FORCE FIELD ANALYSIS OF GUANOSINE-50-DIPHOSPHATE AND ITS BINDING MECHANISM WITH THE 1GIT PROTEIN

By Bhupendra Maharjan

WORD COUNT 16241
TIME SUBMITTED 07-FEB-2024 09:20AM
PAPER ID 106730429

FORCE FIELD ANALYSIS OF GUANOSINE-5′-DIPHOSPHATE AND ITS BINDING MECHANISM WITH THE 1GIT PROTEIN

13

A Dissertation

Submitted to the Department of Physics, Amrit Campus, Lainchour, Tribhuvan University, in the Partial Fulfillment for the Requirement of Master's Degree of Science in Physics

By

Bhupendra Maharjan
Roll No. 3559/PHY-076
Reg. No. 5-1-33-271-2003

February, 2024

RECOMMENDATION

It is certified that **Mr. Bhupendra Maharjan** has carried out the dissertation entitled **"Force Field Analysis of Guanosine-5′-triphosphate and its Binding Mechanism with the 1GIT Protein"** under my guidance.

I recommend the dissertation is the partial fulfillment for the requirement of Master's Degree of Science in Physics.

.....

Asst. Prof. Dr. Manoj Kumar Chaudhary
Department of Physics, Amrit Campus
Lainchaur, Kathmandu
(Supervisor)

Date: .....

ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere gratitude to my supervisor Asst. Prof. Dr. Manoj Kumar Chaudhary for the continuous support of my dissertation study and research, for his patience, motivation, enthusiam and immense knowledge. His guidance helped me in all the time of research and writing of this dissertation. I have not only received a great list of suggestions from him but also learned the way of scientific writing.

There are many personalities whose constant support and encouragement made me to complete this project work successfully. My special acknowledgement goes to Prof. Dr. Rajendra Parajuli, Prof. Dr. Leela Pradhan Joshi and Assoc. Prof. Pitambar Shrestha for their knowledge, valuable suggestions and support during this work.

I would like to extend my acknowledgement to my friends,Om Shree Rijal, Deepak Oli and Puskar Chanda, and thank them so much for their coordination and cooperation.

Last but not the least, I would like to thank my family, my parents for their unconditional love and support.

Bhupendra Maharjan

EVALUATION

We certify that we have read this dissertation and in our opinion it is good in the scope and quality as dissertation in partial fulfillment for the requirement of Master's Degree of Science in Physics.

Evaluation Committee

.....

Asst. Prof. Dr. Manoj Kumar Chaudhary
Department of Physics, Amrit Campus
Lainchaur, Kathmandu
(Supervisor)

.....

Assoc. Prof. Dr. Janak Ratna Malla
Department of Physics, Amrit Campus
Lainchaur, Kathmandu
(HoD)

.....

Asst. Prof. Pitamber Shrestha
Coordinator, Department of Physics
Amrit Campus, Tribhuvan University
Kathmandu

.....
External Examiner

.....
Internal Examiner

Date: .....

Abstract

During the research, we analyzed the 1GIT protein in different mode with its ligand GDP followed by binding mechani44 between them under pyMoL tool. Guanosine-5′-Diphosphate is commonly used as Escherichia coli metabolite, a mouse metabolite and as an uncoupling protein inhibitor.There are so many diseases like associated disorders rachialgia, subarachnoid hemorrhage, epilepsy, stroke and neuroinfection in which GDP is advised to use. To address the effectiveness of this in clinical way, a computational practice was conducted on the force field parameter of GDP and it's binding mechanism. The energy profiles of bond strecthing, bending angles, dihedral angles, van-der Waals and electrostatic interactions of GDP under force field parameter signify the varia15n of bond energy with respect to interaction length and bonds in different nature. The van der Waals and electrostatic interaction between GDP and residues of 1GIT protein are noted to have negative values. The parabolic curve of bond stretching and bending angle clarify that the energy become minimum at equilibrium and increases quadratically upon expansion and contraction of bond length and bond angle. The attractive and repulsive forces of van der Waals curve are seen to cutoff around 6Å . Furthermore the electrostatic curve dies off slowly in quadratic nature but are found harder to predict with cutoff length. The polar bonds are visualized and measured in angstrom range between atoms of ligand and amino acid residues of protein. All these parameters gives idea to explain the binding mechanism fo GDP with 1GIT protein. These analysis are expected to a possble path for computational drug design mechanism.

Key Words:
Guanosine-5′-Diphosphate, CHARMM force field, 1GIT protein, pyMoL, swissparam

शोधसार

यस शोध कार्यमा हामीले 1GIT प्रोटिन र यसको Ligand GDP बिचको अन्तर्किया pyMol सफ्टवेयरको सहायताले अध्ययन गर्‍यौं । Guanosine–5′–Diphosphate (GDP) लाई rachialgia, subarachnoid hemorrhage, epilepsy, stroke र neuroinfection जस्ता विभिन्न विकारहरुमा सिफारिस गरिन्छ । GDP को क्लिनिकल प्रभावकारिता मुल्याङ्कन गर्न हामीले यसको Force Field Parameter हरु र Binding Mechanism को कम्प्युटेशन विश्लेषण गर्‍यौं ।

bond length र bond angle सँग Potential Energy मा देखिएको परिवर्तन bond stretching, bond angle, dihedral angles, van–der Waals र electrostatic अन्तरक्रियाको Energy Profile बाट अध्ययन गरियो । GDP र 1GIT प्रोटिनको अवशेषहरु बिचको Van der Waals तथा Electrostatic अन्तर–क्रियाको मान ऋणात्मक भएको पाइयो । bond stretching र bending angle बिचको ग्राफबाट Potential Energy equilibrium bond angle र bond length मा न्यूनतम भई वर्गीय रुपमा बढेको पाइयो । van–der Waals को आकर्षण र विकर्षण बलहरु 6Å मा न्युनतम भएको पाइयो । यसका साथै electrostatic वक्र वर्गिय स्वरुपले विलिन भएको कारणले cut–off length अनुमान गर्न सकिएन । प्रोटिनमा रहेका Ligand र amino acid residue बिच बनेका polar bond हरु pymol tool मा अवलोकन गरेर angstrom range मा नापियो । यस शोधकार्यबाट 1GIT Protin र यसको Ligand GDP बिच रहेका विभिन्न खालका Potential Energy हरुको परिवर्तन त्यसभित्र रहेका bond length र bond angle सँग दाँजियो । तसर्थ यस शोधकार्यबाट भविष्यमा Computational Drug Design Mechanism को केही सम्भाव्य कार्यहरु अनुमान गर्न सकिन्छ ।

Contents

RECOMMENDATION . . . . . **i**

ACKNOWLEDGEMENTS . . . . . **ii**

EVALUATION . . . . . **iii**

ABSTRACT . . . . . **iv**

LIST OF FIGURES . . . . . **viii**

LIST OF ABBREVIATIONS . . . . . **ix**

LIST OF TABLES . . . . . **x**

INTRODUCTION **1**

1 Introduction **2**
- 1.1 General Background 2
- 1.2 Guanosine-5′-diphosphate 3
- 1.3 1GIT Protein 5
- 1.4 Protein-Drug Binding Mechanism 6
- 1.5 PyMol 9
- 1.6 Open Babel 11
- 1.7 Swissparam 12
- 1.8 PubChem 13
- 1.9 Topology and Parameter Files 14
- 1.10 Objective of Study 15

LITERATURE REVIEW **16**

2 Literature Review **17**

METHODOLOGY **25**

3 Methodology **26**
- 3.1 CHARMM Force Field 26
- 3.2 Protein-Binding Kinetics 29
- 3.3 Computational Approach 30
  - 3.3.1 CHARMM Force Field Parameter of GDP 31
  - 3.3.2 Binding Mechanism of 1GIT Protein with GDP 31

RESULT AND DISCUSSION **33**

4 Result and Discussion **34**
- 4.1 Bond Length Energy 34
- 4.2 Bond Angle Energy 35
- 4.3 Dihedrals 37
- 4.4 Lennard-Jones Potential 39
- 4.5 Coulombs Potential 40

4.6 Protein-Ligand Interactions at 4Å 42
4.7 Protein-Ligand Interactions at 8Å 43

CONCLUSION **45**

5 Conclusion **46**
5.1 Future Work 46

REFERENCES **48**

APPENDIX **50**

List of Figures

1 Stick View of Guanosine -5′-diphosphate (GDP) with its Atom Name in PyMol 3
2 Caption for this figure with two images 4
3 Publication view of 1GIT Protein with it's Ligand 5
4 Schematic illustrations of the three protein-ligand binding models: (a) Lock-and-key; (b) Induced fit; and (c) Conformational selection. 7
5 Pymol view of 1GIT protein in different styles including sticks, cartoon,wire, mesh, dots, surface and sphere. 11
6 Parameters of CHARMM Force Field 27
7 Flow chart showing the steps followed for force field analysis of GDP . . 31
8 Flow chart showing the steps followed for binding mechanism of GDP with 1GIT protein 32
9 Energy Profiles for Bond Energy versus Bond Length 35
10 Energy Profiles for Bond Energy versus Bond Angles 36
11 Energy Profiles for Dihedral Energy versus Dihedral Angles 38
12 Energy Profiles for Lennard-Jones Energy versus Distance 40
13 Energy Profiles for Electrostatic Energy versus Distance 41
14 Amino acids residue of 1git protein forming polar bonds with the ligand at 4Å 42
15 Amino acids residue of 1git protein forming polar bonds with the ligand at 8Å 43

List of Abbreviations

GDP	Guanosine-5′-Diphosphate
MD	Molecular Dynamics
GPI	Glycosylphospho-Inositol
GPCR	G-Protein Coupled Receptors
GTP	Guanosine-5′-Triphosphate
GEF	Guanine Nucleotide Exchange
GAP	GTPase-Activating Proteins
GMP	Guanosine Monophosphate
CAT	Cool-Associated Tyrosine Phosphorylated
GPCR	G-Protein-Coupled Receptor
GPI	Glucose 6-Phosphate Isomerase
G6P	Glucose 6-Phosphate
F6P	Fructose 6-Phosphate
MMFF	Merck Molecular Force Field
NCBI	National Center for Biotechnology Information
PSF	Protein Structure File
PDB	Protein Data Bank
CSD	Cambridge Structural Database
ffTK	Force Field Toolkit
VMD	Visual Molecular Dynamics
ARPI	Androgen Receptor Pathway Inhibition
CRPC	Castration-Resistant Prostate Cancer
RMSE	Root Mean Square Error
SMD	Steered Molecular Dynamics
CHARMM	Chemistry at HARvard Macromolecular Mechanics

List of Tables

4.1 Parameters for Bond Length Interactions 34
4.2 Parameters for Bond Angles Interactions Used in This Work 36
4.3 Parameters for Dihedral Angles Interactions 38
4.4 Parameters for Lennard-Jones (van der Waals) Interactions Used in This Work 39
4.5 Parameters for Electrostatic Interactions Used in This Work 41
A.1 Parameters for Bond Length Interactions Used in This Work 50
A.2 Parameters for Bond Angles Interactions Used in This Work 51
A.3 Parameters for Dihedral Angles Interactions Used in This Work 52
A.4 Parameters for Lennard-Jones (van der Waals) Interactions Used in This Work 54
A.5 Parameters for Electrostatic Interactions Used in This Work 54

Chapter 1
INTRODUCTION

1 Introduction

1.1 General Background

Force field analysis of drug-protein interactions is a fundamental approach in computational drug discovery and design. Force fields are mathematical models that describe the potential energy surface of a molecular system, including the interactions between atoms and molecules. In drug-protein binding studies, force fields are employed to simulate the conformational changes, binding affinities, and dynamics of drug molecules interacting with protein targets.

One common approach to studying drug-protein interactions using force fields is molecular docking, where the drug molecule is computationally positioned within the protein's binding site to predict the most favorable binding orientation and affinity. Force fields are used to calculate the interaction energies between the drug and protein, including contributions from van der Waals forces, electrostatic interactions, hydrogen bonding, and solvation effects. By systematically exploring the conformational space of the drug molecule and evaluating its interactions with the protein, molecular docking provides insights into the structural determinants of binding and guides the rational design of new drugs with improved potency and selectivity.

Molecular dynamics (MD) simulations complement molecular docking by providing a dynamic view of drug-protein interactions over time. In MD simulations, force fields are used to numerically integrate Newton's equations of motion for all atoms in the system, allowing the drug molecule and protein to move and interact with each other under the influence of interatomic forces. Force field parameters govern the behavior of bonded and non-bonded interactions, enabling the simulation of protein conformational changes, ligand binding events, and solvent effects. By analyzing the trajectories generated from MD simulations, researchers can gain insights into the kinetics, thermodynamics, and structural dynamics of drug-protein interactions, which are essential for understanding drug binding mechanisms and predicting ligand binding affinities.

Force field analysis of drug-protein interactions continues to advance with the development of more accurate and sophisticated force fields and computational techniques. By combining molecular docking, molecular dynamics simulations, and other computational methods with state-of-the-art force fields, researchers can systematically explore the binding properties of drug candidates, predict their biological activities, and guide the rational design of novel therapeutics for various diseases.

1.2 Guanosine-5′-diphosphate

Guanosine diphosphate (GDP) is an ester of pyrophosphoric acid linked with the nucleoside guanosine, called a nucleoside diphosphate. The G-proteins are involved in signal transduction and the composition of GDP are nucleobase guanosine, pentose sugar ribose, and pyrophosphate group. The validated condensation of the hydroxy group at the 5′ position of guanosine with pyrophosphoric acid results in a purine ribonucleoside 5′-diphosphate, which is conjugate acid of a GDP(2-). It plays a magnificent bit part as a mouse metabolite, an *Escherichia coli* metabolite, and an uncoupling protein restraint. The IUPAC name of GDP is [(2R, 3S, 4R, 5R)-5-(2-amino-6-oxo-1H-purin-9-yl)-3,4-dihydroxyoxalan-2-yl] methyl phosphono hydrogen phosphate and its molecular formula is $C_{10}H_{15}N_5O_{11}P_2$.

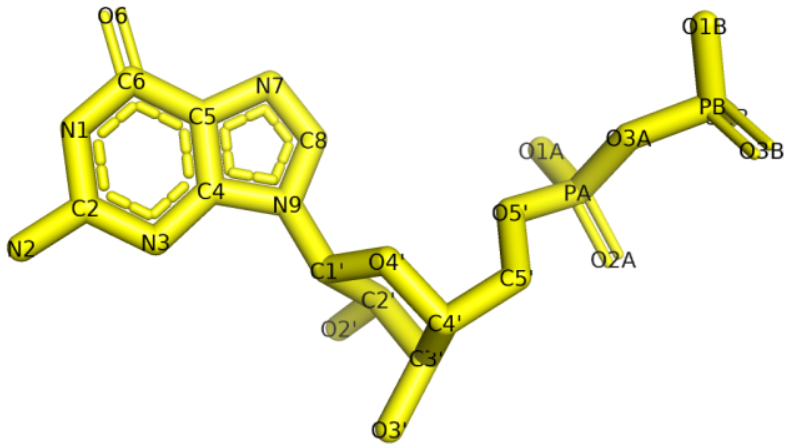


Figure 1: Stick View of Guanosine -5′-diphosphate (GDP) with its Atom Name in PyMol

For carbohydrates, glycosyltransferases are the main enzymes responsible for the assembly. The synthesis of glycosylphospho-inositol (GPI) anchors, eukaryotic N-glycons and bacterial cell-surface polysccharides requires GDP-activated sugar, GDP-mannose (GDP-man) for the biosynthesis of mannosyl donor dolicol phosphate β-D mannnose (Dol-P-man). The kinetics and characteristics of GTPase such as those associated with G-protein coupled receptors (GPCR) are studied by the usage of GDP. It is used as a substrate of Pyruvate kinase to produce GTP in support of RNA biosynthesis. Guanosine-5'-diphosphate (GDP) is a nucleotide crucial for diverse cellular processes, serving as a precursor to guanosine-5'-triphosphate (GTP) and playing a central role in signal transduction and energy transfer. Structurally, GDP comprises a guanine base, a ribose

sugar, and two phosphate groups attached to the 5' carbon of the ribose. It is a key participant in the regulation of cellular signaling pathways, acting as a molecular switch through the controlled conversion between GDP and GTP. This conversion is orchestrated by guanine nucleotide exchange factors (GEFs) and GTPase-activating proteins (GAPs), influencing the activity of GTP-binding proteins (G proteins). Additionally, GDP is an energy carrier, with its hydrolysis to GTP releasing energy utilized in cellular processes. Synthesized from guanosine monophosphate (GMP) and metabolically linked to GTP, GDP is integral to maintaining cellular homeostasis. In the context of RNA and DNA, while GDP itself is not part of their structures, its guanine component is one of the four nitrogenous bases. Understanding the balance between GDP and GTP concentrations is vital, as dysregulation can contribute to diseases. In research and drug development, GDP and related nucleotides are subjects of interest, particularly in studying GTPases and their roles in cellular signaling, offering potential targets for therapeutic interventions in various biomedical applications[1].

