1999

Molecular dynamics simulation study of the stability and conformation of spin-probe labeled DNAs

Eva Darian
West Virginia University

Follow this and additional works at: <https://researchrepository.wvu.edu/etd>

Recommended Citation

Darian, Eva, "Molecular dynamics simulation study of the stability and conformation of spin-probe labeled DNAs" (1999). *Graduate Theses, Dissertations, and Problem Reports*. 1003.
<https://researchrepository.wvu.edu/etd/1003>

This Thesis is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Thesis in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself. This Thesis has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.

Molecular Dynamics Simulation Study of the Stability and Conformation of Spin-probe Labeled DNAs

Eva Darian

**Thesis submitted to the
Eberly College of Arts and Sciences
at West Virginia University
in partial fulfillment of the requirements
for the degree of**

**Master of Science
in
Physics**

**Dr. Larry Halliburton, Ph.D., Chair
Peter M. Gannett, Ph.D.,
Martin Ferer, Ph.D.**

Department of Physics

**Morgantown, West Virginia
1999**

**Keywords: ds DNA, tx DNA, force field, nitroxide, quantum mechanics, molecular
mechanics, molecular dynamics,
Copyright 1999 Eva Darian**

Abstract

Molecular Dynamics Simulation Study of the Stability and Conformation of Spin-labeled DNAs.

Eva Darian

Nitroxide based spin-probes, observable by ESR, have been used to study protein structure and protein-substrate interactions. More recently, spin-probe labeled DNAs have been introduced to study its structure and dynamics. Molecular modeling has proven to be an increasingly powerful tool for studying structure and dynamics of biologically interesting molecules. As part of our work, we have conducted molecular dynamics simulations on both double-stranded (ds) and triple-stranded or triplex (tx) spin-probe labeled DNAs for comparison with unmodified ones and to design new probes. Nitroxide spin-labeled analogue of thymidine in which the methyl group is replaced by an acetylene-tethered nitroxide (T*) was inserted into the middle part of both ds DNA and tx DNA. Molecular mechanics force fields implemented in both Amber and Sybyl programs lack suitable parameters for nitroxides. These force fields were modified to contain new types of atoms designated as N.rad and O.rad with slightly different values than those reported by Barone, *et al.* for MM3 and MM+ force fields (Barone et al., J. Am. Chem. Soc. 1998, 120, 7069-7078). In addition new type of carbon atom designated as C1 (for sp hybridized carbon) with corresponding parameter was developed and added to Amber force field. We conducted MD simulations on both modified and unmodified DNAs to compare their stabilities and relative motion of the nitroxide with respect to a DNA molecule. Some results are in the process of analysis. We expect to see slower motion of nitroxide associated with triplex, rather than with duplex. The simulations will also be correlated with experimental studies conducted on different DNA sequences. (Supported by NIH Grant R15 GM7630).

For my parents

ACKNOWLEDGMENTS

I would like to thank Dr. Peter M. Gannett, my advisor, for his tremendous help in achieving my degree and enlarging my knowledge in this area. Also, I would like to thank my committee members Dr. Larry Halliburton and Dr. Martin Ferer for finding time between their busy schedules to serve on my committee.

I sincerely thank my colleagues Jeannine Powell and Irina Lapshina. Special thanks to Jeannine for helping me with the synthesis of some compounds that I have done up to date.

Finally I would like to thank my husband Nelson for his patience and support toward my work. Also my parents, though being far away, for their encouragement.

Table of Contents

Chapter	page
Abstract	ii
Acknowledgments	iv
Table of Contents	v
List of Figures	vii
List of Tables	ix
1. Introduction	1
1.1 Description of the problem	1
1.2 Antisense strategy	3
1.3 Triplex agent strategy	5
1.4 Description of the solution	8
1.5 Overview of the work	10
2. Modeling of nitroxides	13
2.1. Quantum mechanical background	13
2.2. Molecular mechanics	16
2.3. Force fields for nitroxides	19
2.4. Results and discussion	22
3. Molecular dynamics background	28
3.1. Molecular dynamics (MD)	28
3.2. Minimization technique	31
4. MD simulations of ds DNAs	32

5. Molecular dynamics simulations of tx DNAs	38
6. Discussion and conclusions	50
References	51
Appendix A	53
Appendix B	60

List of Figures

Figure		page
Figure 1.1.	Watson-Crick base-pairing	1
Figure 1.2.	Base triplet formation schemes	2
Figure 1.3.	Modified thymidine, thymidine and 5-membered spin-probe molecules	10
Figure 1.4.	Training nitroxide radicals	11
Figure 2.1.	Side view 1a and front view 1b of MNO molecule.....	23
Figure 2.2.	Side view 2a and front view 2b of TBNO molecule	24
Figure 2.3.	Side view 3a and front view 3b of 5-membered ring nitroxide molecule	26
Figure 4.1.	Equilibration of ds DNA without spin-probe	34
Figure 4.2.	Equilibration of ds DNA with spin-probe	35
Figure 4.3.	RMS deviation of ds DNA from original structure with and without spin-probe	36
Figure 5.1.	Equilibration of tx DNA without spin-probe	39
Figure 5.2.	Equilibration of tx DNA with spin-probe	40
Figure 5.3.	RMS deviation of antiparallel tx DNA from original structure with and without spin-probe	41
Figure 5.4.	RMS deviation of parallel tx DNA from original structure	

with spin-probe	42
Figure 5.5. Fluctuations of two sets of atoms of the nitroxide molecule	43
Figure 5.6. Torsional parameters of the nitroxide	44
Figure 5.7. Plots of torsional parameters for spin-probe labeled ds DNA	46
Figure 5.8. Plots of torsional parameters for <i>antiparallel</i> spin-probe labeled tx DNA	47
Figure 5.9. Plots of torsional parameters for <i>parallel</i> spin-probe labeled tx DNA	48
Figure 5.10. Plots of torsional parameters for spin-probe ds DNA in vacuo	49

List of Tables

Table	page
Table 2.0. Modified parameters for nitroxides	20
Table 2.1. Geometrical parameters for Dimethyl Nitroxide	
Radical (DMNO)	23
Table 2.2. Geometrical parameters for Di(<i>tert</i> -butyl)aminoxyl	
Radical (TBNO)	25
Table 2.3. Geometrical parameters for 5-membered ring nitroxide	27
Table 5.1. Torsional parameters of the nitroxide	44

CHAPTER I

Introduction

1.1. Description of the problem

Much of the drive to understand nucleic acid interactions has come from the interest in understanding the mode of action of existing medicinal agents and from the desire to develop a new generation of superior drugs. Few drugs have been designed from a basic understanding of the biochemical mechanisms that are responsible for the disease state. Besides, most drugs are designed to act at a level of the enzymes, which also tends to make them less than optimal. Since the DNA and messenger RNA (mRNA) are the key species involved in the enzyme production, it would be more efficient to design drugs that could act directly on either of them.

One way to improve drugs is to target mRNA with the single-stranded DNA (ss DNA) thus forming a DNA-RNA hybrid. The binding of the two strands follows Watson-Crick base pairing rules, where purines (Pu) adenine and guanine (A and G) pair with pyrimidines (Py) thymine and cytosine (T and C) by forming two and three hydrogen bonds between A:T and G:C bases, respectively (Figure 1.1). This kind of hybrid formation can lead to inactivation of the complementary mRNA sequence.

Figure 1.1. Watson-Crick base-pairing

The second, and even more efficient way, is to block the double-stranded DNA (ds DNA) with another ss DNA by forming triple-stranded DNA complex (tx DNA) or simply triplex. This, for example, could prevent DNA from producing more mRNAs, binding of the regulatory protein to ds DNA, or could result in inhibition of the human oncogene *c-myc* transcription *in vivo*. But the requirements for triplex formation are more restricted, because it must follow the Hoogsteen base pairing rules, which means, that only certain and fewer base pairing combinations are possible. The resulting base-triplets form the core of a triple helix (Figure 1.2), where C^+ is the N-3 protonated form of cytosine.

Figure 1.2. Base triplet formation schemes

The widespread occurrence of polypurine-polypyrimidine tracts in eukaryotic DNA suggests that these sequences may have a biological function as well as therapeutic significance. Analysis of eukaryotic sequence databases reveals thousands of polypurine-polypyrimidine

tracts, many with the potential for triplex formation. Because of the third strand base pairing rules, a polypurine-polypyrimidine region will define a unique third strand pairing sequence. This can have a great potential for therapeutic applications for treatment of certain diseases and viral infections. Although polypurine-polypyrimidine sequences are important in gene regulation, the nature of their involvement remains to be elucidated. Given a unique sequence of DNA, it is possible to design therapeutic oligonucleotides (ONs) that will specifically bind to it and to inhibit gene expression. Development of new therapies based on ONs is at the stage of intense research. During the past 15 years, ONs have been increasingly used to inhibit gene expression. There are two basic approaches to implementing ON-based therapy, which are described below.

1.2. Antisense strategy

Over the past few years there have been many advances in the design and characterization of antisense ONs for the treatment of various gene based human diseases such as acquired immune disease (AIDs). The promise of controlling gene expression in a specific and efficient manner has spurred a large research effort to develop antisense DNA therapy. To be effective as an antisense drug an ON must possess a number of properties such as nuclease resistance, stable and specific complex formation with the target mRNA and cellular uptake. In this approach, designed ONs are complementary to mRNA, which is synthesized on a DNA template in a process known as DNA transcription. DNA carries the information for protein synthesis, as well as transfer (tRNA), ribosomal (rRNA) and other RNA molecules that have structural and catalytic activities. Depending on the target site, the antisense oligonucleotide can act to inhibit translation (cap, activator protein binding site or AUG, the sequence nucleotides

that initiate peptide synthesis), to block splicing (splice donor or acceptor site) or to induce cleavage by endogenous RNase H, which cleaves RNA-DNA hybrids. An oligonucleotide strand that is longer than the target sequence could fold back into hairpin and form a triplex at a target sequence. Such hairpin could form an especially stable and specific clamp to inhibit translation.

Although naturally occurring nucleic acids are effective templates for the design of antisense drugs, several problems have been identified with their utilization. These include poor absorption through cellular membranes, low affinity for the target mRNA sequences, and degradation by naturally occurring nuclease enzymes. Endonucleases, for example, break the phosphodiester bond at any point along the DNA chain. They can nick and digest covalently closed circular DNA and can introduce breaks internally within a linear DNA molecule. Exonucleases, in contrast, require a free end at which to initiate digestion, which can occur in either direction, 5' to 3' where 5'-PO₄ terminal is required, or 3' to 5', where 3'-OH terminus is required. Depending on the type of nuclease, specificity for either single- or double-stranded DNA is possible.

In trying to solve these problems synthetic modifications of DNA backbone by conversion of phosphates to phosphorothioate or methyl phosphonate and base modifications have been proposed. These methods can impart nuclease resistance and simultaneously improve cellular absorption. Once ONs are absorbed, they can bind to mRNA and inhibit expression of particular genes responsible for disease state. In order to bind to their target, ONs must be designed with Watson-Crick complementary sequence, which in turn provides the selectivity for the desired target. Selectivity depends also on base sequence and ONs length. Almost complete selectivity is predicted for ONs 20-30 bases in length. For shorter chains, the chances of a given

sequence occurring more than once in a mRNA are increased hence selectivity may be decreased.

1.3. Triplex DNA strategy

Alexander Rich, David Davies and Gary Felsenfeld first discovered the triple stranded form of DNA in 1957 while studying synthetic nucleic acids. At the time, triplex DNA seemed to be an interesting anomaly of double helix DNA discovered just a few years earlier by Watson and Crick, without any physiological or practical importance. It rapidly sank in the flood of new information about nucleic acids unleashed by the discovery of the double helix. But triplex DNA was not entirely forgotten. In 1974, Arnott conducted X-ray diffraction studies on **T:A-T** triplets. Later, in 1976, the first calculations were performed on generating the models using a computerized, linked-atom procedure which preserves standard bond lengths, bond angles, sugar rings conformations, and constrains the helices to have the pitches and symmetries observed in X-ray diffraction data. This was fortunate because later work showed that tx DNA may have a biological role and may be put to good use after all.

In 1987, Dervan and his colleagues obtained results indicating that a third strand of DNA could bind to a stretch of natural DNA containing actual genes and form a triplex DNA. At the time, several labs were working on the idea of producing a new type of “molecular scissors” for cutting DNA and that the tx DNA technique could be used to cut a specific DNA sequence. If this could be achieved, it would be possible to cut at a few selected sites giving a much more manageable collection of large DNA pieces. This is in contrast to restriction enzymes, which cut at many sites because they home in on DNA sequences containing few (~ 4-8) base pairs. Thus, restriction enzymes mince the genome into a multitude of small bits that are hard to separate and

analyze. Finally, molecular scissors would be a great help in mapping and sequencing the human genome.

Interest in triple-helical conformations has been stimulated by several recent discoveries. First, in 1987, Mirkin *et al.* showed that an intermolecular triplex could form within a single homopurine-homopyrimidine duplex DNA in supercoiled DNA. This observation reawakened a general interest in triplex DNA because many sequences in the human genome have the potential to form intermolecular triplex structures that are associated with regulatory regions of genes. Then Frank-Kamenetskii and co-workers demonstrated that in addition to a continuous strand of purine bases, the homopurine-homopyrimidine region must contain mirror repeat symmetry. Mirror repeat is a region of DNA that has the same base sequence reading in both the 3' and the 5' direction from a central point in one strand of DNA. For example, a sequence like GAATTAAG, can represent a mirror repeat. These sequences could form triplex by dissociation of half of the mirror repeat duplex followed by triplex formation between the free pyrimidine strand and the remaining half of the mirror repeat. Such conformations could occur in cells in similar regions and may represent a gene control mechanism. In other words, what was shown was the involvement of triplex DNA formation in cellular regulation.

The discovery of the potential for triplexes to affect gene expression lead to the antigene therapy and this may become a crucial step for treatment of gene based diseases such as human immunodeficiency virus (HIV). The triplex-forming oligonucleotides (TFO) can be prepared that target specific sites on naturally occurring genes and this kind of formation can disrupt the copying of the genetic material into messenger RNA, the first step in protein synthesis. Within the past seven years triplex DNA turned out to have therapeutic potential as well.

