTTCAATGATATTTTTAGGCGTTTTTGACGTTT TCAGTTTTTCCGTGTCGCCCTTTTTCCCTTTTTTGCGCATTTTTT CGGCACTTTTTTGCATATATTTTTTGAGTTGTTTTGATCCGTTT TGATTTTCAGTGCGTTTTTGGCCATTTTGCCTAGTTTTTTCGTTT GCTATTTTTTGTTAATTTTTGCCAATTTTCGTATTTTCGCTATTT TTTGGCATTTTTTTGACCATTTTTCTTGTTTTGGATGGTTTTTCGGC CGTTTTTTGGAGTTGAATTTTACGTCCAGATTTGATTTTCAGTGC GTTTTTTGGCCATTTTGCAGTGTTTTTGCCTCTGCTATTTTTTGT TAATTTTTTGCCAATTTTCGAGGTATTTTCGCTATTTTTTGGCATT TTTTGTCTCATTTTTTAGTCGTTTTTGGCCATTTTTCCTGTTTTT GCTCACCCATTTTCGCTGGTGCCGAGTTTTTGATGCTTTTTTGACG TTTTGTGCACGAGTTTTTGACATCGGACTGGTTTTTCACAGCGGTT TTCAGGCTTTTTGCACAACATTTTTCATGTATTTTGAAGGAGAGA AGATTTTGGGCTCAGTTTTGATACCCGACGATTTTGACACCACGA TTTTTGCAGGCGTTTTTGAAGCATTTTTGCGTATAGCATTTTGGA TACGATGTTTTCCATGTTTTAGTGGTTTTTGACGTAGCTTTTCGA AGCTTTAGTCATTTTTTATAGCTGTTTTTGTGTGAGATTTTTATCC GCTCACTTTTCGAATTCCTTTTTACGAGCCGGATTTTTGCGGTGT GGCTTTTTGTCCGTTTTTCTGTTTGAGCT

Circular DNA 4	CTTCTATTCTAATTGTTTGTTGATTTATATGTGTATTGTGTTTCGC TATTATGTTGTGTAATCATGTGTTACTATCTATTGCTTGTATGTT AAGTTGTTGCTTCAGAGTTGTTTCTGATTCATGGTATATGTTGTG TTGTTAATGTTGTTATTGATGTGATGGTCTGTTTCATATTGGTTG TCGGTATTTATGTCTCTCTTGTCCCTTCTATCATTGGTTATGTTCT CCTGTTTATGCTTGGTTATTCACCTTCTGTGTTCTTCTGATTTACT GTATTCTTGGTATTAGTAGTTTATCTTGATTCTTGTGATTACGTA TTGTTGATGCGTTGCTGTGTTCCCTTCGTTGGTTGCATTCTATTCC TGTTTGTAATTGTCCTTGGCTTATCTTCGTTTCGTGTATTTTCGTCT CGCTCTGTCGTATCATTATGATTATCGGTTTGGTTGATGTGTGTG ATTCGTTGATGGAATTCGTAATGGCTGGTGTGTACTTGTCTGGT TAGTATTGCATTTACTTGGTTGCTATTCTCTCGTTTCAGTCGTCT CTCTTGGTGATTTCTCACCTTGATATCCTTATTCCTTTGTCGTGTG TTTATTATTATGTTGTATTGTTGTTGGTAGTTGGAATCGTATTCT GATACTATGATCTTGCTATTCTATGTACTGCTTCGTTGTAGTTTCG TTCTCCTTCATTAGTATCCTTGTCTTCATTAATATGGTATTGATT ATCCTGATATTCAATAGTTATTGCTGTTTCATTTGTTTCTTGATG TGTGTTATTCTAATTAGTTATTGGTTAATTGGTTGTATCATTGCT TATGCTGATTTGTCGTCGCAAGCTTATGATCTTAATCTCTTATCG TGTGTTTCGCGTTGTTCCATTGTTTCGTCACCTTCGTAGATTAGATC TTAGGTTCTTTATCTTGTGATCTTCAGATTTGGTTTCTGCAGCGT TATCTGCTGTATCTTGCTTACATTTATCACCTCCTCTACCTTCTG GTTTGTTTGCCGGATCTTGTGCTATCCTTCTCTTTGGTTCCGTAT GTAATTGGCTTCAGCATTGTTACCTTATTCTGTCTTCTATTGTT GTTGTTGTTAGGCCTCCATTTCTTGTTCTCTGTAGCTCCGCCTAC TTTCCTCGCTCTGCTTATCTTGTGAGCT
----------------	---

As a control experiment, a different circular DNA (Circular DNA 4) was designed and synthesized in our studies (Figure 2.5 and Table 2.2) that possesses the same nucleotide composition as Circular DNA 3 but contains no spaced adenine tract in its sequence.

Figure 2.5 Schematic representation of our design of Circular DNA 4 in relaxed forms. See Table 2.2 for detailed information about the nucleotide sequences of Circular DNA 4.

2.2.3 Design of Double Interwound Structures of DNA

With the purpose of illustrating that more than one self-crossing point of duplex DNA could be generated in some relaxed forms of DNA, an additional circular DNA (Circular DNA 5) was subsequently designed in our lab (Figure 2.6). This new DNA circle contains four segments of consecutive spaced adenine tracts that spread into two opposite strands of the DNA circle in alternate manners (Figure 2.6 and Table 2.3). In addition, Circular DNA 6 was designed in our studies (Figure 2.7 and Table 2.3) as a control experiment.

Figure 2.6 Schematic representation of our design of Circular DNA 5 with double interwound structures. See Table 2.3 for detailed information about the nucleotide sequences of Circular DNA 5.

Table 2.3 Nucleotide sequences of Circular DNA 5 and Circular DNA 6. Only one of the two strands of DNA from 5' end to 3' end is shown in the table. * (1) Junctions between the segments that are highlighted in red and in blue represent the points at which adenine-rich and thymine-rich sequence alternate between the two opposite strands. (2) This DNA contains 4 segments of continuous spaces adenine tracts.

Name of DNA	Nucleotide sequence
Circular DNA 5*	CGAAAAGTGC AAAAAGTCG GAAAAATCCGTGCA AAAATCGTCAAAA GAATTCA AAAAATAGCTA AAAATCGTCGAAAACTGCGTTGAAAAA GCTTAAAAACGATGCA AAAAGTGCATTCA AAAATGGGCAAAAAGTG GCCAAAAAGCTATAAAAAAACGCGCA AAAATCGCAC TTTTTTGCAT ATATTTTTTGGACGTTGTTTTGATCCGTTTTGATTTTCAGTGCGTT TTTGGCCATTTTGGCCTAGTTTTTTGCGTTGCTATTTTTTGTTAAT TTTTGCCAATTTTCGGTATTTTCGCTATTTTTTTGGCATTTTTTTGAC CATTTTTCTTGTTTTGGATGGTTTTGCCGGA AAAAAGTGCGGAAAAG TGCAAAAAGTCG GAAAAAGGACTCA AAAAGTGGCCAAAAATGCCGAA AAATCAGCGAAAAGGATTCA AAAATTCACTAAAACCCAAAATGGCT GAGAAAATGGGCAAAAAGTG GCCAAAAAGCTATAAAAAATCCGTG CAAAAATCGTCAAAAAGGCCCGAAAAAATAGCTAAAAGCAATGAAA AACTGCGTTGAAAAAGGTTAAAAACGATGCA AAAAGTGCATTCAAA AACGCGCAAAAACCGCA AAAATCTCTCA AAAATGAGGTAAAAATGG CCGTAAAAATGGTAAAAGGAGTGAAAAGCCCGAAAAATGGTCCTAA AAATCGCACCGGCCG TTTTTTGGAGTTGAATTTTACGTCCAGATTT GATTTTTCAGTGCGTTTTTGGCCATTTTGCTGTTTTTGCCTCTGCT ATTTTTTGTAAATTTTTGCCAATTTTCGAGGTATTTTCGCTATTTT TTGGCATTTTTTTGTCTCATTTTTTGTAGTCGTTTTTGGCC AAAAAGTG GAAAAAGCAGTGAAAAGTG GCCAAAAATGCCGAAAAATCAGCGAAA AGCCCGAAAAATGGCCGTAAAAATGGTAAAAGGAGTGAAAAATTCA CTAAAACCCAAAATGGCTGAGAAAAGAATTCA AAAATGGTCCTAAA AACCGCA AAAATCTCTCA AAAATGAGCT
Circular DNA 6	CTAGATCATAGTCGCAATTAACAGATTAAGTTGAGTAACACCAGAG TTCACAGTCACGAAGTTGTAATTAACGACGACCAGTCAGTAATACG ACTCACTTAAGACATTGACTAGAGGATAACCAACATAGGTATATAGA ACCAATCTAGAGCCATAACTTCGTATAGAATACATTATACGAAGTT ATATAAGATGTCAAACATGAGAATTATTGTTATAGGTTAATGTAAT GATAATAATGATTTCTTAGAAGTCAGATGACACTTTTCAGAGAAAT GTAAGCAGAACACATATTTATTTATTTCTAAATACATTCAAATATG TATCAGCTCATGAGACAATAACCATGATAAATGATTCAATAATATT GAATTAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCACTTAT TCCCTTATAAGCGACTTATTGCCTTCATGTTCCCTTTGATCACCCAG GAATTCTGGTGAAAGTAAGAGATGATGAAGATAAGTTGGGTGAACG AATGGATTACATAGAACTGGATCTCAACAGAGGTAAGTTAAGATTT GCACAACATGAAGGATCATGTAAC TAGAATTGATAGAAGGAGAGAA GAGATGGAGCTCAATGAAGCCATAACCAAACGACGAGCGTGACACAA CGATGACTGCAGGAATTAATAGAGCCATAACTTAGTATAGCATACA TTATACGAAGTTATCCATGGACTAGTGAGTCGTATTACGTAGATTG GAGTAATAATGGTCATAGCTGTTTACTGTATGAAATTGTTATAAGC TCACAATTACACACAACATACGAGCCGGAAGCATAAAGTGTAAGT GAGAGGAATTAACCATGGATCAGGTAAGTGATATCGAAGACTTAAC GCTAGAATTCGATAACCTATAGTGAGTCGTATTACATGGTCATAGC TGTTCTGGCAGCTCTGACCAATGTCTCAATCAATCTATGATGTTAC ATTGCACAAGATAAAGGAATATATCATCATGAACAATAACCAACTG TCTGATTACATAAACAGTAATACGAGCT

Figure 2.7 Schematic representation of our design of Circular DNA 6 in relaxed forms. See Table 2.3 for detailed information about the nucleotide sequences of Circular DNA 6.

2.2.4 The DNA Sequence of Decatenated Kinetoplast DNA Minicircles

A kinetoplast is a massively catenated network of DNA circles that are found inside a large mitochondrion in protozoa of the class Kinetoplastida. One of the most intriguing characteristics of these kinetoplast DNAs is their possession of high degrees of intrinsic curvatures in their duplex chains. It is our speculation that since the thymine/adenine tracts are widespread in the structures of kinetoplast DNA circles, the overall effects of intrinsic curvature could accumulate along its DNA backbones and lead to the generation of non-zero writhe number in its circular structure. Kinetoplast

DNA minicircles (2561 base pairs in length) were obtained through decatenation of kinetoplast DNA with Human Topoisomerase II and the DNA sequence was shown in Table 2.4.

Table 2.4 Nucleotide sequences of kinetoplast DNA. Only one of the two strands of DNA from 5' end to 3' end is shown in the table.

Name of DNA	Nucleotide sequence
kinetoplast DNA	GTGGATCCTCGTCGCAAAACCTCGAGTGCGATGTTGTGTTGATAGC TTCTTGTAGTTTTTCGTTGTTTGTTAATGTTGGTGTTGGTGTTGGT GTTCTCGGTTGCCACCTGTGGTTTCTTTAAGTGTTGTTGCTGTTT ATTTTGTGTTGTTGTTGGTTATTGGTTTATTGTTTGCATTAGCCTTC TGTGGGTTTGAAACTGTTGTATTCTTGTTTACTTGGGTGGTTTATC TTGATTTGGCTTTATTGTTGGGTACTTGTTGTTGTTTGTGTTGTTT TATGCTCTTTCTTTGTTGCTGGTGCTTGCTGAACTGTTTGTGGTTG TTGGGGCGTGTTGGGTTTGAGGGTGTTTTTTGGGGTGGTTTGGGGTG CCCGCGAAATATCAGAAATGGTCTCGGGTAGGGGCGTTCTGCGAAA ATCGACTTTTGATACAGGAAATCCCGTTCAAAAATGGCAGTTTTCT CGATTTTGGAGGCTCGGCTGGGATTCCGGGGTTGGTGTAGTCATT CCTGGGTCCGGGCGGGTCTGGCGGGGGTTCTGTTAAACGCGGGGGT TGCTTCAGTGCTGTTATTCATCCGCTTCGAAGTTAATTTTCGTTGT TTAGCTTGTAGTTTGCTCTGTGGGGTTCTGAAATTGCCCATTTTGG CGCTTTTTATCGTTGGGTGTGTACGATTGCGCGGCGTCTGCTTTTCA CGACGGGGCCGAGTGTTCTTGACGAGGTCGGGAGCGCTAGCCCGT CGTTGAATGCAAGTGCAACATACGTGAGGCCGCGGACGAGCCCCGT CCCTGAAAGGGGAGGAGGCTAGTTGACGCTAGGCCGGAGCCACGAA TGGCGAGCAAAGCTAGCCCGAGCCATGAACGCGAACGGCCGGGGAG ACTTGCCGGGGAAAGGGGAGGGTCAAGTACCAGGCTCGAACAGTAT ACAACGACAAGACGCCGCTGCATCGCCATACTTTTATCTTTCGCAC ATTCATGTGTGAAGTAGTTTGCTTTAACACGGTGCCTCGTTTAACC TCTTGCGGGTTGGTAGACAGACTCTAAAGCAGATGCGTAGACGTTT AGATTTTGATTTTTGAGTGCGTTTTTGGCCATTTTTTGCCCATTTT TCCCTTAAAATTCAATAAAATTGCGGGATTTTTTACCATTTTTGTC GATTTTTGGGGTATTTTCGCTGTTTTTTGGCATTTTTTTGCCATTT TTCCTTGATTTTGGGCACTTTTCGGGCTCCAAAAAAGTAACCTCGC GATTTTCGCCTGGAATTTTAGGCCTCCTGGCAGGGGGTTTGGCGGG GTTCTAGCCCGATTTTCGGGGCGTTCTGCGGGGGTTTTTTTCTGGTC TGGGCGCGGGTTTGGGCTGGTTTGGGCTGGGTTTGGACTGTTTGTG CTAGTTGGGCGCTACGGAAGCTTTGCGATGGTGCGCGGGGGGGTG GTTTCACCACTATTCTGATTGTTGTTTTCGCTCCTTGGTGGGGTTT ATATGCGCTCCGTTCCGTCGTATTCTGGAATTTTGGGGTTTGCCAA

	AAGTGAAC TTCCGACATTTCTCGCGGGGTTAATATATAGACTAGAC GCGTCGTTGTTAATTTTGCCATGGGTGTGTTTGTGTTGTTCTGGTG CCCGGAGGCTGATTTCCGGGGTCCCGCGAAAAATCAGAAACGGTCT CGGGTAGGGGCGTTCTGCGAAAATCGACTTTTGATACAGGAAATCC CGTTCAAAAATCGTGATTTTTCATTTTGGAGGCAAAC TGGGGAT TTCCGGGGTTGGTGTAGTATTTCTGGGTCCGGGGGTCC TGAGGGGT TCCAATACCTTCTGATAGATTCGCCTTTTATAGGCGTTCTGCTCGT TACTTTTATAACTTTAGTTGCTCTTATGTTTGCTATAAATATATAG CTTTGATTTCTAGACTTGCTTGCGTTTAAAGTTGTTTGCGCGGGCT TCCTGTGGGTTTTGTTTTGGTGTGGCTTGTTATTTGTGATTTTGCT AGTTTCTTTGCGGTTTTGTCTATTTT TAGTTGTTTTGTGTATTTGT ACTTTACGTTTTTTGGTTGTTGTGGCTTTGCGTTTTTATACTGCTT TGCTGGCTTGTTGGTTATGTTGGCTTGTTGGTTTTGTTTTTATTTT GTGTGTTTCGTGGGTGTTGATGTTTTTGTGTTTTTTGGTTGCTTTTG TAGCTTTAGGGTGGTTACTATTAGTTTTTCCTTTTGTTTTCGCTTTT GTTCTGGGGTTTGTGATTAGCTTTGGGGGTTTCGTGGTTGTTGTGC CTGTGTTATTTAGTTGTGTCCACGGTGGGTTCGGCTGCTGGTTGG GTGTGCTTACTGTTTCTTGTTATGTTGGTATGTATGCTATGTTGCT GCTAGTTGTTTTTATGGTTTTGCGCTTGCTGTTGCGTGTGTATGT GTTTATTTATTTGATTGTTTAGATTGTTTAAATAACTTTGTGTTGC ATTTGTTTTAGATTTAAAGGCTTGTTGTTGTGTTGTTGTGTTGTT GCTATTGTTTTGATTGTCTTTGCTGCTCACTGCGTGGTACACATT GATTGCTCGAGGGGGTTAACCATGGATCCGG
--	---

2.2.5 Design of Plasmid DNA Containing the DNA Sequence of kinetoplast DNA

With the aim of examining whether the self-crossing affiliated with kinetoplast DNA could still be upheld, the entire sequence of the kinetoplast DNA is inserted into a plasmid vector (pOK12) that contains no identifiable intrinsic curvature. The newly constructed DNA (Circular DNA 7) contains 4695 base pairs in length, in which 2561 base pairs are from kinetoplast DNA and the rest (2134 base pairs) belongs to pOK12 vector and the DNA sequence was shown in Table 2.5.

Table 2.5 Nucleotide sequences of Circular DNA 7. Only one of the two strands of DNA from 5' end to 3' end is shown in the table. The segment of nucleotide sequence highlighted in red is belongs to kinetoplast DNA while segment of nucleotide sequence highlighted in blue belongs to pOK12 vector.

Name of DNA	Nucleotide sequence
Circular DNA 7	<p>CCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCAATATGCAGCTGGCAGACA GGTTTCCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCA TTAGGCACCCAGGCTTTACACTTTATGCTTCCGCGGCTCGTATGTTGTGGAATTGTGA GCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGCCACTAGTCCGAGGCC TCGAGATCTATCGATGCATGCCATGGTACCCGTGGATCCTCGTCGCAAAACCTCGAGTGCG ATGTTGTGTTGATAGCTTCTTGTAGTTTTTCGTTGTTTGTAAATGTTGGTGTGGTGTGG TGTTCTCGGTTGCCACCTGTGGTTTTCTTAAGTGTGTTGTGCTGTTATTTTGTGTTTGT TGGTTATTGGTTTATTGTTTGATTAGCCTTCTGTGGGTTTGAAACTGTTGTATTCTTGT TACTTGGGTGGTTTATCTTGATTGGCTTTATTGTTGGTACTTGTGTTGTTTGTGTTGT TTTATGCTCTTTCTTGTGCTGGTGCTTGCTGAACGTTTGTGGTGTGGGGCGTGTGG GTTTGAGGGTGTGTTTTGGGGTGGTTGGGGTGCCCGCGAAATATCAGAAATGGTCTCGGG TAGGGGCGTTCTGCGAAATCGACTTTTGATACAGGAAATCCCGTTCAAAAATGGCAGTTT TCTCGATTTTGAGGCTCGGCTGGGATTTCCGGGGTGGTGTAGTCATTCTGGGTCCGGG CGGGTCTGGCGGGGGTCTGTAAACGCGGGGGTTCCTTCAGTGCCTGTTATTCATCCGCTT CGAAGTTAATTTTCGTTGTTTAGCTTGTAGTTTGCTCTGTGGGGTCTGAAATTGCCCAT TTGGCGCTTTTATCGTTGGGTGTGTACGATTGCGCGGCTCGCTTTCGACGACGGGGCCG AGTGTCTTGACAGAGTCCGGAGCGCTAGCCCGTCGTTGAATGCAAGTGCAACATACGTG AGGCCGCGGACGAGCCCCGTCCCTGAAAGGGGAGGAGGCTAGTTGACGCTAGGCCGAGCC ACGAATGGCGAGCAAAGCTAGCCGAGCCATGAACGCGAACGGCCGGGGAGACTTGCCGGG GAAAGGGGAGGTCAAGTACCAGGCTCGAACAGTATACAACGACAAGACGCCGCTGCATCG CCATACTTTTATCTTTCGCACATTCATGTGTGAAGTAGTTTGCTTTAACACGGTGCCCTCGT TTAACCCTCTGCGGGTTGGTAGACAGACTCTAAAGCAGATGCGTAGACGTTTCAAGTTTGA TTTTTGAGTGCGTTTTTGGCCATTTTGGCCATTTTCCCTTAAATTCATAAAATTGCG GGGATTTTTTACCATTTTTGTGCGATTTTGGGGTATTTTCGCTGTTTTTGGCATTTTTTG GCCATTTTCTTGATTTTGGGCACTTTTGGGCTCCAAAAAGTAACCTCGCGATTTTTCG CCTGGAATTTTAGGCCCTCTGGCAGGGGGTTTGGCGGGGTTCTAGCCCGATTTCCGGGCGT TCTGCGGGGGTTTTTCTGGTCTGGGCGCGGGTTTGGGCTGGTTTGGGCTGGGTTTGGAC TGTTTGTGCTAGTTGGGCGCTACGGACTGCTTTCGCGATGGTGCGCGGGGGGTGGTTTAC CACTATTCTGATTGTGTTTTCGCTCCTTGGTGGGTTTATATGCGCTCCGTTCCGGTCGTA TTCTGGAATTTTGGGGTTTGCAAAAAGTGAAGTCCGACATTTCTCGCGGGGTTAATATAT AGACTAGACGCGTCGTTGTTAATTTTGCCATGGGTGTGTTTGTGTTGTTCTGGTGCCCGGA GGCTGATTTCCGGGGTCCCGCAAAAATCAGAAACGGTCTCGGGTAGGGGCGTTCTGCGAA AATCGACTTTTGATACAGGAAATCCCGTTCAAAAATCGTGATTTTTTCAATTTTGGAGGCA AACTGGGGATTTCCGGGGTGGTGTAGTATTTCTGGGTCCGGGGGTCTGAGGGGTCCAA TACCTTCTGATAGATTGCGCTTTTATAGGCGTTCTGCTCGTTACTTTTATAACTTTAGTTG CTCTTATGTTTGTATATAATATATAGCTTTGATTTCTAGACTTGCTTGCGTTTAAAGTTGT TTGCGCGGGCTTCCTGTGGGTTTTGTTTTGGTGTGGCTTGTTATTTGTGATTTTGTAGTT TCTTTGCGGTTTTGTCTATTTTGTAGTTTGTGTATTTGTACTTTACGTTTTTTGGTTG TTGTGGCTTTGCGTTTTTATACTGCTTTGCTGGCTTGGTTGGTTATGTTGGCTTGTGGTTT GTTTTTATTTTGTGTGTTCTGTTGGTGTGATGTTTTTGTGTTTTTGGTTGCTTTGTAG CTTTAGGGTGGTTACTATTAGTTTTCCCTTTTGTGTTTTGCTTTTGTCTGGGGTTTGTGATT AGCTTTGGGGTTTTCGTGGTGTGTGTCCTGTGTTATTTAGTTGTGTCACGGTGGGTTTC GGCTGCTGGTTGGGTGTGCTTACTGTTTCTTGTATGTTGGTATGTAGCTATGTTGCTGC TAGTTGTTTTTATGGTTTTGCGCTTGTCTGTTGCGTGTGTATGTGTTATTTATTTGATTG TTTAGATTGTTTTAATAACTTTGTGTTGCATTTGTTTTAGATTTAAAGGCTTGTGTTGT GTTGTTGTGTTGTTGCTATTGTTTTGATTTGTCTTTGCTGCTCACTGCGTGGTACACATTG ATTGCTCGAGGGGTTAACCATGGATCCGGGGGAGCTCGAATTCGAAGCTTCTGCGAGACGC GTGACGTCATATGGATCCGATATCGCCGTGGCGGGCGCTCTAGAACTAGTGGATCGATCC</p>