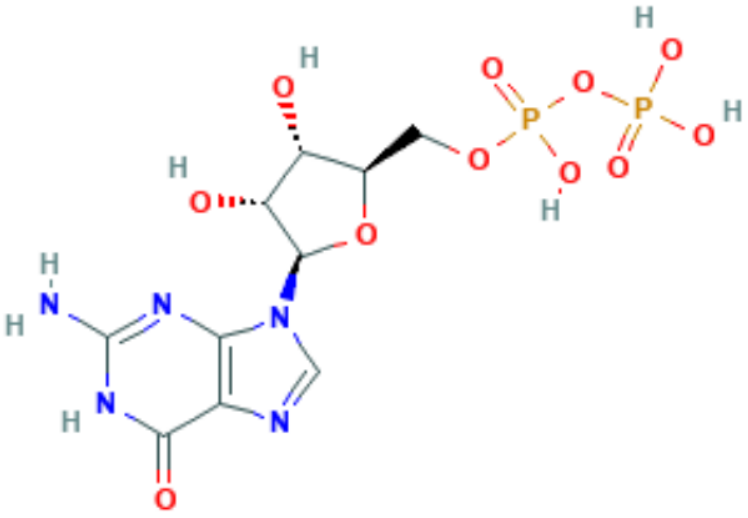


Figure 2: 2D Stucture of GDP with Hydrogen Atoms and Phosphate ligand[2]

1.3 1GIT Protein

The member of the family of multi-functional proteins that are known to balance the activity of certain members of Arf (for ADP ribosylation factor) is called GIT (for G-protein coupled receptor kinase interactor). GIT is also called Cat (for cool-associated tyrosine phosphorylated). 1-GIT is the binding partner for members of a family of threonine/serine protein kinases which is found in a yeast two-hybrid screen [3].

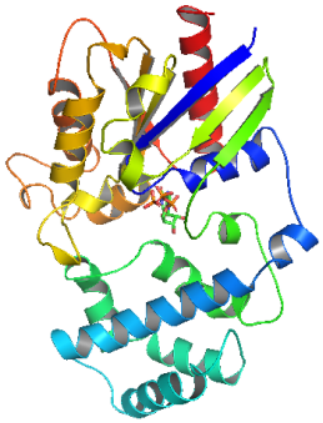


Figure 3: Publication view of 1GIT Protein with it's Ligand

The omnipresent multidomain proteins involved in multiple cellular processes are called G-protein-coupled receptor (GPCR)-kinase-interacting proteins 1 and 2 (1-GIT AND 2-GIT). The dimensional activation of several signaling molecules is controlled by GITs which act as scaffolds and play pathogenic roles in Huntington's disease and HIV infections. There is 65% sequence identity and 85% similarity in human 1-GIT and 2-GIT. The localization of focal complexes at the cell periphery, focal adhesion on the cytoplasmic structures and ventral surfaces, and formation of homo and heterodimers is done by GIT proteins [4].

The 1GIT protein, also known as Glucose-6-phosphate isomerase (GPI), plays a crucial role in cellular metabolism, particularly in the glycolytic pathway. It catalyzes the reversible isomerization of glucose-6-phosphate (G6P) to fructose-6-phosphate (F6P), a pivotal step in glycolysis and gluconeogenesis. This enzymatic activity allows cells to regulate glucose metabolism according to energy demands and nutrient availability. Beyond its metabolic function, 1GIT has been implicated in various non-metabolic roles, including cell signaling, immune response modulation, and even as a moonlighting protein with chaperone-like activities.

Structurally, 1GIT exhibits a TIM barrel fold, a common motif found in many enzymes catalyzing similar reactions. This barrel-shaped structure provides a stable framework for the active site, where the substrate binds and undergoes the isomerization reaction. Moreover, 1GIT's active site contains several conserved amino acid residues critical for substrate binding and catalysis. Understanding the structural intricacies of 1GIT has paved the way for designing inhibitors targeting this protein, offering potential therapeutic avenues for diseases where glucose metabolism dysregulation plays a significant role, such as cancer and diabetes.

Research on 1GIT continues to unravel its diverse functions and implications in health and disease. Beyond its ca[51]nical role in glucose metabolism, recent studies have highlighted its involvement in various cellular processes, including cell proliferation, differentiation, and apoptosis. Furthermore, aberrant expression or activity of 1GIT has been associated with pathological conditions, making it a potential target for therapeutic interventions. Continued exploration of 1GIT's multifaceted roles promises to provide deeper insights into cellular physiology and potentially uncover novel strategies for managing metabolic disorders and other diseases.

1.4 Protein-Drug Binding Mechanism

The [67]otein-drug binding is defined as the complex formation of a drug with protein. The protein binding may be divided into intracellular and extra[8]llular binding where proteins are particularly responsible for such interaction. The weak chemical bonds such as hydrogen bonds, hydrophobic bonds, ionic bonds, and van der Waal's forces are involved under reversible protein-drug binding. The protein-drug binding may be defined through indirect techniques such as equilibrium dialysis, dynamic dialysis, ultrafiltration, etc., and direct techniques such as UV-spectroscopy, fluorimetry etc. It [20]y be affected by two factors such as drug-related factors and protein-related factors [5].

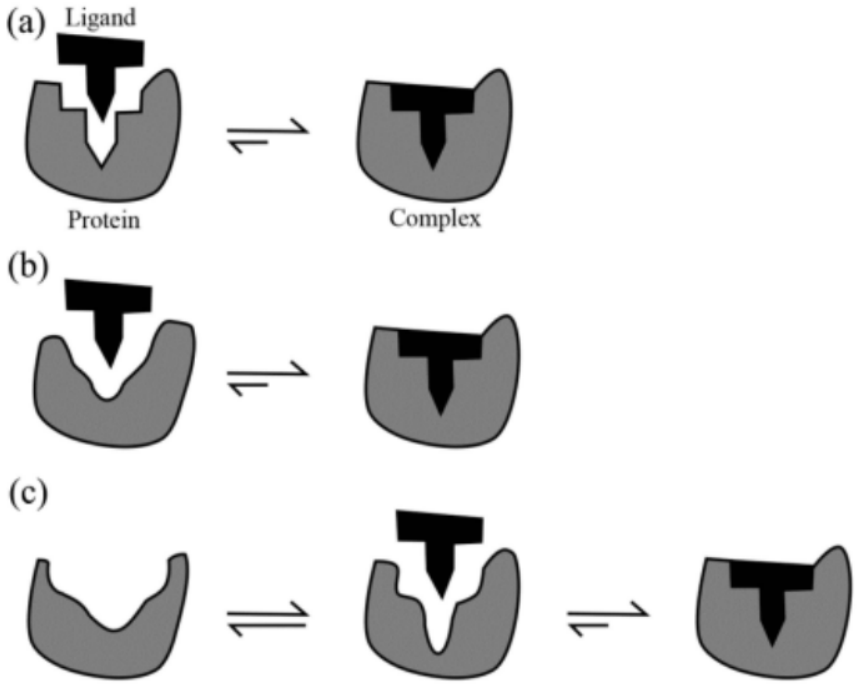


Figure 4: Schematic illustrations of the three protein-ligand binding models: (a) Lock-and-key; (b) Induced fit; and (c) Conformational selection[6]

.

The drug-related factor consists of the physiological characteristics of the drug, the concentration of the drug in the body, and the affinity of a drug for a particular binding component site. Similarly, the protein-related factors consist of the physiological characteristics of the protein or binding component and the number of binding sites on the binding agents. The displacement interaction is the competition between drugs for the common binding site which may result in an unexpected rise in free concentration of the displaced drug. The significance of protein-drug binding is as follows: systematic solubility of drugs, absorption, distribution, displacement interaction and toxicity, diagnosis, therapy, and drug targeting. The protein-drug binding mechanism is a multifaceted process fundamental to the efficacy and specificity of pharmaceutical interventions. This intricate interaction involves specific molecular recognition, primarily at well-defined binding sites on proteins. The binding site's three-dimensional structure, shaped by amino acid residues, dictates the binding affinity and selectivity for particular drugs. Two models, the classical "lock-and-key" and the "induced fit," describe the fitting process between the drug and its protein target. This interaction is characterized by various non-covalent forces, including hydrogen bonding, van der Waals forces, hydrophobic interactions, and electrostatic forces.

Hydrogen bonds contribute significantly to specificity, creating stabilizing links between the drug and amino acid residues in the binding site. Electrostatic interactions, including salt bridges, also play a role, especially for charged drug molecules. Hydrophobic interactions, prominent in the binding of lipophilic drugs, involve hydrophobic regions on the drug interacting with complementary regions on the protein.

Van der Waals forces further stabilize the binding by facilitating the close packing of drug and protein molecules at the binding interface. Allosteric modulation introduces an additional layer of complexity, involving ligand binding at sites other than the active site, inducing conformational changes in the protein.

The kinetics and thermodynamics of protein-drug binding describe the rate of association and dissociation and the energetics of the interaction. Various factors, such as temperature, pH, and ionic strength, can influence the binding affinity. This detailed understanding of the binding mechanism is instrumental in rational drug design, guiding efforts to optimize drug candidates for enhanced binding and therapeutic efficacy.

In a clinical context, disruptions or enhancements of protein-drug binding can significantly impact drug efficacy, safety, and potential side effects. The knowledge of binding mechanisms aids in predicting and understanding drug interactions and potential off-target effects, contributing to informed decision-making in drug development and therapeutic applications. Incorporating computational methods, structural biology, and experimental assays, researchers strive to unravel the complexities of protein-drug interactions, advancing the field of pharmaceutical science and precision medicine[7].

Proteins serve as essential targets for drug action due to their diverse functions in biological systems. The mechanism by which drugs bind to proteins is crucial for understanding their efficacy and potential side effects. The process of protein-drug binding involves several steps, beginning with the recognition and interaction between the drug molecule and specific binding sites on the protein surface.

The initial step in protein-drug binding is molecular recognition, where the drug molecule encounters the protein and forms non-covalent interactions such as hydrogen bonding, hydrophobic interactions, electrostatic interactions, and van der Waals forces. These interactions determine the specificity and affinity of the binding and are influenced by the chemical properties of both the drug and the protein.

Once the drug molecule binds to the protein, conformational changes may occur in either or both molecules to optimize the binding interaction. This induced fit model suggests that the binding event induces structural changes in the protein or drug, leading to a more favorable binding configuration. These conformational changes can alter the

protein's activity or stability and are critical for the drug's pharmacological effects.

The strength of the protein-drug interaction is characterized by the binding affinity, which refers to the degree of attraction between the drug and the protein. High binding affinity ensures tight and long-lasting interactions, increasing the drug's potency. Conversely, low binding affinity may result in weaker interactions and decreased drug efficacy, necessitating higher drug concentrations for therapeutic effects.

The binding kinetics describe the rates of association and dissociation between the drug and the protein. The association rate constant represents the speed at which the drug binds to the protein, while the dissociation rate constant indicates how quickly the drug dissociates from the complex. The ratio of these constants, known as the dissociation constant (KD), quantifies the equilibrium between bound and unbound drugs and determines the strength of the interaction.

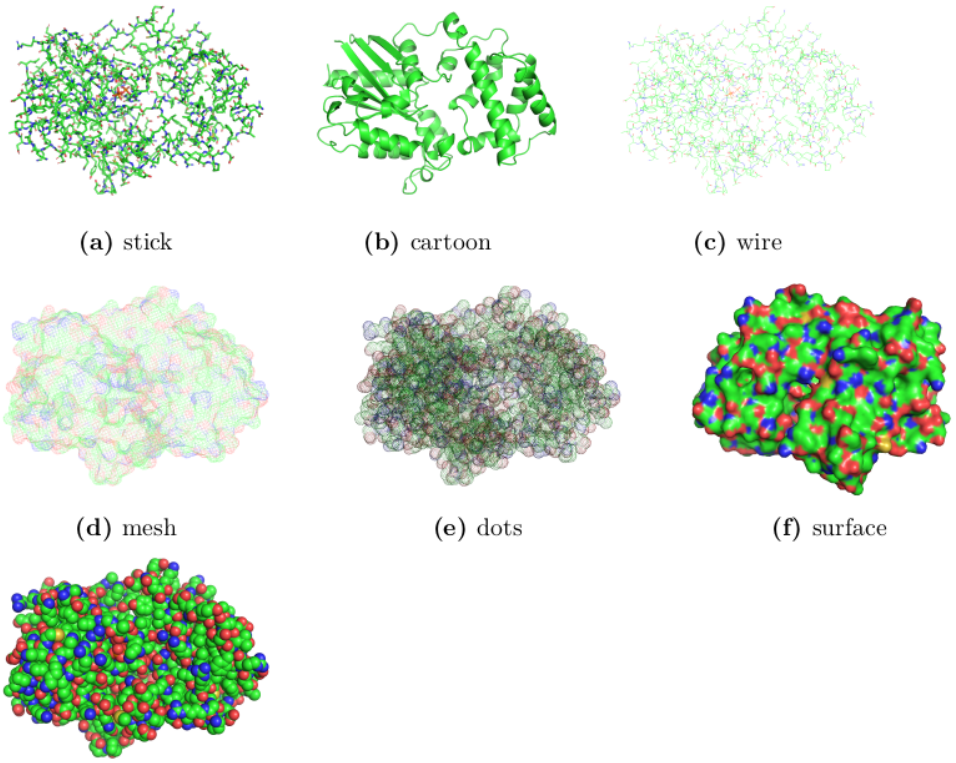
Various factors influence protein-drug binding kinetics, including temperature, pH, solvent composition, and the presence of other molecules. Changes in these environmental conditions can alter the binding affinity and kinetics, affecting the drug's pharmacokinetics and pharmacodynamics in vivo. Understanding these factors is crucial for optimizing drug design and dosage regimens to achieve desired therapeutic outcomes while minimizing adverse effects.

In summary, protein-drug binding involves molecular recognition, conformational changes, and non-covalent interactions that determine the specificity, affinity, and kinetics of the interaction. By elucidating the mechanisms underlying protein-drug binding, researchers can design more effective and safer drugs for treating various diseases.

