Azize SEVİM

M.S. Thesis In Physics

August - 2008

INVESTIGATION OF ENERGETIC AND MECHANISTIC PATHWAYS IN THE ROTATION OF A NICKED DNA

by

Azize SEVİM

August 2008

INVESTIGATION OF ENERGETIC AND MECHANISTIC PATHWAYS IN THE ROTATION OF A NICKED DNA

by

Azize SEVİM

A thesis submitted to

the Graduate Institute of Sciences and Engineering

of

Fatih University

in partial fulfillment of the requirements for the degree of

Master of Science

in

Physics

August 2008
Istanbul, Turkey

APPROVAL PAGE

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Prof. Dr. Mustafa KUMRU
Head of Department

This is to certify that I have read this thesis and that in my opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Assist. Prof. Dr. Levent SARI
Supervisor

Examining Committee Members

Prof. Dr. Mustafa KUMRU ……..

Assist. Prof. Dr. Levent SARI ……..

Assist. Prof. Dr. Serkan ÇALIŞKAN ……..

It is approved that this thesis has been written in compliance with the formatting rules laid down by the Graduate Institute of Sciences and Engineering.

Assist. Prof. Dr. Nurullah ARSLAN

Director

August 2008

INVESTIGATION OF ENERGETIC AND MECHANISTIC PATHWAYS IN THE ROTATION OF A NICKED DNA

Azize SEVİM

M. S. Thesis - Physics
August 2008

Supervisor: Assist. Prof. Dr. Levent SARI

ABSTRACT

Rotational degree of freedom of a nicked DNA molecule is important in many cellular processes and in DNA-protein interactions. The rotation of such a nicked DNA is especially crucial in the relaxation of supercoiled DNA by human topoisomerase I.

Therefore, the dynamic mechanism of a nicked DNA rotation about the its intact strand has been simulated using Molecular Dynamic (MD) method based on the force-field potentials. In order to see large scale movement of DNA atoms, we employ a novel technique known as the Half Quadratic Biased Molecular Dynamics (HQBMD). Rotations in both directions (clockwise and anti-clockwise, that correspond relaxation of negative and positive supercoils respectively) have been performed. Fifty different simulations have been carried out in the present study. As we rotate the downstream part of the DNA (as suggested by a science paper, see reference 5 and 8), we focused on the structural and energetic changes, especially in the nicked region of the DNA. The results we obtained have been compared with the literature studies on the canonic B-DNA.

As a result of a detail discussion on the data produced out of these fifty different simulations, we have proposed the most favorable axes for the rotation of a nicked DNA molecule. The proposed DNA rotations should yield new perspectives in the relaxation mechanism of a supercoiled DNA within the enzyme, as we found different axis of rotations for positive and negative supercoiled DNAs. Also, we propose here that although

a single anti-clockwise DNA downstream rotation is enough to bring DNA to the canonical form, at least two clockwise DNA rotations are needed to bring DNA to its initial structure. Additional new insights have been observed, as summarized in the conclusion part. The current study presents the first report of the data on a nicked DNA molecule while extensive literature studies exist for only intact DNA.

Keywords: DNA topology, supercoiling, Molecular Dynamics Simulations, HQBMD.

**KESİK BİR DNA MOLEKÜLÜNÜN
DÖNMESİ SIRASINDA İZLEDİĞİ ENERJİK VE MEKANİK YOLLARIN
ARAŞTIRILMASI**

Azize SEVİM

Yüksek Lisans Tezi - Fizik
Ağustos 2008

Tez Yöneticisi: Yrd.Doç. Dr. Levent SARI

ÖZ

Kesik bir DNA molekülünün dönme serbestlik derecesi, pek çok hücresel işlemlerde ve protein-DNA etkileşimlerinde önemlidir. Böyle kesik bir DNA'nın dönmesi, özellikle DNA'nın human topoizomeraz I ile relaksasyonunda çok önemlidir.

Bu nedenle, force-field potansiyelleri üzerine kurulu Moleküler Dinamik metodlar uygulanarak DNA'nın kesik olmayan zinciri etrafında dönmesi esnasındaki dinamik mekanizma simule edildi.. DNA atomlarının büyük ölçekli hareketlerini anlayabilmek için, Half Quadratic Biased Molecular Dynamics (HQBMD) olarak bilinen yeni bir teknik kullanıldı. Her iki yönde (negatif ve pozitif süpercoili azaltmaya karşılık gelen, saat yönü ve saat yönünün tersine) dönmeler gerçekleştirildi. Bu çalışmada, elli farklı simulasyon gerçekleştirildi. DNA'nın alt kısmını döndürerek (bir science makalesinde ileri sürüldüğü gibi, referans 5 ve 8'e bakın), özellikle DNA'nın kesik kısmındaki yapısal ve enerjik değişikliklere odaklandık. Elde ettiğimiz sonuçlar literatürdeki kanonik B-DNA üzerine yapılan çalışmalar ile karşılaştırıldı.

Bu elli farklı simulasyon sonucu elde edilen dataların detaylı yorumları sonucu, kesik bir DNA molekülünün dönmesi için en uygun eksenler önerildi. Pozitif ve negatif süpercoil olmuş DNA için farklı dönme eksenleri bulunarak ileri sürülen DNA dönmeleri, süpercoil olmuş bir DNA ile enzimin relaksasyon mekanizmasına yeni bakış açısı

getirebilir. Ayrıca, DNA’nın alt kısmının saat yönünün tersine bir kez dönmesi DNA’nın kanonik formuna dönmesi için yeterli olmasına rağmen, DNA’nın başlangıç yapısına dönebilmesi için saat yönünde en azından iki dönüş gerekli olduğunu öne sürdük. Sonuç kısmında özetlendiği gibi, ekstra yeni bulgular gözlendi. Literatürdeki kapsamlı çalışmaların sadece kesilmemiş DNA üzerine olmasına rağmen bu çalışma, kesik bir DNA molekülü üzerinde yapılan ilk çalışmadır.

Anahtar Kelimeler: DNA topolojisi, süpercoiling, Moleküler Dinamik Simülasyonlar, HQBMD.

ACKNOWLEDGEMENT

I would like to express my gratitude to my supervisor Assist. Prof. Dr. Levent SARI. I thank to him for his contribution, guidance, patience, experience and support throughout the research and writing of my thesis.

Also, my thanks go to Prof. Dr. Mustafa KUMRU for their valuable help.

I express my thanks and appreciation to my family and friends especially Fatma GÖZÜAK for their understanding, motivation and patience.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

SYMBOL/ABBREVIATION

DNA	:	Deoxyribonucleic acid
ccDNA	:	Closed circular DNA
topo	:	Topoisomerase
MD	:	Molecular Dynamics
HQBMD	:	High Quality Biased Molecular Dynamics
Lk	:	Linking number
Lk_0	:	Linking number in relaxed state
Tw	:	Twist
Wr	:	Writhe
ΔLk	:	Linking difference
σ	:	Supercoiling density or specific linking difference
Tyr	:	Tyrosine
Lys	:	Lysine
CPT	:	Camptothecin
NMR	:	Nuclear Magnetic Resonance
k_b	:	Force constant for bond interaction energy
k_θ	:	Force constant for bond angle potential energy
k_Φ	:	Force constant for dihedral angle potential energy
k_ω	:	Force constant for bond improper potential energy
k_u	:	Force constant for Urey-Bradley potential energy
δ	:	Phase of dihedral
ε_0	:	Depth of the minimum energy for Lennard-Jones potential
ε	:	Dielectric constant
α_{hqbmd}	:	Force constant for HQBM

ρ	:	Reaction coordinate
GBMV	:	Generalized Born using Molecular Volume
CHARMM:		Chemistry at HARvard Macromolecular Mechanics
VMD	:	Visual Molecular Dynamics
LSF	:	Load Sharing Facility
PSF	:	Protein Structure File
RTF	:	Residue Topology File
PARAM	:	Parameter File
PSF	:	Protein Structure File
CRD	:	Coordinate File
SD	:	Steepest Descent
ABNR	:	Adapted Basis Newton-Raphson

CHAPTER 1

INTRODUCTION

1.1 DNA TOPOLOGY AND SUPERCOILING

Topology which is important for structure and function of DNA is the study of properties of an object that do not change under continuous deformations. Fuller applied the topology to studying the properties of closed circular DNA by using the results of ribbon theory [1]. According to this theory, besides the topological characteristic of a ribbon, the linking number Lk value plays an important role, the twist, Tw and writhe, Wr of the ribbon. Lk is a strict topological invariant and does not depend on the geometry but Tw and Wr are two differential geometric properties. The mathematical relationship between these quantities is:

$$Lk = Tw + Wr \tag{1.1}$$

which is really interesting because a topological property is equal to the sum of two geometric properties.

A DNA molecule which stores the genetic information in all living cell is composed of two strands of deoxyribonucleotide polymers, in a very special geometric relationship in which one is wound about the other and results an overall helical shape. These two polymer chains attached to one another by weak, noncovalent interactions. This is familiar with double helix, described by Watson and Crick, in which the two helices share a common axis, which is called as helical axis and both are wound in a right-handed manner.

First studies focused on the properties of linear DNA molecules, since this was precisely the form of DNA that could be extracted from cells and virus particles. It was unexpectedly found that DNA existed in a closed circular (cc) form in many viral DNA

and plasmids. Closed circular means DNA contains no breaks in either strand. Many physical properties of linear and closed circular forms are different. The most striking difference between circular and linear DNA is the fact that circular DNA can be knotted [1].

DNA topology helps us to understand function of DNA. There are two topological levels in ccDNA: the double helix as a whole can form knots of different types and the two complementary strands are interlinked, which results DNA supercoiling. Supercoiling, knotting and linking are important in cellular process, so we should understand their geometry and topology.

DNA supercoiling is important for DNA compaction in all cells. A typical human cell contains a total of 2 m of DNA and, the size of a cell is 10 μm. It is obvious that, DNA must be very tightly compacted to fit into the cell. Supercoiling of DNA allows for more DNA packaging by reducing the space. Also, packaging must permit access to the information coded in the DNA for processes such as replication and transcription.

Supercoiling means coiling of a coil. A typical phone cord is a coiled wire. If it is twisted, it results supercoil (Figure 1.1). DNA is coiled in the form of double helix. A bending or twisting of helical axis upon itself is called as DNA supercoiling. If there is no net bending, the DNA is in a relaxed state (the first picture in Figure 1.2).

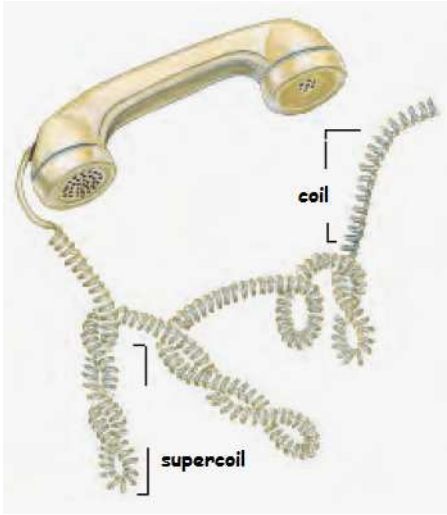

Figure 1.1 Supercoils [2].

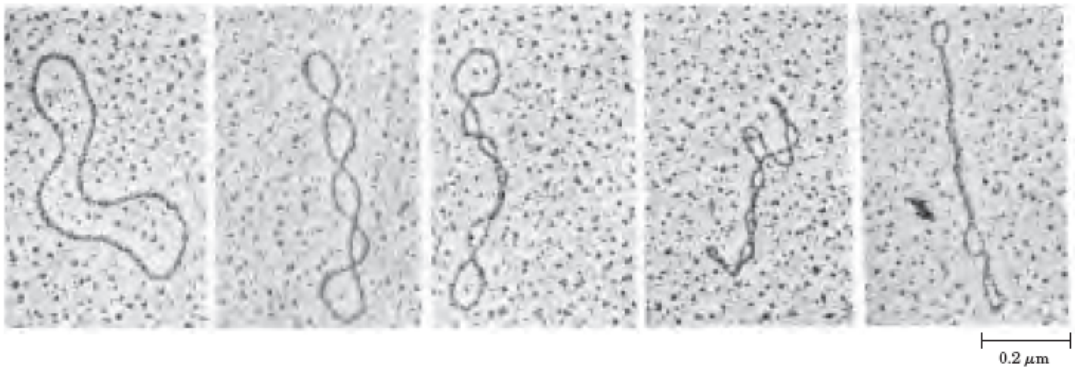

Figure 1.2 Electron micrographs of relaxed and supercoiled plasmid DNAs [2].

Topologically, a ccDNA can be defined by linking number, Lk, the number of times one DNA strand goes around the other, in the absence of any supercoiling. Lk is a strict topological property because it does not vary in the case of double stranded DNA is twisted or deformed, if both DNA strands remain intact. Lk is only defined for covalently closed DNA and its value does not change as long as the molecule remains covalently closed. Lk of a circular DNA can only be changed by breaking a phosphodister bond in one of the two strands, allowing the intact strand to pass through the broken strand and than rejoining the broken strand. If there is a break in either strand, it is possible to untwist the strands and separate them completely. In this case Lk is undefined, because no topological bond exists (Figure 1.3).

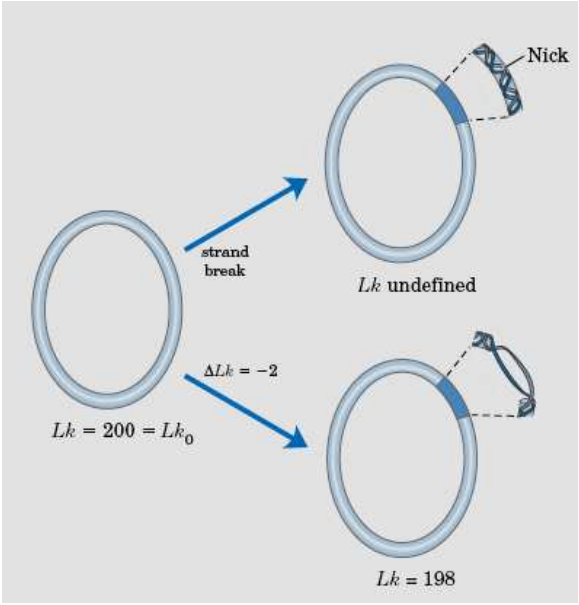

Figure 1.3 Linking number [2]

Linking number is described by the summation of two structural components called writhe and twist. Local inter-winding of the double strands itself results crossings or nodes and they measured by a parameter called twist (Tw). Twist is the number of times that the two strands are twisted about each other. Coiling of the helical axis results nodes and they measured by writhe (Wr). Writhe is the number of times that the DNA helix is coiled about itself in three-dimensional space. Twist and writhe may be changed by deformation of DNA, because they depend on geometry. In addition, Tw and Wr are not necessarily integers, indeed most often they are not.

Lk is always integer for a ccDNA but may be positive or negative depending on the orientation of the two strands. If two strands are interwound in a right-handed helix, the linking number is defined as positive (+). An ordinary B-type DNA has positive linking number, according to the sign convention. Conversely, if they interwound in a left-handed helix, the linking number is negative (-). And if there is no net bending of the DNA axis upon itself or it lies flat on a plane, than the DNA is said to be in a relaxed state (Figure 2). The linking number in relaxed state is called Lk_0, and is the number of base pairs divided by 10.5. Lk_0 is used as a reference [2]. If there is strain on DNA, Lk is less or more than Lk_0, than the DNA will undergo a three-dimensional writhing in space, which is called supercoiling. Supercoiled DNA molecules are torsionally stressed relative to their relaxed counterparts. The linking difference ΔLk

$$\Delta Lk = Lk - Lk_0 \tag{1.2}$$

is used to define the degree of supercoiling.

DNA underwinding produces negative supercoiling and overwinding produces positive supercoiling (Figure 1.4). Circular DNA isolated from most biological systems is underwound and the DNA shows negative supercoiling [3]. In replication process the unwinding of parental DNA duplex produces a positive linking number difference, or superhelical strain. It causes an increase in ΔLk, because separation of the parental strands lowers the value of Lk_0. The replicating DNA can form positive precatenanes, in which the daughter DNAs are intertwined, and positive supercoils. Therefore, the ΔLk of replication increases by about one for every ten base pairs of replicated DNA [4]. In addition, DNA is usually in a supercoiled state, when it is isolated from the cells.

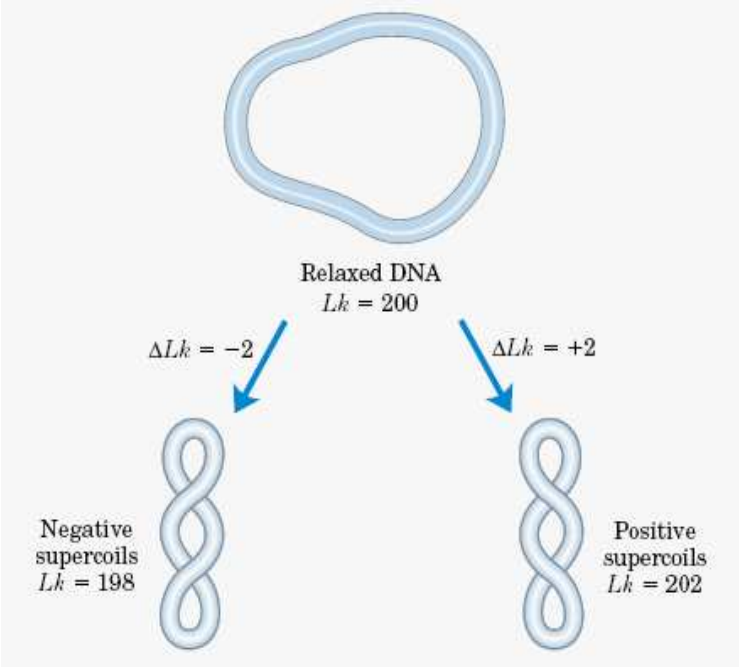

Figure 1.4 Negative and positive supercoiling [2].

The most useful measure of the topological deviation of DNA from the relaxed state is its supercoiling density, or σ. Supercoiling density is more convenient because it can be easily determined experimentally and is therefore independent of DNA length. It is also called superhelical density or specific linking difference and defined as

$$\sigma = \frac{\Delta Lk}{Lk_0} \tag{1.3}$$

DNA is a right handed helix, so it has positive linking number. However, it is generally underwound in the cell. So, Lk is less than Lk_0, σ is negative, and DNA is negatively supercoiled.

The linking number can only be changed by breaking one DNA strand and rejoining the broken ends. The enzymes that increase or decrease the linking number are called topoisomerases. Therefore, they can changed the degree of supercoiling in DNA by cutting one or both strand.

1.2 TOPOISOMERASES AND THEIR CELLULAR IMPORTANCE

DNA topoisomerases are enzymes that can change the topological state of DNA by increasing or decreasing the linking number without changing its primary structure. They regulate the number of topological links between two DNA strands by catalyzing transient breaks one or both strand, passing the strands through one another than resealing the breaks.

There are three main types of topological problems: supercoiling, knotting and catenation (Figure 1.5). Topoisomerases can solve all of these topological problems of DNA in processes such as replication, transcription, recombination and chromosome segregation**.** They can remove DNA supercoiling by catalyzing DNA swiveling and relaxation. Therefore, they are required to relieve torsional stress in DNA caused by supercoiling during transcription or recombination. In addition to remove supercoiling, strand passing activities can alter the catenation and knotting.

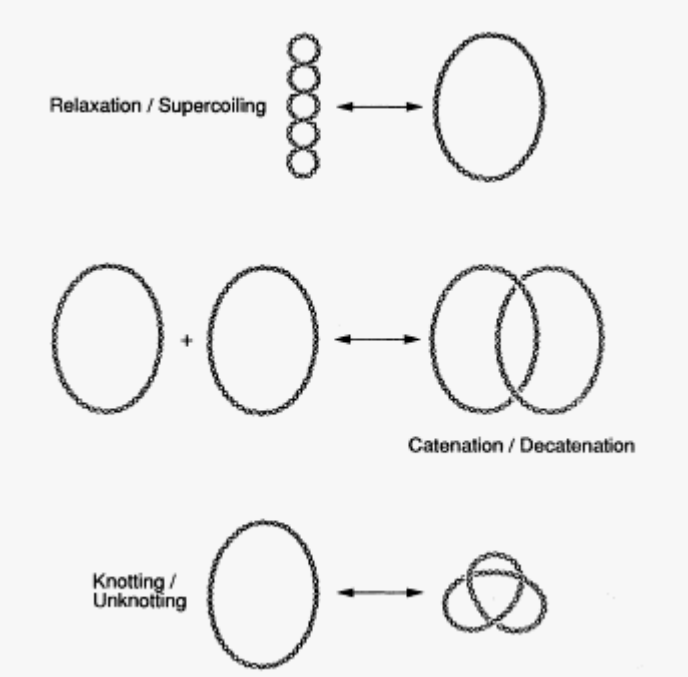

Figure 1.5 The reactions carried out by topoisomerases [4].

In the strand cleavage by all topoisomerases, tyrosyl oxygen of the enzyme attacks the DNA phosphorus and forms a temporary covalent phosphotyrosine bond between the enzyme and one end of the broken strand and also breaks a DNA phosphodiester bond [5]. During the lifetime of the covalent intermediate, DNA topology can be modified. Rejoining of the DNA strand is the reverse of the breaking. The oxygen of the DNA hydroxyl group that is produced in the cleavage attacks the phosphorus of the phosphotyrosine link, breaking the covalent bond, and reforming the DNA backbone bond [6]. The energy of the phosphotyrosine linkage is used for the restoring of DNA, thus, cleaving and rejoining process require no energy cofactor.

Topoisomerases are classified into the two parts: type I and type II based upon their distinct difference in sequence, structure and cellular function. Type I topoisomerases can alter the torsional strain caused by the helical nature of DNA. They are important for DNA replication and transcription. In these processes DNA must be unwound to allow reading of the information and, topoisomerase I releases the stress that occurs in the process. They are monomeric and cleave only one strand of DNA by resulting the linking number to change in increments of one. Type II topoisomerases are dimeric and cleave both strands of DNA to generate a gate through which another region of DNA can be passed. They change the linking number in steps of two. They play an important role in cell divisions.

The two types are further classified into four subfamilies: IA, IB, IIA and IIB. Structure and mechanism are similar for the members of the same subfamily, but there are distinct structural differences between different subfamilies. Further division of the subfamilies is based on structural considerations. Various subfamilies of both prokaryotic and eukaryotic topoisomerases are listed in Table 1.

Table 1.1 Classification of topoisomerases [8]

Topoisomerase[a]	Subfamily type	Subunit structure	Size(s) (aa)[b]
Eubacterial DNA topoisomerase I (*E. coli*)	IA	Monomer	865
Eubacterial DNA topoisomerase III (*E. coli*)	IA	Monomer	653
Yeast DNA topoisomerase III (*S. cerevisiae*)	IA	Monomer	656
Mammalian DNA topoisomerase IIIα (human)	IA	Monomer	1001
Mammalian DNA topoisomerase IIIβ (human)	IA	Monomer	862
Eubacterial and archaeal reverse DNA gyrase (*Sulfolobus acidocaldarius*)	IA	Monomer	1247
Eubacterial reverse gyrase (*Methanopyrus kandleri*)[c]	IA	Heterodimer	A, 358 B, 1221
Eukaryotic DNA topoisomerase I (human)	IB	Monomer	765
Poxvirus DNA topoisomerase (vaccinia)	IB	Monomer	314
Hyperthermophilic eubacterial DNA topoisomerase V (*Methanopyrus kandleri*)[d]	IB	Monomer	—[e]
Eubacterial DNA gyrase (*E. coli*)	IIA	A_2B_2 hetero-tetramer	GyrA, 875 GyrB, 804
Eubacterial DNA topoisomerase IV (*E. coli*)	IIA	C_2E_2 hetero-tetramer	ParC, 752 ParE, 630
Yeast DNA topoisomerase II (*S. cerevisiae*)	IIA	Homodimer	1428
Mammalian DNA topoisomerase IIα (human)	IIA	Homodimer	1531
Mammalian DNA topoisomerase IIβ (human)	IIA	Homodimer	1626
Archaeal DNA topoisomerase VI (*Sulfolobus shibatae*)	IIB	A_2B_2 hetero-tetramer	A, 389 B, 530

[a]The source of the most extensively studied family member is given in parentheses. The top portion of the table lists the type I topoisomerases; the bottom portion the type II enzymes.

[b]The subunit sizes are those corresponding to the most extensively studied family member.

[c]Included as the only known reverse gyrase with a heterodimeric structure.

[d]Only known representative at present. Probably present in other hyperthermophilic eubacteria.

[e]Gene not yet cloned; purified protein has a molecular size of 110 kDa.

Our study of DNA rotation is important in the mechanisms of typeI topos. There are two kinds of type1 topos; Type IA and Type IB. Type IA topoisomerases can relax only negatively supercoiled DNA, require Mg and a single stranded region of DNA for the relaxation activity. They form a covalent intermediate with the 5' phosphate end of the nicked DNA strand (Figure 1.6). Furthermore, these enzymes can catalyze the knotting, unknotting and interlinking of single stranded circles. Nevertheless, Type IB

topoisomerases can relax both negatively and positively supercoiled DNA with equal efficiency, and do not require a single stranded region of DNA or metal ions for function. They attach covalently to the 3' phosphate end of the cleaved strand of DNA (Figure 1.6) rather than 5' phosphate end. In addition, there is no sequence or structural similarity between type IA and IB.

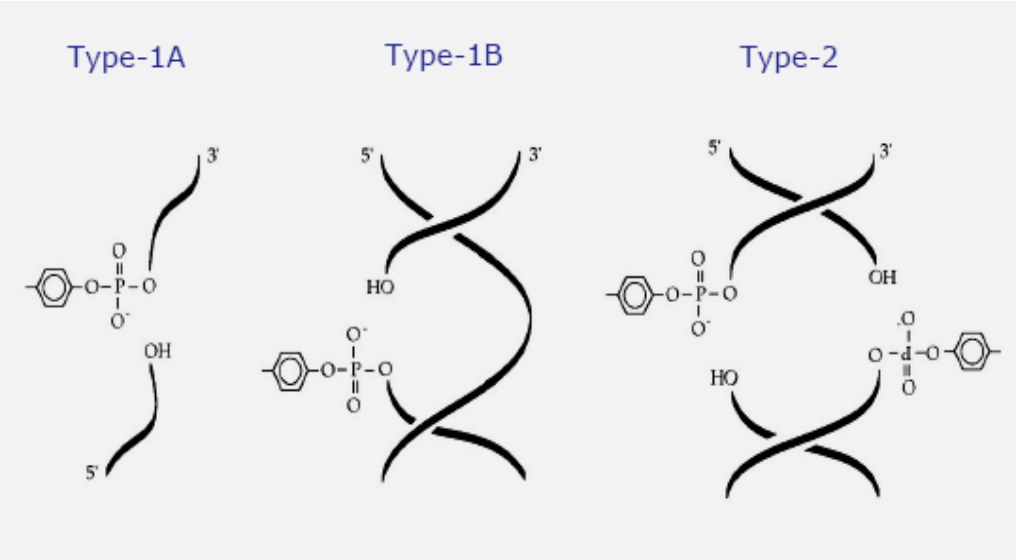

Figure 1.6 Types of topoisomerases [7]

1.3 HUMAN TOPOISOMERASE I

Human topoisomerase I (topo I) is a monomeric type IB protein that composed of 765 residues. Based on amino acid sequence comparison, the 91 kDa topo I protein can be divided into four distinct domains: the N-terminal, core, linker and the C-terminal domains (Figure 1.7).

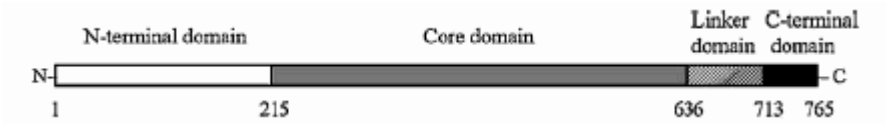

Figure 1.7 Domain structure of human topoisomerase I [8].

The 214 amino acids N-terminal domain of human topo I is highly charged, poorly conserved and largely disordered. The 24 kDa N-terminal domain contains hydrophilic and highly protease-sensitive region of the protein. This domain includes four nuclear localization signals and sites for interaction with other cellular proteins and the first 174 amino acids are responsible for the slow migration and asymmetric shape of the protein. Therefore, N-terminal domain is dispensable for relaxation of DNA supercoils in vitro, because experiments indicate that, full-length protein and amino-terminally deleted enzyme shows identical activity [9].

The highly conserved, 421 amino acids core domain comprising residues 215 to 636 (54 kDa), contains all of the catalytic residues except the active site tyrosine. The crystal structure of topo I reveals that this domain can be divided into three subdomains: I, II and III.

The core domain and C-terminal domain are connected with 77 amino acids linker domain (5kDa) which is poorly conserved, positively charged and protease sensitive. Residues 636 to 712 contribute to relaxation activity in vitro but are not required. The linker domain can be found in one of two states. In the absence of DNA, it is in open state and is sensitive to proteolysis; when topo I is bound to DNA covalently, it is in closed state and less sensitive to proteolysis. Furthermore, during the formation of covalent complex the linker domain may switch from the closed to open state [10].

C-terminal domain composed of 53 amino acids and contains the active site tyrosine (Tyr723). Amino acids which are located at the downstream part of the active site are essential for the activity.

Human topoisomerase I is a bi-lobed protein that clamps tightly around the DNA. The top lobe that contains core subdomains I and II is called upper cap of the protein with helices α5 and α6 forming the nose cone. These V-shaped helices are important for the topoisomerization mechanism [5]. The core subdomain III and C-terminal domain compose the lower cap, and form a base that cradle the DNA. It is connected to the upper cap by a long α-helix flexible hinge called connector. In the opposite side of hinge region, upper and lower caps of the protein are connected with a pair of opposing loops called lips. The upper lip is placed at the core subdomain I and lower lip is located

at the subdomain III. The upper and lower lips are covalently linked and interact with each other via six amino acids and one salt bridge [11]. This interaction is broken during DNA binding and reseal.

These four domains clamp around the DNA and form covalent or noncovalent complex with DNA (Figure 1.8). There is only one base-specific contact that includes hydrogen bond between Lys532 of the protein and the O-2 carbonyl oxygen of the thymine base located at the -1 position of the scissile strand.

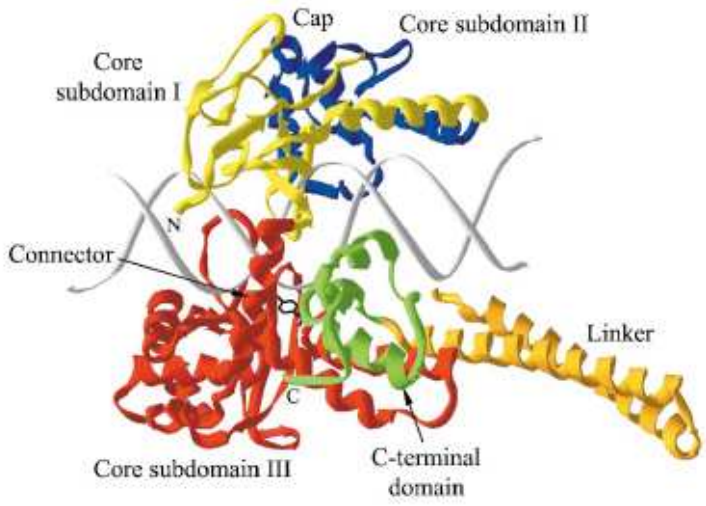

Figure 1.8 Structure of human topoisomerase I with a 22 base pair DNA [8].

1.4 THE CRUCIAL ROLE OF DNA ROTATION WITHIN THE HUMAN TOPOISOMERASE I

Two models have been proposed to explain the mechanism of DNA relaxation in type I topos after DNA is cleaved: free rotation and enzyme-bridging. In free rotation model, protein is attached covalently to one end of the broken strand and downstream of the DNA duplex freely rotate about the unbroken strand in either direction. This model is most favorable for the type IB subfamily. In enzyme-bridging model, protein is not only attached covalently to one end, is also attached noncovalently to the other end to

form a bridge. Then unbroken strand is passed through this bridge. This strand passage model is suited for type IA subfamily [10].

Stewart *et al.* proposed a model for DNA relaxation by human topoisomerases I, called controlled rotation which is modified form of the free rotation model [5]. The relaxation reaction initiates with the binding of topo I to the DNA duplex. Initially, enzyme must be in open state for DNA binding. The upper cap lifts away from the lower cap by upward movement at the hinge point (Figure 1.9 A). Then, lips are closed around the DNA strand and noncovalent complex is formed between enzyme and DNA (Figure 1.9 B). This closed form of the enzyme catalyzes cleavage of the scissile strand. Nucleophilic attacks of the O-4 atom of the active site tyrosine result cleavage of the scissile strand and covalent attachment of the enzyme to the 3' end of the DNA (Figure 1.9 C). After covalent intermediate is formed, downstream part of the nicked DNA starts to rotate about the intact strand due to the supercoiling energy. Linker domain of human topo I is highly flexible [12], therefore during the controlled rotation cap and linker regions move upward and downward to allow the DNA rotate freely within the enzyme as shown small arrows in Figure 1.9 D. Then, superhelical tension is released after one or more rotations. Linking number of DNA changes by one in per rotation. In addition, topo IB enzyme can relax both negative and positive supercoils because of this strand rotation. Finally, religation that is mechanistically just the reverse of the cleavage reaction is occurred and noncovalent complex is regenerated (Figure 1.9 F). The enzyme opens up to release the DNA that is converted to less supercoiled state (Figure 1.9 G).

As seen in the Figure 1.9 D, the **DNA rotation is the fundamental process** in the removal of supercoils. The real-time DNA rotation within the enzyme has recently been simulated [14] as suggested by Stewart *et al* [5]. Later, results of these DNA rotations have been verified experimentally [15]. Therefore, detail investigation of DNA rotations is quite important in understanding the structural deformations of DNA within the protein. Our study focuses on different structural and energetic pathways that the nicked DNA undergoes during rotations, and aims to propose the best rotation scheme.

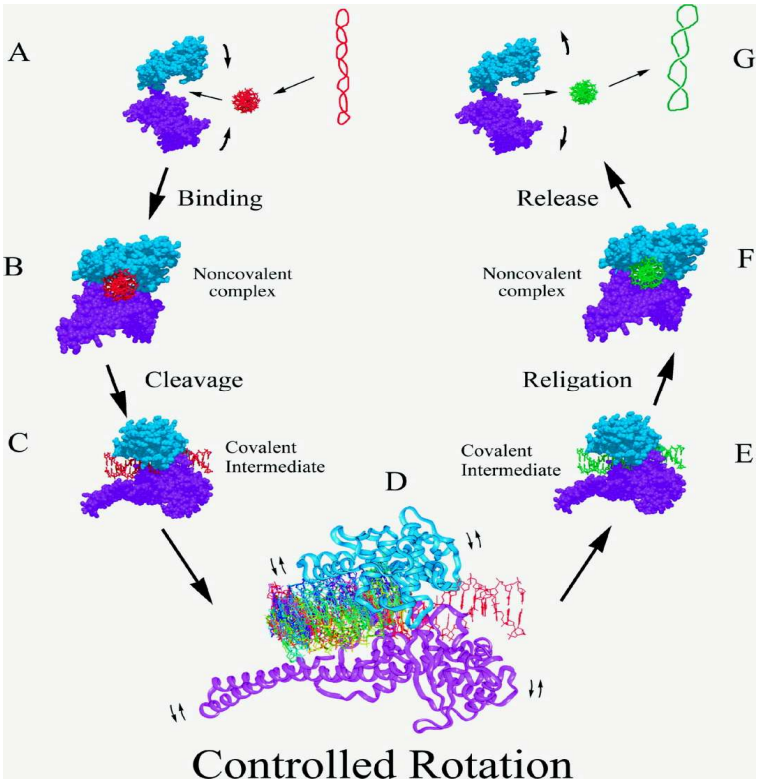

Figure 1.9 The controlled rotation mechanism of human topo I [5].