Triplex DNA formation should be based on the specific rules like those for ds DNA. Thus, there are specific motifs that allow for the formation of DNA triplets (Figure 2). As mentioned earlier, there are four different ways in which homopurine-homopyrimidine mirrors repeat sequence can fold into an intermolecular triplex. The most common structures are *Py-Pu:Py* and *Pu-Pu:Py* configurations (T:A-*T* or C:G-*C*⁺ and C:G-*G* or T:A-*A*) where A:T and G:C corresponds to a normal ds DNA of Watson-Crick base pairing and *T* or *C*⁺ and *G* or *A* are the third bases that associate with ds DNA through hydrogen bonding interactions in Hoogsteen and reverse Hoogsteen base pairing fashion, respectively. With this knowledge, an oligonucleotide sequence, capable of forming a tx DNA complex can be designed to target a specific sequence in a ds DNA molecule. Moreover, since gene-based diseases like AIDs have DNA sequences that are unique to the virus that causes them, an ON can be designed to bind to the viral genomic DNA sequence and thereby inhibit selected processes in a diseased cell, to terminate the result of the disease process, or simply kill the cell. A sufficiently long DNA strand can bind with very high selectivity to a desired duplex gene. Also, additional agents to cross-link or cleave the genomic DNA could be attached to TFOs to permanently inactivate the gene with minimal or no toxic side effects.

There are several difficulties that are hampering the development of the antigene or triplex approach. Although triplexes can be formed with high specificity, they generally have lower stability than duplexes under typical physiological conditions. There are many methods available for studying tx DNA *in vitro*, including thermal denaturation, circular dichroism, x-ray crystallography, NMR and molecular modeling. However, few of these methods can be used *in vivo* and there is no direct methodology available for studying the delivery of antisense or triplex oligonucleotides to their targets directly in cellular systems. Only indirect methods can be used

as the piece of DNA that is being investigated is increased. Thus, the amount of information that can be obtained from more complex molecular systems is limited. Since the ON sequence was designed to be complementary to specific target, then in cells it has been assumed that the binding should occur. However, there are some examples showing the binding of non-sense ONs to the target sequences. A non-sense sequence is a random sequence of bases non-complementary to mRNA sequence, and would not be predicted to bind based on Hoogsteen base-pairing schemes to their target.

A major search is under way to find small molecules that can stabilize triplex formation selectively. Triple helix stability can be enhanced by the use of modified nucleotides. For example, 5-methylcytosine increases stability at neutral pH, and 5-bromouracil can usefully replace thymine. Probes are needed that can directly observe the binding of antisense or triplex ONs in cellular systems and that distinguish a non-specific response of cells from that specific (i.e. DNA-mRNA or tx DNA formation) due to binding. Moreover, the existence of such a probe would be useful for examining processes like cellular uptake of ONs and their migration into the cellular nucleus.

1.4. Description of the solution

To address some of the major problems hampering the progress and development of oligonucleotide based therapies and provide a tool to probe the binding, structure, dynamics and stability of antisense and triplex forming ONs upon binding to their targets, the spin probe technique is introduced. The spin probe labeling technique has been used to investigate, monitor or analyze the environment of the spin-probe labeled species. Spin labels are comparatively stable nitroxide radicals, generated by the reaction between spin traps (nitroso compounds or

nitrones) and short-lived free radicals, and since they possess an unpaired electron, they can be observed by ESR. In addition, most biological molecules are diamagnetic and therefore the ESR signal of the radical attached to a biopolymer can be easily distinguished from the background.

Nitroxide based spin-probes have been used to study protein structure and protein-substrate interactions. Non-ON based spin probes have been used to study enzymatic processes up to cellular level, where the conformational changes of supercoiling of circular DNA upon its interaction with the enzyme and changes of ESR spectrum of spin-labeled T₂ phage DNA in the complex with RNA polymerase were observed. More recently, spin-probe labeled DNAs have been introduced to study its structure and dynamic properties and their use in our work will allow us to monitor the interactions of antisense and triplex agents with their targets. The ESR spectrum is very sensitive to its microenvironment and permits easy registration of even subtle alterations in it. Thus information can be gained about conformational properties of all regions near the point of attachment of a spin-label and about the dynamic behavior of the object as a whole. The size of the DNA and complexity of the system is not a restriction to the spin probe method.

In order to better understand biological information transfer, molecular interactions of nucleic acids and polymorphic character of nucleic acid conformation, it is important to understand the structure, dynamics, stability and relative flexibility of ds DNAs and tx DNAs as well as DNA-RNA duplexes. A large number of theoretical calculations of nucleic acid structure have been performed. The application of computer-based models using analytical potential energy functions within the framework of classical mechanics has proven to be an increasingly powerful tool for studying structure and dynamics of small organic molecules or larger, biologically interesting, molecules. These methods have lead to an increased understanding of

DNA structure, conformation and reactivity. Likewise they can be used to suggest or guide new research. In this work we will discuss some aspects of molecular dynamics simulations performed on several DNA molecules with and without modified nitroxide spin-labeled bases, comparing their stability and relative motion of the nitroxide with respect to the DNA molecule.

1.5. Overview of the work

Molecular dynamics (MD) simulations of double- and triple-stranded unmodified DNA sequences $d[(T_{15}):(A_{15})]$ and $d[(T_{15}):(A_{15})-(T_{15})]$ as well as modified or spin-labeled DNA sequences of $d[(T_7T^*T_7):(A_7A_7)]$ and $d[(T_7TT_7):(A_7AA_7)-(T_7T^*T_7)]$ are presented in the current work. Spin-labeled thymidine (T^*) **1**, represents a nitroxide spin-labeled analogue of thymidine (T) **2**, in which the methyl group is replaced by an acetylene-tethered nitroxide radical **3** (Figure 1.3).

Figure 1.3. Modified thymidine, thymidine and 5-membered spin-probe molecules

But, before initiating any molecular mechanic or molecular dynamics calculations on DNA structures, quantum mechanical calculations of the spin-probe itself should be performed. The optimized structure of a molecule obtained by optimization at the quantum mechanical level

along with the values of the electrostatic potential at each grid point, calculated from quantum mechanical wavefunction, must be used to develop optimized AMBER 5.0 force field parameters and to derive a set of atom-centered point charges. As a starting point in developing parameters from quantum mechanical calculations we will use the training set of two small molecules, which are representative of the different ranges of pyramidalization for the NO group. The structures of the radicals are sketched in Figure 1.4 and the bis-(*tert*-butyl) nitroxide (TBNO) radical is planar, whereas the dimethyl nitroxide (MNO) radical is slightly pyramidal. Upon obtaining geometries close to the reported experimental ones, we will extend similar calculations to the nitroxide radical of our interest **3**. The refined structure of **3** will replace the methyl group of thymidine to produce the spin-labeled thymidine analogue (T*) **1**.

Figure 1.4. Training nitroxide radicals

Incorporation of the modified thymidine **1** should not lead to any significant increase in steric interactions, since this probe is directed away from and perpendicular to the DNA helical axis. There is some experimental evidence, that the acetylene tether of **1** will actually improve the stability of tx DNA formation. With the aid of molecular dynamics simulation the source of this improved stability explored. MD simulation results on both modified and unmodified DNAs will be used to compare their stabilities and relative motion of the nitroxide with respect to a

DNA molecule depending on whether it is in ds DNA or tx DNA. Information about the rate of the motion of the nitroxide extracted from the dynamics runs can be related to the correlation time (τ_c). This is the time required for the label to rotate by an angle of about 40° , and is related to the experimental parameters of line width and line shape of the ESR spectrum and microparameters of the radical. Thus, the simulations will also be correlated with experimental studies conducted on different DNA sequences. Description of theoretical background and computational details are given in the following chapters.

CHAPTER II

Modeling of nitroxides

2.1. Quantum Mechanical background

Nitroxides are one of the few classes of organic free radicals that are stable under ordinary conditions; they are isolable compounds and can be stored indefinitely. This has allowed the determination of a number of molecular structures by using x-ray crystallographic methods and the consequent unequivocal correlation between structural and spectroscopic parameters. This characteristic, in turn, coupled to the strong localization of the unpaired spin in the NO moiety, makes nitroxides ideal "Spin Labels" useful for exploring the structure of short-lived free radicals by Electron Spin Resonance (ESR) spectroscopy.

The use of spin-labeling to study some characteristics of nucleic acids such as their base dynamics, their interactions with complementary nucleic acids, and their interactions with nucleic acid binding proteins and drugs has been well documented in recent years. Until recently available force fields lacked suitable parameters for nitroxide species. Barone et al., studied the structures and spectromagnetic properties of some model nitroxides by self-consistent hybrid of Hartree-Fock and density functional methods (B3LYP) obtaining results close to experimental data. From the computed structures, together with the available experimental data, new parameters for the NO moiety have been developed for two of the most commonly used force fields MM3 and MM+. These force fields are applicable for small molecules, with the maximum number of atoms limited to 500, but they provided the background for computations of reliable structures of large molecules like proteins and nucleic acids, containing nitroxide systems.

The quantum mechanics calculation program Q-Chem, where the quantum-mechanical calculations are based on unrestricted Kohn-Sham approach to density functional theory (DFT),

was used in this work. Density functional methods following the Kohn-Sham formalism are based on parameter-free theory, i.e., they attempt to find solutions "from first principles" to the self-consistent functional (SCF) mean-field model of the electronic structure, while treating the electron correlation problem differently than the post-Hartree-Fock techniques. The basic idea of the DFT is to use the electron density $\rho(\mathbf{r})$ as the variable of the system instead of electronic wavefunction. The total energy of an n-electron system can be expressed as

$$E = E_T + E_V + E_J + E_{XC} \quad 2.1$$

E_T is the kinetic energy of noninteracting electrons with the same total electron density as the actual system of interacting electrons,

E_V is electron-nuclear interaction energy,

E_J is the Coulomb self-interaction of the electron with density $\rho(\mathbf{r})$ and

E_{XC} is the energy of exchange interactions, correlation effects, and the difference between exact kinetic energy and that of the reference system of noninteracting electrons with the density $\rho(\mathbf{r})$. Adopting an unrestricted format, the α and β total electron densities can be written as

$$\rho_\alpha(\mathbf{r}) = \sum_{i=1}^{n_\alpha} |\Psi_i^\alpha|^2 \quad 2.2 \quad \text{and} \quad \rho_\beta(\mathbf{r}) = \sum_{i=1}^{n_\beta} |\Psi_i^\beta|^2 \quad 2.3$$

$$\rho(\mathbf{r}) = \rho_\alpha(\mathbf{r}) + \rho_\beta(\mathbf{r}) \quad 2.4$$

The components of Equation 2.1 can now be written as

$$\begin{aligned} E(\rho) = & -\frac{1}{2} \sum_i \int \Psi_i^*(\mathbf{r}_1) \nabla_1^2 \Psi_i(\mathbf{r}_1) d\mathbf{r}_1 + \sum_A \int \frac{Z_A}{|\mathbf{r}_1 - \mathbf{R}_A|} \rho(\mathbf{r}_1) d\mathbf{r}_1 \\ & + \frac{1}{2} \iint \frac{\rho(\mathbf{r}_1) \rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 + E_{XC}(\rho) \end{aligned} \quad 2.5$$

Although innocuous in appearance, E_{xc} contains, buried within it, all the details of two-body exchange and dynamical correlation and a kinetic energy component as well and requires a more in depth discussion. The B3LYP hybrid functional, which combines Hartree-Fock and Becke exchange terms with Lee-Yang-Parr correlation functional, with the standard 6-31G* basis set have been used for geometry optimization of a training set of three small nitroxide molecules MNO, TBNO and 5-membered ring nitroxide radical. Geometry optimization refers to the determination of stationary points, principally minima and transition states on molecular potential energy surfaces. It is an interactive process, requiring the repeated calculation of energies, gradients at each optimization cycle until convergence is achieved. The form of the B3LYP functional is given in Equation 2.6.

$$E_{xc}^{B3LYP} = (1 - a_0)E_x^{LSDA} + a_0E_x^{HF} + a_x\Delta E_x^{B88} + a_cE_c^{LYP} + (1 - a_c)E_c^{VWN} \quad 2.6$$

where ΔE_x^{LSDA} is local spin-density approximation, which uses the standard local exchange functional and local correlation functional of Vosko, Wilk and Nusair ΔE_c^{VWN} .

ΔE_x^{HF} is exact Hartree-Fock exchange

ΔE_x^{B88} is Becke's gradient correlation to the exchange functional

ΔE_c^{LYP} is Lee-Yang-Parr correlation functional

Becke suggested coefficients $a_0 = 0.2$, $a_x = 0.72$, $a_c = 0.81$ based on fitting to heats of formation of small molecules. Final geometry output is used as input for Gaussian94 single point calculation in order to get electrostatic points in a form, suitable for use in restrained electrostatic potential fit model RESP. RESP will fit the quantum chemically calculated potential at molecular surface using an atom-centered point charge model.

2.2. Molecular Mechanics (MM)

The applications of molecular mechanics (MM) have employed energy minimization, molecular dynamics, and Monte Carlo methods to move on the analytical potential energy surfaces. They are capable of giving insight into the entire spectrum of non-covalent interactions between molecules, and, when combined with quantum mechanical electronic structure calculations, which model covalent bonding changes essentially all molecular reactions and interactions can be examined. Given their importance, much effort has gone into establishment of both the functional form of equations and the parameters in order to apply such analytical energy functions or otherwise called "force fields".

Molecular mechanical force fields are based on the concept that the potential surfaces and physical properties of molecules may be calculated using classical mechanics functions. A molecule is thought of as a collection of atoms, which interact with each other in a way that can be described by simple potential energy harmonic functions or force fields. One of the most common force field components is bond stretching and compression and can be approximated by a Hook's law. The constants, which describe the force constant of a bond and the equilibrium bond length between the atoms, are called force field parameters. These parameters are empirically derived from the physical properties of a prototypical set of molecules.

The total energy of a molecule calculated with a MM force field represents a sum of individual component functions of the potential, such as bond stretching (str), angle bending (bend), electrostatic potential (ele), torsional strain (tor) van der Waals (vdw), and out of plane (oop) bending and is shown in Equation 2.7. Out of plane bending potential term also called

improper torsional strain potential. Using this approach the relative energies of two or more structures can be estimated with given force fields.

$$U_{tot} = \sum_{bend} U + \sum_{str} U + \sum_{tor} U + \sum_{vdw} U + \sum_{ele} U + \sum_{oop} U + \dots \quad 2.7$$

The numerical value of the energy U_{tot} , derived from any empirical force field, for a molecule has no physical meaning. However, energy differences between two or more conformations of a single molecule are physically meaningful, as are comparisons of relative conformational energies, or conformational energy differences, between different molecules. Additional terms, for example cross product terms, energy due to dipole-dipole interactions or twisting about a bond may be included to achieve the desired level of accuracy. Inclusion of all possible potential energy terms to the MM force field can produce high-quality models. But this approach, in turn, becomes expensive in terms of computational resources, so that certain restrictions need to be applied to a model.