1.5 PyMol

The development in internet, hardware, and software technology has launched the free cross-programmed molecular graphics system which provides the majority of the performances and potentiality of traditional molecular graphics packages written in Fortran or C. Without cost or limitation, a researcher can freely acquire PyMol and then dispense their derivative tasks because it is available under an entirely open-source software license. The most usual representations for macromolecular structures: ball-and-stick, dot surfaces, solid surfaces, wire mesh surfaces, backbone ribbons, wire bonds, cylinder spheres and cartoon ribbons are supported in PyMol. It loads molecules from PDB and another file format where hydrogen bonding interaction and distances are indicated by labeling of atoms and dashed bonds. PyMOL stands out as a versatile and user-friendly

molecular visualization tool extensively employed in structural biology, biochemistry, and computational chemistry. Its high-quality graphics and intuitive graphical user interface enable users to interactively explore and manipulate three-dimensional molecular structures, while its support for various file formats, including PDB and CIF, ensures seamless loading and visualization of diverse molecular data. PyMOL's scripting capabilities, powered by the Python programming language, empower users to automate tasks, create custom workflows, and generate complex visualizations. Beyond traditional stick and ball representations, PyMOL offers advanced features such as surface rendering and the addition of annotations, labels, and measurements for detailed analysis. Widely used in education for teaching molecular and structural biology, PyMOL also serves as an effective presentation tool in research settings. Supported by an active user community, tutorials, forums, and documentation, PyMOL is available in both open-source and commercial versions, offering flexibility to users. Its integration capabilities with other computational tools further enhance its utility, solidifying PyMOL as an indispensable resource for researchers, educators, and professionals in the field[8]. RING-PyMOL serves as a PyMOL plugin offering a suite of analytical tools tailored for structural ensembles and molecular dynamic simulations. The plugin integrates residue interaction networks, sourced from the RING software, with structural clustering to enhance the analysis and visualization of conformational complexity. This amalgamation enables the precise computation of non-covalent interactions, leveraging PyMOL's capabilities for the manipulation and visualization of protein structures. RING-PyMOL efficiently identifies and emphasizes correlated contacts and interaction patterns, shedding light on structural allostery, active sites, and the structural diversity associated with molecular function. Known for its user-friendly interface and remarkable speed, the plugin swiftly processes and renders extensive models and lengthy trajectories within seconds. Additionally, RING-PyMOL generates interactive plots and output files compatible with external tools. Significant enhancements have been made to the underlying RING software, rendering it 10 times faster, capable of handling mmCIF files, and proficient in identifying typed interactions, even for nucleic acids.



(a) stick **(b)** cartoon **(c)** wire

(d) mesh **(e)** dots **(f)** surface

(g) sphere

Figure 5: Pymol view of 1GIT protein in different styles including sticks, cartoon,wire, mesh, dots, surface and sphere.

1.6 Open Babel

A set of programming libraries and user applications are used to compose Open Babel which facilitates file format conversion, 2D or 3D structure generation. It is used to manipulate the information in molecule titles, and property fields and has a lot trademarks to handle multi-molecule file formats (such as sdf to mol2). It is freely available under an open-source license and used as a programming library to operate chemical data in areas such as drug design, computational chemistry, organic chemistry, material science, etc. Open Babel is an open-source chemical toolbox that serves as a versatile and powerful resource in the field of cheminformatics. It excels in facilitating the interconversion of chemical file formats and manipulating chemical data, offering a comprehensive solution for researchers and developers engaged in molecular informatics. Key functionalities include format conversion, allowing seamless transformation between a wide range of representations such as SMILES, InChI, and SDF, and the manipulation

of chemical data for tasks like substructure searching, molecular descriptor computation, and molecular similarity analyses. The toolkit's programmatic accessibility makes it adaptable for custom cheminformatics solutions, enabling developers to integrate its capabilities into their applications and workflows. Open Babel's command-line interface further enhances its accessibility, catering to users with diverse programming expertise for quick format conversions and basic cheminformatics tasks. The toolkit's community-driven development ensures continuous updates, compatibility with evolving chemical informatics standards, and active maintenance. It solidifies its role as an essential and evolving tool for those involved in computational chemistry and molecular informatics endeavors [9].

1.7 Swissparam

Swissparam is harmonious with the CHARMM force field which is able to generate, topologies and parameters, for arbitrary small organic molecules based on Merck. molecular force field. In pharmaceutical research, computer-aided drug design (CADD) is indispensable and is partitioned into two categories: structure-based approaches and ligand-based approaches which use the mastery retained from familiar bioactive compounds to derive new ones. The docking software EADock DSS and EADock2 use topologies and parameters generated by Swissparam for docking of small molecules. SwissParam is a valuable web-based tool designed for the generation of parameters and topologies essential in molecular dynamics simulations, particularly for small molecules. Initially released in 2011, SwissParam has evolved to offer a range of new features, including the incorporation of data from over 120 sources, notably integrating Google Patents information into PubChem Patent data. It supports the generation of parameters and topologies based on the Merck molecular force field (MMFF) and is compatible with the CHARMM22/27 and CHARMM36 force fields. The tool allows the setup of covalent ligands, and users can import molecules from various file formats or via a molecular sketcher. Moreover, SwissParam has been updated to provide information on probable alternative tautomers and protonation states of a query molecule, enhancing its utility for designing potent and selective modulators. The latest version of SwissParam is freely accessible at www.swissparam.ch and also features a newly implemented command-line interface, offering a comprehensive solution for researchers and practitioners in computational chemistry and drug design[10].

Many aspects of drug discovery are now routinely facilitated and expedited by computer-aided drug design tools. Among these tools, structure-based approaches leverage the three-dimensional structure of the target biomolecule as a primary source of information.

When calculating the interactions between small molecules and proteins using molecular mechanics equations, it is essential to have parameters for topologies, atom typing, and force field. However, generating parameters for small molecules poses challenges due to the diverse chemical groups. The SwissParam web tool, initially introduced in 2011, aims to generate parameters and topologies for small molecules based on the Merck molecular force field (MMFF) while being compatible with the CHARMM22/27 force field. This updated version of SwissParam introduces several new features, including the ability to set up covalent ligands. Molecules can now be imported from various file formats or through a molecular sketcher. The MMFF-based approach has been enhanced to provide parameters and topologies compatible with the CHARMM36 force field. An option has been added to generate small molecule parametrizations following the CHARMM General Force Field using the multipurpose atom-typer for the CHARMM (MATCH) approach. Additionally, the updated SwissParam now offers information on probable alternative tautomers and protonation states of the query molecule, enabling users to consider all relevant microspecies for their compound. The latest version of SwissParam is freely accessible at www.swissparam.ch and can also be utilized through a newly implemented command-line interface[11].

1.8 PubChem

The information about chemical substances and their biological activities is provided by a public repository called PubChem, which is maintained by the National Institutes of Health (NIH) in the USA. It distributes scientific communication in many areas such as drug discovery, medicinal chemistry, chemical biology, and cheminformatics so it has swiftly developed as a key chemical information resource. PubChem is a vigorous programmatic access that provides its data to RDF to ease data analysis, sharing, and integration with other databases. Using the Chemical Structure Search tool, it permits users to carry out various nontextual searches (such as molecular formula search, 2D and 3D similarity searches, superstructure/substructure search, etc [12].

PubChem (https://pubchem.ncbi.nlm.nih.gov) is a widely utilized chemical information resource catering to diverse needs. Over the past two years, several enhancements have been implemented in PubChem. Data from over 120 sources have been incorporated, with notable developments such as the integration of Google Patents data, significantly expanding the coverage of PubChem Patent information. New additions include the Cell Line and Taxonomy data collections, offering convenient access to chemical details associated with specific cell lines and taxa. The bioassay data model has been updated, and PubChem's programmatic access protocols, PUG-REST and PUG-View, now boast

additional features. These include support for target-centric data downloads related to proteins, genes, pathways, cell lines, and taxa, along with the introduction of the 'standardize' option in PUG-REST, which returns the standardized form of a given chemical structure. A substantial update has also been made to PubChemRDF. This paper provides a comprehensive overview of these recent changes[13].

PubChem, maintained by the National Center for Biotechnology Information (NCBI), is a comprehensive database of chemical compounds, including their biological activities and associated information. It is a valuable resource for researchers, educators, and the general public interested in chemical and biological data. One of PubChem's primary objectives is to facilitate the discovery of new drugs and understand their mechanisms of action by providing access to extensive chemical, biological, and structural data.

The database contains information on millions of chemical substances, including organic and inorganic compounds, natural products, drugs, and their associated bioactivities. Each compound entry in PubChem includes detailed chemical descriptions, such as molecular structure, formula, weight, and physical properties. Additionally, PubChem integrates biological screening data, allowing users to explore the compounds' interactions with biological targets and assess their potential pharmacological effects.

PubChem offers various search and analysis tools to retrieve and analyze chemical and biological data efficiently. Users can search for compounds by name, chemical structure, or biological activity, enabling them to identify compounds of interest for further investigation. Moreover, PubChem provides tools for comparing chemical structures, predicting properties, and visualizing molecular interactions, empowering researchers to explore the relationships between chemical structure, biological activity, and drug efficacy.

Beyond its role as a repository of chemical and biological data, PubChem fosters collaboration and data sharing among researchers worldwide. By providing free and open access to its vast collection of chemical and biological information, PubChem accelerates drug discovery efforts, promotes scientific innovation, and advances our understanding of chemical biology and drug development. As a result, PubChem plays a pivotal role in driving progress in biomedical research and improving public health outcomes.

1.9 Topology and Parameter Files

The whole information needed to convert a list of residue names into a complete protein structure file (PSF) are contained in topology files. The PSF file can be constructed

by defining the type, mass and charge of every atom in every residue by topology file. It helps in the automatic generation of angles and dihedrals for every pair or triple of connected bonds.

All the numerical constants needed to evaluate forces and energy are contained in parameter files into a PSF structure file. It is mainly used to generate the PSF file. Different sets of data for bond length, bond angle, dihedral angles and non-bonded interactions are included in this files. The parameter file is closely linked with the topology file to generate the PSF file. In molecular dynamics (MD) simulations, topology and parameter files are pivotal components that dictate the behavior of proteins and ligands within a computational framework. For proteins, the topology file delineates the connectivity of atoms and residues, encompassing information on the amino acid sequence and secondary structure, while the parameter file specifies force field characteristics governing bond lengths, angles, and non-bonded interactions. Ligand topology files elucidate chemical structure details, including atom types and bond connectivity, while parameter files provide force field parameters for accurate representation during simulations. Choosing compatible force fields, such as AMBER or CHARMM, for proteins and ligands is essential to avoid mismatches and ensure realistic interactions. Water models, like TIP3P or SPC/E, require separate topology and parameter files, contributing to the overall simulation setup's accuracy. Ultimately, the seamless integration of these files ensures a comprehensive representation of proteins, ligands, and solvent molecules, enhancing the reliability and validity of MD simulations in studying the dynamic behavior of biomolecular systems[14].

1.10 Objective of Study

The main objectives of our research work are noted as follows:

- visualization of 1GIT protein with its ligand in different modes in pyMoL.
- study energy profiles of CHARMM force field parameters including bonded and non-bonded terms.
- identification of polar bond interacting amino acid residues
- study the graphical study of energy profiles using mathematica.

Chapter 2
LITERATURE REVIEW

2 Literature Review

In 2001, Sigel, et al., determined the stabilities of the $(M;PuNTP)^{-}$ and $M(PuNTP)^{-2}$ complexes through the potentiometric p^H titrations by neglecting the self-association of purine-nucleoside 5′-triphosphates (PuNTPs) ITP and GTP. The binding site for metal ion was especially favored for N7 of the guanine residue. The application of $\log K^{M}_{MPuNTP} - \log K^{M}_{MPyNTP}$ played a vital role in the increment of percentage of the macro-chelated isomers in the $M(ITP)^{-2}$ and $M(GTP)^{-2}$ system so that the highest and lowest formation degrees were observed as $97 \pm 1\%$ and $17 \pm 1\%$ for $Mg(ITP)^{-2}$ and $Cu(GTP)^{-2}$ respectively [15].

In 2004, the crystal structure of a sample of 510 pharmaceutically relevant protein-ligand complexes were extracted from the Protein Data Bank (PDB) to inspect the quality of Catalyst's conformational model generation algorithm by Kirchmair, et al.,. The implemented algorithm showed the high-quality conformers in silico drug research. The ligands are examined in high energy conformations and the computing time was analyzed for different conformational model generations. The fitting qualities of experimental conformations from the Cambridge Structural Database (CSD) and PDB were compared and properties of the same ligands in various proteins were explored. The small organic ligands were extracted from their respective complexes and stored in the format of MDL MOL files by the software LigandScout. This software is a novel program for visualization and analysis of ligand-protein complexes [16].

In 2009 guanosine-5′-diphosphate (GDP)-L-fucose was studied in the manner of an activated form of nucleotide sugar which plays a significant role in a spacious range of biological functions. The supplement of mannose aimed to enhance the GDP-L-fucose production for a potentially better carbon source to be transformed into GDP-L-fucose than glucose. A 1,3-fold increase in GDP-L-fucose concentration ($52 \cdot 5 \pm 0 \cdot 8$ mgl^{-1}) was done by the supplement of mannose and glucose in a batch-wise fermentation of recombinant E. coli. The control strain overexpressing gmd and wcaG genes was compared with a maximum GDP-L-fucose concentration of ($70 \cdot 3 \pm 2 \cdot 3$ mgl^{-1}). The mass production of GDP-L-fucose was done by optimization of microbial system [17].

In 2015, Sayer, et al., studied the nucleoside bis-phosphate analogs and their metal-ion complexes under various physiological functions. The solution symmetry of 2′-deoxyadenosine-and 2′-deoxyguanosine-3′,5′-biphosphate, 3 and 4, d(pNp) as well as Mg^{+2}/Zn^{+2} binding sites and binding-modes were surveyed. Due to the formation of inner-sphere in the $NDP - Zn^{+2}$ complex versus outer-sphere coordination in the $Zn^{+2} - d(pNp)$, the unpredicted value of low logK values of $Zn^{+2} - NDP$ as compared

to $Zn^{+2} - d(pNp)$ was presumed. The stability of corresponding M^{+2}/Zn^{+2} complexes and the acid-base equilibria of the nucleotides were determined in the absence or presence of Mg^{+2}/Zn^{+2} ions by p^H-titration [18].

The catalyzation of the cleavage of the carbon-sulfur bond from HBPS (2′–2-sulfinate) in the final step of microbial PS pathway desulfurization reactions was done by HBPS in 2017. The accurate molecular mechanics force fields were produced to develop and validate CHARMM-compatible force field parameters for the HBPS substrate and 2-hydroxybiphenyl product. The optimization of charge, angle, bond distances, and torsional parameters were done by using the Force Field Toolkit (ffTK) in Visual Molecular Dynamics (VMD). The quantum mechanical calculations were used to estimate the optimized geometries. The abilities of optimized parameters were assessed through molecular simulations of the molecules in implicit and explicit water solutions to restate optimized geometries. The dihedral angles for HBP (2-hydroxybiphenyl),HBPS and BCA (2-biphenyl carboxylic acid) were found as 53.5°, 59.2° and 59.4° by applying quantum mechanical geometry optimization [19].

In 2019, the trial cure of treatment-induced neuroendocrine prostate cancer (t-NEPC), also called aggressive subtype of Castration-resistant prostate cancer (CRPC) was done by the implementation of potent androgen receptor pathway inhibition (ARPI) therapies. About 20% of CRPC cases were found as t-NEPC which was predicted with the wide application of ARPI therapies. The distinctive event in t-NEPC victims was found which is RNA splicing of the G-protein-coupled receptor kinase-interacting 1-GIT protein. The researcher revealed that t-NEPC patient tumors contain both up-regulation of the 1-GIT-A splice variant expressions. The differential transcriptomers were regulated by 1-GIT-A and 1-GIT-C in prostate cancer. The study revealed that the neural function, morphogenesis and environmental sensing via cell-adhesion processes are regulated by 1-GIT-A protein. The distinct molecular changes were regulated by 1-GIT-A and 1-GIT-C which showed that RNA splicing of 1-GIT is linked with the progression of NEPC [20].

In 2020, Lamichhane, et al., performed a molecular dynamic simulation for the interactions between the inhibitor N3 and MP-LBD residues and investigated the structural features of a ligand binding domain (LBD) in COVID-19 main protease (MP). The superimposed β-barrels and flexible α-helices generated the folding gates in MP-N3 complexes and their degrees were signified through the calculation of root mean square fluctuation (RMSD), the radius of gyration (RG), radial distribution functions (RDF), etc. The flat RDF peak for the MP-N3 atoms pair was seen for GLY143 which was registered as the highest value of residue velocity. Similarly, the sharp RDF peaks for three H-binding atom pairs nearly at 2 Å the radial distance was seen for GLU166 which

was registered as the lowest residue velocity. The mean± SD in RMSDs of N3 and MP were found as $1 \cdot 89 \pm 0 \cdot 36$ Å and $1 \cdot 73 \pm 0 \cdot 16$ Å through 10ns long equilibrations. The compactness of the system was measured by RG which was obtained as $6 \cdot 64 \pm 0 \cdot 23$ Å and $22 \cdot 41 \pm 0 \cdot 10$ Å for N3 and MP respectively [21].