1.5 TOPO TARGETTING ANTICANCER DRUG MOLECULES

Inhibition of the DNA rotation produces a stable tertiary complex (DNA-protein-anti cancer molecule) on the DNA [13]. The replication fork stops as it hits this stable complex on DNA as seen in Figure 1.10 Therefore, topoisomerases have become quite popular targets for cancer chemotherapy treatments, anti viral, anti bacterial and anti tumor drugs. Human topo I is the sole target of the camptothecin (CPT) family of anticancer drugs. Two analogs of CPT, topotecan and irinotecan, have been used successfully in the treatment of several human cancers. CPT inhibits the religation steps, not initial cleavage action. In the topo I reaction, CPT and its derivatives lengthen the lifetime of the cleaved intermediate by blocking the rejoining step of the cleavage. It is assumed that, this resulting covalent intermediate constitute an obstacle to the advancement of transcription and replication process. Finally, it causes DNA damage and cell death [8].

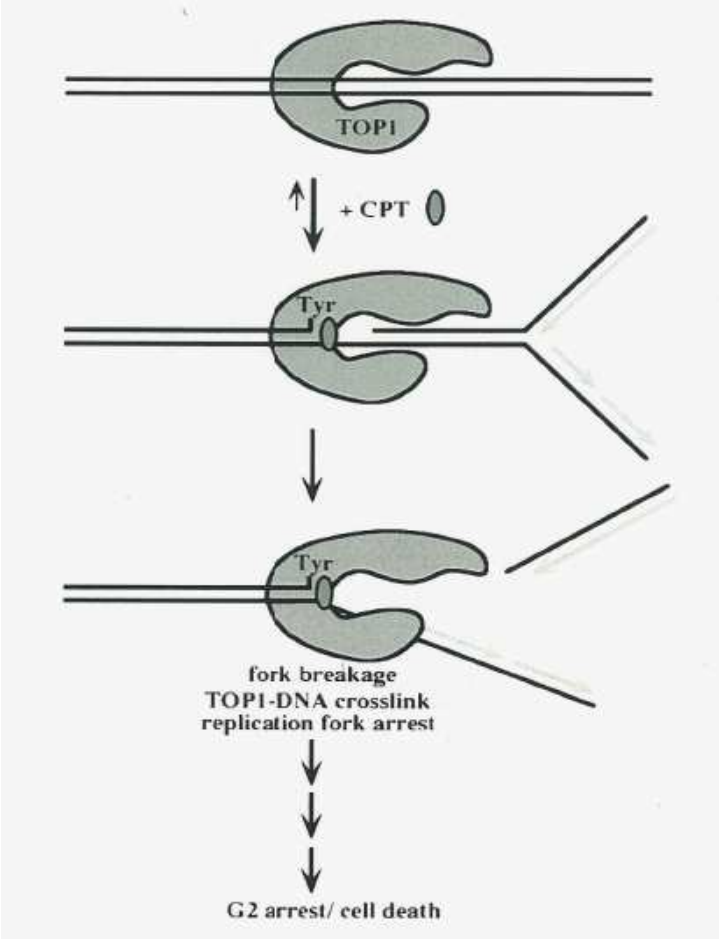

Figure 1.10 A replication collision model for camptothecin cytotoxicity [13]

CHAPTER 2

THEORETICAL APPROACH

2.1 FORCE FIELD AND MOLECULAR DYNAMICS SIMULATIONS

Computational simulations are used to elucidate molecular systems at the atomic level. These computational methods determine the structure and dynamics of molecules by solving the time dependent Schrödinger equation. But, for many molecular systems that contains thousands of atoms, such as proteins and nucleic acids the classical physical approaches are suitable. Molecular dynamics (MD) simulations method is one of these approaches [16].

Molecular dynamics (MD) is defined as the science of simulating the motions of a system of particles. The methods of MD simulations play an important role in the theoretical study of biomolecules. MD simulations are powerful techniques for describing and understanding the relationships of the structure and the function of biomolecules. Also, MD can be used to study fast events that occur on the order of picoseconds to nanosecond time scales. This computational method calculates the time dependent behavior of a molecular system using Newton's equations of motions [17].

MD was first applied to the bovine pancreatic trypsin inhibitor by McCammon *et al.* in 1977 [18]. These methods are now routinely used to investigate the structure, dynamics and thermodynamics of biological molecules and their complexes. They are also used in the determination of structures from x-ray crystallography and from NMR experiments [19].

Quantum mechanical effects can be neglected in most MD simulations, because of the large size of biomolecules and instead empirical potential energy function is used to determine the interaction energy of the particles of a system as a function of the atomic coordinates. And, MD uses this potential energy to calculate the future position of the particles. In general, energy of system is separated into two parts: bonded and non-bonded energies.

$$U_{total} = U_{bonded} + U_{nonbonded} \tag{2.1}$$

where bonded term refers to atoms that are linked by covalent bond and, nonbonded terms describe the long range electrostatic and van der Waals interactions, which is also called noncovalent.

$$U_{bonded} = U_{bonds} + U_{angles} + U_{dihedrals} + U_{impropers} + U_{Urey\text{-}Bradley} \tag{2.2}$$

$$U_{nonbonded} = U_{Lennard\text{-}Jones} + U_{Coulomb} \tag{2.3}$$

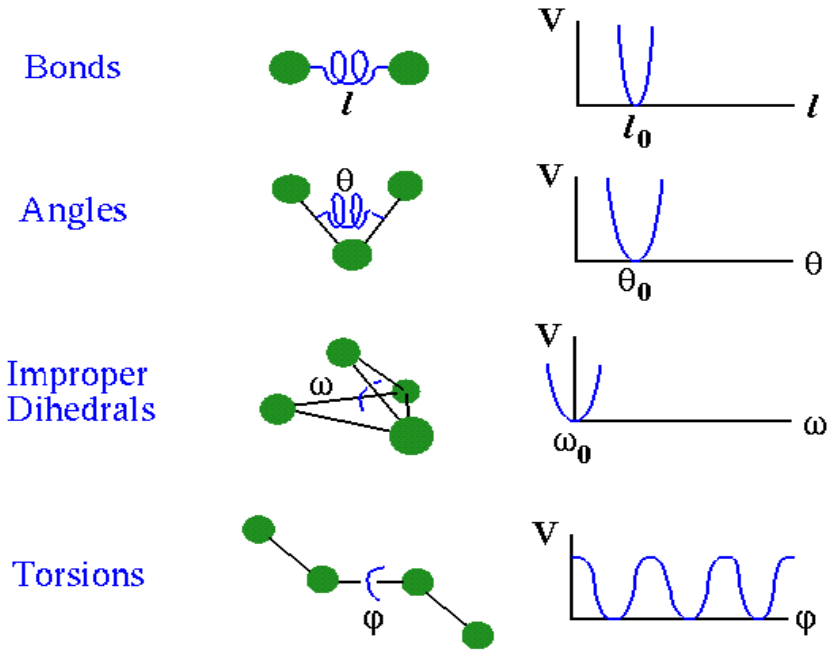

Figure 2.1 Empirical potential energy functions for bonded interactions [20].

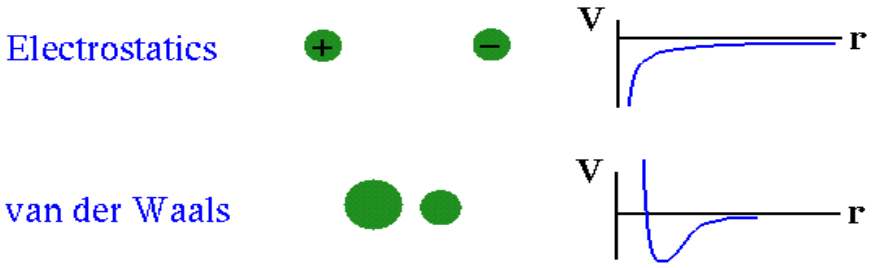

Figure 2.2 Empirical potential energy function for non-bonded interactions [20]

U_{bonds} is the bond interaction energy. It is based on Hook's law and caused by shared electrons:

$$U_{bonds} = \sum_{bonds} k_b (b - b_0)^2 \tag{2.4}$$

where k_b is the force constant, b_0 is the equilibrium bond distance and b is the current bond distance. This equation estimates the energy associated with vibration about the equilibrium bond length.

U_{angles} is the bond angle potential energy. It is also based on Hook's law. Similarly it can be defined as:

$$U_{angles} = \sum_{angles} k_\vartheta (\theta - \theta_0)^2 \tag{2.5}$$

where k_θ is the force constant, θ_0 is the equilibrium bond angle, and θ is the current angle formed by three atoms. This equation estimates the energy associated with vibration about the equilibrium bond angle.

Dihedral angles are formed by four atoms linearly bonded together. $U_{dihedrals}$ is the energy due to the dihedral angle potential which is defined by a periodic function:

$$U_{dihedrals} = \sum_{dihedrals} k_{\phi}[1+\cos(n\phi-\delta)] \tag{2.6}$$

where k_{ϕ} is the force constant, *n* is the multiplicity (number of equivalent minima), ϕ is the dihedral angle value, and δ is the phase of the dihedral.

$U_{impropers}$ is the energy due to the planer dihedral angle which is called improper and harmonic similar to the bond potential:

$$U_{impropers} = \sum_{impropers} k_{\omega}(\omega-\omega_0)^2 \tag{2.7}$$

where k_{ω} is the force constant and ω is the angle between two planes.

$U_{Urey\text{-}Bradley}$ term is also harmonic and defined as:

$$U_{Urey-Bradley} = \sum_{Urey-Bradley} k_u(u-u_0)^2 \tag{2.8}$$

with Urey-Bradley 1.3 distance, u_0.

Non-bonded interactions determine the weak interactions between atoms. Lennard-Jones potential represents the attraction and repulsion force between non-bonded atoms:

$$U_{Lennard-Jones} = \sum_{nonbonded} \varepsilon\left[\left(\frac{\sigma_{i,j}}{r_{i,j}}\right)^{12} - \left(\frac{\sigma_{i,j}}{r_{i,j}}\right)^{6}\right] \tag{2.9}$$

First term describes repulsion at short range and second term describes attraction at long range. ε_0 is the dept of the minimum energy for the interaction, $\sigma_{i,j}$ is the finite distance at which the interparticle potential is zero and *r* is the distance between particles.

$U_{Coulomb}$ term is the electrostatic interaction energy between charged atoms described as:

$$U_{Coulomb} = \sum_{nonbond} \frac{q_i q_j}{\varepsilon r_{i,j}} \tag{2.10}$$

where q is the charge of atoms and ε is the dielectric constant.

These equations together with the energy parameters (*e.g.* $k_b, \sigma_{i,j}$) required to describe the behavior of different kinds of atoms and bonds, is called force-field [21]. Energy parameters, bond lengths and bond angles are usually taken from experimental structural data or from quantum chemical calculations.

To start the MD simulation, initial positions and velocities to all particles in the system should be assigned. The initial coordinates of atoms at time t=0, are read from the crystal structure which are obtained from experimental structure. Then, force-field potential is determined to calculate the energy of a molecular structure. Force-field potential provides to calculate the future position and velocity of the particles. Firstly, the force acting on each atom is found from the derivative of the potential energy function:

$$F(x) = -\nabla U(x) \tag{2.11}$$

Then, the system is moved according to the laws of Newtonian mechanics by integrating equation of motions.

$$F(x) = m\dot{V}(t) \tag{2.12}$$

The potential energy is a function of the atomic positions (3N) of all the atoms in the system. Due to the complicated nature of this function, there is no analytical solution to the equations of motion; they must be solved numerically. Numerous numerical algorithms have been developed for integrating the equations of motion [22]. The Velocity Verlet algorithm is a more commonly used algorithm. This algorithm is derived from Taylor expansion. Thus, the velocities and positions at time $(t + \Delta t)$ are calculated using these equations:

$$x(t + \Delta t) = x(t) + v(t)\Delta t + \frac{1}{2} a(t)(\Delta t)^2 \tag{2.13}$$

$$v(t+\Delta t) = v(t) + \frac{a(t)+a(t+\Delta t)}{2}\Delta t \tag{2.14}$$

The Velocity Verlet algorithm is implemented as follows:

Firstly, positions are obtained using:

$$x(t+\Delta t) = x(t) + v(t)\Delta t + \frac{1}{2}a(t)(\Delta t)^2 \tag{2.15}$$

Then, mid-step velocities are calculated from the acceleration at time (t) by

$$v\left(t+\frac{\Delta t}{2}\right) = v(t) + \frac{a(t)\Delta t}{2} \tag{2.16}$$

The accelerations at time $(t+\Delta t)$ are derived from the interaction potential. After it is computed, the next velocities are obtained using:

$$v(t+\Delta t) = v\left(t+\frac{\Delta t}{2}\right) + \frac{a(t+\Delta t)}{2}\Delta t \tag{2.17}$$

The acceleration comes from the relation:

$$F = -\nabla U = ma \tag{2.18}$$

Nevertheless, this algorithm assumes that acceleration at time $(t+\Delta t)$ does not depend on velocity $v(t+\Delta t)$; only depends on position $x(t+\Delta t)$:

$$a(t+\Delta t) = -\left(\frac{1}{m}\right)\nabla U[x(t+\Delta t)]$$

2.2 HIGH QUALITY BIASED MD

HQBM is an external perturbation designed to induce conformational changes in macromolecules. The time dependent perturbation designed to introduce a very small perturbation to the short time dynamics of the system and does not affect the conservation of the constants of motion of the system (the conservation of the total energy or of the suitable conserved quantity when an extended Lagrangian is used can then be used as a check of the correctness of the forces). The external perturbation needs: (1) a reference (or target) structure, (2) a reaction coordinate which defines a "distance" from the reference structure [23].

HQBM introduce a half quadratic perturbation on a given reaction coordinate [24]. The system was biased half-harmonically, i.e. only when it moves away from the target [14], with an external potential of the form:

$$W(r,t) = \frac{\alpha_{hqbmd}}{2}(\rho - \rho_0)^2 \tag{2.19}$$

where α_{hqbm} is force constant and ρ is reaction coordinate leading from the initial to the final state is given by:

$$\rho(t) = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j \neq i}^{N} \left(r_{ij}(t) - r_{ij}^{R} \right)^2 \tag{2.20}$$

where $r_{ij} = |r_i - r_j|$ is the distance between atoms i and j, and R labels the coordinates of the final reference structure, and N represents the total number of atoms biased.

We have calculated average forces that are needed to rotate DNA around different axes. The forces are obtained from the applied external potentials, as shown below:

$$F(t) = -\frac{d}{dr}\left[W(r,t) \right] \tag{2.21}$$

$$F(t) = 2\alpha_{hqbm} \left[\rho(t) - \rho(t+\Delta t) \right] \left[\rho(t) \right]^{1/2} \tag{2.22}$$

This technique has been used extensively in our research, especially to rotate the DNA in a short period of time. The details of the theoretical set-up of the rotations are given below.

2.3 NOSE-HOOVER CONSTANT TEMPERATURE MD

To investigate the behavior of a system at a specific temperature, it is required to use a thermostat for simulation. Another reason to simulate using thermostat is to avoid steady energy drifts caused by the accumulation of numerical errors during MD simulations [25]. To prevent possible artifacts from temperature variations, the simulations were performed at constant temperature using a Nose-Hoover thermostat.

Nose-Hoover constant temperature MD method performs simulations in the constant temperature situation. This method considers the system is thermally connected with a huge external system (a heat reservoir) and system exchanges energy with the heat reservoir. This method has the advantage that it is a continuous dynamics with well defined conserved quantities [23].

In Nose-Hoover method, two different types of constant temperature methods can be called: (1) simple method: system coupled to one heat bath, (2) multi-bath method: different parts of system can be coupled to different heat baths [23].

2.4 GENERALIZED BORN USING MOLECULAR VOLUME (GBMV)

The GBMV method is a Generalized Born method for revovering the Poisson-Boltzmann (PB) electrostatic solvation energy [26]. The PB procedure for obtaining solvation energies is a benchmark for implicit solvation calculations. However, the PB procedure is slow and the derivatives, i.e. forces, are ill-defined unless we change the definition of the molecular volume [23].

The Generalized Born equation, allows us to compute solvation energies similar to the PB equations. The parameters change very little when optimized for a particular force-field. Hence, forcefields besides those of CHARMM can be used with GBMV without refitting of parameters [23].

2.5 COMPUTATIONAL TOOLS

2.5.1 Software

One of the most popular molecular mechanics programs for biomolecules is CHARMM (Chemistry at HARvard Macromolecular Mechanics) written by Brooks *at al.* in 1983 [27] CHARMM is widely used molecular simulation program with broad application to many- particle system including macromolecular energy, minimization and dynamics calculations. It provides a large suite of computational tools that encompass numerous conformational and path sampling methods, free energy estimates, molecular minimization, dynamics, and analysis techniques, and model-building capabilities. All simulations were performed using CHARMM.

Other than the main software package CHARMM, we have used Visual Molecular Dynamics (VMD), gnuplot program for plotting, and several linux shell scripts that we wrote for data manipulations and job managements.

2.5.2 Hardware

We have used a clustered supercomputer, named ulubatli, in our research. Ulubatli is a linux cluster machine, having a total of 30 CPUs and a total of 30 GB RAM memory. It uses Load Sharing Facility (LSF) for job management and CMU (Cluster Management Utility) for cluster managements. The CPUs are Intel Xeon 5160 dual core, having 1333 MHz Front Side Bus speeds, and all CPUs are talking to each other through a fast networking technology (up to 40 GB/s), *infiniband.* The system has also a disk storage unit, HP MSA 30, which has a total disk capacity of 4.2 TB. Ulubatli is fully financed by The Scientific & Technological Research Council of Turkey (TÜBİTAK) grant no: 107T209.

2.6 DATA STRUCTURE FILES

Data structure files are needed to simulate biomolecules on the computer. Topology file, parameter file, and protein structure file are data structure files that are used in the CHARMM program. These files contain some information about the molecule such as its composition, its chemical connectivity, certain atomic properties, internal coordinates, force constants, energy parameters and more. Topology files and parameter files include these informations for a particular class of molecules, e.g., proteins or nucleic acids. Protein structure files contain the necessary data that is extracted from topology and parameter files, for a specific molecule.

2.6.1 Residue Topology File (RTF)

The residue topology file includes the type, mass, hydrogen bond donors and acceptors, and charge of every atom in a particular residue. This file contains information for the amino acids and nucleic acids which are called as molecular building blocks. Thus, to simulate a protein, DNA or large molecular system these building blocks are used. Furthermore, it contains definitions of the covalent structure for each residue and defines which atoms are connected to one another to form amino acids, DNA bases or lipid molecules. Also, it stores informations about internal coordinates and connectivity.

2.6.2 Parameter File (PARAM)

The parameter file contains force field parameters, e. g., equilibrium bond distances and angles for bond stretching, angle bending and dihedral angle terms in the potential energy function as well as the force constants and the Lennard-Jones radius and well depth. These force field parameters are necessary for calculating the energy of the system.

2.6.3 Protein Structure File (PSF)

The Protein Structure File (PSF) is the most fundamental data structure in the CHARMM program. It is generated for a specific molecule or molecules and it is a concatenation of the information contained in the RTF file. It contains the detailed

composition and connectivity of the molecules of interest. It gives the total number of angles, where triplets of atoms are listed, and for torsion angles, where quadruplets of atoms are specified. Improper dihedral angles are also listed; these are used to maintain planarity and prevent accidental chiral inversion. The PSF can include several molecular entities, defined as segments, which can range from a single macromolecular chain to multiple chains solvated by explicit water molecules [23].

The PSF must be specified before any calculations can be performed on the molecule. The PSF constitutes the molecular topology but does not contain information regarding the bond lengths, angles, etc., so it is necessary to read in the parameter file to add the missing information.

2.6.4 Coordinate File (CRD)

The coordinate file contains the cartesian coordinates of all atoms in the system. The cartesian coordinates of a molecule are most often obtained from x-ray crystal structures or from NMR structures. Alternatively, they can be generated from homology modeling studies or from the library values contained in the topology and parameter files along with some secondary structure information. Missing coordinates can be built within the CHARMM program using the internal coordinate facility; in addition, hydrogen atoms, which are not present in x-ray crystal structures of proteins may be placed using the HBUILD module in CHARMM. Two sets of coordinates can be accommodated in the CHARMM program, which is useful for a variety of calculations. For example, it can be used to calculate the root mean square difference (rmsd) between a structure from the simulation and the experimental structure [23].

2.7 SYSTEM PREPARATION

2.7.1 Building & Equilibrations

The 22 base-pair nicked DNA molecule is extracted from the crystal structure 1a31.pdb [5, 7], and minimized with a 500 step of Steepest Descent (SD), and a following 500 step of Adapted Basis Newton Raphson (ABNR) method. As the crystal

structures are frozen, we have first heated the system to 300 Kelvin and subsequently equilibrated for 0.5 ns at T=300 Kelvin, using an implicit solvation method GBMV (the details of the GBMV is mentioned above).

As we do not have hydrogen atoms in the crystal structure we have built all of the hydrogen atoms according to the topology file (see the above RTF file description in the data structure file section). Also, dummy atoms (atoms that have only coordinates but do not have any physical properties) were added on top of backbone DNA atoms. The reason for such dummies is to use them as some reference points in space as we analyze our results.

2.7.2 Preparing for Biased MD

In this study our purpose is to investigate the dynamic mechanism of a nicked DNA rotation about the intact strand employing MD simulations. In order to see large scale movement of DNA atoms, we employed HQBM. During the simulations, rigid bonds are constrained by the SHAKE algorithm and temperature is controlled using Nose-Hoover thermostat. To perform the simulations in water we used GBMV method.

As we rotate the downstream part of DNA, we are trying to find most suitable rotation axis, both structurally and energetically. To define our rotation axis, we have chosen two points. One point is fixed, which is a point on the phosphate atom between +1 and -1 base-pairs. The second point for the axis is chosen differently for different systems. The collections of the second point form a plane at the bottom of the DNA (see in the figure 2.3).

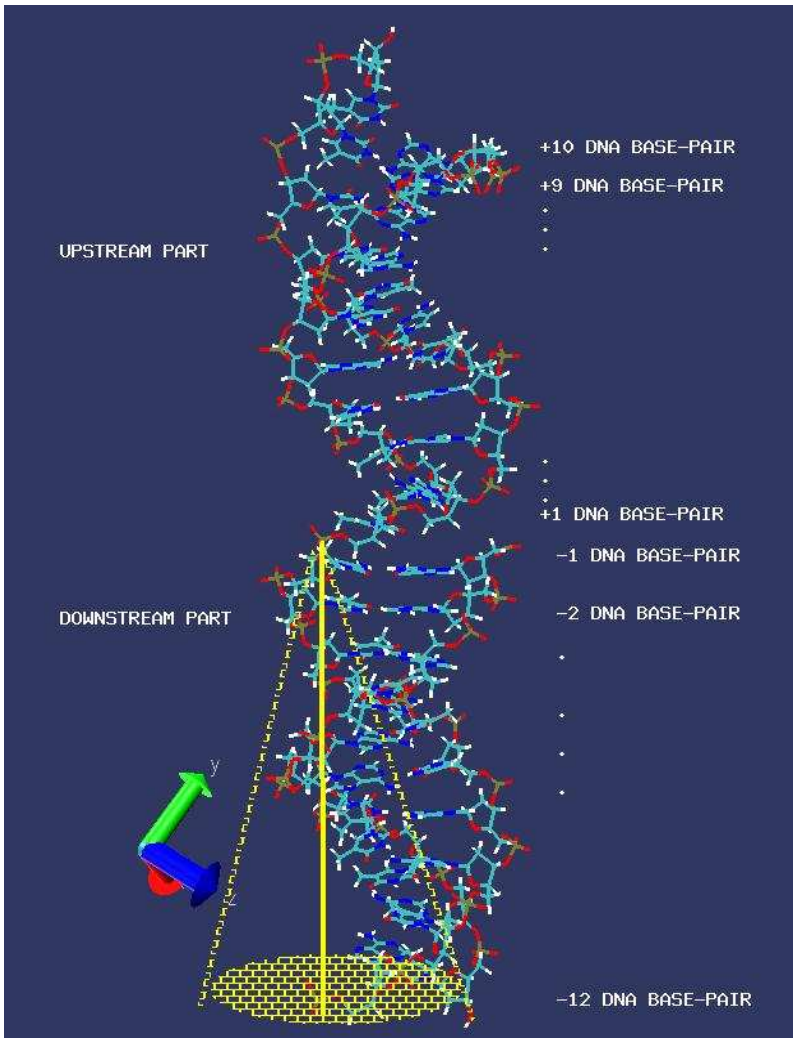

Figure 2.3: The nicked DNA in our study. The yellow line shows our axis of rotation, and the yellow plane at the bottom of the DNA shows the plane of the points that are choosen for the other end of the rotation axis.

By choosing two points, we have determined 25 different axis of rotation. One of them is chosen to be parallel to the helical axis that is called paralel axis. Then, a perpendicular vector to this axis as we called line 1 is found. This vector is rotated 90^0 by using rotation matrix to find the line 2. Again rotating line 1 and line 2 in 45^0, line 3 and line 4 are determined. These lines are on a plane that is perpendicular to the parallel axis of rotation. First point is fixed and all of the second points lie on these lines as shown in figure 2.4.

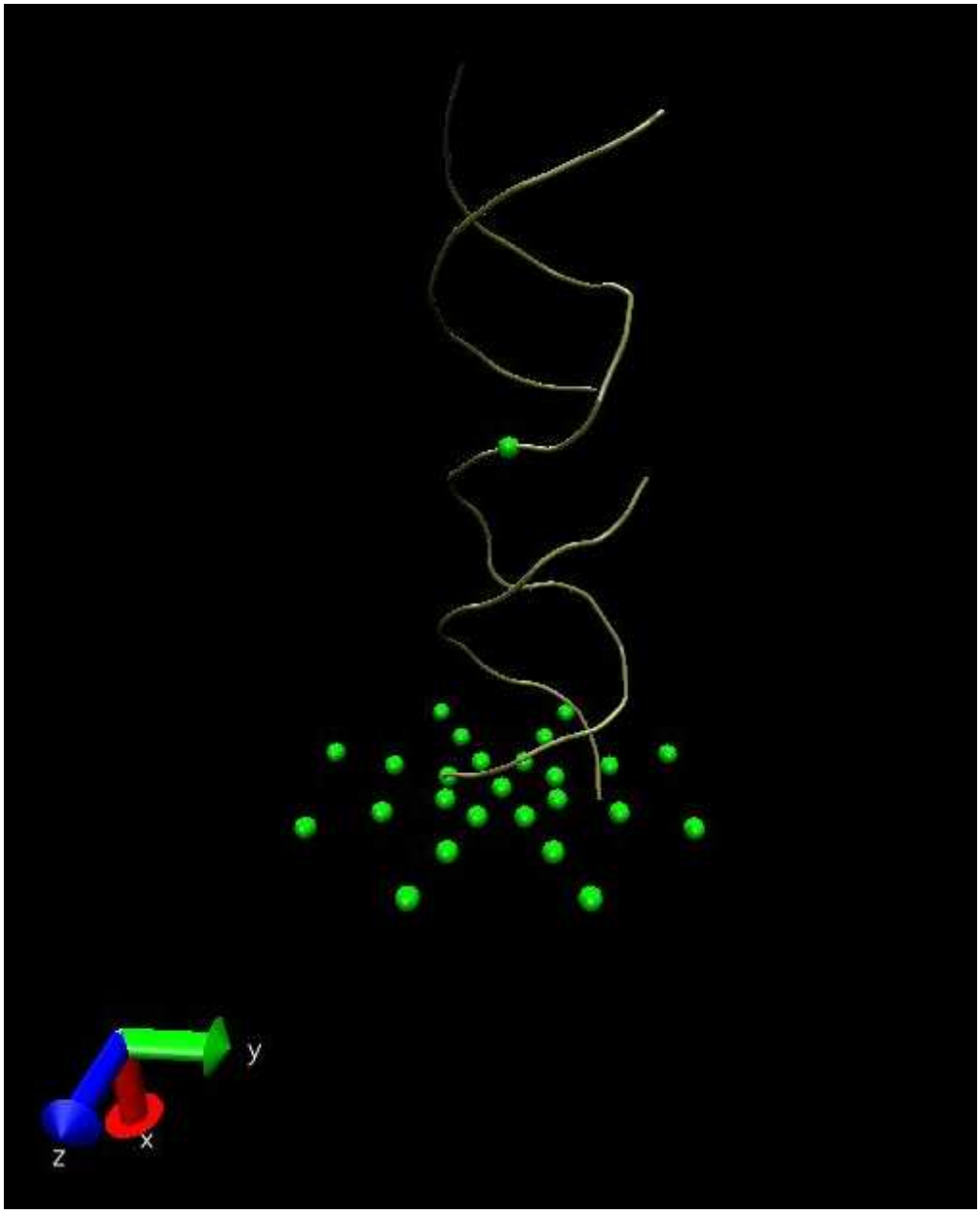

Figure 2.4: The points chosen for the axis of DNA rotations. All points are shown in green color. The top point is fixed, and the points at the bottom corresponds to the diffrent poinst for different systems.

Different axis of rotations along line 1 is shown in figure 2.5. In order to make the angle between two axes is 10^0, the second point is translated required amount of distance along corresponding line. Thus the other axis of rotations toward right and left are assigned. Similarly, the other second points along line 2, line 3 and line 4 are determined as shown in figure 2.6.

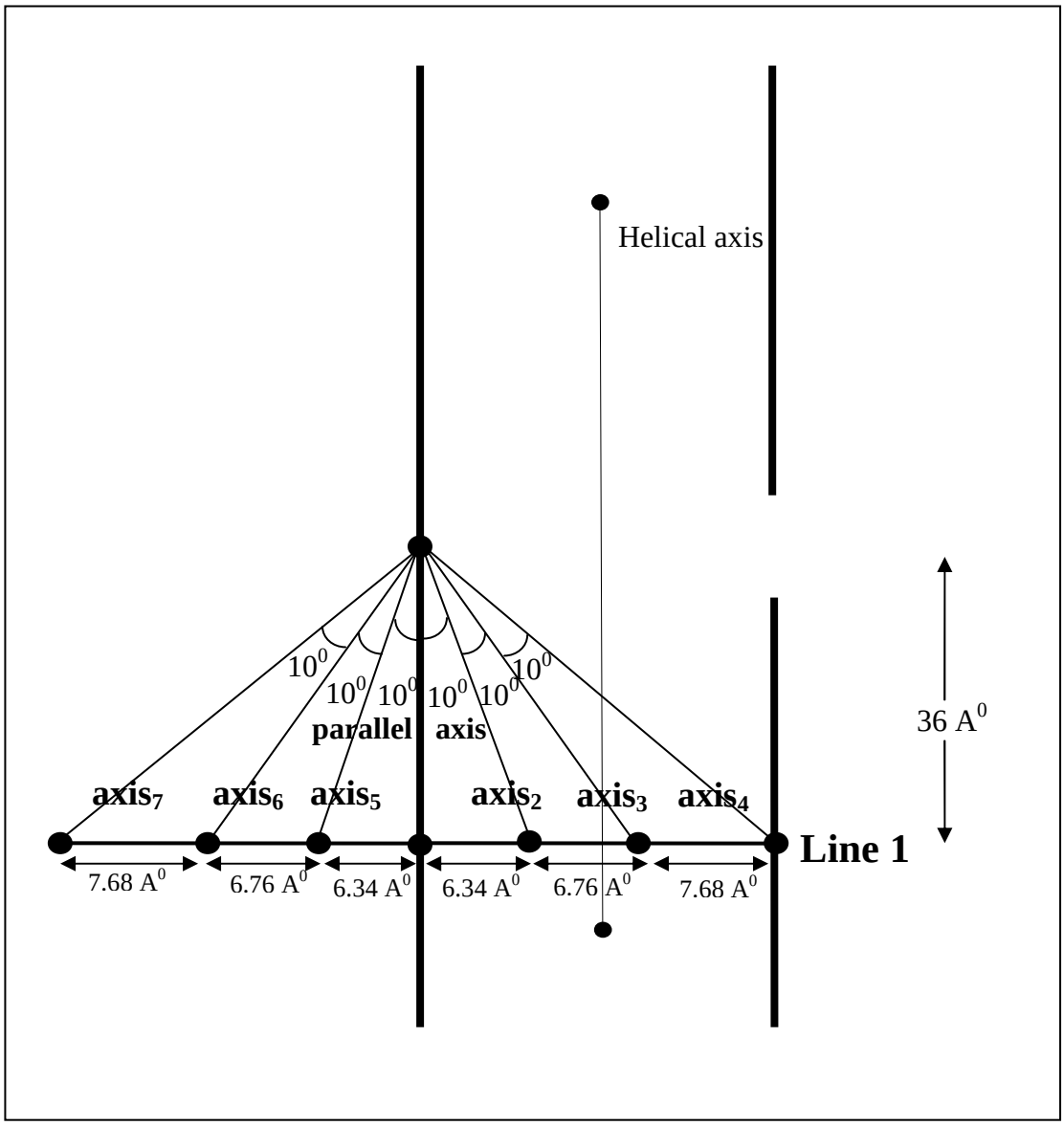

Figure 2.5 Axis of rotations along line 1.

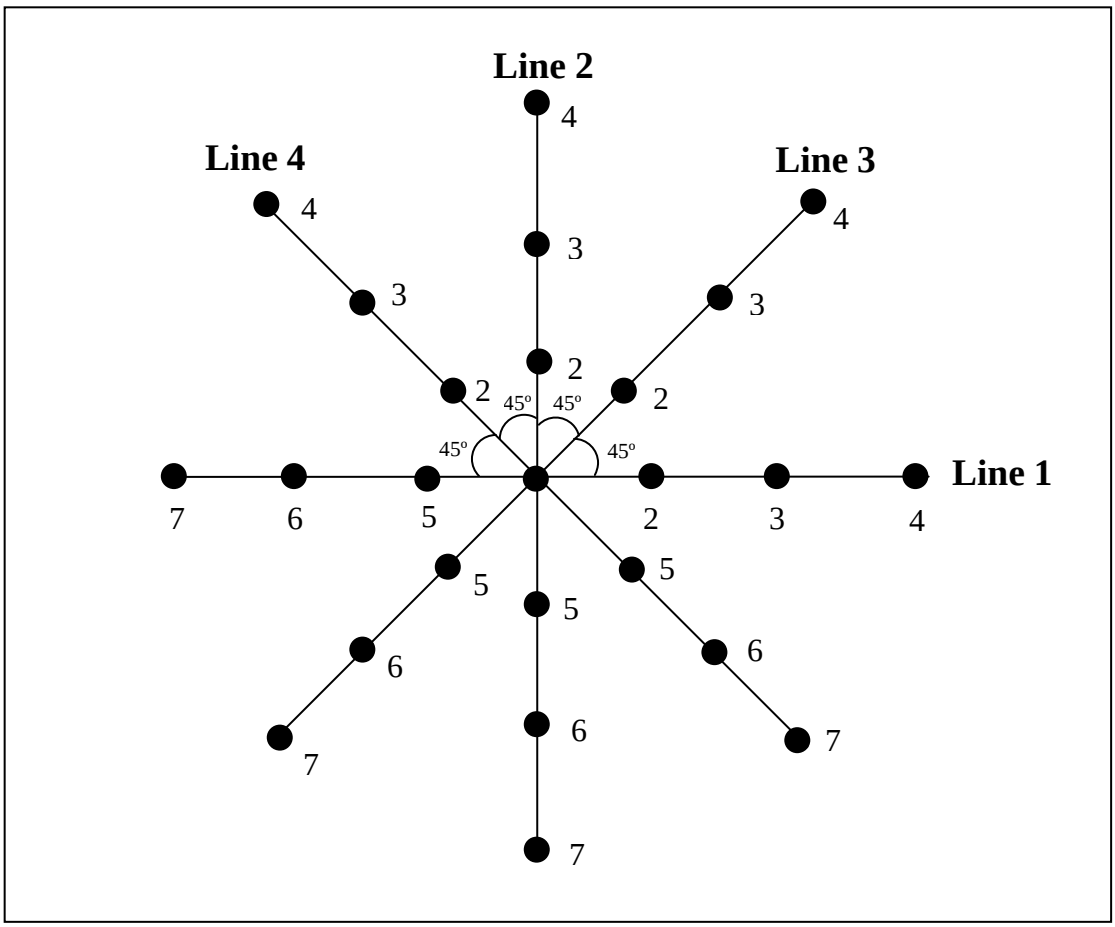

Figure 2.6 The second points that are chosen for the rotation axis along different lines.

