Dissertation Report

On

Exploring NS5 Protein as a Potential Target for Antiviral Development in Yellow Fever: Insights from Molecular Docking & Molecular Dynamics Simulations

Submitted in Partial Fulfilment of the Requirement for the Degree of

B.Sc. Medical Biotechnology

Submitted by

Shalesh Gangwar
Enrollment No. 20SBBS1040013
(Semester VI)

Under the Supervision of

Dr. Neha Sharma
Supervisor & Assistant Professor
Division of Clinical Research
Department of Biosciences
School of Basic and Applied Sciences
Galgotias University,
Greater Noida

Division of Clinical Research, Department of Biosciences

School of Basic and Applied Science
GALGOTIAS UNIVERSITY
Greater Noida, Uttar Pradesh

2023

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the dissertation entitled, "Exploring NS5 Protein as a Potential Target for Antiviral Development in Yellow Fever: Insights from Molecular Docking & Molecular Dynamics Simulations" in fulfillment of the requirement for the award of the Degree of Bachelor of Science in Medical Biotechnology and submitted in the Division of Clinical Research, Department of Biosciences, is an authentic record of my own work carried out during a period from 01 Jan 2023 to 10 May 2023 under the supervision of Dr. Neha Sharma Division of Clinical Research, Department of Biosciences, Galgotias University, Greater Noida.

The matter in this dissertation has not been submitted by me for the award of any other degree of this or any other University/Institute.

Dated: (Shalesh Gangwar)

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

Date:

(Signature)
Dr. Neha Sharma
Supervisor & Assistant Professor
Division of Clinical Research
Department of Biosciences
School of Basic and Applied Sciences
Galgotias University, Greater Noida, U.P.

(Signature)
Professor & Dean
School of Basic and Applied Sciences
Galgotias University, Greater Noida, U.P.

Acknowledgments

I am eternally thankful to Dr. Neha Sharma, my adviser, for her continuous support and assistance during my B.Sc. degree. Their knowledge and patience have been useful to me, and they have played an important part in the achievement of my thesis, as well as acting on my thesis committee and offering significant criticism and ideas. Their ideas and advice were invaluable in shaping my study and writing my thesis.

I am grateful to my friends and family for their love and support during this journey. I could not have finished my adventure without their support and inspiration.

I am thankful to everyone who has helped me along the way. This thesis would not have been feasible without your assistance and direction.

Table of Contents

List of Figures

List of Abbreviations

YFV- Yellow Fever Virus

Kb- Kilo bases

Ps- Pico seconds

Mtase- Methyl transferase

RdRp- RNA dependent RNA polymerase

DAS- Diallyl sulphide

DADS- Diallyl Disulphide

SAC- S-allyl cysteine

GA- Genetic Algorithm

GUI- Graphical User Interface

ADT- AutoDock Tool

GROMACS- Groningen Machine for Chemical Simulations

MD- Molecular Dynamics

RMSD- Root Mean Square Deviation

RMSF- Root Mean Square Fluctuation

Summary

Yellow fever is an infectious disease caused by the Flaviviridae family of viruses. It is spread by the bite of an infected mosquito, Aedes aegypti. It is found in 47 nations throughout Africa and Central and South America, with an estimated 400-500 million individuals living in high-risk regions unvaccinated. Symptoms range from mild to severe, with moderate symptoms such as a fever, headache, aches, pains, nausea, vomiting, and exhaustion. In more severe cases, the illness may develop to a more hazardous stage, known as the toxic phase, which is characterized by high fever, jaundice, organ failure, and hemorrhage. Vaccination and treatment are available, but there is no specific antiviral treatment for yellow fever.