In 2013, Mayne, et al., studied the challenge of quickly and accurately generating parameters for new chemical compounds hinders the widespread use of molecular dynamics simulations in various biological applications, particularly in drug discovery. Despite the availability of generalized versions of common classical force fields like the General Amber Force Field and CHARMM General Force Field, several technical challenges persist, limiting their broad application. This study introduces the Force Field Toolkit (ffTK), which addresses these challenges by employing algorithm and method development, automating repetitive and error-prone tasks, and incorporating a user-friendly graphical interface. Distributed as a VMD plugin, ffTK streamlines the parameterization of ligands, enabling a systematic workflow that produces a comprehensive set of CHARMM-compatible parameters[22]

In the realm of computer-aided drug design, the computational prediction of thermodynamic components has become a commonplace practice. While significant strides have been made in enhancing the calculation of free energy, the prediction of enthalpy (and entropy) has not received ample attention. Additionally, there is a scarcity of literature that comparatively evaluates the performance of different force fields and water models in conjunction with each other. This study addresses these gaps through a comprehensive assessment of force fields and water models using host-guest systems, which mimic key features of protein-ligand systems. The use of computationally inexpensive host-guest systems, owing to their smaller size compared to protein-ligand systems, facilitates this evaluation.

In 2021, Biggin, et al., focuses on absolute enthalpy calculations employing the multi-box approach for a set of 25 cucurbit uril–guest pairs. Eight water models (TIP3P, TIP4P, TIP4P-Ew, SPC, SPC/E, OPC, TIP5P, Bind3P) and five commonly used force fields (GAFFv1, GAFFv2, CGenFF, Parsley, and SwissParam) are considered. The findings reveal a strong sensitivity of host-guest binding enthalpies to the choice of force field and water model. Among the water models, TIP3P and its derivative Bind3P emerge as the best-performing models for this specific host-guest system. Importantly, the performance tends to be superior for aliphatic compounds compared to aromatic ones, underscoring the ongoing challenge of accurately incorporating aromaticity into these simple force fields. This study contributes valuable insights to the literature, shedding light on the intricate interplay between force fields, water models, and the accurate prediction of thermodynamic properties in the context of host-guest systems[23].

In computer-aided drug design, the routine computational prediction of thermodynamic components has garnered attention, particularly in calculating free energy. However, the prediction of enthalpy, and to a lesser extent entropy, remains an area that needs further exploration. Additionally, there is a notable gap in the literature regarding the comparative assessment of different force fields and water models in conjunction with each other. This paper aims to address these gaps by examining a comprehensive study that assesses force fields and water models using host-guest systems, mimicking important aspects of protein-ligand interactions. These host-guest systems are computationally cost-effective due to their smaller size compared to protein-ligand systems.

In 2020, Sonibare et al., focused on absolute enthalpy calculations, employing a multi-box approach for a set of 25 cucurbit[7] uril–guest pairs. Eight water models (TIP3P, TIP4P, TIP4P-Ew, SPC, SPC/E, OPC, TIP5P, Bind3P) and five commonly used force fields (GAFFv1, GAFFv2, CGenFF, Parsley and SwissParam) are considered. The results highlight a significant sensitivity of host-guest binding enthalpies to the chosen force field and water model. Notably, TIP3P and its derivative Bind3P emerge as the top-performing water models for this specific host-guest system. Furthermore, the study reveals that the predictive accuracy is generally better for aliphatic compounds compared to aromatic ones, emphasizing the ongoing challenge of incorporating aromaticity accurately into these simplified force fields. This literature review synthesizes and presents these findings to contribute valuable insights into the complex relationships between force fields, water models, and the accurate prediction of thermodynamic properties in the realm of host-guest systems[24].

In 2012, Gay studied to investigate the applicability of molecular dynamics using the United-Atom approach implemented in the DLPOLY2 software package. The focus was on elucidating polymer behaviors such as melting point temperature, glass transition temperature, crystallization, and density. Additionally, the study aimed to evaluate the impact of factors such as side group length, branching, chemical structure, polymer length, and temperature on these phenomena. The validated model resulting from this investigation is intended to serve as a reliable tool for predicting the properties of polymeric systems. The specific context for applying molecular dynamics simulations was within the domain of polymers relevant to the running shoe industry[25].

In 2011, Gonzalez, et al., aimed to provide a comprehensive introduction to the field of Molecular Dynamics (MD) for those who are not experts in the subject. MD simulations involve creating a trajectory in phase space by numerically solving the classical equations of motion for a system containing N particles. The review covers fundamental concepts necessary for understanding MD, outlines key elements for conducting simulations, and discusses the information that can be derived from them. The exploration begins by

defining a force field, elucidating the components of a classical force field, detailing the optimization of potential parameters, and examining popular force fields currently in use, along with ongoing research to enhance them. The subsequent section introduces the MD technique, offering a general overview of the primary algorithms employed for integrating equations of motion, calculating long-range forces, working with different thermodynamic ensembles, and reducing computational time. Finally, the review briefly touches upon the main properties that can be computed from an MD trajectory[26].

Photosystem II constitutes a sophisticated protein-cofactor apparatus responsible for catalyzing the splitting of water molecules into molecular oxygen, protons, and electrons. Utilizing all-atom molecular dynamics simulations holds the potential to enhance our comprehensive understanding of the workings of photosystem II. In 2018, Adam, et al., conducted dependable simulations at the all-atom level, precise force field parameters for the cofactor molecules are essential. In this study, we provide CHARMM bonded and non-bonded parameters specifically designed for the iron-containing cofactors found in photosystem II. These cofactors include a six-coordinated heme moiety with coordination by two histidine groups and a non-heme iron complex coordinated by bicarbonate and four histidines. The force field parameters presented in this work demonstrate water interaction energies and geometries that align well with the quantum mechanical target data, ensuring the reliability of the simulations[27].

The progressive age-related stiffening and reduced proteolytic digestibility of the extracellular matrix (ECM) are attributed to an increased concentration of advanced glycation end products (AGEs), particularly within collagen molecules. This stiffening is considered a significant factor in various age-related diseases like osteoporosis and cardiovascular disease. However, the specific locations of covalently cross-linked AGE formation within collagen and their impact on collagen's physical and biochemical properties are not well understood. Using molecular dynamics simulations, in 2016, Collier, et al., identified preferential sites for the formation of two lysine-arginine derived AGEs, glucosepane and DOGDIC. These identified sites, particularly in close proximity to the Matrix Metalloproteinase-1 (MMP1) binding site, suggest potential disruption of collagen degradation. Dynamic analysis techniques were employed to explain the site specificity of these AGE cross-links, and steered molecular dynamics were used to investigate the mechanical properties of collagen in the presence of these cross-links. Furthermore, the study explored the influence of the collagen sequence on its mechanical properties due to its heterogeneous response to applied loads. To enhance understanding, a Homo sapiens homology model was developed from the stable Rattus norvegicus crystal structure, and a novel simulation technique was implemented to determine the orientations of collagen molecules within a fibril, which is currently beyond the resolution limit of

experimental techniques[28].

MD Analysis is a Python-based, object-oriented library designed for the structural and temporal analysis of molecular dynamics (MD) simulation trajectories and individual protein structures. The library incorporates some performance-critical code in C to optimize its functionality. Leveraging the robust NumPy package, MD Analysis efficiently exposes trajectory data as NumPy arrays, ensuring swift and effective data manipulation. It has undergone testing on systems comprising millions of particles and offers compatibility with various simulation packages, including CHARMM, Gromacs, Amber, and NAMD, as well as the Protein Data Bank format for both reading and writing. MDAnalysis supports atom selection using a syntax reminiscent of CHARMM's powerful selection commands. This library caters to both novice and experienced programmers, enabling them to quickly develop their analytical tools and access trajectory data for interactive and explorative analysis. MDAnalysis has demonstrated compatibility with most Unix-based platforms, such as Linux and Mac OS X[29].

In 2020, Fan, et al., utilized four distinct force field parametrizations: standard OPLS-AA with transferable charges, OPLS-AA with non-transferable CM1A charges, and CHARMM. To validate octanol parameters for OPLS-AA, GAFF, and CHARMM, the researcher compared their density as a function of temperature with experimental values. The partition coefficients were determined from solvation-free energy calculations for the compounds in water, pure octanol ("dry"), or octanol with 27 mol water dissolved ("wet"). Solvation-free energies were computed using thermodynamic integration (TI) and the Multistate Bennett Acceptance Ratio (MBAR) with uncorrelated samples generated through a protocol involving 5-ns windowed alchemical free energy perturbation (FEP) calculations using the Gromacs molecular dynamics package.

A new measure of convergence, based on the analysis of forward and time-reversed trajectories, was introduced to quantify the equilibration of FEP simulation sets by Fan, et al., in 2020. The accuracy of logPow predictions was assessed using descriptive statistical measures such as the root mean square error (RMSE) compared to experimental values. Discarding the initial 1 ns of each 5-ns window as an equilibration phase significantly improved the RMSE for GAFF data by up to 0.8 log units, while the impact on other datasets varied. CGenFF demonstrated the best prediction overall, achieving an RMSE of 1.2 log units for eight molecules. However, it is noteworthy that the current CGenFF workflow for Gromacs does not generate files for certain halogen-containing compounds. Across all eleven compounds, GAFF yielded an RMSE of 1.5. The inclusion of a mixed water/octanol solvent had a nuanced effect, slightly decreasing accuracy for CGenFF and GAFF, and slightly increasing it for OPLS-AA. Notably, GAFF and OPLS-AA results exhibited a systematic error with molecules appearing excessively hydrophobic,

whereas CGenFF displayed a more balanced performance, at least within this limited dataset[30].

TLR7 and TLR8, essential components of the Toll-like receptor family, play pivotal roles in innate immunity signaling pathways, rendering them attractive targets for therapeutic intervention in various diseases, including infections and cancer. Despite the high sequence homology between TLR7 and TLR8, their distinct responses to small molecule binding have garnered significant interest. To unravel the mechanisms behind the selective behavior of small molecule modulators targeting TLR7 and TLR8, molecular dynamic simulations were conducted on three imidazoquinoline derivatives—Resiquimod (R), Hybrid-2 (H), and Gardiquimod (G)—each acting as selective agonists for TLR7 and TLR8. Analysis of the MD trajectories revealed that, in the TLR7-R and TLR7-G complexes, the chains forming the TLR7 dimer tended to maintain an "open" conformation, whereas other systems retained a closed formation. Conformational deviations in the agonists R, H, and G were primarily observed in the aliphatic tail. Furthermore, the study aimed to quantify the selectivity between TLR7 and TLR8 through binding free energies using the MM-GBSA method. The results indicated that the three selected modulators exhibited greater favorability for TLR7 over TLR8, with the order of strength being H, R, and G, aligning well with experimental data. In TLR7, the flexible and hydrophobic aliphatic side chain of H displayed stronger van der Waals interactions with V351 and F351, in contrast to TLR8, where it only interacted with Y353. As expected, the positively charged side chain of G exhibited less favorable interactions with I585 of TLR7 and V573 of TLR8, explaining its weak agonistic effects on both receptors. All three imidazoquinoline derivatives demonstrated the ability to form stable hydrogen bonds with D555 of TLR7 and the corresponding D543 of TLR8. In summary, the comprehensive 400ns MD studies shed light on potential selectivity mechanisms of agonists towards TLR7 and TLR8, highlighting van der Waals interactions as the driving force for agonist binding by Wang, et al., in 2022. This insight provides valuable information for designing more potent and selective modulators that align with the hydrophobic nature of the binding pocket[31].

The cyclin-dependent kinases 4/6 (CDK4/6) play a pivotal role in regulating the cell cycle, and their aberrant activity can result in uncontrolled cell proliferation, contributing to cancer development. While three CDK4/6 inhibitors have received FDA approval, their use is associated with various side effects. Hence, there is a need to design and develop new CDK4/6 inhibitors with enhanced potency and safety. In 2023, Singh, et al., employed an integrated in-silico approach coupled with steered molecular dynamics (SMD) simulations to identify promising molecules for the development of novel CDK4/6 inhibitors. Among thirty-two 3-methyleneisoindolin-1-one molecules, the re-

searcher identified M18, M24, and M32 as top candidates based on their interaction energies. Through extensive 250 ns molecular dynamics simulations and thermodynamic free energy assessments, M24 emerged as the most promising molecule compared to palbociclib. In SMD, M24 exhibited a higher external pulling force requirement for dissociation from the CDK4 binding pocket compared to palbociclib. The elevated pulling force for M24 suggests a stronger binding affinity with CDK4, positioning M24 as a potential lead structure for the development of effective CDK4 inhibitors[32].

Chapter 3
METHODOLOGY

3 Methodology

3.1 CHARMM Force Field

CHARMM (Chemistry at Havard Macro Molecular Mechanics) is a general, flexible, highly versatile, and widely used molecular simulation program that applies classical and quantum mechanical energy functions for molecular systems. This field includes parameters for proteins, lipids, nucleic acids, carbohydrates, virus membrane protein, and macro-molecular assembles. From the experimental work and high-level quantum mechanical solution, the sources of force field functions and parameters are derived. The estimation of energy within the error is really very correlated with the experimental data. This is the beauty of the force field concept used in macromolecules and makes life much easier and computable. However, this field is insignificant in areas like condensed phase studies such as drug-protein interactions and high dielectric medium. The main demerit of this force field is that it is unable to cover large and vast chemical spaces covered by drug-like molecules to prepare the design of utility.

The functional form and parameter sets are used to describe the potential energy of a system of atoms. It is the negative gradient of a scalar potential and potential energy surface function in chemistry.The CHARMM potential energy function is used to evaluate net potential energy $U(r)$, where r denotes the cartesian co-ordinate of the system, which is expressed as,

$$U_{total} = U_{internal} + U_{external}$$

where $U_{internal}$ includes bonded terms (covalently linked atoms) and is called intramolecular potential energy and $U_{external}$ includes non-bonded terms (long-range electrostatic and van der Waals forces) which is also called intermolecular potential energy.

$$U_{internal} = \sum_{\text{bonds}} k_x(x - x_0)^2 + \sum_{\text{bond angles}} k_\beta(\beta - \beta_0)^2 + \sum_{\text{dihedrals}} k_\alpha[1 + \cos(n\alpha - \delta) \quad (1)$$

where, x is bond length, x_o is equilibrium bond length, β is actual bond angle, β_0 is equilibrium bond angle, α is torsional angle, k_x is bonding force constant, k_α is bending force constant, δ is phase angle in cosine series and n is dihedral multiplicity. The average interaction of each of the bonded atoms is found from crystallographic data within such a minimum error and depicted as x_o and β_0. The values like k_x, k_α, and k_β are defined

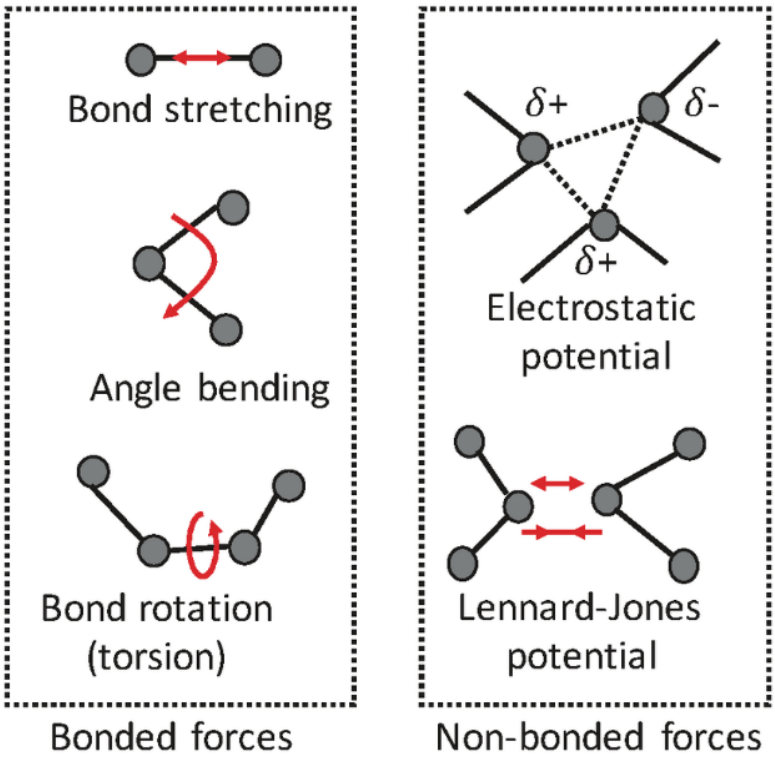


Figure 6: Parameters of CHARMM Force Field including bond stretching, angle bending, bond torsion, Lennard-Jones potential and electrostatic potential [33]

from the knowledge already known from the spectroscopy and other experiments.