To achieve the reference structure, rotation transformations are applied to the downstream backbone atoms of DNA (84 backbone atoms between -12 & -7 base pairs) in increments of 5º (Figure 2.7) until to reach full 360º, rotation both clockwise (positive rotations) and anti-clockwise (negative rotations). These coordinates have taken as a target coordinate to rotate 360º the downstream part of DNA around rotation axis using HQBMD. α_{hqbmd} is chosen as 1000 kcal/mol/A^4.

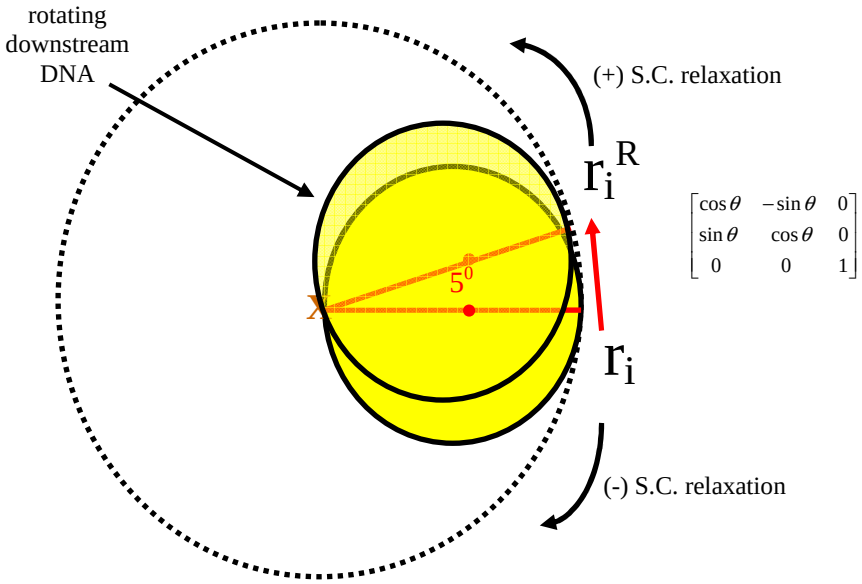

Figure 2.7 Our approach to DNA rotation.

CHAPTER 3

STRUCTURAL AND ENERGETIC ANALYSES

We have investigated the structural and energetic pathways of the rotations of a nicked DNA. Therefore, we have analyzed the results of the simulations in two different aspects; structural and energetic. First, these results will be given separately, and then a discussion section will be included.

25 different axis of rotation have been studied. One of them is chosen to be parallel to the helical axis. Then, a perpendicular vector to this axis as we called line 1 is found. This vector is rotated 90^0 by using rotation matrix to find the line 2. Again rotating line 1 and line 2 in 45^0, line 3 and line 4 are determined. These lines are on a plane that is perpendicular to the parallel axis of rotation, and the descriptions of these lines are shown in figure 2.5 and figure 2.6.

As we look at from the downstream side, downstream DNA is rotated clockwise that results in producing positive supercoils, and anti-clockwise that results in producing negative supercoils. Therefore, all of the analysis we have done are repeated for both kinds of rotations. From now on, we will use the term '**positive rotations**' for the clockwise DNA rotations (as looked from the downstream side), and the term '**negative rotations**' is going to be used for the anti-clockwise rotations.

3.1 STRUCTURAL ANALYSIS

At the present study we are trying to find most feasible axis of rotation by investigation the five backbone dihedral angles that is opposite of the nicked site structurally and energetically. Change of dihedral angles γ ($C_{3'}$– $C_{4'}$ – $C_{5'}$ – $O_{5'}$), β ($C_{4'}$ – $C_{5'}$ – $O_{5'}$ – P), α ($C_{5'}$ – $O_{5'}$ – P – $O_{3'}$), ζ ($O_{5'}$ – P – $O_{3'}$–$C_{3'}$), and ε (P – $O_{3'}$ – $C_{3'}$ – $C_{4'}$) are determined as a function of downstream rotation angle. These angles are shown in Figure 3.1.

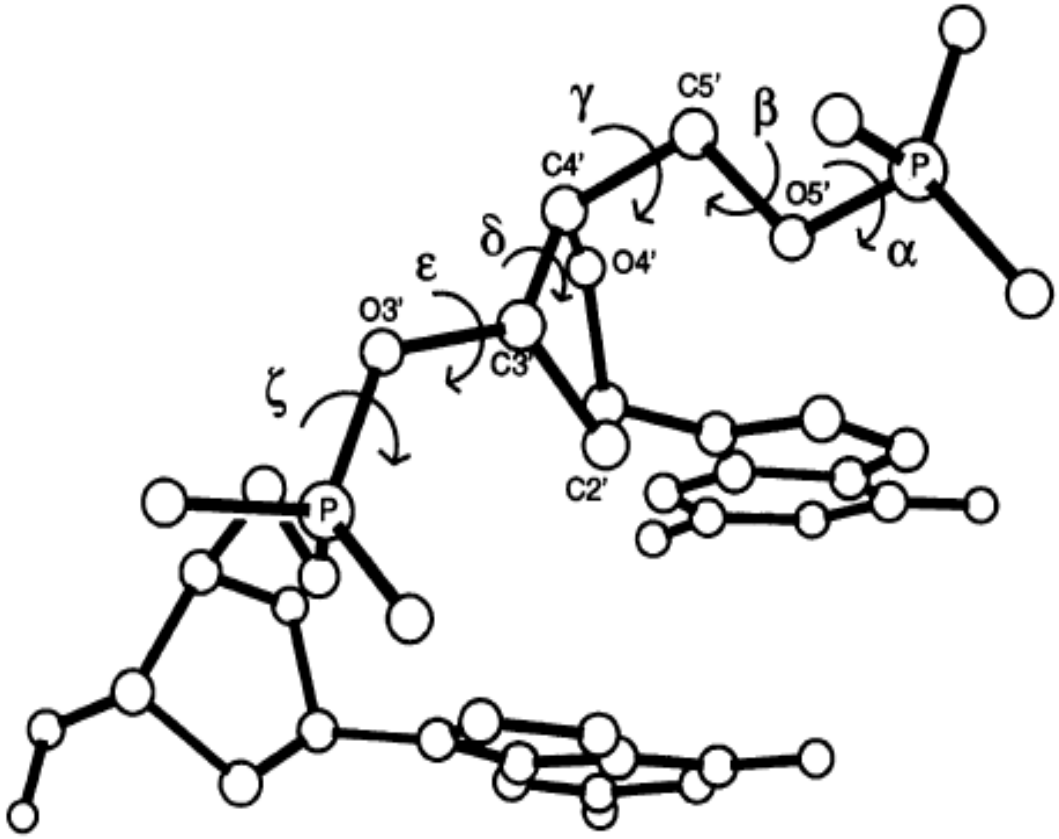

Figure 3.1 Backbone dihedral or torsional angles [28]

Amount of changes in the backbone dihedrals (that are mentioned above) upon DNA rotations are plotted in Figure 3.2 for parallel axis.

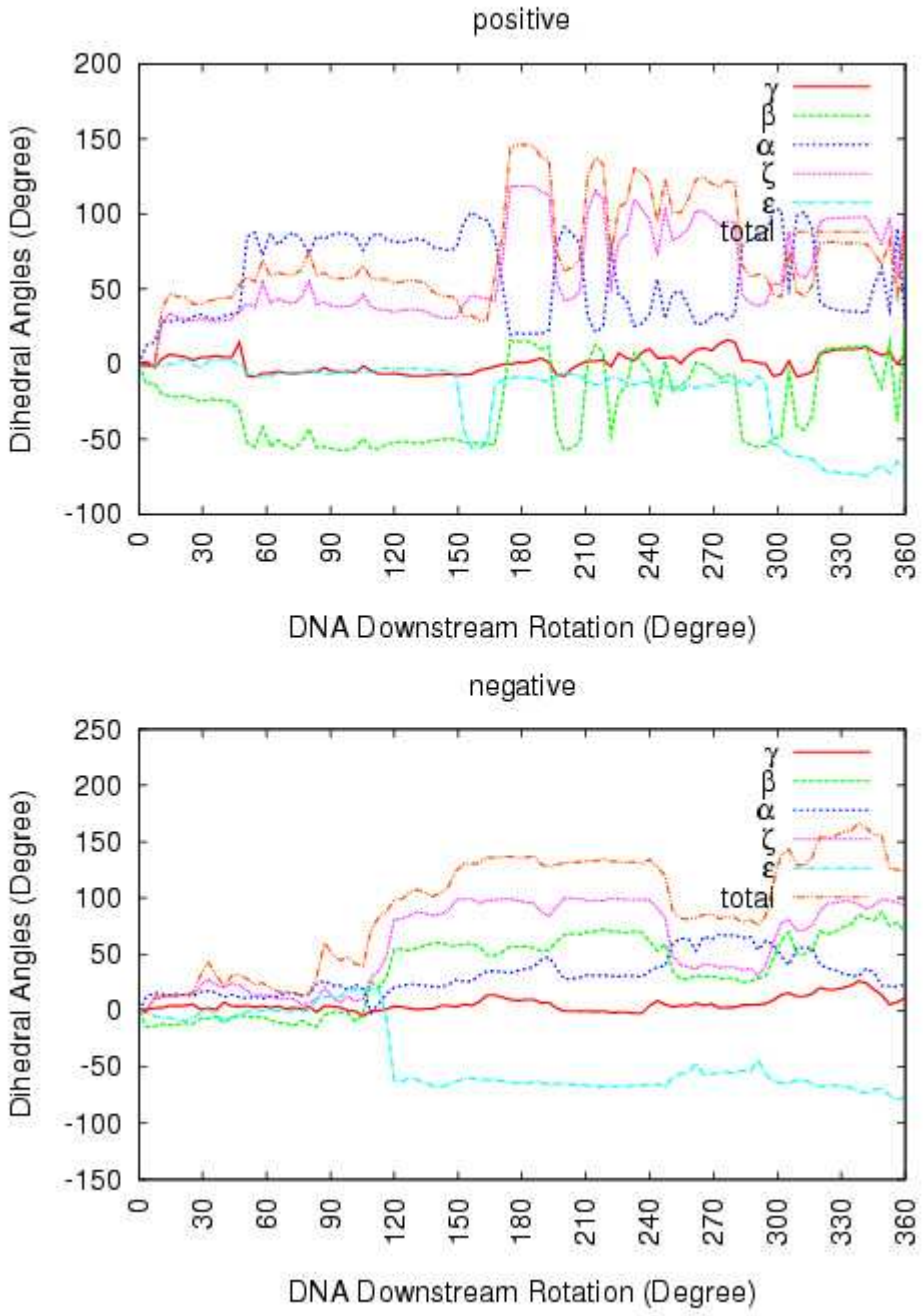

Figure 3.2 Changes in backbone dihedral angles along the parallel axis. The top figure corresponds to the *positve rotations*, and the bottom one corresponds to the *negative rotations.*

The corresponding results for $Axis_2$, $Axis_3$, $Axis_4$, $Axis_5$, $Axis_6$, and $Axis_7$ are given in Figure 3.3 for the *negative rotations* along line 1, in Figure 3.4 for line 2, in Figure 3.5 for line 3, and in Figure 3.6 for line 4.

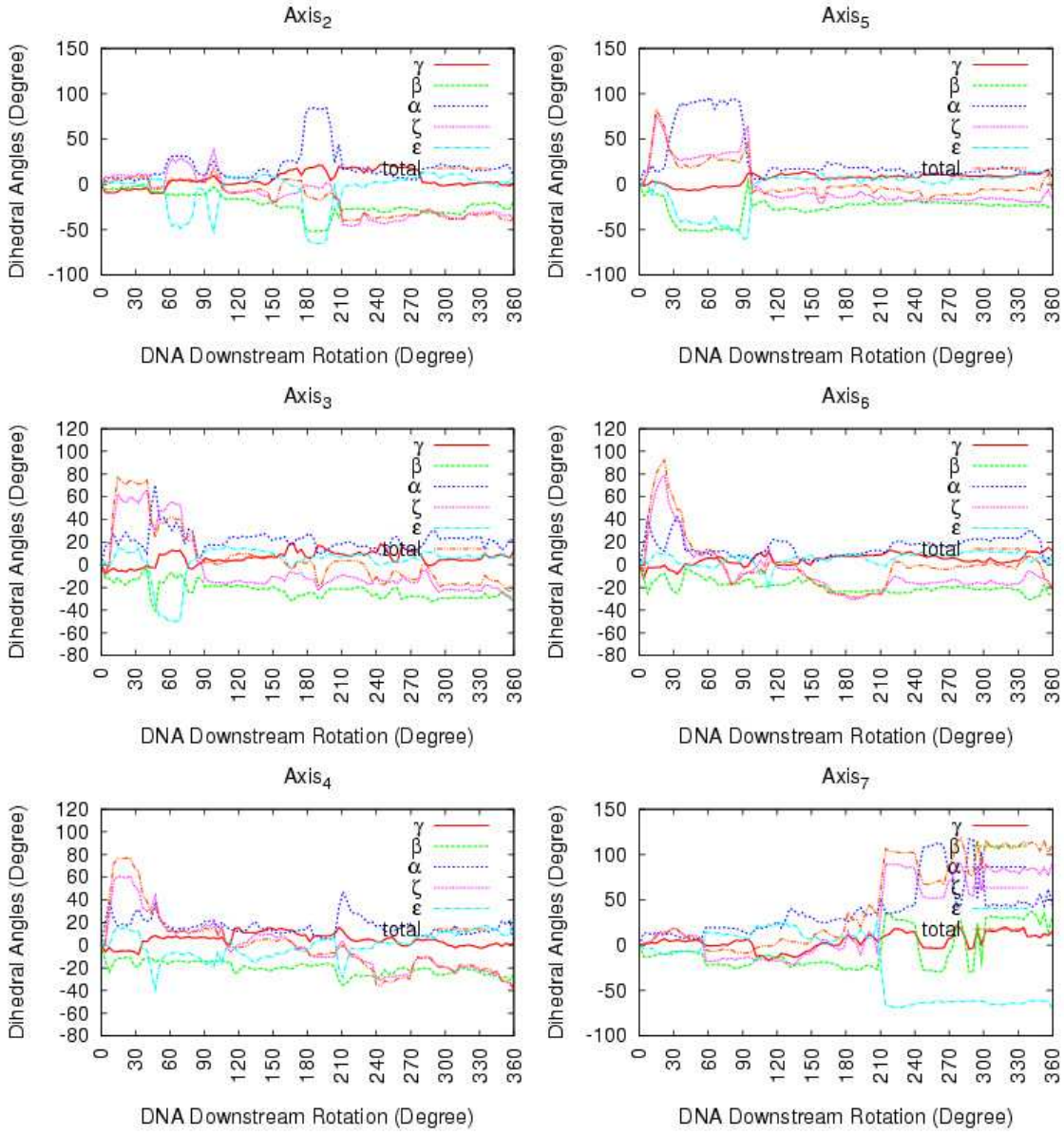

Figure 3.3 Changes in backbone dihedral angles for *negative rotations* along line 1.

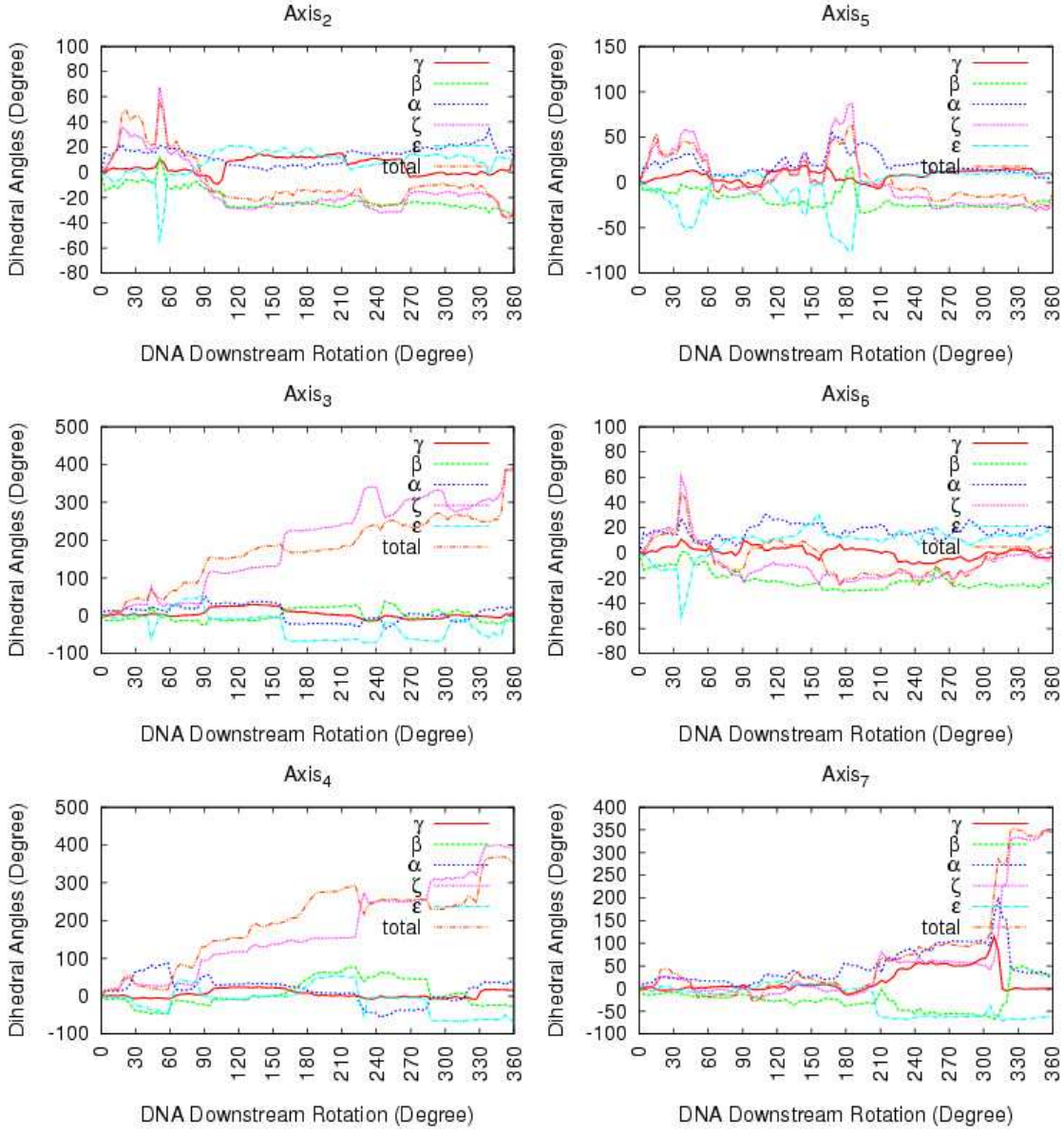

Figure 3.4 Changes in backbone dihedral angles for *negative rotations* along line 2.

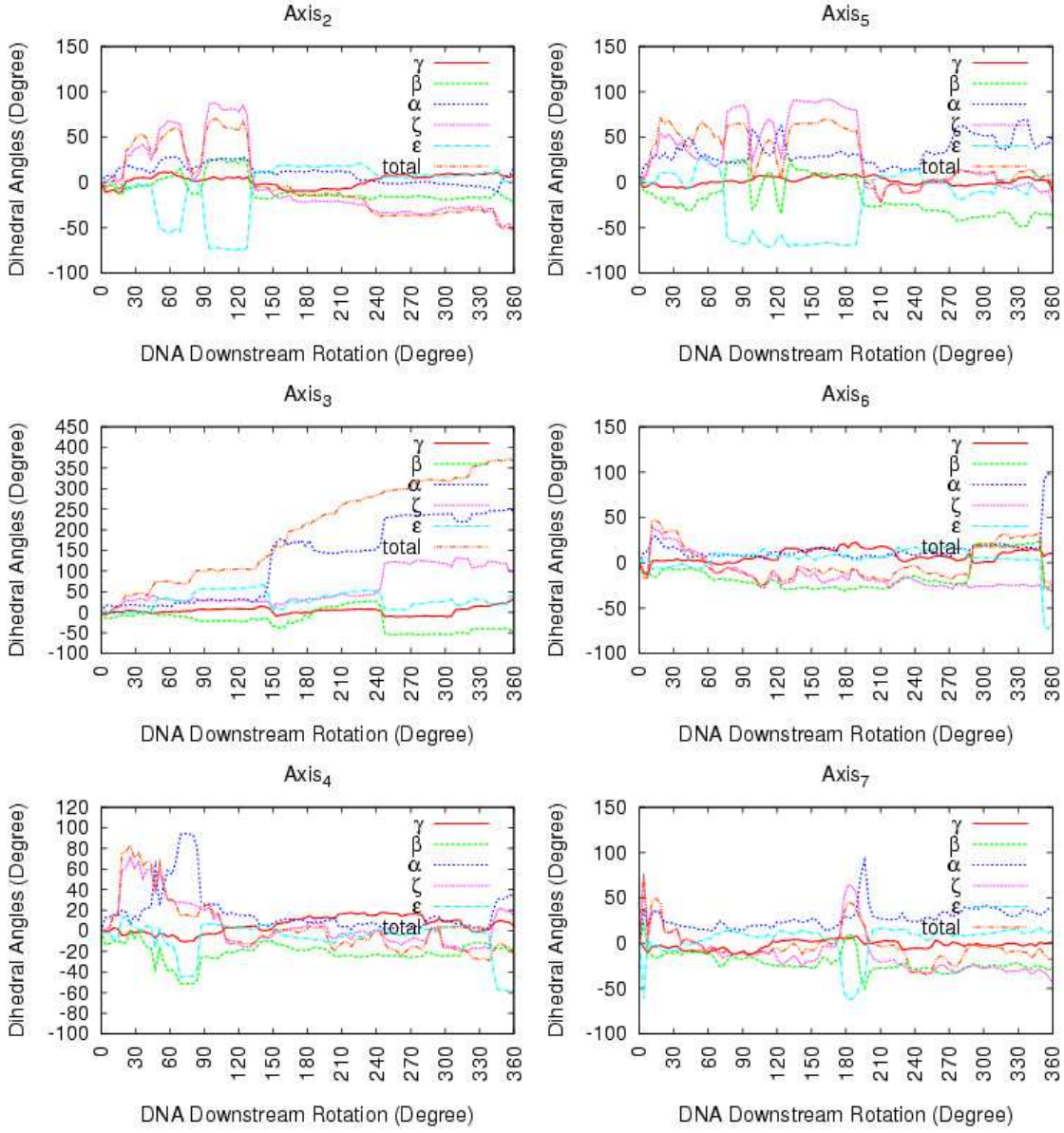

Figure 3.5 Changes in backbone dihedral angles for negative *rotations along* line 3.

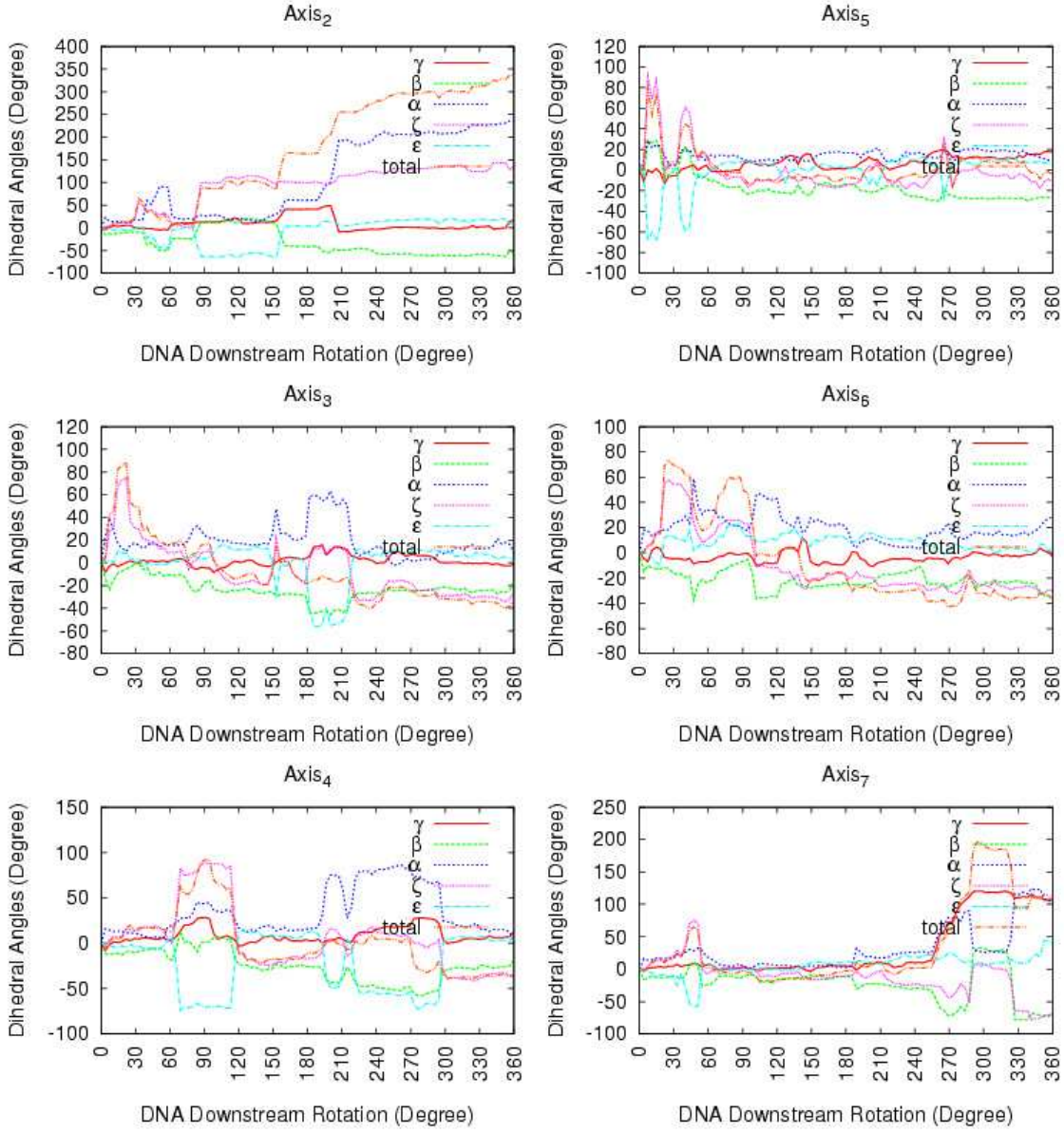

Figure 3.6 Changes in backbone dihedral angles for *negative rotations* along line 4.

The results for $Axis_2$, $Axis_3$, $Axis_4$, $Axis_5$, $Axis_6$, and $Axis_7$ are give in Figure 3.7 for the *positive rotations* along line 1, in Figure 3.8 for line 2, in Figure 3.9 for line 3, and in Figure 3.10 for line 4.

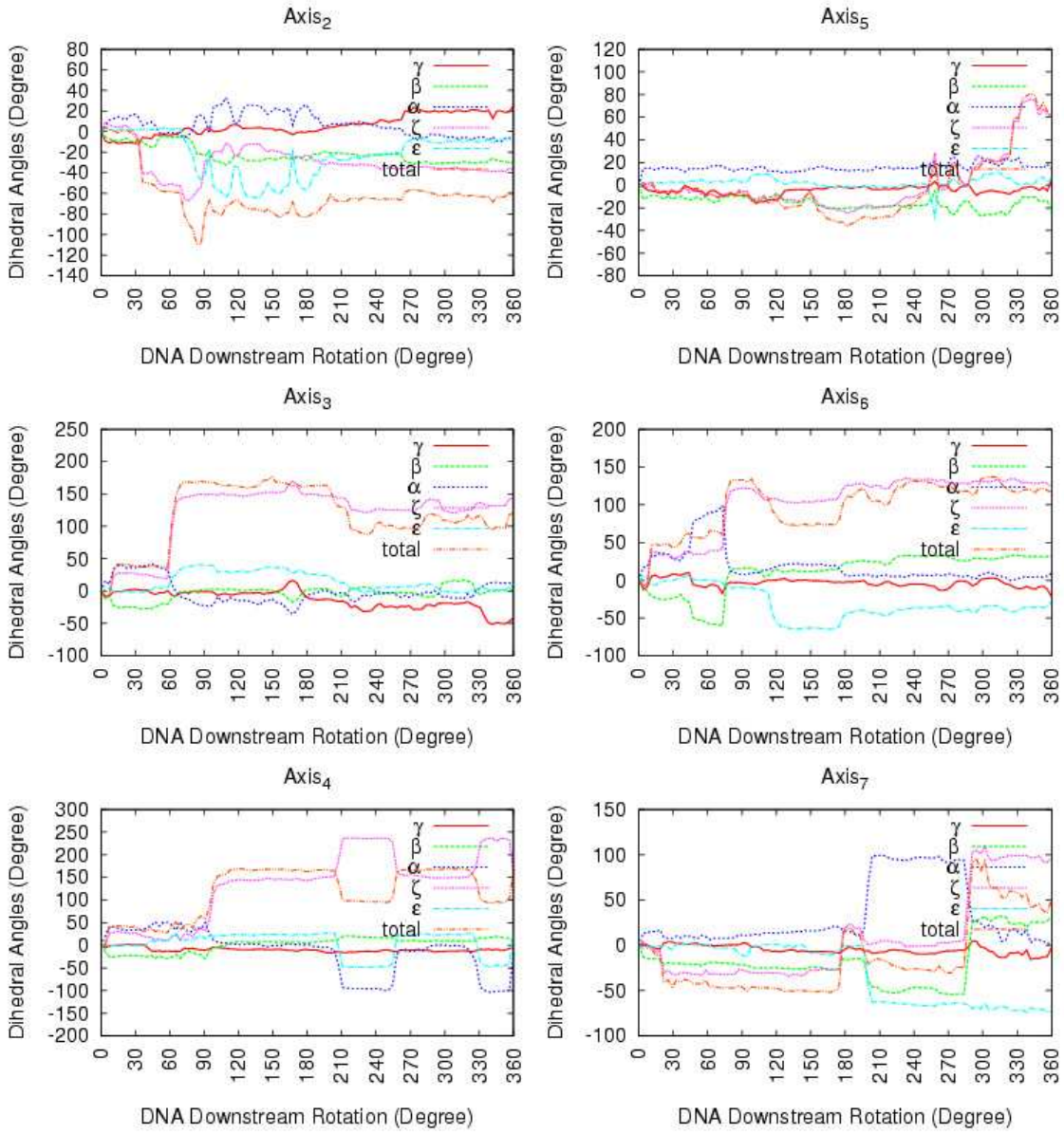

Figure 3.7 Changes in backbone dihedral angles for *positive rotations* along line 1.

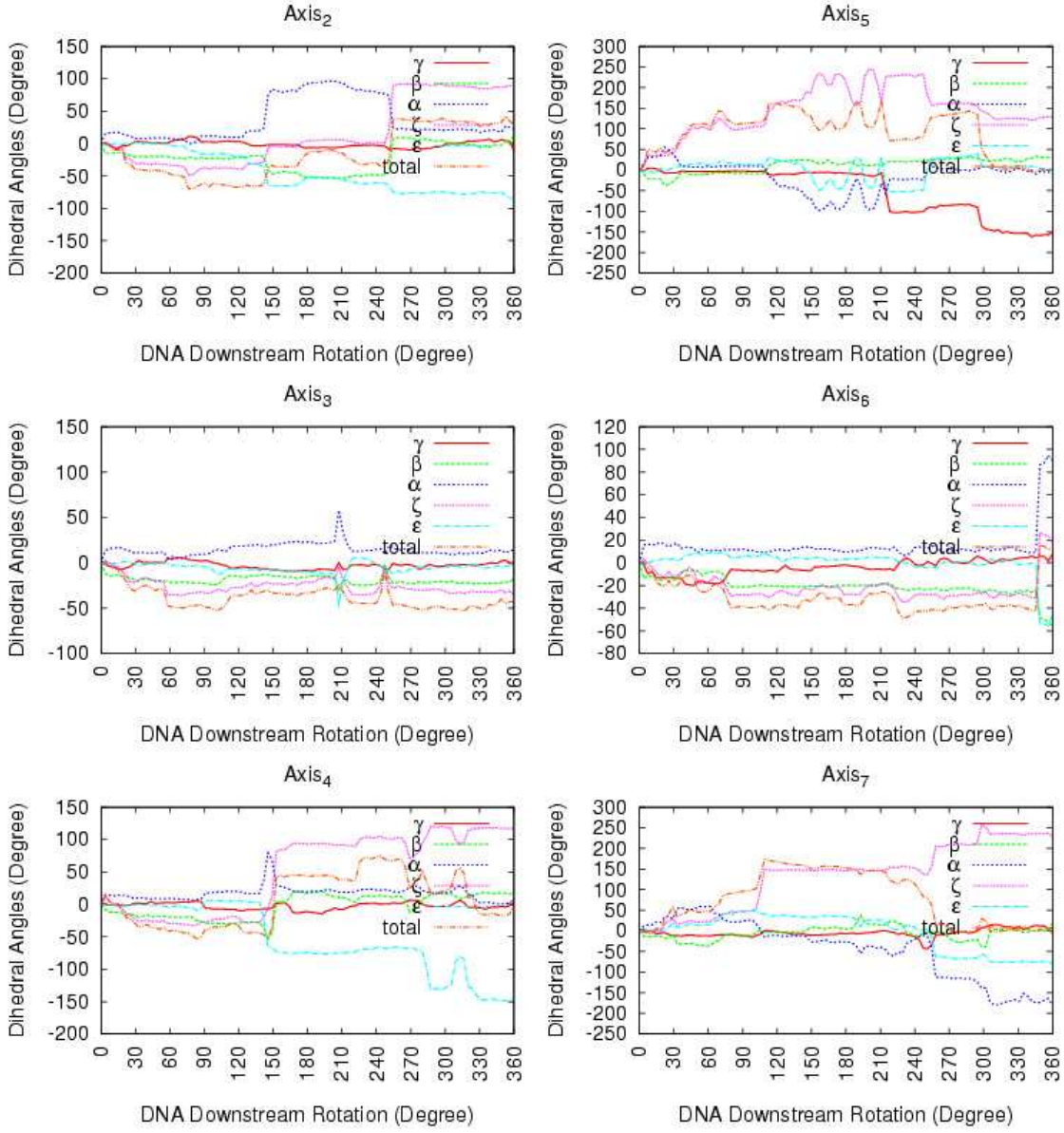

Figure 3.8 Changes in backbone dihedral angles for *positive rotations* along line 2.