	<p> CCAATTGCCCCTATAGTGAGTCGTATTACAATTCCTGCGCGTCGTTTTACAACGTCGTGA CTGGGAAAACCCTGGCGTTACCCAACCTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGC TGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATG GCGAATGAGCTTGCGCCGTCCCGTCAAGTCAGCGTAATGCTCTGCCAGTGTTACAACCAAT TAACCAATTCTGATTAGAAAAAATCATCGAGCATCAAATGAAACTGCAATTTATTTCATATC AGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAAACTCACCG AGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCGTCCAACAT CAATACAACCTATTAATTTCCCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCATG AGTGACGACTGAATCCGGTGAGAATGGCAAAAGGTTATGCATTTCTTCCAGACTTGTTC ACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATCAACCAAAACCGTTATTTCATTC GTGATTGCGCCTGAGCGAGACGAAATACGCGATCGCTGTTAAAAGGACAATTACAAAACAGG AATCGAATGCAACCGCGCAGGAACACTGCCAGCGCATCAACAATATTTTACCTGAATCA GGATATTCTTCTAATACCTGGAATGCTGTTTTCCAGGGATCGCAGTGGTGAGTAACCATG CATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAATCCGTCAGCCA GTTTAGTCTGACCATCTCATCTGTAAATCATTGGCAACGCTACCTTTGCCATGTTTCAGA AACAACTCTGGCGCATCGGGCTTCCCATACAATCAATAGATTGTGCGACCTGATTGCCCCGA CATTATCGCGAGCCCATTTATACCCATATAAATCAGCATCCATGTTGGAATTTAATCGCGG CCTCGACGAGCAAGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGTTT ATGTAAGCAGACAGTTTATTGTTTCATGATGATATATTTTTATCTTGTGCAATGTAAACATC AGAGATTTTGAGACACTCGACAAGATGATCTTCTTGAGATCGTTTTTGGTCTGCGCGTAATC TCTTGCTCTGAAAACGAAAAAACCGCCTTGCAGGGCGGTTTTTCGAAGGTTCTCTGAGCTA CCAACCTCTTTGAACCGAGGTAACCTGGCTTGGAGGAGCGCAGTCACCAAAACTTGTCTTTC AGTTTAGCCTTAACCGCGCATGACTTCAAGACTAACTCCTCTAAATCAATTACCACTGGC TGCTGCCAGTGGTGCTTTTGCATGCTTTCCGGTTGGACTCAAGACGATAGTTACCGGAT AAGGCGCAGCGGTCGGACTGAACGGGGGTTTCGTGCATACAGTCCAGCTTGGAGCGAATG CCTACCCGGAAGTGTGTCAGGCGTGAATGAGACAAACGCGGCCATAACAGCGGAATGA CACCAGTAAACCGAAAGGCAGGAACAGGAGAGCGCACGAGGGAGCCGCCAGGGGAAACGCC TGGTATCTTTATAGTCTGTGCGGTTTCGCCACCACTGATTTGAGCGTCAGATTTCTGTGAT GCTTGTGAGGGGGGCGGAGCCTATGAAAAACGGCTTTGCCGCGGCCCTCTCACTTCCC </p>
--	---

2.2.6 Design of Plasmid DNA Containing the Replication Origins of Bacteriophage λ

It has been well established in the past that besides kinetoplast DNA, intrinsic DNA curvatures could occur at the replication origins of Bacteriophage λ ¹³⁵, Yeast APS1¹³⁷ and Simian Virus¹³⁸. A circular DNA (Circular DNA 8) carrying the repeats of nucleotide sequence at the replication origins of Bacteriophage λ was subsequently designed during our investigations. This new circular DNA is composed of 2641 base pairs in its sequence in which 507 base pairs are the repeats of nucleotide sequence from the replication origins of Bacteriophage λ and 2134 base pairs are the nucleotide sequence from vector pOK12. Unlike kinetoplast DNA in which short adenine tracts

are widespread in its entire DNA circle, only the segment of 507 base pairs (~1/5 of the entire length) in Circular DNA 8 (2641 base pairs in the entire sequence) contains short adenine tracts. One of our aims in designing Circular DNA 8 is to examine whether the potential writhe associated with ~1/5 of the sequence of Circular DNA 8 (507 base pairs of short adenine tracts) could be cancelled out by the rest of ~ 4/5 of non-adenine tract-rich sequence (2134 base pairs). If the curvature associated with 507 base pairs of short adenine tracts could indeed persist, our subsequent aim is to look at (1) whether Circular DNA 8 could display only a single writhe number (one backbone self-crossing section) associated with the short adenine-rich segment of 507 base pairs; and (2) whether one small (507 base pairs) and one big circle (2134 base pairs) could be seen in the AFM images of single Circular DNA 8 molecules. The entire DNA sequence of Circular DNA 8 was shown in Table 2.6. As control experiments, vector pSP73 were used, which possesses a similar length to Circular DNA 8 and has no identifiable intrinsic curvature in its structure. The sequences of vector pSP73 as well as the DNA sequence of vector pOK12 were shown in Table 2.7 and Table 2.8 respectively.

Table 2.6 Nucleotide sequences of Circular DNA 8. Only one of the two strands of DNA from 5' end to 3' end is shown in the table. The segment of nucleotide sequence highlighted in red belongs to the repeats nucleotide sequence of replication origins of Bacteriophage λ while segment of nucleotide sequence highlighted in blue belongs to pOK12 vector.

Name of DNA	Nucleotide sequence
Circular DNA 8	<p> CCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCT GGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAA TGTGAGTTAGCTCACTCATTAGGCACCCAGGCTTTACACTTTATGCTTCCG CGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTACACAGGAAA CAGCTATGACCATGATTACGCCACTAGTCCGAGGCCTCGAGATCTATCGATG CATGCCATGGTACCCGGACCAAATAAAAACATCTCAGAATGGTGCATCCCTC AAAACGAGGGAAAATCCCCTAAAACGAGGGGATAAAACATCCCTCAAATTGGG GGATTGCTATCCCTCAAACAGGGGGACACAAAAGACACTATTACAAAAGAA AAAAGAAAAGATATTCGTAGAGAATTTCGGACCAAATAAAAACATCTCAGAA TGGTGCATCCCTCAAACGAGGGAAAATCCCCTAAAACGAGGGGATAAAACAT CCCTCAAATTGGGGGATTGCTATCCCTCAAACAGGGGGACACAAAAGACAC TATTACAAAAGAAAAAAGAAAAGATATTCGTAGAGAATTTCGGACCAAATAA AAACATCTCAGAATGGTGCATCCCTCAAACGAGGGAAAATCCCTAAAACG AGGGATAAAACATCCCTCAAATTGGGGGATTGCTATCCCTCAAACAGGGGG ACACAAAAGACACTATTACAAAAGAAAAAAGAAAAGATATTCGTAGAGAAT TCGGGAGCTCGAATTCGAAGCTTCTGCAGACGCGTCGACGTCATATGGATCC GATATCGCCGTGGCGGCCGCTCTAGAAGTAGTGGATCGATCCCCAATTCGCC CTATAGTGAGTCGTATTACAATTCAGTGGCCGTCGTTTTTACAACGTCGTGAC TGGGAAAACCCCTGGCGTTACCCAACTTAATCGCCTTGACGACATCCCCCTT TCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACA GTTGCGCAGCCTGAATGGCGAATGAGCTTGCGCCGTCCCGTCAAGTCAGCGT AATGCTCTGCCAGTGTTACAACCAATTAACCAATTCTGATTAGAAAACTCA TCGAGCATCAAATGAACTGCAATTTATTCATATCAGGATTATCAATACCAT ATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAAACCTCACCGAGGCAGTT CCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCGTCCAACA TCAATACAACCTATTAATTTCCCTCGTCAAAAATAAGGTTATCAAGTGAGA AATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAGGTTATGCAT TTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCA CTCGCATCAACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGACGAA ATACGCGATCGCTGTTAAAGGACAATTACAAACAGGAATCGAATGCAACCG GCGCAGGAACACTGCCAGCGCATCAACAATATTTTCACCTGAATCAGGATAT TCTTCTAATACCTGGAATGCTGTTTTCCAGGGATCGCAGTGGTGAGTAACC ATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAA TTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTTGGCAACG CTACCTTTGCCATGTTTCAGAAACAACCTCTGGCGCATCGGGCTTCCCATACA ATCAATAGATTGTGCGACCTGATTGCCCCGACATTATCGCGAGCCCATTTATA CCCATATAAATCAGCATCCATGTTGGAATTTAATCGCGGCCCTCGACGAGCAA GACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGTTTATGT AAGCAGACAGTTTTATTGTTTCATGATGATATATTTTTATCTTGTGCAATGTA ACATCAGAGATTTTGAGACACTCGACAAGATGATCTTCTTGAGATCGTTTTG GTCTGCGCGTAATCTCTTGCTCTGAAAACGAAAAAACC GCCTTGCAGGGCGG TTTTTCGAAGGTTCTCTGAGCTACCAACTCTTTGAACCGAGGTAAC TGGCTT GGAGGAGCGCAGTCACCAAACTTGTCTTTTCAGTTTAGCCTTAACCGGCGC ATGACTTCAAGACTAACTCCTCTAAATCAATTACCAGTGGCTGCTGCCAGTG GTGCTTTTGCATGTCTTTCCGGGTGGACTCAAGACGATAGTTACCGGATAA GGCGCAGCGGTCGGAAGTGAACGGGGGGTTCGTGCATACAGTCCAGCTTGGAG CGAACTGCCTACCCGGAAGTGAAGTGCAGGCGTGGAATGAGACAAACGCGGC CATAACAGCGGAATGACACCGGTAAACCGAAAGGCAGGAACAGGAGAGCGCA CGAGGGAGCCGCCAGGGGAAACGCCTGGTATCTTTATAGTCCTGTCTGGGTTT CGCCACCACTGATTTGAGCGTCAGATTTTCGTGATGCTTGTGAGGGGGCGGA GCCTATGAAAAACGGCTTTGCCGCGGCCCTCTCACTTCCC </p>

Table 2.7 Nucleotide sequences of vector pOK12. Only one of the two strands of DNA from 5' end to 3' end is shown in the table.

Name of DNA	Nucleotide sequence
Vector pOK12	CCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCT GGCACGACAGGTTTCCCGACTGGAAAAGCGGGCAGTGAGCGCAACGCAATTAA TGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTTATGCTTCCG CGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTACACAGGAAA CAGCTATGACCATGATTACGCCACTAGTCCGAGGCCTCGAGATCTATCGATG CATGCCATGGTACCCGGGAGCTCGAATTCGAAGCTTCTGCAGACGCGTCGAC GTCATATGGATCCGATATCGCCGTGGCGGGCCGCTCTAGAACTAGTGGATCGA TCCCCAATTTCGCCCTATAGTGAGTCGTATTACAATTCAGTGGCCGTCGTTTTT ACAACGTCGTGACTGGGAAAACCCGTGGCGTTACCCAACCTAATCGCCTTGCA GCACATCCCCCTTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATC GCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGAGCTTGCGCCGTCCC GTCAAGTCAGCGTAATGCTCTGCCAGTGTTACAACCAATTAACCAATTCTGA TTAGAAAACTCATCGAGCATCAAATGAACTGCAATTTATTTCATATCAGGA TTATCAATACCATATTTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAAAC CACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCC GACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAAATAAGG TTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAAATGGCA AAAGGTTATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTC GTCATCAAAATCACTCGCATCAACCAAACCGTTATTTCATTTCGTGATTGCGCC TGAGCGAGACGAAATACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAA TCGAATGCAACCGGCGCAGGAACACTGCCAGCGCATCAACAATATTTTCACC TGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTCCAGGGATCGCA GTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCG GAAGAGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAAC ATCATTTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACCTCTGGCGCATCG GGCTTCCCATACAATCAATAGATTGTCGCACCTGATTGCCCCGACATTATCGC GAGCCCATTTATACCCATATAAATCAGCATCCATGTTGGAATTTAATCGCGG CCTCGACGAGCAAGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTA TTACTGTTTATGTAAGCAGACAGTTTTATTGTTTCATGATGATATATTTTTAT CTTGTGCAATGTAACATCAGAGATTTTGAGACACTCGACAAGATGATCTTCT TGAGATCGTTTTTGGTCTGCGCGTAATCTCTTGCTCTGAAAACGAAAAAACCG CCTTGCAGGGCGGTTTTTTCGAAGGTTCTCTGAGCTACCAACTCTTGAACCG AGGTAACCTGGCTTGAGAGGAGCGCAGTCACCAAACTTGTCCTTTCAGTTTAG CCTTAACCGGCGCATGACTTCAAGACTAACTCCTCTAAATCAATTACCAGTG GCTGCTGCCAGTGGTGCTTTTGCATGTCTTCCGGGTTGGACTCAAGACGAT AGTTACCGGATAAGGCGCAGCGGTCGGACTGAACGGGGGGTTTCGTGCATACA GTCCAGCTTGAGCGAACTGCCTACCCGGAAGTGAAGTGTAGGCGTGGAATG AGACAAACGCGGCCATAACAGCGGAATGACACCGGTAAACCGAAAGGCAGGA ACAGGAGAGCGCACGAGGGAGCCGCCAGGGGAAACGCCTGGTATCTTTATAG TCCTGTCGGGTTTCGCCACCACTGATTTGAGCGTCAGATTTTCGTGATGCTTG TCAGGGGGGCGGAGCCTATGGAAAAACGGCTTTGCCGCGGCCCTCTCACTTC CC

Table 2.8 Nucleotide sequences of vector pSP73. Only one of the two strands of DNA from 5' end to 3' end is shown in the table.

Name of DNA	Nucleotide sequence
Vector pSP73	GAACCAGATCTGATATCATCGATGAATTCGAGCTCGGTACCCGGGGATCCT CTAGAGTCGACCTGCAGGCATGCAAGCTTCAGCTGCTCGAGGCCGGTCTCC CTATAGTGAGTCGTATTAATTTTCGATAAGCCAGGTTAACCTGCATTAATGA ATCGGCCAACGCGCGGGGAGAGGCGGTTTGCCTATTGGGCGCTCTTCCGCT TCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTTCGGCTGCGGCGAGCGGTA TCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAAC GCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAA AAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCAT CACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAA AGATAACAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCG ACCTTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTG GCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCTGTT CGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTTCAGCCCGACCGCTGC GCCTTATCCGGTAACCTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTA TCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTA GGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGA AGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAA AGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGT TTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAA GATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAACTCA CGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATC CTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAA ACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCG ATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATA ACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCG CGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCC GGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAG TCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGT TTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTCG TTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACA TGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTCCTTCGGTCCCTCCGATC GTTGTCAGAAGTAAGTTGGCCGAGTGTTATCACTCATGGTTATGGCAGCA CTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACT GGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGT TGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACT TTAAAAGTGCTCATCATTTGGAAAACGTTCTTCGGGGCGAAAACCTCTCAAGG ATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAAC TGATCTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACA GGAAGGCAAAATGCCGCAAAAAGGGAATAAGGGCGACACGGAAATGTTGA ATACTCATACTCTTCTTTTCAATATTATTGAAGCATTTATCAGGGTTAT TGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATA GGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACC ATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTT CGTCTCGCGCGTTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTC CCGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCC CGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACCTAT GCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGGACATATTGTCTG TTAGAACGCGGCTACAATTAATACATAACCTTATGTATCATACACATACGA TTTAGGTGACACTATA

2.3 Materials and Methods

2.3.1 Duplex DNA, Enzymes and Chemicals

Product(s)	Manufacturer
Duplex linear DNA precursors	Generay Biotech (Shanghai, China)
Plasmid DNA precursors	Generay Biotech (Shanghai, China)
Kinetoplast DNA	TopoGEN (Columbus, OH)
Vector pSP73	Promega Pte Ltd (Singapore)
Vector pOK12	Generay Biotech (Shanghai, China)
Human topoisomerase I	TopoGEN (Columbus, OH)
Human topoisomerase II	TopoGEN (Columbus, OH)
100 bp DNA ladder	Fermentas (Singapore)
1 Kb DNA ladder	New England Biolabs (Ipswich, MA)
SacI endonuclease	New England Biolabs (Ipswich, MA)
T4 DNA ligase	New England Biolabs (Ipswich, MA)
BAL 31 exonuclease	New England Biolabs (Ipswich, MA)
Biological purity water	1st Base Pte. Ltd (Singapore)

Agarose	Invitrogen (Carlsbad, CA)
Ethidium bromide	Research Biolabs (Singapore)
QIAquick PCR purification kit	Qiagen (Singapore)
QIAquick Gel Extraction Kit	Qiagen (Singapore)
Mini Prep Cell	Bio-Rad (Hercules, CA)
TAE, TBE, TRIS	1st Base Pte. Ltd (Singapore)

Most all the chemicals used in this research were listed above otherwise were obtained from Sigma-Aldrich with analytical grade or molecular biology grade.

2.3.2 Reactions of SacI with Duplex Linear DNA Precursors

The duplex linear DNA precursors containing two SacI digest site in each end were obtained from Generay Biotech (Shanghai, China). In order to create two cohesive ends as shown in the Figure 2.2 – Figure 2.7, those linear DNA were treated with SacI endonuclease. Linear DNA 1 – Linear DNA 6 were obtained as described as follows: A solution containing 10 mM Bis-Tris-Propane-HCl, 10 mM MgCl₂, 1 mM Dithiothreitol, duplex linear DNA precursors (200 ng) and 10 U SacI was incubated at 37 °C for 1 hr. The reaction products were further analyzed using agarose electrophoresis (1.5%) and purified using QIAquick PCR purification kit before the next steps.

2.3.3 Preparations of Circular DNA Using T4 Ligase

As there are two identical cohesive ends digested by SacI in all six linear DNA (Linear DNA 1– Linear DNA 6), Circular DNA 1 – Circular DNA 6 were obtain from ligase reactions¹³⁹⁻¹⁴¹ as described as follows: A 50 µl solution containing 50 mM Tris-HCl, 10 mM MgCl₂, 1 mM ATP, 10 mM dithiothreitol, 500 ng linear DNA and 20 U T4 DNA ligase was incubated at 16 °C for 8 hrs. The obtained circular DNA products were further analyzed using agarose electrophoresis (1.5%) and purified using QIAquick PCR purification kit before the next steps.

2.3.4 Degrade Linear DNA from Ligase Reaction Mixture Using Nuclease BAL-31 Exonuclease

Nuclease BAL-31 exonuclease degrades both 3' and 5' termini of duplex DNA without generating internal scissions.¹⁴²⁻¹⁴⁴ To remove the linear DNA from the mixture of ligase reaction as well as further confirm that there is no nicks, gaps and single-stranded regions in the obtained DNA products, we treated the ligase reaction mixture by nuclease BAL-31 exonuclease as described as follows: A 50 µl solution containing 20 mM Tris-HCl, 600 mM NaCl, 12 mM CaCl₂, 12 mM MgCl₂, 1 mM EDTA, 500 ng reaction products of ligase reactoin and 2 U exonuclease BAL-31 was incubated at 30 °C for 10 hrs. The obtained circular DNA products were further analyzed using agarose electrophoresis (1.5%) and purified next using Mini Prep Cell.

2.3.5 Digest Circular DNA by SacI Endonuclease

To further confirm that the obtained DNA is in a circular conformation, the circular DNA (after treated with exonuclease Bal-31) were digested by SacI. Because the circular DNA were synthesized from connection of the two cohesive ends made by SacI and there should be one SacI site in the whole circular DNA sequence, the digest product form circular DNA should have the same mobility shift as linear precursors when doing gel electrophoresis analysis. The reaction procedure is described as follows: A solution containing 10 mM Bis-Tris-Propane-HCl, 10 mM MgCl₂, 1 mM Dithiothreitol, mixtures of linear DNA and circular DNA (200 ng) and 10 U SacI was incubated at 37 °C for 1 hr. The reaction products were further analyzed using agarose electrophoresis (1.5%).

2.3.6 Reactions of Human Topoisomerase II with Catenated Kinetoplast DNA

Decatenated kinetoplast DNA minicircles were obtained from the reaction of human topoisomerase II¹⁴⁵⁻¹⁴⁷ with kinetoplast DNA. The reaction procedure is described as follows: A 50 µl solution containing 50 mM Tris-Cl (pH 8.0), 150 mM NaCl, 10 mM, MgCl₂, 5 mM ATP, 0.5 mM dithiothreitol, 0.1% BSA, 200 ng catenated kinetoplast DNA and 5 U Human Topo II was incubated at 37 °C for 1 hr. The obtained circular DNA products were further analyzed using agarose electrophoresis (1.5%).

2.3.7 Reactions of Human Topoisomerase I with Circular DNA and Plasmid DNA

To further confirm the obtained circular DNA is indeed in its relaxed form, all the circular DNA and plasmid DNA were treated with human topoisomerase I^{46, 148-149} before the following AFM examination. The reaction procedure is described as follows: A solution containing 10 mM Tris-HCl, 150 mM NaCl, 0.1% BSA, 0.1 mM Spermidine, 5% glycerol, 200 ng DNA (Circular DNA 1–8, Decatenated kinetoplast DNA, vector pOK12 and vector pSP73) and 3 U Human Topo I was incubated at 37 °C for 1 hr. The reaction products were further analyzed using agarose electrophoresis (1.5%) and further purified using QIAquick Gel Purification Kits.

2.3.8 AFM Examination of Obtained Circular DNA

Atomic Force Microscope (AFM) has been known to be a powerful tool for determining certain subtle alternations in DNA topological features.¹⁵⁰⁻¹⁵³ However, only the two-dimensional topological information can be given in the AFM image of DNA. The process of deposition of DNA molecules from the solution onto mica surface is considered to be governed solely by diffusion.¹⁵⁴ By controlling the condition of sample preparation,¹⁵⁵ DNA molecules can adopt kinetic trapping mode to preserve their conformation in 3-D solution and give us useful information of the topological structures.¹⁵⁵⁻¹⁵⁷

To conduct the AFM analysis, DNA molecules need to be fixed to some surface. Crystal mica, which has atomic level smooth surface, is typically used as substrate to study DNA molecules. On the other hand, because DNA and mica are both negatively charged, it is necessary to modify the mica surface or the DNA counter ion to allow binding. As a result, two methods for DNA sample preparation were chosen to absorb the DNA molecules onto the mica surface as described as follows.

(1) Counter ion method

The counter ion method is performed by adsorbing DNA onto the mica in the presence of cations. The divalent ion, Mg^{2+} for example, serves as a counter ion on the negatively charged DNA backbone and also provides additional charge to bind the mica.¹⁵⁸⁻¹⁵⁹ The procedure is described as follows: A solution containing ~ 10 ng DNA sample, 10 mM $MgCl_2$ and 40 mM HEPES (PH = 7.0) was deposited onto freshly cleaved mica and incubated for 5 minutes. Then mica is rinsed with deionized water and dried with nitrogen gas before AFM analysis.

(2) Silanized Mica method

As an alternative, a sample preparation procedure for AFM with the use of functionalized mica substrates (APS- mica) was used during our studies.¹⁵⁵ The major advantage of these sample preparation procedures is that they can work under a wide variety of ionic conditions, pH value and over a wide range of temperatures. The methodology of APS-mica allows routine visualization of topographic studies of alternative DNA structures. Due to the APS (*1-(3-aminopropyl)silatrane*) is not

commercially available, we synthesized it which was used to functionalize mica substrates next. The procedure is described as follows: A solution containing catalytic amount of sodium metal (5 mg) and 15.0 ml (16.8 g, 0.11 mol) of triethanolamine (Aldrich) was prepared in a 250 ml round-bottom flask under nitrogen atmosphere. A rubber balloon was connected to the flask to allow hydrogen to escape without building up pressure. The mixture was heated to 80 °C for 1 hour and cooled to room temperature when no bubble can be observed in the solution. An equivalent amount of (3-aminopropyl)triethoxysilane (26.4 ml or 25.0 g, 0.11 mol) is added to the mixture, then the flask was connected to a rotary evaporator under a 60 °C water bath for more than 24 hours. At the end of the reaction, about 17 g of ethanol was evaporated out and a white solid was obtained. The crude product was purified by crystallization from xylenes for better results and higher stability. The 18 g final product was obtained in a white solid. Minute amounts of sodium hydroxide in the product can hardly affect the performance of the reagent, or change the pH of stock solutions of APS.