Similarly, $U_{external}$ is the sum of electrostatic Coulomb potential and Lennard Jones potential (van der Waals interaction) which is expressed as,

$$U_{external} = \sum_{\text{nonbonded atom pairs}} \left(4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\epsilon r_{ij}} \right) \tag{2}$$

where q_i and q_j are the partial atomic charge of atom i and j respectively, ϵ_{ij} is the well depth, σ_{ij} is the radius in the Lennard-Jones 6-12 terms used to treat the van der Waals interaction and r_{ij} is the distance between i and j atoms [34].

The CHARMM (Chemistry at Harvard Macromolecular Mechanics) force field is a widely used computational method for simulating the structure and dynamics of biological macromolecules, such as proteins, nucleic acids, lipids, and carbohydrates. Developed by Martin Karplus and his colleagues at Harvard University, CHARMM represents one of the most established and extensively validated force fields in computational biophysics.

It incorporates empirical parameters derived from experimental data to describe the interactions between atoms and accurately predict the behavior of biomolecular systems at the atomic level.

The CHARMM force field encompasses parameters for bond lengths, bond angles, dihedral angles, and non-bonded interactions, including van der Waals forces and electrostatic interactions. These parameters are derived from quantum mechanical calculations, experimental data, and statistical analyses of molecular structures and dynamics. By accounting for both bonded and non-bonded interactions, CHARMM provides a comprehensive description of the forces governing the behavior of biomolecules in various environments, such as aqueous solutions, lipid membranes, and protein-ligand complexes.

One of the key strengths of the CHARMM force field is its versatility and applicability to a wide range of biomolecular systems. It has been extensively optimized and validated for simulating diverse biological macromolecules, from small peptides to large protein complexes and nucleic acid assemblies. Additionally, CHARMM is compatible with various molecular dynamics simulation software packages, allowing researchers to perform simulations using different computational platforms and methodologies.

Over the years, the CHARMM force field has undergone several revisions and updates to improve its accuracy and predictive power. These updates often involve refining parameter sets, incorporating new experimental data, and addressing limitations identified through benchmarking studies and empirical observations. As a result, the latest versions of the CHARMM force field, such as CHARMM36 and CHARMM22*, offer improved performance and reliability for simulating complex biomolecular systems and exploring their structure-function relationships.

In conclusion, the CHARMM force field represents a robust and versatile computational tool for studying the structure, dynamics, and interactions of biological macromolecules. Its empirical parameters, derived from experimental data and theoretical calculations, enable accurate simulations of biomolecular systems across different scales and complexities. By providing insights into the atomic-level details of biomolecular behavior, CHARMM contributes to our understanding of fundamental biological processes and informs drug discovery, protein engineering, and other biomedical applications.

3.2 Protein-Binding Kinetics

Let us consider "A" represents the drug and "B'' represents the protein. For reversible protein-drug binding, the law of mass action is applied as follows:

$$A + B \rightleftharpoons AB$$

At equilibrium,

$$K_a = \frac{[AB]}{[A][B]}$$
$$[AB] = K_a[A][B]$$

where, $[A]$ = active mass of free protein, $[B]$ = active mass of free drug, $[AB]$= active mass of free-drug complex and K_a= association rate constant.

If "A_T" represents the net concentration of protein present, unbound and bound, then

$$A_T = [AB] + [A]$$

If "x" represents the number of moles of drug bound to total moles of protein, then

$$x = \frac{AB}{A_T}$$
$$x = \frac{AB}{[AB] + [A]}$$
$$x = \frac{K_a[A][B]}{K_a[A][B] + [A]}$$
$$x = \frac{K_a[B]}{K_a[B] + 1}$$

The above equation is applicable when there is only one binding site on the protein [35].

Protein-drug kinetics refers to the study of the rates and mechanisms by which drugs interact with proteins in biological systems. Understanding these kinetics is crucial for elucidating drug action, optimizing dosage regimens, and predicting therapeutic outcomes. The kinetics of protein-drug interactions involve several key processes, including drug binding, dissociation, and the formation of drug-protein complexes, which are influenced by factors such as drug concentration, protein structure, and environmental conditions.

The kinetics of protein-drug binding are typically described by mathematical models, such as the binding equation derived from the law of mass action. This equation quan-

tifies the rate of association between the drug and protein, governed by the association rate constant (kon), which represents the speed at which the drug binds to the protein's binding site. Conversely, the dissociation rate constant (koff) characterizes the rate at which the drug-protein complex dissociates, releasing the drug back into solution.

The equilibrium between drug binding and dissociation is described by the dissociation constant (KD), which represents the ratio of the dissociation and association rate constants. A lower KD value indicates tighter binding and higher affinity between the drug and protein, while a higher KD value suggests weaker binding and lower affinity. The kinetics of protein-drug binding can be influenced by various factors, including temperature, pH, solvent composition, and the presence of other molecules, which may alter the drug's binding affinity and kinetics in vivo.

Moreover, the kinetics of protein-drug interactions play a critical role in determining the drug's pharmacokinetics, including its absorption, distribution, metabolism, and excretion (ADME). For instance, drugs with rapid association kinetics may exhibit faster onset of action, while those with slow dissociation kinetics may have prolonged effects. Additionally, the kinetics of protein-drug binding can impact the drug's bioavailability, tissue distribution, and clearance rates, affecting its efficacy and safety profiles in vivo.

In summary, protein-drug kinetics encompass the dynamic processes of drug binding, dissociation, and complex formation, which are essential for understanding drug action and pharmacokinetics. By elucidating the mechanisms underlying these kinetics, researchers can optimize drug design, dosage regimens, and therapeutic strategies to enhance drug efficacy and minimize adverse effects. Moreover, advances in computational modeling and experimental techniques continue to improve our ability to characterize and predict protein-drug kinetics, facilitating drug discovery and development efforts.

3.3 Computational Approach

Our computational work is entirely based upon the installation and working of pyMoL-1.20 and Wolfram Mathematica softwares. The installed pyMoL will provide us all the essential tools for setting up, running and analyzing 1GIT protein interaction with ligand GDP. The generation of a complete system using pyMoL, wolfram mathematica for charmm force field and binding mechanism starting from their protein structure files to the analysis of the results are shown in flowcharts (1) and (2).

3.3.1 CHARMM Force Field Parameter of GDP

The systematic analyzing process for bond length, bond angle, dihedral angle, Lennard-Jones potential and electrostatic interactions of atoms of ligand GDP are studied by plotting their corresponding energy profile diagrams which is shown in Figure (7).

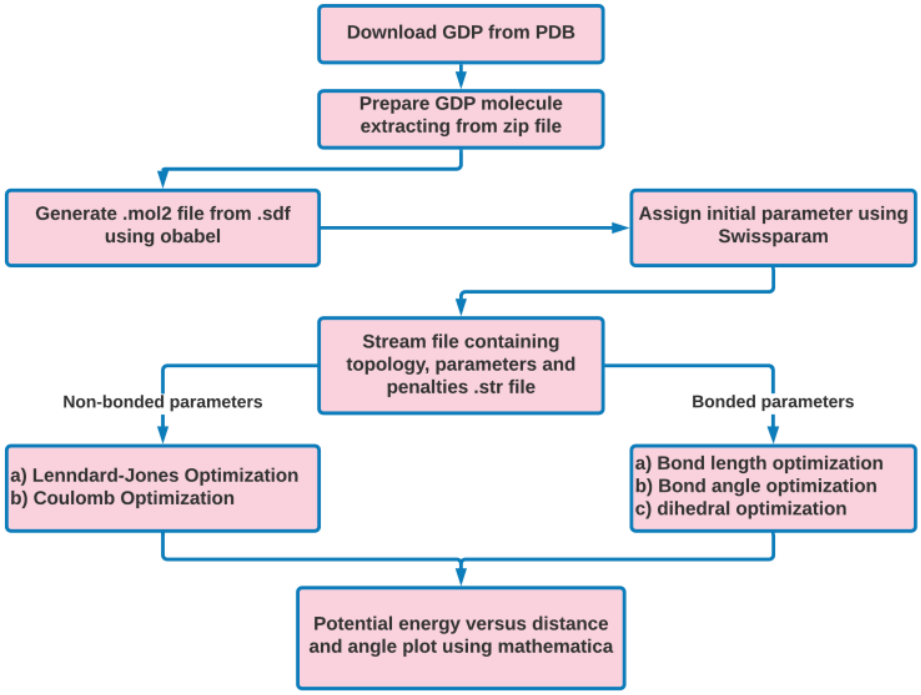


Figure 7: Flow chart showing the steps followed for force field analysis of GDP

3.3.2 Binding Mechanism of 1GIT Protein with GDP

The stepwise process using pyMoL tool for the perfect visualization of different structure of 1GIT protein, GDP, their amino acids, polar bonds between atoms of GDP and amino acids of protein and the bond length measurement in angstrom range are shown in figure (8).

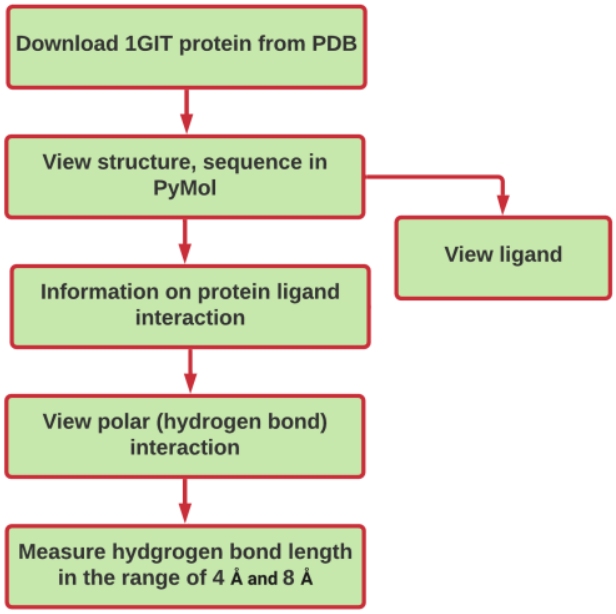


Figure 8: Flow chart showing the steps followed for binding mechanism of GDP with 1GIT protein

Chapter 4

RESULT AND DISCUSSION

4 Result and Discussion

4.1 Bond Length Energy

In CHARMM force fields, bond length energy represents the energy associated with stretching or compressing chemical bonds between atoms. These force fields employ harmonic potentials to model bond stretching, assuming that the bond behaves like a simple spring with a characteristic equilibrium length. When atoms are separated beyond or compressed closer than this equilibrium distance, the bond length energy increases, reflecting the resistance to bond distortion. The parameters defining bond length energy in CHARMM force fields are derived from experimental data, quantum mechanical calculations, and empirical fitting to reproduce the observed bond lengths in various molecular systems accurately.

The accuracy of bond length energy parameters in CHARMM force fields is crucial for capturing the structural and dynamic properties of biomolecules in molecular simulations. Deviations from the correct bond lengths can lead to distorted molecular geometries and inaccurate representations of molecular interactions and energetics. Therefore, extensive validation against experimental data and benchmarking studies is essential to ensure that CHARMM force fields accurately reproduce bond lengths and capture the correct balance of forces governing bond stretching in diverse biological macromolecules and chemical environments.

Table 4.1: Parameters for Bond Length Interactions

S.N.	Atom Name		K_x (kcal/mole/Å^2)	Equilibrium Bond Length (x_0) (Å)
1	C=O	NC=C	439.714	1.370
2	N=C	C5	478.144	1.351
3	C=O	N=C	725.204	1.290

We used the CHARMM force field to describe the time evolution of bond lengths, bond angles, torsional angles, van der Waals, and electrostatic interactions. The different kinds of force field methods consist of various numbers of terms in the expansion using Taylor series expansion. The higher expansions such as cubic and quartic are neglected due to the incorrect behavior with complicated results at large bond lengths. Initially using Hook's law, when the bond is stretched, it becomes a spring. The bond energy became low (near to zero) at the equilibrium distances such as 1·37Å, 1·351Å and 1·29Å for atom pairs C=O, NC=C and N=C, C5 and C=O, N=C respectively. The C=O,

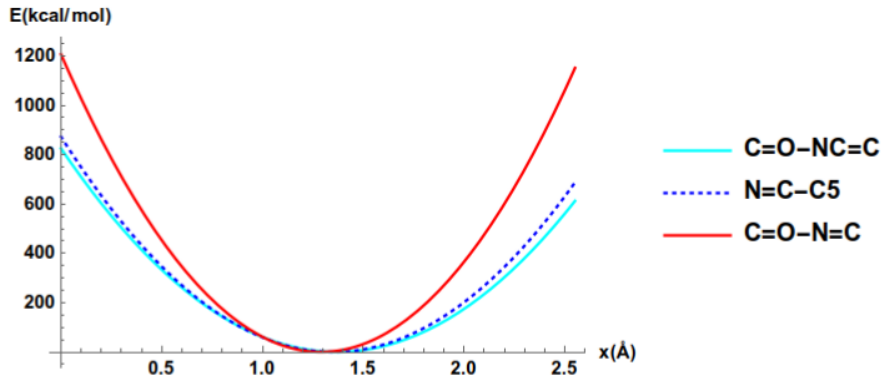


Figure 9: Energy Profiles for Bond Energy versus Bond Length

N=C curve was found bending more inside due to the higher value of force constant in comparison to other atom pairs. The bond energy was seen increased either we increasing or decreasing the bond lengh (x) because the squared term $(x - x_0)^2$ makes the energy positive. The bond energy went quadratically when stretched or compressed the bond. The bond terms were modeled as harmonic oscillators (no bond cleavage). The Morse potential is guessed to be used for long bond length in terms of energy as it has a kind of flattened out. For long bond lengths, the force becomes so small that the optimization will be slow. The parabolic curve was observed in energy profiles of bond energy versus bond length as shown in Figure (9).

4.2 Bond Angle Energy

Bond angle energy in CHARMM force fields represents the energy associated with the bending or stretching of chemical bonds in a molecule. Just like bond length energy, the bond angle energy is also described using harmonic potentials, assuming that bond angles behave like springs with characteristic equilibrium angles. When the angle between bonded atoms deviates from this equilibrium value, the bond angle energy increases, reflecting the resistance to bond angle distortion. The parameters defining bond angle energy in CHARMM force fields are derived from experimental data, quantum mechanical calculations, and empirical fitting to accurately reproduce the observed bond angles in various molecular systems.

The accurate representation of bond angle energy in CHARMM force fields is crucial for maintaining the structural integrity and stability of biomolecular systems in molecular simulations. Incorrect parameters for bond angle energy can lead to distortions in molecular geometries and inaccurate descriptions of molecular conformations and inter-

actions. Therefore, thorough validation against experimental data and benchmarking studies is essential to ensure that CHARMM force fields properly capture the energetics of bond angle bending and stretching across a wide range of molecular systems and environmental conditions.

