ASSESSMENT OF MOLECULAR DYNAMICS FORCE FIELDS FOR DRUG-DNA INTERACTION

Dissertation submitted for the award of the degree of

Master of Philosophy

In

Physics

by

Manas Misra

(Enrollment No. – 551/18)

Under the supervision of

Dr. Anil Kumar Yadav

Department of Physics

School of Physical & Decision Science

Babasaheb Bhimrao Ambedkar University

(A Central University)

Vidya Vihar, Raebareli Road

Lucknow – 226025, U. P., India

2022

DEDICATED
TO
MY PARENTS
AND
NOBLE TEACHERS

DECLARATION

I declare that the dissertation entitled **"ASSESSMENT OF MOLECULAR DYNAMICS FORCE FIELDS FOR DRUG-DNA INTERACTION"** has been prepared by me under the supervision of **Dr. Anil Kumar Yadav**, Assistant Professor, Department of Physics, School of Physical & Decision Science , Babasaheb Bhimrao Ambedkar University, Lucknow. No part of this dissertation has formed the basis for the award of any degree, diploma or fellowship previously. Further, I declare that the material embodied in the present work is based on original research work and the indebtedness to others has been duly acknowledged at relevant places. This is also declared that the dissertation is essentially free from any kinds of plagiarism.

मानस मिश्र

Manas Misra

Department of Physics

School of Physical & Decision Science

Babasaheb Bhimrao Ambedkar University

Vidya Vihar, Raebareli Road,

Lucknow – 226025,

U. P., India

Date:

Place: Lucknow

CERTIFICATE

This is to certify that the M.Phil. dissertation titled **"ASSESSMENT OF MOLECULAR DYNAMICS FORCE FIELDS FOR DRUG-DNA INTERACTION"** submitted by **Mr. Manas Misra** is an original research work and has not been previously submitted in part or full for the award of any other degree or diploma to this or any other university.

The M.Phil. dissertation submitted to Babasaheb Bhimrao Ambedkar University, Lucknow satisfies all the requirements as stipulated in the *Master of Philosophy (M.Phil.)/ Doctor of Philosophy (Ph.D.) regulations amended in 2019* and it is fit for submission and evaluation for the award of the degree of Master of Philosophy of the University.

Date:

Supervisor
29/6/22

22/06/2022
Head of the Department
विभागाध्यक्ष
Head
भौतिकी विभाग
Deptt. of Physics
बाबा साहेब भीमराव अम्बेडकर विश्वविद्यालय
Baba Saheb Bhimrao Ambedkar University
लखनऊ - 226025 उ० प्र०, भारत
Lucknow - 226025, U.P., India

ACKNOWLEDGEMENT

First of all, I would like to convey my sincere thanks and heartiest gratitude to my research supervisor **Dr. Anil Kumar Yadav**. I feel grateful to have a supervisor who has kept his door open for any scientific discussions at any time. He has helped in shaping me through the year not only in the scientific field but also in improving other skills.

I express my deepest gratitude to **Prof. (Dr.) Bal Chandra Yadav**, Head, Department of Physics, Babasaheb Bhimrao Ambedkar University, Lucknow. I always admire his very humble nature towards students.

I would like to express my gratitude to **Prof. (Dr.) Devesh Kumar**, **Dr. Devendra Singh, Dr. Ramesh Chandra**, **Dr. Khem Bahadur Thapa** and **Dr. Ajeet Kumar Maurya** for their constant support and valuable suggestions throughout the entire tenure of M.Phil. I would also take the opportunity to express my gratitude to the **office staff** for their help and kind support at various stages of work.

I'm highly thankful to **Dr. Anwesh Pandey**, Research Associate, ACDSD, CSIR-NEIST, Jorhat, Assam for his continuous experimental support and discussions throughout my dissertation work. I would like to thank **Mr. Diptarka Roy,** Research Fellow, Department of Physics, **Mr. Baldev Singh,** Research Fellow,s Department of Chemistry, BBAU Lucknow, and **Dr. Pawan Singh,** Guest faculty, University of Delhi for their consistent support during my research work.

I would like to acknowledge IIT Delhi for providing access to SCFBio web server.

I am also thankful to **Mr. Mohit Kumar,** scholar, Department of Physics, BBAU Lucknow for his constant support and motivation throughout my research work.

I am highly indebted to my parents **Dr. Awadhesh Kumar Misra** and **Mrs. Pushplata Misra** who played an important role in shaping a lot of me. Their encouraging words constantly inspired me to pursue my dreams. I would also like to express my sincere thanks to my brother **Utkarsh Misra** (Research Fellow, Indian Institute of Technology, Bombay) for encouraging me and for being there to guide me always.

मानस मिश्र

Date: **(Manas Misra)**

Place: Lucknow M.Phil. Student

ABSTRACT

Deoxyribonucleic Acid commonly referred to as DNA is the hereditary material in humans. It has three main features that are Translation, Transcription and Replication and these processes are stimulated by the signals in the from of small molecules. These small molecules can be drugs when activation of DNA is required to cure a particular disease. These drugs are known as DNA targeted drugs. There are several anticancer drugs that comes under this category.

The study of interaction of DNA targeted drugs with the nucleic acid is of great importance in pharmacology. There are some experimental as well as computational techniques to study such types of interactions. Molecular Docking Simulation and Molecular Dynamics (MD) Simulation are the computational techniques used to study the interactions of drugs with the nucleic acids. Molecular Docking Simulation predicts the binding site of the ligand to the target structure whereas the MD Simulation predicts the evolution of the molecular structure with time. Furthermore, it is used to study the effect of the drug on the DNA and to check the stability of the DNA after the binding of the drug.

In the present work, the performance of seven versions of AMBER force field pre-embedded in the GROMACS software suite has been assessed. We have selected a minor groove binder from the literatures and it was docked on to the B-DNA sequence (PDB Id-195D). The Ligand-DNA complex was put under 100ns of MD simulation using all seven versions of AMBER force field. So, a total of 700ns of MD Simulations were done to conclude any result.

The analysis of the MD Simulation was done on the basis of Radius of Gyration, Root Mean Square Deviation, Root Mean Square Fluctuation and number of hydrogen bonds. The results showed that although the DNA remains intact throughout the simulation for all the force fields, but AMBER99SB predicted the best results among all other versions as the DNA-ligand complex was in the most compact form throughout the simulation time and the equilibration of the complex structure was obtained.

PREFACE

This dissertation, entitled **"ASSESSMENT OF MOLECULAR DYNAMICS FORCE FIELDS FOR DRUG-DNA INTERACTION"** sum up the results obtained on research carried out in the Department of Physics, Babasaheb Bhimrao Ambedkar University, Lucknow in between 2020-2021 under the supervision of **Dr. Anil Kumar Yadav**, Assistant Professor, Department of Physics, Babasaheb Bhimrao Ambedkar University, Lucknow.

The work carried out in my dissertation has been divided into 4 chapters.

Chapter 1 contains the basic introduction of the DNA and the interactions of the drug with the DNA. In this chapter, we have briefly discussed about the basic components of the DNA like, Phosphate, sugar, nitrogenous bases. We have also discussed the difference between the ribose and deoxyribose sugar. Difference between nucleotides and nucleosides is discussed in this chapter. Furthermore, the discussion about the interactions between the drug and the DNA has also been discussed. The three most general modes of binding of the drug with the DNA are Minor groove binding, Major groove binding and Intercalation and these modes are discussed herein.

Chapter 2 includes the discussion about some of the computational techniques used for the study of the drug-DNA interactions. These methods include Molecular Docking Simulation and Molecular Dynamics Simulation. Molecular Docking Simulation is used to predict the binding mode of the drug with the DNA whereas MD simulations are used to study the effects of the drug binding by predicting the dynamics of the drug-DNA complex. MD simulations uses force fields to generate the topologies of the molecular structure. This chapter presents the discussion on force fields and the methodology of performing the MD simulation.

Chapter 3 presents the introduction of the identified problem for this dissertation work. The lack of consensus about the selection of force field for the simulation of drug-DNA complex was the identified problem. This chapter includes the discussion about the methods used for the current work. The results of the molecular docking simulation and MD simulation are also discussed in this chapter. The results of the MD simulation are analysed based on the parameters like radius of gyration, root mean square deviation, root mean square fluctuation and the number of hydrogen bonds formed between the drug and the DNA.

Chapter 4 finally gives the general conclusions drawn from the present dissertation and future research work that would be productive in the further understanding of performance of the force fields for the simulations of the drug-DNA complexes.

LIST OF ABBREVIATIONS

A	Adenine
AMBER	Assisted Model Building with Energy Refinement
C	Cytosine
CHARMM	Chemistry at HARward Macromolecular Mechanics
DFT	Density Functional Theory
DNA	De-oxy Ribonucleic Acid
G	Guanine
GROMOS	GROningen Molecular Simulation
MD	Molecular Dynamics
mRNA	Messenger Ribonucleic Acid
OPLS-AA	Optimized Potential for Liquid Simulation-All Atom
OPLS-UA	Optimized Potential for Liquid Simulation-United Atom
PDB	Protein Data Bank
PME	Particle Mesh Ewald
QM/MM	Quantum Mechanics/Molecular Mechanics
RMSD	Root Mean Square Deviation
RMSF	Root Mean Square Fluctuation
RNA	Ribonucleic Acid
T	Thymine
U	Uracil

LIST OF TABLES

LIST OF FIGURES

CONTENTS

CHAPTER

1

INTRODUCTION

1. INTRODUCTION

From the very beginning of the civilization, mankind has been facing various health problems. These health problems ranges from common flue to some severe issues like cancer. Primarily, we have used plant extracts to cure specific diseases. But, with the evolution of civilization, people got to know that the chemical composition present in natural resources were responsible for the drug actions. Later, more advance methods have been used in the field of medicines and drug discovery[1–5]. Computational techniques have also been used in this field. All these technologies aim at aiding the mankind in one or the other way[6–11].

In the field of drug designing, the information about the target for any drug is the primary requirements, apart from that, the information about the interaction between the drug and target is of utmost importance. Modelling studies on DNA, RNA, enzymes, proteins, lipids, etc. are considered as the basic requirement for the computational techniques and these studies are certainly a huge interdisciplinary enterprise including biology, chemistry, physics, computer science and even mathematics [12,13].

Deoxyribonucleic acid (DNA), due to its gene expression and protein formation tendencies, acts as target for a class of drugs known as DNA targeted drugs[14–21]. These class of drugs have been shown to have higher selectivity towards its target. Several DNA targeted drugs inhibit the action of a specific protein whereas several others alter the replication process of the DNA reducing the speed of cell growth[22]. So, the study of interaction of the drug with the DNA is an important aspect in the field of pharmacology. Computational techniques are one of the several ways of studying the interactions of the drug with the target DNA sequence. Some of the most common computational techniques to study these

interactions are Molecular Docking Simulations, Molecular Dynamics (MD) Simulations, and Quantum mechanical/molecular mechanical (QM/MM) calculations etc. Molecular docking simulation predicts the binding site of the ligand with respect to the target structure whereas MD simulations provides the information of the dynamics of the molecular system[17,23–25].

MD simulations predict the time evolution of the molecular structure. MD Simulation has been widely used to study the protein-protein interactions, protein folding and unfolding, protein-DNA interactions and many more such processes[26–31]. The basic idea behind the MD simulation is the representation of the potential energy of the molecular system in terms of atomic coordinates with some intrinsic energy parameters. This type of representation of the potential energy of the molecule is termed as the force field. There are several force fields and the performance of a force field depends significantly on the parameterization of the force fields. So, a force field can perform differently for different systems. We will discuss briefly about the force fields in the upcoming chapters.

The selection of force field for a specific system depends on the past works and literature survey. The lack of consensus in the selection of force field for the MD simulation of the DNA-ligand complexes works as the motivation to carry out this dissertation work. In the following dissertation, we will assess the performance of the several force fields to model the drug-DNA complex. In this chapter, we will discuss in brief about the DNA, the interactions of drugs with the DNA and the several ways of the binding of the drug with the DNA.

1.1. Nucleic Acid

Nucleic acids are the linear polymeric chains of the nucleotides and these are acidic in nature and hence they are termed as the nucleic acids. The two of the main nucleic acids are the DNA and the RNA. They carry the genetic information which is responsible for the formation of the proteins[32]. DNA is the main component of the chromosomes. DNA and RNA are very rare to be found free.

Nucleic acids consist of sugar, phosphate and nitrogenous bases. Together a sugar, a phosphate and a base molecule is termed as nucleotide.

1.1.1. Nitrogenous Bases

These bases are the Nitrogen containing organic rings. For the DNA, these bases are Adenine (A), Thymine (T), Cytosine (C) and Guanine (G) whereas the bases of RNA are Adenine (A), Cytosine (C), Guanine (G) and Uracil (U). These bases are classified into two categories that are purines and pyrimidines. Pyrimidine are the six membered heterocyclic rings, but the purines are the pyrimidines connected with a five membered imidazole ring. Purines can be isolated from the nucleic acid, but pyrimidines have not yet been isolated from the nucleic acids. Adenine (A) and Guanine (G) comes under the category of purines but Thymine (T), Cytosine (C), and Uracil (U) comes under the category of pyrimidines[32]. The figure 1.1 depicts the structure of nitrogenous bases. These structures were drawn using ChemDraw[33].