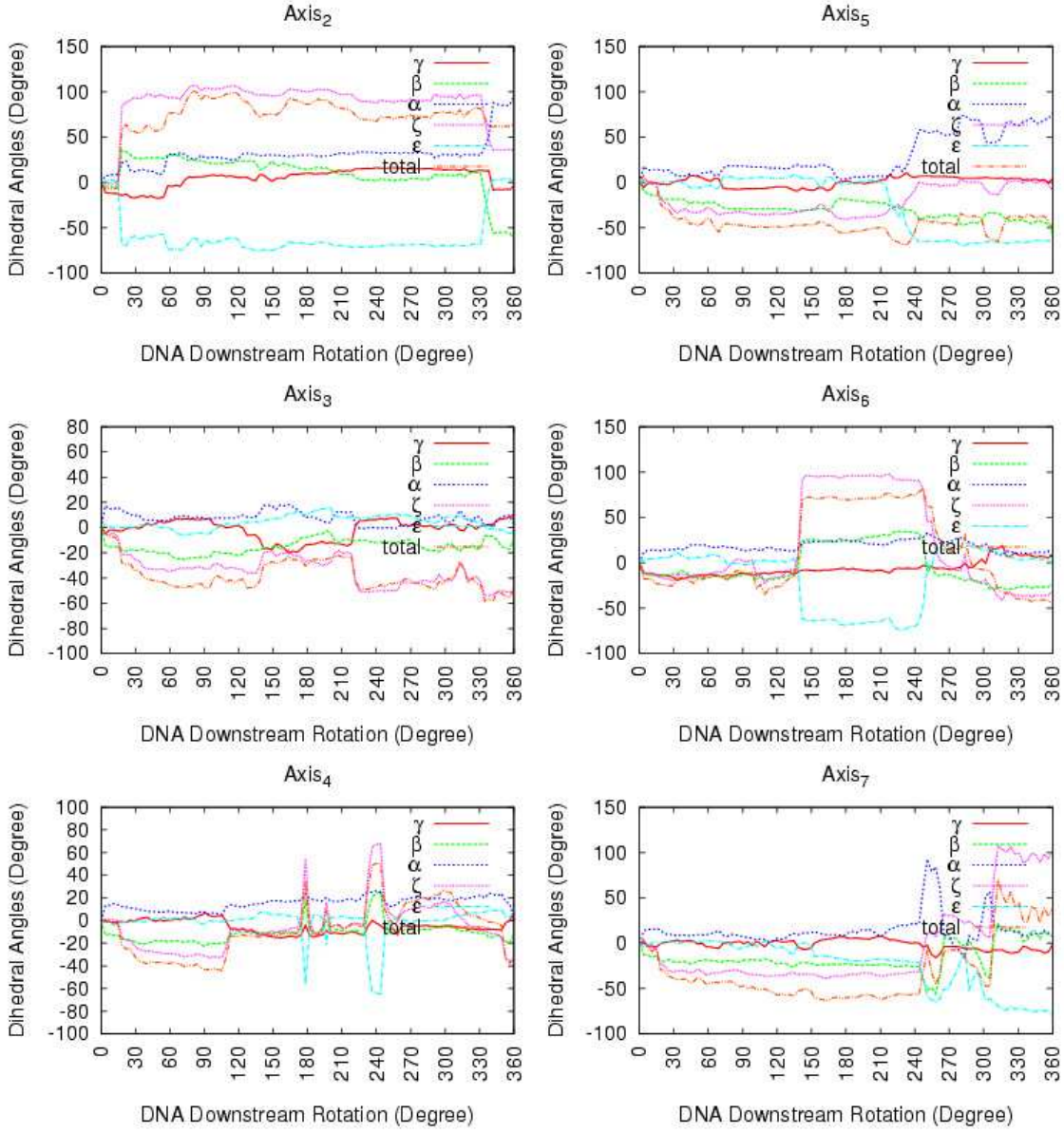

Figure 3.9 Changes in backbone dihedral angles for *positive rotations* along line 3.

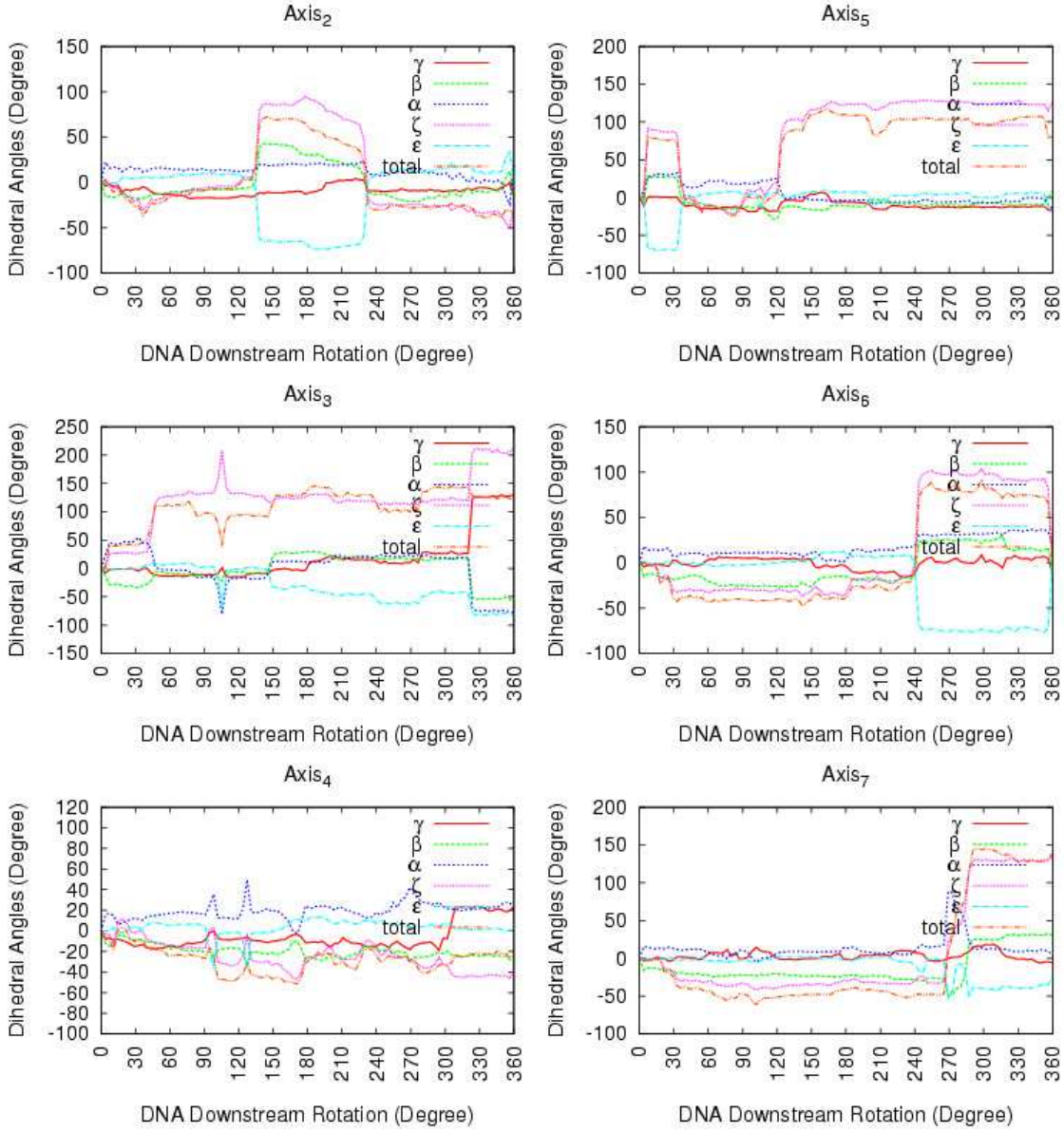

Figure 3.10 Changes in backbone dihedral angles for *positive rotations* along line 4.

The backbone dihedral angles **α** and **γ** has been extensively studied for intact B-DNA. Varnai *et al.* [29] have found that the sum of these two angles is found to minimum around 0 degree, on the free energy profile. Their study was on the intact DNA. Therefore, for a nicked DNA, we have calculated the sum of these two dihedrals, as a function of DNA rotation. The current calculated values represent the first data on the nicked DNA.

The results for *negative* and *positive rotations* along parallel axis are shown in Figure 3.11, and the corresponding results along line 1, line 2, line 3 and line 4 are given in Figure 3.12 for *negative rotations*, in Figure 3.13 for *positive rotations.*

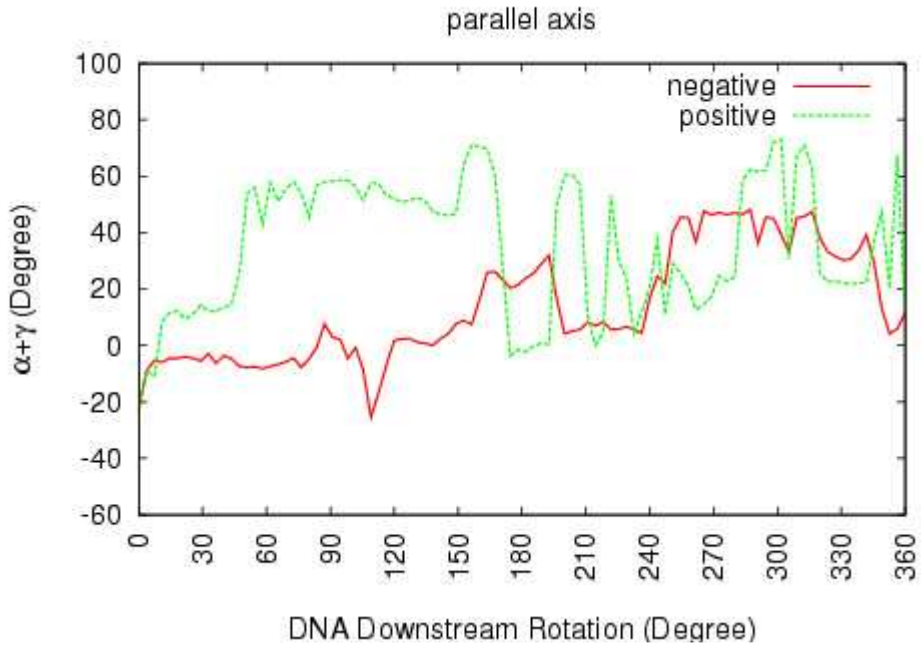

Figure 3.11 $\alpha+\gamma$ as a function of rotation angle for *positive* and *negative rotations* along the parallel axis.

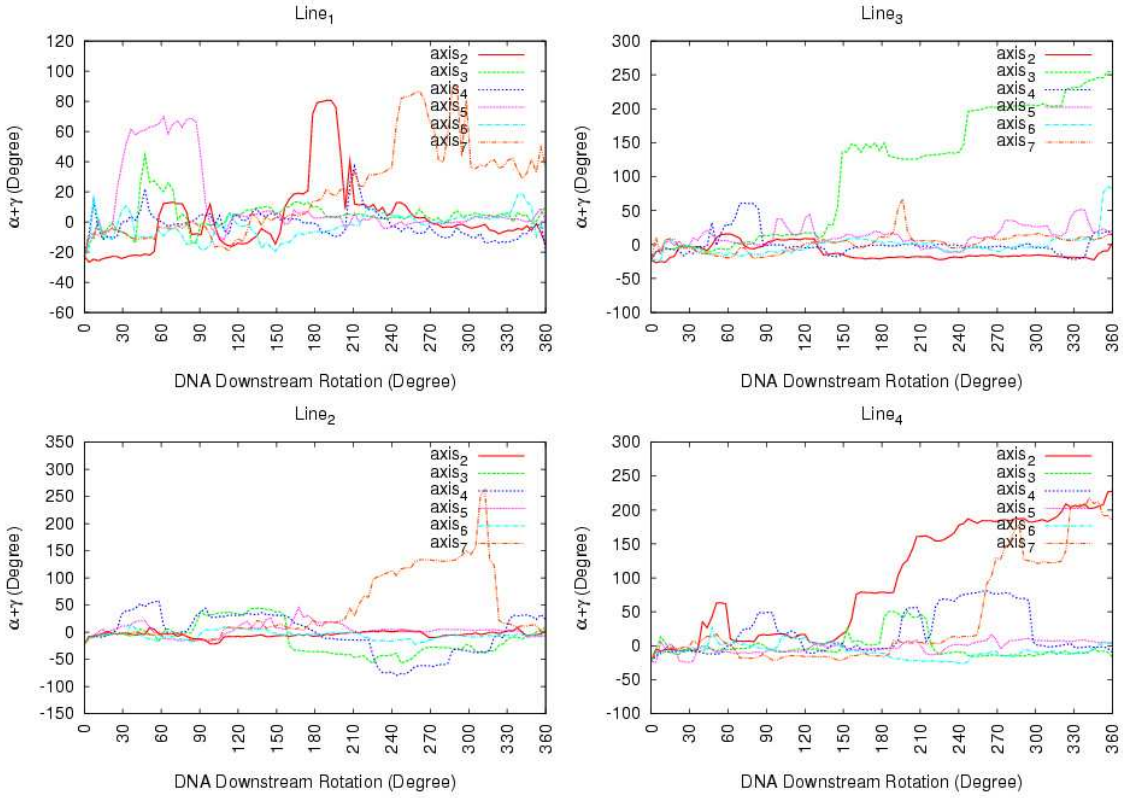

Figure 3.12 $\alpha+\gamma$ as a function of rotation angle for *negative rotations*.

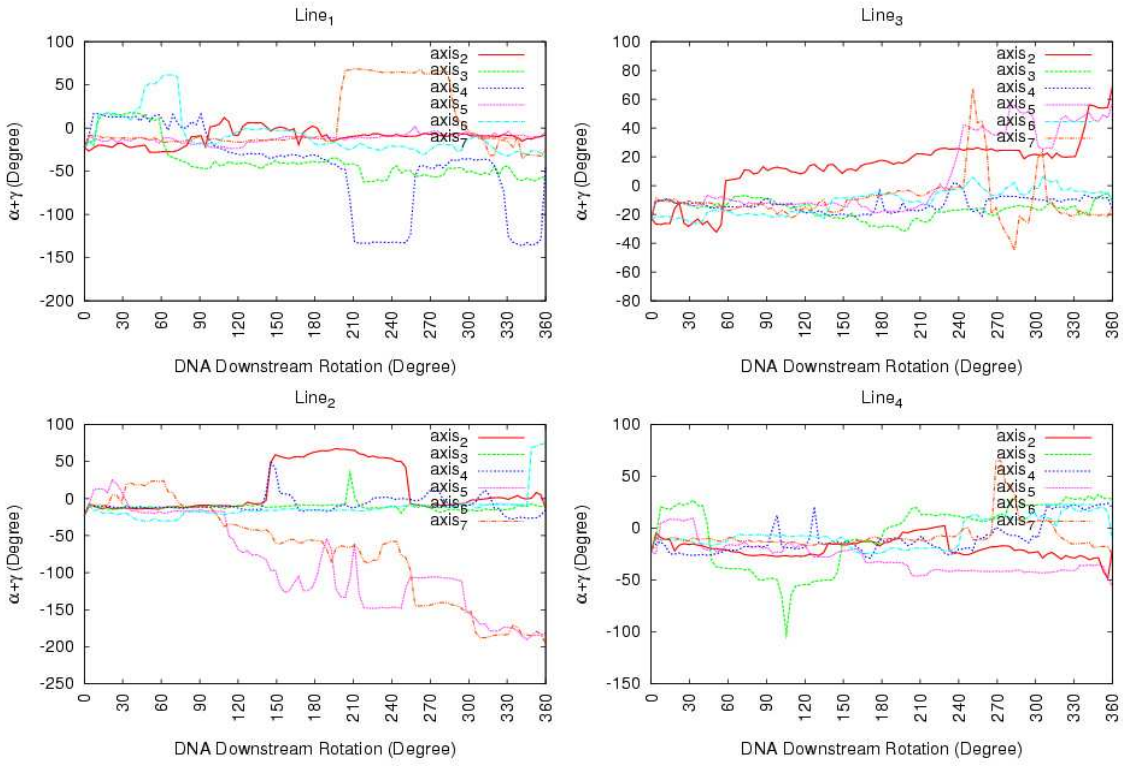

Figure 3.13 $\alpha+\gamma$ as a function of rotation angle for *positive rotations*.

The dihedrals **ε** and ζ are also a focus of interest in many studies. Hartmann *et al*. [30] extensively studied the form of $\varepsilon - \zeta$ for a B-DNA. They reported two main distinct state for B-DNA, as a function of this difference. The 'B_I' state that they called, has a difference of angles around -90, and the other possible state, as called 'B_{II}' state, has allowed angles around +90. They also reported the free energy profiles as a function of the difference. Thus, we have analyzed the difference of such dihedrals as we rotated the downstream DNA. The results are given in Figure 3.14 for the rotations around parallel axis for *positive* and *negative rotations*, in Figure 3.15 for *negative rotations*, in Figure 3.16 for *positive rotations* along line 1, line 2, line 3 and line 4.

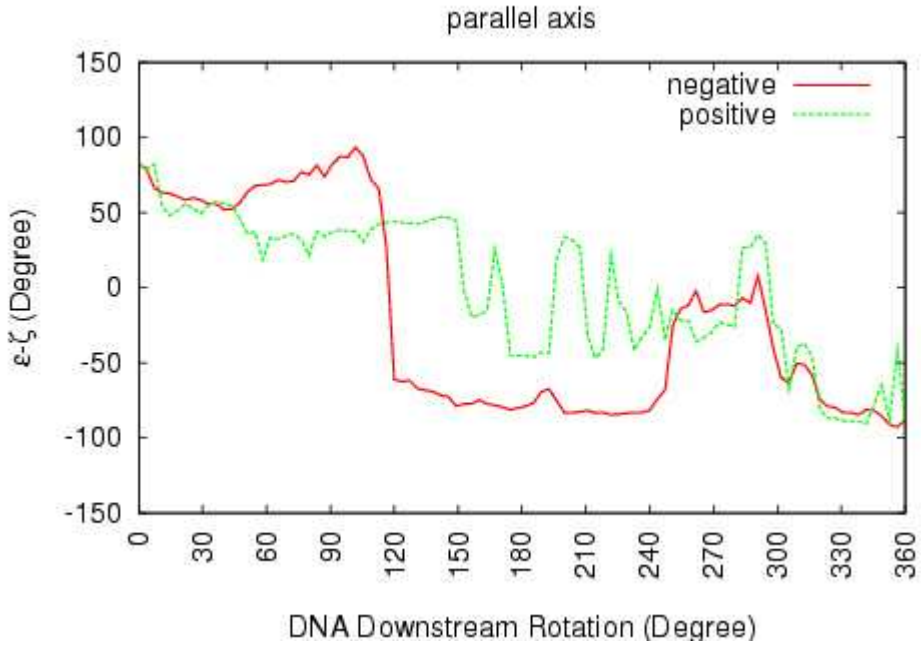

Figure 3.14 $\varepsilon - \zeta$ as a function of rotation angle for *positive* and *negative rotations* along the parallel axis.

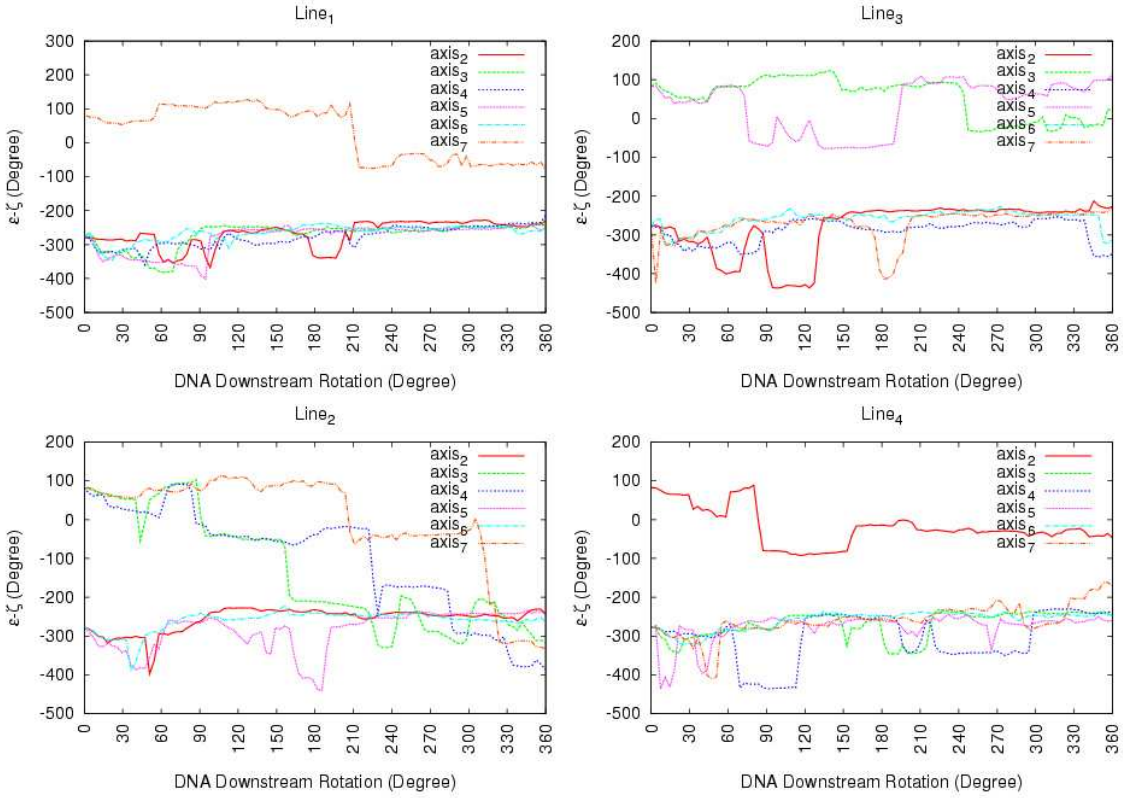

Figure 3.15 $\varepsilon - \zeta$ as a function of rotation angle for *negative rotations*.

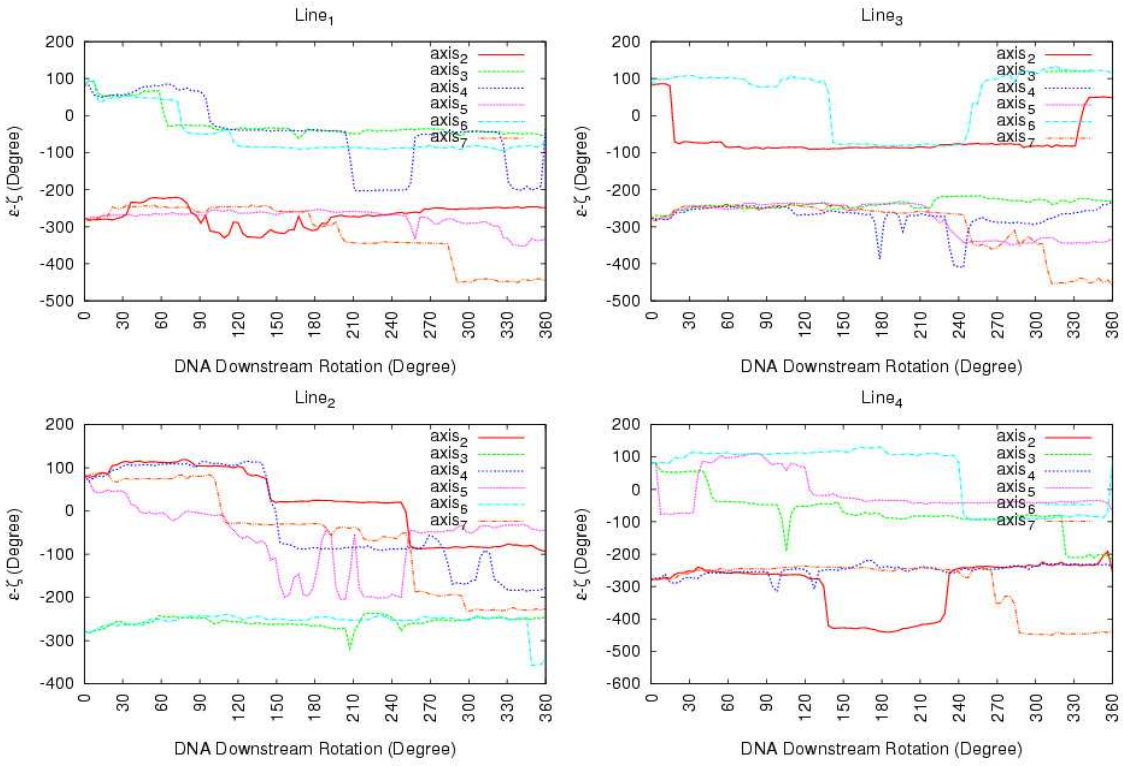

Figure 3.16 $\varepsilon - \zeta$ as a function of rotation angle for *positive rotations*.

The correlation between **ε** and ζ were studied by Hartmann *et al*. [30]. They show that all possible values for these two dihedral angle can be classified as in two different states. As they reported a nice correlation, we also looked for such states where these two dihedral angles are correlated. The states we got nicely include the states previously proposed states. However, we have also obtained additional possible states for these angles. The Table 3.1 summarizes the possible states we obtained for **ε** and ζ. The values of such states are also given in figure 3.17 for *positive* and *negative rotations* along parallel axis, in Figure 3.18, 3.19, 3.20 and 3.21 for *negative rotations*, and in Figure 3.22, 3.23, 3.24 and 3.25 for *positive rotations* along line 1, line 2, line 3 and line 4 respectively.

Table 3.1 Possible states for **ε** and ζ

	ε	ζ
1	-170 ± 30	160 ± 50
2	-100 ± 40	250 ± 40
3	-250 ± 30	210 ± 20
4	-60 ± 25	300 ± 40
5		50 ± 50
6		130 ± 40

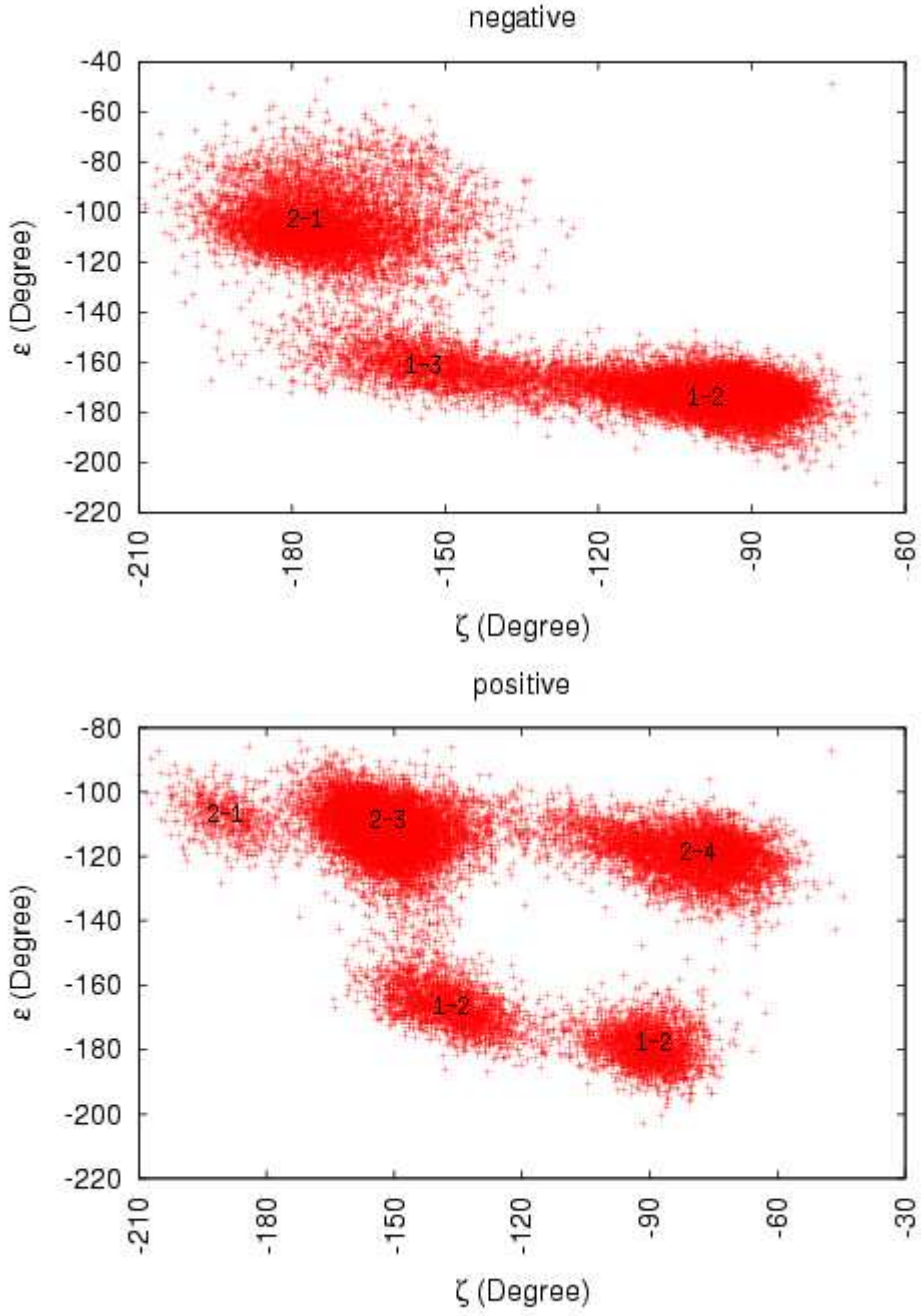

Figure 3.17 Correlation between dihedral angle **ε** and **ζ** for *negative and positive rotations* along the parallel axis.

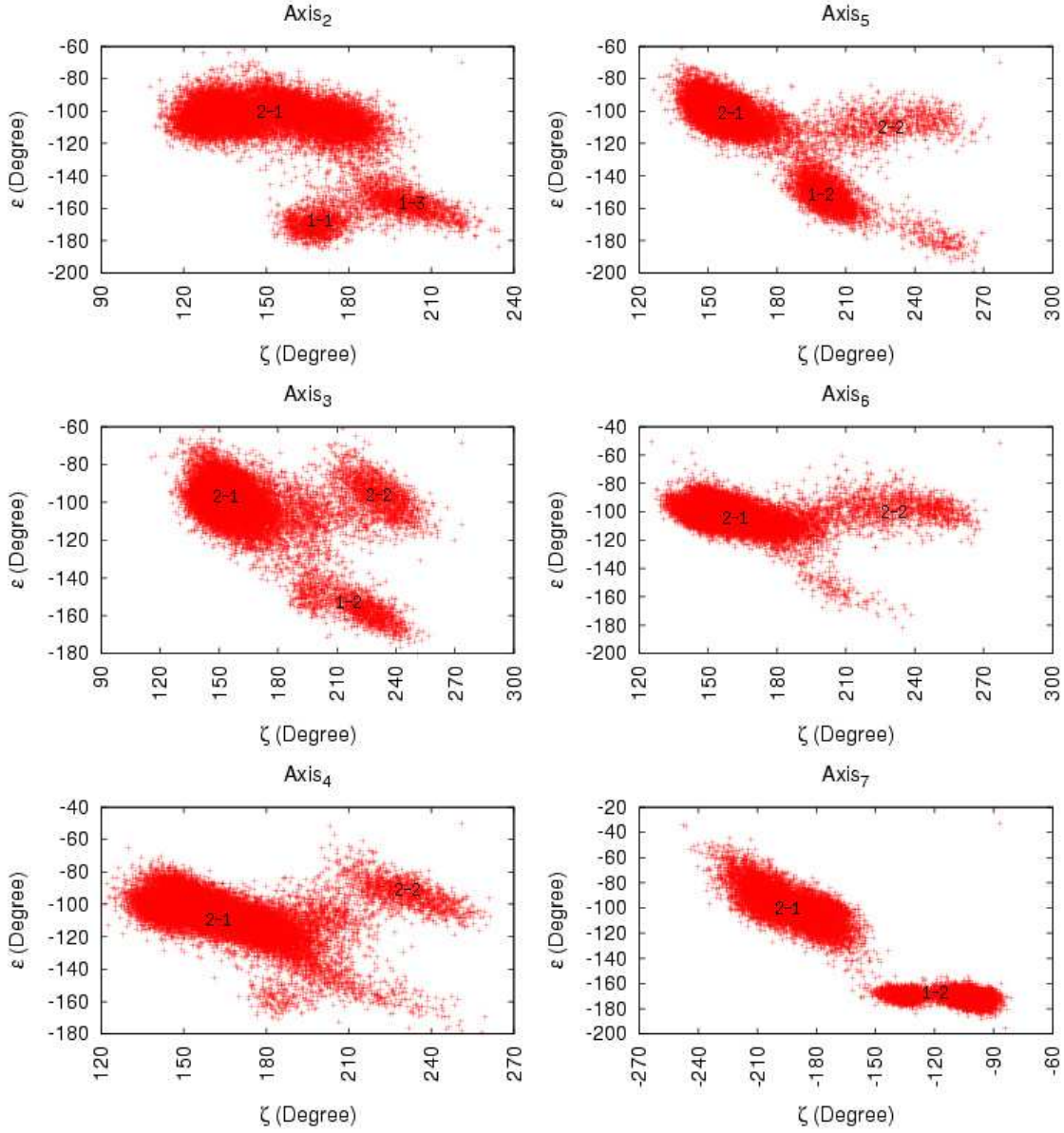

Figure 3.18 Correlation between dihedral angle **ε** and **ζ** for *negative rotations* along line 1.