AMBER is based upon a valence force field, and the potential function, which well describes interactions between atoms for large molecules, is given in Equation 2.8.

$$\begin{aligned} U_{tot} = & \sum_{bonds} K_r (r - r_0)^2 + \sum_{angles} K_\theta (\theta - \theta_0)^2 + \sum_{dihedrals} \sum_n \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] + \\ & + \frac{1}{VDW_{scale}} \sum_{j=1}^{bonds} \sum_{i>j}^{bonds} \epsilon^* \left[\left(\frac{R_{ij}^*}{r_{ij}} \right)^{12} + \left(\frac{R_{ij}^*}{r_{ij}} \right)^6 \right] + \sum_{j=1}^{Hbonds} \sum_{i>j}^{Hbonds} \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) + \quad (2.8)^a \\ & + \frac{1}{EEL_{scale}} \sum_{j=1}^{atoms} \sum_{i>j}^{atoms} \frac{q_i q_j}{r_{ij} \epsilon} + \sum_{constraiats} K_{const} (x - x_0)^2 + \sum_{cap\ atoms} K_{cap} (y - y_0)^2 \end{aligned}$$

^a K_r in (kcal/mole·Å²), K_θ in (kcal/mole·(rad)²), are bond stretching and angle bending constants. r_0 , θ_0 are equilibrium length and angle. V_n is a magnitude of torsion in kcal/mole. n is a periodicity of the torsion. γ is phase offset in degree. ϵ_{ij}^* is van der Waals well depth for given atom, $\epsilon_{ij}^* = (\epsilon_{ii}^* \epsilon_{jj}^*)^{1/2}$. R_{ij}^* is van der Waals

The first three terms that describe bond, angle, and dihedral potentials may be thought of as bonding interactions. Dihedral energies represented in most cases with simple set of parameters, often only specified by the two central atoms and may be thought of as a truncated Fourier series. Improper torsional angles, in which the atoms are not sequentially bonded to each other, use the same potential function in Amber force field as torsional angles. Improper dihedral angles are used to preserve planarity of specific ring fragments. The next three terms, the van der Waals interactions represented by a 6-12 potential, hydrogen-bonding, and electrostatic potentials modeled by a Coulombic interactions of atom-centered point charges, describe the non-bonded interactions. Electrostatic and van der Waals interactions are only calculated between atoms in different molecules or for atoms in the same molecule separated by at least three bonds. Those non-bonded interactions separated by exactly three bonds, "1-4 interactions", are reduced by the application of a scale factor. The charge calculation method is based on the restrained electrostatic potential fit model (RESP-fit). This model involves a least-squares fit of the charges to the electrostatic potential with the addition of hyperbolic restraints on charges on non-hydrogen atoms. Finally, the last two terms in the equation are used to harmonically restrain Cartesian positions. The first of these is used to constrain a set of reference coordinates and is often used during minimization procedures to keep a structure from significantly deviating from a reference structure. The second term is used to restrain solvent molecules spatially.

radius for given atom $R_{ij}^* = (R_{ii}^* R_{jj}^*)^{1/2}$. ϵ is dielectric constant. q_i, q_j are partial atomic charges. C_{ij} and D_{ij} are the coefficients depicting repulsive and attractive hydrogen atom-hydrogen acceptor interactions.

2.3. Force field parameters for nitroxides

As it was mentioned before, we performed quantum mechanical computations to compare the optimized geometries of these radicals with experimental ones reported in the paper by Barone et. al.. We modified these parameters to make them useful for the Molecular Mechanics (MM) force fields implemented in both Sybyl and Amber programs. There are two characteristics of the NO moiety that need special attention for the correct development of the force field. The first aspect is that the unpaired spin of nonconjugated nitroxide is well localized on the diatomic functional NO moiety. Since both force fields lack suitable parameters for nitroxides, addition of two new atom types to current force fields, should be sufficient to account for the particular nature of the bond. The new atom types, referred to as N.rad and O.rad for Sybyl and NR and OR for Amber force fields respectively, correspond to the nitrogen and oxygen atoms of the nitroxide group. The second aspect is that both the preference for a planar or pyramidal environment of the N atom and the partitioning of the spin density between the N and O atoms are strongly dependent on the nature of its substituents. This is where the difficulties in experimental determination of the geometrical parameters are encountered.

The AMBER5.0 program with the Cornell et al. (1995) force field originally was developed for nucleic acids and proteins. In addition to lacking nitroxide parameters, it also does not have any parameters for triple bonds in their force field. This necessitated the development of another new atom type for a *sp* hybridized carbon and introduction to Amber force field. By fitting the MM3 force field equations for E_{str} and E_{bend} energy terms to the appropriate terms in the force field equations represented in Amber, we were able to obtain force constants k_θ and k_r

together with equilibrium bond lengths and angles. Newly developed parameters are given in Table 2.0.

Table 2.0. Modified parameters for nitroxides

MASS	a.u.	
C1	12.01	
LP	0.0001	
NR	14.01	
OR	16.00	
BOND	K_r kcal/(mole·Å ²)	r_{eq} (Å)
C1-C1	700	1.204
C1-HA	50.0	1.056
C1-CM	600.0	1.440
CA-NR	500.0	1.35
CT-NR	450.1	1.47
LP-OR	325.0	0.50
NR-OR	337.0	1.28
ANGLE	K_θ kcal/(mole·rad ²)	θ_{eq} (degree)
C1-C1-HA	32.8	180.0
C1-C1-CM	65.7	180.0
C1-CM-CM	39.4	120.0
C1-CM-CT	39.4	120.0
CM-CT-CT	29.6	109.5
CT-CM-HA	19.7	120.0
NR-CT-CT	80.0	105.6
CT-NR-CT	46.6	120.0
OR-NR-CT	80.2	119.3
HC-CT-NR	50.0	109.5
NR-CA-NR	50.4	122.0
NR-CT-CM	25.2	100.1
CA-NR-OR	50.8	124.0
CA-NR-CT	25.2	122.0
LP-OR-LP	80.0	120.0
LP-OR-NR	80.0	120.0

DIHEDRAL	¹ # of paths	² $V_n/2$	³ γ	⁴ n^I
X -C1-C1-X	1	0.0	180.0	1.
X -CM-C1-X	2	0.0	180.0	1.
NT-CT-CT-X	9	0.2	0.0	3.
X -NR-CT-X	6	0.3	0.0	2.
X -NR-CA-X	6	6.3	0.0	2.
X -NR-OR-X	4	1.0	0.0	2.
OOP				
NR-CT-CT-OR	-	4.0	180.0	2.
X -X -NR-OR	-	10.1	180.0	2.
X -X -NR-CT	-	1.2	180.0	2.
NONB	⁵ R^* (Å)		⁶ ϵ^* (kcal/mole)	
C1	1.908		0.086	
NR	3.66		0.069	
OR	0.00		0.060	
LP	0.00		0.061	

¹ # of bond paths that the total $V_n/2$ is divided into. This is equal to the product of the # of the bonds to each of the middle two atoms. ² Magnitude of torsion in kcal/mole. ³ Phase offset in degree. ⁴ Periodicity of the torsion. ⁵ Van der Waals radius for given atom $R_{ij}^* = (R_{ii}^* R_{jj}^*)^{1/2}$. ⁶ Van der Waals well depth for given atom. $\epsilon_{ij}^* = (\epsilon_{ii}^* \epsilon_{jj}^*)^{1/2}$

2.4. Results and discussions

In tables 2.1, 2.2, and 2.3 are shown parameters of three nitroxide radical molecules obtained from optimization with quantum mechanics (QM) and molecular mechanics (MM) methods implemented in Q-chem and sander module of Amber programs respectively. Results available from experimental data and QM calculation conducted by Barone et. al., are also included in these Tables. These experimental geometries of two radicals shown on Figures 2.1 and 2.2 are compared with the structures issuing from full optimization at the B3LYP/6-31G* level and MM computation. There are no experimental data available for the 5-membered ring nitroxide molecule thus the results obtained from MM calculations should be compared only with results from QM calculation. Comparing the N-O bond lengths for all three molecules obtained from QM and MM calculation, we can see that they fall well into the experimental range (1.26-1.29Å) reported by Baron et. al. Comparison of some of the parameters for DMNO and TBNO molecules shows very little discrepancy with the ones for experimental, B3LYP^a (by Barone et. al.), and B3LYP^c (Q-Chem calculation). However, in some cases, large differences are observed. This suggests that additional refinement for MM parameters needs to be done, so they better reproduce experimental ones. Also x-ray crystallography experiment on 5-membered ring nitroxide molecule will provide better insight on its physical parameters.

Figure 2.1 gives the insight to the side view of DMNO molecule from both, MM minimization using sander module of AMBER5.0 program (**1a**) and QM optimization using Q-Chem program (**2b**) as it was mentioned before. It can be seen from the figure 2.1 1b, that the geometry of DMNO after QM optimization has more pyramidal character than geometry of the molecule after MM minimization. This fact is supported by the value of out-of-plane angle

Figure 2.1. Side view **1a** and front view **1b** of MNO molecule

Table 2.1. Geometrical parameters for Dimethyl Nitroxide Radical (DMNO)

Bonds (Å)	Exp ^a	B3LYP ^a	MM ^b	B3LYP ^c
N2-O4	1.28	1.282	1.278	1.286
N2-C1	-	1.463	1.469	1.548
Angles (°)				
C1-N2-O4	-	-	119.31	117.55
C3-N2-O4	-	119.4	119.37	117.64
C1-N2-C3	118.9	115.5	121.31	118.97
Torsion (°)				
O4-N2-C1-H5	-	-	0.02	38.48
O4-N2-C1-H6	-	-	0.00	11.10
τ	-	-	0.00	25.5

^a From reference # 7

^b From optimization from Sander

^c From QM calculation by Q-Chem program

Figure 2.2. Side view **2a** and front view **2b** of TBNO molecule

$\tau = 25.5^\circ$. The N2-O4 bond length shown in table 2.1 in both cases falls in an experimental range. For the TBNO molecule illustrated on figure 2.2 and geometrical parameters shown in table 2.2 one can see that again QM gave more pyramidal geometry for the N-O moiety than MM minimization. The out-of-plane values $\tau = 0^\circ$ (from MM) and $\tau = 20.5^\circ$ (from QM) obviously confirm that. But the comparison of τ for DMNO and TBNO indicates that former tends to be more planar than DMNO. This geometry can be driven by the fact that nitroxides are in most cases quasi-planar molecules, so that the orbital containing the odd number of electrons has an almost pure π^* character. The structure of the radical either planar or bent, depending on whether stabilization from partial π bond is sufficient to retain a pure sp^2 hybridization of the nitrogen. The competition between π bonding and the preference of nitrogen for sp^3 hybridization explains the occurrence of both planar and pyramidal structures depending on molecular topologies. The C-N-C angle for DMNO and TBNO molecules is larger than the angle O-N-C. The out-of-plane angle of the NO moiety depends on the value of C-N-C angle.

Table 2.2. Geometrical parameters for Di(*tert*-butyl)aminoxyl Radical (TBNO)

Bonds (Å)	Exp ^a	B3LYP ^a	MM ^b	B3LYP ^c
N2-O4	1.280±0.02	1.287	1.281	1.288
N2-C1	1.512±0.02	1.511	1.589	1.517
N2-C3	-	-	1.588	1.517
C3-C9	-	-	1.612	1.540
C3-C8	-	-	1.610	1.543
C3-C10	-	-	1.611	1.542
C1-C5	1.534±0.02	1.545	1.609	1.538
C1-C6	1.534±0.02	1.540	1.612	1.540
C1-C7	1.534±0.02	1.540	1.611	1.542
Angles (°)				
C1-N2-O4	-	-	118.25	112.75
C3-N2-O4	-	-	118.26	115.86
C1-N2-C3	136±3	127.9	123.49	128.12
C9-C3-C8	-	-	102.62	108.71
C9-C3-C10	-	-	102.72	107.90
C8-C3-C10	-	-	103.57	111.13
C5-C1-C7	107±2	110.9	102.68	110.02
C6-C1-C7	107±2	108.6	102.76	109.60
C5-C1-C6	107±2	107.8	102.54	108.28
Torsion				
O4-N2-C1-C5	137±2	138.3	136.00	173.70
O4-N2-C1-C7	-	-	16.50	-63.20
O4-N2-C1-C6	-	-	-102.70	54.30
O4-N2-C3-C8	-	-	136.00	99.40
O4-N2-C3-C9	-	-	16.50	-18.90
O4-N2-C3-C10	-	-	-102.8	-136.50
τ	-	-	0.00	20.5

^a From reference # 7^b From optimization from Sander^c From QM calculation by Q-Chem program

Figure 2.3. Side view 3a and front view 3b of 5-membered ring nitroxide molecule

As τ increases, the C-N-C angle decreases. This does not offer a rationalization for the planarity of 5-membered ring nitroxide molecule, where the C-N-C angle is 115° . For this molecule C-N-C angle is smaller than O-N-C one (C-N-C = 122° , O-N-C = 115°), which is due to the fact that the angles in 5-membered ring are constrained to deviate too much from the value of 108° characterizing a regular pentagon.