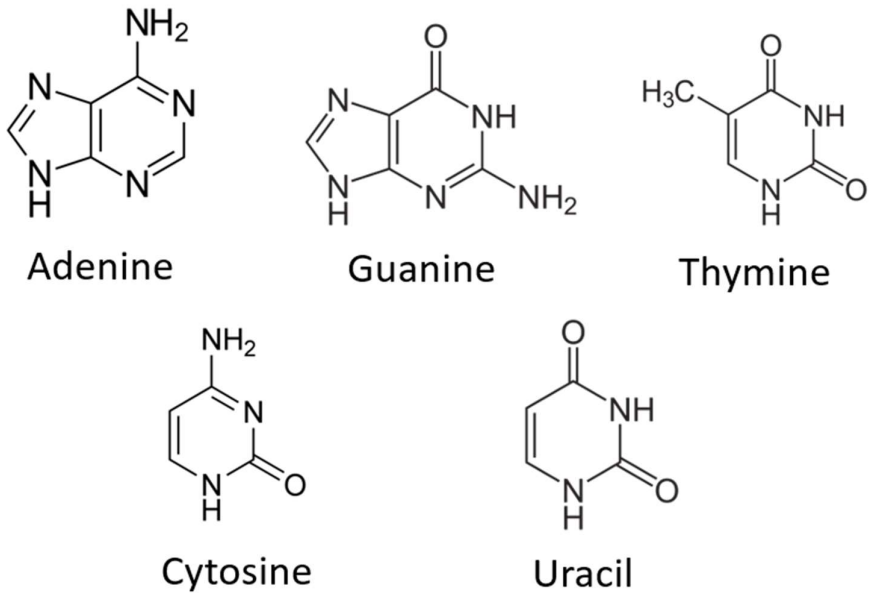

Figure 1.1 *The structure of nucleobases. Adenine and Guanine are purines and Uracil, Thymine and Cytosine are pyrimidines.*

1.1.2. Sugar

The sugar content of the DNA and RNA is also different. DNA consists of the deoxyribose sugar while the RNA has ribose sugars. These sugars are pentoaldose. The structure of the deoxyribose and ribose sugar differ in the sense that in the ribose sugar second carbon atom has a hydroxyl group, a hydrogen atom and two carbon atoms connected to it while in the deoxyribose sugar the second carbon atom has two hydrogen atoms and two carbon atoms connected to it[32].

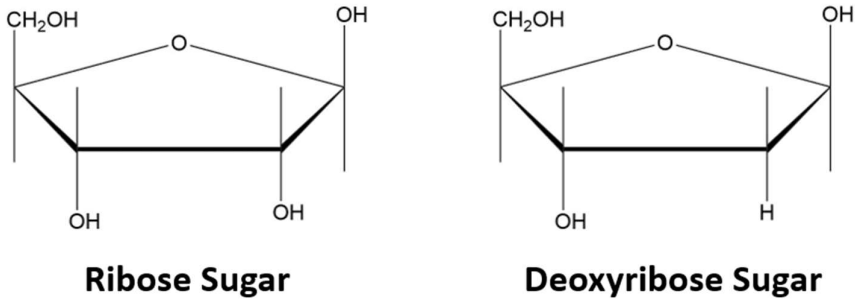

Figure 1.2 *The structural differences between the ribose and deoxyribose sugar*

1.1.3. The Phosphate Group

The phosphate group of the nucleic acids together with the sugar atom forms the backbone of the DNA and RNA. The phosphate group consists of d-electrons and in this group four oxygen atoms are linked with a phosphorous atom via sp^3 hybridization. The structure of the phosphate group is tetrahedral.

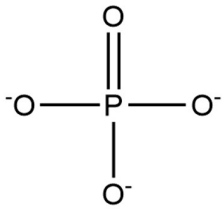

Figure 1.3 *The Structure of the Phosphate group*

1.1.4. Nucleoside and Nucleotide

A nitrogenous base either purine or pyrimidine, attached to a deoxyribose or ribose sugar forms a nucleoside. The nucleoside of the Adenine (A) is termed as Adenosine while the nucleoside of the Guanine (G) is termed as Guanosine. The nucleoside of the Cytosine (C) is Cytidine, nucleoside of Thymine (T) is Thymidine and the nucleoside of Uracil (U) is Uridine. A nucleoside together with a phosphate group is termed as the nucleotide and the nature of the nucleotides are strongly acidic.

Nucleoside = Nucleobase + Sugar

Nucleotide = Nitrogenous base + Sugar + Phosphate

1.2. History and Structure of DNA

DNA was discovered by Friedrich Mischer in the year 1868[34]. After the discovery of the DNA, several studies were carried out to get a deeper insight about the DNA. Avery along with the co-workers reported that the DNA carries the genetic information[35]. The uniqueness of the DNA for each and every species was reported by Chargaff in 1950[36]. The discovery of the double helical structure of DNA was done in the year 1953 when JD Watson and Francis Crick proposed it in their article.[37] Watson and Crick reported that the DNA has two polynucleotide strands which are intertwined to form a double helical structure[38]. The two polynucleotide strands interact via complementary hydrogen bonds which are formed between the nitrogenous bases of the strands. These nitrogenous bases are Adenine (A), Guanine (G), Thymine (T) and Cytosine (C). A is connected with T and G is connected with C. Later in 1980, Dickerson and his co-workers reported the first X-ray crystal structure of DNA which validated the double helical structure of the DNA as proposed by Watson and Crick[39].

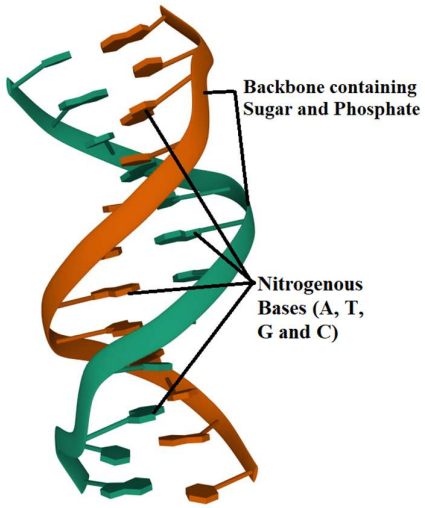

Figure 1.4 Double helical structure of the DNA

Some of the physical properties of the DNA like Helix sense, rise per base pair, the number of residues per base pair, mean rotation per base pair, groove widths and groove depths are presented in the table 1.

Table 2.1 - Some physical properties of A, B and Z-DNA

Property	A-DNA	B-DNA	Z-DNA
Helix sense	Right	Right	Left
Rise/bp (Å)	2.56	3.4	3.8
Residues/bp (Å)	11	10	12
Mean rotation/bp	33.6°	35.9°	-30°
Major groove width (Å)	2.7	11.7	8.8
Major groove depth (Å)	13.5	8.8	3.7
Minor groove width (Å)	11	5.7	-
Minor groove depth (Å)	2.8	7.5	3.7

1.3. Drug-DNA Interaction

DNA has the tendencies of gene expression and protein formulation and the vital life processes involves transcription and replication of DNA. Transcription is the process of transfer of information from the DNA into mRNA and replication of DNA is the process of formation of a copy of DNA. These processes are stimulated by the signals in the form of regulatory proteins[32]. These processes can also be controlled by some drug molecules to cure severe diseases[40].

The study of interactions of drug and DNA is of great importance as it indicates the drug action. Drug can interact with the DNA and bind itself in the vicinity of the DNA thereby altering the processes the DNA[41]. The interaction of the drug with the DNA can also result in the inhibition of the disease by inhibiting the replication process of the diseased DNA. Drugs bind itself to the DNA and thus interfere with DNA transcription factors or DNA topoisomerases[22]. Hence, the design of drugs requires the study of the interaction of drug with the DNA.

1.3.1. Forces Involved in Drug-DNA Interactions

The interaction of the drug with the DNA is responsible for the drug action. The drug interacts with the DNA through several interaction forces. These forces are responsible for the stability of the complex of drug-DNA. These interaction forces are Van der Waals interactions, Electrostatic Interactions, Hydrogen Bonding, Hydrophobic forces, and dispersive forces[42–47]. Each of them is discussed below.

1.3.1.1. Van der Waals Interactions

When two molecules that are neutral otherwise are brought under the field of the each other, the electrons in the outer shell repel each other and hence they experience a force. Such an interaction creates a transient dipole in one of the atoms or molecules which induces the dipole formation in the other. This type of interaction is known as Van der Waal's interaction. Van der Waal's interaction contributes significantly in achieving the stability in some of the drug-DNA complexes.

1.3.1.2. Hydrogen Bonding

When a hydrogen atom is connected with the electronegative elements like Oxygen and Nitrogen, there is a formation of partial positive charge on the Hydrogen atom and partial negative charge on the electronegative atom. The formation of these partial charges is responsible for the hydrogen bonding. This contributes largely to the stability in the drug-DNA complex[47].

In the drug-DNA complex, the hydrogen bonds are formed between the nucleobases and the drug atoms.

1.3.1.3. Electrostatic Forces

Such types of interactions are present between the ionized phosphate group of the DNA and the positively charged groups of the drug molecule. These interactions are long range interactions[42]. These electrostatic interactions have significant strength in the absence of water molecule between the interacting species.

1.3.1.4. Hydrophobic Forces

The hydrophobic forces are originated due to the polar nature of the water molecules. The polar molecules tend to dissolve in the water, but the nonpolar molecules tend to accumulate to reduce the surface area exposed to the water. The water molecules at the interaction surface between the drug and the DNA causes these forces. When a molecule is surrounded by water molecules it tends to create a sharp cusp like curved surface of ordered molecules. As there is accumulation of water molecules at the interface it gets disordered thereby increasing the entropy of the system. The molecules of water that are left at interface of drug-DNA interaction site tend to decrease the system entropy[48].

1.3.1.5. Dispersive Forces Due to Base Stacking

The molecules with zero dipole moment attract each other by transient dipole-induced dipole interactions. This decreases the dispersive forces by inverse of sixth power of the distance between the dipoles. Because they are very much sensitive to the thermal agitation of the molecules involved[49,50]. These dispersive forces play a very significant role in providing the firmness in the DNA by providing support in base pair stacking. Base pair stacking is an interesting feature of the DNA which allows the insertion of planar aromatic molecules between the base pairs. This insertion of planar molecule is termed as the intercalation[51,52].

1.4. Modes of Drug-DNA Interactions

When the drug interacts with the target structure it binds itself in the vicinity of the target structure. Similarly, in the case of DNA targeted drugs, the drugs binds itself to the DNA. There are two types of possible binding between the drug and the DNA one is covalent binding and the other one is non-covalent binding [53]. Non-covalent binding is further divided into

two categories namely the groove binding and the intercalation. Major and minor groove binding are the two types of the groove binding.

1.4.1. Covalent Binding

Covalent binding involves high binding strengths. Such type of binding occurs in two ways, one is inter or intra-strand cross linking and the other is DNA alkylation [54,55]. The drugs which bind through the covalent binding mode are often termed as DNA alkylating agents. DNA alkylation is the process of attachment of an alkyl group to the DNA and it leads to cell death[22]. Some of the common DNA alkylating agents are cis-Platin, Carmustine (BCNU), Cyclophosphamide, Dacarbazine, Temozolomide (TMZ) etc.

1.4.2. Non-Covalent Binding

Hydrogen bonding and base stacking interactions are the non-covalent interactions that determine the conformational structure of the biological molecules. The hydrogen bonding and the π-π stacking significantly contribute in attaining the stability in the structure of nucleic acid[56]. The non-covalent binding tends to produce conformational changes in the DNA and may tend to rupture the DNA[22]. There are three modes of covalent binding that are minor and major groove binding and intercalation.

1.4.2.1. Minor Groove Binding

Minor grooves have narrow widths as compared to the widths of the major grooves. Also, the minor grooves are deeper than the major grooves. The minor groove binders prefer AT-rich regions. The minor groove binders generally interact with the bases of the nucleic acid via hydrogen bonds with the O2 of Thymine (T) and N3 of the Adenine (A). The dominant factor for the higher selectivity of the minor groove binders towards AT-rich regions

is the higher electrostatic potential of AT-rich regions and also the structure of the width[57–59].

Some of the minor groove binders are Gunayl Bisulframidine, Propamidine, Berenil and Netropsin etc. The minor groove binders consist of aromatic rings.

1.4.2.2. Major Groove Binding

Major grooves of the DNA have wider widths as compared to that of minor grooves. Due to wide openings, DNA interacting proteins generally binds itself in the major grooves of the DNA. There are large number of hydrogen bond acceptor and donor species in the major grooves of the DNA[60–62]. The drug action of the major groove binders is achieved by the binding of the drug in the major groove and thereby blocking the access of the groove for the proteins that recognize same groove[22].

Chloroquine, Leinamycin, Azinomycin and Aflatoxins are some of the major groove binders.

1.4.2.3. Intercalation

Intercalation is the process of insertion or stacking of the drug molecule which has a planar heterocyclic structure between the adjacent nucleic acid base pairs. The intercalation of the drug distorts the pitch of the DNA[63]. The stability in the intercalated complex of the drug and the DNA achieves stability through van der Walls forces, hydrogen bonding, hydrophobic interactions, stacking or charge transfer forces and electrostatic force[64].

Intercalation results in significant conformational changes in the DNA viz. lengthening, stiffing and unwinding of the helix[65]. Acridine, Actinomycin D, Ethidium bromide, Quiacrine are some of the intercalating drugs.