The genome of YFV consists of 11 kilobases of RNA and the UTR at 5'-3' end encodes for 10 proteins. The NS5 protein is of interest to researchers as it acts as a major enzyme in the process by which flaviviruses replicate. It is made up of two essential domains, the N-terminal methyltransferase (MTase) domain and the C-terminal RNA-dependent RNA polymerase (RdRp) domain respectively. These domains are necessary for the virus to replicate successfully in its host. Inhibiting the action of NS5 could affect the replication of flaviviruses and offer a way for the development of antiviral medicines.

Gamma-glutamyl-S-allyl-L-cysteine has the highest docking score of -12.5 Kcal/mol with the target NS5 protein of YFV. Ligplot shows 12 H-bonds being formed, and GROMACS molecular dynamics simulation shows an average of 3.09 H-bonds present over 500 ps. The polymerase also shows similarity to Zika and Dengue virus polymerase, suggesting it could be a potent inhibitor of their replication.

Chapter-1

Introduction

The Flaviviridae family of viruses is responsible for the infectious disease known as yellow fever, which is caused by the Yellow Fever Virus. The bite of an infected mosquito is the most common way for the disease to be spread, and the type of mosquito known as *Aedes aegypti* is one of the most common carriers of the virus. South America and Africa's tropical and subtropical areas are home to a significant population of these insects. It is essential to keep in mind that the transmission of yellow fever from one person to another is not possible.[1] The virus is prevalent in 47 countries across Africa (34 countries) and Central and South America (13 countries). In these regions, yellow fever can be found throughout the entire country or in specific sections within those countries.

A research model based on African data sources was conducted to assess the impact of yellow fever in 2013. The findings of the research indicated that there were an estimated 84,000 to 170,000 severe cases of yellow fever during that year. Among those cases, the disease led to approximately 29,000 to 60,000 fatalities.[2] These numbers demonstrate the significant burden that yellow fever poses to affected populations. It is important to note that these figures are specific to the year 2013 and may vary from year to year. Efforts to combat yellow fever, such as vaccination campaigns and mosquito control measures, have been implemented in these regions to reduce the incidence and impact of the disease. Despite these efforts, a considerable number of individuals remain unvaccinated in high-risk regions. It is estimated that there are approximately 400 to 500 million people living in these areas who have not received the yellow fever vaccine. This highlights the potential for further outbreaks and the importance of vaccination to protect individuals and communities from the disease.[3]

Yellow fever can exhibit a range of symptoms, varying from mild to severe. In many cases, individuals may experience symptoms similar to those of the flu, including fever, headache, muscle and joint pain, nausea and vomiting, and fatigue.

These symptoms generally occur in the initial phase of the illness, known as the acute phase. Most individuals recover from this phase without further complications. However, in some cases, the disease progresses to a more severe and potentially life-threatening stage called the toxic phase.

During the toxic phase of yellow fever, the symptoms become more severe, several manifestations may occur such as high fever, jaundice, organ failure, and hemorrhage. This stage of the disease can be life-threatening and requires immediate medical attention. It is important to note that not all individuals infected with the yellow fever virus progress to the toxic phase. Many recover fully from the initial phase

without complications. However, for those who do progress to severe yellow fever, the risk of fatality is significant. Early diagnosis, supportive care, and prompt medical intervention are crucial in managing severe yellow fever cases and improving patient outcomes.[4,5]

Even though a safe and effective vaccine is widely available, the number of individuals that are unvaccinated is very high and hence requires an intervention as there is no specific antiviral treatment for yellow fever. Supportive care is provided to manage the symptoms and complications. Hospitalization may be required for patients with severe disease to receive specialized medical care.

Due to lack of effective antiviral drug against YFV our focus shifts toward phytochemicals as several plants and their products are known to possess antiviral properties such as Elderberry (*Sambucus nigra*), Echinacea (*Echinacea purpurea*), Garlic (*Allium sativum*), Green tea (*Camellia sinensis*), Licorice (Glycyrrhiza glabra), Ginger (*Zingiber officinale*), Turmeric (*Curcuma longa*), Aloe vera, Neem (*Azadirachta indica*)[6]. Among these plants, ayurveda claims that garlic is observed to be effective against the YFV, but there is no concrete proof of the claim.