Table 2.3. Geometrical parameters for 5-membered ring nitroxide

Bonds (Å)	MM ^b	B3LYP ^c
C1-N12	1.555	1.484
C3-C4	1.533	1.522
C4-C5	1.349	1.342
C5-C1	1.522	1.508
C4-C10	1.451	1.422
C10-C11	1.206	1.210
N2-O12	1.277	1.275
C1-C6	1.602	1.540
Angles (°)		
C1-N2-O12	121.3	122.5
O12-N2-C3	120.90	122.35
C1-N2-C3	117.81	115.10
N2-C3-C4	95.69	99.68
N2-C1-C5	95.53	99.90
C1-C5-C4	116.13	113.12
C5-C4-C3	114.81	111.99
C5-C4-C10	120.64	126.61
H13-C5-C4	120.56	124.70
C7-C1-C6	108.30	110.31
C8-C3-C9	107.59	110.60
C6-C1-N2	118.15	109.88
Torsion (°)		
C6-C1-N2-O12	65.47	60.8
τ	0.00	0.00

^b From optimization from Sander^c From QM calculation by Q-Chem program

CHAPTER III

3.1. Molecular dynamics (MD)

Molecular Dynamics is a method that can be used for studying the motions and the conformational space of molecular systems or as a tool to study the structure of molecular systems. Given a potential energy function and its associated force field the classical Newtonian equations of motion can be integrated. In MD, the positions and velocities that a particle will have in future are calculated. This is done by first determining the force on each particle $\mathbf{F}_i$ as a function of time, which is given in Equation 3.1, and acceleration $\mathbf{a}$ of each particle, which can be determined by Equation 3.2.

$$\mathbf{F}_i = -\frac{\partial U}{\partial \mathbf{r}_i} \quad 3.1$$

$$\mathbf{a}_i = \frac{\mathbf{F}_i}{m_i} \quad 3.2$$

The Hamiltonian of the system in terms of kinetic and potential energy is given in Equation 3.3, where $U(\mathbf{q})$ has the form of the AMBER energy function (Equation 2.8).

$$H(\mathbf{q}, \mathbf{p}) = \frac{1}{2} \sum_{i=1}^N \frac{\mathbf{p}_i^2}{m_i} + U(\mathbf{q}) \quad 3.3$$

The Verlet algorithm of Swope and Anderson, which is used to integrate the equations of motion, is given in Equation 3.4, and the velocities, calculated for the next time step, are given in Equation 3.5. This technique involves updating the velocities in two steps. In the first step, the velocities are advanced from time t to time $(t + \frac{\delta t}{2})$. Then forces $\mathbf{F}$ are updated for the new atomic positions and the velocities are updated from time $(t + \frac{\delta t}{2})$ to time $(t + \delta t)$. The

resulting set of positions and velocities, $\mathbf{R} = \{\mathbf{q}(t_1), \mathbf{q}(t_2) \dots \mathbf{v}(t_1), \mathbf{v}(t_2)\}$, is referred to as the dynamics trajectory.

$$\mathbf{q}_i(t + \delta t) = \mathbf{q}_i(t) + \delta t \mathbf{v}_i(t) + \frac{\delta t^2}{2m_i} \mathbf{F}_i(t) \quad 3.4$$

$$\mathbf{v}_i(t + \delta t) = \mathbf{v}_i(t) + \frac{\delta t}{2m_i} [\mathbf{F}_i(t + \delta t) + \mathbf{F}_i(t)] \quad 3.5$$

The complexity of the calculation centers on the evaluation of the pairwise non-bonded and Coulombic interactions. Water is an important and integral part of any nucleic acid, so the explicit introduction of the solvent becomes fairly critical. For the simulations *in vacuo* or simulations in which explicit solvent molecules are represented, which is a case of simulations used in current work, a classical electrostatic function, where the denominator equals qR_{ij} , is usually chosen. Simulations without explicit solvent i.e. *in vacuo* have been performed on many DNA structures. There the use of distance dependent dielectric function for electrostatic energy was used to mimic the polarization effect in attractive interactions, with closer interactions weighted more heavily. Also it compensated for the lack of the explicit solvent by implicitly damping longer-range charge interactions more than short range ones. But in order to keep the base pairs from fraying, additional restraints to Watson-Crick base pairs have been applied which in turn created some discrepancies in charges on phosphates. Many simulations with explicit solvent have been performed on DNA molecules also. They were limited to a short time scale (~100ps) and typically displayed anomalous structure, such as base fraying, which demonstrated the necessity for inclusion of more accurate representation of solvent. More recent simulations on longer ~1 ns time scale suggest the importance of properly treating the long-ranged electrostatic interactions. The application of the particle mesh Ewald method allows simulations of nucleic acids with explicit solvent and counterions in a nanosecond time range. In the particle

mesh Ewald (PME) method, a Gaussian charge distribution of opposite sign is superimposed upon the original point charge, producing a screened charge distribution. The electrostatic interaction between the screened charges is short-ranged. The original distribution is recovered by adding a second Gaussian charge distribution identical to the first, but opposite sign. The interaction between these canceling distributions is calculated in reciprocal or k-space.

When using xleap to solvate a DNA molecule, water is placed randomly from a pre-equilibrated water box around the solute. An appropriate number of sodium ions (Na⁺) are added by replacing water molecules near the solute to neutralize DNA. This water has not felt the influence of the solute and there may be gaps between the solvent and solute and solvent and box edges. Therefore the water should be allowed to relax around solute and come to an equilibrium density before any dynamics run. Unless the water density is perfectly matched to the box size and all gaps around solute are filled, constant pressure and variable volume dynamics should be used to allow the box size to change.

During constant temperature molecular dynamics, the velocities (v) in the system are scaled during each time step of MD. This couples the system, with a temperature relaxation time τ_T , to a heat bath at T_0 . Velocities are scaled by a factor λv , which allows maintaining a desired amount of kinetic energy, where

$$\lambda = 1 + \frac{\delta t}{2\tau_T} \left(\frac{T_0}{T} - 1 \right), \quad \text{where } T \text{ is instantaneous temperature} \quad \mathbf{3.6}$$

In a manner similar to the velocity scaling, the pressure is held constant by scaling the atomic coordinates by a factor μ , so that a desired pressure is maintained:

$$\mu = \left[1 - \frac{\delta t}{\tau_P} (P_0 - P) \right]^{\frac{1}{3}}, \quad \text{where } P \text{ is instantaneous pressure} \quad \mathbf{3.7}$$

3.2. Minimization technique

A widely used technique in computational chemistry is minimization. For the given potential energy function $U(\mathbf{q})$ (Equation 2.2) dependent on coordinate $\mathbf{q}$, one wants to find the minimum potential energy of the system, such that

$$\frac{\partial U}{\partial q_i} = 0 \quad 3.8$$

The steepest-descent method is a first order minimization and is used throughout MM steps in current work. That is, it makes use of the first derivative of the potential energy with respect to the Cartesian coordinates. The descent is accomplished by adding an increment to the coordinates in the direction of the force or the negative gradient of the potential energy and is shown in Equation 3.9.

$$\mathbf{x}_k = \mathbf{x}_{k-1} + \lambda \mathbf{s}_k, \quad \text{where } \mathbf{s}_k = \frac{\mathbf{F}}{F}, \quad \mathbf{F} = -\nabla U \quad 3.9$$

$\mathbf{x}$ are the Cartesian vectors, λ is the variable increment or step size,
 $\mathbf{s}$ is the search vector.

The steepest-descent is notable for its rapid alleviation of large forces on atoms. That makes the technique especially useful for eliminating the large non-bonded interactions often found in initial structures and requires minimal amount of CPU time.

CHAPTER IV

Molecular Dynamics Simulations of ds DNAs

Preliminary structures of canonical B-form 15-mer single-stranded (ss) and duplex (ds) DNAs were built using biopolymer module of Sybyl program. Deoxythymidine (T) base was modified to contain 5-membered ring nitroxide using the LEaP graphical model building program of AMBER 5.0. The modified deoxythymidine (T*) was added to the nucleic acid library (all_nucleic94.lib) for future use. The modified part of all_nucleic94.lib file is given in Appendix A. All three 15-mer ss, ds, and triplex (tx) DNA models were modified with T* at T₇, T₂₃ and T₃₈ positions of thymidine respectively. Hydrogens were added with the EDIT module of AMBER 5.0 and the initial hydrogen positions were minimized *in vacuo* while holding all non-hydrogens fixed. All the calculations were carried out using the all-atom parameters contained in AMBER5.0 program with a Cornell *et al.* (1994) force field. Models were net-neutralized with explicit sodium counterions (Na⁺) placed at the phosphates of all models with the EDIT module, and a periodic box of water molecules described by TIP3P potential surrounded the nucleic acid molecules. The periodic box was extended to the distance of 10 Å in each direction from solute atoms for unmodified ds DNA, and to the distance of 8 Å for modified. This leads to a periodic box size of ~42 Å by ~43 Å by ~73 Å and ~39 Å by ~40 Å by ~66 Å respectively. Simulations were performed in the isothermal isobaric ensemble (300 K, 1 atm) using periodic boundary conditions. MD calculations were carried out using Sander module (2 fs time step) with SHAKE on the hydrogen atoms with a tolerance of 0.0005 Å. A 9 Å cutoff was applied to Lennard-Jones interactions. Simulations were performed using Berendsen temperature coupling algorithm (with a time constant of 0.2 ps) and constant pressure

with isotropic molecule based scaling (with a time constant of 0.2 ps). In this way two DNA strands were treated as if they were two different molecules and the shifting of these two strands was avoided during equilibration. The non-bonded pair list was updated every 10 steps.

Equilibration was started by 1000 steps of minimization with the positional restraints on the nucleic acids and 500 kcal/mole·Å of force constant. Then, dynamics run for 25 ps with a cutoff of 9 Å on all interactions was followed. All subsequent simulations were performed using the particle mesh-Ewald method (PME) for inclusion of long-range electrostatic interactions without truncation. The particle mesh-Ewald method is a fast method for performing Ewald summation of Coulombic interactions. It means that the system is treated as if it was truly periodic and includes all electrostatic interactions within the unit cell and with all image cells.

Equilibration was continued by 25 ps of PME dynamics with position restrains on DNA. Next, 1000 steps of minimization were carried out with 25 kcal/mole·Å of positional restrains on DNA followed by 3 ps dynamics run, which allowed water and Na⁺ ions to relax around the solute. Subsequently, 600 steps of equilibration with gradual removal of positional restraints by 5 kcal/mole·Å on DNA molecule for each run were performed. On the final dynamics step the system was heated from 100 K to 300 K over 20 ps. Production runs of 1 ns duration, at constant temperature (T=300 K), were performed after the final dynamics step. After each run the PME box information was updated to match the final box coordinates from previous runs. The results were analyzed using the Carnal, Mdanal, and rdparm modules of AMBER 5.0.

Figure 4.1. Equilibration of ds DNA without spin-probe

Figure 4.1 shows the equilibration process of ds DNA without spin-probe in a water box for 20 ps molecular dynamics equilibration run with consequent 1.0 ns production run. It can be seen that the 20 ps equilibration, during which the system was heated from $T=100\text{K}$ to 300K , reached rather rapidly, within 5 ps time range, which suggests that 20 ps is a reasonable time for the solvent to relax around solute. During the production run the temperature was kept constant ($T=300\text{K}$). These plots show that although density, pressure and volume are fluctuating

Figure 4.2. Equilibration of ds DNA with spin-probe

in a somewhat wider range than other parameters (0.01 g/cm^3 , 1000 atm and a $2 \cdot 10^3 \text{ \AA}^3$ respectively), they remain near their constant values throughout all production run. These fluctuations do not affect the overall stability of a system. Temperature, KE (kinetic), PE (potential) and E (total) energies of the system are well stabilized and stay constant throughout the run. The similar argument can be conducted for the figure (Figure 4.2) with one exception, that it shows the equilibration for 20 ps and production run for additional 1.0 ns run for ds DNA

bearing 5-membered spin-probe nitroxide molecule at the middle (8 position) of the T₁₅ strand. These plots suggest that the presence of nitroxide does not affect the stability of the DNA, but on the contrary brings more order to the vicinity of the spin-probe. This can be seen from the intensity of RMSd picks for atoms in a range 223-254 on figure 4.3.

Figure 4.3. RMS deviation of ds DNA from original structure with and without spin-probe

Root-mean-square deviations (RMSd) of snapshots from initial structure are calculated over all DNA bases. Figure 4.3 shows atomic positional fluctuations of both modified and

unmodified ds DNAs. The large peaks at the left, middle and right correspond to the beginning and ends of each strands of the double helix. Larger peaks indicate more motion. The bottom plot (2) of Figure 4.3 has a large peak that includes atoms 702 - 754, which indicates positional fluctuations of the thymidine base bearing a spin-probe. From the intensity of the peaks around the spin-probe it can be noticed from the picture that there is some stability in that region.

CHAPTER V

Molecular dynamics simulations of tx DNAs

Triple stranded (tx) DNA was also built in Sybyl program. Using earlier constructed ss and ds DNAs, the docking module of Sybyl program was used to anneal the ss DNA with ds DNA to form a triplex. Two types of tx DNAs are constructed and used for MD simulations. One, with the third ($T_7T^*T_7$) strand running parallel to A_{15} strand ($T \uparrow \uparrow A$), and another, running antiparallel to it ($T \uparrow \downarrow A$). The first model is of the most interest, because it is widely found in nature. A short minimization of the tx DNA was performed in Sybyl to eliminate particularly bad steric interactions between ds DNA and the third strand. Minimization was done by Powell method for 5000 steps, before transferring the structures into AMBER5.0 program for the further simulations.

All the simulations for both modified and unmodified tx DNAs were proceeded in the same manner as for ds DNA, with the only difference in the periodic box size, which lead to a $\sim 42 \text{ \AA}$ by $\sim 42 \text{ \AA}$ by $\sim 70 \text{ \AA}$ for both triplexes.

A production run of 1 ns duration, at constant temperature ($T=300 \text{ K}$), was performed after the final dynamics step for unmodified tx DNA. After each run the PME box information was updated to match the final box coordinates from previous runs. The results were analyzed using the Carnal, Mdanal, and rdparm modules of AMBER 5.0.

Figures 5.1 and 5.2 show the equilibration process of tx DNA without and with the spin-probe in a water box for 20 ps molecular dynamics equilibration run with consequent 1.0 ns and 0.5 ns production runs correspondingly. Again, as in the case of ds DNAs, it can be seen that the

20 ps equilibration, during which the system was heated from $T=100\text{K}$ to 300K , reached rather rapidly.

Figure 5.1. Equilibration of tx DNA without spin-probe

During the production run the temperature was kept constant ($T=300\text{K}$). These plots show that although density, pressure and volume are fluctuating in a somewhat wider range than other parameters, they remain near their constant values throughout all production run. These fluctuations do not affect the overall stability of a system. Temperature, KE (kinetic), PE

(potential) and E (total) energies of the system are well stabilized and stay constant throughout the run. The units of all parameters are the same as in Figure 4.1.

Figure 5.2. Equilibration of tx DNA with spin-probe

Figure 5.3. RMS deviation of *antiparallel* tx DNA.

In Figure 5.3 are given atomic positional fluctuations for antiparallel tx DNAs. Base numbering is somewhat different for these triplexes. On plot 1 adenosine strand includes bases 1-15 (left portion of a plot), while on plot 2, adenosine strand includes bases 16-30 (middle portion of a plot). Plot 2 of this figure has a large peak, which includes atoms 223 - 254, and indicates fluctuations of the spin-labeled thymidine base. Comparing plots for both triplexes we can see that over all DNA bases, fluctuations of spin-labeled triplex from initial structure are more stable than the fluctuations of unlabeled triplex, but there is some instability at the ends of

strands. The RMSd plot for parallel tx DNA from its starting structure, shown on Figure 5.4 indicates that over 0.6 ns run this molecule seems to be less stable than the antiparallel tx DNA. Comparing Figure 5.4 with plot 2 of Figure 5.3 that even though fluctuations of parallel triplex are larger than those for antiparallel triplex, the fluctuations of spin-probe bearing thymidine is slightly less in parallel ($\sim 1\text{-}3\text{ \AA}$) than in antiparallel ($\sim 1\text{-}4\text{ \AA}$). Also looking at the positional fluctuations over time for different atoms of the nitroxide (Figure 5.5) we can see that fluctuations over 0.4 ns time period involving atoms C36-C40, N41 (green line) do not differ for both triplexes (1' and 3') compared to the duplex (2'). But for another set of atoms C5, C36-C40, N41, O42 (black lines) the fluctuations of parallel triplex (2'') are larger than the ones for antiparallel (3'').

Figure 5.4. RMS deviation of *parallel* tx DNA.