In addition to accurately representing bond angle energy, CHARMM force fields often incorporate terms to account for the anharmonicity of bond angle potentials. These anharmonic terms allow for deviations from idealized harmonic behavior, particularly at large bond angles where the assumption of linearity breaks down. By including anharmonic contributions, CHARMM force fields can more accurately describe the flexibility and conformational dynamics of biomolecules, providing a more realistic representation of their behavior in molecular simulations.

Table 4.2: Parameters for Bond Angles Interactions Used in This Work

S.N.	**Atoms**			**Force Constant (k_β) (kcal/mole/deg^2)**	**Equilibrium Bond Angle (β_0) (deg)**
1	C5	C5	N5M	83.408	111.900
2	N=C	C5	N5M	62.106	129.501
3	OR	CR	CR	71.390	108.133

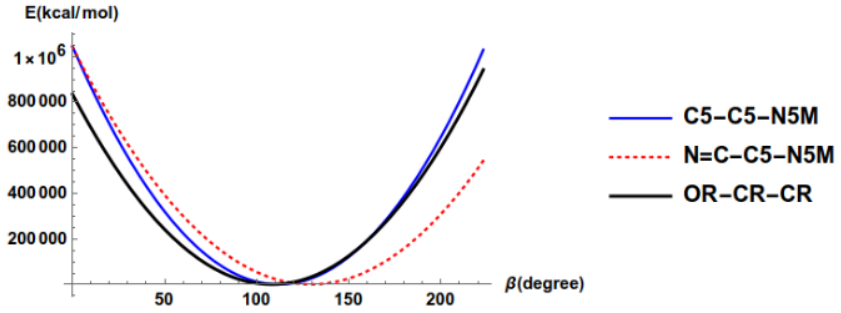


Figure 10: Energy Profiles for Bond Energy versus Bond Angles

The bond angle energy is a three-body problem that was treated very similarly to bond length energy because in this case energy was also seen to increase quadratically with the displacement of the bond angle from the equilibrium bond angle. The bond energy was seen increased either we increasing or decreasing the bond angle (β) because the squared term $(\beta-\beta_0)^2$ makes the energy positive. In this case, also, the higher expansion terms of Taylor series expansion were also neglected due to the incorrect behavior with complicated results at large bond angles. The parabolic curve was seen in the energy profiles of bond energy versus bond angles as shown in Figure (10). The variation of

bond energy with compression or expansion of bond angles was expected to depend upon the value of the force constant and the identities of three chosen atoms. The bond energy became low (near to zero) at the equilibrium bond angle 111.900°,129.501° and 108.133° for C5-C5-N5M, N=C-C5-N5M and OR-CR-CR atom pairs respectively. The parabolic curve was seen to go up and up at large angles beyond equilibrium bond angles.

4.3 Dihedrals

In CHARMM force fields, the dihedral angle energy accounts for the torsional rotation around covalent bonds in a molecule. This energy term describes the resistance to bond rotation and is modeled using a combination of Fourier series potentials. The dihedral angle energy depends on the torsional angle between four consecutive atoms, and its functional form includes terms representing the periodicity of the energy barrier, phase shifts, and amplitude coefficients. These parameters are optimized to accurately reproduce the potential energy surface and capture the energetics of molecular conformations.

An accurate representation of dihedral angle energy is crucial for describing the conformational flexibility and dynamics of biomolecules in molecular simulations. Incorrect parameters for dihedral angle potentials can lead to inaccuracies in predicting molecular conformations and transitions between different states. Therefore, extensive validation against experimental data, quantum mechanical calculations, and conformational analyses of biomolecular structures is performed to refine and optimize the parameters defining dihedral angle energy in CHARMM force fields.

56

In addition to the harmonic terms describing the torsional energy profile, CHARMM force fields often incorporate additional terms to account for non-bonded interactions between atoms involved in dihedral angles. These additional terms help to capture the steric effects and electrostatic interactions that influence the energetics of molecular conformations. By including these contributions, CHARMM force fields can provide a more comprehensive description of the energy landscape governing dihedral angle rotations, improving the accuracy and reliability of molecular simulations for studying biomolecular structure and dynamics.

Table 4.3: Parameters for Dihedral Angles Interactions

S.N.	Atoms (Dihedrals)				Bending Force Constant (k_α) (kcal/mol)	Multi-plicity (n)	Phase Angle (δ) (deg)
1	NM	C=O	C5	N5M	1.250	2	180
2	C5	N=C	C=O	NC=C	0.900	2	180
3	N5M	C5	N5M	CR	1.800	2	180

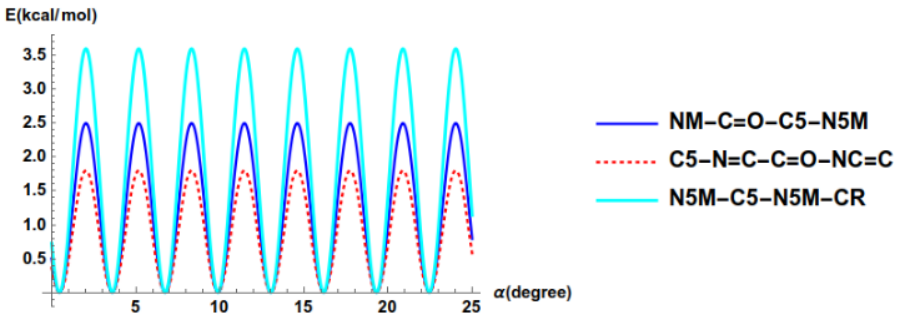


Figure 11: Energy Profiles for Dihedral Energy versus Dihedral Angles

The steric and electrostatic non-bonded interaction between two atoms (NM and N5M) connected through an intermediate bond say (C=O and C5) was expected to be captured by torsional energy. The torsional angle went far from equilibrium so that the torsional potential is not expanded as Taylor series. The small perturbation was done around some preferred angle at 180°. The energy was not seen as negative even though we frequently chopp it off at about three terms n equals 3 as the maximum. The $n\alpha$ showed that there occur n maxima and minima within certain limits of angle. The Fourier series was implemented in the formula of torsional energy due to the reason that torsion can freely rotate in molecules unless there exists a double bond or other molecules. The periodic function having an oscillatory nature i.e. sinusoidal curve was seen in the energy profile diagram of dihedral energy versus dihedral angle which is shown in the figure (11).

4.4 Lennard-Jones Potential

Lennard-Jones potentials are fundamental components of force fields used in molecular modeling, including CHARMM force fields. These potentials describe the non-bonded interactions between pairs of atoms and are based on the intermolecular forces arising from van der Waals interactions. The Lennard-Jones potential is characterized by two parameters: the van der Waals radius and the depth of the potential energy well. The parameter represents the distance at which the potential energy between two atoms is zero, while represents the energy minimum of the potential energy curve.

In CHARMM force fields, Lennard-Jones potentials are employed to model both steric repulsion and dispersion interactions between atoms. The repulsive term of the Lennard-Jones potential, which arises from the Pauli exclusion principle, prevents atomic overlap and accounts for the steric hindrance between atoms. Conversely, the attractive term, governed by the dispersion forces, describes the van der Waals attraction between atoms at distances beyond their van der Waals radii. By balancing these repulsive and attractive interactions, Lennard-Jones potentials accurately capture the intermolecular forces that influence the spatial arrangement and packing of atoms in molecular systems.

The parameters defining Lennard-Jones potentials in CHARMM force fields are typically optimized to reproduce experimental data and quantum mechanical calculations of intermolecular interactions. These parameters are specific to different atom types and are adjusted to account for variations in atomic size and electronegativity. Additionally, CHARMM force fields often include modification terms to account for long-range electrostatic interactions, such as the use of Coulombic potentials or explicit treatment of electrostatic forces using particle mesh Ewald (PME) or other methods. By incorporating Lennard-Jones potentials with optimized parameters, CHARMM force fields provide an accurate description of non-bonded interactions in molecular simulations, enabling the study of biomolecular structure, dynamics, and interactions with high fidelity.

Table 4.4: Parameters for Lennard-Jones (van der Waals) Interactions Used in This Work

S.N.	Atom/Group	Well Depth (ϵ_{ij}) (kcal/mol)	Radius in Lennard Jones (σ_{ij}) (Å)
1	NM	-0.200	1.850
2	OR	-0.152	1.770
3	O2CM	-0.120	1.700
4	HCMM	-0.022	1.320

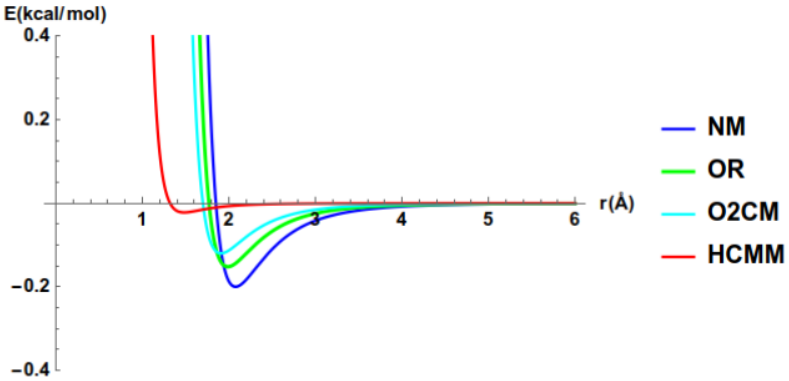


Figure 12: Energy Profiles for Lennard-Jones Energy versus Distance

The attractive and repulsive forces were seen cutoff to each other at lowest energy -0.2 kcal/mol, -0.1521 kcal/mol, -0.12 kcal/mol, and -0.022 kcal/mol for NM, OR, O2CM and HCMM atom pairs respectively. The short, intermediate, and long-range energies were seen through the interaction between electron clouds around two non-bonded atoms which are strongly repulsive, attractive, and go to zero respectively. The two close atoms collided and became repulsive. At short distances, the van der Waals curve was rised away too fast so it was assumed to going to prevent the molecule from getting as close as it might be in reality. The intermediate attraction became attractive in the 4Å range. The long-range interaction appeared to be zero because the atoms were too far apart to interact. The attraction curve was seen to be proportional to $\frac{1}{r^6}$ at intermediate to long ranges. The van der Waals terms were speeded up substantially by a more economic expression called Lennard-Jones potential. From figure (12), it was observed that the van der Waals terms die off relatively quickly i.e. inversely proportional to r_{ij}^6, and seen as cutoff around 6Å

4.5 Coulombs Potential

In CHARMM force fields, Coulomb potentials play a crucial role in describing electrostatic interactions between charged atoms or groups within a molecular system. These potentials are based on Coulomb's law, which describes the force between two charged particles as inversely proportional to the square of the distance between them. The Coulomb potential energy between two point charges is calculated as the product of their charges divided by the distance between them, with an additional term to account for the dielectric constant of the surrounding medium.

CHARMM force fields typically use Coulomb potentials to model electrostatic interac-

tions between charged atoms, such as those in ionic bonds or polar covalent bonds. The parameters defining Coulomb potentials include the charges on atoms or groups and the dielectric constant of the solvent environment. These parameters are often derived from quantum mechanical calculations, experimental data, or empirical fitting to reproduce the observed behavior of charged species in various molecular systems.

In addition to modeling electrostatic interactions between charged atoms, CHARMM force fields may incorporate modification terms to account for the long-range nature of electrostatic forces. One common approach is the use of cutoff distances or switching functions to approximate the decay of electrostatic interactions with distance. Alternatively, explicit treatment of long-range electrostatic interactions can be achieved using methods such as particle mesh Ewald (PME), which efficiently computes the electrostatic potential over large spatial scales. By accurately describing Coulomb potentials and their long-range effects, CHARMM force fields provide a realistic representation of electrostatic interactions in molecular simulations, contributing to our understanding of biomolecular structure, dynamics, and function.

Table 4.5: Parameters for Electrostatic Interactions Used in This Work

S.N.	Atom 1	Atom 2	Charge (q_i) (C)	Charge (q_j) (C)	Dielectric Constant (ϵ_r)	Constant (K)
1	C6	C5	0.423	0.202	5	332
2	C5	N3	0.202	-0.588	5	332
3	C1'	C2'	0.723	0.280	5	332
4	C5	N7	0.202	-1.050	5	332

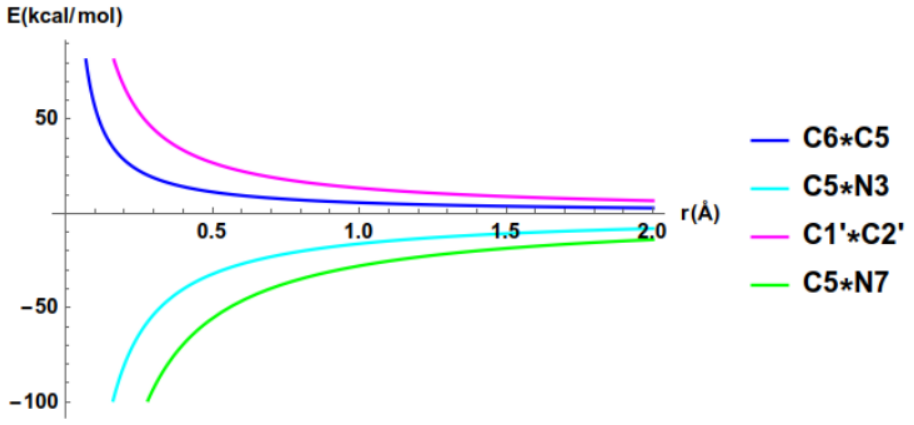


Figure 13: Energy Profiles for Electrostatic Energy versus Distance

The electrostatic interaction between two atoms (C6 and C5), (C5 and N3) ,(C1′ and

C2′) and (C5 and N7) were calculated using coulomb law which were seen only falling as $\frac{1}{r_{ij}}$. The number of non-bonded electrostatic interactions between two atoms in protein grew quadratically with their molecular size. It was observed that the electrostatic curve dies off slower i.e. inversely proportional to r_{ij}, although sometimes faster in practice, and were much harder to treat with cutoff distance. A sharp discontinuity was seen between atoms inside and outside due to the large cutoff radius. The electrostatic curve was seen not falling so rapidly in contrast to the van der Waals curve. The slowly falling quadrative curve was seen in the energy profile of electrostatic energy versus distance between two non-bonded atoms which is shown in figure (13).

4.6 Protein-Ligand Interactions at 4Å

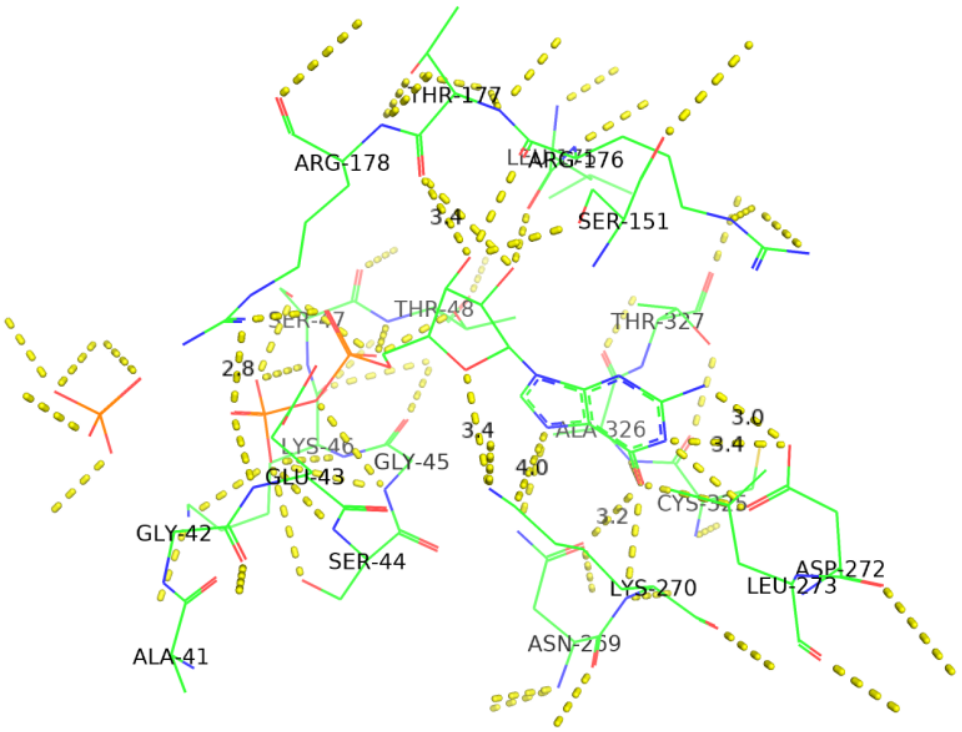


Figure 14: Amino acids residue of 1git protein forming polar bonds with the ligand at 4Å

The pyMol tool was used for 1GIT protein interaction with GDP ligand. The non-interacting amino acids and water molecules were removed from 1GIT protein to get perfect visualization of hydrogen (polar) bonds. We were able to visualize polar interactions under 4Å. The amino acids that are not attached with hydrogen bonds were termed

as non-polar interactions. The ionic and van der Waals interactions were not observed in pyMol view. The yellow dotted lines were found which are GDP ligand interactions with 1GIT protein residues called polar bonds. The measured lengths such as 2.8Å, 3.4Å, 3.2Å, 4.0Å, etc. were polar bonds between the amino acid residue of protein and GDP ligand. The GDP was designed as a non-standard amino acid and THR-177, ARG-178, ARG-176, SER-151, THR-327, THR-48, SER-47, ALA-326, CYS-326, GLY-45, LYS-46, GLU-43 were designed as standard amino acids as shown in figure (14).