Garlic (*Allium sativum*) is known for its distinctive taste and smell, and many of its health advantages are due to organosulfur compounds found in the plant. Allicin, DAS, DADS, Ajoene, and SAC are some of the most significant organosulfur compounds found in garlic. Allicin is responsible for fresh garlic's intense odor and flavor and has antimicrobial, antioxidant, and anticancer effects. Diallyl disulfide (DADS) is the source of the pungent odor associated with garlic ingestion, and Ajoene has been studied for potential health advantages. SAC is a stable and bioavailable organosulfur molecule generated during garlic aging. Given the diverse bioactive compounds present in garlic, researchers have investigated their potential inhibitory properties against the replication of Yellow Fever virus (YFV). Yellow Fever is a viral infection transmitted by mosquitoes and can cause severe illness. The assessment of inhibitory properties against YFV replication involved the selection of 12 organosulfur compounds from garlic for the assessment of inhibitory properties against YFV replication.

Review of Literature

2.1 Overview of Yellow Fever Virus

The yellow fever virus (YFV) is an enveloped, single-stranded RNA virus belonging to the Flaviviridae family causing around 60,000 deaths. It has a spherical shape with an average diameter of approximately 40-50 nanometers. The virus is composed of several structural and non-structural proteins that contribute to its overall structure and function. The Basic structure of YGV includes 4 major parts: Envelope is the outermost layer of the virus is the envelope, which is derived from the host cell membrane during viral budding. The envelope is studded with viral glycoproteins known as envelope proteins (E proteins). These proteins play a crucial role in viral attachment, entry into host cells, and induction of immune responses causing symptoms such as high fever (The fever becomes more pronounced and can reach dangerous levels), jaundice (One of the defining characteristics of yellow fever is jaundice, which is the yellowing of the skin and eyes, this occurs due to liver damage caused by the virus), organ failure (yellow fever can lead to the dysfunction of multiple organs, particularly the liver and kidneys, hemorrhage (In rare cases, yellow fever can cause bleeding disorders and result in internal bleeding or bleeding from the gums, nose, or other parts of the body.)

Capsid is present beneath the envelope, the virus has a protein shell called the capsid, which encloses the viral genetic material. The capsid is made up of repeating units of a single protein called the capsid protein (C protein). It provides protection to the viral RNA, RNA Genome is a positive-sense, single-stranded RNA genome. The RNA genome is approximately 11,000 nucleotides long and contains a single open reading frame encoding a polyprotein precursor. This polyprotein is processed by viral and host proteases to generate individual viral proteins & 7 Non-Structural Proteins that, after the viral polyprotein is processed, it gives rise to several non-structural proteins (NS proteins) that are essential for viral replication and evasion of host immune responses. These NS proteins include NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5.The yellow fever virus structure enables its attachment to host cells, entry into cells, replication of its RNA genome, and assembly of new viral particles. The envelope proteins (E proteins) play a crucial role in viral entry and are the primary targets for host immune responses.

2.2 Essential proteins of Yellow Fever Virus (YFV)

The genome of YFV consists of around 11 kilobases (kb) of RNA and the UTR at 5’-3’ end encodes for 10 proteins. These proteins include three structural proteins (capsid, pre-membrane, and envelope) and seven nonstructural proteins (designated NS1, NS2A, NS2B, NS3, NS4A, and NS5).[7]. Among these NS5 protein is of keen interest to researchers as it acts as a major enzyme, the NS5 protein is an extremely important component in the process by which flaviviruses replicate. It is made up of two essential

domains, that are the N-terminal methyltransferase (MTase) domain and the C-terminal RNA-dependent RNA polymerase (RdRp) domain respectively. These domains are necessary for the virus to replicate successfully in its host.[8]. Guanine-N7 methylation and nucleoside-2'-O methylation are the two enzymatic processes that are carried out by the MTase domain that is included inside the NS5 protein[8–10]. These alterations prevent the freshly generated RNA from being broken down by enzymes called 5'-3' exoribonucleases, which is a result of an increase in the RNA's stability. The innate immunological response of the host may also be avoided by the virus with the help of RNA methylation[11]. The innate immunity factor known as Ifit1 (IFN-induced RNA-binding protein) has the ability to identify non-methylated RNA, which in turn leads to the suppression of viral RNA translation. Because of this, the actions of the MTase domain are essential for the viral replication and immune evasion processes.