Figure 5.5. Fluctuations of two sets of atoms of the nitroxide molecule.

These fluctuations do not give any information about rotational motion of the nitroxide itself. In order to extract rotational information from the dynamics performed on all DNAs torsional parameters in the backbone angles, helicoidal parameters, sugar puckers and newly defined torsions for nitroxide molecule are calculated. In the Figure 5.6 the following torsional angles for nitroxide molecule are shown. Also in table 5.1 the average torsional angles for DNA helices are listed.

Figure 5.6. Torsional parameters of the nitroxide.

Table 5.1. Torsional angles for DNA helices and nitroxide.

Torsion°	Atoms	Average Torsion°
α	C5' - O5' - P - O3'	-50
β	C4' - C5' - O5' - P	172
γ	C3' - C4' - C5' - O5'	41
δ	O3' - C3' - C4' - C5'	79
ϵ	C4' - C3' - O3' - P	-146
χ	C2 - N1 - C1' - O4'	-78
ζ	C3' - O3' - P - O5'	-154
a	O4 - C4 - C5 - C35	-
b	H6 - C6 - C5 - C35	-
c1	C4 - C5 - C35 - C36	-
c2	C6 - C5 - C35 - C36	-
d	C5 - C35 - C36 - C37	-
e1	C35 - C36 - C37 - C38	-
e2	C35 - C36 - C37 - C39	-
f1	C36 - C37 - C38 - C45	-
f2	C36 - C37 - C38 - C46	-
g	C36 - C37 - C39 - H59	-
h	C6 - C5 - C37 - C39	-

The time course of torsional angles for only the modified thymidine of ds DNA and the histograms are shown in Figure 5.7. Analysing the data in Figure 5.7a it can be seen that torsional parameters ($\beta, \gamma, \delta, \epsilon, \chi$, and ζ) are held at their average values as given in the table 5.1, which suggests that throughout the simulation DTM base retains its geometry. Similar diagrams are obtained for all other bases (results not shown). Torsions of the most interest for us are the ones on the nitroxide. We can see that torsional angles defined in the rings (a, b, f1, f2, g) are kept constant as it was expected. On the other hand, torsions around the tether (c1, c2, d, e1, e2) connecting the thymidine base to the nitroxide (triple bond and adjacent single bonds) are covering the whole range of possible angles. This result suggests that the torsional barrier defined for the single bonds was too small, and that the nitroxide could easily overcome these barriers, and freely rotate around them without reaching an average equilibrium value. The similar conclusion can be drawn with respect to the triplexes (Figure 5.8 and 5.9). ESR studies by Bobst et al. on duplexes with different nitroxides showed that the correlation times τ for the nitroxide motion are longer than 3 ns. In fact for spin-labeled RNA a line shape simulation gives a value of 20 ns. This fact tells us, that much longer dynamics simulations will give the results comparable with the experimental ones. In fact, from Figure 5.7 we can notice that during 1 ns run of the ds DNA some range of possible angles is detectable. Reviewing torsional barrier defined in the force fields we noticed that slight changes to the torsional force constant gave us much better result (i.e. better correlation with experiment) for values c1, c2, d, f1, and f2 (Figure 5.10). This short run was done for the ds DNA in vacuo to confirm the validation of the new force field.

Figure 5.7. Plots of torsional parameters for spin-probe labeled ds DNA.

1.

502.

Figure 5.8. Plots of torsional parameters for *antiparallel* spin-probe labeled tx DNA.

Figure 5.8. Plots of torsional parameters for *parallel* spin-probe labeled tx DNA.

Figure 5.9. Plots of torsional parameters for spin-probe labeled ds DNA in vacuo.

CHAPTER VI

Discussion and conclusions

We have conducted Molecular Dynamics simulations on double- and triple-stranded DNAs with and without modified thymidine base at eighth base position with the spin-probe labeled nitroxide molecule. We have shown that the presence of the nitroxide does not disrupt DNA strands but rather brings some stability to the structure. Torsional analysis of modified base (DTM) showed that the rotation of the nitroxide can be in the reasonable angle range for much longer than 1 ns time. Also the further refinement of force field parameters can drastically change the outcome of the run.

Charges on nitroxide molecule were not included in the dynamics calculations. They were derived later using gaussian 98 program. Partial atomic charges were evaluated by fitting electrostatic points obtained from gaussian single point calculation with the RESP algorithm at a molecular surface using an atom-centered point charge model. The inclusion of charges may affect solute-solvent interactions in a vicinity of the modified base and may play role in stability of that region in DNAs. For that reason further dynamics simulations (over 1 ns period of time) will be conducted with inclusion of partial charges.

Animation of the trajectories obtained over 1 ns runs for ds DNAs and 0.5 ns runs for antiparallel tx DNAs revealed in all cases that the DNA structures were maintained throughout. The average RMSd measured from the starting structure is 2.79 Å for ds DNA, 1.445 Å for antiparallel and 2.149 Å for parallel tx DNAs. The average RMSd measured for spin-probe labeled thymidine is 0.934 Å, 0.464 Å, and 1.149 Å, respectively.

Reference

1. Moffat, A. S. (1991) Triplex DNA finally comes of age. *Science*, **252**, 1374-1375.
2. G. T. Pauly, I. E. Thomas, and A. M. Bobst. (1987) Base Dynamics of Nitroxide-Labeled Thymidine Analogues Incorporated into (dA-dT) by DNA Polymerase I from *Escherichia coli*. *Biochemistry*, **26**, 7304-7310.
3. V. Mohan, Y. K. Cheng, G. E. Marlow, and B. M. Pettitt. (1993) Molecular Recognition of Watson-Crick Base-Pair Reversals in Triple-Helix Formation: Use of Nonnatural Oligonucleotide Bases. *Biopolymers*, **33**, 1317-1325.
4. Y. Duan, P. Wilkosz, M. Crowley, and J. M. Rosenberg. (1997) Molecular Dynamics Simulation Study of DNA Dodecamer d(CGCGAATTCGCG) in Solution: Conformation and Hydration. *J. Mol. Biol.*, **272**, 553-572.
5. J. Robles and L. W. McLaughlin. (1997) DNA Triplex Stabilization Using a Tethered Minor Groove Binding Hoechst 33258 Analogue. *J. Am. Chem. Soc.*, **119**, 6014-6021.
6. J. A. Suburana and T. Faria. (1997) Influence of Sequence on the Conformation of the B-DNA Helix. *Biophysical Journal*, **73**, 333-338.
7. V. Barone, A. Bencini, M. Cossi, A. Di Matteo, M. Mattesini, and F. Totti. (1998) Assessment of Combined QM/MM Approach for the Study of Large Nitroxide Systems in Vacuo and in Condensed Phases. *J. Am. Chem. Soc.*, **120**, 7069-7078.
8. K. Sakakibara, H. Kawamura, T. Nagata, J. Ren, M. Hirota, K. Shimazaki, and T. Chuman. (1994) MM2 Molecular Mechanics Studies of Periplanone B and Nitroxides. Empirical Methods to evaluate the Stretching (k_s) and Bending (k_b) Force Constants of the MM2 Force Field. *Bull. Chem. Soc. Jap.* **67**, 2768-2778.
9. A. Spaltenstein, B. H. Robinson, and P. B. Hopkins. (1989) DNA structural data from (1989) Sequence- and structure-dependent DNA base dynamics: synthesis, structure, and dynamics of site and sequence specifically spin-labeled DNA. *Biochemistry*, **28**, 9484-9495.
10. Y.-K. Cheng and B. M. Pettit. (1992) Stabilities of double- and Triple-Stranded Helical Nucleic Acids. *Prog. Biophys. Molec. Biol.* **58**, 225-257.
11. A. Bobst, S.-C. Kao, R. C. Toppin, J. C. Ireland, and I. E. Thomas. (1984) Dipsticking the Major Groove of DNA with Enzymatically Incorporated Spin-labeled Deoxyuridines by Electron Spin Resonance Spectroscopy. *J. Mol. Biol.*, **173**, 63-74.
12. B. H. Robinson, L. S. Lerman, A. H. Beth, H. L. Fisher, L. R. Dalton, and C. Auer. (1980) Analysis of Double-helix Motions with Spin-labeled Probes: Binding Geometry and the Limit of Torsional Elasticity. *J. Mol. Biol.*, **139**, 19-44.
13. U. Essmann, L. Perera, and M. L. Berkowitz. (1995) A smooth particle mesh Ewald method. *J. Chem. Phys.*, **103**, 8577-8593.

14. Gee, JE and Miller, DM. (1992) Structure and applications of intermolecular DNA triplexes. *American Journal of the Medical Sciences*, **304**(6), 366-372.
15. Strobel, SA; Moser, HE; and Dervan, PB. (1988) Double-strand cleavage of genomic DNA at a single site by triple-helix formation. *Journal of the American Chemical Society*, **110**, 7927-7929.
16. Hartman, DA; Kuo, SR; Broker, TR; Chow, LT; and Wells, RD. (1992) Intermolecular triplex formation distorts the DNA duplex in the regulatory region of human papillomavirus type-II. *Journal of Biological Chemistry*, **267**(8), 5488-5494.
17. A. Spaltenstein, B. H. Robinson, and P. B. Hopkins. (1989) DNA structural data from a dynamics probe; the dynamic signatures of single-stranded, hairpin-looped, and duplex forms of DNA are distinguishable. *Journal of the American Chemical Society*, **111**, 2303-2305.
18. Becke, AD; (1993) A new mixing of Hartree-Fock and local density-functional theories. *J. Chem. Phys.*, **98**, 1372-1376.
19. Stephens, PJ, Devlin, FJ, Chabalowski, CF, Frisch, MJ. (1994) Ab initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.*, **Vol. 98**, No 45, 11623-11627.

Appendix A

Modified all_nucleic94.lib file

!!index array str

"5SP"

"DA"

"DA3"

"DA5"

"DAN"

"DC"

"DC3"

"DC5"

"DCN"

"DG"

"DG3"

"DG5"

"DGN"

"DT"

"DT3"

"DT5"

"DTH"

"DTM"

"DTN"

"RC"

"RC3"

"RC5"

"RCN"

"RG"

"RG3"

"RG5"

"RGN"

"RU"

"RU3"

"RU5"

"RUN"

!entry.DTM.unit.atoms table str name str type int typex int resx int flags int seq int elmnt dbl chg

"P" "P" 0 1 196608 1 15 1.165900

"O1P" "O2" 0 1 196608 2 8 -0.776100

"O2P" "O2" 0 1 196608 3 8 -0.776100

"O5'" "OS" 0 1 196608 4 8 -0.495400

"C5'" "CT" 0 1 196608 5 6 -0.006900

"H5'1" "H1" 0 1 196608 6 1 0.075400

"H5'2" "H1" 0 1 196608 7 1 0.075400

"C4'" "CT" 0 1 2293760 8 6 0.162900

"H4'" "H1" 0 1 196608 9 1 0.117600

"O4'" "OS" 0 1 2293760 10 8 -0.369100

"C1'" "CT" 0 1 2293760 11 6 0.068000

"H1'" "H2" 0 1 196608 12 1 0.180400

"N1" "N*" 0 1 2293760 13 7 -0.023900

"C6" "CM" 0 1 2293760 14 6 -0.220900

"H6" "H4" 0 1 196608 15 1 0.260700

"C5" "CM" 0 1 2293760 16 6 0.002500

"C4" "C" 0 1 2293760 21 6 0.519400

"O4" "O" 0 1 196608 22 8 -0.556300

```

"N3" "NA" 0 1 2293760 23 7 -0.434000
"H3" "H" 0 1 196608 24 1 0.342000
"C2" "C" 0 1 2293760 25 6 0.567700
"O2" "O" 0 1 196608 26 8 -0.588100
"C3'" "CT" 0 1 2293760 27 6 0.071300
"H3'" "H1" 0 1 196608 28 1 0.098500
"C2'" "CT" 0 1 2293760 29 6 -0.085400
"H2'1" "HC" 0 1 196608 30 1 0.071800
"H2'2" "HC" 0 1 196608 31 1 0.071800
"O3'" "OS" 0 1 196608 32 8 -0.523200
"C35" "C1" 0 1 196610 35 6 0.0
"C36" "C1" 0 1 196610 36 6 0.0
"C37" "CM" 0 1 2293762 37 6 0.0
"C38" "CT" 0 1 2293762 38 6 0.0
"C39" "CM" 0 1 2293762 39 6 0.0
"C40" "CT" 0 1 2293762 40 6 0.0
"N41" "NR" 0 1 2293762 41 7 0.0
"O42" "OR" 0 1 196610 42 8 0.0
"C43" "CT" 0 1 196610 43 6 0.0
"C44" "CT" 0 1 196610 44 6 0.0
"C45" "CT" 0 1 196610 45 6 0.0
"C46" "CT" 0 1 196610 46 6 0.0
"H47" "HC" 0 1 196610 47 1 0.0
"H48" "HC" 0 1 196610 48 1 0.0
"H49" "HC" 0 1 196610 49 1 0.0
"H50" "HC" 0 1 196610 50 1 0.0
"H51" "HC" 0 1 196610 51 1 0.0
"H52" "HC" 0 1 196610 52 1 0.0
"H53" "HC" 0 1 196610 53 1 0.0
"H54" "HC" 0 1 196610 54 1 0.0
"H55" "HC" 0 1 196610 55 1 0.0
"H56" "HC" 0 1 196610 56 1 0.0
"H57" "HC" 0 1 196610 57 1 0.0
"H58" "HC" 0 1 196610 58 1 0.0
"H59" "HA" 0 1 196610 59 1 0.0
!entry.DTM.unit.atomsptinfo table str pname str ptype int ptypex int pelmnt dbl pchg
"P" "P" 0 -1 0.0
"O1P" "O2" 0 -1 0.0
"O2P" "O2" 0 -1 0.0
"O5'" "OS" 0 -1 0.0
"C5'" "CT" 0 -1 0.0
"H5'1" "H1" 0 -1 0.0
"H5'2" "H1" 0 -1 0.0
"C4'" "CT" 0 -1 0.0
"H4'" "H1" 0 -1 0.0
"O4'" "OS" 0 -1 0.0
"C1'" "CT" 0 -1 0.0
"H1'" "H2" 0 -1 0.0
"N1" "N*" 0 -1 0.0
"C6" "CM" 0 -1 0.0
"H6" "H4" 0 -1 0.0
"C5" "CM" 0 -1 0.0
"C4" "C" 0 -1 0.0
"O4" "O" 0 -1 0.0
"N3" "NA" 0 -1 0.0
"H3" "H" 0 -1 0.0