References: -

1. Nicolaou KC. Advancing the drug discovery and development process. Angew Chemie. 2014;126(35):9280–92.
2. Gershell LJ, Atkins JH. A brief history of novel drug discovery technologies. Nat Rev Drug Discov. 2003;2(4):321–7.
3. Gawehn E, Hiss JA, Brown JB, Schneider G. Advancing drug discovery via GPU-based deep learning. Expert Opin Drug Discov. 2018;13(7):579–82.
4. Bahuguna A, Rawat DS. An overview of new antitubercular drugs, drug candidates, and their targets. Med Res Rev. 2020;40(1):263–92.
5. Seyhan AA, Carini C. Are innovation and new technologies in precision medicine paving a new era in patients centric care? J Transl Med. 2019;17(1):1–28.
6. David B, Wolfender J-L, Dias DA. The pharmaceutical industry and natural products: historical status and new trends. Phytochem Rev. 2015;14(2):299–315.
7. Baskin II, Winkler D, Tetko I V. A renaissance of neural networks in drug discovery. Expert Opin Drug Discov. 2016;11(8):785–95.
8. Bolis G, Di Pace L, Fabrocini F. A machine learning approach to computer-aided molecular design. J Comput Aided Mol Des. 1991;5(6):617–28.
9. Chen H, Engkvist O, Wang Y, Olivecrona M, Blaschke T. The rise of deep learning in drug discovery. Drug Discov Today. 2018;23(6):1241–50.
10. Kononenko I. Machine learning for medical diagnosis: history, state of the art and perspective. Artif Intell Med. 2001;23(1):89–109.
11. Korotcov A, Tkachenko V, Russo DP, Ekins S. Comparison of deep learning with multiple machine learning methods and metrics using diverse drug discovery data sets. Mol Pharm. 2017;14(12):4462–75.
12. Schlick T. Molecular dynamics: basics. In: Molecular modeling and simulation: an interdisciplinary guide. Springer; 2010. p. 425–61.
13. Yuan H, Ma Q, Ye L, Piao G. The traditional medicine and modern medicine from natural products. Molecules. 2016;21(5):559.

14. Ramakrishnan S, Suresh E, Riyasdeen A, Akbarsha MA, Palaniandavar M. DNA binding, prominent DNA cleavage and efficient anticancer activities of Tris(diimine)iron(ii) complexes. Dalt Trans. 2011;40(14):3524–36.

15. Heydari M, Moghadam ME, Tarlani AA, Farhangian H. DNA as a Target for Anticancer Phen-Imidazole Pd(II) Complexes. Appl Biochem Biotechnol. 2017;182(1):110–27.

16. Dwivedi A, Kumari A, Aarthy M, Singh SK, Ojha M, Jha S, et al. Spectroscopic and molecular docking studies for the binding and interaction aspects of curcumin-cysteine conjugate and rosmarinic acid with human telomeric G-quadruplex DNA. Int J Biol Macromol. 2021;182:1463–72.

17. Mestizo PD, Narváez DM, Pinzón-Ulloa JA, Di Bello DT, Franco-Ulloa S, Macías MA, et al. Novel complexes with ONNO tetradentate coumarin schiff-base donor ligands: x-ray structures, DFT calculations, molecular dynamics and potential anticarcinogenic activity. BioMetals. 2021;34(1):119–40.

18. Mulliri S, Laaksonen A, Spanu P, Farris R, Farci M, Mingoia F, et al. Spectroscopic and in silico studies on the interaction of substituted pyrazolo[1,2-a]benzo[1,2,3,4]tetrazine-3-one derivatives with c-myc g4-dna. Int J Mol Sci. 2021;22(11):6028.

19. Parveen S, Fatima K, Zehra S, Arjmand F. RNA-targeted Cu(II)-based potential antitumor drug entity: comprehensive structural, biological {DNA/RNA binding, cleavage, cytotoxicity} and computational studies. J Biomol Struct Dyn. 2021;39(16):6070–83.

20. Heidarpoor Saremi L, Dadashi Noshahr K, Ebrahimi A, Khalegian A, Abdi K, Lagzian M. Multi-stage screening to predict the specific anticancer activity of Ni(II) mixed-ligand complex on gastric cancer cells; biological activity, FTIR spectrum, DNA binding behavior and simulation studies. Spectrochim Acta - Part A Mol Biomol Spectrosc. 2021;251:119377.

21. Scoditti S, Dabbish E, Sicilia E. Is the cytotoxic activity of phenanthriplatin dependent on the specific size of the phenanthridine ligand π system? J Inorg Biochem. 2021;219:111447.

22. Mishra R, Kumar A, Chandra R, Kumar D. A review on theoretical studies of various types of Drug-DNA Interaction. Int J Sci Technol Soc. 2017;3(September):11–27.

23. Zheng S, Sahimi A, Shing KS, Sahimi M. Molecular Dynamics Study of Structure, Folding, and Aggregation of Poly-PR and Poly-GR Proteins. Biophys J. 2021;120(1):64–72.

24. Chong WL, Chupradit K, Chin SP, Khoo MM, Khor SM, Tayapiwatana C, et al. Protein-protein interactions: Insight from molecular dynamics simulations and nanoparticle tracking analysis. Molecules. 2021;26(18):5696.

25. Arnittali M, Rissanou AN, Harmandaris V. Structure of biomolecules through molecular dynamics simulations. Procedia Comput Sci. 2019;156:69–78.

26. Guvench O, MacKerell AD. Comparison of protein force fields for molecular dynamics simulations. In: Methods in Molecular Biology. Springer; 2008. p. 63–88.

27. Mackerell AD. Empirical force fields for biological macromolecules: Overview and issues. J Comput Chem. 2004 Oct;25(13):1584–604.

28. Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, et al. A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations. J Comput Chem. 2003;24(16):1999–2012.

29. Cornell WD. A second Generation force field for the simulation of proteins. [Internet]. Vol. 117, J. Am. Chem. Soc. 1995. p. 5179–97. Available from: https://www.docsity.com/pt/a-second-generation-force-field-for-the-simulation-of/4719622/

30. Kollman PA. Advances and continuing challenges in achieving realistic and predictive simulations of the properties of organic and biological molecules. Acc Chem Res. 1996;29(10):461–9.

31. Wang J, Cieplak P, Kollman PA. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? J Comput Chem. 2000;21(12):1049–74.

32. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. The structure and function of DNA. In: Molecular Biology of the Cell 4th edition. Garland Science; 2002.

33. Brown T. ChemDraw. Sci Teach. 2014;81(2):67.

34. Dahm R. Friedrich Miescher and the discovery of DNA. Dev Biol. 2005;278(2):274–88.

35. Avery OT, MacLeod CM, McCarty M. Studies on the chemical nature of the substance inducing transformation of pneumococcal types: induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. J Exp Med. 1944;79(2):137–58.

36. Chargaff E. Some recent studies on the composition and structure of nucleic acids. J Cell Comp Physiol. 1951;38(S1):41–59.

37. Watson JD, Crick FHC. Molecular structure of nucleic acids: A Structure for deoxyribose nucleic acid. 50 Years DNA. 2016;171(4356):83–4.

38. Watson JD, Crick FHC. Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. Nature. 1953;171(4356):737–8.

39. Dickerson RE, Drew HR, Conner BN, Wing RM, Fratini A V, Kopka ML. The anatomy of a-, b-, and z-dna. Science (80-). 1982;216(4545):475–85.

40. Chaires JB. Energetics of drug–DNA interactions. Biopolym Orig Res Biomol. 1997;44(3):201–15.

41. Chaires JB. Drug—DNA interactions. Curr Opin Struct Biol. 1998;8(3):314–20.

42. Mikheikin AL, Zhuze AL, Zasedatelev AS. Molecular Modelling of Ligand—DNA Minor Groove Binding: Role of Ligand—Water Interactions. J Biomol Struct Dyn. 2001;19(1):175–8.

43. Singh SB, Wemmer DE, Kollman PA. Relative binding affinities of distamycin and its analog to d (CGCAAGTTGGC). d (GCCAACTTGCG): comparison of simulation results with experiment. Proc Natl Acad Sci. 1994;91(16):7673–7.

44. Misra VK, Honig B. On the magnitude of the electrostatic contribution to ligand-DNA interactions. Proc Natl Acad Sci. 1995;92(10):4691–5.

45. Sharp KA, Friedman RA, Misra V, Hecht J, Honig B. Salt effects on polyelectrolyte–ligand binding: Comparison of Poisson–Boltzmann, and limiting law/counterion binding models. Biopolym Orig Res Biomol. 1995;36(2):245–62.

46. Seibel GL, Singh UC, Kollman PA. A molecular dynamics simulation of double-helical B-DNA including counterions and water. Proc Natl Acad Sci. 1985;82(19):6537–40.

47. Shaikh SA, Ahmed SR, Jayaram B. A molecular thermodynamic view of DNA–drug interactions: a case study of 25 minor-groove binders. Arch Biochem Biophys. 2004;429(1):81–99.

48. Dashnau JL, Sharp KA, Vanderkooi JM. Carbohydrate intramolecular hydrogen bonding cooperativity and its effect on water structure. J Phys Chem B. 2005;109(50):24152–9.

49. Mak CH. Unraveling base stacking driving forces in DNA. J Phys Chem B. 2016;120(26):6010–20.

50. Gellman SH, Haque TS, Newcomb LF. New evidence that the hydrophobic effect and dispersion are not major driving forces for nucleotide base stacking. Biophys J. 1996;71(6):3523.

51. Yakovchuk P, Protozanova E, Frank-Kamenetskii MD. Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. Nucleic Acids Res. 2006;34(2):564–74.

52. Friedman RA, Honig B. A free energy analysis of nucleic acid base stacking in aqueous solution. Biophys J. 1995;69(4):1528–35.

53. Pindur U, Jansen M, Lemster T. Advances in DNA-Ligands with Groove Binding, Intercalating and/or Alkylating Activity: Chemistry, DNA-Binding and Biology. Curr Med Chem. 2005;12(24):2805–47.

54. Paul A, Bhattacharya S. Chemistry and biology of DNA-binding small molecules. Curr Sci. 2012;212–31.

55. Liu H-K, Sadler PJ. Metal complexes as DNA intercalators. Acc Chem Res. 2011;44(5):349–59.

56. Manalo MN, Pérez LM, LiWang A. Hydrogen-Bonding and π− π Base-Stacking Interactions Are Coupled in DNA, As Suggested by Calculated and Experimental Trans-Hbond Deuterium Isotope Shifts. J Am Chem Soc. 2007;129(37):11298–9.

57. Nelson SM, Ferguson LR, Denny WA. Mutation research/fundamental and molecular mechanisms of mutagenesis. Mutat Res-DNA Repair. 2007;623:24–40.

58. Khan GS, Shah A, Barker D. Chemistry of DNA minor groove binding agents. J Photochem Photobiol B Biol. 2012;115:105–18.

59. Privalov PL, Dragan AI, Crane-Robinson C, Breslauer KJ, Remeta DP, Minetti CASA. What drives proteins into the major or minor grooves of DNA? J Mol Biol. 2007;365(1):1–9.

60. Schleif R. DNA binding by proteins. Science (80-). 1988;241(4870):1182–7.

61. Thuong NT, Hélène C. Sequence-specific recognition and modification of double-helical DNA by oligonucleotides. Angew Chemie Int Ed English. 1993;32(5):666–90.

62. Jain AK, Bhattacharya S. Groove binding ligands for the interaction with parallel-stranded ps-duplex DNA and triplex DNA. Bioconjug Chem. 2010;21(8):1389–403.

63. Remers WA. chemistry of antitumor antibiotics. Wiley; 1979.

64. Williams LD, Egli M, Qi G, Bash P, van der Marel GA, van Boom JH, et al. Structure of nogalamycin bound to a DNA hexamer. Proc Natl Acad Sci. 1990;87(6):2225–9.

65. Pigram WJ, Fuller W, Hamilton LD. Stereochemistry of intercalation: interaction of daunomycin with DNA. Nat New Biol. 1972;235(53):17–9.

CHAPTER

2

COMPUTATIONAL METHODS

2. COMPUTATIONAL METHODS

2.1. Molecular Docking Simulation

As the name suggests, the molecular docking simulation predicts the possible binding modes of the ligand with respect to the target macro biomolecules. The docking programs generates several binding modes of the ligand and then rank them based on their binding affinities[1–5].

The three main features of the docking simulations are: -

a) The representation of the molecular system,

b) Search of the conformational space via some inbuilt search algorithm,

c) Ranking the several possible binding modes based on scoring function.

The molecular docking software predicts the binding poses of the ligand with respect to the target structure in such a way that the overall free energy of the complex formed is minimized. Molecular docking simulation involves large numbers of the degrees of freedom, six translational and rotational and certain conformational degrees of freedom. With such a large degree of freedom the number of the predicted binding modes will be larger and computationally expensive. To tackle this problem several search algorithms have been developed[6–10].

The treatment of two molecules as rigid bodies so that only six translational and rotational degrees of freedom are left is the most general search algorithm and it was used by Kuntz and coworkers in the docking program DOCK[11]. This algorithm neglected the conformational degrees of freedom.

Several other algorithms assume the target to be rigid and treat the ligand flexible. Genetic algorithms are also used as a search algorithm[12–14]. The idea is the same as the Darwin's theory of evolution. Each chromosome codes are for the geometrical conformation of the ligand and the orientation of the ligand within the domains of the binding site of the target. score of each docked structure within the site act as the fitness function used to select the individuals for the next iteration[14].

2.2. Molecular Dynamics Simulation

Molecular Dynamics (MD) Simulation is a tool to study the structure activity relationship of a molecular structure. It predicts the dynamics i.e. time evolution of the molecular structure. MD simulation has been widely used to study the protein-protein interaction, protein folding and unfolding, DNA-protein interactions etc.[15]. The use of MD simulation is not limited to the study of biomolecules, but it has been used in the field of material science as well. MD simulation becomes more important for the studies where the laboratory experiments are tough to be performed or rather the experiments are too expensive.