On the other side, the replication of the viral RNA is the responsibility of the RdRp domain. This particular enzymatic activity can only be found in the RdRp domain, and it cannot be replicated by any of the host enzymes. As a direct consequence of this, the RdRp domain is required in order for flaviviral RNA to be replicated. Given the significance of the MTase and RdRp domains in flaviviral replication, the NS5 enzyme emerges as a desirable target for the investigation and development of new antiviral medications. Inhibiting the action of NS5 might possibly affect the replication of flaviviruses and offer a way for the development of antiviral medicines.[12]

2.3 Current Preventive and Post-infective Intervention

Since there is no treatment for this disease, preventive intervention such as vaccination is used against yellow fever. It provides long-term immunity and is highly effective in preventing the disease. Routine vaccination campaigns, particularly in endemic areas, are essential for reaching as many individuals as possible and reducing the risk of outbreaks. Additionally, vaccination is often required or recommended for travelers visiting countries where yellow fever is endemic, as it protects both the individual traveler and helps prevent the importation of the virus to non-endemic areas. It is important to note that there is no specific antiviral treatment available for yellow fever (WHO, 2019). Once a person becomes infected, supportive care is provided to manage the symptoms and complications associated with the disease. This typically involves symptomatic treatment such as acetaminophen (paracetamol) to alleviate fever, pain, and other symptoms, such as), may be administered, fluid management, monitoring and management of complications, preventing secondary infections (Due to the weakened immune system in severe cases, precautions should be taken to prevent secondary infections. This may involve the use of antibiotics or antifungal medications if necessary). But due to lack of conventional allopathic antiviral agents against YFV infection alternate products are analyzed such as phytochemicals

2.4 Antiviral Properties of Plant and their Products

Several plants and their products are known to possess antiviral properties such as Elderberry (*Sambucus nigra*) whose extract has been used traditionally to treat viral infections, including influenza. It contains compounds that can inhibit the replication of certain viruses. Echinacea (*Echinacea purpurea*) is commonly used as an immune-boosting herb. It has been studied for its antiviral activity against respiratory viruses, including the common cold. Garlic (*Allium sativum*) has been recognized for its antimicrobial properties, including antiviral effects. It contains compounds like allicin, which have shown inhibitory effects against various viruses. Green tea (*Camellia sinensis*) is rich in polyphenols, particularly catechins, which possess antiviral properties. It has been studied for its effectiveness against viruses like influenza and hepatitis, Licorice (*Glycyrrhiza glabra*) root contains glycyrrhizin, a compound known for its antiviral and immune-stimulating properties. It has demonstrated activity against viruses like herpes simplex and respiratory viruses. Ginger (*Zingiber officinale*) possesses antiviral properties and has been used traditionally to treat respiratory infections. It has been shown to inhibit the replication of certain viruses, including respiratory syncytial virus (RSV). Turmeric (*Curcuma longa*) contains curcumin, a compound with antiviral activity. It has been investigated for its potential against viruses like hepatitis B, herpes simplex, and HIV. Neem (*Azadirachta indica*) has been used in traditional medicine for its antiviral properties. It contains compounds that have shown inhibitory effects against various viruses, including herpes simplex and dengue virus.[6]