Figure 2.8 Schematic illustration of the formation of functionalized mica substrates with APS. The initial adduct reacted with one hydroxyl group can reach with a second surface OH group forming the indicated product in a reversible equilibrium.¹⁵⁵

Figure 2.8 depicted the formation of APS-mica. The amino groups in solution become positively charged in a rather broad range of pH value. DNA, On the other hand, possesses a negatively charged backbone, which should adhere strongly to this functionalized mica substrates. The general procedure in our studies is described as follows: A 50 mM stock solution of APS was prepared and it can be stored at $-20\text{ }^{\circ}\text{C}$ for several months. A working solution was obtained by dissolving the stock solution in 1:300 ratio in water, which can be decomposed in the room temperature within 2 days. The working solution of APS was deposited onto freshly cleaved mica glued on a steel disc for about 30 minutes, and then the mica was rinsed with deionized water and dried with nitrogen gas. The newly prepared APS-mica sheets should be stored under nitrogen atmosphere and it can be used within several days. Once the APS-mica was ready to use, A solution containing $\sim 10\text{ ng}$ DNA sample, 10 mM Tris-EDTA (PH = 7.0) and 20 mM NaCl was deposited onto freshly cleaved mica and incubated for 5 minutes. Then mica is rinsed with deionized water and dried with nitrogen gas before AFM analysis.

AFM images 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.

2.4 Results and Discussion

2.4.1 Synthesis and Confirmation of Interwound Structures of DNA that Possesses Writhe Number of ± 1

With the aim of demonstrating that fabrication of an interwound DNA structure with writhe number of one in the absence of gyrase or reverse gyrase is feasible, a linear duplex DNA sequence was designed during our investigation that contains 676 base pairs in length as discussed above. The newly designed sequence possesses two consecutive spaced tracts of adenines with nearly equal lengths that spread in its two opposite strands respectively. Linear DNA 1 with two cohesive ends were obtained by the reaction of SacI endonuclease digestion. As demonstrated in Figure 2.2, a circular DNA (Circular DNA 1) was obtained in our studies after the action of T4 DNA ligase on these linear DNA precursors. Synthesis of Circular DNA 1 was examined by electrophoresis analysis as shown in Figure 2.9.

Figure 2.9 Electrophoretic analysis of synthesis of intrinsic curvature-containing Circular DNA 1 (676 bp in length) from Linear DNA 1. Lane 1: molecular weight markers; Lane 2: Linear DNA 1 alone; Lane 3: reaction mixture of Linear DNA 1 + T4

DNA ligase; Lane 4: reaction mixture of Linear DNA 1 + T4 DNA ligase followed by Nuclease BAL-31 hydrolysis; Lane 5: reaction mixture of Linear DNA 1 + T4 DNA ligase followed by SacI cleavage (see Figure 2.2 and Table 2.1 for detailed information about the nucleotide sequences of Circular DNA 1).

Figure 2.10 AFM image of intrinsic curvature-containing Circular DNA 1. The DNA product was purified from the band in Lane 4 in Figure 2.9.

Atomic Force Microscope (AFM) makes it possible to obtain images with the resolution of several nanometers and it has been successfully used for studying the morphologies of biological molecules for its less required manipulation of sample and height resolution.¹⁵⁴ Our newly synthesized 676 base pair circular DNA was subsequently examined using AFM in our studies. As shown in Figure 2.10, these relaxed forms of 676 base pair circular DNA molecules display implausibly interwound structures in their AFM images that contain one self-crossing (Figure-8 shape) around their duplex backbones. The emergence of Figure-8 shaped structures

from Circular DNA 1 in its AFM image signifies that the writhe of the relaxed circular DNA is either positive or negative one ($Wr = \pm 1$).

With the intention of further verifying that the self-crossings of the circular DNA shown in Figure 2.10 is indeed associated with the relaxed forms of DNA, Circular DNA 1 was incubated with topoisomerase I (topo I), an enzyme that transforms supercoiled DNA into its relaxed form (Figure 2.10). The reaction mixture was analyzed by gel electrophoresis, which showed no mobility shift occurred after Circular DNA 1 was incubated with topo I (Figure 2.11). The DNA sample which was purified from Lane 4 in Figure 2.11 was tested by AFM. As shown in Figure 2.12, the self-crossings in Circular DNA 1 still remain in their AFM images after the DNA circles were acted on by topo I.

Figure 2.11 Electrophoretic analysis of Circular DNA 1 with Topo I. Lane 1: molecular weight markers; Lane 2: Linear DNA 1 alone; Lane 3: Circular DNA 1 alone; Lane 4: Circular DNA 1 + topo I.

Figure 2.12 AFM image of intrinsic curvature-containing Circular DNA 1 after reacting with Topo I. The DNA product purified from the band in Lane 4 in Figure 2.11.

In order to further confirm that the observed self-crossings of DNA in Figure 2.10 and Figure 2.12 arose from sequence dependent effect and are indeed associated with the intrinsic curvatures of DNA, a different 676 base pair circular DNA (Circular DNA 2) was designed and synthesized (Figure 2.13A) next during our studies. Circular DNA 2 possesses the same nucleotide composition and the same length as Circular DNA 1 (Figure 2.3). Different from Circular DNA 1, however, Circular DNA 2 possesses no apparent spaced tracts of adenines or other recognizable curvature-forming segment in its duplex sequence. As shown in Figure 2.13B, there was no self-crossing noticeable in the AFM images of the new DNA circles. The observations shown in Figure 2.10 and Figure 2.12 reveal that formation of DNA writhe is indeed maneuverable by manipulating nucleotide sequences of DNA. In addition, it is shown in our studies that the intrinsic curvature-containing circular DNA (Circular DNA 1 in Lane 4 in Figure 2.9) and the non curvature-containing circular DNA (Circular DNA 2

in Lane 4 in Figure 2.13A) migrate faster and slower than their linear precursors respectively. This happens most likely because Circular DNA 1 holds a more compact structure than Circular DNA 2.

Figure 2.13 Synthesis and examination of non-supertwisted structures of Circular DNA 2 that are in their relaxed forms. (A) Electrophoretic analysis of non-curvature-containing Circular DNA 2 (as controls, 676 bp in length) from Linear DNA 2 (see Figure 2.3 and Table 2.1 for detailed information about the nucleotide sequences of Circular DNA 2). (B) Obtained AFM images of non-intrinsic curvature-containing circular DNA 2 (as controls) in our studies (the DNA product purified from the band in Lane 4 in Figure 2.13A).

2.4.2 Synthesis and Confirmation of Toroidal Structures of DNA that Possesses Writhe Number of ± 1

Figure 2.14 Synthesis and examination of toroidal structures of Circular DNA 3 that are in their relaxed forms. (A) Electrophoretic analysis of formation of Circular DNA 3. Lane 1: molecular weight markers; Lane 2: Linear DNA 3 alone; Lane 3: reaction mixture of Linear DNA 3 + T4 DNA ligase; Lane 4: reaction mixture of Linear DNA 3 + T4 DNA ligase followed by Nuclease BAL-31 hydrolysis; Lane 5: reaction mixture of Linear DNA 3 + T4 DNA ligase followed by SacI cleavage (see Figure 2.4 and Table 2.2 for detailed information about the nucleotide sequences of Circular DNA 3). (B) Obtained AFM images of intrinsic curvature-containing Circular DNA 3 in our studies (the DNA product purified from the band in Lane 4 in Figure 2.14B).

Toroidal forms of DNA are known to be alternative conformations which can be adopted by supercoiled DNA *in vivo* and *in vitro*.¹ With the aim of examining whether intrinsic curvature-containing DNA could adopt the topological conformations beyond the interwound forms shown in Figure 2.10 and Figure 2.12, a new 1154 base pair

circular DNA (Circular DNA 3) was synthesized (Figure 2.14A) next during our studies. Circular DNA 3 possesses spaced tracts of adenines that occur in the same strand of its duplex structure (Figure 2.4). As expected, this new circular DNA displays toroidal shapes (Figure 2.14B) rather than interwound structures in its AFM images. As a control experiment, a different circular DNA (Circular DNA 4) was synthesized in our studies (Figure 2.15A) that possesses the same nucleotide composition as Circular DNA 4 but contains no spaced adenine tract in its sequence (Figure 2.5). As shown in Figure 2.15B, there is no self-crossing observable in the AFM images of the control circular DNA (Circular DNA 4). The observations shown in Figure 2.14B could be considered as the evidences that besides the interwound forms, formation of toroidal structure of DNA could be achieved through varying nucleotide sequences of DNA as well.

Figure 2.15 Synthesis and examination of non-supertwisted structures of Circular DNA 4 that are in their relaxed forms. (A) Electrophoretic analysis of formation of Circular DNA 4 (as controls); Lane 1: molecular weight markers; Lane 2: Linear DNA 4 alone; Lane 3: reaction mixture of Linear DNA 4 + T4 DNA ligase; Lane 4: reaction mixture of Linear DNA 4 + T4 DNA ligase followed by Nuclease BAL-31 hydrolysis; Lane 5: reaction mixture of Linear DNA 4 + T4 DNA ligase followed by SacI cleavage (see Figure 2.5 and Table 2.2 for detailed information about the nucleotide sequences of Circular DNA 4). (B) Obtained AFM images of non-intrinsic curvature-containing Circular DNA 4 (as controls) in our studies (the DNA product purified from the band in Lane 4 in Figure 2.15A).

2.4.3 Synthesis and Confirmation of Double Interwound Structures of DNA that Possesses Writhe Number of ± 2

With the aim of demonstrating that more than one self-crossing point of duplex DNA could be formed non-enzymatically, an additional circular DNA (Circular DNA 5) was subsequently prepared in our lab (Figure 2.16A). This new DNA circle contains four segments of consecutive spaced adenine tracts that spread into two opposite strands of the DNA circle in alternate manners (Figure 2.6). As anticipated, the relaxed forms of Circular DNA 5 exhibit two self-crossings (double interwound structures) in their AFM images (Figure 2.16B) while such a self-crossing is not visible in a control DNA (Circular DNA 6, Figure 2.17A and Figure 2.17B) in which there is no apparent adenine tract present in its circular backbones (Figure 2.7). The

occurrence of double interwound structure shown in Figure 2.16B is the indication that DNA structures with writhe number beyond + 1 is achievable as well in the absence of gyrase and reverse gyrase.

Figure 2.16 Synthesis and examination of double interwound structures of Circular DNA 5 that are in their relaxed forms. (A) Electrophoretic analysis of formation of Circular DNA 5; Lane 1: molecular weight markers; Lane 2: Linear DNA 5 alone; Lane 3: reaction mixture of Linear DNA 5 + T4 DNA ligase; Lane 4: reaction mixture of Linear DNA 5 + T4 DNA ligase followed by Nuclease BAL-31 hydrolysis; Lane 5: reaction mixture of Linear DNA 5 + T4 DNA ligase followed by SacI cleavage (see Figure 2.6 and Table 2.3 for detailed information about the nucleotide sequences of Circular DNA 5). (B) Obtained AFM images of intrinsic curvature-containing Circular DNA 5 in our studies (the DNA product purified from the band in Lane 4 in Figure 2.16A).

Figure 2.17 Synthesis and examination of non-supertwisted structures of Circular DNA 6 that are in their relaxed forms. (A) Electrophoretic analysis of formation of Circular DNA 6 (as controls); Lane 1: molecular weight markers; Lane 2: Linear DNA 6 alone; Lane 3: reaction mixture of Linear DNA 6 + T4 DNA ligase; Lane 4: reaction mixture of Linear DNA 6 + T4 DNA ligase followed by Nuclease BAL-31 hydrolysis; Lane 5: reaction mixture of Linear DNA 6 + T4 DNA ligase followed by SacI cleavage (see Figure 2.7 and Table 2.3 for detailed information about the nucleotide sequences of Circular DNA 6). (B) Obtained AFM images of non-intrinsic curvature-containing Circular DNA 6 (as controls) in our studies (the DNA product purified from the band in Lane 4 in Figure 2.17A).

2.4.4 Observation of Backbone Self-crossings of Kinetoplast DNA as well as Plasmid DNA Containing Kinetoplast DNA sequences

Figure 2.18 Electrophoretic analysis of catenated kinetoplast DNA and decatenated kinetoplast DNA. Lane 1: molecular weight markers; Lane 2: catenated kinetoplast DNA (Topogen); Lane 3: reaction mixture of catenated kinetoplast DNA and Human Topoisomerase II; Lane 4: reaction mixture of decatenated kinetoplast DNA minicircles and Human Topoisomerase I. See Table 2.4 for detailed information about the nucleotide sequences of decatenated kinetoplast DNA minicircles.

A kinetoplast is a massive catenated network of DNA circles that are found inside a large mitochondrion in protozoa of the class Kinetoplastida.⁷⁶ It is known that kinetoplast DNA exists as a giant network of thousands of catenated DNA circles. There are two types of DNA circles: maxicircles and minicircles. The length of maxicircles usually ranges from 20 to 40 kb while minicircles appears from 0.5 to 10 kb in length depending on the species. Moreover, kinetoplast DNAs possesses a high degree of intrinsic curvatures in their duplex chains as we discussed above. In our studies, kinetoplast DNA minicircles (2561 base pairs in length) were obtained

through decatenation of kinetoplast DNA with Human Topoisomerase II (Topo II) which can break one duplex and pass another duplex through the break (Lane 3 in Figure 2.18). The newly obtained kinetoplast DNA minicircles were accordingly examined in our studies using atomic force microscopy. As shown in Figure 2.19A, the intrinsic curvature-containing kinetoplast DNA circles displayed well-recognizable backbone self-crossing in their AFM images. With the purpose of confirming that the self-crossings shown in Figure 2.19A are in effect affiliated with their relaxed forms, decatenated kinetoplast DNA minicircles were incubated with Human Topoisomerase I (Topo I), an enzyme that converts supercoiled DNA into its relaxed form. As seen in Figure 2.18, no mobility shift difference can be observed between Lane 3 and Lane 4 in the gel electrophoresis analysis. The later AFM image also revealed that the self-crossings in kinetoplast DNA still remain in their AFM images after these DNA circles had been allowed to act on by Topo I (Figure 2.19B). The observations of non-zero writhe number in Figure 2.19A and Figure 2.19B suggest that the occurrence of backbone self-crossings is the manifestation of intrinsic curvatures possessed by kinetoplast DNA.

Figure 2.19 AFM image of decatenated kinetoplast DNA minicircles in their relaxed forms. (A) AFM images of decatenated kinetoplast DNA minicircles (purified DNA sample from Band 1 in Lane 3 in Figure 1A); (B) AFM images of decatenated kinetoplast DNA minicircles that was pre-treated with Human Topoisomerase I (purified DNA sample from Band 1 in Lane 4 in Figure 1A).

With the aim of examining whether the self-crossing affiliated with kinetoplast DNA could still be upheld if the entire sequence of the DNA is inserted into a plasmid vector (pOK12), Circular DNA 7 was obtained in our lab. The newly constructed DNA contains 4695 base pairs in length, in which 2561 base pairs are from kinetoplast DNA and the rest (2134 base pairs) belongs to pOK12 vector that contains no identifiable intrinsic curvature in its duplex structures (See Table 2.5 for detailed information about the nucleotide sequences of Circular DNA 7). AFM examination on Circular DNA 7 was next carried out during our investigations under the same sample preparation condition as those described for kinetoplast DNA (Figure 2.19). As illustrated in Figure 2.20A, this newly constructed circular DNA displays duplex backbone self-crossings in its AFM images as well. In addition, as a control, vector pOK12 was examined as well in our studies (See Table 2.7 for detailed information about the nucleotide sequences of vector pOK12). It turned out that this vector alone exhibited no duplex backbone self-crossing in its AFM images (Figure 2.20B). The observations shown in both Figure 2.20A and Figure 2.20B are consistent with the suggestion that non-curved segment of vector pOK12 has little effect on the manifestation of backbone self-crossing associated with intrinsic curvatures in kinetoplast DNA.

Figure 2.20 AFM images of Circular DNA 7 and pOK12 vector in their relaxed forms. (A) Relaxed form of DNA 1 that contains both kinetoplast DNA sequence (2561 bp in length) and pOK12 vector sequence (2134 bp in length) in its structure; (B) Relaxed form of pOK12 vector DNA alone (2134 bp in length). Both DNA 1 and pOK12 vector were incubated with Human Topoisomerase I prior to our AFM examinations.

2.4.5 Observation of Backbone Self-crossings of Plasmid DNA Containing Repeats of Replication Origins of Bacteriophage λ Sequence

It has been well known that intrinsic DNA curvatures could exist at the replication origins of Bacteriophage λ , Yeast APS1 and Simian Virus.^{135, 137-138} With the aim of illustrating that those DNA segments associated with naturally occurred intrinsic DNA curvatures could introduce the backbone self-crossings as well, a circular DNA (Circular DNA 8) carrying the repeats of nucleotide sequence at the

replication origins of of Bacteriophage λ was constructed during our investigations. This new circular DNA is composed of 2641 base pairs in its sequence in which 507 base pairs are the repeats of nucleotide sequence from the replication origins of Bacteriophage λ and 2134 base pairs are the nucleotide sequence from vector pOK12 (See Table 2.6 for detailed information about the nucleotide sequences of Circular DNA 8). To obtain the DNA circles in their relaxed conformations, all DNA samples were treated with Human Topoisomerase I. Our subsequent AFM examination revealed the potential writhe associated with $\sim 1/5$ of the sequence of Circular DNA 8 (507 base pairs of short adenine tracts) cannot be cancelled out by the rest of $\sim 4/5$ of non-adenine tract-rich sequence (2134 base pairs) and there is indeed one backbone self-crossing in each of DNA molecules in their AFM images (Figure 2.21A). In addition, the sizes of two circles (507 base pairs and 2134 base pairs) on the opposite sides of the self-crossing points are evidently different, which corresponds to our design. As a control, AFM examination on vector pSP73 was carried out as well, which possesses a similar length to DNA 2 and has no identifiable intrinsic curvature in its structure (See Table 2.8 for detailed information about the nucleotide sequences of vector pSP73). As seen in Figure 2.21B, there is no backbone self-crossing visible in the AFM images of pSP73. The observations shown in Figure 2.21A and Figure 2.21B illustrate that similar to kinetoplast DNA, the intrinsic curvatures affiliated with replication origins of Bacteriophage λ could manifest themselves in the form of backbone self-crossings in space.

Figure 2.21 AFM images of Circular DNA 8 and pSP73 vector. (A) Relaxed form of Circular DNA 8 that contains the repeats of replication origins of Bacteriophage λ sequence (507 bp in length) and pOK12 vector sequence (2134 bp in length); (B) Relaxed form of pSP73 vector DNA alone. Both DNA 2 and pSP73 vector were incubated with Human Topoisomerase I prior to our AFM examinations.

2.4.6 General DNA Topological Conservation Law of DNA

The mathematical equation $Lk - Tw = Wr$ was formulated by Fuller^{48, 52} in 1970s based in part on the conception of canonical B-form of DNA, the only type of DNA conformations that was known to the scientific community during that early time period. This equation was in turn named “DNA Topological Conservation Law” by Miller and Benham in 1996 in their discussions of writhe and linking densities for closed circular DNA.⁵⁴ Not long after this conservation law was formulated, varieties of non-canonical B-forms of DNA have been identified, which include Z-DNA, A-

DNA, intrinsically curved DNA, anisomorphic DNA, cruciform, triple helix DNA and G-quadruplex. Among them, the intrinsic curvatures of DNA are known nowadays to occur ubiquitously in organismal genomic DNA.¹⁶⁰⁻¹⁶¹ In our studies, a series of curvature-containing circular DNAs were artificially designed,⁶¹ which exhibited intrinsic curvature-affiliated self-crossings in their backbones as shown above. In addition, we demonstrate that besides those artificially designed DNA, the intrinsic curvatures in organismal DNAs could lead to the generation of backbone self-crossing as well. Since DNA writhe could be affiliated with (1) underwinding/overwinding and (2) intrinsic curvatures as illustrated in our studies, we introduce the following formula in order to differentiate between the writhe number associated with the two types of causes:

$$Lk - Tw + Nb = Wb + Wn = Wr \quad (\text{Equation 2.1})$$

Equation 1, which is tentatively named as “General DNA Topological Conservation Law of DNA”, is in effect an amended form of $Lk - Tw = Wr$ (original “DNA Topological Conservation Law”),⁵³⁻⁵⁴ in which Nb represents (1) the curvature parameters affiliated with non-canonical B structures (*e.g.* length of a curved DNA segment and its degree of curvature; rigidity and length of poly(dA) chain, and the presence of cruciform or G-quadruplex) and (2) the environmental parameters such as the change of temperature, nature of salts and salt concentration that could cause structural deviation from the ideal canonical-B conformation. Wb in this equation is the writhe number of DNA that is associated with the difference between linking

number and twist number in the canonical-B DNA while W_n is the writhe number that is correlated with non-canonical B structures and other environmental factors.

Figure 2.22 Schematic illustration of relationship among the parameters in canonical B-form DNA.

Schematic illustrations of correlations among Lk , Tw , Nb , W_n , W_b , and W_r described in Equation 2.1 as well as definitions of each term in Equation 2.1 are shown in Figure 2.22 and Figure 2.23. The definition of each parameter in Equation 2.1 corresponds to those in original “DNA Topological Conservation Law”: Lk is linking number which is the number of intertwines between two complementary DNA strands; Tw is twist number which is the total number of turns of double stranded DNA around its helical axis; Nb represents the curvature parameters affiliated with

non-canonical B structures and environmental factors; W_r is apparent writhe number which is observed number of times the double helix crosses over on itself; W_b is writhe number associated with difference between linking number and twist number; and W_n is writhe number associated with accumulation of non-B DNA structures. If there is no accumulable non-B segments present in the nucleotide sequence of a DNA (Structure 1-3 in Figure 2.22), N_b and W_n will both be equal to zero ($N_b = 0$ and $W_n = 0$), which will then allow $Lk - Tw + N_b = W_b + W_n = W_r$ (Equation 2.1) to resume to $Lk - Tw = W_r$ (original “DNA Topological Conservation Law”, see Figure 2.22).

Figure 2.23 Schematic illustration of correlations among Lk , Tw , N_b , W_n , W_b and W_r in DNA that contain accumulable non-canonical B structures.

If some non-B segments occur in a circular DNA (Structure 4-5 in Figure 2.23), accumulation of the non-B segments could lead to number of writhe (W_n) that cannot be resolved by the action of topoisomerases I (from Structure 5 to Structure 6 in Figure 2.23). Consequently, the writhes observed in the DNA molecules shown in our studies are corresponding to W_n in Equation 2.1.

2.4.7 Significance of Our Studies

Demonstration of the manipulable nature of DNA writhe as well as General DNA Topological Conservation Law of DNA in our studies could possibly have certain impacts on our perception and understanding of the topological features and roles of DNA *in vivo* and *in vitro*. Firstly, for example, as demonstrated in the current studies, exhibition of self-crossing around the axis of double helix of DNA does not always represent the manifestation of difference between linking number and twist number in DNA structures any longer (Figure 1.8). In addition, poly(dA) tracts are known to commonly exist in some eukaryotes and to resist the formation of curved structures of DNA.¹⁶⁰ It can be therefore speculated that when such bending-resisting segments are present in DNA sequences, supercoiled structures in some cases might not even be generated even if linking number and twist number are different. Consequently, the concepts of “supercoiled form” and “relaxed form” of DNA as well as the correlation between writhe number and supercoiling may need to be precisely re-defined by experts in the fields in the near future.

Secondly, there is an inconsistency between theoretical deduction of negative two superhelical turns of DNA wrapped around octamer histones and experimental detection of negative one supercoil after the histone proteins are removed. Contradiction between the theoretical prediction and experimental observation was recognized more than thirty years ago and has been known as “linking number paradox”.¹⁶² It is known nowadays that (i) the intrinsically bent structures commonly occurs in both prokaryotic and eukaryotic cells¹³³⁻¹³⁴ and (ii) besides adenine-tract-containing segments, various non-adenine-tract-containing DNA sequences are capable of forming intrinsically curved structures.¹⁶³⁻¹⁶⁵ Therefore, it is our speculation that the abovementioned negative two superhelical turns of DNA wrapped around each unit of octamer histones could be composed of both forcibly formed curvatures and the widespread intrinsically curved segments. Determination of the number of superhelical turns that are associated with DNA, on the other hand, is commonly conducted through relaxation of supercoiled structures of DNA by topo I.¹⁶² If both forcibly formed curvatures and intrinsically curved segments indeed co-exist in the formation of nucleosomes, the sections of superhelical turns associated with the intrinsic curvatures of DNA are not relaxable by topo I (see illustrations in Figure 2.22 and Figure 2.23), which might lead to certain writhe number unnoticed. Further studies on the co-existence of the forcible and intrinsic curvatures in nucleosome structures as well as development of new writhe-number detecting methods might possibly provide a partial solution to the more than thirty-year- old paradox.