The key feature of the MD simulation is the representation of the molecular interaction potential, in terms of atomic coordinates and some intrinsic energy parameters. This representation of potential is termed as force field[16]. The atoms in the molecules are allowed to interact for a duration of time. The forces on individual atoms are calculated using the force field whereas the trajectory is calculated using the Newton's laws of motion. The forces on atoms are calculated for a step of time and the system is evolved for the time step. In this way the atomic coordinates of the atoms get evolved and the forces are further calculated for the evolved coordinates. The velocity and the position of the atoms at each time step is calculated using integration algorithms like Verlet algorithm, Velocity Verlet algorithm, Leap Frog and

corrected Velocity Verlet algorithms[17]. The equations for the evolution of the velocity and position after a time step for different algorithms are comprised in the following table 2.

Table 2.1 - Details of integration algorithms

S. No.	**Algorithm**	**Equation for position**	**Equation for Velocity**
1.	Verlet	$x(t+\Delta t) = 2x(t) - x(t-\Delta t) + \frac{f(t)}{m}(\Delta t)^2$	$v(t) = \frac{x(t-\Delta t) - x(t-\Delta t)}{2x(t)} + o(\Delta t)^2$
2.	Velocity Verlet	$x(t+\Delta t) = x(t) + v(t) + \frac{f(t)}{m}(\Delta t)^2$	$v(t+\Delta t) = v(t) + \frac{f(t+\Delta t) + f(t)}{2x(t)}(\Delta t)$
3.	Leap frog	$x(t+\Delta t) = x(t) - x(t-\Delta t) + \frac{f(t)}{m}(\Delta t)^2$	$v\left(t+\frac{\Delta t}{2}\right) = v\left(t-\frac{\Delta t}{2}\right) + (\Delta t)\frac{f(t)}{m}$
4.	Corrected velocity Verlet	$x(t+\Delta t) = x(t) + v(t).\Delta t + \frac{4f(t) - f(t-\Delta t)}{6m}(\Delta t)^2$	$v(t+\Delta t) = v(t) + \frac{2f(t+\Delta t) + 5f(t) - f(t+\Delta$ $6m$

2.2.1. Force Fields

The representation of the interaction potential among the atoms of the molecule in terms of the atomic coordinate is and parameters is termed as the 'force field'. There are two types of interactions present in the molecular systems, one is bonded interactions and the other is nonbonded interactions. Stretching of bond lengths, bending of bond angle and dihedral

rotation are the types of bonded interaction. Lennard-Jones interactions and Coulombic interactions are the types of nonbonded interactions. The interaction potential can be written as

$$E_{bonded} = \sum_{bonds} K_b(b - b_0)^2 + \sum_{angles} K_\theta(\theta - \theta_0)^2 + \sum_{dihedrals} K_\chi[1 + \cos(n\chi - \sigma)]$$

$$E_{nonbonded} = \sum_{\substack{nonbonded \\ pairs\ i,j}} \left(\varepsilon_{ij} \left[\left(\frac{R_{min,ij}}{r_{ij}} \right)^{12} - 2 * \left(\frac{R_{min,ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{r_{ij}} \right)$$

This representation of interaction potential is also known as class 1 type of potential function. This is the most general and simplest form of the potential function. Simple in the sense that some other representations also include the cross terms between the individual interactions to model phenomenon like the stretching of the bonds on bending of bond angle. The inclusion of cross terms enhances the performance of force field but on the cost of computational time.

The first term of the bonded interaction provides the contribution of the stretching of bond lengths about equilibrium bond length (interaction between two bonded atoms, 1-2 interaction) in the overall potential energy. Here K_b and b_0 are the parameters which describes the rigidity towards stretching of bond lengths and the equilibrium bond lengths respectively. The second term on the R.H.S. of bonded potential term is the potential energy due to the bending of bond angles (interaction between 3 bonded atoms, 1-3 interactions). Here K_θ and θ_0 are the parameters that represent the rigidity towards bending of bond angles and the equilibrium bond angle respectively. The last term of the bonded interaction potential models the rotation of dihedral angle (interaction between 4 bonded atoms, 1-4 interaction). Here χ is

the dihedral angle and $\boldsymbol{K_\chi}$, $\boldsymbol{n}$ & $\boldsymbol{\sigma}$ are the parameters representing barrier height, periodicity & phase respectively.

The nonbonded interactions are the interactions present between the nonbonded atoms. The atoms separated by three or more bonds also interact like the interaction between the nonbonded atoms. These are known as 1-4 nonbonded interactions. Generally, they occur with some scaling function. The two types of nonbonded interactions are Lennard-Jones interaction and the Coulombic interaction.

Lennard-Jones interaction models the Pauli's exclusion repulsion at shorter interatomic distances and diffusion attraction at larger distances. The general form of the L-J potential is given by

$$\sum_{\substack{\boldsymbol{nonbonded} \\ \boldsymbol{pairs\ i,j}}} \boldsymbol{\varepsilon_{ij}} \left[\left(\frac{\boldsymbol{R_{min,ij}}}{\boldsymbol{r_{ij}}} \right)^{\boldsymbol{12}} - \boldsymbol{2} * \left(\frac{\boldsymbol{R_{min,ij}}}{\boldsymbol{r_{ij}}} \right)^{\boldsymbol{6}} \right]$$

Here $\boldsymbol{r_{ij}}$ is the separation between two atoms and $\boldsymbol{\varepsilon_{ij}}$, $\boldsymbol{R_{min,ij}}$ are the well depth and separation between atoms at minimum energy.

The Coulombic interactions model the electrostatic interactions between the nonbonded charged atoms. The charges used here are partial charges. The potential function for the Coulombic interaction is given by

$$\sum_{\substack{\boldsymbol{nonbonded} \\ \boldsymbol{pairs\ i,j}}} \frac{\boldsymbol{q_i q_j}}{\boldsymbol{r_{ij}}}$$

Here $\boldsymbol{q_i}$ and $\boldsymbol{q_j}$ are the partial charges on the atoms and $\boldsymbol{r_{ij}}$ is the separation between the atoms.

2.2.2. Popular Force Fields

Some of most generally force fields for the simulation of the biomolecules are AMBER, CHARMM, OPLS-AA and GROMOS.

Assisted **M**odel **B**uilding with **E**nergy **R**efinement (**AMBER**) was developed by Weiner et al. in late 80s. It incorporated the class 1 type of potential function to model the molecular interactions. In the early stages of the development of AMBER force field, the hydrogen atoms present in the molecule were not considered as an interacting site because it reduced the computational cost. But later on, with the evolution of the computer power, the hydrogen atoms connected with the electronegative atoms like oxygen, nitrogen were considered as an individual interaction site, but all the other hydrogen atoms were combined with the atoms with which they were connected to form extended atom.

Chemistry at **HAR**ward using **M**olecular **M**echanics (**CHARMM**) was developed by Karplus et al. Similar to the AMBER force field, CHRAMM force field also incorporated the extended atom approach for the hydrogen atoms. This force field has been used for the simulation of nucleic acids, lipids, proteins and carbohydrates.

Optimized **P**otential for **L**iquid **S**imulation (**OPLS**) was developed by Jorgensen et al. In the early models of OPLS, the hydrogen atoms bonded to the aliphatic carbons were combined with the carbons to form extended atoms but all other hydrogen atoms were treated as an individual interaction site. These models of OPLS were termed as OPLS-UA (United Atom). But recent models treated all the hydrogen atoms as in individual interaction site and hence termed as OPLS-AA (All Atom).

GROningen **MO**lecular **S**imulation (**GROMOS**) was developed by University of Groningen and ETH of Zurich collectively. It has been used to predict the dynamical motions of molecules and bulk liquids. GROMOS is also used for the simulation of the biomolecules.

2.2.3. Steps in Molecular Dynamics Simulation

The primary requirement of the MD simulation is an initial set of coordinates for the system of interest. These primary structures can be the structure downloaded from the protein data bank (PDB)[18] for the cases of DNA and proteins. These structures are minimized energetically to obtain a structure with lowest possible energy state, which will correspond to the structure near 0 K temperature[19].

Atoms are then assigned some initial velocities corresponding to Gaussian distribution in order to raise the temperature to the desired level[19]. Then the Newton's equations are utilized to evolve the system with time. After assigning the velocities to the atoms the system maybe in a vibrational kind of state where the kinetic energy and potential energy of the molecular system are not regular but the sum of them is a constant. This state may not be an equilibrium state and hence the molecule may transform into the state with lower temperatures[20]. To overcome this we equilibrate the system so that the kinetic energy and potential energy gets well distributed and the velocities are rescaled for the values at desired temperature. This process is repeated till the fluctuations in the energy and temperature is sufficiently small for a long time. The production MD run should start only after equilibrium is achieved.

References: -

1. Abagyan R, Totrov M. High-throughput docking for lead generation. Curr Opin Chem Biol. 2001;5(4):375–82.
2. Kuntz ID. Structure-based strategies for drug design and discovery. Science (80-). 1992;257(5073):1078–82.
3. Lengauer T, Rarey M. Computational methods for biomolecular docking. Curr Opin Struct Biol. 1996;6(3):402–6.
4. Gane PJ, Dean PM. Recent developments in rational structure-based drug design. Curr Opin Struct Biol. 2000;10:401–4.
5. Joseph-McCarthy D. Computational approaches to structure-based ligand design. Pharmacol Ther. 1999;84(2):179–91.
6. Schneider G, Böhm H-J. Virtual screening and fast automated docking methods. Drug Discov Today. 2002;7:64–70.
7. Waszkowycz B. Structure-based approaches to drug design and virtual screening. Curr Opin Drug Discov Devel. 2002;5(3):407–13.
8. Toledo-Sherman LM, Chen D. High-throughput virtual screening for drug discovery in parallel. Curr Opin Drug Discov Devel. 2002;5(3):414–21.
9. Stahl M, Rarey M. Detailed analysis of scoring functions for virtual screening. J Med Chem. 2001;44(7):1035–42.
10. Doman TN, McGovern SL, Witherbee BJ, Kasten TP, Kurumbail R, Stallings WC, et al. Molecular docking and high-throughput screening for novel inhibitors of protein tyrosine phosphatase-1B. J Med Chem. 2002;45(11):2213–21.
11. Kuntz ID, Blaney JM, Oatley SJ, Langridge R, Ferrin TE. A geometric approach to macromolecule-ligand interactions. J Mol Biol. 1982;161(2):269–88.
12. Judson RS, Jaeger EP, Treasurywala AM. A genetic algorithm based method for docking flexible molecules. J Mol Struct THEOCHEM. 1994;308:191–206.
13. Jones G, Willett P, Glen RC. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. J Mol Biol. 1995;245(1):43–53.

14. Oshiro CM, Kuntz ID, Dixon JS. Flexible ligand docking using a genetic algorithm. J Comput Aided Mol Des. 1995;9(2):113–30.

15. Guvench O, MacKerell AD. Comparison of protein force fields for molecular dynamics simulations. In: Methods in Molecular Biology. Springer; 2008. p. 63–88.

16. Orozco M, Luque FJ. Theoretical methods for the description of the solvent effect in biomolecular systems. Chem Rev. 2000;100(11):4187–226.

17. Hill DRC. Theory of Modelling and Simulation: Integrating Discrete Event and Continuous Complex Dynamic Systems: by BP Zeigler, H. Praehofer, TG Kim, Academic Press, San Diego, CA, 2000. Wiley Online Library; 2002.

18. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. Nucleic Acids Res. 2000;28(1):235–42.

19. Hinchliffe A. Molecular modelling for beginners. John Wiley & Sons; 2003.

20. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. J Comput Chem. 1983;4(2):187–217.

CHAPTER

3

ASSESSMENT OF AVAILABLE AMBER FORCE FIELDS TO MODEL DRUG-DNA INTERACTION

3. Assessment of Available Amber Force Fields to Model Drug-DNA Interaction

3.1. Introduction

Molecular Dynamics Simulation has been widely used for the study of dynamical features of the molecular structures. MD simulations has been used to study protein-protein interactions, DNA-protein interaction, protein folding, unfolding etc[1–6]. MD simulation has also been used to study the interactions of the DNA targeted drugs like anticancer and antitumor drugs[7–13]. The success of molecular dynamics simulation is based upon the idea of representing the potential of the molecular structure in terms of its atomic coordinates and some additional energy parameters. This representation of potential energy of the system is termed as the force field[1].

The most generally used biomolecular force fields incorporates class 1 type of potential function which has been discussed in brief in the second chapter. The most general form of the class 1 type of potential function is the following-

$$E_{bonded} = \sum_{bonds} K_b(b - b_0)^2 + \sum_{angles} K_\theta(\theta - \theta_0)^2 + \sum_{dihedrals} K_\chi[1 + \cos(n\chi - \sigma)]$$

$$E_{nonbonded} = \sum_{\substack{nonbonded \\ pairs\ i,j}} \left(\varepsilon_{ij} \left[\left(\frac{R_{min,ij}}{r_{ij}} \right)^{12} - 2 * \left(\frac{R_{min,ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{r_{ij}} \right)$$

As discussed earlier that the parameterization of the force field may vary among different force fields. So, the performance may also vary from one force field to other[1,14].