2.5 *Allium sativum* (Garlic): A potential NS5 inhibitor

For millennia, garlic (*Allium sativum*) has been used for its therapeutic benefits. It includes a number of chemicals, including sulfur compounds such as allicin, which are thought to contribute to its health benefits. While garlic is often used in cooking, it has also been researched for potential medicinal purposes. Here are some of the potential health advantages and applications of garlic:

- Heart health: Garlic has been linked to potential cardiovascular benefits. It may aid in lowering blood pressure, lowering cholesterol, and improving blood circulation. These effects may help to minimize the risk of heart disease and stroke.[13,14]
- Immune system support: Garlic has long been used to improve the immune system. According to certain research, garlic may stimulate some immune cells and have antibacterial and antiviral properties.[15]
- Antioxidant properties: Garlic includes antioxidants that help protect the body from oxidative stress, which is linked to a variety of chronic diseases and aging. Antioxidants may aid in the neutralization of damaging free radicals and the reduction of inflammation in the body.[16]

- Potential antimicrobial effects: Garlic has been proven in laboratory experiments to have antibacterial activity against certain bacteria and fungi. It may be beneficial against common pathogens, though the degree of its effects in clinical settings is still being researched.[17]
- Cancer prevention: Garlic consumption may be connected with a lower risk of certain types of cancer, notably malignancies of the digestive system, such as stomach and colorectal cancer.[18–21]

Many of its characteristic properties are due to organosulfur compounds found in the plant. When garlic is crushed, diced, or otherwise broken, enzymes are released that transform the sulfur-containing chemicals contained in garlic into other beneficial compounds.[22] The following are some of the most significant organosulfur compounds identified in garlic:

- Allicin: Allicin is a well-known organosulfur chemical found in garlic. It is responsible for fresh garlic's intense odor and flavor. Antimicrobial, antioxidant, and anticancer effects have been shown for allicin. However, since it is unstable and rapidly degrades into other compounds, its content in garlic decreases with heating or processing.
- DAS (diallyl sulfide): DAS is a significant organosulfur component found in garlic. It is produced as a byproduct of the breakdown of allicin and adds to the distinctive scent and taste of cooked garlic. DAS seems to have anticancer, antibacterial, and antioxidant effects.
- Another significant organosulfur molecule generated from allicin is diallyl disulfide (DADS). It is the source of the pungent odor associated with garlic ingestion. DADS has a wide range of biological actions, including anticancer, antioxidant, and antibacterial properties.
- Ajoene is a distinct organosulfur chemical found in garlic. It is generated from allicin and has been studied for possible health advantages such as antithrombotic (anti-blood clotting), antibacterial, and antifungal characteristics.[23]
- S-allyl cysteine (SAC): SAC is a stable and bioavailable organosulfur molecule generated during garlic aging. It is quickly absorbed by the body and has been proven to have potential cardiovascular, anticancer, and antioxidant properties.

Ayurveda has hinted that Garlic may be effective against YFV but there is no concrete proof to back the claim.

2.6 Molecular Docking using AutoDock Vina

To assess the claims of ayurveda and find the compound responsible for its action molecular docking is to be performed as preliminary in-silico analysis.

AutoDock Vina is a widely used software program for molecular docking, which is a computational method used to predict and study the binding of small molecules (ligands) to a target protein or macromolecule. AutoDock Vina is an improved version of the original AutoDock software and provides

several enhancements and features. Here are some key points about AutoDock Vina:

- Molecular Docking: Molecular docking involves predicting the preferred binding orientation and affinity between a small molecule ligand and a target protein. It is useful for studying protein-ligand interactions, virtual screening of compound libraries, and drug discovery.
- Scoring Function: AutoDock Vina employs a scoring function that estimates the binding energy between the ligand and protein. The scoring function evaluates various energetic contributions, such as van der Waals interactions, electrostatic interactions, and hydrogen bonding. The goal is to predict the most favorable binding pose and affinity for a given ligand.
- Search Algorithm: AutoDock Vina employs an efficient search algorithm based on a modified Lamarckian genetic algorithm (GA). The GA explores the conformational space of the ligand and performs a global search to find low-energy binding poses.
- Flexible Ligand and Receptor: AutoDock Vina can handle ligands with flexible torsional degrees of freedom, allowing the ligand to adopt different conformations during docking. It can also accommodate protein flexibility by allowing side-chain movements in the receptor.
- Graphical User Interface (GUI): AutoDock Vina provides a command-line interface for advanced users, but it also has a graphical user interface called AutoDockTools (ADT). ADT allows users to prepare ligand and receptor structures, set up docking parameters, visualize and analyze docking results, and perform various other tasks related to molecular docking.
- Output and Analysis: AutoDock Vina generates output files that contain information about the predicted binding poses, predicted binding affinities (energies), and other relevant data. Users can analyze the results to understand the binding interactions, evaluate different poses, and prioritize compounds based on their predicted binding affinities.

2.7 GROMACS for Molecular Dynamic Simulations

GROMACS (Groningen Machine for Chemical Simulations) is a widely used software package for molecular dynamics (MD) simulations of biomolecular systems. It provides a comprehensive set of tools for simulating the motion and behavior of atoms and molecules over time. Here are some key features and steps involved in performing molecular dynamics simulations using GROMACS:

- System setup: The first step is to prepare the system for simulation. This includes obtaining the initial structure of the biomolecule or complex of interest (such as a protein or a protein-ligand

complex) and solvating it in a suitable solvent model, typically a water model like TIP3P or SPC. The system may also require the addition of counterions to maintain neutrality.

- Force field selection: GROMACS utilizes force fields to describe the interactions between atoms and molecules in the system. Force fields define parameters such as bond lengths, bond angles, dihedral angles, and non-bonded interactions (van der Waals and electrostatic forces). GROMACS supports various force fields, including Amber, CHARMM, and OPLS, among others.
- Energy minimization: Before starting a molecular dynamics simulation, the system is typically subjected to energy minimization to relax any steric clashes or unfavorable contacts. This process involves iteratively adjusting the atomic positions to find a low-energy conformation.
- Equilibration: After energy minimization, the system undergoes equilibration to bring it to a stable state before the production run. Equilibration involves a series of simulation steps, including position restraints on the solute, gradual temperature and pressure adjustments, and possibly additional steps like solvent equilibration.
- Production run: Once equilibration is complete, the production run begins. In this phase, the system is simulated over a longer time period to capture the dynamics of the biomolecular system. GROMACS utilizes numerical integration algorithms to solve the equations of motion and simulate the behavior of the atoms based on the defined force field.
- Analysis: After the molecular dynamic simulation, GROMACS provides various analysis tools to extract information from the trajectory data. These tools can calculate properties such as root mean square deviation (RMSD), root mean square fluctuation (RMSF), hydrogen bond analysis, and many other structural and dynamic properties.

2.8 Similarity of NS5 Protein among Flavivirus genus

Similarity with Zika and Dengue Viruses: The similarity between the NS5 protein of YFV and the polymerases of Zika and Dengue viruses suggests that gamma-glutamyl-S-allyl-L-cysteine could potentially inhibit the replication of these viruses as well. This similarity hints at the compound's broad-spectrum antiviral potential, but experimental studies would be required to confirm its efficacy against Zika and Dengue viruses.[24]

Objectives:

1) Create a library of organo-sulfur compounds present in *Allium sativum*.
2) Docking the compounds with the crystal structure of NS5 protein of YFV for potential inhibitory effect.
3) Molecular dynamic simulation of the highest binding energy configuration to find RMSD value, RMSF value, H-bonding, bond energy of the protein-ligand complex.