Thirdly, topoisomerases (e.g. gyrase, reverse gyrase, topo I and topo II) are a group of enzymes that catalyze the interconversion between different topological

forms of DNA.⁴⁶ Among them, topoisomerase II is the most abundant enzyme associated with chromosome assembly *in vivo*.^{59, 129} This enzyme has been known to recognize and bind to the self-crossing section occurred in negative DNA supercoils. It is not known, however, whether topoisomerase II could recognize the non-supercoil-associated self-crossing structure of DNA reported in the current studies. It is thus anticipated that further examination on the interaction between topoisomerase II and intrinsic curvature-based self-crossing could possibly reveal new information about the mechanistic action of this nuclear enzyme.

Fourthly, it has been established in the past that during its actions, gyrase wraps DNA around its own protein complex and alters the writhe number of its substrate DNA by two in the end.⁴⁷ Based on the information about the relationship between intrinsic curvatures and self-crossings of duplex DNA backbones reported in the current studies, this enzyme may not necessarily alter the writhe number of DNA by two exactly in each step of its action, depending on whether or not intrinsically curved segments or bending-resisting poly(dA) are present adjacently. Further investigation on correlations between the self-crossings of duplex backbones in relaxed form of DNA and catalytic activity of topoisomerases might possibly reveal new knowledge about the mechanisms of these topological enzymes.

Moreover, in the process of genetic recombination *in vivo*, two remote DNA segments in the same molecule need to be brought into proximity so that a DNA exchange reaction could proceed. It is demonstrated in our current studies that the proximity of certain two remote DNA segments could be achieved by forming

intrinsic DNA-curvature-based toroidal and interwound structures (Figure 2.14 and Figure 2.16). Since intrinsic DNA curvatures are known nowadays to exist ubiquitously in eukaryotic and prokaryotic genomes,¹³³⁻¹³⁴ further studies on the correlation between genetic recombination and curvature-based toroidal/interwound structures could be beneficial to our understanding of the mechanisms of genetic recombination.

2.5 Conclusion

In conclusion, it was well established that local DNA geometry could substantially contribute to the global conformation of DNA. It is demonstrated in our current studies that manipulation of toroidal and interwound structures of DNA can be readily carried out solely through maneuvering these local DNA geometries. Our experimental results further illustrate that the magnitude of DNA writhe could be accurately engineered as well by carefully altering T-rich and A-rich segments on a target DNA.⁶¹

On the other hand, it is also demonstrated in our current studies that besides the DNAs that possess artificially edited nucleotide sequences, some organismal DNAs could produce non-zero writhe number through their accumable intrinsic backbone curvatures.⁷⁷

In addition, the original DNA Topological Conservation Law was formulated by mathematicians more than thirty years ago on the basis of a closed ribbon model and

conception of ideal canonical B-form of DNA. Our current studies demonstrate that this original conservation law is violated by the actions of some organismal as well as designed structures of DNA. Consequently, “General DNA Topological Conservation Law of DNA” was tentatively introduced by us in order to differentiate between the writhe number caused by canonical-B forms and non-canonical-B conformations.⁷⁷

It is our hope that further examination of correlation between accumulable intrinsic curvatures and the writhe number associated with them could provide important information for our understanding of the topological principles of DNA utilized by prokaryotic and eukaryotic cells in the future.

Chapter 3

Precise Engineering and Visualization of Signs and Magnitudes of DNA Writhe on the Basis of PNA Invasion

3.1 Introduction

Supercoiling of DNA (Figures 3.1) refers commonly to a physical arrangement of topologically closed double helical structure that exists in space in a self-twisted fashion.² Generation of such supercoiled entity of DNA in eukaryotic cells is associated with replication, transcription, and packing of DNA into chromatin while in prokaryotic cell, supercoiled DNA can form with the assistance of gyrase and reverse gyrase.¹⁶⁶⁻¹⁶⁸ The degree of supercoiling of DNA is often described by “superhelical density” (σ), which is defined as the number of turns that have been added or subtracted in the supercoiled DNA, compared to its relaxed state, divided by the total number of turns in the DNA if it were relaxed (see Equation 1.3 in Chapter 1).⁶⁵ Typically, the magnitude of superhelical density in both prokaryotic and eukaryotic cells is -0.06.¹⁶⁹⁻¹⁷⁰ On the other hand, spatial arrangement of superhelicity of DNA belongs conceptually to the category of tertiary structure of molecules in chemistry, which determines physical, chemical and biological properties of the corresponding macromolecule to a high extent. In addition, the supercoiling structures of DNA that

possess the same sequences but different superhelical densities are called topological isomers or topoisomers.³³ Similar to other isomerisms identified in chemistry, topoisomers of different superhelical densities of DNA should in theory display different physical and biological properties in some way. Consequently, construction of supercoiled structures of oligonucleotides (DNA) that possess desirable sequences and controllable superhelical densities has attracted more attention in both chemist and biologist. Unfortunately, certain fundamental issues regarding certain physical and biological properties of topoisomers of DNA with different superhelical densities have still partly remained unknown.

Figure 3.1 Illustration of supercoiled structure of DNA present in cells.

Peptide nucleic acids (PNAs) are analogs of DNA in which the original sugar-phosphate backbones are replaced with electrically neutral pseudo peptide linkage (Figure 3.2).¹⁷¹⁻¹⁷³ From the time when they were invented in the early 1990s, these new types of DNA mimics have been shown to be capable of invading and opening up

duplex structures of DNA effectively through forming new triplex or duplex assemblies with one of the target duplex strands.¹⁷⁴⁻¹⁷⁶ Homopyrimidine PNA, on the other hand, can bind to complementary DNA normally by formation of unusually stable (PNA)₂ DNA triplex conformations which was known as the P-loops¹⁷⁷⁻¹⁷⁸ (Figure 3.3). Due to extraordinary high sequence-specificity and stability, the P-loop structures are quite different from other triplex structures such as (DNA)₃ or (DNA)₂ RNA and has attracted much more attentions in scientist.

Figure 3.2 Chemical structures of DNA and PNA.

Figure 3.3 Pictorial illustration of P-loop structures.

To enhance strand invasion efficiency, two homopyrimidine oligomers are covalently linked “head-to-tail” by a flexible 8-amino-3,6-dioxaoctanoic acid linker, which was name bis-PNAs.¹⁷⁹⁻¹⁸¹ The structure of bis-PNAs facilitate locally opens the DNA double helix via forming Hoogsteen base pairing by one PNA strands and constructed Watson-Crick base pairing through another strand as shown in Figure 3.4. In addition, bis-PNAs also reduce a trimolecular reaction of PNA to DNA binding to a bimolecular reaction.¹⁸²⁻¹⁸⁴ Highly pH-dependent is known to be existed in the PNA-DNA triplex C-containing homopyrimidine.¹⁷⁷ This happens because the formation of Hoogsteen base pairing needs to protonation of cytosine bases. On the other hand, carrying a hydrogen atom at the N3 position, pseudoisocytosine (J base) is allowed to form Hoogsteen pairing with a guanine base without protonation¹⁷² as illustrated in Figure 3.5. If the pseudoisocytosine was employed instead of cytosine in the N-terminal half of bis-PNAs, the invasion of PNA can be carried out at a wide range of pH value.

Figure 3.4 Schematic illustrations of two possible routes for formation P-loop from bis-PNA.¹⁷⁷

Figure 3.5 Hoogsteen binding with protonated cytosine (I) and with pseudoisocytosine (II). J indicated pseudoisocytosine.¹⁷²

Owing to its unique mode of action, PNAs have been widely utilized to modulate gene expression and to perform diagnostic functions.¹⁸⁵⁻¹⁸⁹ Based on our recent analysis on the available information about the properties of PNA, we speculated that PNA invasion could be taken as an action that alters the linking number of a double helix of DNA since such invasion could interrupt the integrity of helicity of the target duplex structure.¹⁹⁰⁻¹⁹¹ Consequently, our attempts to precisely engineer DNA supercoils at the macromolecular level have been made recently on the basis of PNA invasion principle.

Figure 3.6 Schematic representation of reduction of linking number in linear DNA duplex by PNA.

Figure 3.6-3.8 depicts the general strategies in manipulating the signs ($\pm$) and magnitude of writhe (e.g. 0, 1, 2) of DNA supercoils utilized in our lab. If a linear duplex DNA possesses 156 base pairs in length, for instance, 15 helical turns (156 base pairs/10.4 base pairs = 15 helical turns) would in theory appear around its duplex backbones. When a designed PNA of 10 bases in length invades the target duplex DNA of 156 base pairs that contains a complementary segment to the designed PNA, the linking number of the target duplex will be reduced to 14 from 15 (Figure 3.6). This reduction occurs because the helical turns in the complementary segment of target DNA is interrupted and incapable of maintaining its regular double helicity any longer.

Similarly, if a linear duplex DNA contains both two cohesive ends and a complementary segment to a designed PNA, the linking number of the duplex DNA

will be reduced to 14 from 15 upon a PNA invasion (Figure 3.7). After two cohesive ends are subsequently joined covalently by the action of DNA ligase followed by the removal of the 10 base PNA from the duplex DNA circle, the remaining linking number of 14 will redistribute into the entire 156 base pairs. A negative supercoil with a writhe number of -1 would consequently be generated by the 156 base pair circular DNA because this DNA circle would require 15 helical turns in its structure in order to maintain its low energy conformation.²

Figure 3.7 Schematic representation of engineering of negatively supercoiled DNA by PNA invasion approach.

Besides the possible manipulation of negative DNA supercoils illustrated in Figure 3.6 and Figure 3.7, it is our further speculation that positive supercoils could be engineered as well on the basis of PNA invasion (Figure 3.8). A circular DNA of 156

base pairs that contains no nick site in its sequence, for example, will in theory possess 15 helical turns when it exists in its relaxed form. Without the invasion of a PNA, the linking number of 15 distribute over the entire 156 base pairs in its circular backbones. When a designed PNA of 10 bases invades this circular DNA, the double helicity in the segment that the PNA invades into is interrupted. Consequently, the linking number of 15 has to gather together in the rest 146 base pairs. Since a sequence of DNA with 146 base pairs would need 14 helical turns to maintain its low energy state, the linking number of 15 that are forcibly to reside in the sequence of 146 base pairs will lead to the generation of positive supercoiling of DNA. Supercoiled DNA is known to possess higher energy than its relaxed counterpart. In addition, it has been well established that the free energy of a DNA supercoil is proportional to the square of its linking number difference and inversely proportional to the size of the DNA circle.⁷ Consequently, the free energy associated with the negative and positive DNA supercoils designed in the current studies should be higher than those of their relaxed circular precursors.¹ It is our anticipation that circular DNA supercoils with the writhe number of ± 2 , ± 3 , ± 4 and beyond ± 4 could be achievable as well on the basis of the PNA principle shown in Figure 3.7 and Figure 3.8 as long as two or more PNA molecules are applied.

Figure 3.8 Schematic representation of engineering of positively supercoiled DNA on the base of PNA invasion.

3.2 Design of DNA Sequences

3.2.1 Design of Linear DNA Precursors with One PNA Binding Site

In our studies, linear DNA precursors (558 bp) with one PNA binding site were produced by means of polymerase chain reaction. A plasmid DNA (X2420G) composed of pGH vector and duplex segment of Linear DNA 9 was accordingly designed for the purpose of engineering the supercoils of DNA with writhe numbers of 0, -1 and +1. After two cohesive ends were obtained by *SacI* digestion, T4 DNA ligase was used to join the paired ends and a covalently closed Circular DNA 9 (530 bp) was synthesized, in which one PNA binding site can be find as shown in Figure 3.9. In addition, the nucleotide sequence of Linear DNA 9 can be found in Table 3.1.

Figure 3.9 Schematic illustrations of the routes for synthesis of Circular DNA 9 with writhe number of 0.

Table 3.1 Nucleotide sequences of Linear DNA 9. Only one of the two strands of DNA from 5' end to 3' end is shown in the table. The segment highlighted in red is designed PNA binding site.

Name of DNA	Nucleotide sequence
Linear DNA 9	CCGAGCTCCCGTAATACGACTCACTTAAGGCCTTGACTAGAG GGTACCAACCTAGGTATCTAGAACCGGTCTCGAGCCATAACT TCGTATAGCATACATTATACGAAGTTATATAAGCTGTCAAAC ATGAGAATTCTTGTTATAGGTTAATGTCATGATAATAATGGT TTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGG AACCCCTATTTGTTTATTTTCTAAATACATTCAAATATGTA TCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATA TTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTGCG CCTTATTCCCTTTTTTGCGGCATTTTGCTTCCTGTTTTTGC TCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCA GTTGGGTGCACGAGTGGGTACATCGAACTGGATCTCAACAG CGGTAAGTTAAGCTTTTTGCACAACATGGGGGATCATGTAAC TCGCCTTGATCGAAGGAGAGAGAGCTGGAGCTCAATGAAGC CATACCAAACGA

If PNA invasion occurred after SacI digestion, a linear PNA-DNA complex with (1) two cohesive ends and (2) linking number of 50 can be produced. Upon the action of T4 DNA ligase on this linear PNA-DNA complex, a circular PNA-DNA complex will be obtained. After the final removal of PNA is carried out by heating the reaction mixture followed by cooling it to room temperature in the presence of high salt concentration,¹⁹²⁻¹⁹³ designed circular DNA (Circular DNA N9) with writhe number of -1 can be produced as shown in Figure 3.10.

Figure 3.10 Schematic illustrations of engineering of Circular DNA N9 with writhe number of -1. The structure of PNA are given: PNAs are written from the N terminus to the C terminus using normal peptide conventions: H is a free amino group; NH₂ is a terminal carboxamide; Lys is the lysine residue; J denotes pseudoisocytosine and eg1 denotes the linker unit, 8-amino-3,6-dioxaoctanoic acid.

If PNA is allowed to invade Circular DNA 9, an action will force the linking number of 51 from 530 base pairs in the overall DNA circle to gather into the section of 520 base pairs in its duplex backbones. This accumulation of the linking number of

51 in the 520 base pair segment drives the PNA-containing circular DNA to form a positive DNA supercoil (Circular DNA P9) as shown in Figure 3.11.

Figure 3.11 Schematic illustrations of engineering of Circular DNA P9 with writhe number of +1.

3.2.2 Design of Linear DNA Precursors with Two PNA Binding Sites

With the aim of engineering the DNA supercoils with more than one of writhe number, a linear DNA (Linear DNA 10) with two PNA binding sites was designed in our studies. After polymerase chain reaction upon the plasmid DNA (W2054E) carried out with specific primers, a 1068 base pair linear duplex DNA (Linear DNA 10) can be obtained, in which two identical PNA binding sites exists. Similar to the synthesis of Linear DNA 9, two cohesive ends can be produced by the actions of digestion by restriction endonuclease SacI. After that, the same strategy could be used to construct the DNA molecules with writhe number of 0, -1 and +1 respectively if two of PNA molecules invade one DNA duplex backbone as illustrated in Figure 3.12. In addition, the nucleotide sequence of Linear DNA 10 is given in Table 3.2.

Figure 3.12 Schematic representation of molecular engineering of DNA supercoils with writhe number of ± 2 . (A) Construction of Circular DNA 10 with writhe number of 0. (B) Engineering of Circular DNA N10 with writhe number of -2. (C) Engineering of Circular DNA P10 with writhe number of +2.

Table 3.2 Nucleotide sequences of Linear DNA 10. Only one of the two strands of DNA from 5' end to 3' end is shown in the table. The segments highlighted in red are designed PNA binding sites.

Name of DNA	Nucleotide sequence
Linear DNA 10	GTGGATCCTCGTCGCAAAACGAGCTCCGATTAAGTTGGGTA ACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGC CAGTCCGTAATACGGCTCACTTAAGGCCTTGACTAGAGGGT ACCAACCTAGGTATCTAGAACCGGTCTCGAGCCATAACTTC GTATAGCATACATTATACGAAGTTATATAAGCTGTCAAACA TGAAACCTCTTGTTATAGGTTAATGTCATGATAATAATGGT TTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCG GAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATG TATCCGCTCATAATAACCCTGATAAATGCTTCAATAATATT GAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCC CTTATTCCCTTTTTTGCTCTCCCTATAGTGAGTCGTATTAA TACCCTCAGCTTCACCCATGAGAAGATTGACATCACATAAA CTATTCATACAGGATAATTGGGAGGCTTTATTGAAAGCCCA CTCACTGATTAACGGGCCTTCCTGTTTTTGCTCACCCAGAA ACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGC ACGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAGT TAAGCTTTTTGCACAACATGGGGGATCATGTAACTCGCCTT GATCGAAGGAGAGAATCCAAGAGAGGAATAGCTCTCCTTTT GAGGTGTTGCTCAATGAAGCCATACCAAACGACGAGCGTGA CACCACGATGCCTGCAGTGATTCCTCGAGCCATAACTTCGT ATAGCATACATTATACGAAGTTATCCATGGACTAGTGTATT ACGTAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGT GAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCC GGAAGCATAAAGTGTAAGCCTGGGGGAATTCGGGGTTAAC CATGGATCCGGGGGATATCACGTGAAGCTTGCAAGCTCCAG CTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTGAGCTC GG

3.3 Materials and Methods

3.3.1 Duplex DNA, Enzymes and Chemicals

Product(s)

Manufacturer

Plasmid DNA (X2420G)

Generay Biotech (Shanghai, China)

Peptide nucleic acids	Biosynthesis (Lewisville, Texas)
Oligodeoxyribonucleotides	Sigma-Proligo (Singapore)
100 bp DNA ladder	Fermentas (Singapore)
<i>Taq</i> Polymerase	New England Biolabs (Ipswich, MA)
1 Kb DNA ladder	New England Biolabs (Ipswich, MA)
SacI endonuclease	New England Biolabs (Ipswich, MA)
T4 DNA ligase	New England Biolabs (Ipswich, MA)
BAL 31 exonuclease	New England Biolabs (Ipswich, MA)
Biological purity water	1st Base Pte. Ltd (Singapore)
Agarose	Invitrogen (Carlsbad, CA)
Ethidium bromide	Research Biolabs (Singapore)
QIAquick PCR purification kit	Qiagen (Singapore)
QIAquick Gel Extraction Kit	Qiagen (Singapore)
Mini Prep Cell	Bio-Rad (Hercules, CA)
TAE, TBE, TRIS	1st Base Pte. Ltd (Singapore)

3.3.2 Polymerase chain reactions for synthesis of Linear DNA 9 and Linear DNA 10

Polymerase chain reaction was carried out following standard procedures with Taq DNA Polymerase as described as follows: A reaction mixture containing 1 ng plasmid DNA (X2420G), 0.25 μ M forward primer, 0.25 μ M reverse primer, 200 μ M dNTP, 1 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 processed as below protocol:

95 $^\circ\text{C}$ for 180 sec (denature);

56 $^\circ\text{C}$ for 40 sec (anneal);

72 $^\circ\text{C}$ for 40 sec (elongate) (60 sec per kb target sequence length);

29 cycles only (otherwise enzyme decay causes artifacts);

72 $^\circ\text{C}$ for 10 min (allow complete elongation of all DNA products).

The ssODN-1 and ssODN-2 are forward and reverse primers for Linear DNA 9 while the ssODN-3 and ssODN-4 are forward and reverse primers for Linear DNA 10. 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. In addition, the nucleotide sequence of Linear DNA 9 and Linear DNA 10 were given in Table 3.3.

Table 3.3 Nucleotide sequences of primers used in polymerase chain reactions.

Name of DNA	Nucleotide sequence
ssODN-1	5'-CCGAGCTCCCGTAATACGACTCACTTA-3'
ssODN-2	5'-TCGTTTGGTATGGCTTCATT-3'
ssODN-3	5'-GTGGATCCTCGTCGCAAAAC-3'
ssODN-4	5'CCGAGCTCAGCGCGCAATTAACCCTCAC-3'

3.3.3 Reactions of SacI with Duplex Linear DNA Precursors

The duplex linear DNA precursors containing two SacI digest site in each end were obtained from polymerase chain reactions. In order to create two cohesive ends as shown in the Figure 3.19 - Figure 3.12, those linear DNA were treated with SacI endonuclease. Linear DNA 9 and Linear DNA 10 were obtained as described as follows: A solution containing 10 mM Bis-Tris-Propane-HCl, 10 mM MgCl₂, 1 mM Dithiothreitol, duplex linear DNA precursors (200 ng) and 10 U SacI was incubated at 37 °C for 1 hr. The reaction products were further analyzed using agarose electrophoresis (1.5%) and purified using QIAquick PCR purification kit before the next steps.

3.3.4 Preparations of Circular DNA Using T4 Ligase

As there are two identical cohesive ends digested by Sac I in Linear DNA 9 and Linear DNA 10, Circular DNA 9 and Circular DNA 10 were obtain from ligase reactions as described as follows: A 50 µl solution containing 50 mM Tris-HCl, 10

mM MgCl₂, 1 mM ATP, 10 mM dithiothreitol, 500 ng linear DNA and 20 U T4 DNA ligase was incubated at 16 °C for 8 hrs. The obtained circular DNA products were further analyzed using agarose electrophoresis (1.5%) and purified using QIAquick PCR purification kit before the next steps.

3.3.5 Degrade Linear DNA from Ligase Reaction Mixture Using Nuclease BAL-31 Exonuclease

Nuclease BAL-31 exonuclease degrades both 3' and 5' termini of duplex DNA without generating internal scissions. To remove the linear DNA from the mixture of ligase reaction as well as further confirm that there is no nicks, gaps and single-stranded regions in the obtained DNA products, we treated the ligase reaction mixture by nuclease BAL-31 exonuclease as described as follows: A 50 µl solution containing 20 mM Tris-HCl, 600 mM NaCl, 12 mM CaCl₂, 12 mM MgCl₂, 1 mM EDTA, 500 ng reaction products of ligase reaction and 2 U exonuclease BAL-31 was incubated at 30 °C for 10 hrs. The obtained circular DNA products were further analyzed using agarose electrophoresis (1.5%) and purified using purified next using Mini Prep Cell.

3.3.6 PNA Invasion

In the binding reactions of PNA to target sites, the concentration of PNA was kept at a large excess over the DNA concentration, and the binding was performed at in pre-siliconized tubes at 37 °C for 6 hr in 10mM Sodium-Phosphate Buffer (PH = 6.9). In addition, the gel-mobility-shift experiments were performed in 1.5% agarose gels in 1 x TAE buffer. The electrophoresis was run at 100 V for four hours at room temperature and the gels were then stained with ethidium bromide.

3.3.7 AFM Studies of Obtained Circular DNA

To immobilize DNA for AFM imaging, a mica surface modified with APS was used as a substrate. DNA adheres to the imaging surface through electrostatic attraction as discussed in section of 2.3.8. Prepare the solution of the DNA sample in appropriate buffer. DNA concentration should be between 0.1 and 0.01 $\mu\text{g/ml}$ depending on the size of the molecules. Place 5–10 μl of the solution in the middle of APS-mica substrate (usually $1 \times 1 \text{ cm}^2$) for 2–3 min. Rinse the surface thoroughly with water (2–3 ml per sample) to remove all buffer components. AFM images 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.

3.4 Results and Discussion

3.4.1 Engineering of DNA Supercoils with Writhe Number of -1 and +1

A plasmid DNA (X2420G) containing duplex segment of linear DNA 9 was accordingly designed in our studies for the purpose of engineering the supercoils of DNA with writhe numbers of 0 and ± 1 (Figure 3.9 - Figure 3.11). After polymerase

chain reaction was carried out with specific primers (ssODN-1 and ssODN-2), a 558 base pair linear duplex DNA (Linear DNA 9) was produced. Linear DNA 9 was then digested by SacI (a restriction endonuclease) to create two cohesive ends in the termini of its duplex sequence. The two resultant cohesive ends was subsequently joined together covalently by DNA ligase to form a 530 base pair circular DNA (Circular DNA 9) which in theory possesses 51 helical turns in its relaxed structure. The new produced Circular DNA 9 should exist in its relaxed form because T4 DNA ligase does not have any activity of topoisomerase. Synthesis of Circular DNA 9 was examined by electrophoresis analysis and AFM as shown in Figure 3.13.

Figure 3.13 Synthesis and confirmation of Circular DNA 9 (530 bp in length) from Linear DNA 9 (558 bp in length). (A) Agarose gel electrophoretic analysis of DNA products; Lane 1: Molecular weight markers; Lane 2: Linear DNA 9 obtained by PCR amplification; Lane 3: Digested product of Linear DNA 9 by SacI; Lane 4: Circular

DNA (Circular DNA 9) promoted by T4 DNA Ligase; Lane 5: reaction mixture of ligase reaction followed by Nuclease BAL-31 hydrolysis. (B) AFM image of covalently closed Circular DNA 9 with writhe numbers of 0 (scale bar 100 nm).