```



```

"C2" "C" 0 -1 0.0
"O2" "O" 0 -1 0.0
"C3'" "CT" 0 -1 0.0
"H3'" "H1" 0 -1 0.0
"C2'" "CT" 0 -1 0.0
"H2'1" "HC" 0 -1 0.0
"H2'2" "HC" 0 -1 0.0
"O3'" "OS" 0 -1 0.0
"C35" "C1" 0 -1 0.0
"C36" "C1" 0 -1 0.0
"C37" "CM" 0 -1 0.0
"C38" "CT" 0 -1 0.0
"C39" "CM" 0 -1 0.0
"C40" "CT" 0 -1 0.0
"N41" "NR" 0 -1 0.0
"O42" "OR" 0 -1 0.0
"C43" "CT" 0 -1 0.0
"C44" "CT" 0 -1 0.0
"C45" "CT" 0 -1 0.0
"C46" "CT" 0 -1 0.0
"H47" "HC" 0 -1 0.0
"H48" "HC" 0 -1 0.0
"H49" "HC" 0 -1 0.0
"H50" "HC" 0 -1 0.0
"H51" "HC" 0 -1 0.0
"H52" "HC" 0 -1 0.0
"H53" "HC" 0 -1 0.0
"H54" "HC" 0 -1 0.0
"H55" "HC" 0 -1 0.0
"H56" "HC" 0 -1 0.0
"H57" "HC" 0 -1 0.0
"H58" "HC" 0 -1 0.0
"H59" "HA" 0 -1 0.0
!entry.DTM.unit.boundbox array dbl
-1.000000
0.0
0.0
0.0
0.0
!entry.DTM.unit.childsequence single int
2
!entry.DTM.unit.connect array int
1
28
!entry.DTM.unit.connectivity table int atom1x int atom2x int flags
40 50 1
40 51 1
40 52 1
39 47 1
39 48 1
39 49 1
38 44 1
38 45 1
38 46 1
37 41 1
37 42 1

```

```

37 43 1
35 36 1
34 37 1
34 38 1
34 35 1
33 34 1
33 53 1
32 35 1
32 39 1
32 40 1
31 32 1
31 33 1
30 31 1
29 30 1
25 26 1
25 27 1
23 24 1
23 25 1
23 28 1
21 22 1
19 20 1
19 21 1
17 18 1
17 19 1
16 17 1
16 29 1
14 15 1
14 16 1
13 14 1
13 21 1
11 12 1
11 13 1
11 25 1
10 11 1
8 9 1
8 10 1
8 23 1
5 6 1
5 7 1
5 8 1
4 5 1
1 2 1
1 3 1
1 4 1
!entry.DTM.unit.hierarchy table str abovetype int abovex str belowtype int belowx
"U" 0 "R" 1
"R" 1 "A" 53
"R" 1 "A" 52
"R" 1 "A" 51
"R" 1 "A" 50
"R" 1 "A" 49
"R" 1 "A" 48
"R" 1 "A" 47
"R" 1 "A" 46
"R" 1 "A" 45
"R" 1 "A" 44

```

```

"R" 1 "A" 43
"R" 1 "A" 42
"R" 1 "A" 41
"R" 1 "A" 40
"R" 1 "A" 39
"R" 1 "A" 38
"R" 1 "A" 37
"R" 1 "A" 36
"R" 1 "A" 35
"R" 1 "A" 34
"R" 1 "A" 33
"R" 1 "A" 32
"R" 1 "A" 31
"R" 1 "A" 30
"R" 1 "A" 29
"R" 1 "A" 28
"R" 1 "A" 27
"R" 1 "A" 26
"R" 1 "A" 25
"R" 1 "A" 24
"R" 1 "A" 23
"R" 1 "A" 22
"R" 1 "A" 21
"R" 1 "A" 20
"R" 1 "A" 19
"R" 1 "A" 18
"R" 1 "A" 17
"R" 1 "A" 16
"R" 1 "A" 15
"R" 1 "A" 14
"R" 1 "A" 13
"R" 1 "A" 12
"R" 1 "A" 11
"R" 1 "A" 10
"R" 1 "A" 9
"R" 1 "A" 8
"R" 1 "A" 7
"R" 1 "A" 6
"R" 1 "A" 5
"R" 1 "A" 4
"R" 1 "A" 3
"R" 1 "A" 2
"R" 1 "A" 1
!entry.DTM.unit.name single str
"DTM"
!entry.DTM.unit.positions table dbl x dbl y dbl z
2.402000 1.446000 -0.496000
2.133000 2.724000 -1.190000
3.402000 0.507000 -1.043000
2.895000 1.801000 0.985000
2.225000 2.769000 1.769000
1.188000 2.469000 1.914000
2.248000 3.731000 1.253000
2.916000 2.899000 3.129000
2.363000 3.610000 3.745000
2.912000 1.624000 3.756000

```

```

4.238000 1.294000 4.135000
4.401000 1.622000 5.165000
4.452000 -0.175000 4.029000
4.230000 -0.840000 2.839000
3.899000 -0.272000 1.978000
4.431000 -2.180000 2.730000
4.883000 -2.961000 3.873000
5.094000 -4.172000 3.892000
5.091000 -2.225000 5.026000
5.413000 -2.723000 5.835000
4.907000 -0.860000 5.161000
5.150000 -0.318000 6.238000
4.371000 3.369000 3.007000
4.574000 3.781000 2.017000
5.178000 2.099000 3.238000
5.360000 1.600000 2.286000
6.117000 2.313000 3.748000
4.675000 4.309000 4.022000
4.238000 -2.848000 1.467000
4.118000 -3.432000 0.421000
4.027000 -4.162000 -0.833000
2.904000 -4.034000 -1.865000
5.004000 -5.035000 -1.160000
4.851000 -5.739000 -2.500000
3.527000 -5.027000 -2.892000
2.968000 -5.246000 -4.022000
6.174000 -5.413000 -3.340000
4.816000 -7.314000 -2.213000
2.784000 -2.494000 -2.284000
1.530000 -4.450000 -1.153000
6.999000 -5.922000 -2.831000
6.550000 -4.370000 -3.357000
6.227000 -5.847000 -4.356000
4.728000 -7.966000 -3.102000
4.166000 -7.696000 -1.401000
5.806000 -7.582000 -1.830000
2.033000 -2.244000 -3.057000
3.718000 -1.927000 -2.468000
2.385000 -1.968000 -1.411000
1.321000 -3.681000 -0.402000
1.510000 -5.365000 -0.526000
0.620000 -4.406000 -1.781000
5.852000 -5.198000 -0.489000
!entry.DTM.unit.residueconnect table int c1x int c2x int c3x int c4x int c5x int c6x
1 28 0 0 0 0
!entry.DTM.unit.residues table str name int seq int childseq int startatomx str restype int imagingx
"DTM" 1 60 1 "n" 0
!entry.DTM.unit.residuesPdbSequenceNumber array int
1
!entry.DTM.unit.solventcap array dbl
-1.000000
0.0
0.0
0.0
0.0
!entry.DTM.unit.velocities table dbl x dbl y dbl z

```

[illegible]

Appendix B

Control files for runs

mindd.in

```
min (no shake, periodic). Minimization of ds DNA to clean sterically bad interactions
&cntrl imin = 1, maxcyc = 5000, nsnb = 25,
  scee=2.0, idiel=0, cut=12.0, ichdna=1,
  ntp=5, ntr=0,
  ntb=0,
  ntc = 1, ntf = 1,
&end
END
```

fixit.in

5DNB, initial minimization on ds DNAs, 9.0 cut. (Fixing hydrogens)

```
&cntrl
timlim = 999999.,
ntx  = 1,   irect = 0,   ntrx  = 1,   nt xo  = 1,
ntpr  = 10,  ntwx  = 0,   ntwv  = 0,   ntwe  = 0,
ntwxm = 999999, ntwvm = 999999, ntwem = 999999, ioutfm = 0,
ntwprt = 0,

ntf  = 1,   ntb  = 0,   idiel = 1,   dielc = 1.0,
cut  = 9.0,  ntnb = 1,   nsnb  = 10,  ntid  = 0,
scnb  = 2.0,  scee = 1.2,  cut2nd = 0.0,  ichdna = 0,

ibelly = 1,   ntr  = 0,

imin  = 1,
maxcyc = 500,
ncyc  = 400,
ntmin = 1,   dx0  = 0.1,  dxm   = 0.5,  drms  = 0.05,

ntc   = 1,   tol  = 0.0005,
```

```
&end
Fix damn H1' placement
FIND
* H * *
* HO * *
* HS * *
* HA * *
* HC * *
SEARCH
RES 1 99
END
Fix hydrogens continued
FIND
* H1 * *
* H2 * *
```

```

* H3 * *
* HP * *
* H4 * *
* H5 * *
SEARCH
RES 1 99
END
END

```

ta.in

Initial minimization with position restraints on ds DNAs, 9.0 cut. (Allowing only the water and Na⁺ to move)

```

&cntrl
timlim = 999999.,
ntx = 1,  irest = 0,  ntrx = 1,  nt xo = 1,
ntpr = 10,  ntwx = 0,  ntwv = 0,  ntwe = 0,
ntwxm = 999999, ntwvm = 999999, ntwem = 999999, ioutfm = 0,
ntwpvt = 0,

ntf = 1,  ntb = 2,  idiel = 1,  dielc = 1.0,
cut = 9.0,  ntnb = 1,  nsnb = 10,  nt id = 0,
scnb = 2.0,  scee = 1.2,  cut2nd = 0.0,  ichdna = 0,

isftrp = 0,  rwell = 0.0,

ipol = 0,

ibelly = 0,  ntr = 1,

imin = 1,
maxcyc = 1000,
ncyc = 5000,
ntmin = 1,  dx0 = 0.1,  dxm = 0.5,  drms = 0.0001,

nrun = 0,
nstlim = 10000,
ndfmin = 0,  ntc m = 0,  nsc m = 0,
init = 3,  t = 0.0,  dt = 0.002,

temp0 = 300.0,  tempi = 300.0,
ig = 71277,  heat = 0.0,
ntt = 0,  isolv p = 0,  dtemp = 0.0,
tautp = 0.2,  tauts = 0.2,
vlimit = 20.0,

tauv0 = 0.0,  tauv = 0.1,  vzero = 0.0,

ntp = 1,  pres0 = 1.0,  comp = 44.6,
taup = 0.2,  npscal = 1,

ntc = 1,  tol = 0.0005,

imgslt = 0,  iftres = 1,
jfastw = 0,

```

ivcap = 0, matcap = 0, fcap = 1.5,

&end
Hold the DNA fixed
500.0
RES 1 30
END
END

ta_md.in

5DNB, initial dynamics on ds DNAs, model1, 9.0 cut. (Allowing only water and Na+ to move freely by heating the system from 100 K to 300 K)
&cntrl

timlim = 999999., nmropt = 1,
ntx = 1, irst = 0, ntrx = 1, nt xo = 1,
ntpr = 100, ntwx = 500, ntwv = 0, ntwe = 0,
ntwxm = 999999, ntwvm = 999999, ntwem = 999999, ioutfm = 0,
ntwppt = 0,

ntf = 2, ntb = 2, idiel = 1, dielc = 1.0,
cut = 9.0, ntnb = 1, nsnb = 10, nt id = 0,
scnb = 2.0, scee = 1.2, cut2nd = 0.0, ichdna = 0,

isfrp = 0, rwell = 0.0,

ipol = 0,

ibelly = 0, ntr = 1,

imin = 0,
nrun = 1,
nstlim = 12500,
ndfmin = 0, ntc m = 0, nsc m = 0,
init = 3, t = 0.0, dt = 0.002,

temp0 = 300.0, tempi = 100.0,
ig = 71277, heat = 0.0,
ntt = 1, isolv p = 0, dtemp = 0.0,
tautp = 0.2, tauts = 0.2,
vlimit = 15.0,

tauv0 = 0.0, tauv = 0.1, vzero = 0.0,

ntp = 1, pres0 = 1.0, comp = 44.6,
taup = 0.2, npscal = 1,

ntc = 2, tol = 0.00001,

imgslt = 0, iftres = 1,
jfastw = 0,


```

ivcap = 0,    matcap = 0,    fcap = 1.5,

&end
&wt
  type='TEMP0', istep1=0,  istep2=500,
    value1=100.0, value2=300.0,
&end
&wt
  type='TEMP0', istep1=500, istep2=12500,
    value1=300.0, value2=300.0,
&end
&wt
  type='END',
&end
&rst
  iat=0,
&end
Hold the DNA fixed
500.0
RES 1 30
END
END

```

ta_md_ew.in

5DNB, initial dynamics with Ewald mesh particle method on ds DNAs, model1, 9.0 cut. (Allowing only water and Na⁺ to move freely)

```

&cntrl

timlim = 999999., nmropt = 0,
ntx  = 1,  irest = 0,  ntrx = 1,  nt xo = 1,
ntr  = 100,  ntwx = 500,  ntwv = 0,  nt we = 0,
ntwxm = 999999, ntwvm = 999999, ntwem = 999999, ioutfm = 0,
ntwp rt = 0,

ntf  = 2,  ntb  = 2,  idiel = 1,  dielc = 1.0,
cut  = 9.0,  ntnb = 1,  nsnb = 10,  nt id = 0,
scnb = 2.0,  scee = 1.2,  cut2nd = 0.0,  ichdna = 0,

isftrp = 0,  rwell = 0.0,

ipol = 0,

ibelly = 0,  ntr  = 1,

imin = 0,

nrun = 1,
nstlim = 12500,
ndfmin = 0,  ntc m = 0,  nsc m = 0,
init  = 3,  t    = 0.0,  dt   = 0.002,

temp0 = 300.0,  tempi = 100.0,
ig     = 71277,  heat  = 0.0,

```

```

ntt = 1,    isolv = 0,    dtemp = 0.0,
tautp = 0.2,    tauts = 0.2,
vlimit = 20.0,

tauv0 = 0.0,    tauv = 0.1,    vzero = 0.0,

ntp = 1,    pres0 = 1.0,    comp = 44.6,
taup = 0.2,    npscal = 1,

ntc = 2,    tol = 0.00001,

imgslt = 0,    iftres = 1,
jfastw = 0,

ivcap = 0,    matcap = 0,    fcap = 1.5,

iewald = 1,

&end
41.9970599 42.9345782 73.2162581 90.0 90.0 90.0
45 45 80 4 0 0 0
0.00001
Allowing only the water and counterions to move in the belly
500.0
RES 1 30
END
END

```

equil_min1.in

5DNB, equilibration step 1. Minimization of ds DNAs imposing 25 kcal/mole·Å position restraints on DNA.
&cntrl

```

timlim = 999999., nmropt = 0,
ntx = 1,    irest = 0,    ntrx = 1,    nt xo = 1,
ntpr = 500,    nt wx = 0,    nt wv = 0,    nt we = 0,
ntwxm = 999999, nt wvm = 999999, nt wem = 999999, ioutfm = 0,
ntwp rt = 0,

ntf = 1,    ntb = 2,    idiel = 1,    dielc = 1.0,
cut = 9.0,    ntnb = 1,    nsnb = 10,    nt id = 0,
scnb = 2.0,    scee = 1.2,    cut2nd = 0.0,    ichdna = 0,

isftrp = 0,    rwell = 0.0,

ipol = 0,

ibelly = 0,    ntr = 1,

imin = 1,
maxcyc = 1000,
ncyc = 5000,
ntmin = 1,    dx0 = 0.1,    dxm = 0.5,    drms = 0.0001,