Chapter-3

Materials and Methods

3.1 Optimizing the target protein

- The protein structure with the PDB ID 6QSN was downloaded from the Protein Data Bank (PDB) database. The PDB is a repository for 3D structural data of biological macromolecules.
- The downloaded structure was opened in UCSF Chimera v1.16(Build: 42360). UCSF Chimera is a molecular visualization and analysis software commonly used in structural biology.
- In order to prepare the protein structure for further analysis, all heteroatoms were removed. Heteroatoms are non-standard atoms present in the structure, such as ligands or cofactors. The removal of heteroatoms simplifies the analysis by focusing only on the protein itself.
- Additionally, any missing non-terminal residues in both the A and B chains of the protein were added. Non-terminal residues are the amino acid building blocks that make up the protein chain. The addition of missing residues helps to complete the protein structure, ensuring that it is more accurate and complete.
- To add the missing residues, the Modeller web server was utilized. Modeller is a software tool commonly used for homology modeling, which predicts the missing residues based on known protein structures with similar sequences.
- The Modeller web server generated a total of 5 models for each chain (A and B). These models represent different possible conformations of the missing residues.
- Among the generated models, the one with the lowest zDOPE (Discrete Optimized Protein Energy) score was selected. The zDOPE score is a measure of the quality and reliability of a protein model, with lower scores indicating better quality.
- The selected model for each chain was further processed and prepared for docking. Docking is a computational method used to predict the binding interactions between a protein and other molecules.
- The preparation of the repaired structure for docking was done using the default settings in the Dock

Prep tool of UCSF Chimera. The Dock Prep tool performs various tasks, such as adding missing hydrogen atoms, optimizing bond lengths and angles, and assigning charges to prepare the protein structure for docking simulations.

3.2 Preparation of ligand library and their optimization

Dr. Duke's Phytochemical and Ethnobotanical Databases, provided by the USDA (United States Department of Agriculture), were utilized to search for organo-sulfur molecules found in garlic. These databases contain information on various phytochemicals and their sources. Once the specific organo-sulfur molecules were identified, their 3D structure files in .sdf (Structure-Data File) format were obtained from PubChem. PubChem is a public database that provides access to a vast collection of chemical compounds and their properties. The .sdf files, representing the 3D structures of the organo-sulfur molecules, were opened using UCSF Chimera. UCSF Chimera is a versatile software tool used for molecular visualization and analysis. To prepare the organo-sulfur molecules for further analysis, missing hydrogen atoms were added. Hydrogen atoms play a crucial role in determining the 3D structure and interactions of molecules. Adding hydrogen atoms ensures that the structures are complete and accurately represent the molecules. In addition to adding hydrogen atoms, charges were assigned to the atoms in the molecules. Charges are important for understanding the electrostatic interactions and behavior of molecules. Assigning charges is typically done based on empirical or theoretical methods. After adding hydrogens and charges, the modified structures of the organo-sulfur molecules were saved in .pdb (Protein Data Bank) format. The PDB format is widely used for storing and sharing 3D structural information of biological molecules, including small molecules like the organo-sulfur compounds

3.3 .Molecular Docking using AutoDock Vina v1.1.2

Molecular docking is a computational method used to predict how a small molecule, known as a ligand, interacts with a target macromolecule, typically a protein. It helps to determine the binding mode and affinity of the ligand to the protein. Molecular docking is widely used in drug discovery and design processes to identify potential drug candidates and understand their interactions with the target protein. In order to perform molecular docking, the dimensions of a grid box are determined. The grid box defines the search space within the protein where the ligand can bind. The size and position of the grid box are crucial as they affect the accuracy and efficiency of the docking simulation. In this case, Discovery Studio Visualizer v21.1.0.20298 was used to determine the dimensions of the grid box. Once the grid box dimensions are determined, a configuration file is prepared for docking. The configuration file contains parameters that define the docking protocol, including the coordinates and size of the grid box, exhaustiveness (which controls the thoroughness of the docking search), and other settings specific to the docking software being used. (center_x = 57.342733 center_y = 18.398052 center_z = 48.695976