In several literatures, AMBER force field has been reported to perform better than the other force field for the simulation of the DNA. The results obtained using AMBER force field were in good agreement with the experimental results for the simulation of the DNA[15–20]. So, we will assess the performance of all the versions of AMBER force fields, available with the framework of GROMACS software[21] suite, for the simulation of the drug-DNA complex. We will select a minor groove binding ligand from the literature. We will then optimize the structure of the ligand geometrically using GAUSSIAN software. After that, the optimized structure of the ligand will be docked on to a DNA sequence using the AutoDock4 software. The complex of the ligand-DNA will then be put under the Molecular Dynamics simulation using all the versions of AMBER force fields available with GROMACS for 100 ns each. The parameters like radius of gyration, root mean square fluctuation (RMSF), root mean square deviation (RMSD), and number of hydrogen bonds will be studied for the analysis of the MD results.

3.2. Computational Methodology

3.2.1. System Selection and Preparation

We have selected a minor groove binding ligand (3,6-bis(4,5-dihydro-1H-imidazol-2-yl)-9H-carbazole) from the literature[22]. The structure of the target DNA i.e. B-DNA dodecamer has been downloaded from the PDB (PDB Id- 195D)[23,24].

The water molecules present in the downloaded structure of the DNA were removed using UCSF Chimera[25].

3.2.2. Geometry Optimization

Geometry optimization of a molecular structure restricts the structure to be in the lowest allowed potential energy state. The structure obtained after the geometry optimization have least steric hindrances, equilibrium bond lengths, equilibrium bond angels and dihedrals. The charges also get well distributed after the geometry optimization[26].

Gaussian09 software[27] has been utilized to perform the geometry optimization of the ligand structure using B3LYP functional of DFT incorporating 6-31G as the basis set.

3.2.3. Molecular Docking Simulation

Molecular docking simulation predicts the possible docking positions of the ligand towards the target structure and rank them based on the binding affinities. The best pose of the ligand is the pose with the least binding affinity having rank 1. AutoDock4 software[28] has been used to perform the molecular docking simulation of the ligand. Lamarckian Gene Algorithm will be used as the search algorithm.

3.2.4. Molecular Dynamics Simulation

Molecular Dynamics (MD) Simulation predicts the dynamics i.e. the time evolution of the molecular structure. MD simulation investigated the structure-activity relationship of a molecular structure[29]. MD simulation has been widely used to study the protein-protein interactions, DNA-protein interactions, protein folding and unfolding etc[30–32]. MD simulation uses force field to generate topologies of the molecular structure. These force field may perform differently for different force fields and the selection of force field for a system is purely based on the literature survey and past experiences. Literatures reported that AMBER force field performs well for the DNA system. But there is a lack of literature assessing the performance of force field for the simulation of drug-DNA complex[15–19]. So, we have

performed the MD simulation of drug-DNA complex using the seven versions of AMBER force fields AMBER03, AMBER94, AMBER96, AMBER99, AMBER99SB, AMBER99SB-ILDN, and AMBERGS available with the GROMACS software. These force fields are termed as ff1, ff2, ff3, ff4, ff5, ff6, and ff7 respectively for further discussion. We will perform the MD simulations using GROMACS 5.1.1 software.

3.3. Results and Discussion

3.3.1. Geometry Optimization

The geometry optimization was performed using Gaussian 09 software. The DFT method with B3LYP functional and 6-31G basis set was used to perform the geometry optimization. The optimized structure of the ligand is given in the following figure.

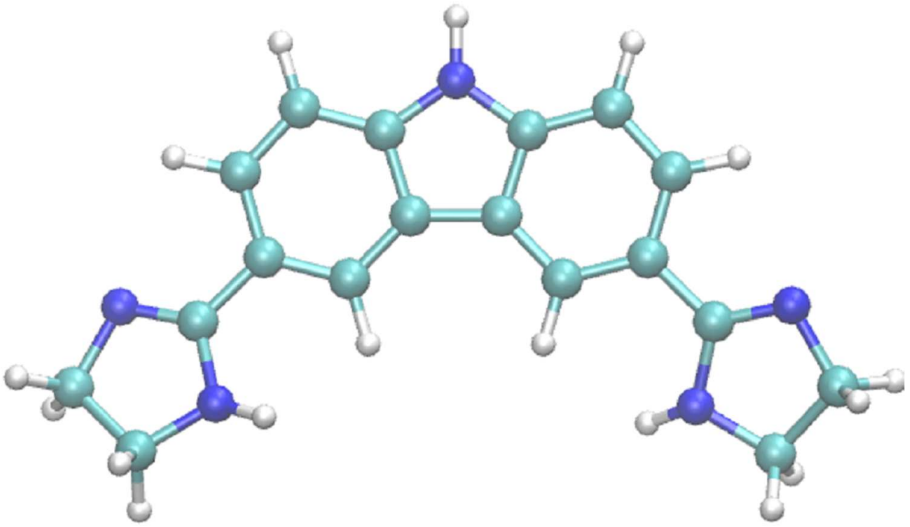

Figure 3.1 *The optimized structure of the ligand.*

3.3.2. Molecular Docking Simulation

The pose of the ligand with the rank 1 and hence least binding affinity was extracted and aligned with the structure of the DNA to obtain the best docked pose. The result of the Molecular Docking simulation clearly indicated that the ligand binds itself in the minor groove of the DNA. The figure shows the best docked pose with binding energy= -9.74 Kcal/mol . Figure shows the hydrogen bond donor and acceptor regions while figure shows the distribution of charges about the interaction site. The following table 3.1 gives the information about the hydrogen bonds between the ligand and the DNA structure

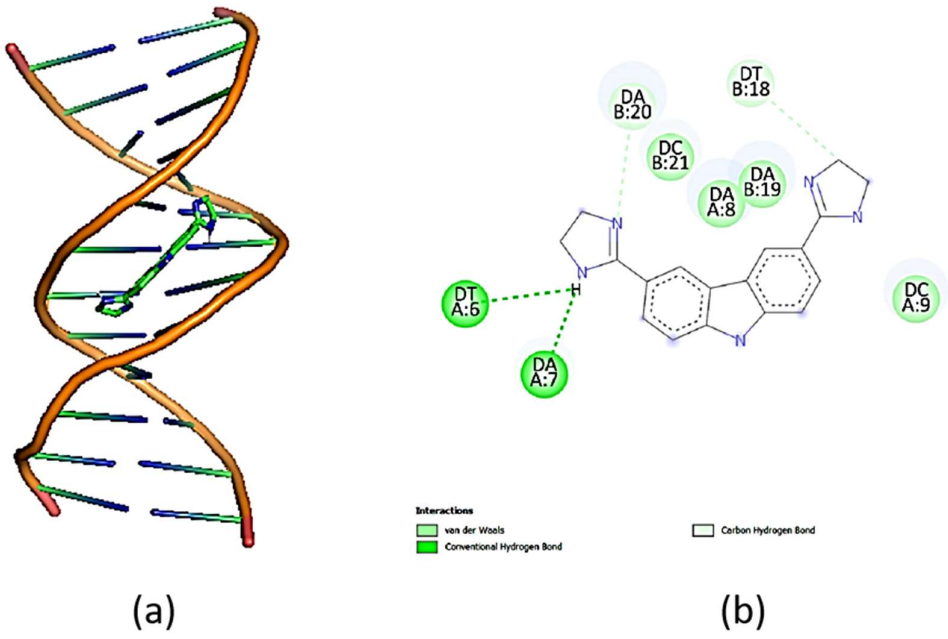

Figure 3.2 *(a) Docked Pose of the Ligand with 195D and (b) 2d representation of interaction between them*

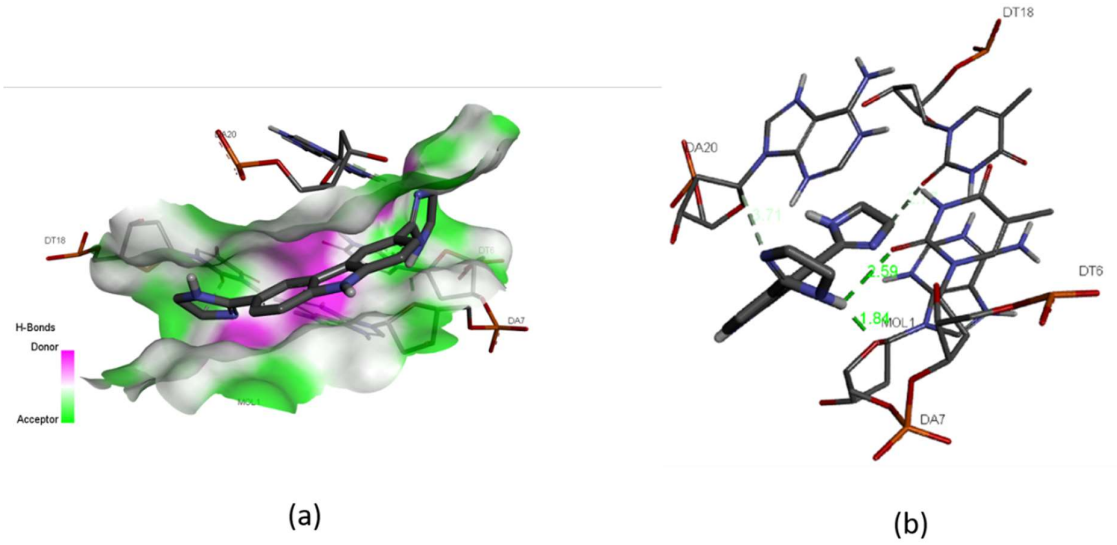

Figure 3.3 *(a) h-bond donor and acceptor site about the ligand (b) h-bond between ligand and DNA*

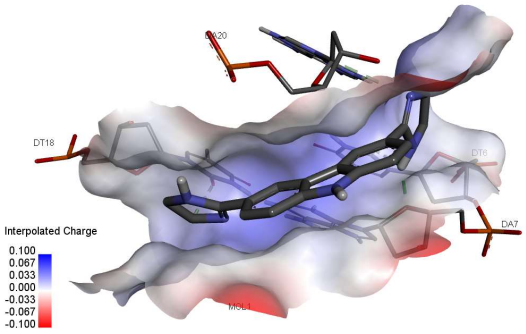

Figure 3.4*. Charge distribution about the interaction site*

Table 3.1 - Details of Hydrogen Bonds between DNA and Ligand

Sl. No.	**INTERACTING ATOMS**	**BOND LENGTH (Å)**
1.	LIG1:H10 – DT6:O2	2.594087
2.	LIG1:H10 – DA7:O4'	1.843068
3.	DA20:C1' – LIG1:N1	3.705932
4.	LIG1:C14 – DT18:O2	2.766987

3.3.3. Molecular Dynamics Simulation

The MD simulation was performed using GROMACS 5.1.1 and the versions of AMBER force fields were incorporated to generate topologies of the DNA. The time scale for the simulation was100ns each for each force field. The topologies of the DNA was generated using the versions of AMBER force field and that for ligand was generated using ANTECHAMBER module of the AMBER program through a python script "acpype.py"[33]. Solvation of the complex was done in conjunction with the TIP3P water model at 298K[34]. Sodium ions were the added to neutralize the system by replacing water molecules. Particle Mesh Ewald (PME)[35] handled the long-range interactions. Energy minimization was carried out implementing Steepest descent leap-frog integration method in 25000 steps. NVT ensemble equilibration was done incorporating Berendsen thermostat as 300 K for 50s[36]. Steepest descent leap-frog integration was again implemented to perform NPT ensemble equilibration at 1atm in 25000 steps[36]. LINCS algorithm[37] constrained the bonds with hydrogen atoms. The results of the MD simulations were analysed on the parameters of radius of gyration, root mean square fluctuation (RMSF), root mean square deviation (RMSD), and number of hydrogen bonds. The graphs were plotted using XMGRACE software[38].

3.3.3.1. Radius of Gyration

Radius of gyration predicts whether the molecular structure has achieved the equilibrium structure or not. It also predicts the compactness of the molecular structure. The lesser the value of the radius of gyration the compact is the structure. The formula for the radius of gyration is

$$R_g = \sqrt{\frac{\sum_i^N m_i r_i^2}{\sum_i^N m_i}}$$

here, m_i is the mass of the ith atom, r_i is the distance of the ith atom from the center of mass of the system[39]. The figure 3.5 shows the comparison of the different force fields whereas the figure 3.6 represents the graphs for individual force fields. The table 3.2 represents the statistical values of the radius of gyration. It is clear from the figure and table that AMBER99SB predicts most compact structure for the complex and the equilibration of the structure of complex was also observed.