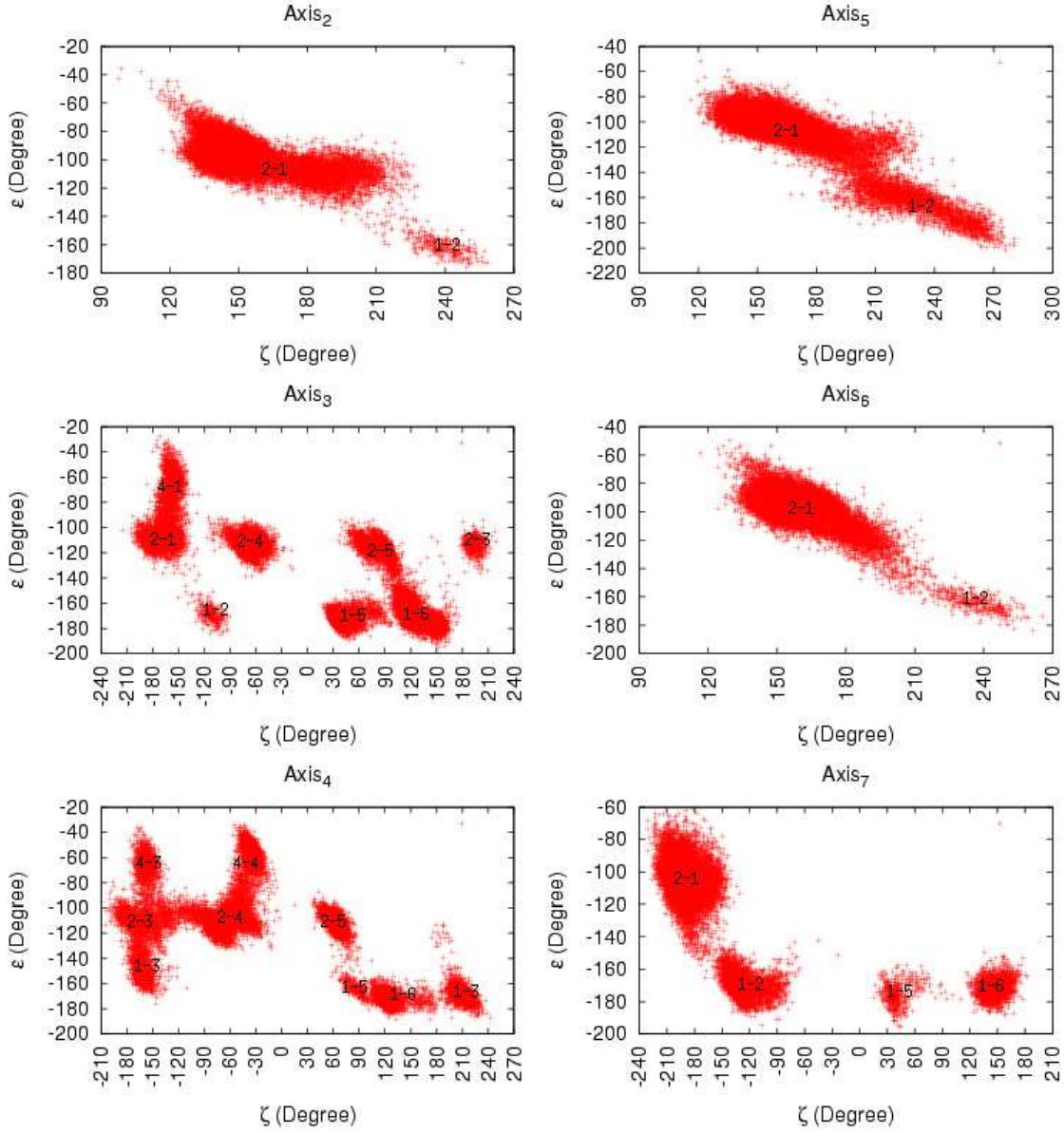

Figure 3.19 Correlation between dihedral angle **ε** and **ζ** for *negative rotations* along line 2.

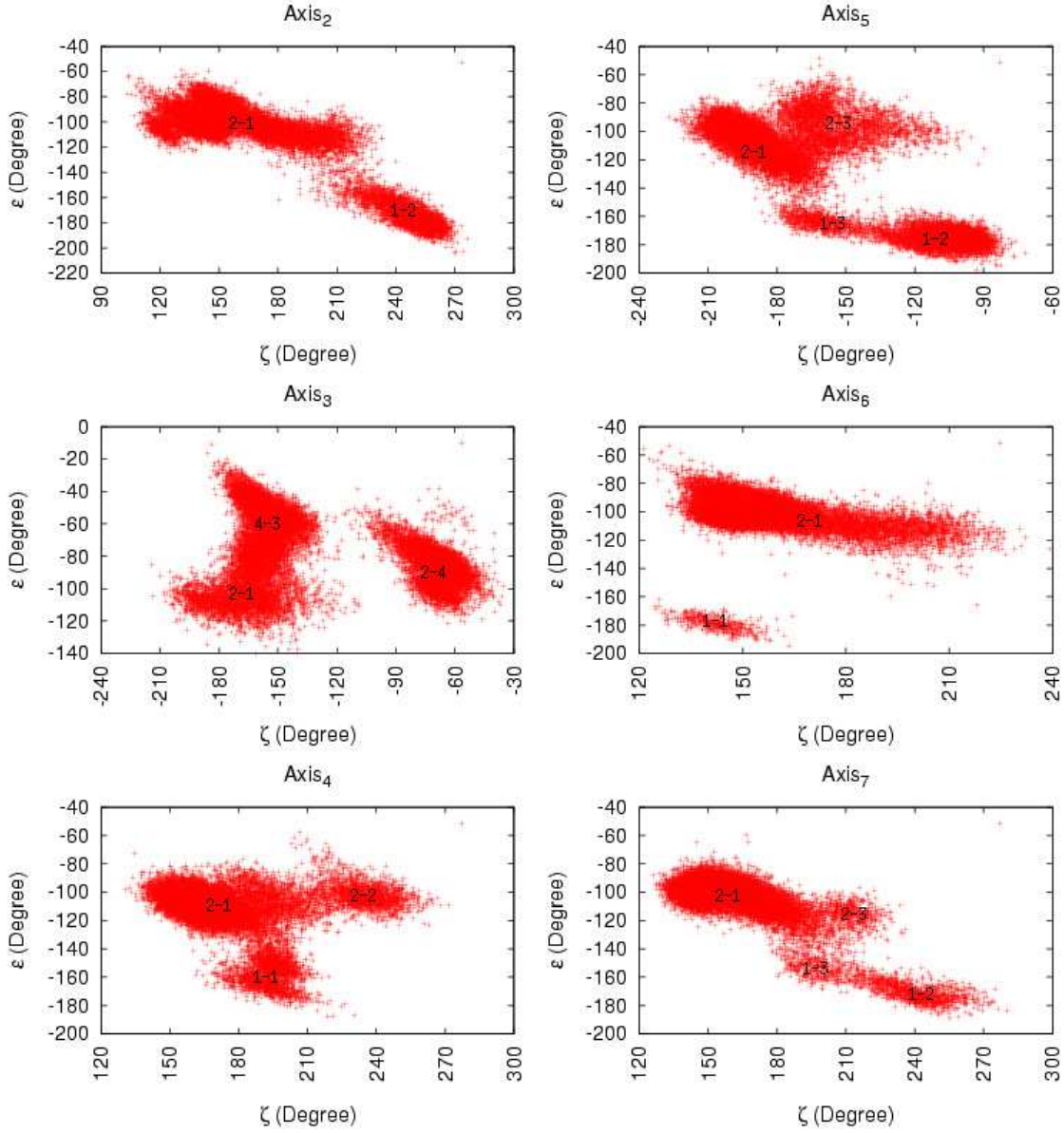

Figure 3.20 Correlation between dihedral angle **ε** and **ζ** for *negative rotations* along line 3.

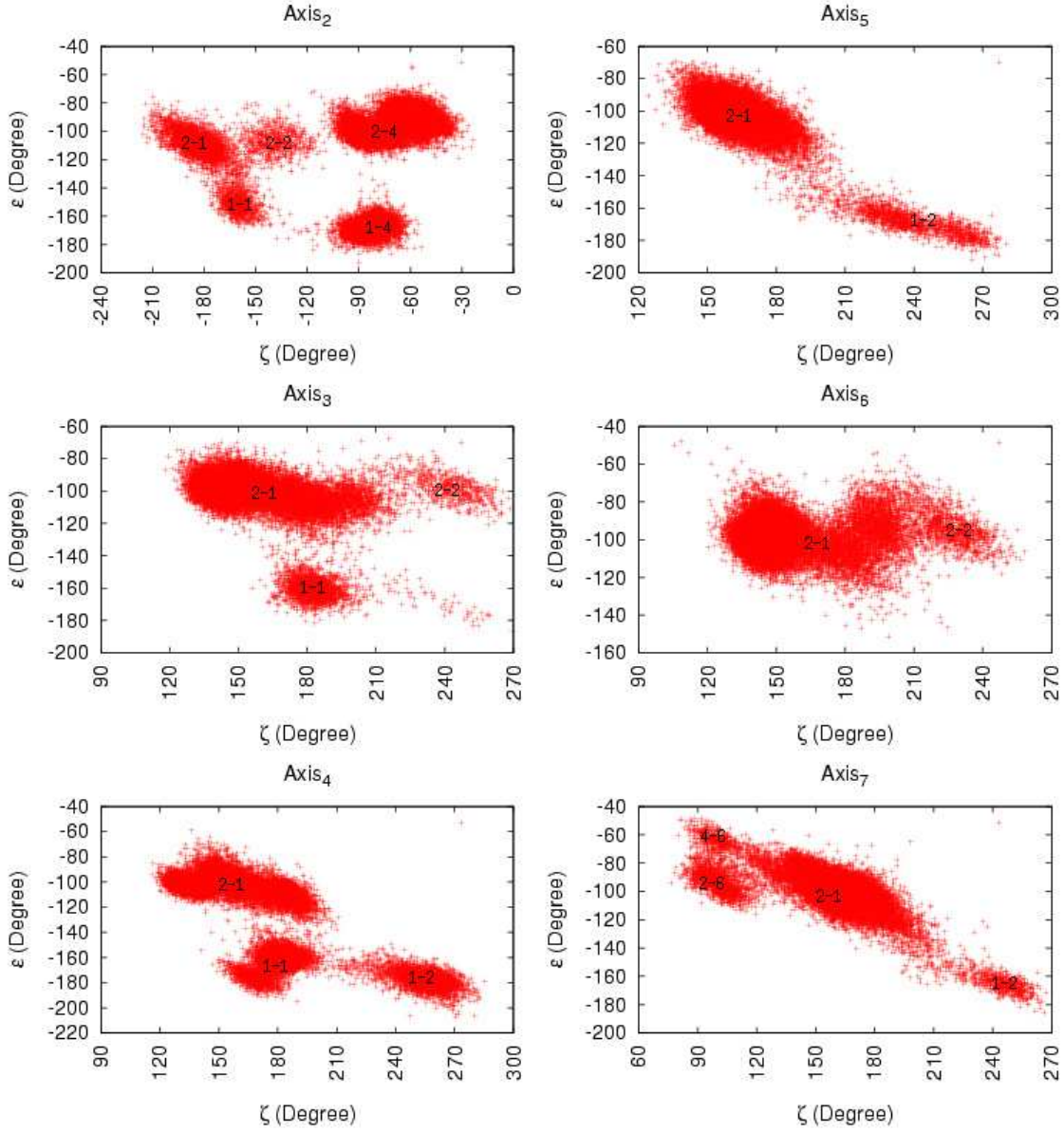

Figure 3.21 Correlation between dihedral angle **ε** and **ζ** for *negative rotations* along line 4.

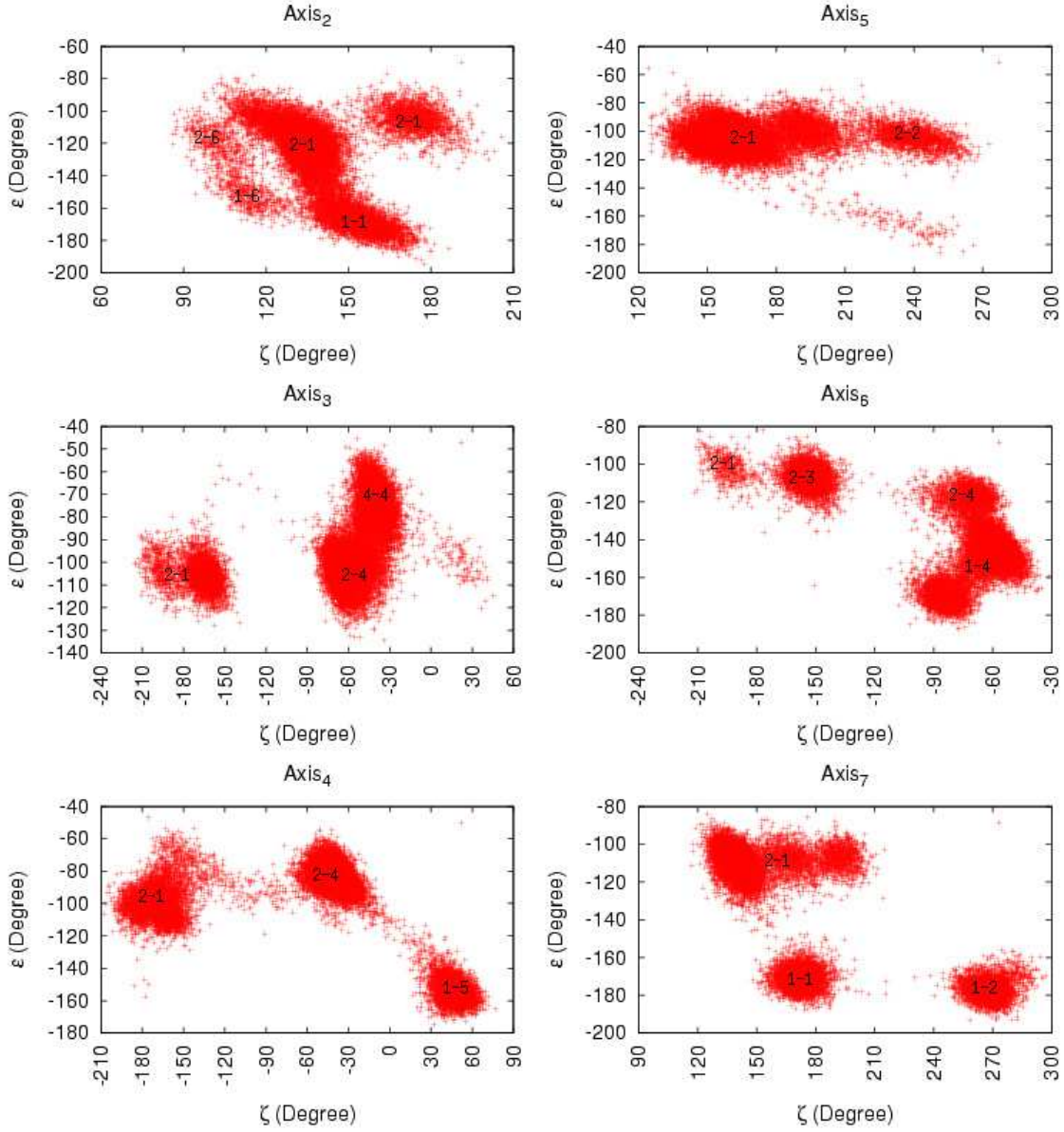

Figure 3.22 Correlation between dihedral angle **ε** and **ζ** for *positive rotations* along line 1.

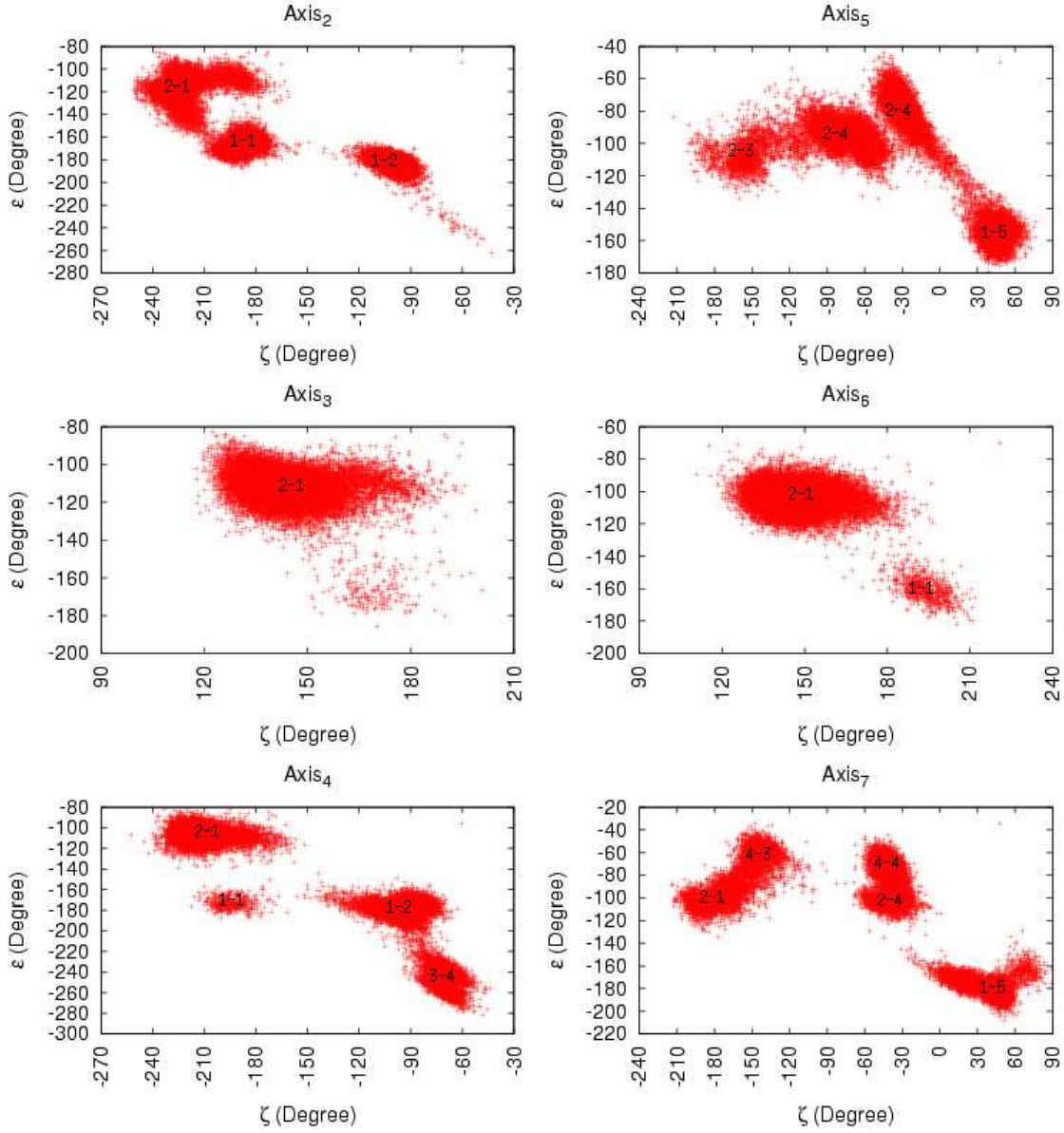

Figure 3.23 Correlation between dihedral angle **ε** and **ζ** for *positive rotations* along line 2.

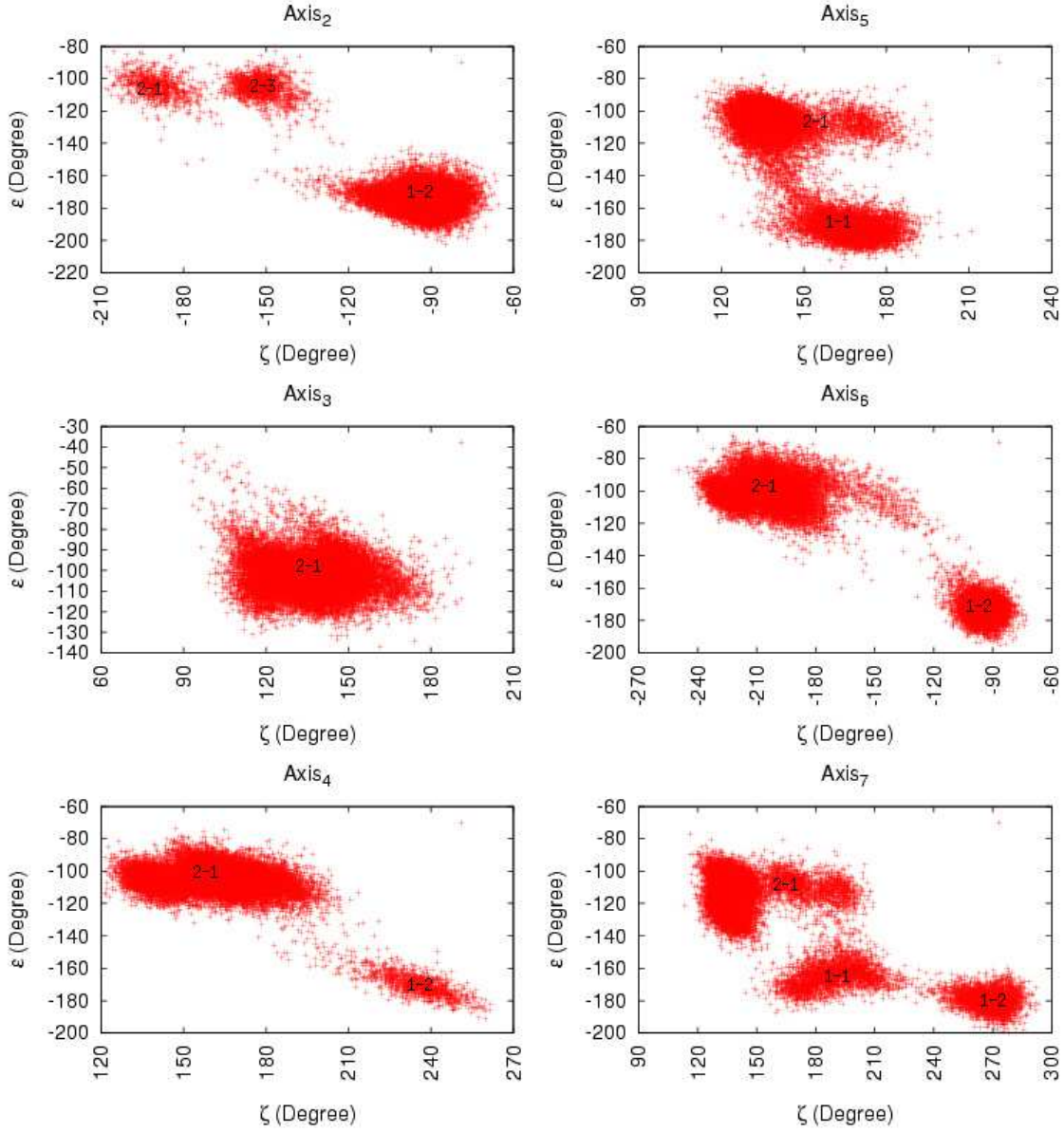

Figure 3.24 Correlation between dihedral angle **ε** and **ζ** for *positive rotations* along line 3.

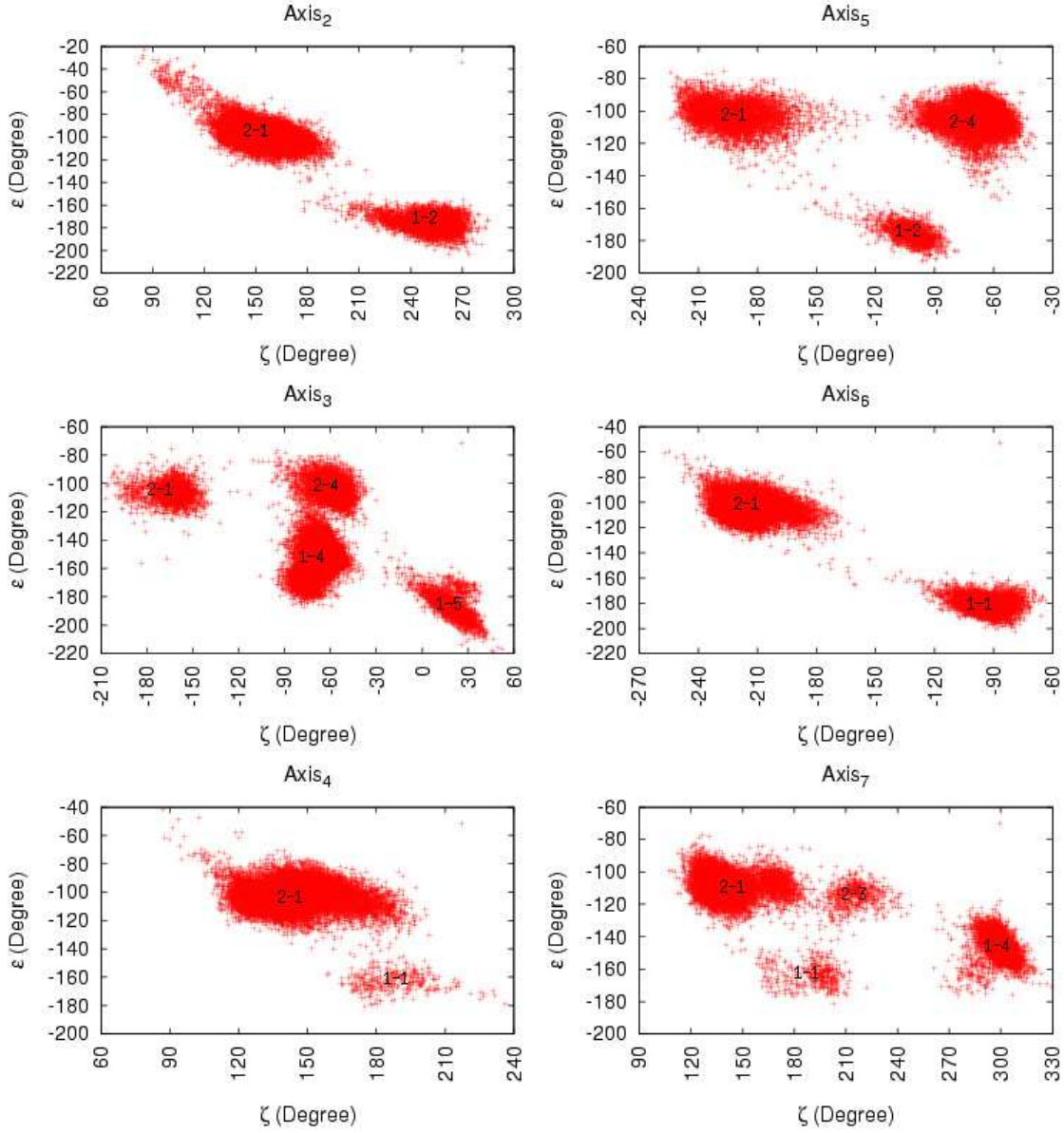

Figure 3.25 Correlation between dihedral angle **ε** and **ζ** for *positive rotations* along line 4.

The other important dihedral angles are **γ** and **α**. In B-DNA γ and α are only the canonical conformations which **γ** is around 60 degree and **α** is around -60 degree [29]. Also these angles are inversely proportional, when **γ** increases, **α** decreases. And for minimum energy state summation of the value of these anges are zero [29]. Therefore, we observed another possible conformations for **γ** and **α**. The Table 3.2 shows the possible conformations that we obtained for **γ** and **α**. In Figure 3.26 the correlation between dihedral angle **γ** and **α** for *negative* and *positive rotations* along parallel axis is given. In Figure 3.27, 3.28, 3.29, and 3.30 the correlation between these angles for *negative rotations* along line 1, line 2, line 3 and line 4, and in Figure 3.31, 3.32, 3.33 and 3.34 for *positive rotations* along line 1, line 2, line 3 and line 4 are given.

Table 3.2 Possible conformations for **γ** and **α**

	γ	**α**
1	60±40	-60±40
2	175±30	10±40
3	120±30	-150±40
4	0±40	-120±30
5	-50±50	80±40
6		150±30

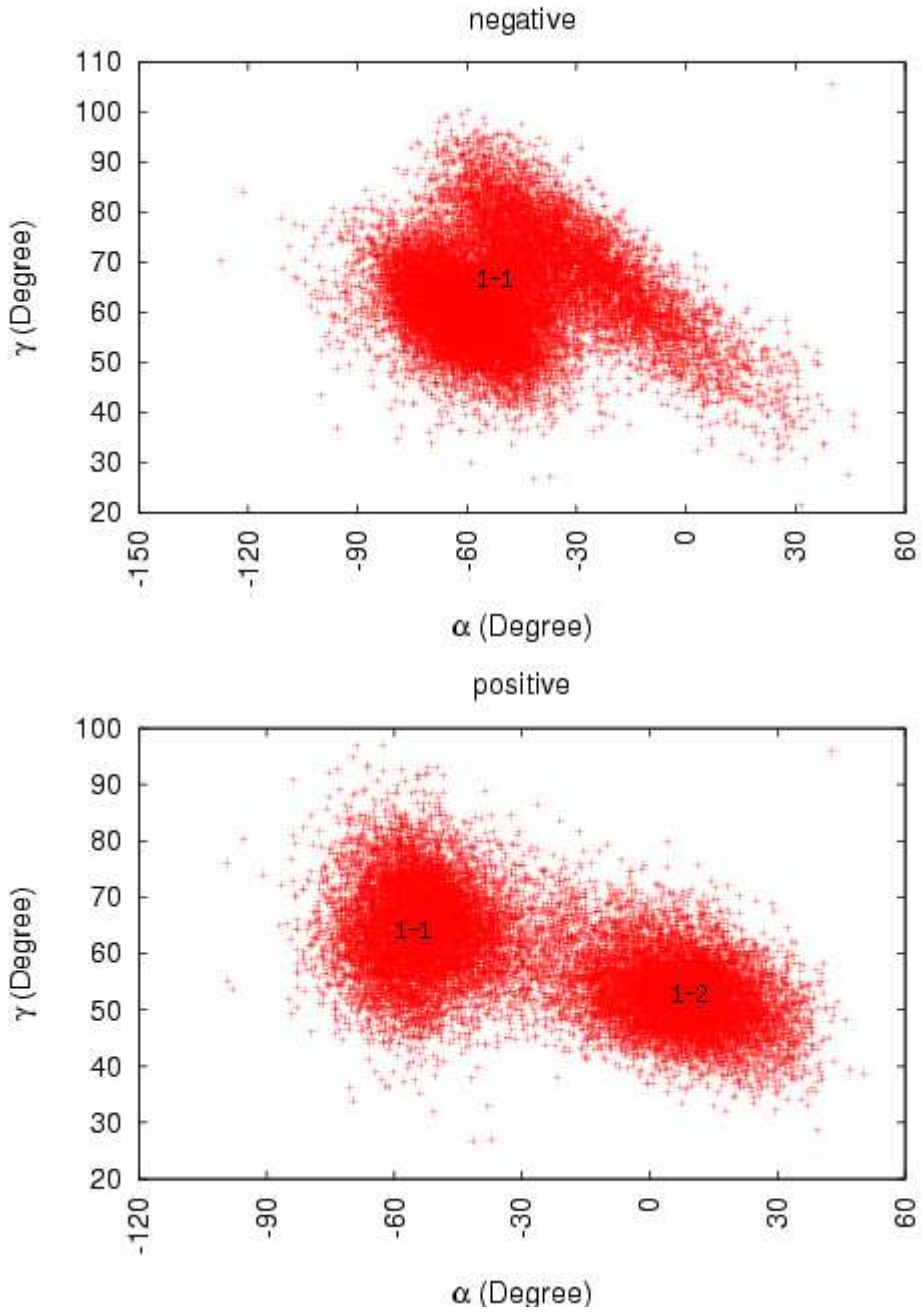

Figure 3.26 Correlation between dihedral angle **γ** and **α** for *negative* and *positive rotations* around parallel axis.

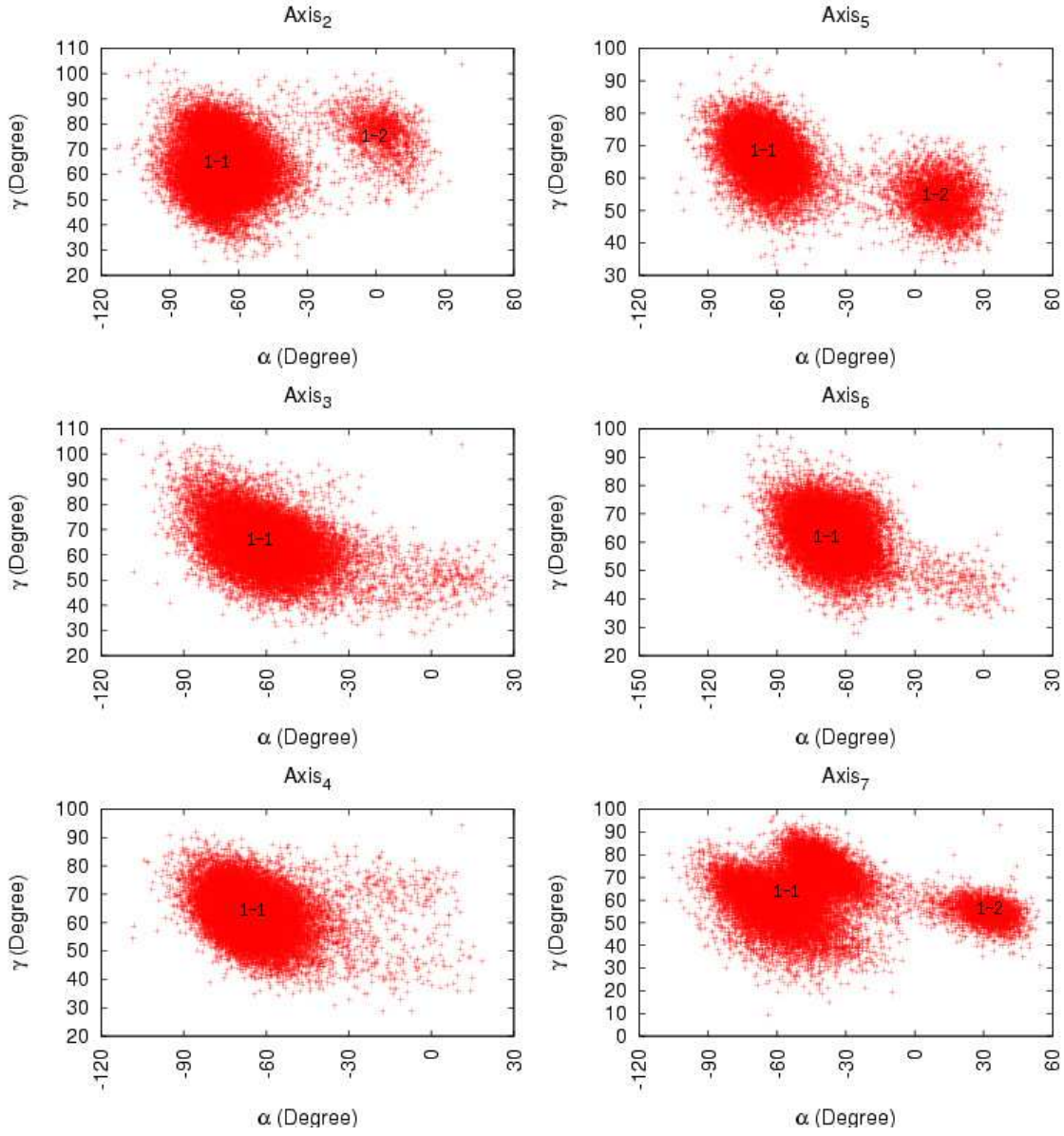

Figure 3.27 Correlation between dihedral angle **γ** and **α** for *negative rotations* along line 1.