A designed 10-mer PNA (Structure of PNA is shown in Figure 3.10) was next allowed to invade the linear DNA (digest products of Linear DNA 9 by SacI) with two cohesive ends followed by the action of T4 DNA ligase to catalyze the formation of a phosphodiester bond between juxtaposed 5' phosphate and 3' hydroxyl termini in duplex DNA. The final removal of PNA 1 was carried out by heating the reaction mixture at 90 °C for 5 minutes followed by cooling it to room temperature in the presence 100 mM NaCl, which ended up with the designed DNA supercoils with writhe number of -1. The reaction products were examined by electrophoresis analysis as shown in Figure 3.14.

Figure 3.14 Agarose gel electrophoretic analysis of the synthesis of Circular DNA N9. Lane 1: Molecular weight markers; Lane 2: Linear DNA 9 obtained by PCR amplification; Lane 3: Digested product of Linear DNA 9 by SacI; Lane 4: PNA-containing linear DNA produced upon the invasion of PNA; Lane 5: PNA-containing circular DNA obtained by the action of T4 DNA ligase; Lane 6: Supercoiled DNA

(Circular DNA N9) obtained after removal of PNA from PNA-containing circular DNA with relaxed form.

Figure 3.15 AFM image of Circular DNA 9 with writhe number of -1. The scale bar indicated 100 nm.

AFM has been known to be a very useful tool for studying certain subtle structural changes of DNA and was subsequently used in our investigations with the aim of examining the topological features of our newly designed supercoiled DNA. Our next examination on the circular DNA obtained after PNA invasion (Circular DNA N9), on the other hand, unveiled an interwound appearance of the DNA molecule (Figure 3.15), which indicates that Circular DNA N9 possessed indeed a supercoiled structure. Our further analysis on the AFM images of Circular DNA N9 revealed that the writhe signs that are associated with each supercoils formed by Circular DNA N1 are negative and the overall structures of these DNA supercoils are right-handed (writhe = -1) as initially designed (Figure 3.16). The observations shown

in Figure 3.16 suggest that both signs and magnitudes of DNA supercoils are indeed engineerable at the molecular level through utilizing PNA invasion principles.

Figure 3.16 Detail analysis of AFM images obtained from Circular DNA N9. (A) Amplitude and (B) 3D image showing the ring crossing path more clearly. (C) Theoretical definition of negative DNA supercoils. (D) Section image analysis of the self-crossing.

In addition to the negative supercoil of DNA shown in Figure 3.15 and Figure 3.16, a positively supercoiled assembly of DNA was engineered as well during our investigations (Figure 3.11). Circular DNA 9 was accordingly reconstructed in our studies, which contains 530 base pairs and 51 helical turns in its backbones. PNA is

next allowed to invade this circular DNA, an action that will force the linking number of 51 from 530 base pairs in the overall DNA circle to gather into the section of 520 base pairs in its duplex backbones. This accumulation of the linking number of 51 in the 520 base pair segment drives the PNA-containing circular DNA to form a positive DNA supercoil (Circular DNA P9), which was examined by electrophoresis analysis and AFM as shown in Figure 3.17. Our subsequent detail AFM examination reveals that the PNA-containing Circular DNA P9 is a left-handed single interwound structure (writhe = +1) as designed originally (Figure 3.18). Such an interwound structure is, however, not observable in the relaxed precursor DNA (Circular DNA 9), an observation that is consistent with the suggestion that the positive supercoil in Circular DNA P9 is closely associated with PNA invasion.

Figure 3.17 Synthesis and confirmation of Circular DNA P9. (A) Agarose gel electrophoretic analysis of DNA products. Lane 1: Molecular weight markers; Lane 2:

Linear DNA 9 obtained by PCR amplification; Lane 3: Digested product of Linear DNA 9 by SacI; Lane 4: Circular DNA (Circular DNA 9) promoted by T4 DNA Ligase; Lane 5: Reaction mixture of ligase reaction followed by Nuclease BAL-31 hydrolysis; Lane 6: PNA-containing Circular DNA P9 produced upon the invasion of PNA. (B) Obtained AFM image of Circular DNA P9 (scale bar 100 nm).

Figure 3.18 Detail analysis of AFM images obtained from Circular DNA P9. (A) Amplitude and (B) 3D image showing the ring crossing path more clearly. (C) Theoretical definition of positive DNA supercoils. (D) Section image analysis of the self-crossing.

3.4.2 Engineering of DNA Supercoils with Writhe Number of -2 and +2

Besides the DNA supercoils with writhe numbers of -1 and +1 (Figure 3.14 – Figure 3.18), a new circular DNA (Circular DNA 10) was designed and synthesized next in our studies (Figure 3.12) in order to construct a DNA supercoil with a writhe number of -2 and +2. After polymerase chain reaction was carried out with specific primers (ssODN-3 and ssODN-4), a 1068 base pair linear duplex DNA (Linear DNA 10) was produced. Linear DNA 10 was digested by SacI (a restriction endonuclease) and then incubated with T4 DNA ligase to form a covalently closed circular DNA (Circular DNA 10 in Figure 3.12A). This new circular DNA possesses 1040 base pairs in length and holds in theory 100 helical turns ($1040 \text{ base pairs} / 10.4 \text{ base pairs} = 100$ helical turns) in its overall backbones. In addition, different from Linear DNA 9, there are two identical PNA binding sites in the duplex backbones of Linear DNA 10 (see Table 3.2 for detail information), which extends over 21 base pairs successively, which can be invaded by two molecules of PNA. Similar to Circular DNA 9, the synthesis of Circular DNA 10 was accordingly examined by our electrophoresis analysis and AFM image as shown in Figure 3.19. When Circular DNA 10 was analyzed using AFM, it shows regular round-shaped structures as shown in Figure 3.19B (writhe = 0).

Figure 3.19 Synthesis and confirmation of Circular DNA 10 from Linear DNA 10. (A) Agarose gel electrophoretic analysis of DNA products; Lane 1: Molecular weight markers; Lane 2: Linear DNA 10 obtained by PCR amplification; Lane 3: Digested product of Linear DNA 10 by *SacI*; Lane 4: Circular DNA (Circular DNA 10) promoted by T4 DNA Ligase; Lane 5: reaction mixture of ligase reaction followed by Nuclease BAL-31 hydrolysis. (B) AFM image of covalently closed Circular DNA 10 with writhe numbers of 0 (scale bar 100 nm).

Figure 3.20 Synthesis and confirmation of Circular DNA P10. (A) Agarose gel electrophoretic analysis of DNA products. Lane 1: Molecular weight markers; Lane 2: Linear DNA 10 obtained by PCR amplification; Lane 3: Digested product of Linear DNA 10 by Sac I; Lane 4: PNA-containing linear DNA produced upon the invasion of PNA; Lane 5: PNA-containing circular DNA obtained by the action of T4 DNA ligase; Lane 6: Supercoiled DNA (Circular DNA N10) obtained after removal of PNA from PNA-containing circular DNA with relaxed form; (B) AFM image of negatively supercoiled structure with writhe number of -2 (scale bar 100 nm).

After Linear DNA 10 was digested by SacI, PNA invasion was allowed to take place, which leads to the opening up of a segment of about 21 base pairs in Linear DNA 10 simultaneously. The cohesive end termini is then joined together covalently by T4 DNA ligase. Removal of two molecules of PNA gives rise to the desired supercoiled DNA (Circular DNA N10) with writhe number of -2 (Figure 3.20). When Circular DNA N2 was examined next using AFM, it turns out that the obtained AFM images of the new circular DNA exhibit double self-crossings in their backbones as anticipated.

Moreover, it is clearly perceptible from the obtained AFM images that the writhe sign for each of the two backbone self-crossings in Circular DNA N2 is negative and overall double interwound structure of Circular DNA N2 is right-handed, which signifies that the writhe number of the newly engineered supercoil (Circular DNA N2) is -2 as intended originally. The detail analysis of AFM images as shown in Figure 3.21.

Figure 3.21 Detail analysis of AFM images obtained from Circular DNA N10. (A) Amplitude and (B) 3D image showing the ring crossing path more clearly. (C) Theoretical definition of negative DNA supercoils with writhe number of -2. (D) Section image analysis of the self-crossings.

Besides Circular DNA N10 that holds writhe number of -2, a circular DNA with writhe of +2 was constructed as well in our studies starting with Circular DNA 10 that contains two PNA-binding sites in its duplex sequence (Table 3.2). Prior to PNA invasion, this circular DNA exhibits regular round-shaped arrangement in its AFM images (Figure 3.19B). After two molecules of PNA invaded Circular DNA 10, the mobility shift of newly obtained PNA-DNA complex is different from that of Circular

DNA 10 and the resultant new DNA circular structure exhibit two left-handed self-crossings in its AFM images (writhe = +2) as it was intended in the beginning (Figure 3.22). This observation along with the results shown in Figure 3.22 implies that positive DNA supercoils (Circular DNA P10) could be readily engineerable as well through utilization of PNA invasion principle. In addition, we placed the measurement error data and the numbers of each type of DNA samples measured in our studies in a new table as shown in Table 3.4.

Figure 3.22 Engineering of positive supercoiled Circular DNA P10 with writhe number of +2. (A) Agarose gel electrophoretic analysis of DNA products. Lane 1: Molecular weight markers; Lane 2: Linear DNA 10 obtained by PCR amplification;

Lane 3: Digested product of Linear DNA 10 by SacI; Lane 4: Circular DNA (Circular DNA 10) promoted by T4 DNA Ligase; Lane 5: Reaction mixture of ligase reaction followed by Nuclease BAL-31 hydrolysis; Lane 6: PNA-containing Circular DNA P10 produced upon the invasion of PNA. (B) Theoretical definition of negative DNA supercoils with writhe number of -2. (C) AFM image of positively supercoiled structure with writhe number of +2 (scale bar 100 nm): I) height image; II) amplitude and III) 3D image showing the ring crossing path more clearly; IV) section image analysis of the self-crossing.

Table 3.4 Statistical data of DNA molecules examined using AFM and their measurement errors.

	Total number of DNA molecules measured	Height of duplex DNA molecules (nm)	Height of ring crossing (nm)	Number of DNA molecules with no self-crossing	Number of DNA molecules identified as positive supercoils	Number of DNA molecules identified as negative supercoils	Identifiable rate of self-crossing
Circular DNA 9	51	0.7 ± 0.1	—	51	—	—	—
Circular DNA N9	61	0.8 ± 0.1	1.3 ± 0.2	16	—	37	82% (37/45)
Circular DNA P9	69	0.6 ± 0.1	1.0 ± 0.3	17	35	—	67% (35/52)
Circular DNA 10	57	0.7 ± 0.1	—	55	—	—	—
Circular DNA N10	68	0.7 ± 0.1	1.2 ± 0.2	1	—	30	71% (30/42)
Circular DNA P10	65	0.7 ± 0.2	1.1 ± 0.2	2	29	—	73% (29/40)

3.4.3 Significance of Our Studies

The precise engineering of DNA supercoils demonstrated in the current studies could have certain implications in our understanding of the topological features of this biomacromolecule.¹⁹⁴⁻¹⁹⁶ Firstly, it is known that topoisomerase, DNA gyrase for example, converts relaxed forms of DNA merely into the forms of DNA that hold fixed superhelical density of -0.06 while the supercoils of DNA with superhelical densities between 0 and -0.06 *in vivo*.¹⁹⁷⁻¹⁹⁸ Our newly developed approaches could in theory be used to precisely engineer DNA supercoils with desired superhelical densities and writhe signs. Consequently, the new strategy reported in the current studies could offer the possibility to study the physical properties of the DNA supercoils with their superhelical densities that cannot be achieved by gyrase and reverse gyrase. On the other hand, the correlations between topoisomerase (e.g. Topoisomerase I and Topoisomerase II) and supercoiled DNA with different superhelical density (σ) engineered by our PNA invasion approach should be studied in the future, which could facilitate designing topoisomers as inhibitors of human topoisomerase I (a known anticancer drug target). In addition, some experiments concerning with sequence dependence of formation of different structures of DNA could be investigated. DNA, for example, was known to exist in either interwound or toroidal forms. DNA that possess different nucleotide sequences and compositions (but in the same length) could be engineered to form a fixed writhe number, which would help us to find out what are the sequences and compositions of nucleotides in a supercoiled DNA that are needed for the formation of interwound or toroidal conformations respectively.

Secondly, for example, a mathematical equation of $Lk - Tw = Wr$ was introduced to the field of molecular biology in 1970s for describing the topological features of DNA formed in chromosomes, which was subsequently named “DNA Topological Conservation Law”.^{48, 52-53} According to this law, the action of underwinding and overwinding on a relaxed form of DNA would lead to the formation of right-handed and left-handed supertwisting of the corresponding DNA respectively. The projected connection between the underwinding and right handedness as well as between the overwinding and left handedness in “DNA Topological Conservation Law” are now verified experimentally in our studies (Figure 3.14–3.22), which exemplifies the power and value of mathematics in biology.

Thirdly, it is known that 10.4~10.5 base pairs occur in average in each helical turn of DNA as per “Watson Crick Model”.^{20, 199} According to “DNA Topological Conservation Law”, on the other hand, alternation of one helical turn would lead to the generation of one writhe number. Our current studies demonstrate that PNA with 10 bases in length (10.4~10.5 bases) and two molecules of PNA with 20 bases in length (~ 21 bases) led to the generation of one and two new writhes respectively, which confirms experimentally the existence of a truthful connection between “Watson Crick Model” and “DNA Topological Conservation Law”.

3.5 Conclusion

DNA is stored either as a right handed or a left handed supercoil with a fixed magnitude of writhe in prokaryotic and eukaryotic cells. Our current studies demonstrate for the first time that the right and left handedness of DNA supercoils can

be engineered precisely and readily at the molecular level *in vitro* through utilization of the invading property of peptide nucleic acid. In addition, unlike the cellular process in which DNA can merely be converted into its supercoil with a fixed superhelical density, the PNA-invasion action can be utilized to engineer DNA supercoils with desired magnitudes of its writhes.

In conclusion, a PNA-based new approach has been established in the current investigations that can be used to precisely engineer the DNA supercoils with desired writhe signs and magnitudes. The manipulable DNA supercoils could be used for examining the correlation between the degree of forcible DNA curvature and DNA writhe number as well as could serve as molecular probes for unveiling the precise mechanism of actions of topoisomerases. Since the conversion between supercoiled and relaxed forms of DNA is constantly associated replication, transcription and transformation, it is our hope that the discoveries presented in this thesis could be beneficial to our further comprehension of the topological properties of DNA associated with its biological functions *in vivo*.

Chapter 4

Positive Supercoiling Affiliated with Nucleosome Repairs Non-B Structures of DNA

4.1 Introduction

DNA damage refers generally to chemical irreversible alternations of DNA structures in the prokaryotic and eukaryotic cells that are caused by endogenous metabolites and exogenous chemical agents or irradiations.^{8, 200-201} UV-B light is electromagnetic radiation with a wavelength of 280 - 320 nm and high UV-B radiation directly damage DNA, membranes and proteins in all organisms. UV-B light as an exogenous cause, for example, could lead to cycloaddition reactions between two carbon-carbon double bonds in two adjacent pyrimidines in organismal DNA.²⁰²⁻²⁰⁴ Since the generated covalent dimeric entities of pyrimidines are not able to resume to its original monomeric forms spontaneously, the molecular structures and subsequent cellular functions of the DNA at the sites of these two pyrimidines are considered as damaged by the UV-B light.²⁰⁴⁻²⁰⁵ In addition, endogenous reactive oxygen species produced from cellular metabolic pathways could result in chemical alteration of DNA structures as well in the forms of DNA alkylation, methylation and deamination, which frequently end up with the permanent loss of heterocyclic bases.²⁰⁶⁻²¹⁰ It is

estimated that spontaneous damage to DNA in human cells by endogenous and exogenous causes takes place at a rate of approximately 10,000 sites in every cell each day.⁸ If DNA damages occur in some vital sites of human genomes, these irreversible chemical modifications of DNA could obstruct the innate functions of cells and lead to mutations, transformation of normal cells to malignant cells as well as other detrimental biological consequences.^{201, 211-215}

In response to the attack by endogenous metabolites and exogenous causes, all organisms on earth have evolved delicate DNA repair mechanisms that are able to detect DNA damage, to activate the productions of related enzymes and proteins for DNA repair, and to further repair the damaged DNA.^{9, 12, 216} Once UV-B light-induced pyrimidine dimer is detected, for example, the signal of presence of structural alterations of DNA will be sent out in the prokaryotic cells and certain eukaryotic cells. DNA photolyase will be subsequently produced to repair the dimeric pyrimidines through “Direct Reversal” pathway in the presence of visible light.²¹⁷⁻²²⁰ In addition, the apurine and apyrimidine sites induced by endogenous reactive oxygen species, acid and ionizing radiation could be repaired by “General Excision Repair Pathway”, in which sequential actions of endonucleases, helicases, polymerase and DNA ligase take place.²²¹⁻²²³ Besides the single-strand DNA damage by formation of thymine dimers as well as by generation of apurine and apyrimidine sites, double-strand breaks could occur in organismal DNA.²²⁴⁻²²⁶ The mechanisms of non-homologous end joining,²²⁷⁻²²⁹ microhomology-mediated end joining,²³⁰⁻²³¹ and homologous recombination²³²⁻²³⁴ have been evolved in certain cells for effectively repairing the double-strand breaks of DNA. It is known nowadays that the ratio of repaired DNA

damages to overall damaged DNA (the efficiency of repairing of damaged DNA) relies on cell types, age of the cells, as well as other environmental factors.²³⁵⁻²³⁶ When cells are not capable of effectively repairing their damaged DNA, they will enter one of the three stages of senescence, apoptosis and unregulated cell division. DNA repair machineries are therefore vitally essential for maintaining genome integrity and for cellular and organismal functions.

In addition to the abovementioned chemical damage of DNA resulted in by endogenous and environmental causes, physical impairment of canonical B-form of DNA (e.g. formation of G-quadruplex¹⁰³, cruciform²³⁷, H-DNA²³⁸⁻²³⁹ and slipped DNA²⁴⁰⁻²⁴¹) often occur in organismal DNA as well. Similar to chemical modification of DNA, some of these physically altered canonical-B structures of DNA (non-B DNA) are incapable of returning to their initial Watson-Crick base-pairing in spontaneous manners once they are generated.²⁴²⁻²⁴⁴ Many of the non-B structures of DNA that have been discovered up till now are known to be stable under physiological conditions and to be utilized by organisms in a widespread manner as signals for cellular functions. G-quadruplexes among non-B structures, for instance, occur *in vivo* in the promoter region of c-MYC gene and served as a transcriptional repressor element for the expression of the gene²⁴⁵⁻²⁴⁶ while human genome contains ~376000 sites that have the potential to generate these types of non-B DNA structures²⁴⁷. In addition, G-quadruplexes are known to be very stable structural entities, melting points of which can be as high as 80 to 90 °C.^{101, 248}

Besides G-quadruplex conformations, cruciform DNA as a type of non-B DNA holds two stem-loop structures in its opposite strands, in which the length of their stems ranges from several to thousand base pairs.²⁴⁹⁻²⁵⁰ It has been well confirmed in the past that these hairpin types of DNA conformations are involved in a wide range of biological processes such as replication, regulation of gene expression, recombination and transposition.^{92, 97-98} In addition, sticky H-DNA and slipped DNA are known to possess stable structures. It is conceivable that if G-quadruplex, cruciform and other stable non-B structures of DNA cannot be repaired in time after they service as a cellular signal in a living organism are completed, these physical structures will obstruct the innate functions of cells as chemically damaged DNA does.

Figure 4.1 Pictorial illustration of topological relationship between circular DNA and nucleosome.

After each of cellular transactions of DNA (e.g. replication, transcription and recombination) completes, on the other hand, the DNA segments involved will wrap themselves around histone proteins to form nucleosome and chromatin for storage.^{9, 251-252} Since the DNA sequence bound around histone proteins is negatively

supercoiled in overall⁸¹, positive supercoil must be generated in the adjacent DNA segments in order to maintain an invariable writhe number on the whole according to “DNA Topological Conservation Law”^{48, 52-54}. As shown in Figure 4.1, the topological relationship between covalently closed circular DNA and nucleosome promoted by histone octamer has been illustrated.

Unlike a negative DNA supercoil that is underwound, positive DNA supercoils is overwound, which is anticipated to hold more backbone constraints than its negative counterpart does.⁵⁵ Accordingly, we speculate that the constraints associated with positive supercoils could provide “driving power”, which can repair the stable non-B structures of DNA. Here we report that our examination of repairing of stable G-quadruplex and DNA cruciform through formation of nucleosome. Our results confirm that positive supercoiling affiliated with nucleosome formation is indeed capable of disintegrating G-quadruplex and cruciform of DNA. In addition, our studies show that the PNA in the PNA-DNA duplex can be detached from DNA as well by the positive supercoiling generated in the assembly of nucleosome. Our finding suggest that generation of positive supercoiling through wrapping DNA around histone proteins could be an “ingenious” strategy adopted by eukaryotic cells for repairing non-B structures of DNA produced during dynamic transactions of DNA.

4.2 Design of DNA Sequences

4.2.1 Design of Circular DNA with G-quadruplex Structures

G-quadruplex was known as one of the most stable structures DNA can adopt in physiological conditions.¹⁰¹ With the purpose of examining whether non-B DNA structures could be indeed disintegrated by the action of wrapping DNA around histone proteins, the G-quadruplex-containing circular DNA is designed to serve as the substrates for our subsequent studies.

Figure 4.2 Pictorial illustration of generating G-quadruplex from duplex DNA using DNA gyrase.

Table 4.1 Nucleotide sequences of Circular DNA 11. The segment highlighted in red indicated the guanine-rich parts.

Name of DNA	Nucleotide sequence
Circular DNA 11	5'GAGCTCAGGATCCGGATGATCCCCAAAACCCCAAACCCAA AACCCAGTCCGTAATACGACTCACTTAAGGCCTTGACTAGAG GGTACCAACCTAGGTATCTAGAACCGGTCTCGAGCCATAACTT CGTATAGCATACATTATACGAAGTTATATAAGCTGTCAAACAT GAGAATTCTTGTTATAGGTTAATGTCATGATAATAATGGTTTC TTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACC CCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGC TCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAA AAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATT CCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGTCTACCCAG AAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGC ACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGTTA AGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATC GAAGGAGAGAAGAGCTG 3'

Two routes were designed for our synthesis circular DNA with G-quadruplex structures. Figure 4.2 depicts our first strategy for constructing G-quadruplex-containing circular DNA that possesses 575 base pairs in length (see Table 4.1 for the detail information of Circular DNA 11). The linear DNA 11 containing 603 base pairs in length can be obtain form polymerase chain reaction, in which (i) plasmid DNA (X2420G) served as the template and (ii) Primer 1 and Primer 2 were used as the forward primer and reverse primer to generate a duplex linear DNA (Structure 1 in Figure 4.2). Similar to the strategies for synthesis of circular DNA as mentioned above, two cohesive ends can be generated by the reaction of SacI digestion (Structure 2 in Figure 4.2). After the reaction of T4 DNA ligase on the linear DNA with paired ends, a circular DNA (Circular DNA 11) could be obtained, in which the guanine-rich

segment is contained (Structure 3 in Figure 4.2). Incubation of Circular DNA 11 with DNA gyrase could lead to the production of a negatively supercoiled DNA (Structure 4 in Figure 4.2), which then promoted the formation of a G-quadruplex-containing negative DNA supercoils (Structure 5 in Figure 4.2) in the presence of K^+ . If Structure 5 could be incubated with nicking endonucleases Nt.BsmAI (Structure 6 in Figure 4.2) and then the nick site could be subsequently sealed covalently by the action of DNA ligase, a desired relaxed form of G-quadruplex-containing circular DNA was obtained (Structure 7 in Figure 4.2).

Figure 4.3 Pictorial illustration of generating G-quadruplex from duplex DNA by alternative methods.

On the other hand, another alternative synthesis route could also be carried out for synthesis of circular DNA with G-quadruplex structures. As shown in Figure 4.3, Circular DNA 11 could be produced by the same strategy stated in Figure 4.2. After

nicking endonucleases Nt.BsmAI is used to generate one nick site in the duplex backbone of the Circular DNA 11, KCl could be added in the reaction mixture for the formation of G-quadruplex structures. The final product (Circular DNA G11) could be obtained through the ligase reaction which can covalently close the nicked site in one strand of the duplex of DNA circles.