Table 3.2 - Statistical values of the Graphs of Radius of Gyration

FORCE FIELD	**RANGE**	**MEAN**	**STANDARD DEVIATION**
FF1	1.45594 - 2.40820	2.08783	0.1809400
FF2	1.40796 - 1.89447	1.70169	0.1350980
FF3	1.41531 - 2.39154	1.95687	0.1630220
FF4	1.35889 - 3.88906	2.78326	0.4878010
FF5	**1.24221 - 1.74083**	**1.36525**	**0.0531752**
FF6	1.33761 - 2.35564	2.00462	0.1918220
FF7	1.44100 - 2.59832	2.15995	0.1691940

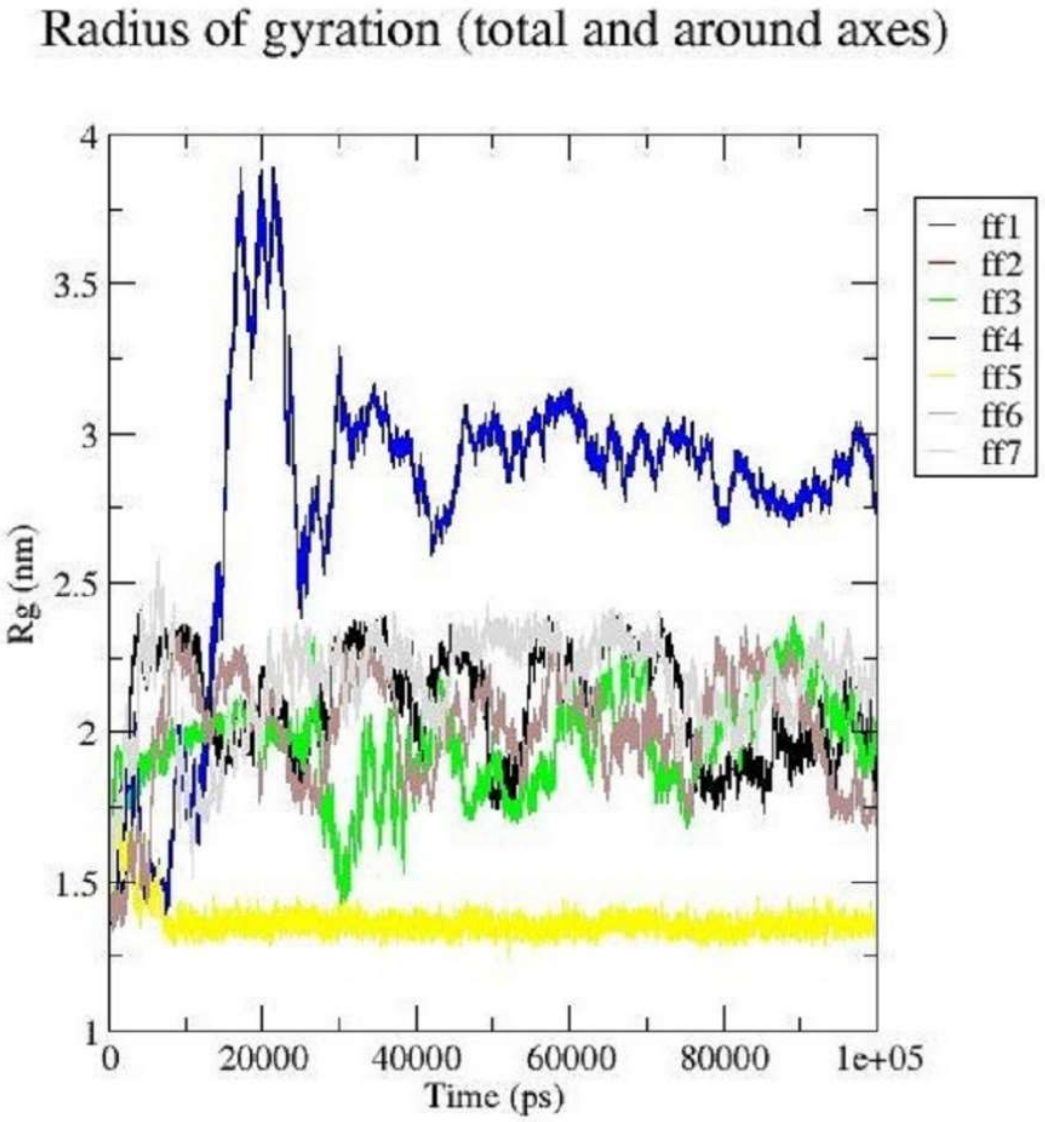

Figure 3.5 *Comparison of Radius of gyration with simulation time for AMBER force fields*

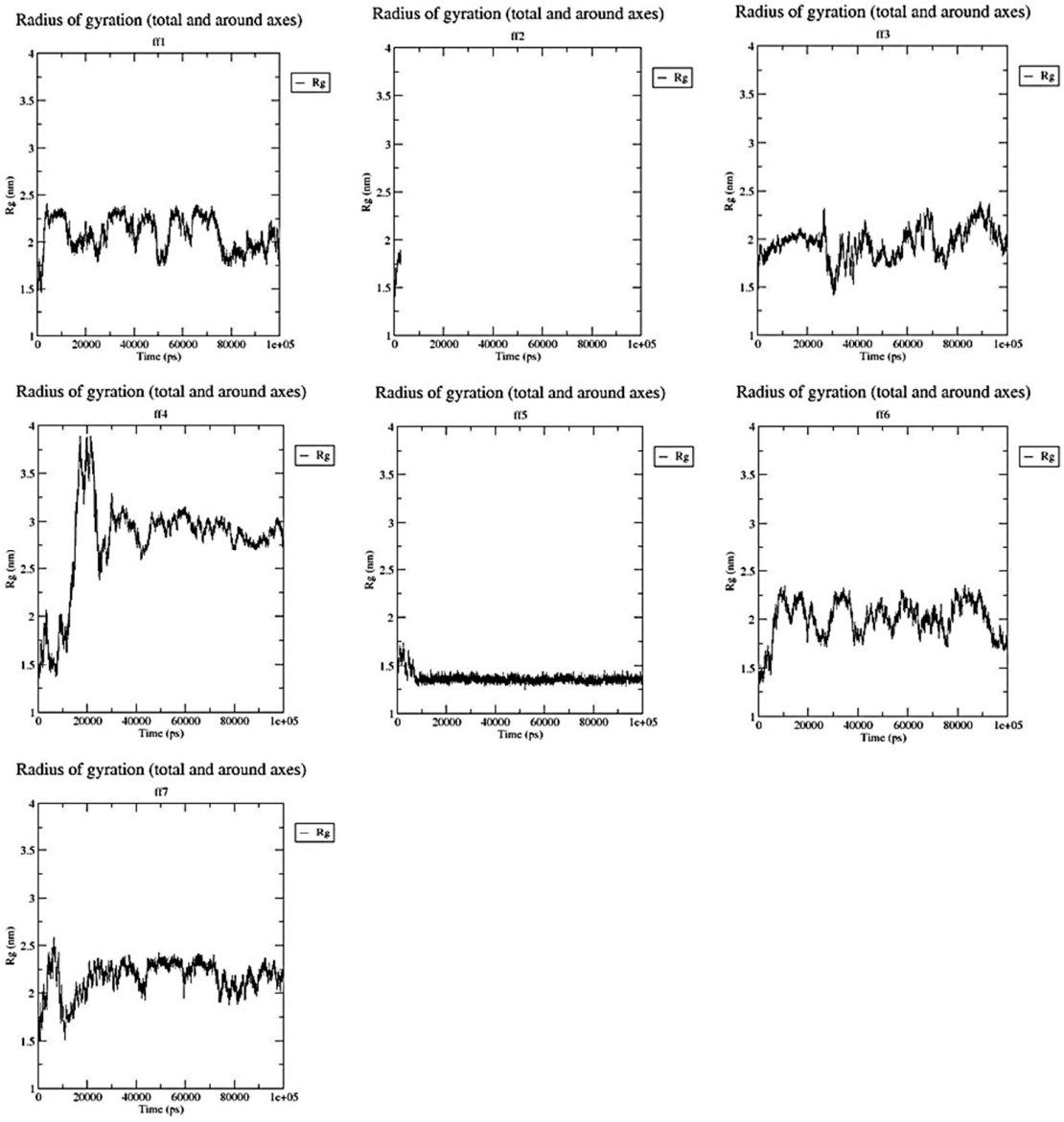

Figure 3.6 *Graphs of Radius of Gyration vs Simulation Time using AMBER force fields*

3.3.3.2. Root Mean Square Deviation

RMSD measured the deviation of the structure from its native or reference state. It indicates whether the equilibrium is achieved or not. The formula for calculating the RMSD is

$$\sigma_{RMSD} = \sqrt{\frac{\sum_i^N m_i \left(r_i - r_i^{ref}\right)^2}{\sum_i^N m_i}}$$

where $\boldsymbol{m_i}$ is the mass of i[th] atom, $\boldsymbol{r_i}$ and $\boldsymbol{r_i^{ref}}$ are the position of i[th] atom at any instance and in the native or reference structure, respectively[39,40]. The figure 3.7 shows the comparison of the different force fields whereas the figure 3.8 represents the graphs for individual force fields. It is evident from the figure that AMBER99SB outperforms all other versions.

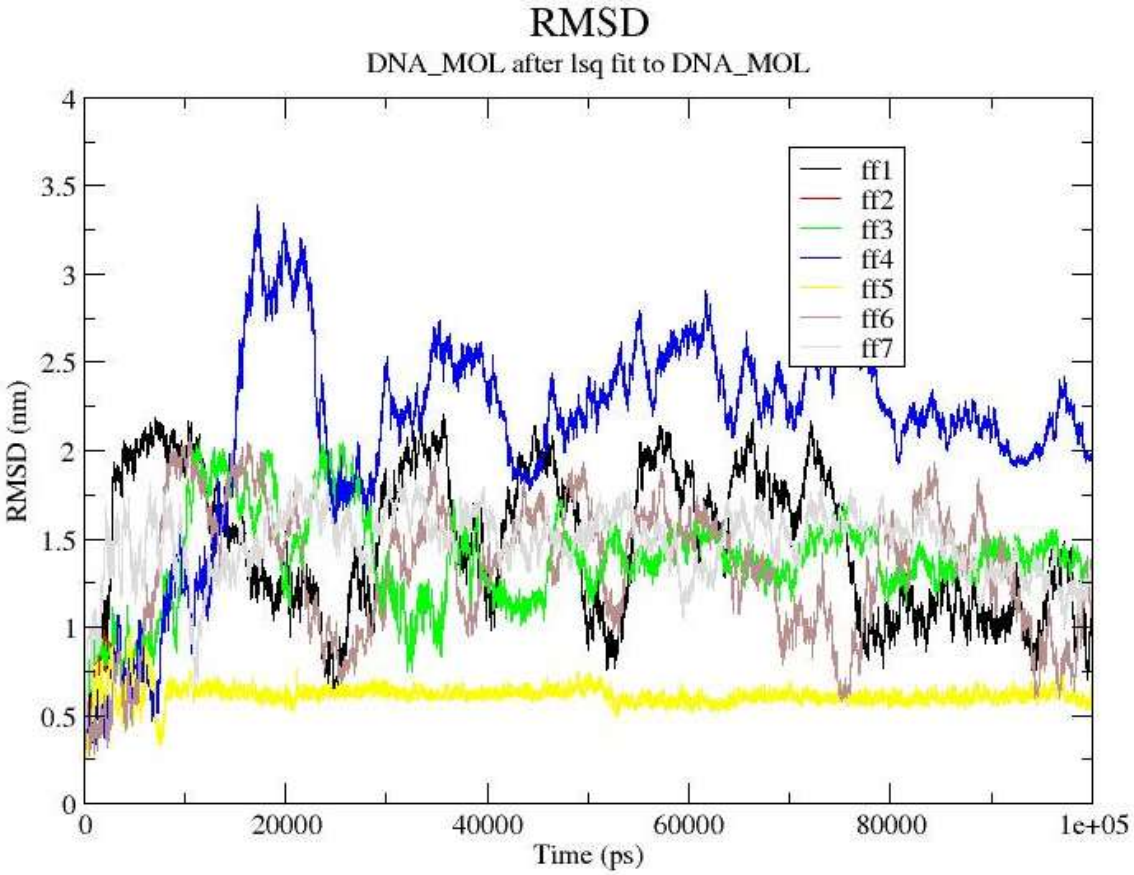

Figure 3.7 *Comparison of RMSD (complex with respect to complex) with simulation time*

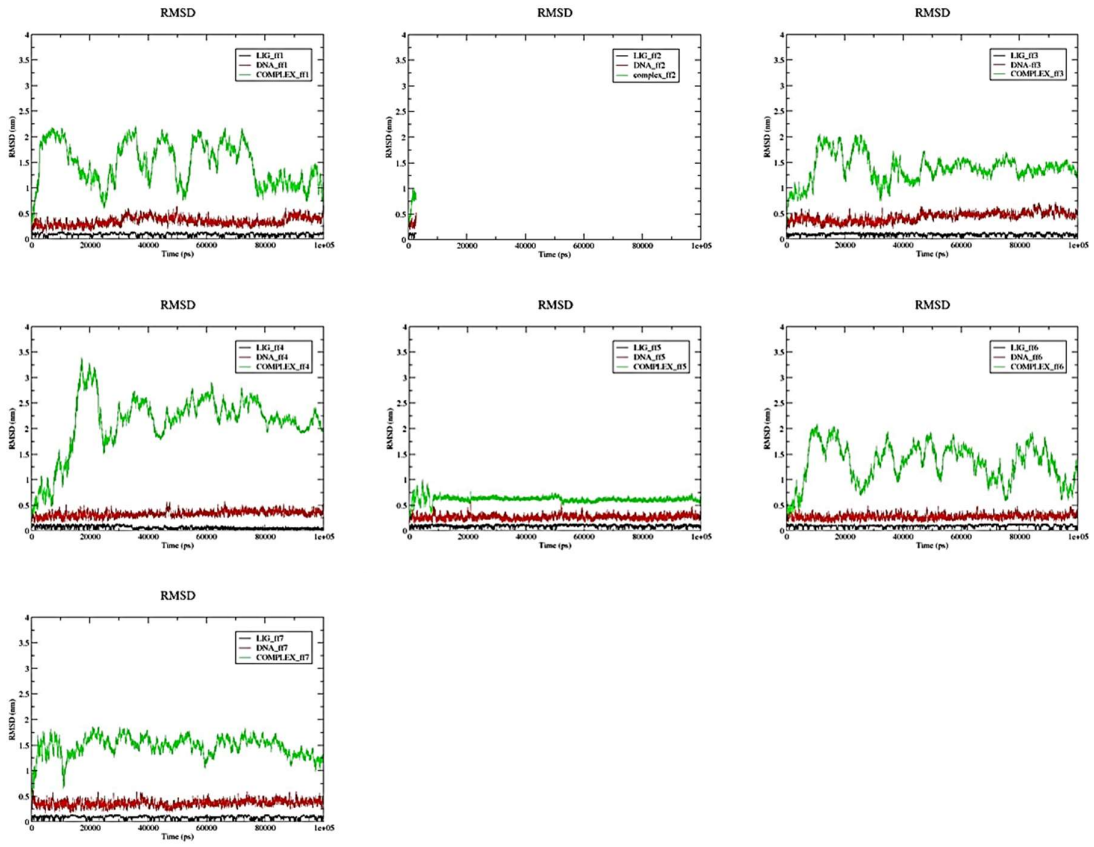

Figure 3.8 *Graphs of RMSD using the mentioned Force Fields*

3.3.3.3. Root Mean Square Fluctuation

RMSF measured the fluctuation of subunits of the molecular structure from the native state. The RMSF graphs showed that for all force fields, the fluctuation in the ligand atoms were greater than the fluctuations in DNA atoms. The figure 3.9 shows the comparison of the different force fields whereas the figure 3.10 represents the graphs for individual force fields. The table 3.3 represents the statistical values of the RMSF. It is clear from the graphs and table that AMBER99SB predicted the values with the least standard deviations and it predicted the least mean value of the RMSF. Hence it performs better than other versions.