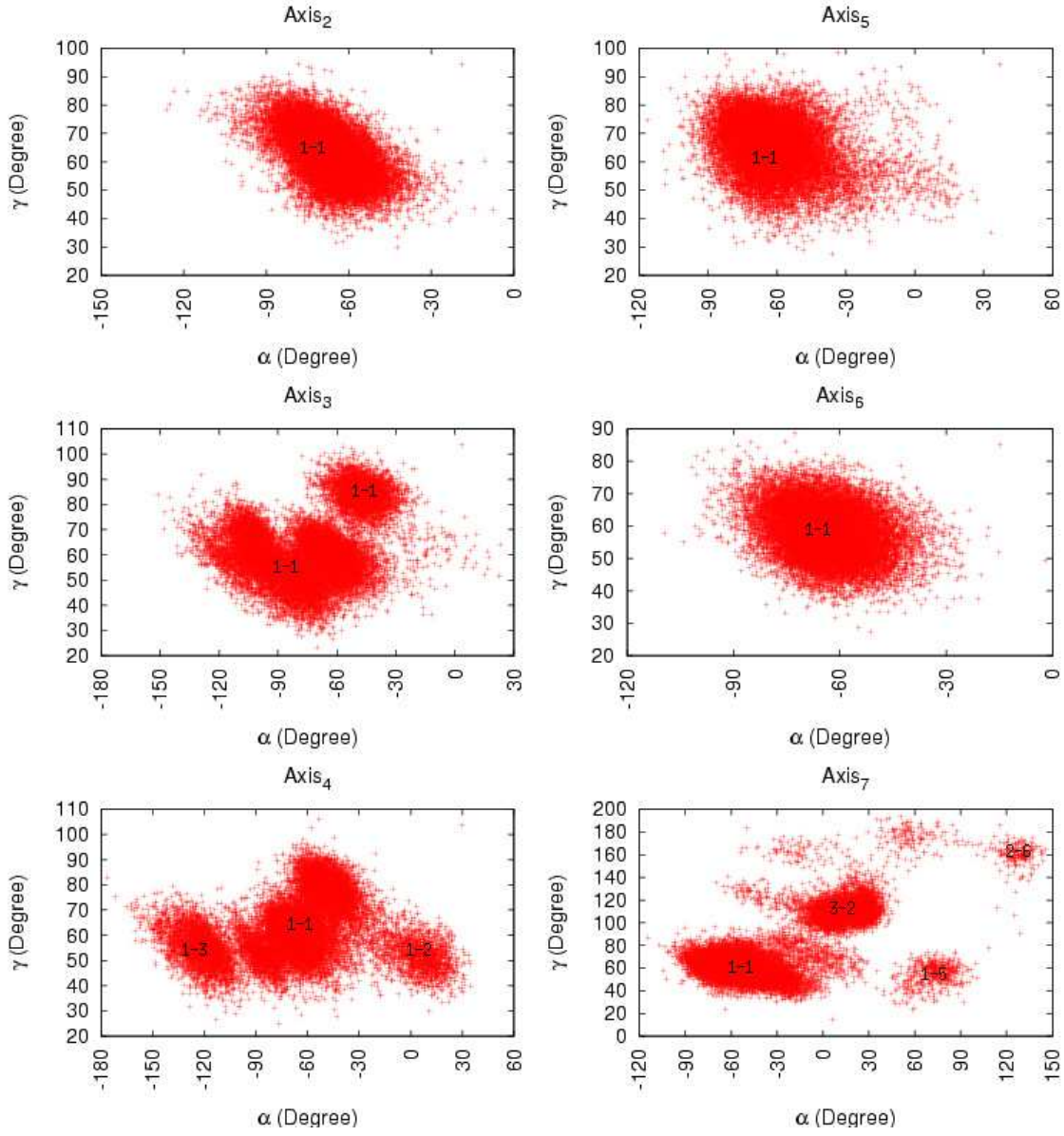

Figure 3.28 Correlation between dihedral angle **γ** and **α** for *negative rotations* along line 2.

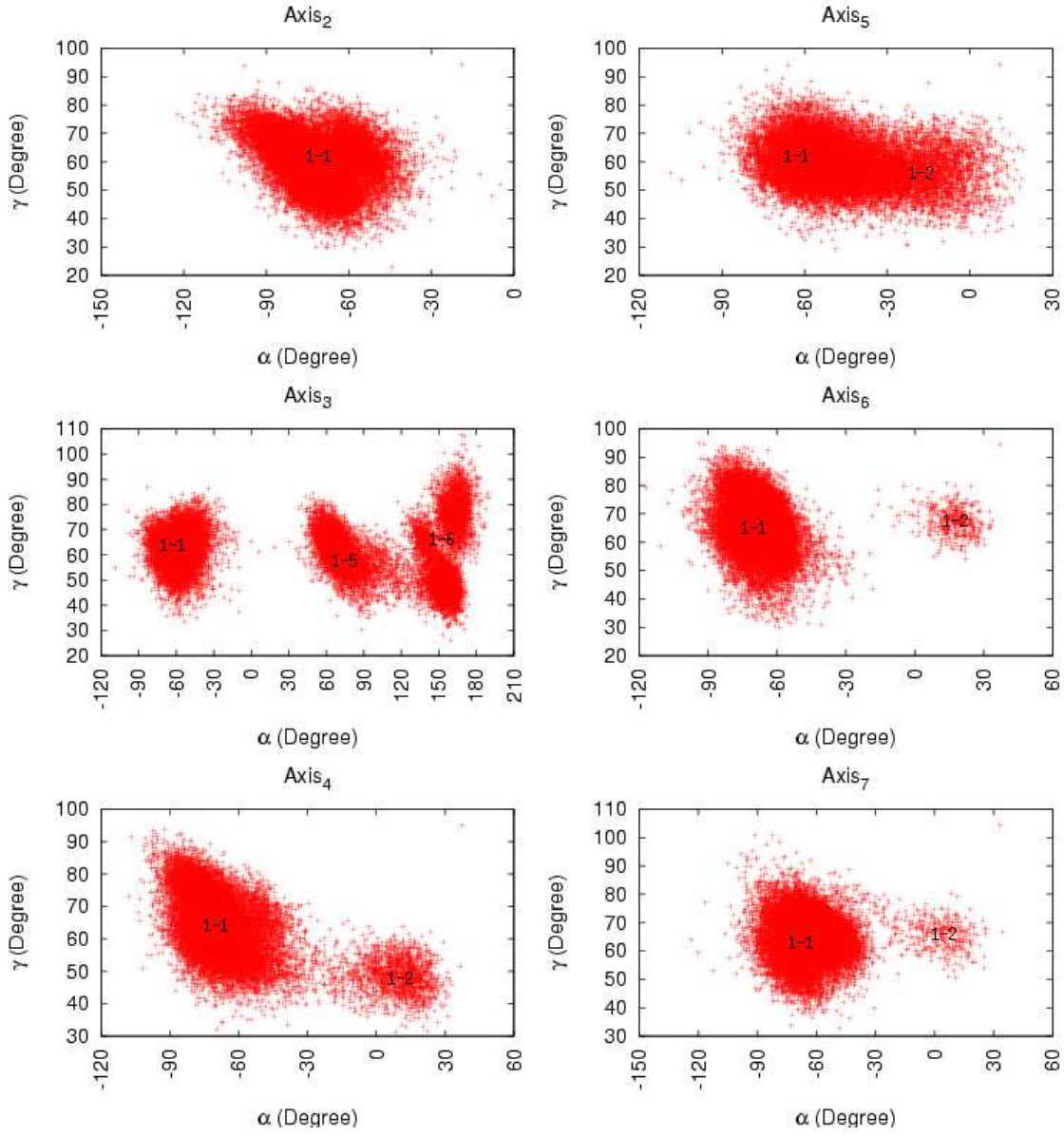

Figure 3.29 Correlation between dihedral angle **γ** and **α** for *negative rotations* along line 3.

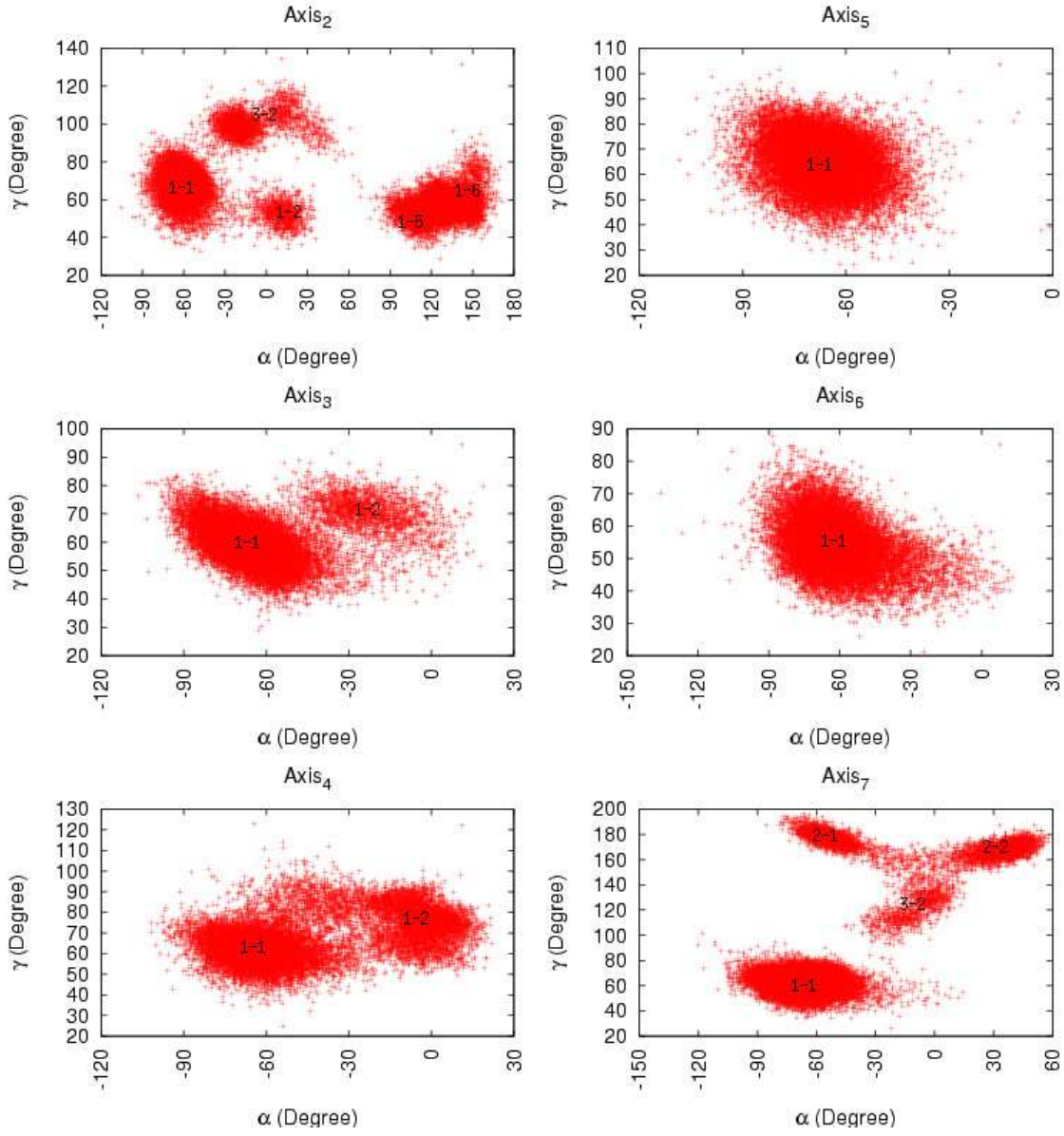

Figure 3.30 Correlation between dihedral angle **γ** and **α** for *negative rotations* along line 4.

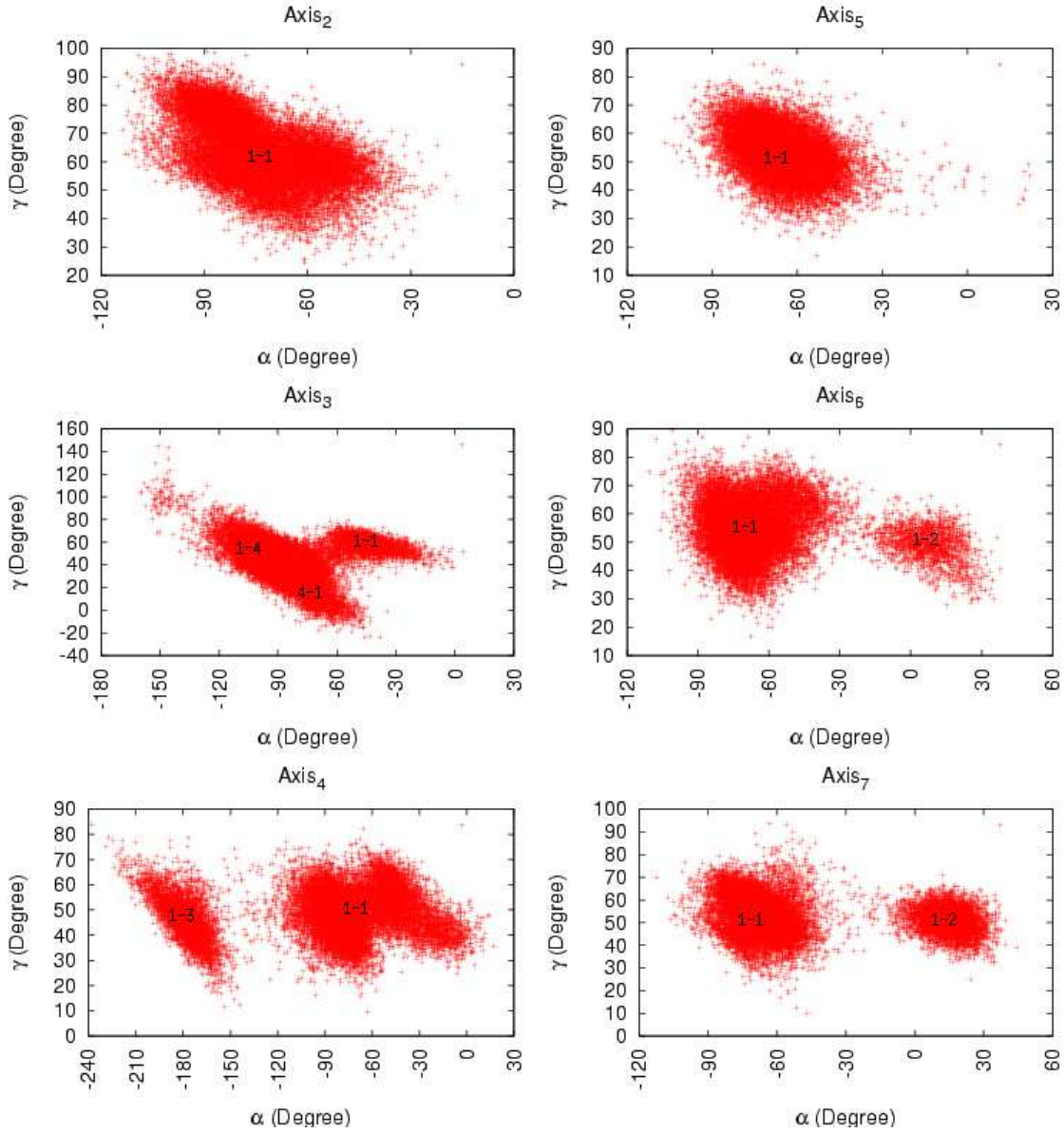

Figure 3.31 Correlation between dihedral angle **γ** and **α** for *positive rotations* along line 1.

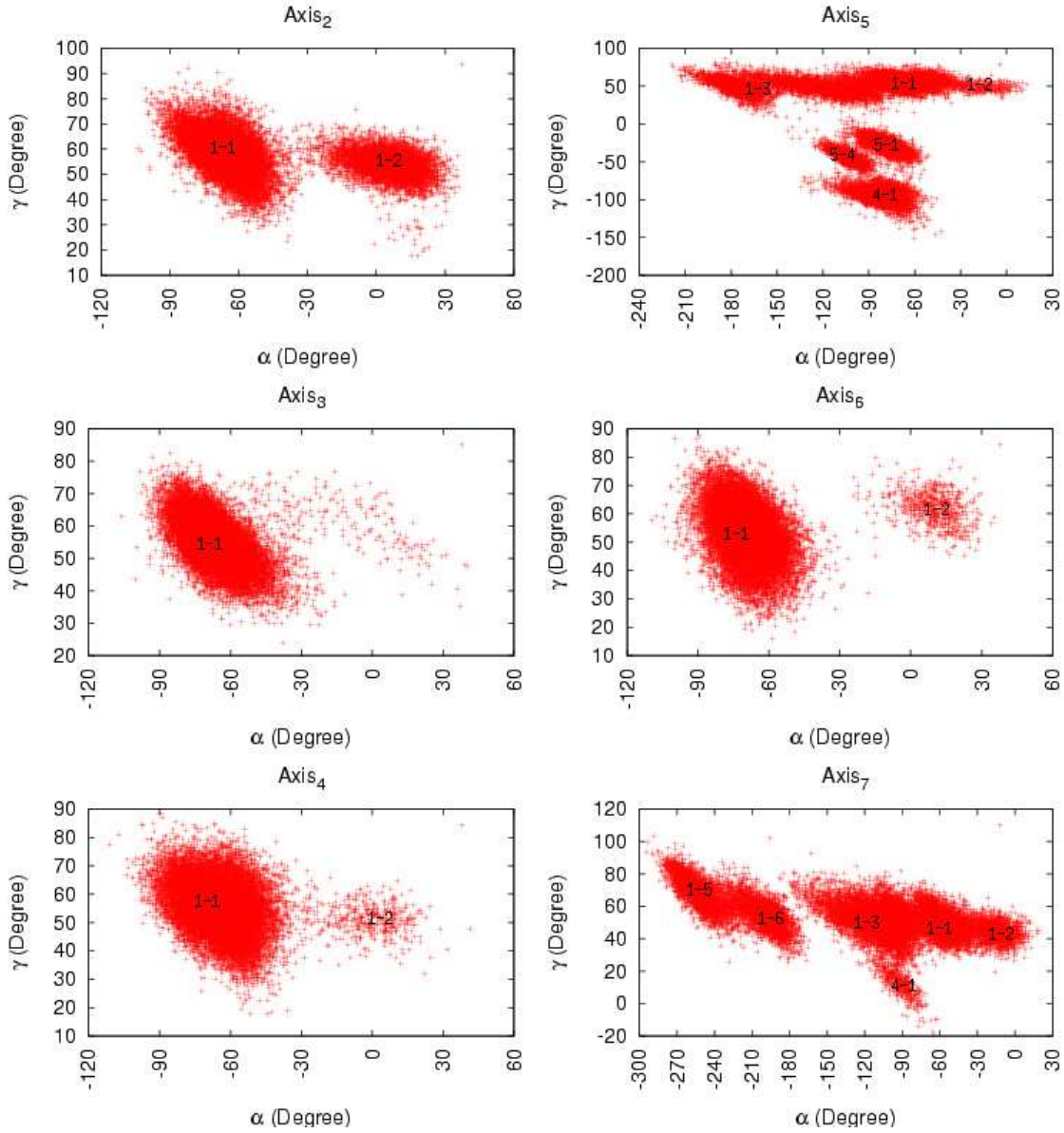

Figure 3.32 Correlation between dihedral angle **γ** and **α** for *positive rotations* along line 2.

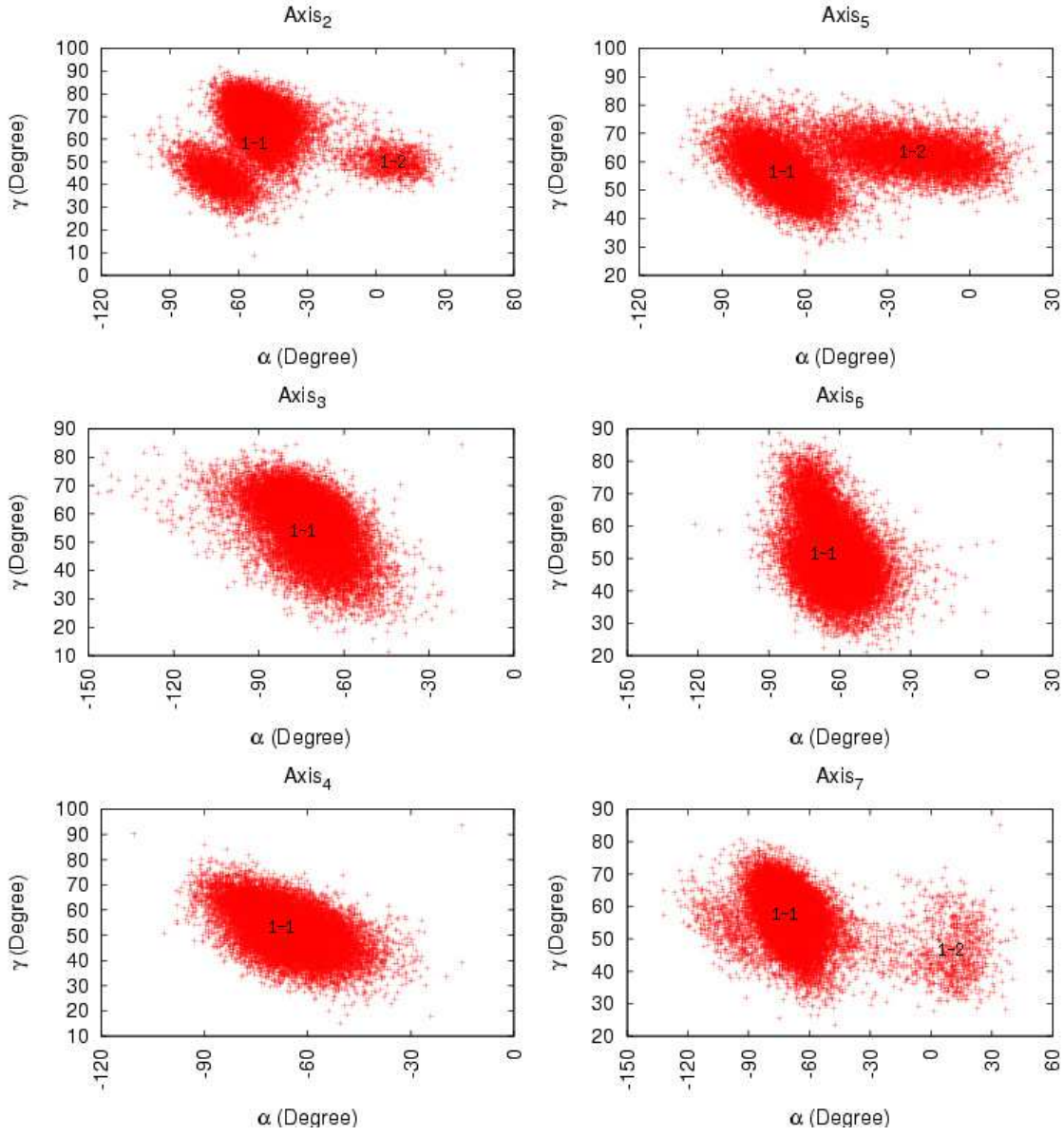

Figure 3.33 Correlation between dihedral angle **γ** and **α** for *positive rotation*s along line 3.

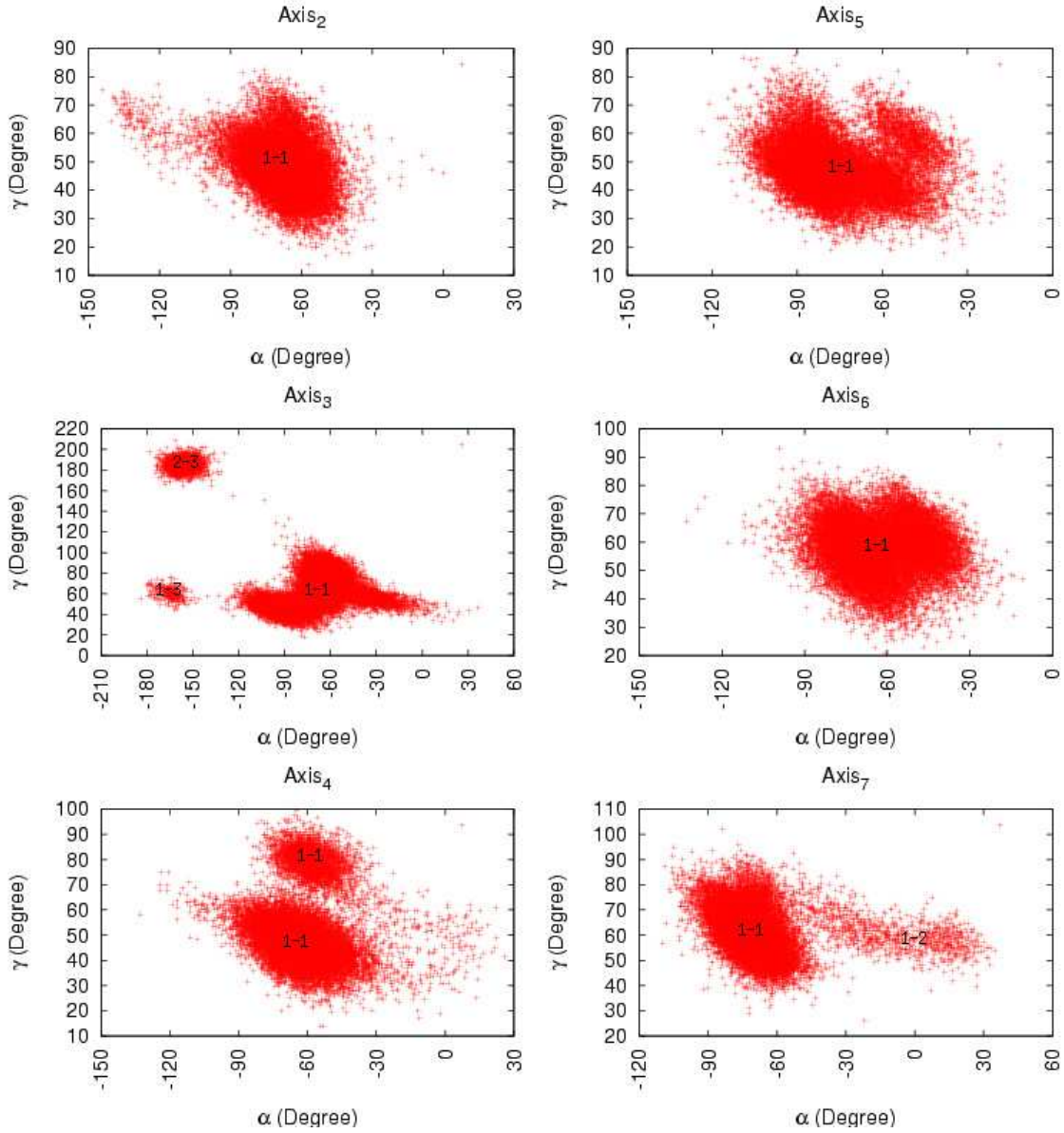

Figure 3.34 Correlation between dihedral angle **γ** and **α** for *positive rotations* along line 4.

3.2 ENERGETIC ANALYSIS

In the energetic analysis, we have calculated average force that is applied to bring the DNA rotations. The required mean forces for rotations are given in Figure 3.35 for rotations along parallel axis, and in Figure 3.36 for *negative rotations*, and in Figure 3.37 for *positive rotations*.

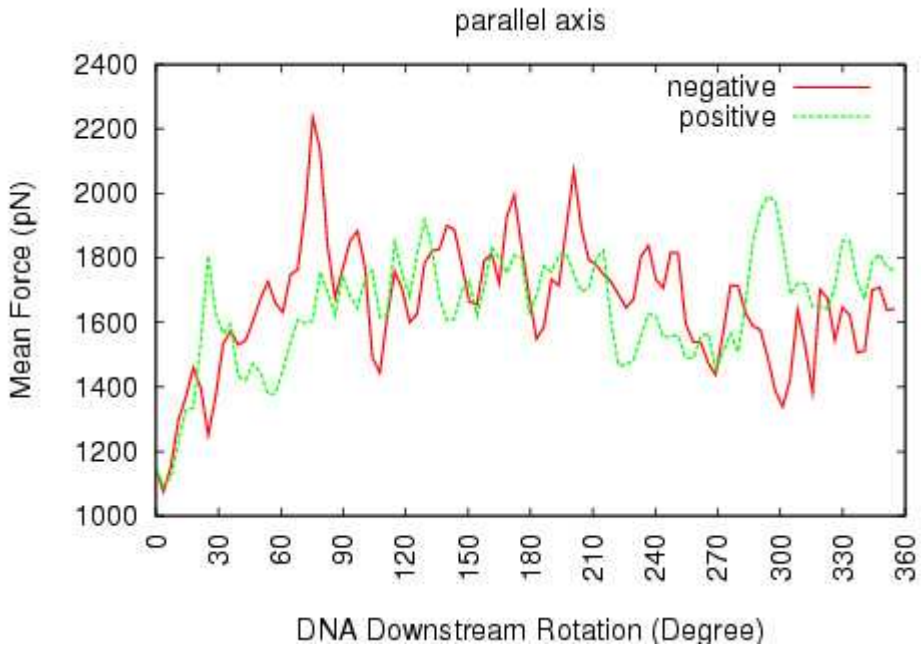

Figure 3.35 Mean force as a function of rotation angle for *negative*and *positive rotations* along the parallel axis.

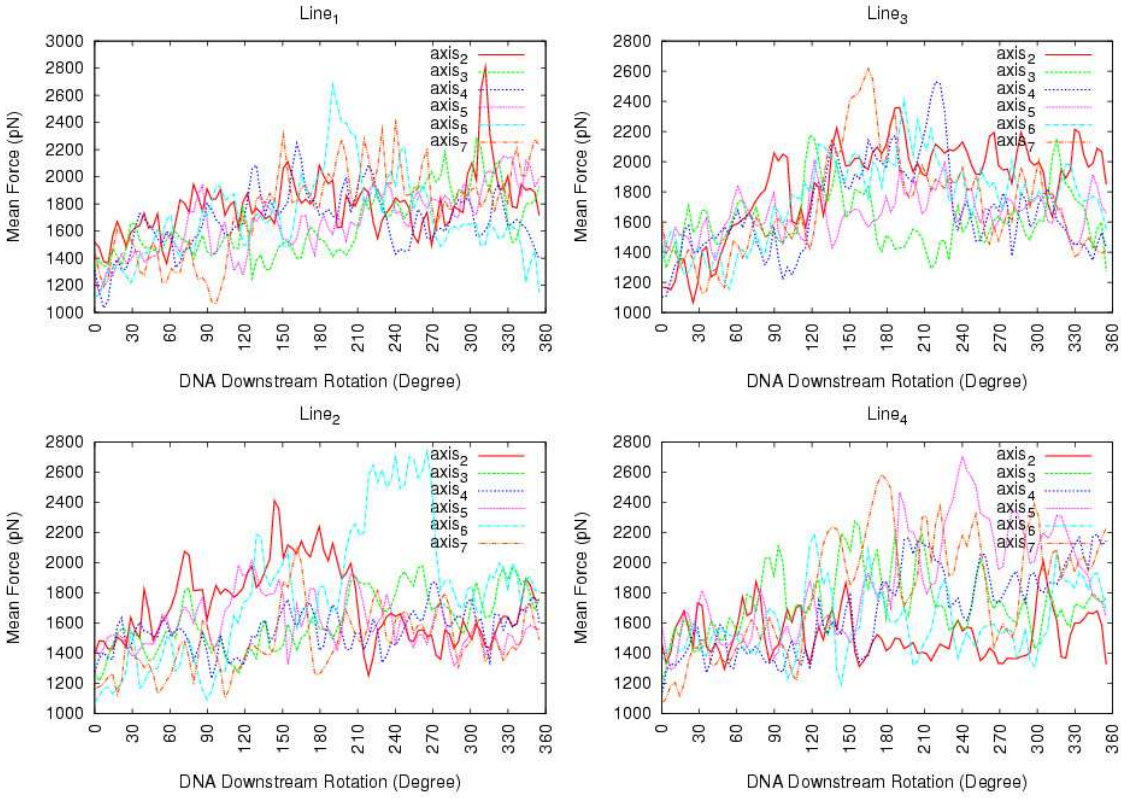

Figure 3.36 Mean force as a function of rotation angle for *negative rotations*.

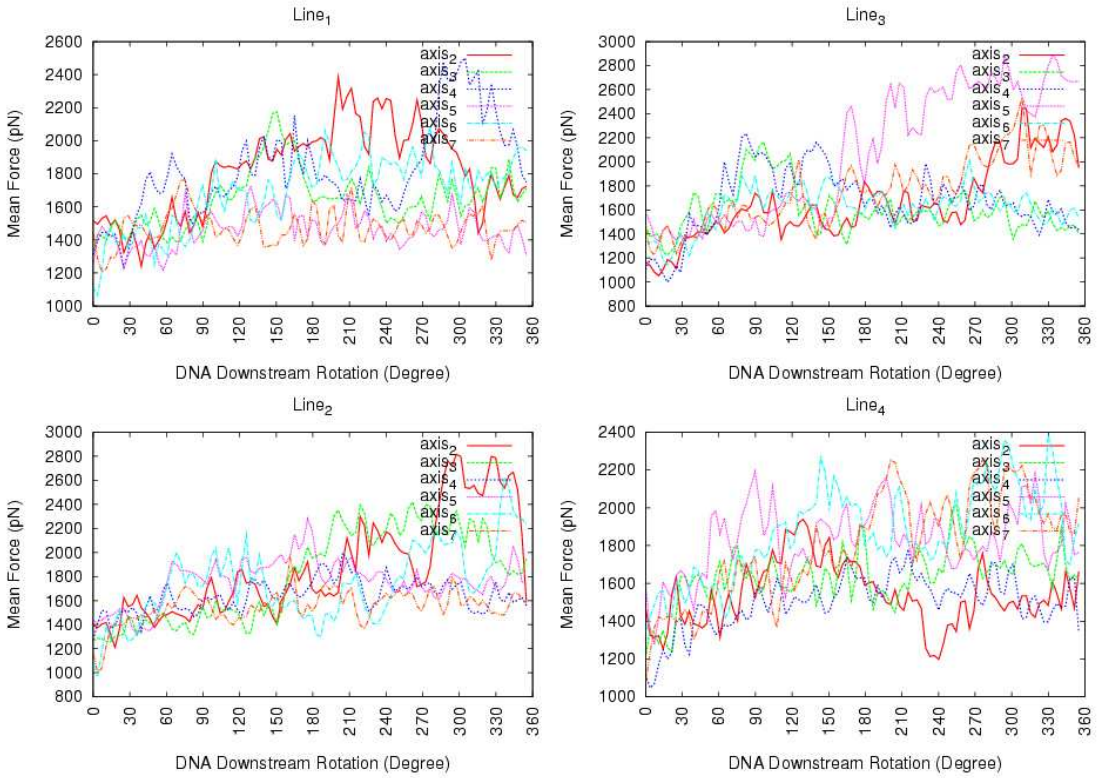

Figure 3.37 Mean force as a function of rotation angle for *positive rotations*.

Two dimensional potential energy surfaces as a function of **α** and **γ** are calculated, and given in Figures 3.38 to 3.46. The potential energies correspond to the total bonded and non-bonded interactions for the atoms in the +1 and -1 DNA base-pairs. These potential energy surfaces are quite important in studying the energetic pathways for the rotations. Such energy graphs for intact DNA have been studied extensively [29, 30] in literature, but the corresponding energy maps for a ***nicked DNA*** is reported first time here in our study.