```

```

nrun = 0,
nstlim = 10000,
ndfmin = 0,   ntcn = 0,   nscn = 0,
init = 3,    t    = 0.0,   dt   = 0.002,

temp0 = 300.0, tempi = 300.0,
ig    = 71277, heat = 0.0,
ntt   = 0,   isolv = 0,   dtemp = 0.0,
tautp = 0.2,   tauts = 0.2,
vlimit = 20.0,

tauv0 = 0.0,   tauv = 0.1,   vzero = 0.0,

ntp   = 1,   pres0 = 1.0,   comp = 44.6,
taup  = 0.2,   npscal = 1,

ntc   = 1,   tol   = 0.0005,

imgslt = 0,   iftres = 1,
jfastw = 0,

ivcap = 0,   matcap = 0,   fcap = 1.5,

iewald = 1,

&end
41.8090432 42.7423643 72.8884761 90.0000000 90.0000000 90.0000000
45 45 80 4 0 0 0
0.000001
Constraints
25.0
RES 1 30
END
END

```

equil_md1.in

5DNB, equilibration step 1. MD on ds DNAs imposing 25 kcal/mole·Å position restraints on DNA.
&cntrl

```

timlim = 999999., nmropt = 0,
ntx   = 1,   irest = 0,   ntrx = 1,   ntco = 1,
ntpr  = 500,   ntwx = 0,   ntwv = 0,   ntwe = 0,
ntwxm = 999999, ntwvm = 999999, ntwem = 999999, ioutfm = 0,
ntwpri = 0,

ntf   = 2,   ntb = 2,   idiel = 1,   dielc = 1.0,
cut   = 9.0,   ntnb = 1,   nsnb = 10,   ntid = 0,
scnb  = 2.0,   scee = 1.2,   cut2nd = 0.0,   ichdna = 0,

isfrp = 0,   rwell = 0.0,

ipol  = 0,

```

```

ibelly = 0,    ntr  = 1,

imin  = 0,
maxcyc = 1000,
ncyc  = 5000,
ntmin = 1,    dx0  = 0.1,   dxm  = 0.5,   drms  = 0.0001,

nrun  = 0,
nstlim = 1500,
ndfmin = 0,    ntcn  = 0,    nscn  = 0,
init  = 3,    t    = 0.0,   dt    = 0.002,

temp0 = 298.0, tempi = 298.0,
ig    = 71277, heat  = 0.0,
ntt   = 1,    isolv = 0,    dtemp = 0.0,
tautp = 0.2,  tauts = 0.2,
vlimit = 20.0,

tauv0 = 0.0,  tauv  = 0.1,  vzero = 0.0,

ntp   = 1,    pres0 = 1.0,   comp  = 44.6,
taup  = 0.2,  npscal = 1,

ntc   = 2,    tol   = 0.0005,

imgslt = 0,    iftres = 1,
jfastw = 0,

ivcap = 0,    matcap = 0,    fcap  = 1.5,

iewald = 1,

&end
41.8090432 42.7423643 72.8884761 90.0000000 90.0000000 90.0000000
45 45 80 4 0 0 0
0.000001
Constraints
25.0
RES 1 30
END
END

```

equil_min2a.in

5DNB, equilibration step 2a. Minimization of ds DNAs imposing 20 kcal/mole·Å position restraints on DNA.
&cntrl

```

timlim = 999999., nmrmax = 0,
ntx   = 1,    irest = 0,    ntrn  = 1,    ntco  = 1,
ntpr  = 500,  ntwn  = 0,    ntww  = 0,    ntwe  = 0,
ntwxm = 999999, ntwwm = 999999, ntwev = 999999, ioutfm = 0,
ntwprt = 0,

ntf   = 1,    ntb   = 2,    idiel = 1,    dielc = 1.0,

```

```

cut  = 9.0,  ntnb  = 1,   nsnb  = 10,  ntid  = 0,
scnb  = 2.0,  scee  = 1.2,  cut2nd = 0.0,  ichdna = 0,

isftrp = 0,   rwell  = 0.0,

ipol  = 0,

ibelly = 0,   ntr   = 1,

imin  = 1,
maxcyc = 600,
ncyc  = 5000,
ntmin  = 1,   dx0   = 0.1,   dxm   = 0.5,   drms  = 0.0001,

nrun  = 0,
nstlim = 10000,
ndfmin = 0,   ntcn  = 0,   nscn  = 0,
init  = 3,   t     = 0.0,   dt    = 0.002,

temp0 = 300.0,  tempi  = 300.0,
ig    = 71277,  heat  = 0.0,
ntt   = 0,     isolv = 0,   dtemp = 0.0,
tautp = 0.2,   tauts = 0.2,
vlimit = 20.0,

tauv0 = 0.0,   tauv  = 0.1,   vzero = 0.0,

ntp   = 1,     pres0 = 1.0,   comp  = 44.6,
taup  = 0.2,   npscal = 1,

ntc   = 1,     tol   = 0.0005,

imgslt = 0,    iftres = 1,
jfastw = 0,

ivcap  = 0,    matcap = 0,    fcap  = 1.5,

iewald = 1,

&end
41.5828858 42.5111583 72.4942009 90.0000000 90.0000000 90.0000000
45 45 75 4 0 0 0
0.000001
Constraints
20.0
RES 1 30
END
END

```

equil_min2b.in

5DNB, equilibration step 2b. Minimization of ds DNAs imposing 15 kcal/mole·Å position restraints on DNA.
&cntrl

```

timlim = 999999., nmropt = 0,
ntx  = 1,   irest = 0,   ntrx = 1,   nt xo = 1,
ntpr = 500,  ntwx = 0,   ntwv = 0,   ntwe = 0,
ntwxm = 999999, ntwvm = 999999, ntwem = 999999, ioutfm = 0,
ntwppt = 0,

ntf  = 1,   ntb  = 2,   idiel = 1,   dielc = 1.0,
cut  = 9.0,  ntnb = 1,   nsnb = 10,  nt id = 0,
scnb = 2.0,  scee = 1.2,  cut2nd = 0.0,  ichdna = 0,

isftrp = 0,   rwell = 0.0,

ipol  = 0,

ibelly = 0,   ntr  = 1,

imin  = 1,
maxcyc = 600,
ncyc  = 5000,
ntmin = 1,   dx0  = 0.1,   dxm  = 0.5,   drms  = 0.0001,

nrun  = 0,
nstlim = 10000,
ndfmin = 0,   ntc m = 0,   nscm = 0,
init  = 3,   t    = 0.0,   dt   = 0.002,

temp0 = 300.0,  tempi = 300.0,
ig    = 71277,  heat  = 0.0,
ntt   = 0,     isolv = 0,   dtemp = 0.0,
tautp = 0.2,   tauts = 0.2,
vlimit = 20.0,

tauv0 = 0.0,   tauv  = 0.1,   vzero = 0.0,

ntp   = 1,     pres0 = 1.0,   comp  = 44.6,
taup  = 0.2,   npscal = 1,

ntc   = 1,     tol   = 0.0005,

imgslt = 0,    iftres = 1,
jfastw = 0,

ivcap = 0,     matcap = 0,    fcap  = 1.5,

iewald = 1,

&end
41.5828858 42.5111583 72.4942009 90.0000000 90.0000000 90.0000000
45 45 75 4 0 0 0
0.000001
Constraints
15.0
RES 1 30
END
END

```

equil_min2c.in

5DNB, equilibration step 2c. Minimization of ds DNAs imposing 10 kcal/mole·Å position restraints on DNA.
&cntrl

timlim = 999999., nmropt = 0,
ntx = 1, irest = 0, ntrx = 1, nt xo = 1,
ntpr = 500, ntwx = 0, ntwv = 0, ntwe = 0,
ntwxm = 999999, ntwvm = 999999, ntwem = 999999, ioutfm = 0,

ntf = 1, ntb = 2, idiel = 1, dielc = 1.0,
cut = 9.0, ntnb = 1, nsnb = 10, nt id = 0,
scnb = 2.0, scee = 1.2, cut2nd = 0.0, ichdna = 0,

isftrp = 0, rwell = 0.0,

ipol = 0,

ibelly = 0, ntr = 1,

imin = 1,
maxcyc = 600,
ncyc = 5000,
ntmin = 1, dx0 = 0.1, dxm = 0.5, drms = 0.0001,

nrun = 0,
nstlim = 10000,
ndfmin = 0, ntc m = 0, nsc m = 0,
init = 3, t = 0.0, dt = 0.002,

temp0 = 300.0, tempi = 300.0,
ig = 71277, heat = 0.0,
ntt = 0, isolv p = 0, dtemp = 0.0,
tautp = 0.2, tauts = 0.2,
vlimit = 20.0,

tauv0 = 0.0, tauv = 0.1, vzero = 0.0,

ntp = 1, pres0 = 1.0, comp = 44.6,
taup = 0.2, np scal = 1,

ntc = 1, tol = 0.0005,

imgsl t = 0, iftres = 1,
jfastw = 0,

ivcap = 0, matcap = 0, fcap = 1.5,

iewald = 1,

&end
41.5828858 42.5111583 72.4942009 90.0000000 90.0000000 90.0000000
45 45 75 4 0 0 0
0.000001
Constraints

```
10.0
RES 1 30
END
END
```

equil_min2d.in

5DNB, equilibration step 2d. Minimization of ds DNAs imposing 5 kcal/mole·Å position restraints on DNA.
&cntrl

```
timlim = 999999., nmropt = 0,
ntx  = 1,   irest = 0,   ntrx = 1,   nt xo = 1,
ntpr  = 500,   nt wx = 0,   nt wv = 0,   nt we = 0,
ntwxm = 999999, nt wvm = 999999, nt wem = 999999, ioutfm = 0,
```

```
ntf  = 1,   ntb  = 2,   idiel = 1,   dielc = 1.0,
cut  = 9.0,   ntnb = 1,   nsnb = 10,   nt id = 0,
scnb = 2.0,   scee = 1.2,   cut2nd = 0.0,   ichdna = 0,
```

```
isftrp = 0,   rwell = 0.0,
```

```
ipol  = 0,
```

```
ibelly = 0,   ntr  = 1,
```

```
imin  = 1,
maxcyc = 600,
ncyc  = 5000,
ntmin = 1,   dx0  = 0.1,   dxm  = 0.5,   drms  = 0.0001,
```

```
nrun  = 0,
nstlim = 10000,
ndfmin = 0,   ntcn = 0,   nscn = 0,
init  = 3,   t    = 0.0,   dt   = 0.002,
```

```
temp0 = 300.0, tempi = 300.0,
ig     = 71277, heat  = 0.0,
ntt    = 0,   isolvp = 0,   dtemp = 0.0,
tautp  = 0.2,   tauts = 0.2,
vlimit = 20.0,
```

```
tauv0 = 0.0,   tauv  = 0.1,   vzero = 0.0,
```

```
ntp    = 1,   pres0 = 1.0,   comp  = 44.6,
taup   = 0.2,   npscal = 1,
```

```
ntc    = 1,   tol   = 0.0005,
```

```
imgslt = 0,   iftres = 1,
jfastw = 0,
```

```
ivcap  = 0,   matcap = 0,   fcap  = 1.5,
```

```
iewald = 1,
```



```

&end
41.5828858 42.5111583 72.4942009 90.0000000 90.0000000 90.0000000
45 45 75 4 0 0 0
0.000001
Constraints
5.0
RES 1 30
END
END

```

equil_min2e.in

5DNB, equilibration step 2e. Minimization of ds DNAs with no position restraints on DNA.
&cntrl

```

timlim = 999999., nmropt = 0,
ntx  = 1,  irest = 0,  ntrx = 1,  nt xo = 1,
ntpr = 500,  ntwx = 0,  ntwv = 0,  ntwe = 0,
ntwxm = 999999, ntwvm = 999999, ntwem = 999999, ioutfm = 0,
ntwp rt = 0,

```

```

ntf  = 1,  ntb  = 2,  idiel = 1,  dielc = 1.0,
cut  = 9.0,  ntnb = 1,  nsnb = 10,  nt id = 0,
scnb = 2.0,  scee = 1.2,  cut2nd = 0.0,  ichdna = 0,

```

```

isftrp = 0,  rwell = 0.0,

```

```

ipol  = 0,

```

```

ibelly = 0,  ntr  = 0,

```

```

imin  = 1,
maxcyc = 600,
ncyc  = 5000,
ntmin = 1,  dx0  = 0.1,  dxm  = 0.5,  drms = 0.0001,

```

```

nrun  = 0,
nstlim = 10000,
ndfmin = 0,  ntc m = 0,  nscm = 0,
init  = 3,  t    = 0.0,  dt   = 0.002,

```

```

temp0 = 300.0,  tempi = 300.0,
ig    = 71277,  heat  = 0.0,
ntt   = 0,  isolv p = 0,  dtemp = 0.0,
tautp = 0.2,  tauts = 0.2,
vlimit = 20.0,

```

```

tauv0 = 0.0,  tauv  = 0.1,  vzero = 0.0,

```

```

ntp  = 1,  pres0 = 1.0,  comp  = 44.6,
taup = 0.2,  npscal = 1,

```

```

ntc  = 1,  tol  = 0.0005,

```

```

imgslt = 0,    iftres = 1,
jfastw = 0,

ivcap = 0,    matcap = 0,    fcap = 1.5,

iewald = 1,

&end
41.5828858 42.5111583 72.4942009 90.0000000 90.0000000 90.0000000
45 45 75 4 0 0 0
0.000001

```

equil_md.in

5DNB, warm it up. MD on ds DNAs by heating the system from 100 K to 300 K
&cntrl

```

timlim = 999999., nmropt = 1,
ntx = 1,    irect = 0,    ntrx = 1,    ntco = 1,
ntr = 100,  ntwx = 500,  ntwv = 0,    ntwe = 0,
ntwxm = 999999, ntwvm = 999999, ntwem = 999999, ioutfm = 0,
ntwppt = 0,

ntf = 2,    ntb = 2,    idiel = 1,    dielc = 1.0,
cut = 9.0,  ntnb = 1,    nsnb = 10,  ntid = 0,
scnb = 2.0,  scee = 1.2,  cut2nd = 0.0,  ichdna = 0,

isftrp = 0,    rwell = 0.0,

ipol = 0,

ibelly = 0,    ntr = 0,

imin = 0,
maxcyc = 500,
ncyc = 5000,
ntmin = 1,    dx0 = 0.1,    dxm = 0.5,    drms = 0.0001,

nrun = 1,
nstlim = 10000,
ndfmin = 0,    ntcn = 0,    nscn = 0,
init = 3,    t = 0.0,    dt = 0.002,

temp0 = 300.0,  tempi = 100.0,
ig = 71277,  heat = 0.0,
ntt = 1,    isolv = 0,    dtemp = 0.0,
tautp = 0.2,  tauts = 0.2,
vlimit = 20.0,

tauv0 = 0.0,  tauv = 0.1,  vzero = 0.0,

ntp = 1,    pres0 = 1.0,  comp = 44.6,
taup = 0.2,  npscal = 1,