Table 3.3 - Statistical Values of the graphs of RMSF

FORCE FIELD	RANGE OF RMSF	MEAN	STANDARD DEVIATION
FF1	0.1921-3.4384	0.830303	0.615188
FF2	0.1080-1.5350	0.276464	0.281719
FF3	0.2356-4.2439	0.783406	0.783653
FF4	0.2134-4.7748	0.896153	0.897097
FF5	**0.0923-1.0225**	**0.216315**	**0.182600**
FF6	0.1849-3.9137	0.845027	0.695093
FF7	0.2105-4.2037	0.863774	0.775285

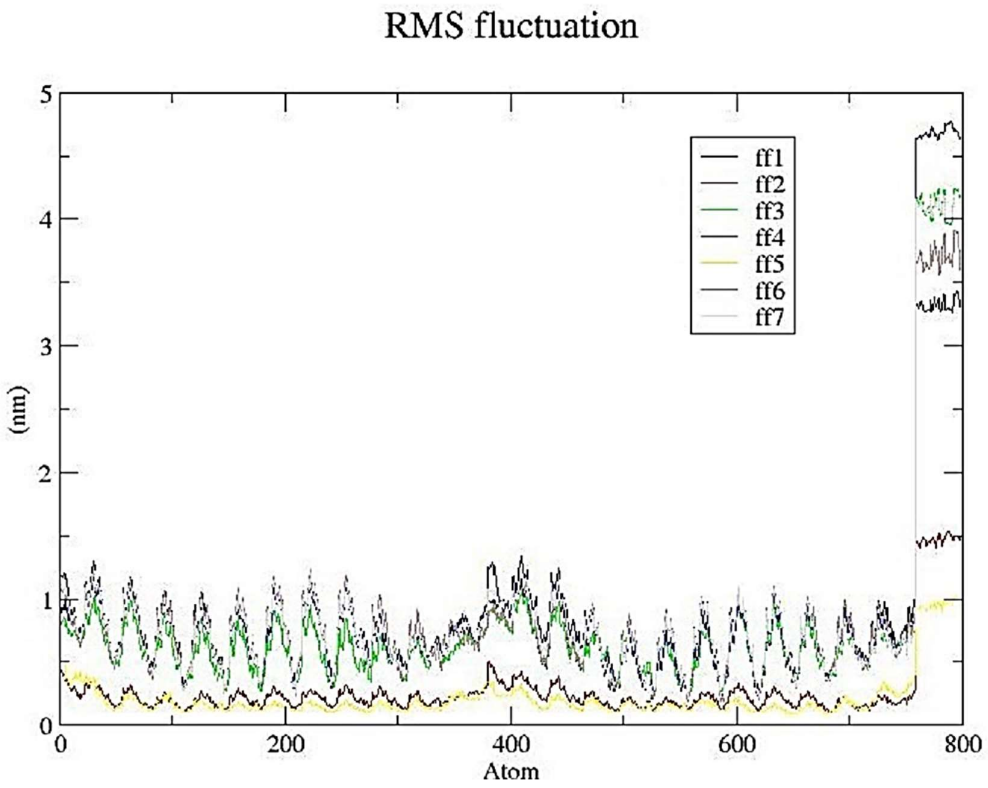

Figure 3.9 *Variation of RMSF with simulation time*

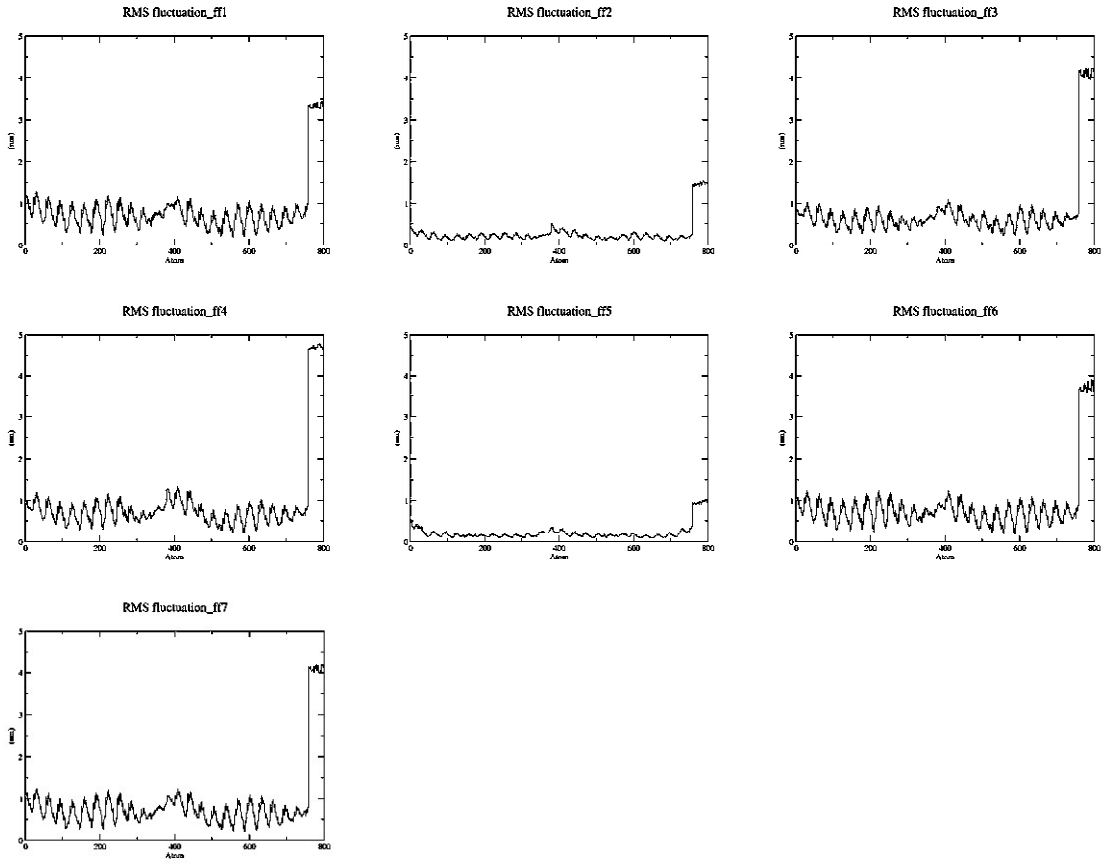

Figure 3.10 *Graphs of RMSF using the mentioned Force Fields*

3.3.3.4. Hydrogen Bonds

The number of hydrogen bonds between the DNA and the ligand was plotted against the simulation time using XMGRACE software. The greater the number of hydrogen bonds the better is the binding.