4.2.2 Design of Circular DNA with Cruciform Structures

Cruciform structures are another typical non-canonical B conformations adopted by DNA. With the purpose of examining whether those cruciform structures could be repaired by the actions associated with positive supercoils and nucleosome assembly, a circular DNA with inverted repeat sequence is designed in our studies.

Figure 4.4 depicted our design of synthesis circular DNA with cruciform structures. To construct the circular DNA with cruciform structures, a linear (Linear DNA 12) could be designed and synthesized through the polymerase chain reaction using plasmids X4510E and two primers (primer 3 and primer 4), which comprise the inverted repeat sequences as shown in Table 4.2. The same strategy is applied for the synthesis of the circular DNA as discussed above (restriction endonucleases digestion by SacI and ligase reaction) and a circular DNA (Circular DNA 12) could be obtained, which has the inverted repeat sequences in its duplex DNA backbones. It has been well established in the past that the cruciform structures can be formed in negative supercoiled circular DNA which promotes breathing effect in the double helix.²⁸ Therefore, Circular DNA 12 can be treated with DNA gyrase which can introduce the negative supercoiling into DNA circles. Our final desired cruciform-containing

circular DNA (Circular DNA C12) will be produced in the presence of 20 mM MgCl_2 according to the literature reports.¹⁰⁰ A cruciform structure of DNA can be divided into stem and loop regions separately while the length of the stem is known to affect the stability of the cruciform to a high degree. Our newly designed Circular DNA C12 possesses a stem with 30 base pairs and a loop with 3 base pairs, which can stabilize the cruciform structures once it forms. In addition, the stems with 30 base pairs can be identified under the AFM analysis.

Figure 4.4 Pictorial illustration of our strategy for synthesis of cruciform-containing circular DNA.

Table 4.2 Nucleotide sequences of Circular DNA 12. The segment highlighted in red indicated the inverted repeat sequences.

Name of DNA	Nucleotide sequence
Circular DNA 12	5'GAGCTCCTCGATGAAAGATCCTTTCCGGAGATCCTTGATTGCG AGCATAGCTGGCTGGTGTGCGGCAGTCCGCCTTGACTAGAGG GTACCAACCTAGGTATCTAGAACGAATTCCGGAGCCTGAATCG GCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTC CGCTTCCTCGCTCACTGATTCGCTGCGCTCGGTCGTTCCGGCTG CGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTAT CCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAT CAAGGCCAGCAAAAGGCCAGGAACCGTAAACAAGGCCGCGTTG CTGGCGTGACGAGCATCACAACAATCGACGCTCAAGTCAGAG GTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCG ACTAGTGCCCTGGAAGCTCCCTCGTGCGCTCATAAGAAGGAGA GAAGCTAAGAGAGGAACTGGACTCTCAAACATGAAACGTTTTG TTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAG GTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTAA ATACATTCAAATATGTATCCGCTCATGATACAATAAGTCTCCC CTGATAAATGCTTCAATGAAGGAAGAGTATGAGTATTCAACAT TTCCGTGTCGCCCTTATTCCCTTTTGCACAACATGGGGGATCA TGTAACCTCGCCTTGATCGGAGCTGAATGAAGCCATAACAAACG ACGAGCGTGACACCACGATGCCTGCAGCTCGAGCCCTGAATGT ATTTAGCGCCAGGGTTTTCCCAGTCACGACCGCACATTTCCCC GAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGAC TCCTGTGTGAAATTGTTATCCGCTCACGAGGCCCTTTCGCCTC GCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGC TCCCGGAGGCGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAG CAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTGGCGGGTGT CGGGGCTGGCTTAACATATGCGGCATCAGAGCAGATTGTACTGA GAGTGCACCATATGGACATATTGTCGTTACCGAATTCATGGAC TAGTGAATCGTATTACGTCTGTGTGATTGTTATCCGAGCTTAT CAAACCACCGCTCGCCAAAGGATCTCCGGAAAGGATCTTTCA TC 3'

4.2.3 Design of Covalently Closed PNA-containing Circular DNA

With the aim of examining whether PNA molecules could be removed from PNA-containing circular DNA by the action of wrapping DNA around histone proteins, a PNA-containing circular DNA is designed in our studies. As shown in Figure 4.5, Circular DNA 9 was reconstructed as the starting material which has one PNA binding site in its duplex backbones. After one nicking site is obtained by the action of nicking endonuclease (Nt.BsmAI), PNA invasion could be carried out. Finally, ligase reaction seal the nicking sites and produce the covalently closed PNA-containing circular DNA (Circular PNA-DNA 9).

Figure 4.5 Schematic illustrations of our strategy for synthesis of covalently closed PNA-containing circular DNA. The structure of PNA are given: PNAs are written from the N terminus to the C terminus using normal peptide conventions: H is a free

amino group; NH₂ is a terminal carboxamide; Lys is the lysine residue; J denotes pseudoisocytosine and eg1 denotes the linker unit, 8-amino-3,6-dioxaoctanoic acid.

4.3 Materials and Methods

4.3.1 Duplex DNA, Enzymes and Chemicals

Product(s)	Manufacturer
Plasmid DNA (X2420G)	Generay Biotech (Shanghai, China)
Plasmid DNA (X4510E)	Generay Biotech (Shanghai, China)
Peptide nucleic acids	Biosynthesis (Lewisville, Texas)
Primers	Sigma-Proligo (Singapore)
100 bp DNA ladder	Fermentas (Singapore)
Proteinase k	Fermentas (Singapore)
<i>Taq</i> Polymerase	New England Biolabs (Ipswich, MA)
1 Kb DNA ladder	New England Biolabs (Ipswich, MA)
SacI endonuclease	New England Biolabs (Ipswich, MA)
Nt.BsmAI	New England Biolabs (Ipswich, MA)
T4 DNA ligase	New England Biolabs (Ipswich, MA)

BAL 31 exonuclease	New England Biolabs (Ipswich, MA)
Nucleosome Assembly Kit	New England Biolabs (Ipswich, MA)
Biological purity water	1st Base Pte. Ltd (Singapore)
Agarose	Invitrogen (Carlsbad, CA)
SYBER Gold	Invitrogen (Carlsbad, CA)
QIAquick PCR purification kit	Qiagen (Singapore)
QIAquick Gel Extraction Kit	Qiagen (Singapore)
Mini Prep Cell	Bio-Rad (Hercules, CA)
TAE, TBE, TRIS	1st Base Pte. Ltd (Singapore)

4.3.2 Polymerase chain reactions for synthesis of Linear DNA 11 and Linear DNA 12

Polymerase chain reaction was carried out following standard procedures with Taq DNA Polymerase as described as follows: A reaction mixture containing 1 ng plasmid DNA (X2420G for Linear DNA 11 and X4510E for Linear DNA 12), 0.25 μ M forward primer, 0.25 μ M reverse primer, 200 μ M dNTP, 1 U *Taq* polymerase in a total volume of 50 μ l reaction buffer (20 mM Tris-HCl, 10 mM (NH₄)₂SO₄, 10 mM KCl, 2 mM MgSO₄, 0.1% Triton X-100, pH 8.8 @ 25 °C) was processed as below protocol:

95 °C for 180 sec (denature); 60 °C for 40 sec (anneal for Linear DNA 11); 58 °C for 40 sec (anneal for Linear DNA 12) and 72 °C for 40 sec (elongate) (60 sec per kb target sequence length); 29 cycles only (otherwise enzyme decay causes artifacts); 72 °C for 10 min at end to allow complete elongation of all product DNA

The primer 1 and primer 2 are forward and reverse primers for Linear DNA 11 while the primer 3 and primer 4 are forward and reverse primers for Linear DNA 12. 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. In addition, the nucleotide sequence of Linear DNA 11 and Linear DNA 12 were given in Table 4.3.

Table 4.3 Nucleotide sequences of primers used in polymerase chain reactions.

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

4.3.3 Reactions of SacI with Duplex Linear DNA Precursors

The duplex linear DNA precursors containing two SacI digest site in each end were obtained from polymerase chain reactions. In order to create two cohesive ends as shown in the Figure 4.3 and Figure 4.4, the linear DNA were treated with SacI

endonuclease. Linear DNA 11 and Linear DNA 12 were obtained as described as follows: A solution containing 10 mM Bis-Tris-Propane-HCl, 10 mM MgCl₂, 1 mM Dithiothreitol, purified PCR products (200 ng) and 10 U SacI was incubated at 37 °C for 1 hr. The reaction products were further analyzed using agarose electrophoresis (1.5%) and purified using QIAquick PCR purification kit before the next steps.

4.3.4 Preparations of Circular DNA Using T4 Ligase

As there are two identical cohesive ends digested by Sac I in Linear DNA 11 and Linear DNA 12, Circular DNA 11 and Circular DNA 12 were obtained from ligase reactions as described as follows: A 50 µl solution containing 50 mM Tris-HCl, 10 mM MgCl₂, 1 mM ATP, 10 mM dithiothreitol, 500 ng linear DNA with cohesive ends and 20 U T4 DNA ligase was incubated at 16 °C for 8 hrs. The obtained circular DNA products were further analyzed using agarose electrophoresis (1.5%) and purified using QIAquick PCR purification kit before the next steps.

4.3.5 Degrade Linear DNA from Ligase Reaction Mixture Using Nuclease BAL-31 Exonuclease

Nuclease BAL-31 exonuclease degrades both 3' and 5' termini of duplex DNA without generating internal scissions. To remove the linear DNA from the mixture of ligase reaction as well as further confirm that there is no nicks, gaps and single-stranded regions in the obtained DNA products, we treated the ligase reaction mixture by nuclease BAL-31 exonuclease as described as follows: A 50 µl solution containing 20 mM Tris-HCl, 600 mM NaCl, 12 mM CaCl₂, 12 mM MgCl₂, 1 mM EDTA, 500 ng

reaction products of ligase reaction and 2 U exonuclease BAL-31 was incubated at 30 °C for 10 hrs. The obtained circular DNA products were further analyzed using agarose electrophoresis (1.5%) and purified using purified next using Mini Prep Cell.

4.3.6 Reactions of Nt.BsmAI with Circular DNA

To create one nicking site in one strand of DNA duplex, nicking endonuclease (Nt.BsmAI) was used as described as follows: A solution containing 20 mM Tris-acetate, 50 mM potassium acetate, 10 mM Magnesium Acetate, 1 mM Dithiothreitol, Circular DNA 9 (200 ng) and 10 U Nt.BsmAI was incubated at 37 °C for 1 hr. The reaction products were further analyzed using agarose electrophoresis (1.5%).

4.3.7 PNA Invasion

In the binding reactions of PNA to target sites, the concentration of PNA was kept at a large excess over the DNA concentration, and the binding was performed at in pre-siliconized tubes at 37 °C for 6 hr in 10mM Sodium-Phosphate Buffer (PH = 6.9). In addition, the gel-mobility-shift experiments were performed in 1.5% agarose gels in 1 x TAE buffer. The electrophoresis was run at 100 V for four hours at room temperature and the gels were then stained with SYBER Gold.

4.3.8 Nucleosome Assembly

The nucleosome assembly was conducted using “EpiMark® Nucleosome Assembly Kit” (NEB). The protocol is described as follows: A 10 µl solution containing 0~1.5 µl H₂O, 1 µl NaCl, 5 pmol DNA, 3.75 µl Dimer (20µM) and 3.75 µl

Tetramer (10 μ M) was mixed and incubated in room temperature for 30 minutes. Then (1) Add 3.5 μ l dilution buffer (10 mM Tris, pH 8.0) to the mixture mentioned above at room temperature. This brings the reactions to 1.48 M NaCl, 13.5 μ l total volume. Incubate at room temperature for 30 minutes. (2) Add 6.5 μ l dilution buffer (10 mM Tris, pH 8.0) to the mixture at room temperature. This brings the reactions to 1.0 M NaCl, 20 μ l total volume. Incubate at room temperature for 30 minutes. (3) Add 13.5 μ l dilution buffer (10 mM Tris, pH 8.0) to the mixture at room temperature. This brings the reactions to 0.6 M NaCl, 33.5 μ l total volume. Incubate at room temperature for 30 minutes. (4) Add 46.5 μ l dilution buffer (10 mM Tris, pH 8.0) to the mixture at room temperature. This brings the reactions to 0.25 M NaCl, 80 μ l total volume. Incubate at room temperature for 30 minutes. (5) Store samples at 4 $^{\circ}$ C.

4.3.9 AFM Studies of Obtained Circular DNA

To immobilize DNA for AFM imaging, a mica surface modified with APS was used as a substrate. DNA adheres to the imaging surface through electrostatic attraction as discussed in section of 2.3.8. Prepare the solution of the DNA sample in appropriate buffer. DNA concentration should be between 0.1 and 0.01 μ g/ml depending on the size of the molecules. Place 5–10 μ l of the solution in the middle of APS-mica substrate (usually 1 x 1 cm²) for 2–3 min. Rinse the surface thoroughly with water (2–3 ml per sample) to remove all buffer components. AFM images 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.

4.4 Results and Discussion

4.4.1 Construct Covalently Closed Circular DNA with G-quadruplex Structures

Since G-quadruplex is widespread eukaryotic DNA, construction of circular DNA that contains well defined G-quadruplex as well as disintegration of the tetraplex structure by formation of nucleosome have been accordingly carried out in our studies. Primer 1 was designed to contain guanine-rich segment from which G-quadruplex structure could be formed in the later stages of the synthetic process as described above. Polymerase Chain Reaction was carried out to produce Linear DNA 11 at first using template and primers. 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.²⁵³⁻²⁵⁶ With the purpose of investigating whether the G-quadruplex structures could be formed through double strands conformation of DNA, Linear DNA 11 which has guanine-rich segment was incubated with KCl and PEG 200. The reaction mixture was examined by electrophoresis analysis and AFM (Figure 4.6), which indicated that the G-quadruplex structures are indeed formed (Linear DNA

G11) through incubation of duplex DNA with guanine-rich segment in the presence of KCl and molecular crowding condition.

Figure 4.6 Examination of formation of G-quadruplex structures from duplex linear DNA with guanine-rich segment. (A) Gel electrophoresis analysis. Lane 1: Molecular weight markers; Lane 2: Linear DNA 11 obtained by PCR amplification; Lane 3 Linear DNA G11 formation in double stranded DNA in solution containing 150 mM KCl and 40% PEG 200. (B) AFM image of Linear DNA 11. (C) AFM image of Linear DNA G11 (scale bar 200 nm).

After Linear DNA 11 was digested by SacI endonuclease, a circular structure was (Circular DNA 11) formed through complementarity of two cohesive ends by the

action of T4 DNA ligase. Then Circular DNA 11 was digested by Nt.BsmAI, a nicking endonuclease which can cleave one strands of duplex DNA and release the free energy caused by supercoiling form DNA circles. Incubation of nicked DNA circles with 150mM KCl and 40% PEG 200 led to the formation of G-quadruplex structures in one strands of duplex DNA. A desired relaxed form of G-quadruplex-containing circular DNA (Circular DNA G11) was obtained when T4 DNA ligase had been used to seal the nicked site in double strand DNA backbones. The formations of Circular DNA 11 as well as Circular DNA G11 were accordingly examined by the electrophoresis as shown in Figure 4.7.

Figure 4.7 Gel electrophoresis analysis of formation of G-quadruplex structures in circular DNA. (A) Synthesis of Circular DNA 11. Lane 1: Molecular weight markers; Lane 2: Linear DNA 11 obtained by PCR amplification; Lane 3: Digested product of Linear DNA 11 by SacI; Lane 4: Circular DNA (Circular DNA 11) promoted by T4 DNA Ligase; Lane 5: reaction mixture of ligase reaction followed by Nuclease BAL-

31 hydrolysis. (B) Synthesis Circular DNA G11 with G-quadruplex structures. Lane 1: Molecular weight markers; Lane 2: Circular DNA 11; Lane 3: Nicked Circular DNA11 using Nt.BsmAI; Lane 4: Incubation of nicked circular DNA in 150 mM KCl and 40% PEG 200; Lane 5: Circular DNA G11 obtained from ligase reaction.

Figure 4.8 AFM image of circular DNA with and without G-quadruplex structures. (A) Left: AFM images of Circular DNA 11 (Lane 5 in Figure 4.7A). Right: Section image analysis of the height of DNA backbones (scale bar 200 nm). (B) Left: AFM image of Circular DNA G11 (Lane 5 in Figure 4.7B). Right: Section image analysis of the height of DNA backbones (scale bar 200 nm).

On the other hands, atomic force microscopy was used to detect the formation of G-quadruplex and C-rich strand structures in backbone of DNA circles. As shown in Figure 4.8, Circular DNA 11 showed a regular round-shaped structure while Circular DNA G11 exhibited a different conformation. The further section analysis reveals that

the height of duplex DNA backbone ranges from 0.6~0.8 nm. On the other hand, the G-quadruplex structure constructed by four-strand nucleic acid showed the height of backbone ranged from 1.1~1.3 nm. The observation of structure difference between Circular DNA 11 and Circular DNA G11 implied that G-quadruplex structures indeed formed in the presence of K^+ and molecular crowding conditions.

4.4.2 Disintegrate G-quadruplex Structures from Circular DNA through the Nucleosome Assembly Associated with Positive Supercoiling

DNA segments will wrap themselves around histone proteins to form nucleosome and chromatin for storage after the dynamic cellular transactions of DNA (e.g. replication, transcription and recombination) completes. Because wrapping DNA around histone octamers is in a negative supercoiled conformation, the opposite forms (positive supercoiling) must be introduced in the rest part of DNA circles (covalently closed). It is well known that positive supercoils hold more backbone constraints than its negative counterpart. As shown in Figure 4.9, if a covalently closed circular DNA containing G-quadruplex structure (Circular DNA G11) will be treated with histone octamers, the positive supercoil could be generated in the duplex backbones of DNA circle. The Structure 2 (Figure 4.9) is a high energy intermediate with higher constraints that will provide a “driving power” to disintegrate G-quadruplex structures from duplex DNA backbones (Structure 3 in Figure 4.9). After the histone proteins are

digested with proteinase K and followed by the action of Topo I relaxation, Circular DNA 11 could be obtained, which has a regular B structures in overall duplex backbones.

Figure 4.9 Schematic illustrations of the disintegration of non-B structure (G-quadruplex) of DNA by nucleosome's positive-supercoil-introducing activity.

Based on the properties of nucleosome in DNA circles as discussed above, our newly synthesized Circular DNA G11 was accordingly treated with histone octamers to examine whether the G-quadruplex structures could be disintegrated from the circular DNA. As anticipated, Circular DNA G11 was transformed in to Circular DNA 11 after the three steps of reactions (histone octamer binding; proteinase K digestion and Topo I relaxation). The disintegration of G-quadruplex structures from Circular DNA G11 was examined by the gel electrophoresis analysis and AFM (Figure 4.10).

Figure 4.10 Examination of the disintegration of G-quadruplex structures from DNA circles. (A) Gel electrophoresis analysis. Lane 1: Molecular weight markers; Lane 2: Circular DNA G11 with G-quadruplex structures; Lane 3: Nucleosome assembly products from Circular DNA G11; Lane 4: Digestion products with proteinase K; Lane 5: Relaxation products with TopoI. (B) AFM image of DNA circles obtained from disintegration of Circular DNA G11 (Samples purified from Lane 5 in Figure 4.10A). (C) Section image analysis of the height of DNA backbones (scale bar 200 nm).

As a control experiment, Circular DNA G11 was treated with the same procedures as illustrated in Figure 4.10 but in the absence of histone proteins during the binding step. The results of gel electrophoresis and AFM analysis indicated the G-quadruplex structures still remained in the DNA circles (Figure 4.11). The observation

of disintegration of G-quadruplex structures suggested that the nucleosome formation associated with positive supercoiling is indeed capable of remove the non-B structures from duplex DNA circles.

Figure 4.11 Examination of the disintegration of G-quadruplex structures from DNA circles but in the absence of histone proteins. (A) Gel electrophoresis analysis. Lane 1: Molecular weight markers; Lane 2: Circular DNA G11 with G-quadruplex structures; Lane 3: Reaction mixture obtained from the procedures of nucleosome assembly in the absence of histone proteins; Lane 4: Reaction mixture digested products by proteinase K; Lane 5: Relaxation products with TopoI. (B) AFM image of DNA circles obtained from disintegration of Circular DNA G11 (Samples purified from Lane 5 in Figure 4.10A). (C) Section image analysis of the height of DNA backbones (scale bar 200 nm).

4.4.3 Construct Covalently Closed Circular DNA with Cruciform Structures

Cruciform structure, another typical non-canonical B conformation, is important for the critical biological processes of DNA recombination and repair that occur in the cell.⁹⁰⁻⁹² With the aim of examine whether cruciform structures could be disintegrated by the action of nucleosome assembly, a circular DNA (Circular DNA 12) was synthesized in our studies, which possesses inverted repeat sequences in its circular structures. The construction of Circular DNA 12 adopted the same strategy as those for synthesis of Circular DNA 11 as shown in the section of 4.2.2. PCR amplification gave the Linear DNA 12 and SacI digestion produced two paired cohesive ends. After the ligase reaction was carried out, Circular DNA 12 was obtained as shown in Figure 4.12.

Figure 4.12 Examination of synthesis of Circular DNA 12. (A) Gel electrophoresis analysis. Lane 1: Molecular weight markers; Lane 2: Linear DNA 12; Lane 3: Digested product of Linear DNA 12 by SacI; Lane 4: Reaction mixture of ligase

reaction to give the DNA circles; Lane 5: reaction mixture of ligase reaction followed by Nuclease BAL-31 hydrolysis. (B) AFM image of Circular 12 (DNA samples were purified from Lane 5 in Figure 4.12A. Scale bar 100 nm).

To form a cruciform structure, negative supercoils could be introduced into the DNA circles due to the underwinding of double helix and breathing effect as illustrated in Figure 4.4. As a result, DNA gyrase was used to generate negative supercoils in Circular DNA 12 (Circular DNA N12). After the final addition of 20 mM $MgCl_2$ was carried and incubation of the reaction mixtures at room temperature for 24 hours, the cruciform-containing circular DNA (Circular DNA C12) was obtained. The synthesis of cruciform-containing circular DNA (Circular DNA C12) was examined by Gel electrophoresis and AFM analysis as shown in Figure 4.13.

Figure 4.13 Examination of synthesis of Circular DNA C12. (A) Gel electrophoresis analysis. Lane 1: Molecular weight markers; Lane 2: Negative supercoils were introduced by DNA gyrase with Circular DNA 12; Lane 3: Negative supercoiled DNA circles were incubated in the presence of 20 mM $MgCl_2$ to form cruciform structures.

(B) AFM image of Circular N12. (C) AFM image of cruciform-containing DNA circles (Circular DNA C12) (Scale bar 100 nm).

4.4.4 Disintegrate Cruciform Structures from Covalently Closed Circular DNA through Introduction of Positive Supercoils Affiliated with Nucleosome Assembly

Our newly synthesized cruciform-containing circular DNA (Circular DNA C12) was accordingly treated with histone octamer to examine whether the introduced positive supercoils associated with nucleosome assembly could disintegrate the cruciform structures in covalent closed DNA circles. As shown in Figure 4.14A, after Circular DNA C12 (Lane 2 in Figure 4.14A) was treated with histone octamers, nucleosomes were obtained (Lane 3 in Figure 4.14A), in which positive supercoils were introduced. After the histone proteins were digested by proteinase K, a new band different from Circular DNA C12 could be observed, which has a faster mobility shift (higher superhelical density) than its former conformation (Circular DNA C12). As a result, we speculate that cruciform structures were disintegrated and Circular DNA C12 was transformed into a circular DNA with negative supercoils because it has the same mobility shift as Circular DNA N12 (Lane 2 in Figure 4.13A). The later AFM experiment showed that the cruciform structures were indeed disintegrated after the actions of nucleosome assembly and digestion of proteins. As shown in Figure 4.14B,

no cruciform conformation was observed in the AFM image of DNA circles which were obtained from two steps of reaction upon Circular DNA C12.

Figure 4.14 Examination of the disintegration of cruciform structures from DNA circles. (A) Gel electrophoresis analysis. Lane 1: Molecular weight markers; Lane 2: Circular DNA C12 with cruciform structures; Lane 3: Nucleosome assembly products from Circular DNA C12; Lane 4: Digestion products with proteinase K. Lane 5: Negative supercoiled DNA circles obtained from DNA gyrase with Circular DNA 12. (B) AFM image of DNA circles obtained from disintegration of Circular DNA C12 (Samples were purified from Lane 4 in Figure 4.14A and were immobilized on mica surface immediately once they were obtained. Scale bar 100 nm).