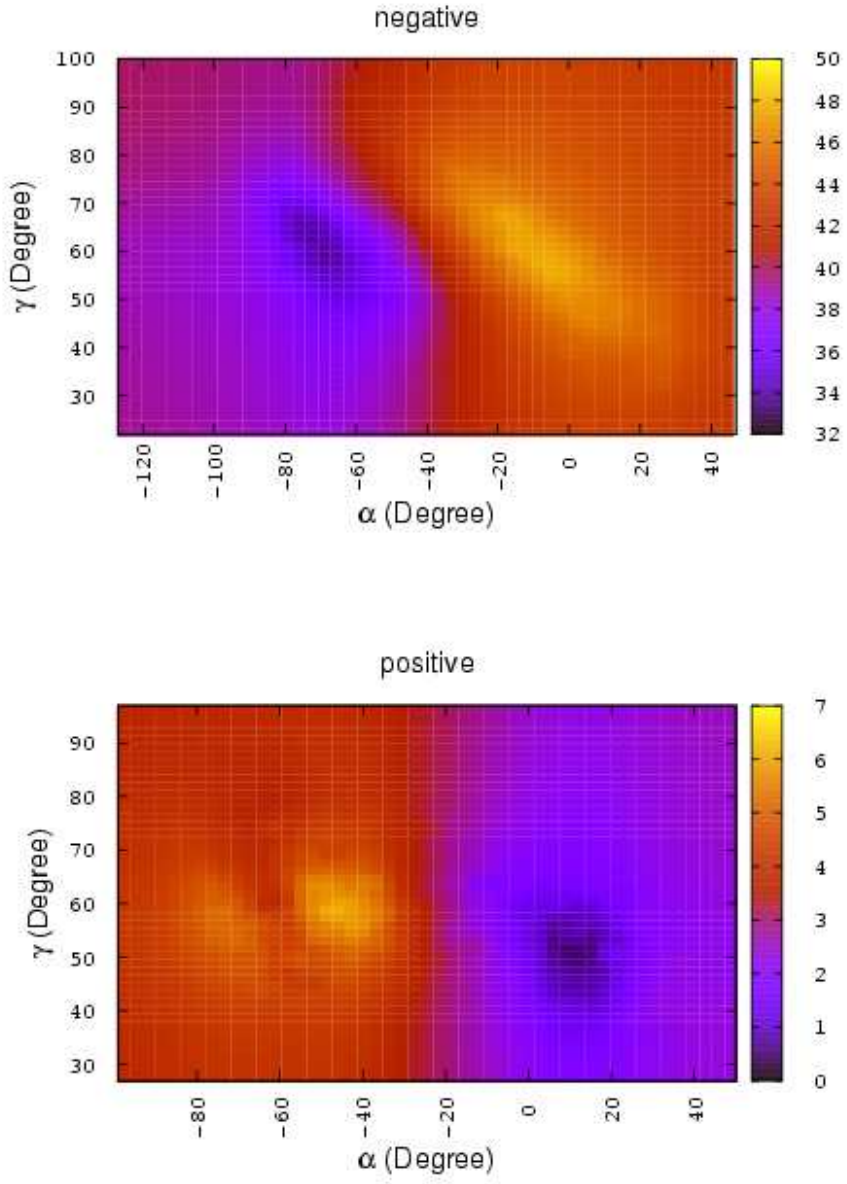

Figure 3.38 Energy map (kcal/mol) as a function of **α** and **γ** dihedral angle values along parallel axis for *negative* and *positive rotations*.

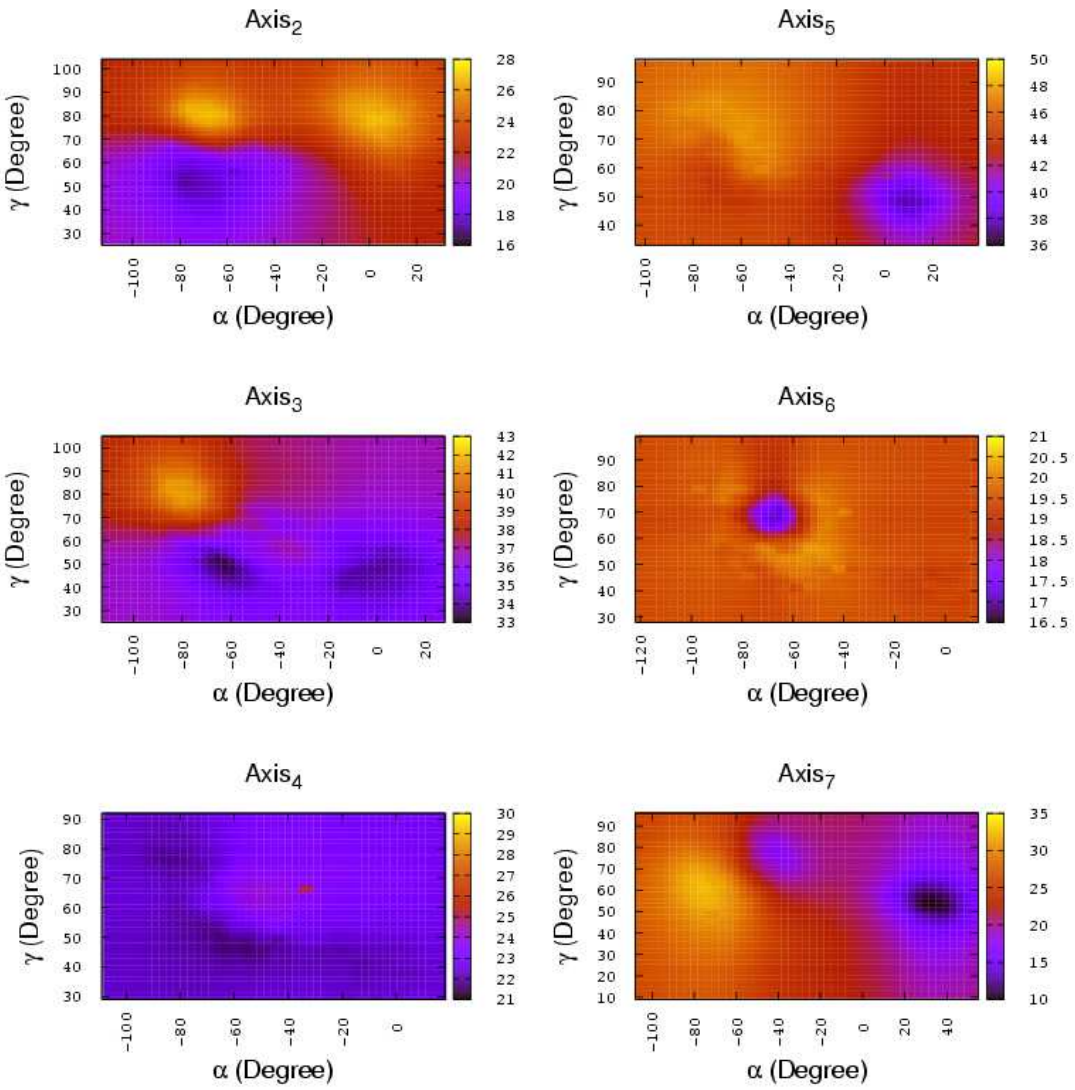

Figure 3.39 Energy map (kcal/mol) as a function of **α** and **γ** dihedral angle values for *negative rotations* along line 1.

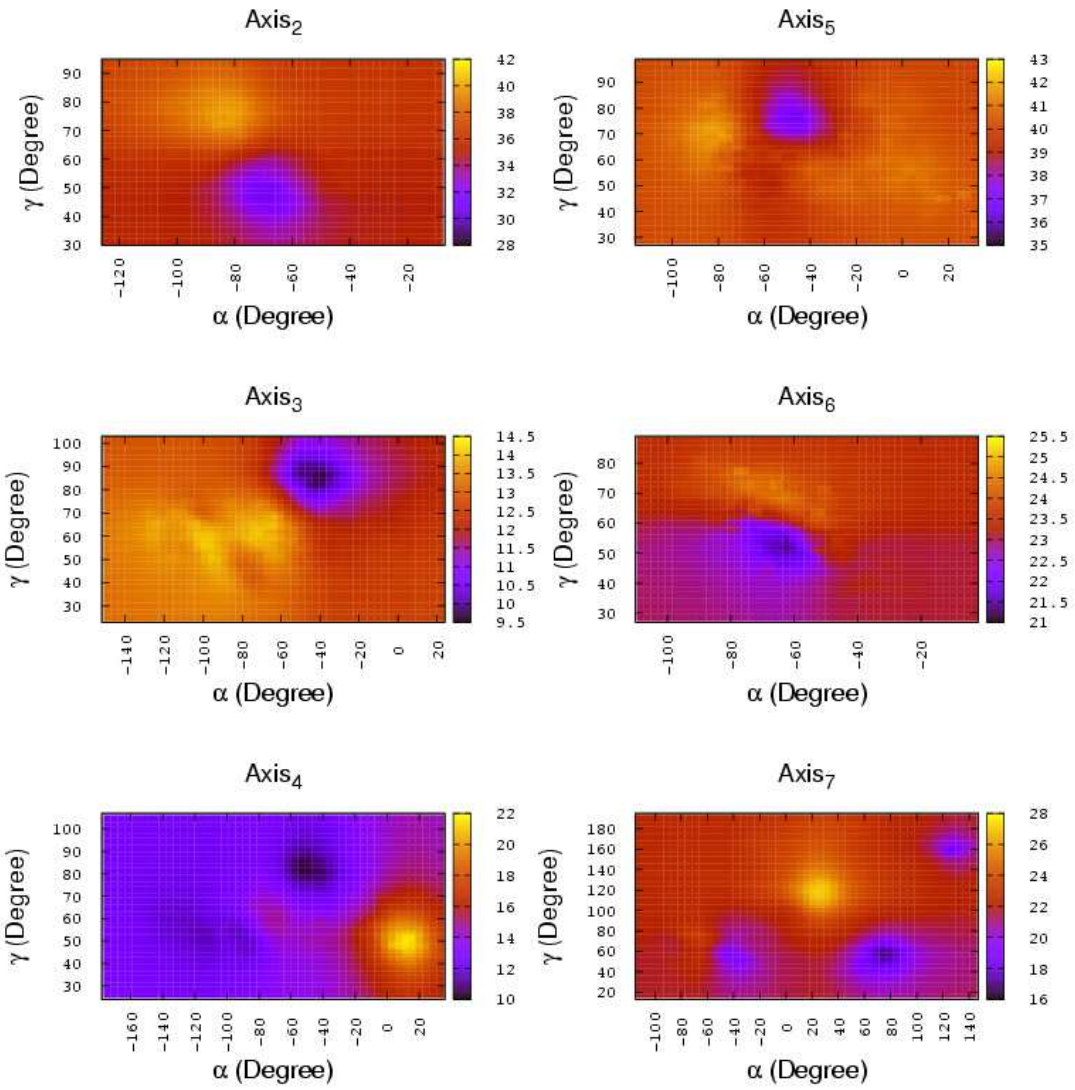

Figure 3.40 Energy map (kcal/mol) as a function of **α** and **γ** dihedral angle values for *negative rotations* along line 2.

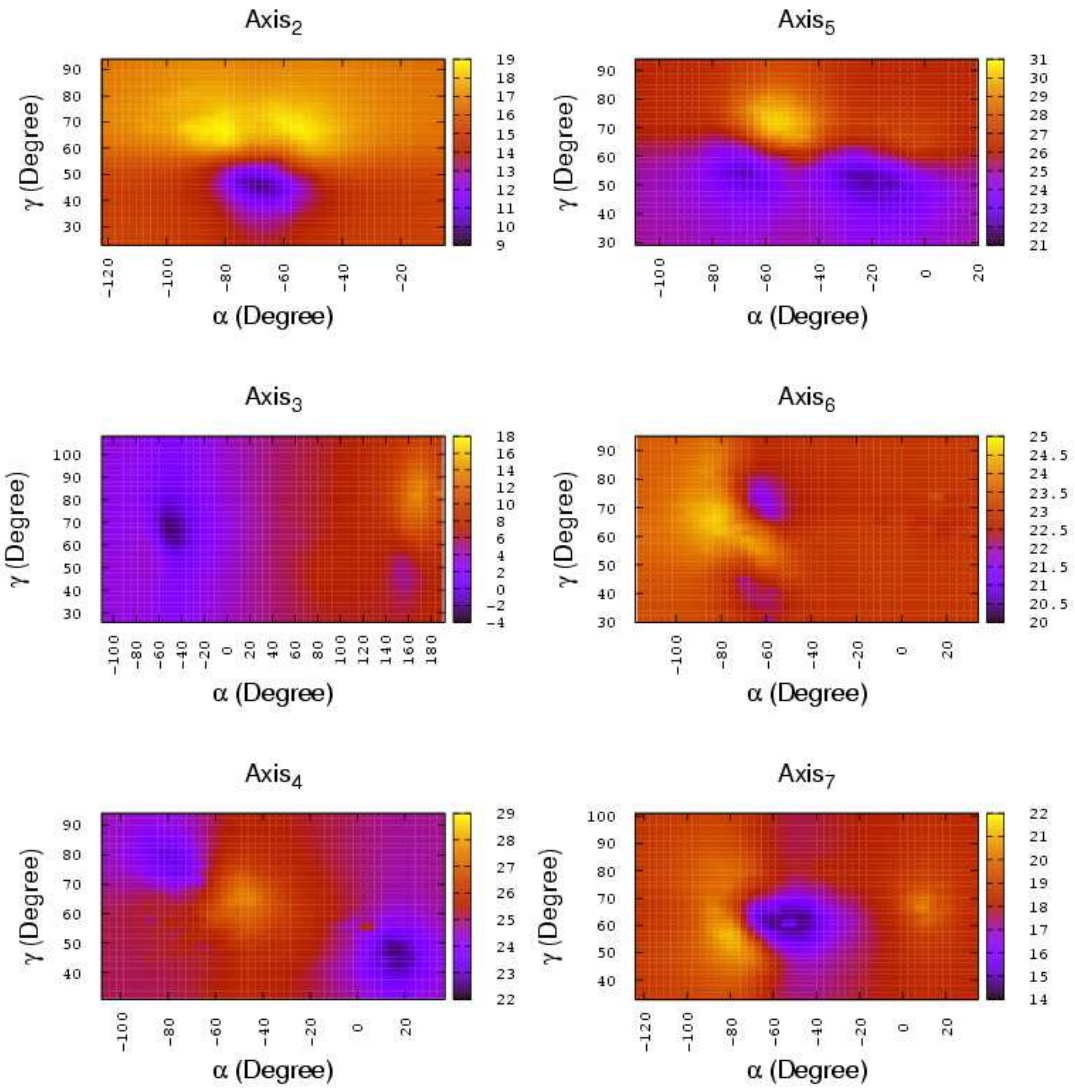

Figure 3.41 Energy map (kcal/mol) as a function of **α** and **γ** dihedral angle values for *negative rotations* along line 3.

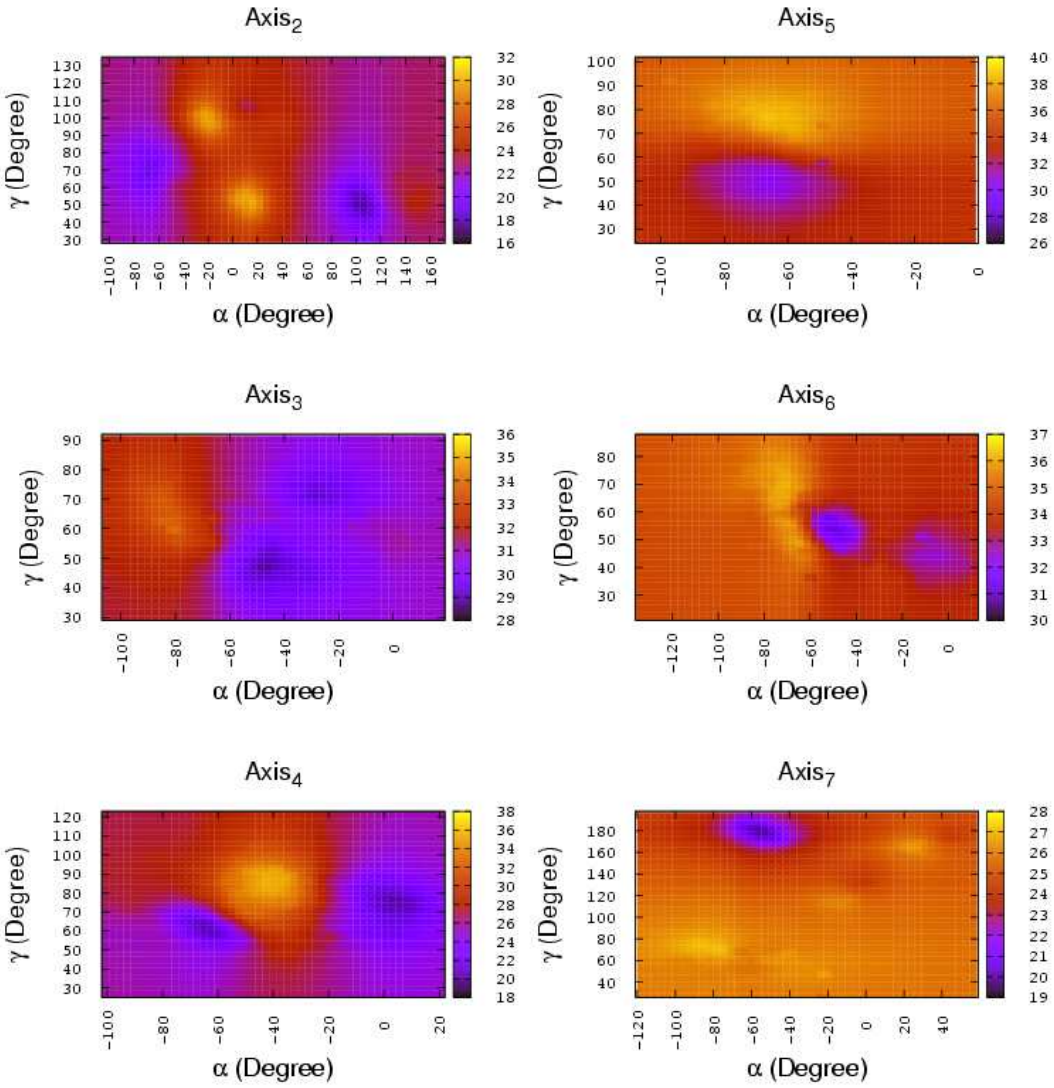

Figure 3.42 Energy map (kcal/mol) as a function of **α** and **γ** dihedral angle values for *negative rotations* along line 4.

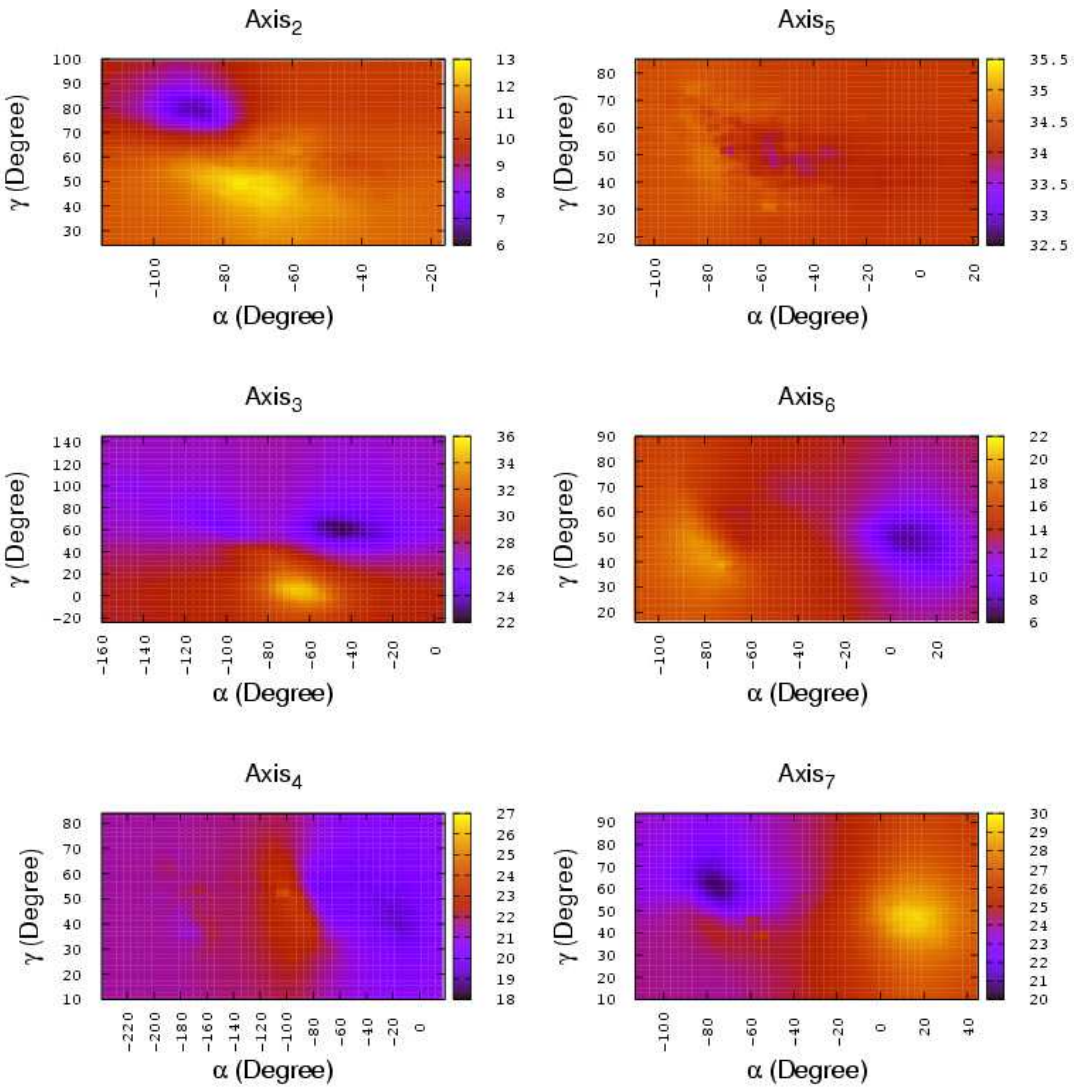

Figure 3.43 Energy map (kcal/mol) as a function of **α** and **γ** dihedral angle values for *positive rotations* along line 1.

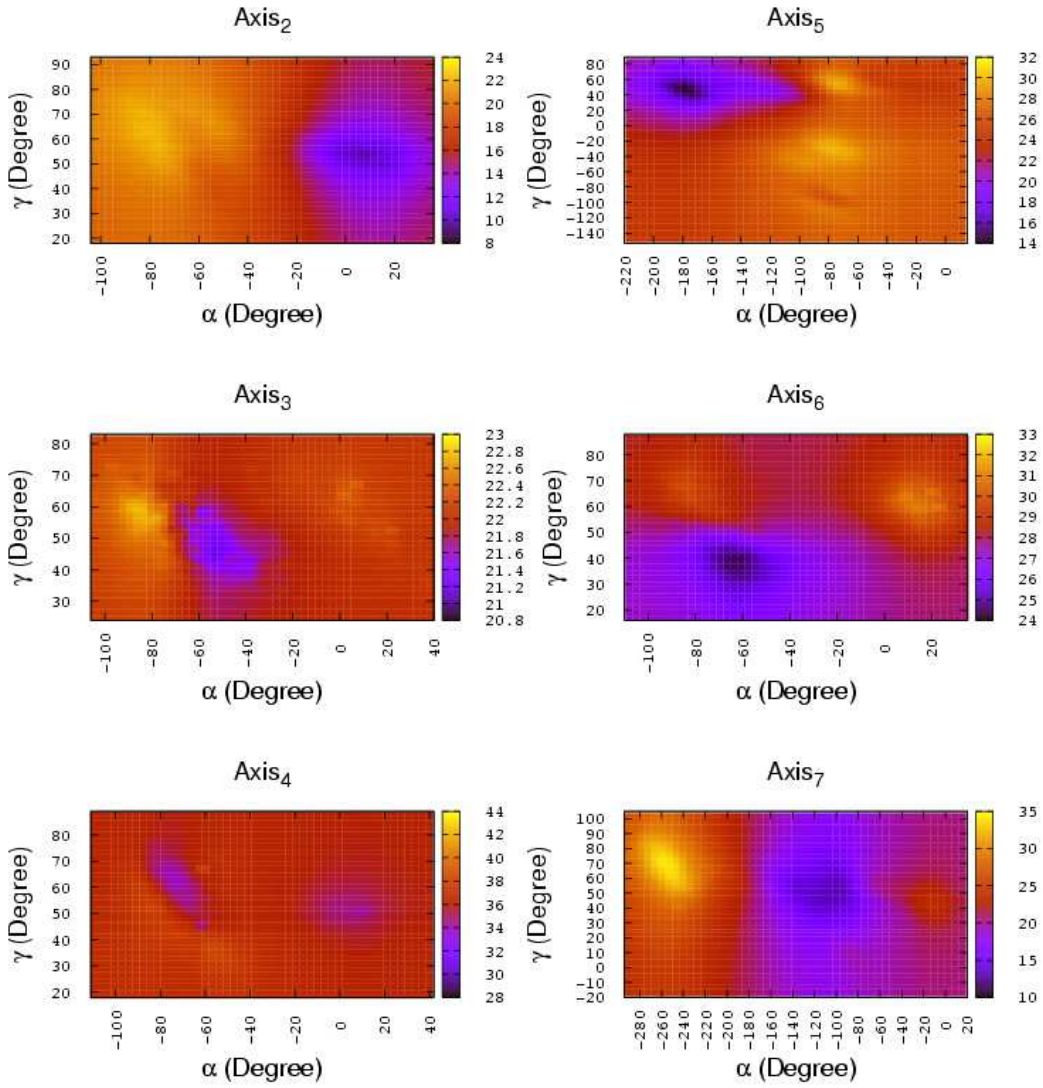

Figure 3.44 Energy map (kcal/mol) as a function of **α** and **γ** dihedral angle values for *positive rotations* along line 2.

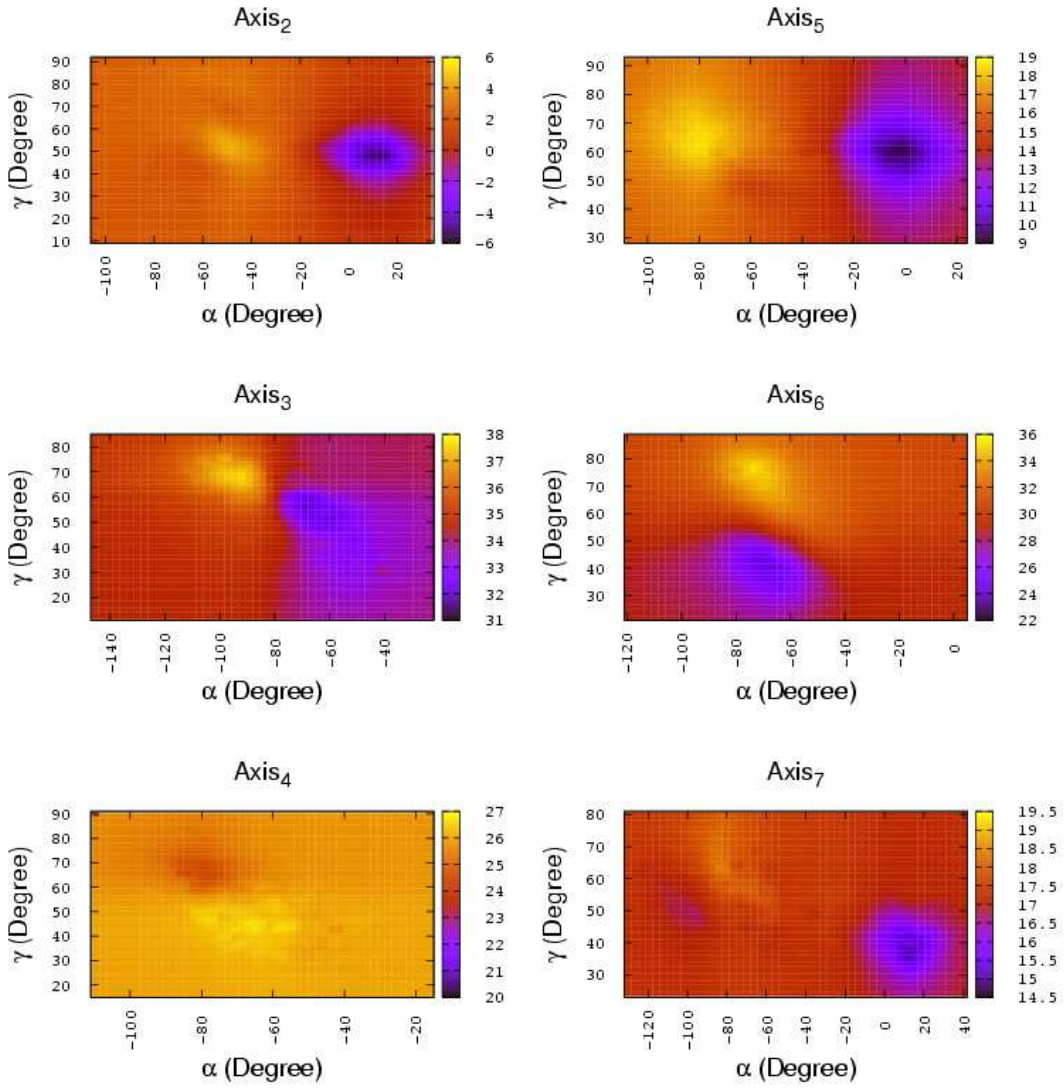

Figure 3.45 Energy map (kcal/mol) as a function of **α** and **γ** dihedral angle values for *positive rotations* along line 3.

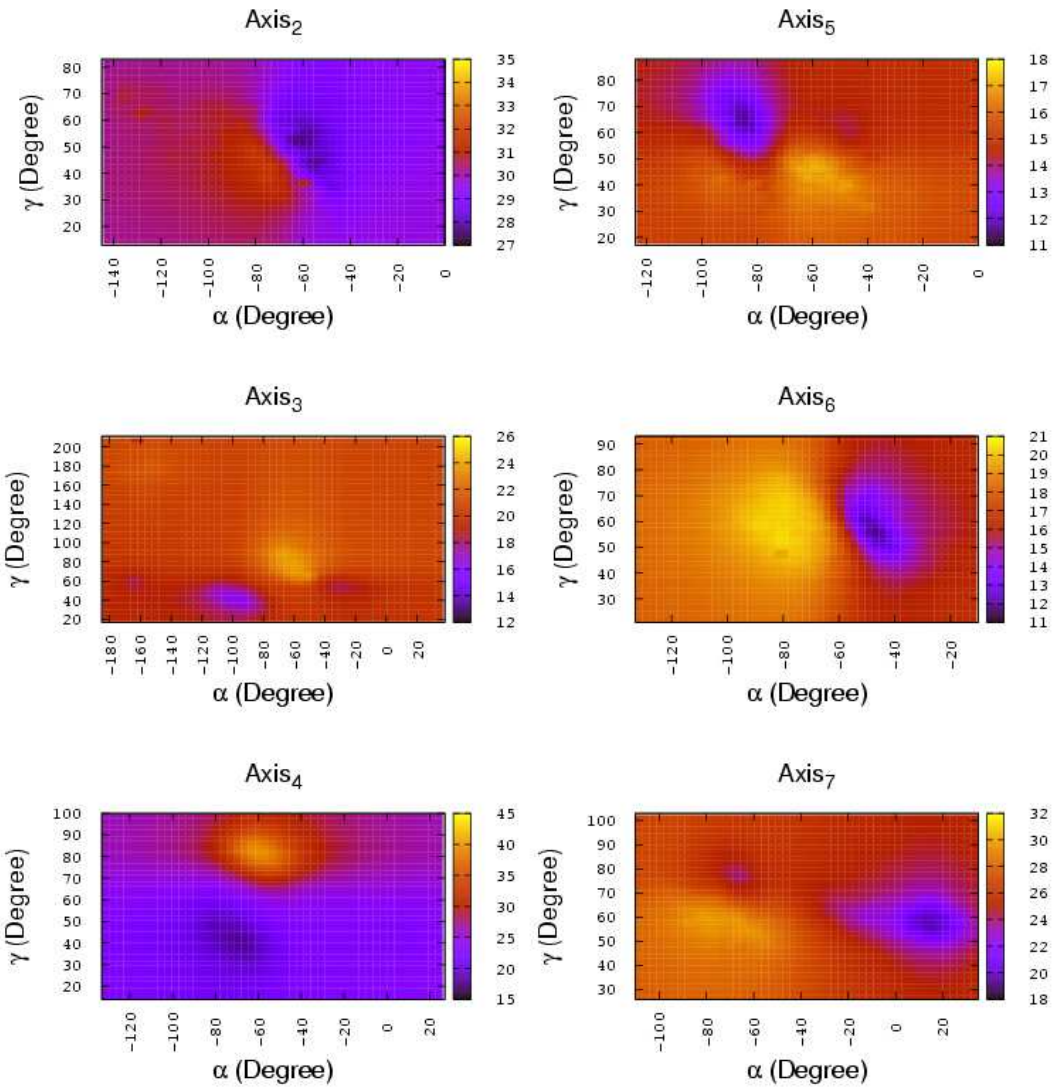

Figure 3.46 Energy map (kcal/mol) as a function of **α** and **γ** dihedral angle values for *positive rotations* along line 4.

CHAPTER 4

RESUTLS AND DISCUSSION

In this study fifty different simulations are performed using MD as mentioned in Chapter 2. Our purpose is to find the most favorable simulations by analyzing all of these fifty simulations. We focused on the structural and energetic changes in the nicked region of the DNA. Structural analysis includes informations about the changes of backbone dihedral angles at the nicked site and deformations of DNA during the rotations. So we investigate the five backbone dihedral angles in the opposite of the niked site. Changes of total potential energy and applied mean forces are also analyzed.

One of the most important questions that we had at the beginning of this research was to see whether the full 360 degree DNA rotations are accommodated by partial dihedral rotations or a full 360 degree rotation(s) of specific dihedrals. Our results surprisingly showed us some quite important data in regards to this question. Basically, we have seen that the full DNA rotation is provided by different paths for different rotations.

4.1 Changes in backbone dihedral angles

For the *negative rotations*, mainly three different scenarios are found to be followed by the backbone dihedrals. In the *first scenario*, the dihedral angle ζ makes a full rotation as the downstream DNA makes a full rotation. The *second scenario* observed for *negative rotations* is that the full DNA rotation is provided by partial dihedral rotations, and the dihedrals that bring the DNA rotation DO end in a different state than they started. The *last scenario* that brings the DNA rotation is that all of the dihedral angles stay unchanged as the downstream DNA rotates.

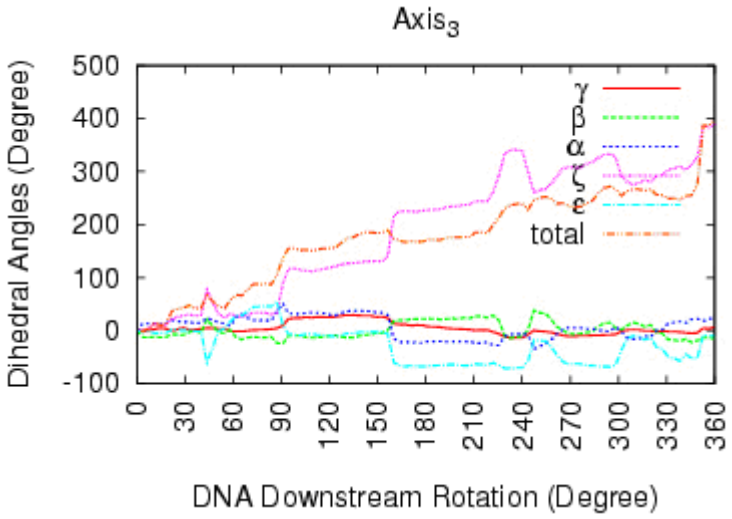

Figure 4.1 Changes in backbone dihedral angles for *negative rotation* around $axis_3$ along line 2.

Figure 4.1 is an example for the *first scenario*. As seen in this figure all of the backbone dihedral angles stay unchanged at the end of the simulation except ζ. The dihedral angle ζ (lilac color) changes around 380 degree and total dihedral angle (orange color) changes depending on the ζ. The full rotation is seen in the *negative rotations* around $axis_3$, $axis_4$, and $axis_7$ along line 2 (Figure 3.4- lilac color). The graphs of these rotations that are given in Chapter 3 are similar with this graph.