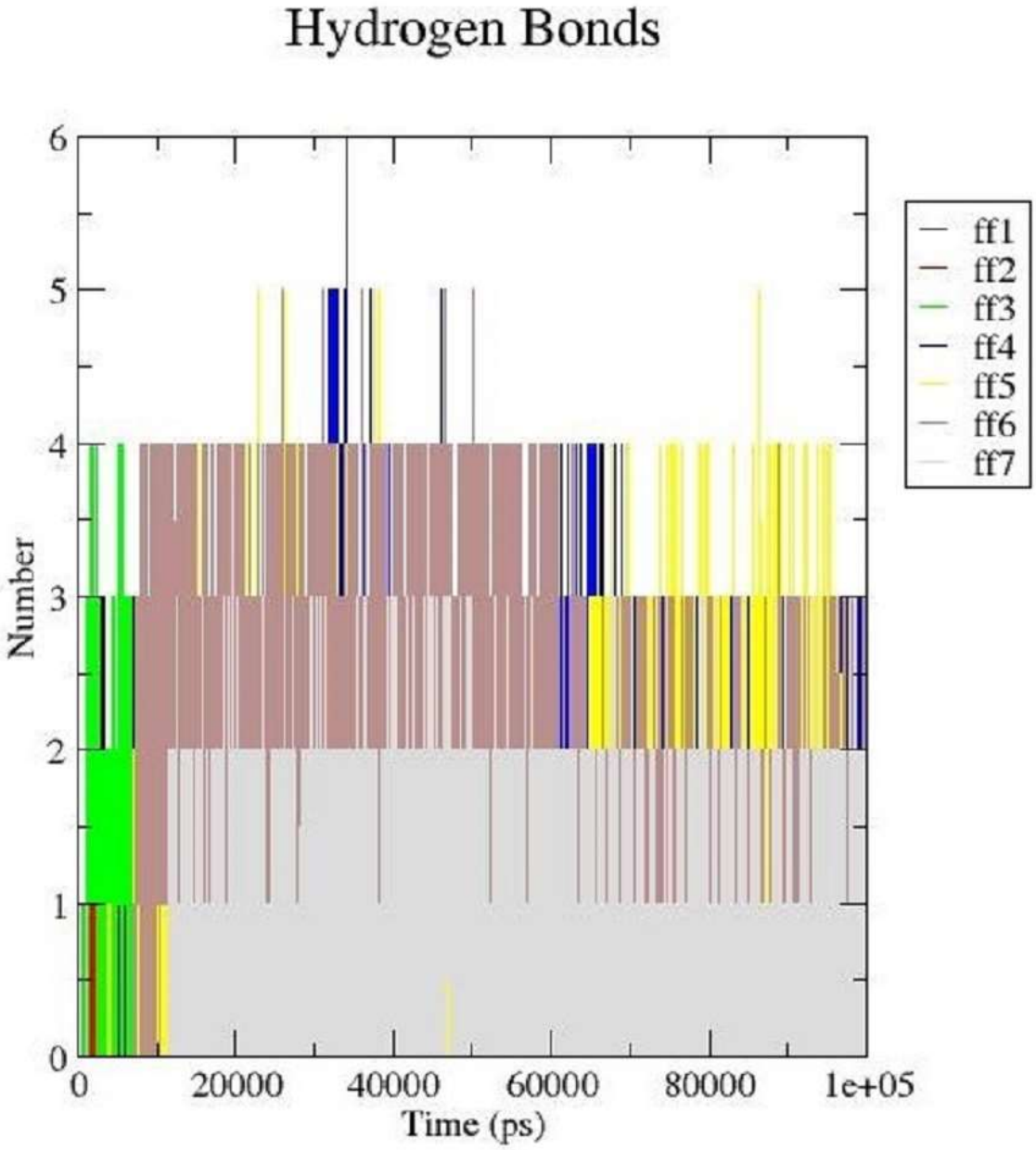

Figure 3.11 *Variation of number of hydrogen bonds with simulation time*

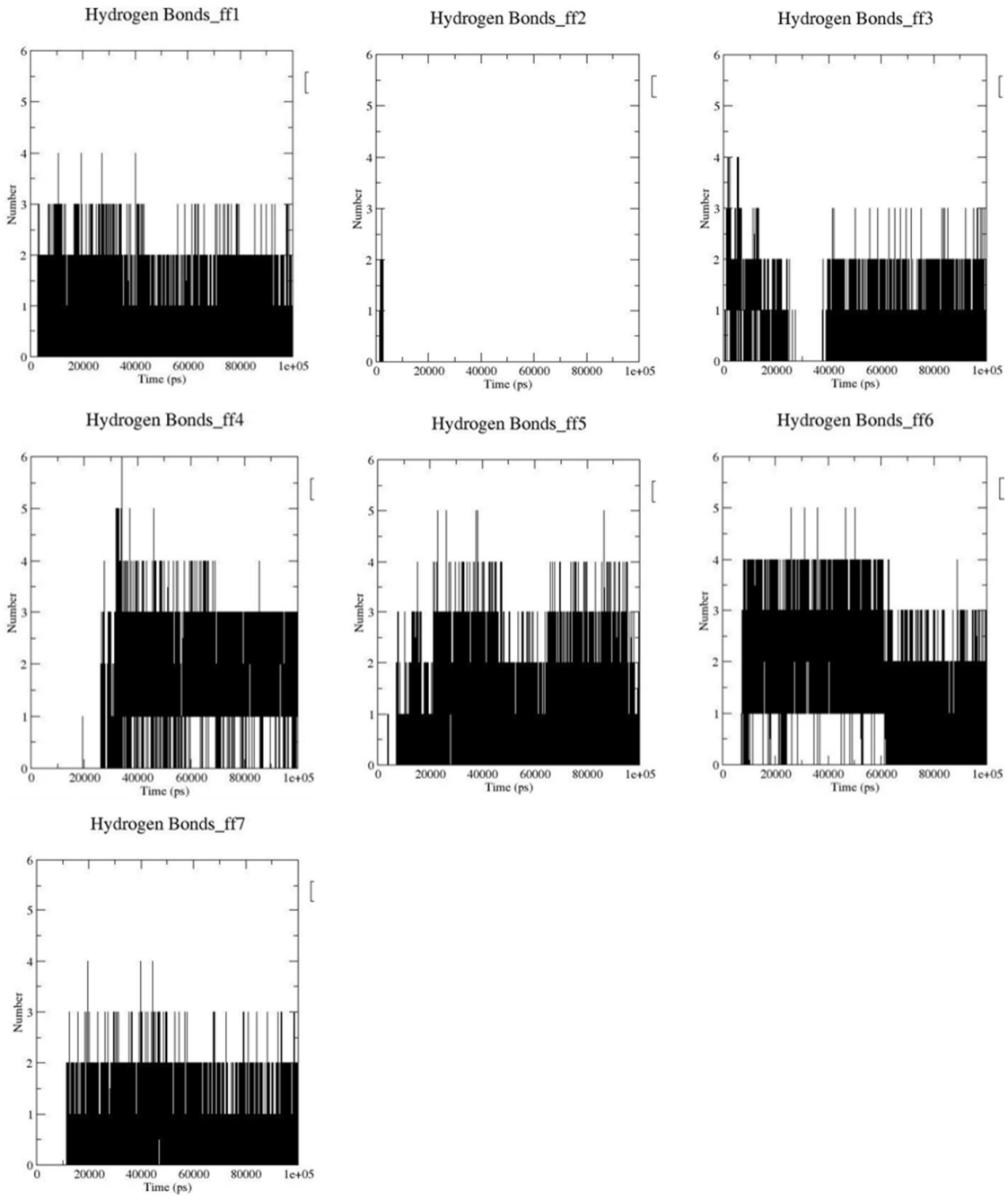

Figure 3.12 *Graphs of the hydrogen bond between ligand and 195D*

References: -

1. Guvench O, MacKerell AD. Comparison of protein force fields for molecular dynamics simulations. In: Methods in Molecular Biology. Springer; 2008. p. 63–88.
2. Scoditti S, Dabbish E, Sicilia E. Is the cytotoxic activity of phenanthriplatin dependent on the specific size of the phenanthridine ligand π system? J Inorg Biochem. 2021;219:111447.
3. Salehpour M, Azizian J. Exploration of interaction behavior between spiro[indene-2,2'-[1,3,5]oxathiazine]-1,3-diones and DNA with the help of DFT, molecular docking, and MD simulations. J Biomol Struct Dyn. 2021;1–21.
4. Stanzione F, Giangreco I, Cole JC. Use of molecular docking computational tools in drug discovery. In: Progress in Medicinal Chemistry. Elsevier; 2021. p. 273–343.
5. Jakhmola S, Jonniya NA, Sk MF, Rani A, Kar P, Jha HC. Identification of Potential Inhibitors against Epstein-Barr Virus Nuclear Antigen 1 (EBNA1): An Insight from Docking and Molecular Dynamic Simulations. ACS Chem Neurosci. 2021;12(16):3060–72.
6. Matias-Barrios VM, Radaeva M, Song Y, Alperstein Z, Lee AR, Schmitt V, et al. Discovery of New Catalytic Topoisomerase II Inhibitors for Anticancer Therapeutics. Front Oncol. 2021;10:3293.
7. Ramakrishnan S, Suresh E, Riyasdeen A, Akbarsha MA, Palaniandavar M. DNA binding, prominent DNA cleavage and efficient anticancer activities of Tris(diimine)iron(ii) complexes. Dalt Trans. 2011;40(14):3524–36.
8. Dwivedi A, Kumari A, Aarthy M, Singh SK, Ojha M, Jha S, et al. Spectroscopic and molecular docking studies for the binding and interaction aspects of curcumin-cysteine conjugate and rosmarinic acid with human telomeric G-quadruplex DNA. Int J Biol Macromol. 2021;182:1463–72.
9. Mestizo PD, Narváez DM, Pinzón-Ulloa JA, Di Bello DT, Franco-Ulloa S, Macías MA, et al. Novel complexes with ONNO tetradentate coumarin schiff-base donor ligands: x-ray structures, DFT calculations, molecular dynamics and potential anticarcinogenic activity. BioMetals. 2021;34(1):119–40.

10. Mulliri S, Laaksonen A, Spanu P, Farris R, Farci M, Mingoia F, et al. Spectroscopic and in silico studies on the interaction of substituted pyrazolo[1,2-a]benzo[1,2,3,4]tetrazine-3-one derivatives with c-myc g4-dna. Int J Mol Sci. 2021;22(11):6028.

11. Heydari M, Moghadam ME, Tarlani AA, Farhangian H. DNA as a Target for Anticancer Phen-Imidazole Pd(II) Complexes. Appl Biochem Biotechnol. 2017;182(1):110–27.

12. Rahiminezhad A, Moghadam ME, Divsalar A, Mesbah AW. How can the cisplatin analogs with different amine act on DNA during cancer treatment theoretically? J Mol Model. 2022;28(1):1–12.

13. Parveen S, Fatima K, Zehra S, Arjmand F. RNA-targeted Cu(II)-based potential antitumor drug entity: comprehensive structural, biological {DNA/RNA binding, cleavage, cytotoxicity} and computational studies. J Biomol Struct Dyn. 2021;39(16):6070–83.

14. Kaminski GA, Friesner RA, Tirado-Rives J, Jorgensen WL. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. J Phys Chem B. 2001;105(28):6474–87.

15. Ponder JW, Case DA. Force fields for protein simulations. Adv Protein Chem. 2003;66:27–85.

16. Beveridge DL, Cheatham TE, Mezei M. The ABCs of molecular dynamics simulations on B-DNA, circa 2012. J Biosci. 2012;37(3):379–97.

17. Dans PD, Ivani I, Hospital A, Portella G, González C, Orozco M. How accurate are accurate force-fields for B-DNA? Nucleic Acids Res. 2017;45(7):4217–30.

18. Galindo-Murillo R, Robertson JC, Zgarbová M, Šponer J, Otyepka M, Jurečka P, et al. Assessing the Current State of Amber Force Field Modifications for DNA. J Chem Theory Comput. 2016;12(8):4114–27.

19. Oweida TJ, Kim HS, Donald JM, Singh A, Yingling YG. Assessment of AMBER Force Fields for Simulations of ssDNA. J Chem Theory Comput. 2021;17(2):1208–17.

20. Misra M, Yadav AK. Assessment of Available AMBER Force Fields to Model DNA-

Ligand Interactions. 2022;

21. M.J. Abraham, D. van der Spoel, E. Lindahl, B. Hess and the G development team. GROMACS user manual version 5.0.2 [Internet]. GROMACS development team; 2014. Available from: www.gromacs.org

22. Tidwell RR, Boykin DW. Dicationic DNA Minor Groove Binders as Antimicrobial Agents. Small Mol DNA RNA Bind. 2004;414–60.

23. Balendiran K, Rao ST, Sekharudu CY, Zon G, Sundaralingam M. X-ray structures of the B-DNA dodecamer d(CGCGTTAACGCG) with an inverted central tetranucleotide and its netropsin complex. Acta Crystallogr - Sect D Biol Crystallogr. 1995;51(2):190–8.

24. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. Nucleic Acids Res. 2000;28(1):235–42.

25. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera - A visualization system for exploratory research and analysis. J Comput Chem. 2004;25(13):1605–12.

26. Dubey DK, Chaubey KA, Parveen A, Ojha PR. Comparative study of inhibition of drug potencies of c-Abl human kinase inhibitors: a computational and molecular docking study. J Biophys Struct Biol. 2010;2(4):47–54.

27. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenb DJ. Gaussian 09. Wallingford CT: Gaussian Inc.; 2009.

28. Morris GM, Huey R, Lindstrom W. MF Sanner RK Belew DS Goodsell and AJ Olson. J Comput Chem. 2009;16:2785–91.

29. Mackerell AD. Empirical force fields for biological macromolecules: Overview and issues. J Comput Chem. 2004 Oct;25(13):1584–604.

30. Zheng S, Sahimi A, Shing KS, Sahimi M. Molecular Dynamics Study of Structure, Folding, and Aggregation of Poly-PR and Poly-GR Proteins. Biophys J. 2021;120(1):64–72.

31. Kaur H, Shekhar N, Sharma S, Sarma P, Prakash A, Medhi B. Ivermectin as a potential drug for treatment of COVID-19: an in-sync review with clinical and computational attributes. Pharmacol Reports. 2021;73(3):736–49.

32. Chong WL, Chupradit K, Chin SP, Khoo MM, Khor SM, Tayapiwatana C, et al. Protein-protein interactions: Insight from molecular dynamics simulations and nanoparticle tracking analysis. Molecules. 2021;26(18):5696.

33. Sousa Da Silva AW, Vranken WF. ACPYPE - AnteChamber PYthon Parser interfacE. BMC Res Notes. 2012;5(1):367.

34. Mark P, Nilsson L. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. J Phys Chem A. 2001;105(43):9954–60.

35. Darden T, York D, Pedersen L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. J Chem Phys. 1993;98(12):10089–92.

36. Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys. 1977;23(3):327–41.

37. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS: A Linear Constraint Solver for molecular simulations. J Comput Chem. 1997;18(12):1463–72.

38. Turner PJ. XMGRACE, Version 5.1.19. Center For Coastal and Land-Margin Research, Oregon Graduate Institute of Science and Technology, Beaverton, Ore, USA; 2005.

39. Arnittali M, Rissanou AN, Harmandaris V. Structure of biomolecules through molecular dynamics simulations. Procedia Comput Sci. 2019;156:69–78.

40. Bolhuis PG. Sampling kinetic protein folding pathways using all-atom models. In: Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology Volume 1. Springer; 2006. p. 393–433.

CHAPTER

4

Conclusions and Scope for the

further research

4. Conclusions and Future Scope

This chapter presents the conclusions of the current dissertation work and the scope for the further research.

4.1. Conclusions

The main conclusions of the work are given below:-

1. The ligand was docked on the DNA using AutoDock4 software.
2. After that, the MD simulation of the Ligand-DNA complex was performed using versions of AMBER force fields pre-embedded in GROMACS software.
3. The results of the MD simulation were analyzed on the basis of parameters like Radius of Gyration, Root Mean Square Deviation, Root Mean Square Fluctuation and Number of Hydrogen Bonds between the DNA and the ligand.
4. Radius of Gyration measured the compactness of the structure and among all other versions of AMBER force fields AMBER99SB performs the best results by predicting the most compact structure.
5. Radius of Gyration also indicated the formation of equilibrium structure of the DNA-Ligand structure.
6. RMSD showed the fluctuation of the structure from its reference state and it was observed that AMBER99SB outperforms all the other versions by producing the lowest values of RMSD with the lowest standard deviation.
7. RMSF indicated the fluctuations of subunits of the molecular structure and it was observed that the ligand atoms were fluctuating more than the DNA atoms.

8. On the behalf of above points, it is observed that AMBER99SB outperforms all the other versions of AMBER force field for the simulation of DNA-Ligand complex.

4.2. Future Scope

The scope for the further research in the field of the present research work are listed below:-

1. One can test the reproducibility of the conclusions for a different DNA-Ligand system.
2. One can perform MD simulation of the current DNA-Ligand complex using other versions of AMBER force fields available with softwares other than GROMACS and compare it with our results of AMBER99SB to obtain the overall best version of AMBER force field.

List of Publications

Misra Manas, Yadav, A.K., 2022. Assessment of Available AMBER Force Fields to Model DNA-Ligand Interactions. *Biointerface Research in Applied Chemistry*, *13*(2) 156, https://doi.org/10.33263/BRIAC132.156 .

List of conference attended

Attended an international workshop on '*Tools and Techniques to Perform Molecular Modelling and Computer-Aided Drug Design (MMTT-2021)*' organised by the Department of Medical Chemistry, NIPER Guwahati between January 11-17, 2021. (Annexed)

ANNEXURE

National Institute of Pharmaceutical Education and Research Guwahati

(Dept. of Pharmaceuticals, Ministry of Chemicals and Fertilizers, Govt. of India)

Sila Katamur (Halugurisuk), Changsari, Kamrup, Guwahati 781101, Assam (India)

NIPER GUWAHATI

Dept. of Pharmaceuticals
Ministry of Chemicals and Fertilizers,
Govt. of India

International Workshop on

"Tools and Techniques to Perform Molecular Modelling and Computer-Aided Drug Design".

(MMTT-2021)

JANUARY 11-17, 2021

CERTIFICATE for ATTENDANCE

*This is to certify that Dr./Mr./Mrs./*MANAS MISRA

from BABASAHEB BHIMRAO AMBEDKAR UNIVERSITY LUCKNOW

have attended the seven days international workshop 'MMTT-2021' organized by the Department of Medicinal Chemistry, NIPER Guwahati between January 11-17, 2021.

We wish him/her a good luck for the future.

H K Srivastava
(Convener, MMTT-2021)

V G M Naidu
(Dean, NIPER-G)

U S N Murty
(Director, NIPER-G)