It should be pointed out that the transformation of Circular DNA C12 to Circular DNA N12 is a equilibrium process, which depends on many factors (For example, the concentration of Mg^{2+} or Na^{+} , temperature and incubation time). The newly obtained Circular DNA N12 (Lane 5 in Figure 4.14A) could be converted into cruciform-

containing circular DNA in the presence of Mg^{2+} after the reaction mixture was incubated for more than 2 hours. On the other hand, the control experiment was carried out using the same protocol as illustrated in Figure 4.14 but in the absence of histone proteins during the binding step. The results of gel electrophoresis and AFM analysis showed that the cruciform structures still remained in the DNA circles (Figure 4.15). The experiment data from Figure 4.14 and Figure 4.15 indicated that the disintegration of cruciform structures could be achieved by the introduction of positive supercoils associated with nucleosome assembly in a covalently closed DNA circle.

Figure 4.15 Examination of the disintegration of cruciform structures from DNA circles but in the absence of histone proteins. (A) Gel electrophoresis analysis. Lane 1: Molecular weight markers; Lane 2: Circular DNA C12 with cruciform structures; Lane 3: Reaction mixture obtained according to the procedures of nucleosome assembly but in the absence of histone proteins; Lane 4: Reaction mixture digested products by proteinase K. Negative supercoiled DNA circles obtained from DNA gyrase with Circular DNA 12. (B) AFM image of DNA circles obtained from

disintegration of Circular DNA C12 in the absence of histone proteins (Samples purified from Lane 4 in Figure 4.15A. Scale bar 100 nm).

4.4.5 Construct and disintegrate Covalently Closed PNA-containing Circular DNA

Figure 4.16 Examination of synthesis of PNA-containing circular DNA. (A) Gel electrophoresis analysis. Lane 1: Molecular weight markers; Lane 2: Linear DNA 9; Lane 3: Digested product of Linear DNA 9 by SacI; Lane 4: Reaction mixture of ligase reaction; Lane 5: Reaction mixture of ligase reaction followed by Nuclease BAL-31 hydrolysis; Lane 6: Nicked product of Circular DNA 9 by Nt.BsmAI; Lane7: Incubation of nicked Circular DNA9 with PNA; Lane 8: Invasion products was treated by T4 DNA ligase (Circular PNA-DNA 9). (B) AFM image of Circular PNA-DNA 9. (C) Section image analysis of the height of DNA backbones (scale bar 200 nm).

In addition to G-quadruplex, cruciform and other physically altered structures formed by certain particular sequences of DNA itself, single-stranded DNA binding proteins and DNA-binding and intercalating small molecules could interrupt the integrity of Watson-Crick base pairing as well to form non-DNA molecule-assisted non-B DNA structures²⁵⁷⁻²⁵⁸. Peptide nucleic acids (PNAs) are analogs of DNA, which can invade DNA duplex via forming triplex structures. In chapter 3, we demonstrated positive and negative DNA supercoils can be engineered precisely and readily at the molecular level *in vitro* through utilization of the invading property of peptide nucleic acid. In this section, on the other hand, the removal of PNA (disintegration of P-loop structures) from DNA circles will be discussed. To examine whether bis-PNA could be removed from acircular PNA-DNA complex through action of nucleosome assembly associated with positive supercoils, a PNA-containing circular DNA (Circular PNA-DNA 9) was synthesized in our studies. Circular DNA 9 was reconstructed firstly (Lane 5 in Figure 4.16A) and then a nicked site was produced by the addition of Nt.BsmAI (Lane 6 in Figure 4.16A). PNA invasion was carried out next (Lane 7 in Figure 4.16A) and it was followed by the ligase reaction for sealing the nicking site in the DNA duplex backbone to give a covalently closed DNA circle (Lane 8 in Figure 4.16A). The newly synthesized Circular PNA-DNA 9 was also examined by AFM, which showed that triplex structures (0.9 ± 0.2 nm) existed in the duplex backbone of DNA circles.

Our newly synthesized PNA-containing DNA circles (Circular PNA-DNA 9) were treated with histone octamers, which promote DNA circles wrap themselves around histone proteins to form nucleosomes (Lane 3 in Figure 4.17A). After the

digestion conducted by proteinase K, DNA circles were treated by Topo I to obtain the circular DNA with relaxed forms (Lane 5 in Figure 4.17A), which showed a different mobility shift from that of Circular PNA-DNA 9. Similar to our studies before, AFM analysis was accordingly carried out as well. There was no identifiable non-B structure (P-loop shown in Figure 4.16B and C) observed in the AFM image of DNA samples purified from Lane 5 in Figure 4.17A.

Figure 4.17 Examination of the disintegration of P-loop structures from DNA circles.

(A) Gel electrophoresis analysis. Lane 1: Molecular weight markers; Lane 2: Circular PNA-DNA 9 with P-loop structures; Lane 3: Nucleosome assembly products from Circular PNA-DNA9; Lane 4: Digestion products with proteinase K; Lane 5: Relaxation products with TopoI. (B) AFM image of DNA circles obtained from

disintegration of Circular PNA-DNA 9 (Samples purified from Lane 5 in Figure 4.10A). (C) Section image analysis of the height of DNA backbones (scale bar 200 nm).

As a control experiment, Circular DNA PNA-DNA 9 was also treated with the same procedures as illustrated in Figure 4.17 but in the absence of histone proteins during the binding step. As shown in Figure 4.18, no mobility shift difference can be observed in the gel electrophoresis and non-B structures (P-loop) still remained in the DNA circles after the final relaxation of DNA products (Lane 5 in Figure 4.18A). The observation of disintegration of PNA-DNA complex structures gave us evidence that the nucleosome formation associated with positive supercoiling is indeed capable of remove the non-B structures from duplex DNA circles.

Figure 4.18 Examination of the disintegration of P-loop structures from DNA circles but in the absence of histone proteins. (A) Gel electrophoresis analysis. Lane 1: Molecular weight markers; Lane 2: Circular PNA-DNA 9 with P-loop structures; Lane 3: Reaction mixture from the procedures of nucleosome assembly in the absence of histone proteins; Lane 4: Reaction mixture digested products by proteinase K; Lane 5: Relaxation products with TopoI. (B) AFM image of DNA circles obtained from disintegration of Circular PNA-DNA 9 (Samples purified from Lane 5 in Figure 4.10A). (C) Section image analysis of the height of DNA backbones (scale bar 200 nm).

4.4.6 Significance of Our Studies

DNA damages refer commonly to chemical modifications of DNA structures in the prokaryotic and eukaryotic cells that make the DNA molecules incapable of resuming their original B conformations in a spontaneous manner⁴⁵. In response to the attack of cellular DNA by endogenous metabolites and exogenous causes, all organisms have evolved delicate DNA repairing mechanisms that are able to detect DNA damages, to activate productions of related enzymes and proteins, and to further repair their damaged DNA^{201, 259}.

Besides these well-known chemical damages to DNA, physical alterations of canonical B-form of DNA such as formations of G-quadruplex, cruciform and sticky DNA routinely occur in organismal DNA that serve as signals for specified cellular

actions²⁶⁰⁻²⁶¹. Similar to chemical damages of DNA, many of the non-B DNA structures, once formed, are incapable of resuming their original Watson-Crick base pairings in a spontaneous manner, which could cause damages to DNA in a physical fashion. It is conceivable that if stable non-B DNA structures cannot be repaired in time after their services as cellular signals in living organisms complete, these physically damaged DNA will obstruct the subsequent innate functions of cells in the same ways as chemically damaged DNA does. Unlike the repairing mechanisms of chemically damaged DNA, however, the driving forces and pathways for repairing physically damaged DNA in living organisms have not yet been well understood. In our studies, we demonstrated that positive supercoiling affiliated with nucleosome formation can act as the driving force to repair G-quadruplex, cruciform as well as a stable non-B DNA structure caused by peptide nucleic acid. Our discoveries of the new roles of DNA positive supercoiling affiliated with nucleosome formations may be relevant to the repairing mechanisms of physically damaged DNA in the living organisms.

4.5 Conclusion

Chemical damages to DNA and their repairing mechanisms are the typical topics discussed in nearly every textbook of genetics, biology, biochemistry and nucleic acids. Even though formation of stable non-B DNA structures have been recognized for a few decades, on the other hand, the repairing mechanisms for the physically damaged DNA in organisms have not yet been known. It is conceivable that if the

physically damaged DNA cannot be repaired in time *in vivo*, they will obstruct the subsequent innate functions of cells in the same ways as chemically damaged DNA does. We discovered now that physically damaged DNA (G-quadruplex, cruciform as well as a stable non-B DNA structure caused by peptide nucleic acid) can be repaired by positive supercoiling affiliated with nucleosome formation.

Since our new discoveries are related to DNA repair, nucleosome formation, DNA topology and thermodynamic stability of non-B DNA structures, it is our hope that the discoveries presented in this thesis could be beneficial to our further comprehension of the topological properties of DNA associated with its biological functions *in vivo*.

References

1. R. R. Sinden, *DNA structure and function*, Academic Press, San Diego, 1994.
2. A. D. Bates and A. Maxwell, *DNA topology*, Oxford University Press, Oxford ; New York, 2005.
3. M. McCarty, *J Exp Med*, 1994, **179**, 385-394.
4. S. C. Case-Green, K. U. Mir, C. E. Pritchard and E. M. Southern, *Curr Opin Chem Biol*, 1998, **2**, 404-410.
5. J. E. Zull and S. K. Smith, *Trends Biochem Sci*, 1990, **15**, 257-261.
6. J. Nordman and T. L. Orr-Weaver, *Development*, 2012, **139**, 455-464.
7. J. F. Diffley, *Philos Trans R Soc Lond B Biol Sci*, 2011, **366**, 3545-3553.
8. S. Lagerwerf, M. G. Vrouwe, R. M. Overmeer, M. I. Fousteri and L. H. Mullenders, *DNA Repair (Amst)*, 2011, **10**, 743-750.
9. S. Messner and M. O. Hottiger, *Trends Cell Biol*, 2011, **21**, 534-542.
10. A. M. Pyle, *Annu Rev Biophys*, 2008, **37**, 317-336.
11. M. D. Szczelkun, *Biochem Soc Trans*, 2011, **39**, 589-594.
12. N. O. Knudsen, S. D. Andersen, A. Lutzen, F. C. Nielsen and L. J. Rasmussen, *DNA Repair (Amst)*, 2009, **8**, 682-689.
13. P. B. Moens, *Genome*, 2003, **46**, 936-937.
14. J. Y. Chin, E. B. Schleifman and P. M. Glazer, *Front Biosci*, 2007, **12**, 4288-4297.
15. Y. Murakami, *Tanpakushitsu Kakusan Koso*, 2003, **48**, 698-699.
16. H. Ogawa, *Tanpakushitsu Kakusan Koso*, 2003, **48**, 678-680.
17. G. M. Blackburn, *Nucleic acids in chemistry and biology*, RSC Pub., Cambridge, UK, 2006.
18. N. R. Cozzarelli and J. C. Wang, *DNA topology and its biological effects*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1990.
19. J. D. Watson and F. H. Crick, *Nature*, 2003, **421**, 397-398; discussion 396.
20. J. D. Watson and F. H. Crick, *Nature*, 1953, **171**, 737-738.
21. D. M. Gowers and K. R. Fox, *Nucleic Acids Res*, 1999, **27**, 1569-1577.
22. R. Wing, H. Drew, T. Takano, C. Broka, S. Tanaka, K. Itakura and R. E. Dickerson, *Nature*, 1980, **287**, 755-758.
23. R. E. Franklin and R. G. Gosling, *Nature*, 1953, **171**, 740-741.
24. Z. Shakked, D. Rabinovich, W. B. Cruse, E. Egert, O. Kennard, G. Sala, S. A. Salisbury and M. A. Viswamitra, *Proc R Soc Lond B Biol Sci*, 1981, **213**, 479-487.
25. J. Klysik, S. M. Stirdivant, J. E. Larson, P. A. Hart and R. D. Wells, *Nature*, 1981, **290**, 672-677.
26. C. K. Singleton, J. Klysik, S. M. Stirdivant and R. D. Wells, *Nature*, 1982, **299**, 312-316.
27. L. J. Peck, A. Nordheim, A. Rich and J. C. Wang, *Proc Natl Acad Sci U S A*, 1982, **79**, 4560-4564.
28. J. C. Wang, L. J. Peck and K. Becherer, *Cold Spring Harb Symp Quant Biol*, 1983, **47 Pt 1**, 85-91.
29. A. H. J. Wang, G. J. Quigley, F. J. Kolpak, J. L. Crawford, J. H. Vanboom, G. Vandermarel and A. Rich, *Nature*, 1979, **282**, 680-686.
30. H. Drew, T. Takano, S. Tanaka, K. Itakura and R. E. Dickerson, *Nature*, 1980, **286**, 567-573.
31. F. M. Pohl and T. M. Jovin, *J Mol Biol*, 1972, **67**, 375-396.
32. R. Liu, H. Liu, X. Chen, M. Kirby, P. O. Brown and K. Zhao, *Cell*, 2001, **106**, 309-318.

33. A. V. Vologodskii and N. R. Cozzarelli, *Annu Rev Biophys Biomol Struct*, 1994, **23**, 609-643.
34. D. A. Koster, A. Crut, S. Shuman, M. A. Bjornsti and N. H. Dekker, *Cell*, 2010, **142**, 519-530.
35. T. Furuyama and S. Henikoff, *Cell*, 2009, **138**, 104-113.
36. G. Witz and A. Stasiak, *Nucleic Acids Res*, 2010, **38**, 2119-2133.
37. C. J. Dorman, *Sci Prog*, 2006, **89**, 151-166.
38. S. Deng, R. A. Stein and N. P. Higgins, *Mol Microbiol*, 2005, **57**, 1511-1521.
39. H. Wang and C. J. Benham, *Plos Comput Biol*, 2008, **4**.
40. T. C. Boles, J. H. White and N. R. Cozzarelli, *J Mol Biol*, 1990, **213**, 931-951.
41. A. V. Vologodskii, *Topology and physics of circular DNA*, CRC Press, Boca Raton, 1992.
42. J. E. Deweese, M. A. Osheroff and N. Osheroff, *Biochem Mol Biol Educ*, 2008, **37**, 2-10.
43. R. Dulbecco and M. Vogt, *Proc Natl Acad Sci U S A*, 1963, **50**, 236-243.
44. J. C. Wang, *Nat Rev Mol Cell Biol*, 2002, **3**, 430-440.
45. T. S. Hsieh and J. C. Wang, *Biochemistry-Us*, 1975, **14**, 527-535.
46. J. C. Wang, *Annu Rev Biochem*, 1996, **65**, 635-692.
47. S. C. Kampranis, A. D. Bates and A. Maxwell, *P Natl Acad Sci USA*, 1999, **96**, 8414-8419.
48. F. B. Fuller, *Proc Natl Acad Sci U S A*, 1971, **68**, 815-819.
49. C. Benham and D. Miller, *J Knot Theor Ramif*, 2000, **9**, 577-585.
50. J. H. White, *Am J Math*, 1969, **91**, 693-&.
51. W. F. Pohl, *J Math Mech*, 1968, **17**, 975-&.
52. F. B. Fuller, *Proc Natl Acad Sci U S A*, 1978, **75**, 3557-3561.
53. F. H. C. Crick, *P Natl Acad Sci USA*, 1976, **73**, 2639-2643.
54. D. Miller and C. Benham, *J Knot Theor Ramif*, 1996, **5**, 859-866.
55. A. V. Vologodskii, S. D. Levene, K. V. Klenin, M. Frank-Kamenetskii and N. R. Cozzarelli, *J Mol Biol*, 1992, **227**, 1224-1243.
56. L. A. Britton, W. K. Olson and I. Tobias, *Journal of Chemical Physics*, 2009, **131**.
57. J. M. Fogg, N. Kolmakova, I. Rees, S. Magonov, H. Hansma, J. J. Perona and E. L. Zechiedrich, *J Phys-Condens Mat*, 2006, **18**, S145-S159.
58. Y. Zhang, A. E. McEwen, D. M. Crothers and S. D. Levene, *Biophys J*, 2006, **90**, 1903-1912.
59. R. E. Depew and J. C. Wang, *P Natl Acad Sci USA*, 1975, **72**, 4275-4279.
60. J. Bednar, P. Furrer, A. Stasiak, J. Dubochet, E. H. Egelman and A. D. Bates, *Journal of Molecular Biology*, 1994, **235**, 825-847.
61. D. W. Li, Z. Q. Yang, G. J. Zhao, Y. Long, B. Lv, C. Li, S. Hiew, M. T. T. Ng, J. J. Guo, H. Tan, H. Zhang and T. H. Li, *Chem Commun*, 2011, **47**, 7479-7481.
62. D. W. Li, Z. Q. Yang, Y. Long, G. Zhao, B. Lv, S. Hiew, T. T. N. Magdeline, J. J. Guo, H. Tan, H. Zhang, W. X. Yuan, H. B. Su and T. H. Li, *Chem Commun*, 2011, **47**, 10695-10697.
63. J. C. Wang, *Adv Pharmacol*, 1994, **29B**, 257-270.
64. M. Napierala, A. Bacolla and R. D. Wells, *J Biol Chem*, 2005, **280**, 37366-37376.
65. J. H. Crosa, L. K. Luttrupp and S. Falkow, *Nature*, 1976, **261**, 516-519.
66. P. R. Cook and I. A. Brazell, *Eur J Biochem*, 1977, **74**, 527-531.
67. K. Kran, *Arch Mikrobiol*, 1962, **44**, 1-22.
68. M. Adrian, B. Tenheggelerbordier, W. Wahli, A. Z. Stasiak, A. Stasiak and J. Dubochet, *Embo Journal*, 1990, **9**, 4551-4554.
69. H. M. Wu and D. M. Crothers, *Nature*, 1984, **308**, 509-513.
70. P. J. Hagerman, *Biochemistry-Us*, 1985, **24**, 7033-7037.

71. C. H. Laundon and J. D. Griffith, *Cell*, 1988, **52**, 545-549.
72. E. N. Trifonov and J. L. Sussman, *Proc Natl Acad Sci U S A*, 1980, **77**, 3816-3820.
73. L. Ulanovsky, M. Bodner, E. N. Trifonov and M. Choder, *P Natl Acad Sci USA*, 1986, **83**, 862-866.
74. D. M. Crothers, T. E. Haran and J. G. Nadeau, *J Biol Chem*, 1990, **265**, 7093-7096.
75. H. S. Koo, H. M. Wu and D. M. Crothers, *Nature*, 1986, **320**, 501-506.
76. L. Simpson, *Annu Rev Microbiol*, 1987, **41**, 363-382.
77. D. W. Li, Z. Q. Yang, B. Lv and T. H. Li, *Bioorg Med Chem Lett*, 2012, **22**, 833-836.
78. Q. Du, A. Kotlyar and A. Vologodskii, *Nucleic Acids Res*, 2008, **36**, 1120-1128.
79. F. Lankas, R. Lavery and J. H. Maddocks, *Structure*, 2006, **14**, 1527-1534.
80. Z. Q. Yang, D. W. Li, S. H. Hiew, M. T. Ng, W. X. Yuan, H. B. Su, F. W. Shao and T. H. Li, *Chem Commun*, 2011, **47**, 11309-11311.
81. P. Forte, L. Leoni, B. Sampaiolese and M. Savino, *Nucleic Acids Research*, 1989, **17**, 8683-8694.
82. K. Hizume, S. Araki, K. Yoshikawa and K. Takeyasu, *Nucleic Acids Research*, 2007, **35**, 2787-2799.
83. K. Luger, A. W. Mader, R. K. Richmond, D. F. Sargent and T. J. Richmond, *Nature*, 1997, **389**, 251-260.
84. L. C. Welsh, D. A. Marvin and R. N. Perham, *J Mol Biol*, 1998, **284**, 1265-1271.
85. D. J. Patel, L. Shapiro and D. Hare, *Annu Rev Biophys Biophys Chem*, 1987, **16**, 423-454.
86. S. Campagne, V. Gervais and A. Milon, *J R Soc Interface*, 2011, **8**, 1065-1078.
87. P. S. Belton, *Cell Mol Life Sci*, 2000, **57**, 993-998.
88. A. Bacolla, M. Wojciechowska, B. Kosmider, J. E. Larson and R. D. Wells, *DNA Repair (Amst)*, 2006, **5**, 1161-1170.
89. S. C. Raghavan and M. R. Lieber, *Cell Cycle*, 2004, **3**, 762-768.
90. A. Dayn, S. Malkhosyan and S. M. Mirkin, *Nucleic Acids Res*, 1992, **20**, 5991-5997.
91. T. Allers and D. R. Leach, *J Mol Biol*, 1995, **252**, 70-85.
92. V. Brazda, R. C. Laister, E. B. Jagelska and C. Arrowsmith, *Bmc Mol Biol*, 2011, **12**, 33.
93. M. Zannis-Hadjopoulos, W. Yahyaoui and M. Callejo, *Trends Biochem Sci*, 2008, **33**, 44-50.
94. O. Novac, D. Alvarez, C. E. Pearson, G. B. Price and M. Zannis-Hadjopoulos, *J Biol Chem*, 2002, **277**, 11174-11183.
95. D. M. J. Lilley, *P Natl Acad Sci-Biol*, 1980, **77**, 6468-6472.
96. E. Caffarelli, C. Franzini, L. Leoni and M. Savino, *Cell Biophys*, 1984, **6**, 23-31.
97. D. M. J. Lilley, *Chemical Society Reviews*, 1989, **18**, 53-83.
98. A. I. H. Murchie and D. M. J. Lilley, *Method Enzymol*, 1992, **211**, 158-180.
99. D. M. J. Lilley, G. W. Gough, L. R. Hallam and K. M. Sullivan, *Biochimie*, 1985, **67**, 697-706.
100. M. Y. Vologodskaya and A. V. Vologodskii, *J Mol Biol*, 1999, **289**, 851-859.
101. J. L. Huppert, *Chem Soc Rev*, 2008, **37**, 1375-1384.
102. S. Neidle and S. Balasubramanian, *Quadruplex nucleic acids*, RSC Pub., Cambridge, 2006.
103. T. Simonsson, *Biol Chem*, 2001, **382**, 621-628.
104. V. Dapic, V. Abdomerovic, R. Marrington, J. Peberdy, A. Rodger, J. O. Trent and P. J. Bates, *Nucleic Acids Res*, 2003, **31**, 2097-2107.
105. J. T. Davis, *Angew Chem Int Ed Engl*, 2004, **43**, 668-698.
106. S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd and S. Neidle, *Nucleic Acids Res*, 2006, **34**, 5402-5415.
107. S. Neidle and G. N. Parkinson, *Biochimie*, 2008, **90**, 1184-1196.