4.7 Protein-Ligand Interactions at 8Å

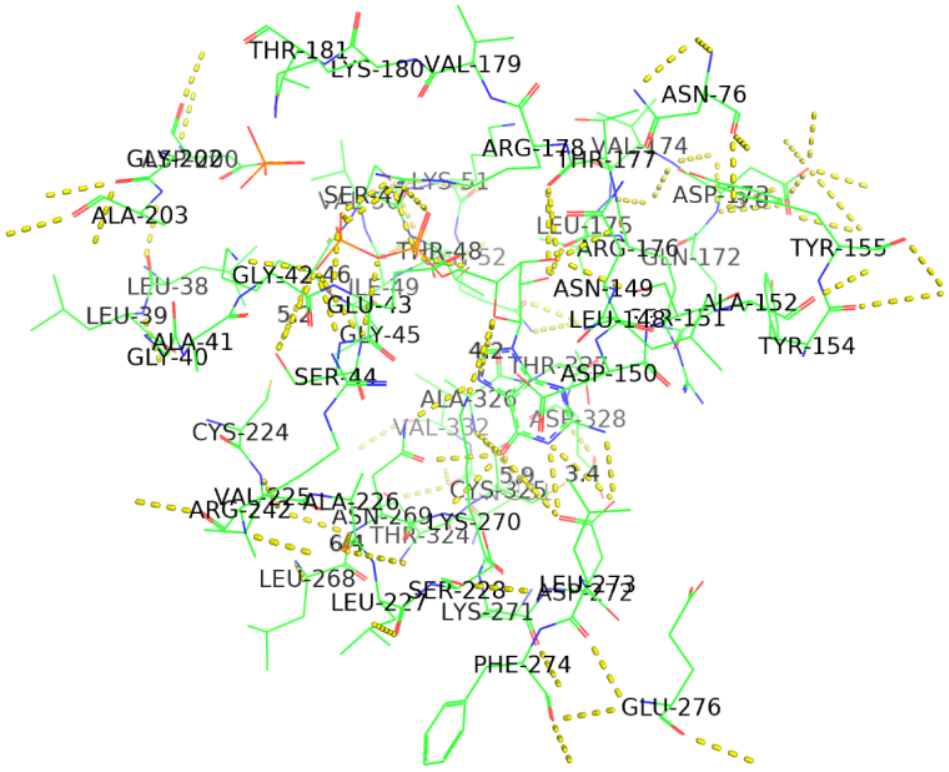


Figure 15: Amino acids residue of 1git protein forming polar bonds with the ligand at 8Å

The same process that was implemented for the protein-ligand interaction at 4Å was also applied for 8Å visualization. In comparison to 4Å, there occurred a large number of hydrogen bonds in the 8Å range. We removed the non-standard amino acids such as HOH and PO4 and only considered GDP as a protein ligand for the perfect visualization of polar bonds. The highlighted dotted yellow lines were GDP ligand interactions with 1GIT protein residue. The measured lengths such as 5.2Å, 4.2Å, 6.4Å, 5.9Å, etc. were

polar bonds between the amino acid residue of protein and GDP ligand. The amino acids such as ALA-203, ASN-76, SER-47, ASN-149, LEU-148, CYS-325, LEU-268, etc. were designed as standard amino acids as shown in figure (15).

Chapter 5
CONCLUSION

5 Conclusion

The force field parameters of GDP and its binding mechanism were examined and studied by using Wolfram Mathematica and the pyMoL tool. The CHARMM force field was used efficiently and accurately for the calculations to attack problems in molecular structure. The various data used in drawing the energy profiles of this force field components are shown in the appendix table. The 1GIT protein containing ligand was downloaded from the PDB bank and visualized in the pyMoL tool. The different measurements and analyses of this research work are concluded as follows.

- The 1GIT protein was visualized in the pyMoL tool in stereo mode such as wires, sticks, spheres, surfaces, mesh, dots, ribbon, and cartoon form.
- The bond energy was found minimum at equilibrium bond length and found to increase quadratically when we increase or decrease the bond length.
- The bond energy was found minimum at the equilibrium bond angle and rose quadratically when we increased or decreased the bond angle.
- The bond energy was found to vary sinusoidally with n maxima and minima depending upon the value of dihedral angles.
- The short, intermediate, and long-range Lennard-Jones interaction energies were found strongly repulsive, attractive, and go to zero respectively.
- The Lennard-Jones curves were found to be cut off around 6Å.
- The electrostatic curves were seen falling slowly in quadratic nature and were found harder to predict with cutoff distance.
- The polar (hydrogen) bonds were observed between atoms of GDP and amino acid residues of protein and measured in 4Åand 8Å range using the pyMoL tool.

5.1 Future Work

In this research work, the force field analysis of GDP and its polar binding with amino acid residue has been conducted computationally. In the spirit of this topic, future works may be like these:

- Researcher may continue the force field analysis of different ligands associated with respective proteins.
- The binding of macromolecule into the binding site of a ligand and its binding affinity may be estimated for structure-based drug design.

- The new compounds may be generated by exploring the knowledge of binding pocket and electrostatic properties.
- The modern computational drug design shall be carried out by using pyMoL as homology, macromolecular analysis, protein-ligand docking, pharmacophore modeling, etc.
- More research and innovation related to computational modeling, specifically virtual screening can be performed to discover effective drugs so that the governmental and non-governmental institutions can support and make policy for contribution in science and technology.

References

[1] L. Li, Y. Liu, Y. Wan, Y. Li, X. Chen, W. Zhao, and P. G. Wang, Org. Lett. **15**, 5528 (2013).

[2] \bibfield journal \bibinfo journal Biotechnol. J.\ \textbf \bibinfo volume 16,\ \bibinfo pages 2170073 (\bibinfo year 2021).

[3] S. M. Yoo, M. A. Antonyak, and R. A. Cerione, J. Biol. Chem **287**, 31462 (2012).

[4] R. J. Hoefen and B. C. Berk, J. Cell Sci. **119**, 1469 (2006).

[5] D. Brahankar and S. B. Jaiswal, *Biopharmaceutics and Pharmacokinetics: A treatise* (Vallabh Prakashan, 2009).

[6] X. Du, Y. Li, Y.-L. Xia, S.-M. Ai, J. Liang, P. Sang, X.-L. Ji, and S.-Q. Liu, Int. J. Mol. Sci. **17**, 144 (2016).

[7] L. Shargel, B. Andrew, and S. Wu-Pong, *Applied biopharmaceutics & pharmacokinetics*, Vol. 264 (Appleton & Lange Stamford, 1999).

[8] W. L. DeLano *et al.*, CCP4 Newsletter on protein crystallography **40**, 82 (2002).

[9] N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch, and G. R. Hutchison, J. Cheminformatics **3**, 1 (2011).

[10] V. Zoete, M. A. Cuendet, A. Grosdidier, and O. Michielin, J. Comput. Chem. **32**, 2359 (2011).

[11] M. Bugnon, M. Goullieux, U. F. Röhrig, M. A. Perez, A. Daina, O. Michielin, and V. Zoete, Journal of Chemical Information and Modeling **63**, 6469 (2023).

[12] S. Kim, P. A. Thiessen, E. E. Bolton, J. Chen, G. Fu, A. Gindulyte, L. Han, J. He, S. He, B. A. Shoemaker, *et al.*, Nucleic acids research **44**, D1202 (2016).

[13] S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, Q. Li, B. A. Shoemaker, P. A. Thiessen, B. Yu, *et al.*, Nucleic acids research **51**, D1373 (2023).

[14] J. Phillips, T. Isgro, M. Sotomayor, and E. Villa, (2003).

[15] H. Sigel, E. M. Bianchi, N. A. Corfù, Y. Kinjo, R. Tribolet, and R. B. Martin, Chem. Eur. J. **7**, 3729 (2001).

[16] J. Kirchmair, C. Laggner, G. Wolber, and T. Langer, J. Chem. Inf. Model **45**, 422 (2005).

[17] W.-H. Lee, N.-S. Han, Y.-C. Park, and J.-H. Seo, Bioresour. Technol. **100**, 6143 (2009).

[18] A. H. Sayera, E. Bluma, D. T. Major, A. Vardi-Kilshtaina, B. L. Hevronia, B. Fischera, and B. Fischer, .

[19] Y. Yu, I. A. Fursule, L. C. Mills, D. L. Englert, B. J. Berron, and C. M. Payne, J. Mol. Graph. **72**, 32 (2017).

[20] A. R. Lee, Y. Gan, N. Xie, V. R. Ramnarine, J. M. Lovnicki, and X. Dong, Cancer Sci. **110**, 245 (2019).

[21] T. R. Lamichhane and M. P. Ghimire, Chemrxiv.org (2020) .

[22] C. G. Mayne, J. Saam, K. Schulten, E. Tajkhorshid, and J. C. Gumbart, Journal of computational chemistry **34**, 2757 (2013).

[23] S. S. Çınaroğlu and P. C. Biggin, The Journal of Physical Chemistry B **125**, 1558 (2021).

[24] K. Sonibare, L. Rathnayaka, and L. Zhang, Construction and Building Materials **236**, 117577 (2020).

[25] J. Gay, (2012).

[26] M. A. González, École thématique de la Société Française de la Neutronique **12**, 169 (2011).

[27] S. Adam, M. Knapp-Mohammady, J. Yi, and A.-N. Bondar, Journal of Computational Chemistry **39**, 7 (2018).

[28] T. Collier, *Fully Atomistic Modelling of Collagen Cross-linking*, Ph.D. thesis, UCL (University College London) (2016).

[29] N. Michaud-Agrawal, E. J. Denning, T. B. Woolf, and O. Beckstein, Journal of computational chemistry **32**, 2319 (2011).

[30] S. Fan, B. I. Iorga, and O. Beckstein, Journal of computer-aided molecular design **34**, 543 (2020).

[31] X. Wang, Y. Chen, S. Zhang, and J. N. Deng, Plos one **17**, e0260565 (2022).

[32] R. Singh and R. Purohit, Computer Methods and Programs in Biomedicine **231**, 107367 (2023).

[33] H. Waidyasooriya, M. Hariyama, and K. Kasahara, \bibfield journal \bibinfo journal Int. J. Networked Distrib.\ \textbf \bibinfo volume 5,\ \bibinfo pages 52 (\bibinfo year 2017).

[34] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, *et al.*, J. Comput. Chem. **31**, 671 (2010).

[35] D. Brahmankar and S. B. Jaiswal, *Biopharmaceutics and pharmacokinetics: A treatise* (Vallabh prakashan, 2005).

Appendix

Table A.1: Parameters for Bond Length Interactions Used in This Work

S.N.	Atom Name		K_x	Equilibrium Bond Length (x_0)
1	PO4	OR	377.319	1.63
2	PO4	O2CM	597.032	1.51
3	CR	CR	306.432	1.508
4	CR	HCMM	342.991	1.093
5	CR	OR	363.214	1.418
6	OR	HOR	560.905	0.972
7	CR	N5M	302.258	1.5
8	NM	C=O	544.641	1.322
9	OR	HOCO	532.766	0.981
10	C=O	NC=C	439.714	1.37
11	C=O	N=C	725.204	1.29
12	NC=C	HNCO	473.25	1.018
13	N=C	C5	478.144	1.351
14	C5	C5	401.068	1.374
15	C5	C=O	410.567	1.413
16	C=O	O=C	931.963	1.222
17	C5	N5M	491.098	1.345
18	C5	HCMM	396.246	1.08

Table A.2: Parameters for Bond Angles Interactions Used in This Work

S.N.	Atoms			Force Constant (k_β)	Equilibrium Bond Angle (β_0)
1	C=O	NM	C=O	94.851	106.821
2	NM	C=O	N=C	71.966	120
3	NM	C=O	NC=C	71.966	120
4	N=C	C=O	NC=C	60.739	128.078
5	C=O	N=C	C5	95.211	106.641
6	N=C	C5	C5	62.106	129.501
7	N=C	C5	N5M	62.106	129.501
8	C5	C5	N5M	83.408	111.9
9	C5	C5	C=O	59.516	125.468
10	C=O	C5	N5M	59.516	125.468
11	NM	C=O	C5	71.966	120
12	NM	C=O	O=C	81.249	129.349
13	C5	C=O	O=C	68.727	132.047
14	C5	N5M	C5	88.878	109.421
15	N5M	C5	N5M	89.597	113.179
16	N5M	C5	HCMM	42.028	123.407
17	C5	N5M	CR	42.028	123.407
18	OR	PO4	O2CM	108.02	109.688
19	OR	PO4	OR	127.307	99.311
20	O2CM	PO4	O2CM	89.813	122.857
21	OR	CR	CR	71.39	108.133
22	OR	CR	HCMM	56.205	108.577
23	CR	CR	HCMM	45.77	110.549
24	HCMM	CR	HCMM	37.134	108.836
25	PO4	OR	CR	78.802	115.581
26	CR	CR	CR	61.243	109.608
27	CR	OR	CR	86.143	106.926
28	CR	OR	HOR	57.069	106.503
29	N5M	CR	OR	71.966	109.5
30	N5M	CR	CR	71.966	109.5
31	N5M	CR	HCMM	71.966	109.5
32	PO4	OR	HOCO	43.683	118.533
33	C=O	NC=C	HNCO	50.376	114.808
34	HNCO	NC=C	HNCO	40.301	109.16
35	PO4	OR	PO4	55.917	129.375