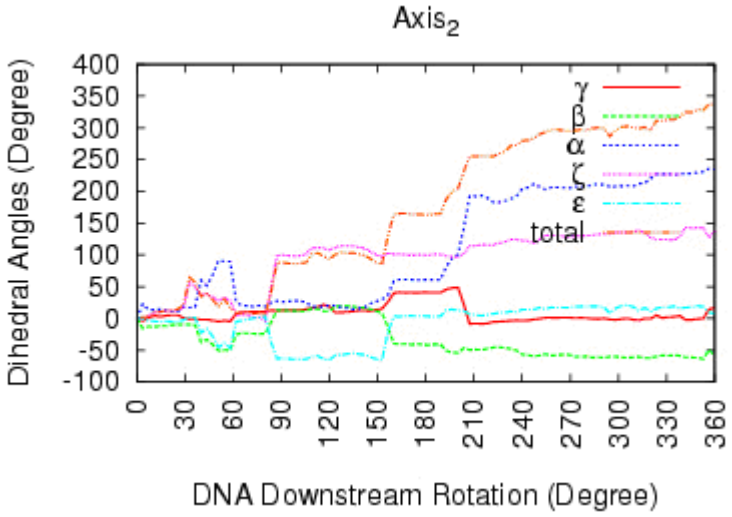

Figure 4.2 Changes in backbone dihedral angles for *negative rotation* around $axis_2$ along line 4.

In the *second scenario*, partial dihedral angle rotations are observed as shown in Figure 4.2. At the end of the full DNA downstream rotation the dihedral angles **γ**, **ε** and **β** show no changes. The only changing dihedral angles are **ζ** and **α**. **ζ** changes to 130 degree and **α** changes to 230 degree. The change of the total dihedral at the end of the rotation is 340 degree. Also this observation is seen in the *negative rotations* around parallel axis (Figure 3.2), $axis_7$ along line 1 (Figure 3.3), $axis_3$ along line 3 (Figure 3.5), $axis_2$ and $axis_7$ along line 4 (Figure 3.6). Their dihedral angle values at the end of the rotation are listed in Table 4.1.

Table 4.1 Dihedral angles at the end of the rotation for *negative rotations* in the second scenario.

rotation axis	**γ**	**β**	**α**	**ζ**	**ε**	total	
parallel	15	70	20	95	-80	125	Fig. 3.2
$axis_7$	20	30	50	40	-70	10	Fig. 3.3
$axis_3$	35	-50	250	100	35	20	Fig. 3.5
$axis_2$	35	-50	235	135	20	335	Fig. 3.6
$axis_7$	105	-70	105	-70	35	105	Fig. 3.6

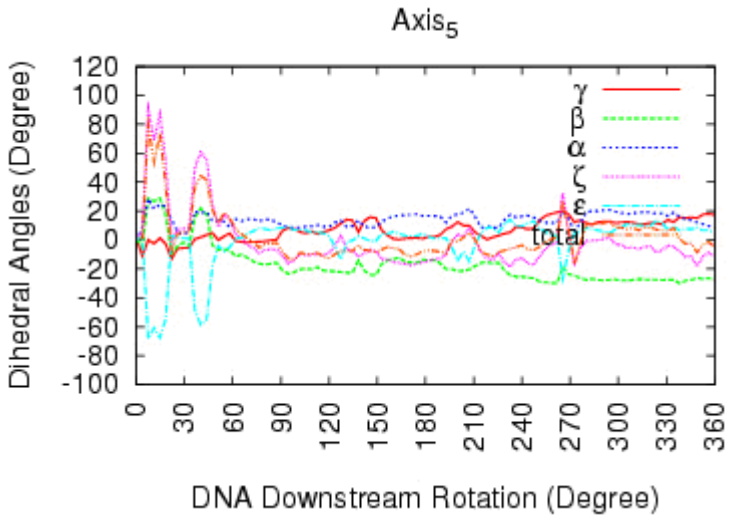

Figure 4.3 Changes in backbone dihedral angles for *negative rotation* around $axis_5$ along line 4

The most common behavior for the *negative rotations* is that all of the dihedral angles show no changes as seen in Figure 4.3. This is called as *last scenario* and observed in these *negative rotations*: rotations around $axis_2$, $axis_3$, $axis_4$, $axis_5$, $axis_6$ along line 1 (Figure 3.3), rotations around $axis_2$, $axis_5$, $axis_6$ along line 2 (Figure 3.4), rotations around $axis_2$, $axis_4$, $axis_5$, $axis_6$, $axis_7$ along line 3 (Figure 3.5) and rotations around $axis_3$, $axis_4$, $axis_5$, $axis_6$ along line 4 (Figure 3.6).

In contrast to the *negative rotations*, we have not observed a full rotation of any dihedral angle in any simulation for *positive rotations* (Figure 3.7, 3.8, 3.9, 3.10). In *positive rotations*, we only see two scenarios. The *first one* is that the changes in the total dihedral angles are summed to have a value of around zero (0±60) and one or two dihedral angle changes or all of them stay unchanged at the end of the downstream DNA rotations. Similar to the first scenario, in the *second scenario*, we see that one or two dihedral angle changes again, but in this case, changes in the total dihedrals are summed to have non-zero value, somewhere between 80 and 140 degrees.

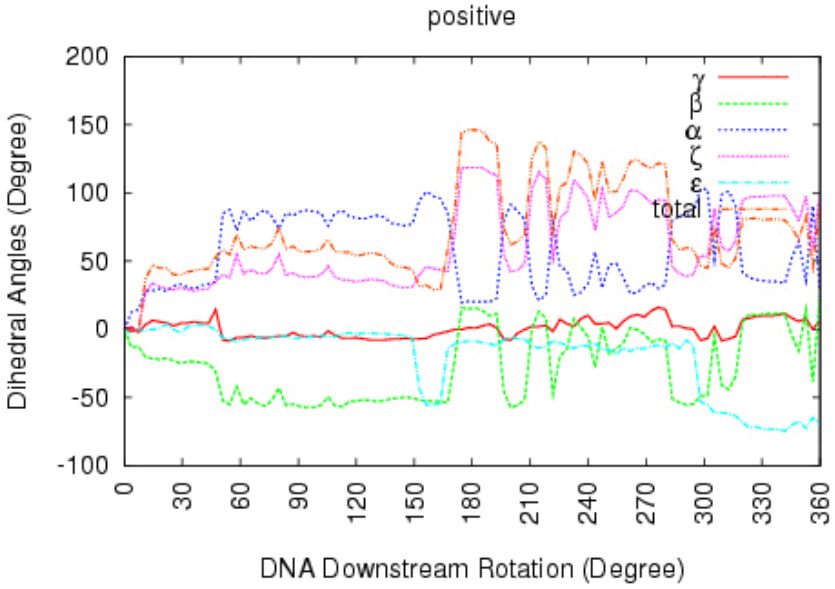

Figure 4.4 Changes in backbone dihedral angles for *positive rotation* around parallel axis

In the *first scenario*, total dihedral angles stay almost unchanged and this behavior is frequently observed in the *positive rotations*. The rotation around parallel axis is suitable for the *first scenario* and changes in dihedral angles are shown in Figure 4.4. The dihedral angles rotate partially and the change of the total dihedral at the end of the downstream DNA rotation is 50 degree. This is also observed in the *positive rotations* around $axis_2$, $axis_5$, $axis_7$ along line 1 (Figure 3.7), rotations around $axis_2$, $axis_3$, $axis_4$, $axis_5$, $axis_6$, $axis_7$ along line 2 (Figure 3.8), rotations around $axis_2$, $axis_3$, $axis_4$, $axis_5$, $axis_6$, $axis_7$ along line 3 (Figure 3.9) and rotations around $axis_2$, $axis_4$, $axis_6$, along line 4 (Figure 3.10).

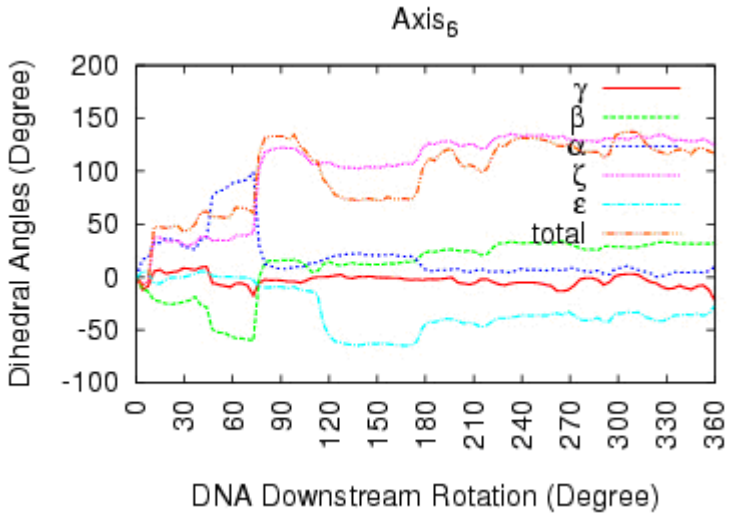

Figure 4.5 Changes in backbone dihedral angles for *positive rotation* around $axis_6$ along line 1

The *positive rotation* around $axis_6$ along line 1 obeys the *second scenario*. As shown in Figure 4.5 only the ζ angle changes and total change in dihedrals is around 120 degree at the end of the rotation. In the other *positive rotations* that compatible with this case, the change of the total dihedral is around 120 degree for the rotation around $axis_3$ and $axis_6$ along line 1 (Figure 3.7), 150 degree for the rotation around $axis_4$ along line 1 (Figure 3.7), 125 degree for the rotation around $axis_3$ along line 4 (Figure 3.10), 80 degree for the rotation around $axis_5$ along line 4 and 130 degree for the rotation around $axis_7$ along line 4 (Figure 3.10).

The changes in dihedral angles are given above in Figures from 3.2 to 3.10. As seen in these figures, the dihedral angle ζ has observed to be the most flexible dihedral angle in all kinds of rotations. The only additional change that contributes to the total rotation comes from the small changes in the dihedrals of **α** and **ε**. The other dihedrals of **γ** and **β** show no changes in all rotations. This observation is strongly in agreement with the literature studies [29, 30]. The only exceptions to this general observation are seen to be for $axis_7$ along line 4 of *negative rotations* (Figure 3.6), for $axis_5$ along line 2 of *positive rotations* (Figure 3.8), and $axis_3$ along line 4 of *positive rotations* (Figure 3.10). In these exceptions, we have observed that the **γ** and **β** dihedral angles changes as correlates to each other. The changes for these exceptions are around +90 degree for **γ** and -70 for **β** in Figure 3.6 and in Figure 3.10. The bad behavior in Figure 3.8 is different from the other two. As seen in Figure 3.8, only the **γ** angle is changing around -150 degree; however **β** value does not change at all.

4.2 Changes in α+γ in the rotations of the nicked DNA

Varnai *et al.* shows that **α+γ** should be around zero (0±60) to have minimum energy for an intact DNA [28]. For most of the rotations our results are convenient with this study. So we propose that this is also valid for a nicked DNA. As shown in Figure 4.6 the value of **α+γ** is observed to be around zero during the downstream DNA rotation. We have observed this property for both *positive and negative rotations*. The sum of these two angles are depicted in Figure 3.11 for the *positive and negative rotations* around parallel axis, in Figure 3.12 for all *negative rotations* and in Figure 3.13 for all *positive rotations*.

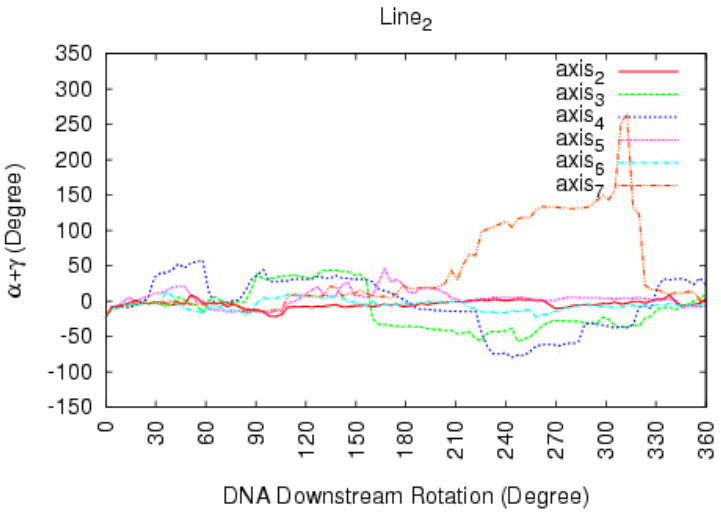

Figure 4.6 α+γ as a function of rotation angle for *negative rotations* along line 2

4.3 Changes in ε-ζ in the rotations of the nicked DNA

The difference of **ε** and ζ determines the state of DNA. **ε**-ζ is around -90 for B_I state, and +90 for B_{II} state [30]. At the beginning of the simulations, our nicked DNA is in B_{II} state as shown in Figure 3.15 and 3.16. For *negative and positive rotations*, we see that it passes B_I state and returns to B_{II} state or stays in B_I. Figure 4.7 is depicted **ε**-ζ values as we rotated the downstream DNA. As shown in this figure the difference of such dihedrals are around +90 at the beginning. In the *negative rotations* around $axis_3$, $axis_4$ and $axis_7$ DNA goes from B_I state and returns to B_{II} state and rotations around the other axes it stays in B_I state. We have observed this behavior in all of the rotations (Figure 3.14, 3.15 and 3.16).

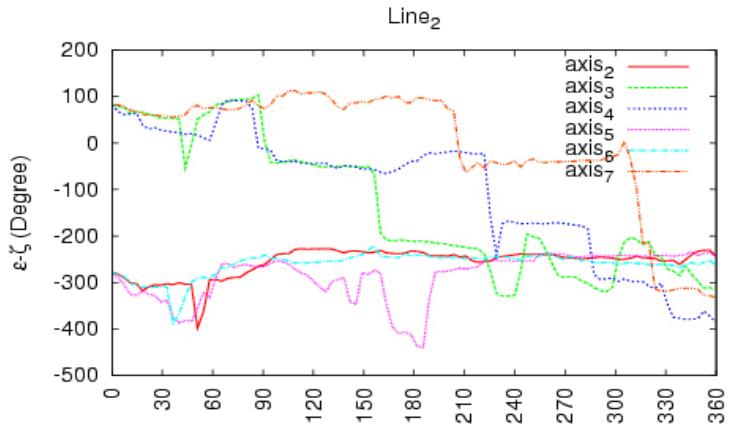

Figure 4.7 ε-ζ as a function of rotation angle for *negative rotations* along line 2

4.4 Correlation between dihedral angle ε and ζ

Hartman *et al*. [30] studied the correlation between **ε** and **ζ**. They showed two different state that two dihedral angle can be taken. Our simulations includes these states that we called as (1-2) state and (2-1) state. However, we have also obtained additional possible states for **ε** and **ζ** as listed in table 3.1. In Figure 4.8 these possible states that we observed are shown. The correlation between backbone dihedral angle between **ε** and **ζ** are depicted in Figure 3.17 for *negative* and *positive rotations* around parallel axis, in Figure 3.18, 3.19, 3.20, 3.21 for *negative rotations* around the other axes, in Figure 3.22, 3.23, 3.24, 3.25 for *positive rotations* around the other axes.

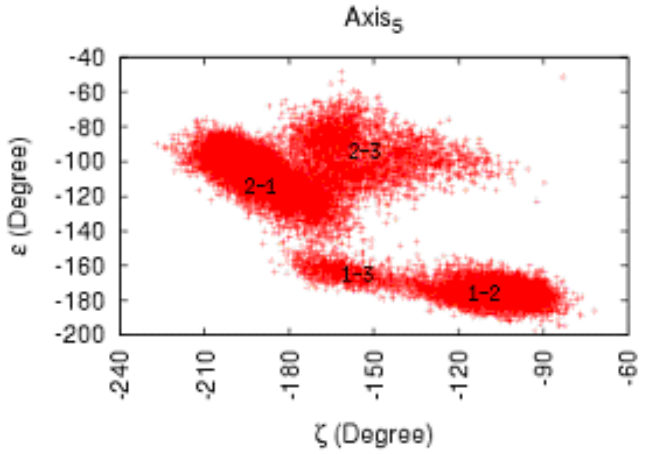

Figure 4.8 Correlation between dihedral angle **ε** and **ζ** for *negative rotation* around $axis_5$ along line 3

4.5 Correlation between dihedral angle γ and α

The correlation between dihedral angles **γ** and **α** and possible configurations are shown in Figure 3.26 to 3.34 for *negative* and *positive rotations*. (1-1) conformation is called as canonical conformation for B-DNA in literature [28, 29]. As seen in these figures all systems contain this conformation. Beyond the (1-1) conformation we have been observed the other conformations that we showed in table 3.2. (1-2) conformation is the second one that we see frequently in many simulations. In Figure 4.9 these conformations are shown for *positive rotations* around parallel axis. The correlation between dihedral angles **γ** and **α** is given in Figure 3.26 for *negative* and *positive rotations* around parallel axis, in Figure 3.27, 3.28, 3.29, 3.30 for *negative rotations* around the other axes, in Figure 3.31, 3.32, 3.33, 3.34 for *positive rotations* around the other axes.

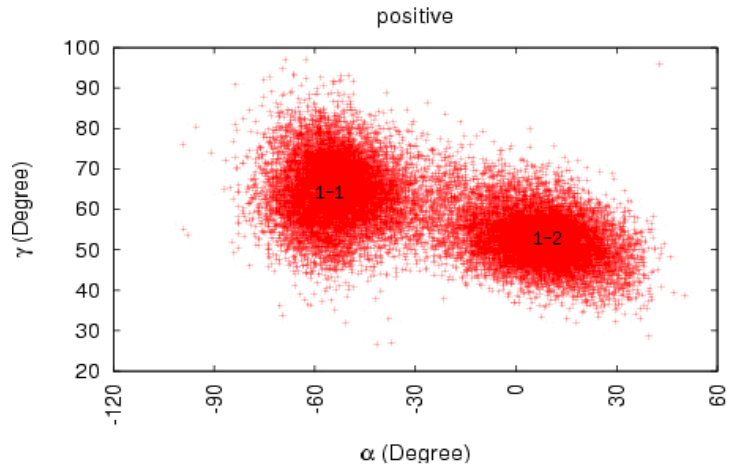

Figure 4.9 Correlation between dihedral angle **γ** and **α** for *positive rotation* around parallel axis

Structural analysis also includes deformations and base pair openings of DNA during the rotations. *Negative rotations* around parallel axis, around all axes along line 1, line 3, line 4, around $axis_2$, $axis_5$ and $axis_6$ along line 2 show base pair openings and large DNA internal deformations. In addition, we have not observed full DNA rotation at the nicked site. *Negative rotations* around $axis_3$, $axis_4$ and $axis_7$ along line 2 do not

show base pair openings. In these three simulations, the largest opening in the nick side (around 30 A^0) has been observed, and the DNA internal structure has been very smoothly conserved during whole downstream rotations. As a result, these rotations could be favorable *negative rotations*.

In the *positive rotations*, the DNA internal structure has been observed to deform more than that of *negative rotations*. In addition to this, *positive rotations* produced quite small amount of openings compared to the *negative ones*. We have observed openings up to 30 A^0 between the ends of the nick side for *negative rotations*, while the highest openings that we observed for *positive rotations* are around 10-15 A^0. These openings were observed in four different rotations: Parallel axis rotation (~13 A^0), rotations around $axis_4$ along line 2 (~16 A^0), around $axis_7$ along line 2 (~13 A^0), and $axis_3$ along line 4 (~15 A^0). In addition, these rotations do not show base pair openings. Therefore, they can be chosen as favorable *positive rotations*.

The most obvious difference between the *negative* and *positive rotations* is that *negative rotations* go into completion to the full rotation while *positive rotations* produce only a half rotation at the nick point. This is because, in all the systems that we propose as the most favorable DNA rotations; we see this behavior very clearly. Therefore, one of the most important result of our study is that although a single *negative rotation* at the downstream part of DNA brings the nick region to its original structure, and so after a single rotation DNA can religate back to the intact form, but for the *positive rotations* we need at least two downstream rotation to bring the DNA to its original form. The *negative rotations* here in our study are the rotations that relax positive supercoils, while the *positive rotations* are the ones that relax negative supercoils. Our result that we just mentioned suggest a quite important fact that in the relaxation mechanism of DNA within the human topoisomerase I, the linking number (Lk) can change by any integer value for the relaxation of positive supercoiled DNAs, while by only even numbers for negatively supercoiled DNAs.

The second important difference that we found between the *negative* and *positive rotations* is that the *negative rotations* are brought by the rotations around a single dihedral angle (which is the angle ζ), while *positive rotations* are produced through the

partial rotations around at least ζ and **ε**. This difference is very clear in our simulations, and first time reported here in our study.

In all of the systems that we can chose as the favorable ones, we see that **ε** and ζ make large amount of changes, and these changes are quite nicely correlated. This means that when **ε** decreases ζ increases and vice versa.

4.6 Mean force that is applied to bring the DNA rotations

Mean forces that is obtained from the applied external potentials have calculated as we mentioned in Chapter 2. The required mean forces are depicted in Figure 3.35 for *negative* and *positive rotations* around the parallel axis, in Figure 3.36 for *negative rotations* around all axes, in Figure 3.37 for *positive rotations* around all axes. An example of these graphs is given in Figure 4.10. The average forces needed to bring DNA rotation must be lowest. To find the most feasible axis of rotation(s), rotations that show peak value can be eliminated. In Figure 4.10 rotations around $axis_5$ shows a peak around 2700 pN and $axis_7$ shows a peak around 2600 pN. Also average mean forces of the rotations around $axis_3$, $axis_4$ and $axis_6$ are high. These systems also show base pair openings and large DNA internal deformations compared to the others. So these *negative rotations* along line 4 can not be good rotations.

The lowest average forces that we observed are around 1700 pN for *positive rotation* around parallel axis (Figure 3.35- green color), 1600 pN (Figure 3.36- green, blue and orange color) for *negative rotations* along line 2, 1600 pN (Figure 3.36- green and purple color) for *negative rotations* along line 3, 1600 pN (Figure 3.36- red color) for *negative rotations* along line 4, 1500 pN (Figure 3.37- purple and orange color) for *positive rotations* along line 1, 1700 pN (Figure 3.37- blue and orange color) for *positive rotations* along line 2, 1700 pN (Figure 3.37- aqua color) for *positive rotations* along line 3, 1500 pN (Figure 3.37- blue and red color) for *positive rotations* along line 4.

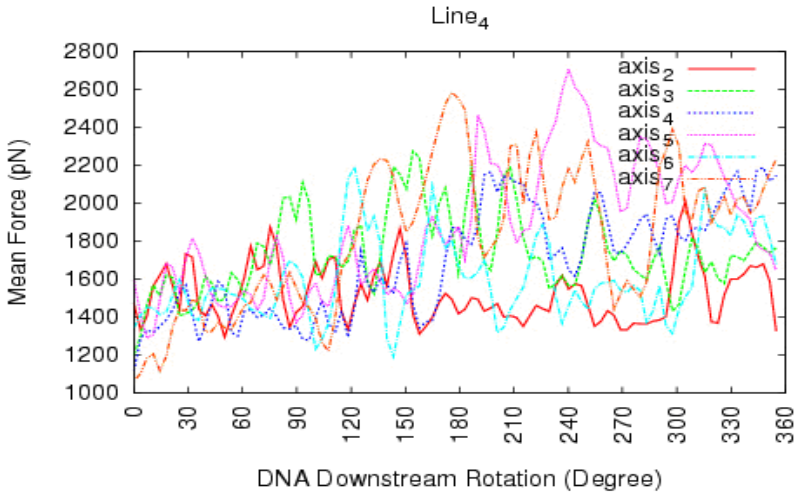

Figure 4.10 Mean force as a function of rotation angle for *negative rotations* along line 4

4.7 Two dimensional potential energy surfaces

The potential energies correspond to the total bonded and non-bonded interactions for the atoms in the +1 and -1 DNA base-pairs are calculated and given in Figure 3.38 for *negative* and *positive rotations* around parallel axis, in Figure 3.39, 3.40, 3.41, 3.42 for *negative rotations*, in Figure 3.43, 3.44, 3.45, 3.46 for *positive rotations* along line 1, 2, 3 and 4 respectively. These potential energy surfaces are quite important in studying the energetic pathways for the rotations. Such energy graphs for intact DNA have been studied extensively [29, 30] in literature, but the corresponding energy maps for a ***nicked DNA*** is reported first time here in our study. The potential energy of the +1 and -1 base pair must be follow lowest energy hypersurface for the feasible rotation. So we eliminated the rotations that follow higher energy hypersurface comparing the other rotations. As seen in Figure 4.11 the energy values are in between 9.5 kcal/mol to 14.5 kcal/mol. This rotation follows lowest energy. Hence, it could be feasible axis of rotation.

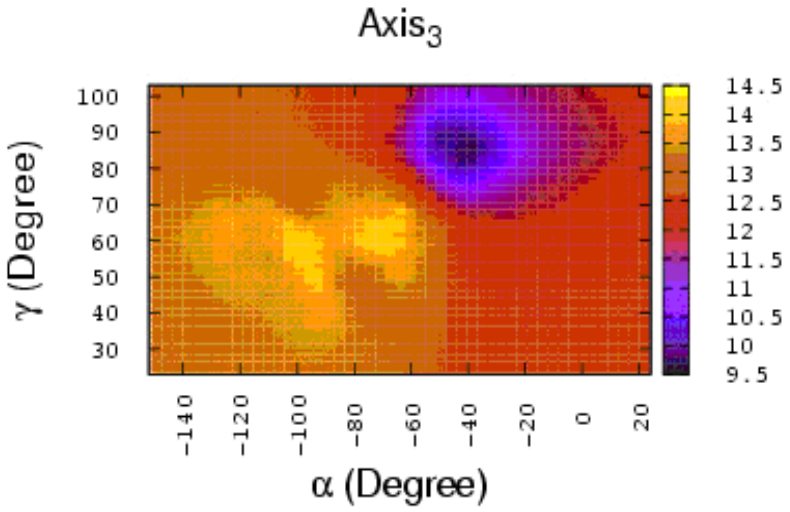

Figure 4.11 Energy map (kcal/mol) as a function of **α** and **γ** dihedral angle values for *negative rotation* around $axis_3$ along line 2

As a result of all structural and energetic analysis, we proposed that the most favorable *negative DNA rotations* are the rotations around $axis_3$, $axis_4$, and $axis_7$ along the line 2. We proposed that the *first scenario* is the most suitable case for the *negative rotations*. That is, the average forces needed to bring DNA rotation is lowest for those simulations where just only **ζ** makes a full rotation while the other dihedrals fluctuates within very small values. This observation is clear in Figure 3.36, where the green color that shows $axis_3$ rotations along line 2 has a minimum path compared to the others. The situation surprisingly is the same for $axis_4$ rotations along line 2 (Figure 3.36- blue color). This result is also supported by another observation, which is the minimum force path for rotations around $axis_7$ along line 2, in which we again observe the full rotation of dihedral **ζ**. This result is very nicely supported by the shape of DNA during the rotations. In these three simulations, the largest opening in the nick side (around 30 A^0) have been observed, and the DNA internal structure has been very smoothly conserved during whole downstream rotations. In addition, this result is nicely supported by the energetic considerations. The energy values are between 9.5 kcal/mol and 14.5 kcal/mol for $axis_3$, 10 kcal/mol and 14.5 kcal/mol for $axis_4$, 16 kcal/mol and 28 kcal/mol for $axis_7$. Among these three good systems, we can refer the rotation around $axis_3$ along line2 as

the best one for *negative rotations*, as it follows quite low potential energy path compared to the other two. Therefore, the mean force is minimum (around1600 pN) for this rotation by comparing the others (Figure 3.36). In addition, for the best *negative rotation* DNA passes B_I state and returns to B_{II} state (Figure 3.15 for line 2- green color) as the downstream part rotates.

However, as we look at the energetic analysis, we see that the *last scenario* requires much higher forces in many cases for *positive rotations*. For example, the graph of $axis_7$ in Figure 3.10 has shown total dihedral ends around 140 degree, and it has a peak value of mean force of around 2200 pN (Figure 3.40 for line 4- orange color). The similar observations are also seen for rotations around $axis_3$, $axis_4$, and $axis_6$ along line 1 as shown in Figure 3.7, the value of total dihedrals at the end of the rotation are 120, 100, and 120 degree respectively. Also, in Figure 3.36 for line 1 graphs show a peak around 2200 pN (green color) for $axis_3$, 2500 pN (blue color) for $axis_4$, and 2100 pN (aqua color) for $axis_6$. Similarly, for rotation around $axis_5$ along line 4 total dihedral ends around 80 degree (Figure 3.10 for $axis_5$), and has a peak force value of around 2200 pN (Figure 3.37 for line 4- purple color). Therefore, we suggest that the *first scenario* is the case for the *positive DNA rotations*. Among the systems that follow the *first scenario*, we can suggest that the following *positive rotations* are the favorable ones, as they require comparable less force for DNA rotations: rotations around parallel axis, around $axis_4$ and $axis_7$ along line 2, and around $axis_3$ along line 4. It is also very interesting to observe that the most flexible dihedral angles **ε** and **ζ** are the angles that bring the DNA rotation. However, among these systems, only rotations along line 2 show nice conservation of DNA internal structure, as the others brought base pair openings.

Therefore, among the rotations that end in zero degree total dihedral angle, the following rotation is chosen as the best one for positive rotations: rotations around parallel axis. The only noticeable difference between the three good axes of *positive rotations* is that the parallel axis rotation show quite low potential energy paths compare to the others. As seen in Figure 3.38, the potential energy of the +1 and -1 base pair follow the lowest energy hypersurface through which energy changes from 0 kcal/mol and 7 kcal/mol. The second lowest energy path was seen to be for $axis_3$ along line 4, which takes values between 12 kcal/mol and 26 kcal/mol. This system also show nice

opening of the nick side around 15 A^0. In addition, for the best *positive rotation* DNA passes B_I state and stays in B_I state (Figure 3.14- green color) as the downstream part rotates. It is indicate that for *positive rotations*, DNA must be rotated one more time to return its original position.

CHAPTER 5

CONCLUSIONS

The DNA rotation is the fundamental process in the removal of supercoils. The real-time DNA rotation within the enzyme has recently been simulated [14] as suggested by Stewart *et al* [5]. Later, results of these DNA rotations have been verified experimentally [15]. Therefore, we thought that detail investigation of a nicked DNA rotation around its intact backbone is a quite important study in understanding the structural deformations of DNA within the protein. As such a study is completely missing in literature; we have decided to work on these rotations. Our study focuses on different structural and energetic pathways that the nicked DNA undergoes during rotations, and aims to propose the best rotation scheme. In this regard, we have carried out 50 different simulations, in each one of them the DNA is rotated around a different axis.

At the end of our study, we propose the following novel conclusions (see the discussion section for the details):

(1) The DNA rotations are topology dependent. Linking number (Lk) can change by any integer value for the relaxation of *positive* supercoiled DNAs, while by only even numbers for *negatively* supercoiled DNAs.

(2) In the *negative rotations* DNA rotates around a single dihedral angle ζ, but in the *positive rotations* it rotates around at least two dihedral angles; ζ and $\boldsymbol{\varepsilon}$.

(3) The dihedral angle ζ has been found as the *most flexible dihedral* angle in all kinds of rotations.

(4) For the *negative rotations;* full DNA rotation is provided by *a full rotation* of the dihedral **ζ**, while for positive rotations the dihedral angles should rotate *partially* to bring a full DNA downstream rotation.

(5) The rotations around $axis_3$, $axis_4$, and $axis_7$ along the line 2 have been chosen as the candidates for the *negative* DNA rotations, and the rotation around *$axis_3$ along line 2* is chosen to be the best one.

(6) For *positive ones,* rotations around parallel axis, around $axis_4$ and $axis_7$ along line 2, and rotations around $axis_3$ along line 4 are observed to be the most logical ones, as we propose the parallel axis to be the best one in this case.

(7) For the best *negative* and *positive rotations* that we propose, the value of **α+γ** is observed to be around zero, which is perfectly in agreement with the literature studies.

(8) For *negative rotations,* DNA fallow a *structural path* from B_{II} state to B_I state and then back to B_{II} state. However, for *positive rotations* DNA goes from B_{II} state to B_I at the end of the downstream rotations. This means again that positive rotaitons should occur in even numbers, in agreemnt with our first conclusion.

REFERENCES

[1] F. B.Fuller, *Proc. Natl. Acad. Sci. U. S. A.* 1971, 68, 815

[2] D.L.Nelson, M. M. Cox *Principles of Biochemistry*

[3] L. M. Fisher et al., *Phil. Trans. R. Soc. Lond. B* 1992, 336, 83

[4] L. Postow, N. J. Crisona, B. J. Peter, C. D. Hardy, N. R. Cozarelli, *PNAS* 2001, 98, 8219

[5] L. Srewart, M. R. Redinbo, X. Qiu, W. G. J. Hol, J. J. Champoux, *Science* 1998, 279, 1534

[6] J. C. Wang, *Nature Reviews* 2002, 3, 430

[7] http://www-dna2006.cea.fr/lectures/Roca1.pdf

[8] M. R. Redinbo, L. Stewart, P. Kuhn, J. J. Champoux, W. G. J. Hol *Science* 1998, 279, 1504

[9] L. Stewart, G. C. Ireton, L. H. Parker, K. R. Madden, J. J. Champoux *The Journal of Biological Chemistry* 1996, 271, 7593

[10] L. Stewart, G. C. Ireton, J. J. Champoux *The Journal of Biological Chemistry* 1996, 271, 7602

[11] M. R. Redinbo, J. J. Champoux, W. G. J. Hol *Curr. Opin. Struct. Biol.* 1999, 9, 29

[12] M. R. Redinbo, J. J. Champoux, W. G. J. Hol *Biochemistry* 2000, 39, 6832

[13] T. K. Li, L. F. Liu, *Annu. Rev. Pharmacol. Toxicol.* 2001, 41, 53

[14] L. Sari, I. Andricioaei, *Nucleic Acids Research* 2005, 33, 6621

[15] R. F. Frohlich, C. Veigaard, F. F. Andersan *et al. Nucleic Acids Research* 2007, 35, 6170

[16] B. J. Alder, T. E. Wainwright *J. Chem. Phys.* 1959, 31, 2, 459

[17] W. B. Streett, D. J. Tildesley, G. Saville *Mol. Phys.*, 1978, 35, 3, 639

[18] A. McCammon, B. R. Gelin, M. Karplus *Nature* 1977, 267, 585

[19] M. Karplus, J. A. McCammon *Nature Structural Biology* 2002, 9, 646

[20] http://cmm.info.nih.gov/intro_simulation/node15.html

[21] MacKerell *et el. J. Phys. Chem. B.* , 1998, 102, 3586

[22] http://vit-embnet.unil.ch/MD_tutorial/

[23] http://www.charm.org

[24] M. Marchi, P. Ballone *J. Chem. Phys*., 1999, 110, 8, 3697

[25] P. H. Hünenberg *Adv. Polymer. Sci.* 2005, 173, 105

[26] M. S. Lee, F. R. Salsbury, Jr., and C. L. Brooks *J. Chem. Phys*., 2002, 116, 10606

[27] B.R.Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, M. Karplus *J. Comp. Chem.* 1983, 4, 187

[28] J. Mazur, R. L. Jernigan, A. Sarai *Biophysical Journal*, 1996, 71, 1493

[29] P. Varnai, D. Djuranovic, R. Lavery, B. Hartman *Nucleic Acid Research*, 2002, 30,5398

[30] B. Hartman, D. Piazzola, R. Lavery *Nucleic Acid Research*, 1993, 21, 561