108. E. Henderson, C. C. Hardin, S. K. Walk, I. Tinoco, Jr. and E. H. Blackburn, *Cell*, 1987, **51**, 899-908.
109. R. F. Macaya, P. Schultze, F. W. Smith, J. A. Roe and J. Feigon, *Proc Natl Acad Sci U S A*, 1993, **90**, 3745-3749.
110. W. I. Sundquist and A. Klug, *Nature*, 1989, **342**, 825-829.
111. M. A. Keniry, G. D. Strahan, E. A. Owen and R. H. Shafer, *European Journal of Biochemistry*, 1995, **233**, 631-643.
112. D. Sen and W. Gilbert, *Nature*, 1988, **334**, 364-366.
113. G. Laughlan, A. I. Murchie, D. G. Norman, M. H. Moore, P. C. Moody, D. M. Lilley and B. Luisi, *Science*, 1994, **265**, 520-524.
114. C. C. Hardin, E. Henderson, T. Watson and J. K. Prosser, *Biochemistry-Us*, 1991, **30**, 4460-4472.
115. P. Balagurumoorthy, S. K. Brahmachari, D. Mohanty, M. Bansal and V. Sasisekharan, *Nucleic Acids Res*, 1992, **20**, 4061-4067.
116. R. Jin, B. L. Gaffney, C. Wang, R. A. Jones and K. J. Breslauer, *Proc Natl Acad Sci U S A*, 1992, **89**, 8832-8836.
117. M. C. Shiber, E. H. Braswell, H. Klump and J. R. Fresco, *Nucleic Acids Research*, 1996, **24**, 5004-5012.
118. S. Neidle and G. N. Parkinson, *Curr Opin Struct Biol*, 2003, **13**, 275-283.
119. T. Miyoshi, J. Kanoh, M. Saito and F. Ishikawa, *Science*, 2008, **320**, 1341-1344.
120. M. J. McEachern, A. Krauskopf and E. H. Blackburn, *Annu Rev Genet*, 2000, **34**, 331-358.
121. Y. S. Cong, W. E. Wright and J. W. Shay, *Microbiol Mol Biol Rev*, 2002, **66**, 407-425, table of contents.
122. S. Bassham, A. Beam and J. Champay, *Mol Cell Biol*, 1998, **18**, 269-275.
123. M. E. Lee, S. Y. Rha, H. C. Jeung, T. S. Kim, H. C. Chung and B. K. Oh, *Cancer Lett*, 2008, **264**, 107-118.
124. D. Gomez, R. Paterski, T. Lemarteleur, K. Shin-Ya, J. L. Mergny and J. F. Riou, *J Biol Chem*, 2004, **279**, 41487-41494.
125. J. L. Mergny, J. F. Riou, P. Mailliet, M. P. Teulade-Fichou and E. Gilson, *Nucleic Acids Res*, 2002, **30**, 839-865.
126. A. Arola and R. Vilar, *Curr Top Med Chem*, 2008, **8**, 1405-1415.
127. L. Oganessian and T. M. Bryan, *Bioessays*, 2007, **29**, 155-165.
128. D. J. Patel, A. T. Phan and V. Kuryavyi, *Nucleic Acids Res*, 2007, **35**, 7429-7455.
129. E. L. Zechiedrich and N. Osheroff, *EMBO J*, 1990, **9**, 4555-4562.
130. J. C. Wang, *Biochim Biophys Acta*, 1987, **909**, 1-9.
131. V. A. Bloomfield, D. M. Crothers and I. Tinoco, *Nucleic acids : structures, properties, and functions*, University Science Books, Sausalito, Calif., 2000.
132. K. Hu, *Bull Math Biol*, 2007, **69**, 1019-1030.
133. F. Gimenes, K. I. Takeda, A. Fiorini, F. S. Gouveia and M. A. Fernandez, *Genet Mol Res*, 2008, **7**, 549-558.
134. Y. Wada-Kiyama and R. Kiyama, *DNA Res*, 1996, **3**, 25-30.
135. K. Zahn and F. R. Blattner, *Science*, 1987, **236**, 416-422.
136. J. F. Marko and E. D. Siggia, *Science*, 1994, **265**, 506-508.
137. M. Snyder, A. R. Buchman and R. W. Davis, *Nature*, 1986, **324**, 87-89.
138. F. B. Dean, J. A. Borowiec, Y. Ishimi, S. Deb, P. Tegtmeier and J. Hurwitz, *Proc Natl Acad Sci U S A*, 1987, **84**, 8267-8271.
139. E. Remaut, H. Tsao and W. Fiers, *Gene*, 1983, **22**, 103-113.
140. G. J. Lohman, L. Chen and T. C. Evans, Jr., *J Biol Chem*, 2011, **286**, 44187-44196.
141. X. Liang, K. Fujioka and H. Asanuma, *Chemistry*, 2011, **17**, 10388-10396.

142. H. B. Gray, Jr., D. A. Ostrander, J. L. Hodnett, R. J. Legerski and D. L. Robberson, *Nucleic Acids Res*, 1975, **2**, 1459-1492.
143. R. J. Legerski, J. L. Hodnett and H. B. Gray, Jr., *Nucleic Acids Res*, 1978, **5**, 1445-1464.
144. C. F. Wei, G. A. Alianell, G. H. Bencen and H. B. Gray, Jr., *J Biol Chem*, 1983, **258**, 13506-13512.
145. M. T. Muller, J. R. Spitzner, J. A. DiDonato, V. B. Mehta and K. Tsutsui, *Biochemistry-Us*, 1988, **27**, 8369-8379.
146. J. R. Spitzner and M. T. Muller, *Nucleic Acids Res*, 1988, **16**, 5533-5556.
147. J. C. Wang and K. Kirkegaard, *Gene Amplif Anal*, 1981, **2**, 455-473.
148. E. Cretaio, L. Pattarello, Y. Fontebasso, P. Benedetti and C. Losasso, *Ital J Biochem*, 2007, **56**, 91-102.
149. J. B. Leppard and J. J. Champoux, *Chromosoma*, 2005, **114**, 75-85.
150. Y. L. Lyubchenko, L. S. Shlyakhtenko, T. Aki and S. Adhya, *Nucleic Acids Res*, 1997, **25**, 873-876.
151. Y. L. Lyubchenko, A. A. Gall, L. S. Shlyakhtenko, R. E. Harrington, B. L. Jacobs, P. I. Oden and S. M. Lindsay, *J Biomol Struct Dyn*, 1992, **10**, 589-606.
152. L. S. Shlyakhtenko, L. Miloseska, V. N. Potaman, R. R. Sinden and Y. L. Lyubchenko, *Ultramicroscopy*, 2003, **97**, 263-270.
153. Y. L. Lyubchenko, P. I. Oden, D. Lampner, S. M. Lindsay and K. A. Dunker, *Nucleic Acids Res*, 1993, **21**, 1117-1123.
154. Y. L. Lyubchenko and L. S. Shlyakhtenko, *Proc Natl Acad Sci U S A*, 1997, **94**, 496-501.
155. Y. L. Lyubchenko and L. S. Shlyakhtenko, *Methods*, 2009, **47**, 206-213.
156. Y. L. Lyubchenko, L. S. Shlyakhtenko and A. A. Gall, *Methods Mol Biol*, 2009, **543**, 337-351.
157. L. Hamon, D. Pastre, P. Dupaigne, C. Le Breton, E. Le Cam and O. Pietrement, *Nucleic Acids Res*, 2007, **35**, e58.
158. J. Zheng, Z. Li, A. Wu and H. Zhou, *Biophys Chem*, 2003, **104**, 37-43.
159. P. R. Dahlgren and Y. L. Lyubchenko, *Biochemistry-Us*, 2002, **41**, 11372-11378.
160. B. Suter, G. Schnappauf and F. Thoma, *Nucleic Acids Research*, 2000, **28**, 4083-4089.
161. R. Ohki, M. Hirota, M. Oishi and R. Kiyama, *Nucleic Acids Res*, 1998, **26**, 3026-3033.
162. J. Zlatanova and S. H. Leuba, *Chromatin structure and dynamics : state-of-the-art*, Elsevier, Amsterdam London, 2004.
163. A. Matsumoto, I. Tobias and W. K. Olson, *J Chem Theory Comput*, 2005, **1**, 130-142.
164. V. Y. Katritch and A. V. Vologodskii, *Biofizika*, 1996, **41**, 536-538.
165. M. Vologodskaia and A. Vologodskii, *J Mol Biol*, 2002, **317**, 205-213.
166. J. Gore, Z. Bryant, M. D. Stone, M. Nollmann, N. R. Cozzarelli and C. Bustamante, *Nature*, 2006, **439**, 100-104.
167. A. C. Declais, C. B. de La Tour and M. Duguet, *Methods Enzymol*, 2001, **334**, 146-162.
168. A. C. Rodriguez and D. Stock, *EMBO J*, 2002, **21**, 418-426.
169. D. A. Collier and R. D. Wells, *J Biol Chem*, 1990, **265**, 10652-10658.
170. N. Figueroa-Bossi, M. Guerin, R. Rahmouni, M. Leng and L. Bossi, *EMBO J*, 1998, **17**, 2359-2367.
171. P. E. Nielsen, *Pure Appl Chem*, 1998, **70**, 105-110.
172. E. Uhlmann, A. Peyman, G. Breipohl and D. W. Will, *Angew Chem Int Edit*, 1998, **37**, 2797-2823.
173. P. E. Nielsen, *Accounts Chem Res*, 1999, **32**, 624-630.
174. P. E. Nielsen, M. Egholm, R. H. Berg and O. Buchardt, *Science*, 1991, **254**, 1497-1500.

175. M. Egholm, O. Buchardt, P. E. Nielsen and R. H. Berg, *Journal of the American Chemical Society*, 1992, **114**, 1895-1897.
176. A. Ray and B. Norden, *Faseb Journal*, 2000, **14**, 1041-1060.
177. H. Kuhn, V. V. Demidov, P. E. Nielsen and M. D. Frank-Kamenetskii, *J Mol Biol*, 1999, **286**, 1337-1345.
178. N. O. Bukanov, V. V. Demidov, P. E. Nielsen and M. D. Frank-Kamenetskii, *Proc Natl Acad Sci U S A*, 1998, **95**, 5516-5520.
179. H. Kuhn, V. V. Demidov, M. D. Frank-Kamenetskii and P. E. Nielsen, *Nucleic Acids Res*, 1998, **26**, 582-587.
180. K. Kaihatsu, R. H. Shah, X. Zhao and D. R. Corey, *Biochemistry-Us*, 2003, **42**, 13996-14003.
181. V. V. Demidov, H. Kuhn, I. V. Lavrentieva-Smolina and M. D. Frank-Kamenetskii, *Methods*, 2001, **23**, 123-131.
182. J. Lohse, O. Dahl and P. E. Nielsen, *Proc Natl Acad Sci U S A*, 1999, **96**, 11804-11808.
183. V. V. Demidov and M. D. Frank-Kamenetskii, *Methods*, 2001, **23**, 108-122.
184. B. Liu, Y. Han, D. R. Corey and T. Kodadek, *J Am Chem Soc*, 2002, **124**, 1838-1839.
185. J. Hu and D. R. Corey, *Biochemistry-Us*, 2007, **46**, 7581-7589.
186. F. Pellestor, P. Paulasova and S. Hamamah, *Curr Pharm Design*, 2008, **14**, 2439-2444.
187. A. F. Saleh, A. Arzumanov, R. Abes, D. Owen, B. Lebleu and M. J. Gait, *Bioconjugate Chem*, 2010, **21**, 1902-1911.
188. S. El-Andaloussi, H. J. Johansson, P. Lundberg and U. Langel, *J Gene Med*, 2006, **8**, 1262-1273.
189. S. Lusvarghi, C. T. Murphy, S. Roy, F. A. Tanious, I. Sacui, W. D. Wilson, D. H. Ly and B. A. Armitage, *Journal of the American Chemical Society*, 2009, **131**, 18415-18424.
190. S. Abes, G. D. Ivanova, R. Abes, A. A. Arzumanov, D. Williams, D. Owen, B. Lebleu and M. J. Gait, *Methods Mol Biol*, 2009, **480**, 85-99.
191. T. Ratilainen and B. Norden, *Methods Mol Biol*, 2002, **208**, 59-88.
192. V. V. Demidov, M. V. Yavnilovich, B. P. Belotserkovskii, M. D. Frank-Kamenetskii and P. E. Nielsen, *Proc Natl Acad Sci U S A*, 1995, **92**, 2637-2641.
193. A. Kurakin, H. J. Larsen and P. E. Nielsen, *Chem Biol*, 1998, **5**, 81-89.
194. X. H. Peng, H. Li and M. Seidman, *Eur J Org Chem*, 2010, 4194-4197.
195. E. Wickstrom, M. Choob, K. A. Urtishak, X. Tian, N. Sternheim, S. Talbot, J. Archdeacon, V. A. Efimov and S. A. Farber, *Journal of Drug Targeting*, 2004, **12**, 363-372.
196. K. E. Lundin, L. Good, R. Stromberg, A. Graslund and C. I. Smith, *Adv Genet*, 2006, **56**, 1-51.
197. J. H. M. Cabral, A. P. Jackson, C. V. Smith, N. Shikotra, A. Maxwell and R. C. Liddington, *Nature*, 1997, **388**, 903-906.
198. J. G. Heddle, S. Mittelheiser, A. Maxwell and N. H. Thomson, *J Mol Biol*, 2004, **337**, 597-610.
199. A. M. MacMillan, *Pure Appl Chem*, 2004, **76**, 1521-1524.
200. M. P. Stone, H. Huang, K. L. Brown and G. Shanmugam, *Chem Biodivers*, 2011, **8**, 1571-1615.
201. C. J. Lord and A. Ashworth, *Nature*, 2012, **481**, 287-294.
202. T. Douki and J. Cadet, *Biochemistry-Us*, 2001, **40**, 2495-2501.
203. R. P. Sinha and D. P. Hader, *Photochem Photobiol Sci*, 2002, **1**, 225-236.
204. J. Cadet, M. Berger, T. Douki, B. Morin, S. Raoul, J. L. Ravanat and S. Spinelli, *Biol Chem*, 1997, **378**, 1275-1286.

205. D. W. Brown, L. J. Libertini, C. Suquet, E. W. Small and M. J. Smerdon, *Ann N Y Acad Sci*, 1994, **726**, 292-294.
206. C. Richter, *Mutat Res*, 1992, **275**, 249-255.
207. E. M. Hammond, M. R. Kaufmann and A. J. Giaccia, *Curr Opin Cell Biol*, 2007, **19**, 680-684.
208. R. A. Costa, C. D. Romagna, J. L. Pereira and N. C. Souza-Pinto, *J Bioenerg Biomembr*, 2011, **43**, 25-29.
209. C. Bertram and R. Hass, *Biol Chem*, 2008, **389**, 211-220.
210. T. Ploszaj, A. Robaszkiewicz and H. Witas, *Postepy Biochem*, 2010, **56**, 139-146.
211. P. Mao and P. H. Reddy, *Biochim Biophys Acta*, 2011, **1812**, 1359-1370.
212. P. A. Nikitin and M. A. Luftig, *Br J Cancer*, 2012, **106**, 429-435.
213. G. Pratviel, *Met Ions Life Sci*, 2012, **10**, 201-216.
214. H. C. Reinhardt and B. Schumacher, *Trends Genet*, 2012, **28**, 128-136.
215. D. Fu, J. A. Calvo and L. D. Samson, *Nat Rev Cancer*, 2012, **12**, 104-120.
216. F. R. de Gruijl, H. J. van Kranen and L. H. Mullenders, *J Photochem Photobiol B*, 2001, **63**, 19-27.
217. H. Werbin, *Photochem Photobiol*, 1977, **26**, 675-678.
218. A. Yasui and S. J. McCready, *Bioessays*, 1998, **20**, 291-297.
219. A. Sancar, C. Thompson, R. J. Thresher, F. Araujo, J. Mo, S. Ozgur, E. Vagas, L. Dawut and C. P. Selby, *Cold Spring Harb Symp Quant Biol*, 2000, **65**, 157-171.
220. F. Thoma, *EMBO J*, 1999, **18**, 6585-6598.
221. I. Alseth, H. Korvald, F. Osman, E. Seeberg and M. Bjoras, *Nucleic Acids Res*, 2004, **32**, 5119-5125.
222. J. Tuo, M. Muftuoglu, C. Chen, P. Jaruga, R. R. Selzer, R. M. Brosh, Jr., H. Rodriguez, M. Dizdaroglu and V. A. Bohr, *J Biol Chem*, 2001, **276**, 45772-45779.
223. J. T. Reardon, L. H. Thompson and A. Sancar, *Nucleic Acids Res*, 1997, **25**, 1015-1021.
224. J. W. Shay and W. E. Wright, *Mol Cell*, 2004, **14**, 420-421.
225. B. Kaina, *Biochem Pharmacol*, 2003, **66**, 1547-1554.
226. M. Frankenberg-Schwager and D. Frankenberg, *Radiat Res*, 1994, **138**, S97-100.
227. E. Mladenov and G. Iliakis, *Mutat Res*, 2011, **711**, 61-72.
228. T. Poplawski, E. Stoczynska and J. Blasiak, *Postepy Biochem*, 2009, **55**, 36-45.
229. T. E. Wilson, L. M. Topper and P. L. Palmbo, *Trends Biochem Sci*, 2003, **28**, 62-66.
230. Y. Liu, L. Deng, S. C. Nguyen, C. L. Au, C. Shao, J. A. Tischfield and L. Liang, *Mutat Res*, 2012, **731**, 140-144.
231. L. Liang, L. Deng, S. C. Nguyen, X. Zhao, C. D. Maulion, C. Shao and J. A. Tischfield, *Nucleic Acids Res*, 2008, **36**, 3297-3310.
232. G. Bordeianu, F. Zugun-Eloae and M. G. Rusu, *Rev Med Chir Soc Med Nat Iasi*, 2011, **115**, 1189-1194.
233. L. A. Cahoon and H. S. Seifert, *Mol Microbiol*, 2011, **81**, 1136-1143.
234. W. D. Heyer, K. T. Ehmsen and J. Liu, *Annu Rev Genet*, 2010, **44**, 113-139.
235. M. E. Lomax, S. Cunniffe and P. O'Neill, *Biochemistry-Us*, 2004, **43**, 11017-11026.
236. E. Haghazari and W. D. Heyer, *Nucleic Acids Res*, 2004, **32**, 4257-4268.
237. N. Panayotatos and R. D. Wells, *Nature*, 1981, **289**, 466-470.
238. S. M. Mirkin and M. D. Frank-Kamenetskii, *Annu Rev Biophys Biomol Struct*, 1994, **23**, 541-576.
239. H. Htun and J. E. Dahlberg, *Science*, 1989, **243**, 1571-1576.
240. R. R. Sinden, M. J. Pytlos-Sinden and V. N. Potaman, *Front Biosci*, 2007, **12**, 4788-4799.
241. G. Levinson and G. A. Gutman, *Mol Biol Evol*, 1987, **4**, 203-221.
242. A. Bacolla and R. D. Wells, *J Biol Chem*, 2004, **279**, 47411-47414.

243. R. D. Wells, *J Biol Chem*, 2009, **284**, 8997-9009.
244. H. J. Lipps and D. Rhodes, *Trends Cell Biol*, 2009, **19**, 414-422.
245. L. H. Hurley, D. D. Von Hoff, A. Siddiqui-Jain and D. Yang, *Semin Oncol*, 2006, **33**, 498-512.
246. A. Siddiqui-Jain, C. L. Grand, D. J. Bearss and L. H. Hurley, *Proc Natl Acad Sci U S A*, 2002, **99**, 11593-11598.
247. J. L. Huppert and S. Balasubramanian, *Nucleic Acids Res*, 2005, **33**, 2908-2916.
248. P. L. Tran, J. L. Mergny and P. Alberti, *Nucleic Acids Res*, 2011, **39**, 3282-3294.
249. F. O. Losch, A. Bredenbeck, V. M. Hollstein, P. Walden and P. Wrede, *Human Genetics*, 2007, **122**, 337-343.
250. H. Yoshimura, T. Yoshino, T. Hirose, Y. Nakamura, M. Higashi, T. Hase, K. Yamaguchi, H. Hirokawa and Y. Masamune, *J Gen Appl Microbiol*, 1986, **32**, 393-404.
251. M. Ivanchenko, J. Zlatanova and K. van Holde, *Biophys J*, 1997, **72**, 1388-1395.
252. M. de la Cruz Munoz-Centeno, G. Millan-Zambrano and S. Chavez, *Methods Mol Biol*, 2012, **824**, 51-64.
253. Z. Y. Kan, Y. Lin, F. Wang, X. Y. Zhuang, Y. Zhao, D. W. Pang, Y. H. Hao and Z. Tan, *Nucleic Acids Res*, 2007, **35**, 3646-3653.
254. J. Zhou, C. Wei, G. Jia, X. Wang, Z. Feng and C. Li, *Biochimie*, 2009, **91**, 1104-1111.
255. D. Miyoshi, S. Matsumura, S. Nakano and N. Sugimoto, *J Am Chem Soc*, 2004, **126**, 165-169.
256. D. Miyoshi, H. Karimata and N. Sugimoto, *J Am Chem Soc*, 2006, **128**, 7957-7963.
257. S. A. Wasserman and N. R. Cozzarelli, *Proc Natl Acad Sci U S A*, 1985, **82**, 1079-1083.
258. C. Bailly and J. P. Henichart, *Bioconjug Chem*, 1991, **2**, 379-393.
259. E. Soutoglou and T. Misteli, *Science*, 2008, **320**, 1507-1510.
260. A. Bacolla and R. D. Wells, *Mol Carcinog*, 2009, **48**, 273-285.
261. I. Voineagu, C. H. Freudenreich and S. M. Mirkin, *Mol Carcinog*, 2009, **48**, 309-318.

Li Dawei, Ph.D. (candidate)
Division of Chemistry and Biological Chemistry
School of Physical & Mathematical Sciences
Nanyang Technological University
21 Nanyang Link
Tel.: (65) 9613-3546
Email: lida0006@e.ntu.edu.sg

EDUCATION:

- 2009 - present Graduate Student (Ph.D. degree will be awarded in 2013)
Chemical Biology
Division of Chemistry and Biological Chemistry
School of Physical & Mathematical Sciences
Nanyang Technological University
- 2003 - 2006 Graduate Student, Master of Science
Organic Synthesis
Department of Chemistry
Lanzhou University
- 1999 - 2003 Undergraduate Student, Bachelor of Science
Chemistry
Department of Chemistry
Lanzhou University

PROFESSIONAL EXPERIENCE:

- 2006 - 2009 Advanced Medicinal Chemistry Researcher
Department of Medicinal Chemistry
Shanghai Chemexplorer Co., Ltd.

RESEARCH INTERESTS:

DNA topology and Topoisomerase
DNA nanotechnology
Medicinal Chemistry (Drug discovery, Building library for drug candidate)

PUBLICATION:

1. **Li, D. W.;** Lv, B.; Zhang, H.; Li, Y. Q. J; Li, T. H., Positive supercoiling affiliated with nucleosome repairs non-B structures of DNA. Submitted.
2. **Li, D. W.;** Lv, B.; Zhang, H.; Li, Y. Q. J; Li, T. H., Gyrase-assisted formation of G-quadruplex from duplex DNA. Submitted.

3. **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 (2), 833-836.
4. Yang, Z. Q.; **Li, D. W.**; Guo, J. J.; Shao, F. W.; Li, T. H., Intrinsic curvature in duplex DNA inhibits Human Topoisomerase I. *Bioorg Med Chem Lett* **2012**, 22 (3), 1322-1325.
5. Tan, H. K.; **Li, D. W.**; Gray, R. K.; Yang, Z. Q.; Ng, M. T. T.; Zhang, H.; Tan, J. M. R.; Hiew, S. H.; Lee, J. Y.; Li, T. H., Interference of intrinsic curvature of DNA by DNA-intercalating agents. *Organic & Biomolecular Chemistry* **2012**, 10 (11), 2227-2230.
6. Xu, W.; Xie, X. J.; **Li, D. W.**; Yang, Z. Q.; Li, T. H.; Liu, X. G., Ultrasensitive Colorimetric DNA Detection using a Combination of Rolling Circle Amplification and Nicking Endonuclease-Assisted Nanoparticle Amplification (NEANA). *Small* **2012**, DOI: 10.1002/smll.201200263
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 (38), 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 (26), 7479-7481.
9. Yang, Z. Q.; **Li, D. W.**; Hiew, S. H.; Ng, M. T.; Yuan, W. X.; Su, H. B.; Shao, F. W.; Li, T. H., Recognition of forcible curvature in circular DNA by human topoisomerase I. *Chem Commun* **2011**, 47 (40), 11309-11311.
10. Yang, Z. Q.; **Li, D. W.**; Li, T. H., Design and synthesis of catenated rings based on toroidal DNA structures. *Chem Commun* **2011**, 47 (43), 11930-11932.
11. Wang, H. B.; Xu, W.; Zhang, H.; **Li, D. W.**; Yang, Z. Q.; Xie, X. J.; Li, T. H.; Liu, X. G., EcoRI-Modified Gold Nanoparticles for Dual-Mode Colorimetric Detection of Magnesium and Pyrophosphate Ions. *Small* **2011**, 7 (14), 1987-1992.
12. **Li, D. W.**; Li, W. L.; Wang, Q. A.; Yang, Z. Q.; Hou, Z. J., Concise synthesis of Cannabisin G. *Bioorg Med Chem Lett* **2010**, 20 (17), 5095-5098.
13. Wang, Q.; He, K. K.; Li, Y. Z.; **Li, D. W.**; Li, Y.; Hou, Z. J., Enantioselective synthesis and absolute configuration of the natural threo-3-chloro-1-(4-hydroxy-3-methoxyphenyl)propane -1,2-diol. *J Chem Res* **2004**, 504-505.
14. **Li, D. W.**; Wang, Q.; Hou, Z. J., First Synthesis of Natural Dihydroconiferyl Ferulate. *Chin J Org Chem Suppl* **2004**, 159.
15. Wang, Q., He, K. K.; **Li, D. W.**; Li, Y.; Hou, Z. J., Enantioselective synthesis of the naturally Phenylpropanoid. *Chin J Org Chem Suppl* **2004**, 157.

AWARDS:

1. "Team Milestone Award" for valuable contribution to Eli Lilly and Company.
2. "Certificate of Achievement" for outstanding performance and last contribution on candidate selection. Granted by Eli Lilly and Company