Table A.3: Parameters for Dihedral Angles Interactions Used in This Work

S.N.	Atoms (Dihedrals)				Bending Force Constant (k_α)	Multi-plicity (n)	Phase Angle (δ)
1	NM	C=O	N=C	C5	8	2	180
2	NM	C=O	NC=C	HNCO	1.95	2	180
3	NM	C=O	C5	C5	1.25	2	180
4	NM	C=O	C5	N5M	1.25	2	180
5	C=O	NM	C=O	C5	1.8	2	180
6	C=O	NM	C=O	O=C	1.8	2	180
7	C=O	N=C	C5	C5	0.9	2	180
8	C=O	N=C	C5	N5M	0.9	2	180
9	N=C	C=O	NM	C=O	1.8	2	180
10	N=C	C=O	NC=C	HNCO	0.748	1	0
11	N=C	C=O	NC=C	HNCO	2.184	2	180
12	N=C	C=O	NC=C	HNCO	-0.208	3	0
13	N=C	C5	C5	C=O	0	1	0
14	N=C	C5	C5	N5M	0	1	0
15	N=C	C5	N5M	C5	0	1	0
16	N=C	C5	N5M	CR	0	1	0
17	C5	N=C	C=O	NC=C	0.9	2	180
18	C5	C5	C=O	O=C	1.25	2	180
19	C5	C5	N5M	C5	1.8	2	180
20	C5	N5M	C5	N5M	1.8	2	180
21	C5	N5M	C5	HCMM	1.8	2	180
22	C5	N5M	CR	OR	0	1	0
23	C5	N5M	CR	CR	0	1	0
24	C5	N5M	CR	HCMM	0	1	0
25	C5	C5	N5M	CR	1.8	2	180
26	C=O	NM	C=O	NC=C	1.8	2	180
27	C=O	C5	C5	N5M	0	1	0
28	C=O	C5	N5M	C5	0	1	0
29	N5M	C5	C5	N5M	3.5	2	180
30	N5M	C5	C=O	O=C	1.25	2	180
31	N5M	C5	N5M	CR	1.8	2	180
32	N5M	CR	OR	CR	0.1	3	0
33	N5M	CR	CR	CR	0.15	3	0
34	N5M	CR	CR	OR	0.15	3	0
35	N5M	CR	CR	HCMM	0.15	3	0
36	PO4	OR	CR	CR	0.1	3	0
37	PO4	OR	CR	HCMM	0.03	3	0
38	PO4	OR	PO4	OR	0.325	3	0
39	PO4	OR	PO4	O2CM	0.325	3	0
40	CR	OR	PO4	O2CM	0.603	1	0
41	CR	OR	PO4	O2CM	0.457	2	180
42	CR	OR	PO4	O2CM	0.306	3	0
43	CR	OR	PO4	OR	0.389	3	0

44	CR	CR	OR	CR	-0.341	1	0
45	CR	CR	OR	CR	0.378	2	180
46	CR	CR	OR	CR	0.378	3	0
47	CR	CR	CR	OR	-0.344	1	0
48	CR	CR	CR	OR	0.879	2	180
49	CR	CR	CR	OR	0.238	3	0
50	CR	CR	CR	CR	0.051	1	0
51	CR	CR	CR	CR	0.341	2	180
52	CR	CR	CR	CR	0.166	3	0
53	CR	CR	CR	HCMM	0.32	1	0
54	CR	CR	CR	HCMM	-0.315	2	180
55	CR	CR	CR	HCMM	0.132	3	0
56	OR	CR	CR	OR	0.204	1	0
57	OR	CR	CR	OR	0.699	2	180
58	OR	CR	CR	OR	0.48	3	0
59	OR	CR	CR	HCMM	-0.327	1	0
60	OR	CR	CR	HCMM	0.536	2	180
61	OR	CR	CR	HCMM	0.14	3	0
62	CR	OR	CR	HCMM	0.285	1	0
63	CR	OR	CR	HCMM	0.16	2	180
64	CR	OR	CR	HCMM	0.285	3	0
65	CR	CR	OR	HOR	0.135	2	180
66	CR	CR	OR	HOR	0.118	3	0
67	CR	N5M	C5	HCMM	1.8	2	180
68	O2CM	PO4	OR	HOCO	-2.946	1	0
69	O2CM	PO4	OR	HOCO	-1.666	2	180
70	O2CM	PO4	OR	HOCO	0.145	3	0
71	OR	PO4	OR	HOCO	-1.605	1	0
72	OR	PO4	OR	HOCO	-3.811	2	180
73	OR	PO4	OR	HOCO	0.532	3	0
74	HCMM	CR	CR	HCMM	0.142	1	0
75	HCMM	CR	CR	HCMM	-0.693	2	180
76	HCMM	CR	CR	HCMM	0.157	3	0
77	HCMM	CR	OR	HOR	0.298	1	0
78	HCMM	CR	OR	HOR	-0.138	2	180
79	HCMM	CR	OR	HOR	0.173	3	0

Table A.4: Parameters for Lennard-Jones (van der Waals) Interactions Used in This Work

S.N.	Atom/Group	Well Depth (ϵ_{ij})	Radius in Lennard Jones (σ_{ij})
1	NM	-0.2	1.85
2	C=O	-0.11	2
3	N=C	-0.2	1.85
4	C5	-0.05	2.04
5	N5M	-0.2	1.85
6	PO4	-0.585	2.15
7	OR	-0.1521	1.77
8	O2CM	-0.12	1.7
9	NC=C	-0.2	1.85
10	HCMM	-0.022	1.32
11	HOR	-0.046	0.2245
12	HOCO	-0.046	0.2245
13	HNCO	-0.046	0.2245

Table A.5: Parameters for Electrostatic Interactions Used in This Work

S.N.	Atoms		Charge (q_i, q_j)	Dielectric Constant (ϵ_r)	Constant (K)
1	N1	NM	-0.56	5	332
2	C2	C=O	0.28	5	332
3	N3	N=C	-0.588	5	332
4	C4	C5	0.413	5	332
5	C5	C5	0.202	5	332
6	C6	C=O	0.423	5	332
7	N7	N5M	-1.05	5	332
8	C8	C5	0.4	5	332
9	N9	N5M	-1.493	5	332
10	PA	PO4	1.4424	5	332
11	PB	PO4	1.4424	5	332
12	C5'	CR	0.28	5	332
13	O5'	OR	-0.5512	5	332
14	C4'	CR	0.28	5	332
15	O4'	OR	-0.56	5	332
16	C3'	CR	0.28	5	332
17	O3'	OR	-0.68	5	332
18	C2'	CR	0.28	5	332
19	O2'	OR	-0.68	5	332
20	C1'	CR	0.723	5	332
21	O1A	O2CM	-0.95	5	332
22	O1B	OR	-0.7712	5	332

23	N2	NC=C	-0.85	5	332
24	O2A	O2CM	-0.95	5	332
25	O2B	O2CM	-0.95	5	332
26	O3A	OR	-0.5424	5	332
27	O3B	O2CM	-0.95	5	332
28	O6	O=C	-0.57	5	332
29	H01	HCMM	0	5	332
30	H02	HCMM	0	5	332
31	H03	HCMM	0	5	332
32	H04	HCMM	0	5	332
33	H05	HOR	0.4	5	332
34	H06	HCMM	0	5	332
35	H07	HOR	0.4	5	332
36	H08	HCMM	0	5	332
37	H09	HOCO	0.5	5	332
38	H10	HNCO	0.4	5	332
39	H11	HNCO	0.4	5	332
40	H12	HCMM	0.15	5	332

FORCE FIELD ANALYSIS OF GUANOSINE-50-DIPHOSPHATE AND ITS BINDING MECHANISM WITH THE 1GIT PROTEIN

ORIGINALITY REPORT

18%
SIMILARITY INDEX

PRIMARY SOURCES

1	www.researchgate.net Internet	508 words — 4%
2	pubmed.ncbi.nlm.nih.gov Internet	399 words — 3%
3	nepjol.info Internet	157 words — 1%
4	www.science.gov Internet	137 words — 1%
5	elibrary.tucl.edu.np Internet	111 words — 1%
6	onlinelibrary.wiley.com Internet	77 words — 1%
7	hal.archives-ouvertes.fr Internet	71 words — < 1%
8	pt.scribd.com Internet	68 words — < 1%
9	chemrxiv.org Internet	57 words — < 1%

10	core.ac.uk Internet	47 words — < 1%
11	utpedia.utp.edu.my Internet	40 words — < 1%
12	mdpi.com Internet	38 words — < 1%
13	moam.info Internet	37 words — < 1%
14	www.ncbi.nlm.nih.gov Internet	35 words — < 1%
15	worldwidescience.org Internet	32 words — < 1%
16	Sunghwan Kim, Jie Chen, Tiejun Cheng, Asta Gindulyte et al. "PubChem 2023 update", Nucleic Acids Research, 2022 Crossref	30 words — < 1%
17	Sungsoo M. Yoo, Marc A. Antonyak, Richard A. Cerione. "The Adaptor Protein and Arf GTPase-activating Protein Cat-1/Git-1 Is Required for Cellular Transformation", Journal of Biological Chemistry, 2012 Crossref	29 words — < 1%
18	docplayer.net Internet	29 words — < 1%
19	Kirchmair, Johannes, Christian Laggner, Gerhard Wolber, and Thierry Langer. "Comparative Analysis of Protein-Bound Ligand Conformations with Respect to Catalyst's Conformational Space Subsampling Algorithms", Journal of Chemical Information and Modeling, 2005.	23 words — < 1%

Crossref

20 repositorio.ufla.br
Internet
22 words — < 1%

21 Helmut Sigel, Emanuela M. Bianchi, Nicolas A. Corfù, Yoshiaki Kinjo, Roger Tribolet, R. Bruce Martin. "Stabilities and Isomeric Equilibria in Solutions of Monomeric Metal-Ion Complexes of Guanosine 5′-Triphosphate (GTP4−) and Inosine 5′-Triphosphate (ITP4−) in Comparison with Those of Adenosine 5′-Triphosphate (ATP4−)", Chemistry - A European Journal, 2001
Crossref
21 words — < 1%

22 bmrb.io
Internet
21 words — < 1%

23 Eltareb, Ali. "The Role of Nuclear Quantum Effects in Supercooled Water and Amorphous Ice", City University of New York, 2023
ProQuest
17 words — < 1%

24 M. C. García-Cuesta, A. M. Lozano, J. J. Meléndez-Martínez, F. Luna-Giles et al. " Structure determination of nitrato-κ -bis[2-(2-pyridyl-κ)amino-5,6-dihydro-4 -1,3-thiazine-κ]copper(II) nitrate molecular modelling coupled with X-ray powder diffractometry ", Journal of Applied Crystallography, 2004
Crossref
17 words — < 1%

25 Rahul, Falaq Naz, Smita Jyoti, Yasir Hasan Siddique. "Effect of kaempferol on the transgenic Drosophila model of Parkinson's disease", Scientific Reports, 2020
Crossref
17 words — < 1%

26 dl.lib.mrt.ac.lk
Internet

17 words — < 1%

27 www.frontiersin.org
Internet
17 words — < 1%

28 Jiang, Hao. "Dynamic Asset Allocation Decisions Under Policy and Economic Uncertainty: A Macroeconomic News-Based Study.", University of Delaware, 2020
ProQuest
16 words — < 1%

29 ntnuopen.ntnu.no
Internet
16 words — < 1%

30 becksteinlab.physics.asu.edu
Internet
15 words — < 1%

31 www.slideshare.net
Internet
15 words — < 1%

32 Yu Gan, Yinan Li, Zhi Long, Ahn R. Lee, Ning Xie, Jessica M. Lovnicki, Yuxin Tang, Xiang Chen, Jiaoti Huang, Xuesen Dong. "Roles of Alternative RNA Splicing of the Bif-1 Gene by SRRM4 During the Development of Treatment-induced Neuroendocrine Prostate Cancer", EBioMedicine, 2018
Crossref
14 words — < 1%

33 hmdb.ca
Internet
14 words — < 1%

34 www.americanchemicalsuppliers.com
Internet
14 words — < 1%

35 www.stressmarq.com
Internet
14 words — < 1%

36	helda.helsinki.fi Internet	13 words — < 1%
37	www.mdpi.com Internet	13 words — < 1%
38	M. Bernetti, A. Cavalli, L. Mollica. "Protein–ligand (un)binding kinetics as a new paradigm for drug discovery at the crossroad between experiments and modelling", MedChemComm, 2017 Crossref	12 words — < 1%
39	link.springer.com Internet	12 words — < 1%
40	www.researchsquare.com Internet	12 words — < 1%
41	Carmen Domene, Christian Jorgensen, Kenno Vanommeslaeghe, Christopher J. Schofield, Alexander MacKerell. "Quantifying the Binding Interaction between the Hypoxia-Inducible Transcription Factor and the von Hippel–Lindau Suppressor", Journal of Chemical Theory and Computation, 2015 Crossref	11 words — < 1%
42	Jingjing Guo, Yiqiong Bao, Mengrong Li, Shu Li, Lili Xi, Pengyang Xin, Lei Wu, Huanxiang Liu, Yuguang Mu. "Application of computational approaches in biomembranes: From structure to function", WIREs Computational Molecular Science, 2023 Crossref	11 words — < 1%
43	Maria Anastassiadou, Maria Arena, Domenica Auteri, Alba Brancato et al. "Peer review of the	11 words — < 1%

pesticide risk assessment of the active substance Purpureocillium lilacinum strain PL11", EFSA Journal, 2022
Crossref

44 isotope.bocsci.com
Internet
11 words — < 1%

45 Marine Bugnon, Mathilde Goullieux, Ute F. Röhrig, Marta A. S. Perez, Antoine Daina, Olivier Michielin, Vincent Zoete. "SwissParam 2023: A Modern Web-Based Tool for Efficient Small Molecule Parametrization", Journal of Chemical Information and Modeling, 2023
Crossref
10 words — < 1%

46 Won-Heong Lee, Nam-Soo Han, Yong-Cheol Park, Jin-Ho Seo. "Modulation of guanosine 5′-diphosphate-d-mannose metabolism in recombinant Escherichia coli for production of guanosine 5′-diphosphate-l-fucose", Bioresource Technology, 2009
Crossref
10 words — < 1%

47 Zonghui Wei, Robert Sinko, Sinan Keten, Erik Luijten. "Effect of Surface Modification on Water Adsorption and Interfacial Mechanics of Cellulose Nanocrystals", ACS Applied Materials & Interfaces, 2018
Crossref
10 words — < 1%

48 dokumen.pub
Internet
10 words — < 1%

49 escholarship.org
Internet
10 words — < 1%

50 eurchembull.com
Internet
10 words — < 1%

51 iovs.arvojournals.org
Internet

10 words — < 1%

52 nap.nationalacademies.org
Internet
10 words — < 1%

53 pubag.nal.usda.gov
Internet
10 words — < 1%

54 webzila.com
Internet
10 words — < 1%

55 www.omicsdi.org
Internet
10 words — < 1%

56 "MOLECULAR DYNAMICS SIMULATION OF NUCLEIC ACIDS", Annual Review of Physical Chemistry, 10/2000
Crossref
9 words — < 1%

57 Jang, Hyesu. "Quantum Chemical Studies of Four-Center Iron CO2 Reduction Electrocatalyst and Multifaceted Development of a Small Molecule Force Field", University of California, Davis, 2021
ProQuest
9 words — < 1%

58 Nicolas Tielker, Lukas Eberlein, Oliver Beckstein, Stefan Güssregen, Bogdan I. Iorga, Stefan M. Kast, Shuai Liu. "Perspective on the SAMPL and D3R Blind Prediction Challenges for Physics-Based Free Energy Methods", American Chemical Society (ACS), 2021
Crossref
9 words — < 1%

59 Ryan J. Hoefen, Bradford C. Berk. "The multifunctional GIT family of proteins", Journal of Cell Science, 2006
Crossref
9 words — < 1%

60	Sayer, Alon Haim, Eliav Blum, Dan Thomas Major, Alexandra Vardi-Kilshtain, Bosmat Levi Hevroni, and Bilha Fischer. "Adenosine/guanosine-3′,5′-bis-phosphates as biocompatible and selective Zn2+-ion chelators. Characterization and comparison with adenosine/guanosine-5′-di-phosphate", Dalton Transactions, 2015. Crossref	9 words — < 1%
61	ebin.pub Internet	9 words — < 1%
62	foodb.ca Internet	9 words — < 1%
63	ira.lib.polyu.edu.hk Internet	9 words — < 1%
64	pubs.rsc.org Internet	9 words — < 1%
65	scholarsjunction.msstate.edu Internet	9 words — < 1%
66	tesisenred.net Internet	9 words — < 1%
67	www.authorstream.com Internet	9 words — < 1%
68	www.biocristalografia.df.ibilce.unesp.br Internet	9 words — < 1%

EXCLUDE QUOTES	ON	EXCLUDE SOURCES	OFF
EXCLUDE BIBLIOGRAPHY	ON	EXCLUDE MATCHES	< 9 WORDS