```

```

ntc = 2,    tol = 0.0005,

imgslt = 0,    iftres = 1,
jfastw = 0,

ivcap = 0,    matcap = 0,    fcap = 1.5,

iewald = 1,

&end
41.5828858 42.5111583 72.4942009 90.0000000 90.0000000 90.0000000
45 45 75 4 0 0 0
0.000001
&wt
type='TEMP0', istep1=0, istep2=1000,
value1=100.0, value2=300.0,
&end
&wt
type='TEMP0', istep1=1000, istep2=10000,
value1=300.0, value2=300.0,
&end
&wt
type='END',
&end
&rst
iat=0,
&end
END
END

```

equil_md_prod.in

5DNB, warm it up. MD production runs for 1 ns on ds DNAs by heating the system from 100 K to 300 K
&cntrl

```

timlim = 999999., nmropt = 1,
ntx = 1,    irest = 0,    ntrx = 1,    nt xo = 1,
ntpr = 100,    ntwx = 500,    ntwv = 0,    ntwe = 0,
ntwxm = 999999, ntwvm = 999999, ntwem = 999999, ioutfm = 0,
ntwp rt = 0,

ntf = 2,    ntb = 2,    idiel = 1,    dielc = 1.0,
cut = 9.0,    ntnb = 1,    nsnb = 10,    nt id = 0,
scnb = 2.0,    scee = 1.2,    cut2nd = 0.0,    ichdna = 0,

isftrp = 0,    rwell = 0.0,

ipol = 0,

ibelly = 0,    ntr = 0,

imin = 0,
maxcyc = 500,

```

```

ncyc = 5000,
ntmin = 1,   dx0 = 0.1,   dxm = 0.5,   drms = 0.0001,

nrun = 1,
nstlim = 500000,
ndfmin = 0,   ntcn = 0,   nscn = 0,
init = 3,   t = 0.0,   dt = 0.002,

temp0 = 300.0,   tempi = 100.0,
ig = 71277,   heat = 0.0,
ntt = 1,   isolv = 0,   dtemp = 0.0,
tautp = 0.2,   tauts = 0.2,
vlimit = 20.0,

tauv0 = 0.0,   tauv = 0.1,   vzero = 0.0,

ntp = 1,   pres0 = 1.0,   comp = 44.6,
taup = 0.2,   npscal = 1,

ntc = 2,   tol = 0.0005,

imgslt = 0,   iftres = 1,
jfastw = 0,

ivcap = 0,   matcap = 0,   fcap = 1.5,

iewald = 1,

&end
41.5828858 42.5111583 72.4942009 90.0000000 90.0000000 90.0000000
45 45 75 4 0 0 0
0.000001
&wt
type='TEMP0', istep1=0,   istep2=1000,
value1=100.0, value2=300.0,
&end
&wt
type='TEMP0', istep1=1000, istep2=10000,
value1=300.0, value2=300.0,
&end
&wt
type='END',
&end
&rst
iat=0,
&end
END
END

```

Note: PME box size information must be updated everytime to match the final box size from the previous run.
For triple stranded DNAs appropriate changed must be made.

process_mdout.perl

```
#!/bin/perl

if ($#ARGV < 0) {
    print " Incorrect usage...\n";
    exit;
}

foreach $i ( 0..$#ARGV ) {
    $filein = $ARGV[$i];
    $checkfile = $filein;
    $checkfile =~ s/\./Z/;
    if ( $filein ne $checkfile ) {
        open(INPUT, "zcat $filein |") ||
            die "Cannot open compressed $filein -- $!\n";
    } else {
        open(INPUT, $filein) || die "Cannot open $filein -- $!\n";
    }
    print "Processing sander output file ($filein)...\n";
    &process_input;
    close(INPUT);
}

print "Starting output...\n";
@sortedkeys = sort by_number keys(%TIME);
@sortedavgkeys = sort by_number keys(%AVG_TIME);

foreach $i ( TEMP, TSOLUTE, TSOLVENT, PRES, EKCM, ETOT, EKTOT, EPTOT, DENSITY, VOLUME ) {
    print "Outputting summary.$i\n";
    open(OUTPUT, "> summary.$i");
    %outarray = eval "%$i";
    foreach $j ( @sortedkeys ) {
        print OUTPUT "$j ", $outarray{$j}, "\n";
    }
    close (OUTPUT);

    print "Outputting summary_avg.$i\n";
    open(OUTPUT, "> summary_avg.$i");
    %outarray = eval "%AVG_$i";
    foreach $j ( @sortedavgkeys ) {
        print OUTPUT "$j ", $outarray{$j}, "\n";
    }
    close (OUTPUT);

    print "Outputting summary_rms.$i\n";
    open(OUTPUT, "> summary_rms.$i");
    %outarray = eval "%RMS_$i";
    foreach $j ( @sortedavgkeys ) {
        print OUTPUT "$j ", $outarray{$j}, "\n";
    }
    close (OUTPUT);
}
```

```

}

sub by_number {
  if ($a < $b) {
    -1;
  } elsif ($a == $b) {
    0;
  } elsif ($a > $b) {
    1;
  }
}

sub process_input {

  $status = 0;
  $debug = 0;
  while ( <INPUT> ) {
    $string = $_;

    print $_ if ( ! /NB-upda/ && $debug );

    if (/A V E R A G E S/) {
      $averages = 1;
      ($averages_over) = /. *O V E R. *(\d*). *S T E P S/;
    }

    $rms = 1 if (/R M S/);

    if (/NSTEP/) {
      ($time, $temp, $pres) =
        /NSTEP. *=(.*\d*\.\d*). *TEMP. *=(.*\d*\.\d*). *PRESS = (.*\d*\.\d*)/;
      if ( $debug ) {
        print $_;
        print "time is $time, temp is $temp, pres is $pres\n";
      }
      $_ = <INPUT>;
    }

    if (/Etot/) {
      ($etot, $ektot, $eptot) =
        /Etot. *=(.*\d*\.\d*). *EKtot. *=(.*\d*\.\d*). *EPtot. *=(.*\d*\.\d*)/;
      if ( $debug ) {
        print $_;
        print "Etot is $etot, ektot is $ektot, eptot is $eptot\n";
      }
      $_ = <INPUT>;
    }

    if (/BOND. *ANGLE. *DIHED/) {
      ($bond, $angle, $dihedral) =
        /BOND. *=(.*\d*\.\d*). *ANGLE. *=(.*\d*\.\d*). *DIHED. *=(.*\d*\.\d*)/;
      if ( $debug ) {
        print $_;
        print "bond is $bond, angle is $angle, dihedral is $dihedral\n";
      }
      $_ = <INPUT>;
    }
  }
}

```

```

}
if (/1-4 NB/) {
    ($nb14, $eel14, $nb) =
        /1-4 NB.*=(.*\d*\.\d*).1-4 EEL.*=(.*\d*\.\d*).VDWAALS.*=(.*\d*\.\d*)/;
    if ( $debug ) {
        print $_;
        print "nb14 is $nb14, eel14 is $eel14, vdwaals is $nb\n";
    }
    $_ = <INPUT>;
}
if (/EELEC/) {
    ($eel, $ehbond, $constraint) =
        /EELEC.*=(.*\d*\.\d*).EHBOND.*=(.*\d*\.\d*).CONSTRAINT.*=(.*\d*\.\d*)/;
    if ( $debug ) {
        print $_;
        print "eel is $eel, ehbond is $ehbond, constraint is $constraint\n";
    }
    $_ = <INPUT>;
}

#
# check to see if EAMBER is in the mdout file (present when
# NTR=1)
#
    if ( /EAMBER/ ) {
        $_ = <INPUT>;
    }
}
if (/EKCMT/) {
    ($ekcmt, $virial, $volume) =
        /EKCMT.*=(.*\d*\.\d*).VIRIAL.*=(.*\d*\.\d*).VOLUME.*=(.*\d*\.\d*)/;
    if ( $debug ) {
        print $_;
        print "Ekcmt is $ekcmt, virial is $virial, volume is $volume\n";
    }
    $_ = <INPUT>;
}
if (/T_SOLUTE/) {
    ($tsolute, $tsolvent) =
        /T_SOLUTE.*=(.*\d*\.\d*).T_SOLVENT.*=(.*\d*\.\d*)/;
    if ( $debug ) {
        print $_;
        print "Temp solute is $tsolute, temp solvent is $tsolvent\n";
    }
    $_ = <INPUT>;
}

if (/Density/) {
    ($density) = /.Density.*=(.*\d*\.\d*)/;
    if ( $debug ) {
        print $_;
        print "Density is $density\n";
    }
    $_ = <INPUT>;
}

#    update arrays

```

```

if ( $averages == 1 ) {
    $AVG_TIME{$time}    = $time;
    $AVG_TEMP{$time}    = $temp;
    $AVG_PRES{$time}    = $pres;
    $AVG_ETOT{$time}    = $etot;
    $AVG_EKTOT{$time}   = $ektot;
    $AVG_EPTOT{$time}   = $eptot;
    $AVG_BOND{$time}    = $bond;
    $AVG_ANGLE{$time}   = $angle;
    $AVG_DIHEDRAL{$time} = $dihedral;
    $AVG_NB14{$time}    = $nb14;
    $AVG_EEL14{$time}   = $eel14;
    $AVG_NB{$time}      = $nb;
    $AVG_EEL{$time}     = $eel;
    $AVG_EHBOND{$time}  = $ehbond;
    $AVG_CONSTRAINT{$time} = $constraint;
    $AVG_EKCMT{$time}   = $ekcmt;
    $AVG_VIRIAL{$time}  = $virial;
    $AVG_VOLUME{$time}  = $volume;
    $AVG_TSOLUTE{$time} = $tsolute;
    $AVG_TSOLVENT{$time} = $tsolvent;
    $AVG_DENSITY{$time} = $density;

    $averages = 0;
} elseif ( $rms == 1 ) {
    $RMS_TIME{$time}    = $time;
    $RMS_TEMP{$time}    = $temp;
    $RMS_PRES{$time}    = $pres;
    $RMS_ETOT{$time}    = $etot;
    $RMS_EKTOT{$time}   = $ektot;
    $RMS_EPTOT{$time}   = $eptot;
    $RMS_BOND{$time}    = $bond;
    $RMS_ANGLE{$time}   = $angle;
    $RMS_DIHEDRAL{$time} = $dihedral;
    $RMS_NB14{$time}    = $nb14;
    $RMS_EEL14{$time}   = $eel14;
    $RMS_NB{$time}      = $nb;
    $RMS_EEL{$time}     = $eel;
    $RMS_EHBOND{$time}  = $ehbond;
    $RMS_CONSTRAINT{$time} = $constraint;
    $RMS_EKCMT{$time}   = $ekcmt;
    $RMS_VIRIAL{$time}  = $virial;
    $RMS_VOLUME{$time}  = $volume;
    $RMS_TSOLUTE{$time} = $tsolute;
    $RMS_TSOLVENT{$time} = $tsolvent;
    $RMS_DENSITY{$time} = $density;

    $rms = 0;
} else {
    $TIME{$time}    = $time;
    $TEMP{$time}    = $temp;
    $PRES{$time}    = $pres;
    $ETOT{$time}    = $etot;
    $EKTOT{$time}   = $ektot;
    $EPTOT{$time}   = $eptot;
    $BOND{$time}    = $bond;

```



```

$ANGLE{$time}    = $angle;
$DIHEDRAL{$time} = $dihedral;
$NB14{$time}     = $nb14;
$EEL14{$time}    = $eel14;
$NB{$time}       = $nb;
$EEL{$time}      = $eel;
$EHBOND{$time}   = $ehbond;
$CONSTRAINT{$time} = $constraint;
$EKCMT{$time}    = $ekcmt;
$VIRIAL{$time}   = $virial;
$VOLUME{$time}   = $volume;
$TSOLUTE{$time}  = $tsolute;
$TSOLVENT{$time} = $tsolvent;
$DENSITY{$time}  = $density;
    }
}
}
}

```

carnal_rms_start.in

```

FILES_IN
  PARM p1 tat_fsi.top;
  STREAM s1 tat_md_prod.traj;
  STREAM s2 tat_md_prod_imaged.traj;
  STREAM s3 tat_md_prod_strip.traj;
FILES_OUT
  TABLE tab1 rms_of_notimagaed;
  TABLE tab2 rms_of_imaged;
  TABLE tab3 rms_of_strip;
DECLARE
  GROUP gALL (RES 1 - 45);
  RMS r1 FIT gALL s1;
  RMS r2 FIT gALL s2;
  RMS r3 FIT gALL s3;
OUTPUT
  TABLE tab1 r1;
  TABLE tab2 r2;
  TABLE tab3 r3;
END
process_carnal.perl rms_of_notimagaed.plt < rms_of_notimagaed
process_carnal.perl rms_of_imaged.plt < rms_of_imaged
process_carnal.perl rms_of_strip.plt < rms_of_strip

```

process_carnal.perl

```

#!/usr/sbin/perl

$output = "1.data";

```

```

if ($#ARGV >= 0) {
    $output = $ARGV[0];
}

$counter = 0;
$acounter = 0;
while ( <STDIN> ) {
    $string = $_;
    if ( /^L[0-9]*/ ) {
        @a = split(' ', $string);
        @data1[$acounter++] = $a[1];
    }
    $counter++;
}

print "Successfully read in $counter lines\n";
print "representing $acounter values...\n";
print "Output will be directed to $output\n";

$start_time = $ENV{'START_TIME'};
if ( ! $start_time ) {
    $start_time = 0;
}
print "Starting time in ps is $start_time\n";

$interval = $ENV{'INTERVAL'};
if ( ! $interval ) {
    $interval = 1;
}

open(FILE1, "> $output") || die "Could not open $output -- $!";

for ( $i = 0; $i < $acounter; $i++ ) {
    $time = $i * $interval + $start_time;
    print FILE1 "$time $data1[$i]\n";
}

close(FILE1);

```

mdanal.in

```

MDANAL input
1 0 0 0 12 12 1
1 0
0 0 0 0
0 1 0 2 0 0
1
group spec - all DNA atoms
RES 1 30
END
END

```