size_x = 70 size_y = 70 size_z = 70 exhaustiveness= 50). In this case, AutoDock Vina, a widely used molecular docking software, was employed to perform the docking simulations. AutoDock Vina uses a search algorithm to explore the conformational space of the ligand and protein, searching for the optimal binding mode. The molecular docking was performed for each organo-sulfur molecule obtained earlier, using AutoDock Vina. The ligand molecules were docked into the binding site of the target protein (obtained from the repaired protein structure) based on the defined grid box. During the docking simulations, AutoDock Vina calculates docking scores to evaluate the binding affinity of each ligand towards the protein. The docking scores represent the predicted strength of the ligand-protein interaction, with lower scores indicating a potentially better binding affinity.

3.4 Ligand-protein interaction observed using Ligplot+ v2.2.5

4 LigPlot is a software tool used for the visual analysis and representation of protein-ligand interactions. It allows the generation of 2D schematic diagrams that depict the interactions between a protein and its ligands, such as small molecules, peptides, or nucleic acids. LigPlot provides insights into the types of interactions occurring within the binding site, including hydrogen bonds, hydrophobic contacts, and other non-covalent interactions.

3.5 Molecular Dynamic Simulation using GROMACS 2023.1

Molecular dynamics (MD) simulation is a computational method used to study the behavior of atoms and molecules over time. It is widely used in various fields of science, including chemistry, physics, materials science, and biology.

- CHARMM27 all-atom force field is selected
- Ligand is converted to mol2 format and using SwissParam server topology and parameters were successfully generated.
- Box around the target protein is formed for the simulation.
- Fill the box with solvate i.e., water molecules
- Ions such as Chloride and Sodium were added to neutralize the system.
- Energy Minimization: Before starting the simulation, it is necessary to perform an energy minimization step to relax the system and remove any unfavorable contacts or clashes. During energy minimization, the positions of atoms are adjusted iteratively to find a local minimum of the potential energy.
- Equilibration: Next, the system is gradually brought to the desired temperature and pressure through an equilibration process. This step allows the system to adjust and reach a stable state before the production simulation. Equilibration typically involves a series of simulation runs, such as heating

or cooling the system, followed by a gradual release of any constraints applied.

- Production Run: Once the system is equilibrated, the production simulation begins. During this phase, the equations of motion are integrated numerically, using Verlet algorithm. The system evolves in time, and coordinates, velocities, and energies are recorded at regular intervals for analysis.
- Analysis: After the simulation is completed, the generated trajectory and other output files are analyzed to extract meaningful information. This include the RMSD, RMSF values, hydrogen bonding & bond energies.

Chapter-4

Results and Discussion

4.1 Molecular docking using AutoDock Vina

In molecular docking, a docking score is a numerical value that quantitatively represents the predicted binding affinity or energy of a ligand-protein complex. The docking score provides a ranking of different ligands based on their potential binding strength to the target protein. The docking score is typically calculated using a scoring function, which incorporates various terms to estimate the interaction energy between the ligand and protein. Different scoring functions may have distinct components and weighting schemes, but some common factors considered in scoring functions include Van der Waals Interactions: that account for the attractive and repulsive forces between non-bonded atoms. The scoring function penalizes steric clashes and rewards favorable contacts between ligand and protein atoms. Electrostatic Interactions that arise from the charges of atoms and contribute to the binding energy. The scoring function evaluates the electrostatic interactions between charged groups in the ligand and protein. Hydrogen Bonding of Hydrogen bonds form between hydrogen atoms and electronegative atoms like oxygen or nitrogen. The scoring function may consider the presence and quality of hydrogen bonds between the ligand and protein. Desolvation Effects: When a ligand binds to a protein, it displaces water molecules from the binding site, resulting in desolvation effects. The scoring function accounts for the cost of desolvating the ligand and protein surface areas.

12 organo-sufur molecules of *Allium sativum* were docked using the parameters (center_x = 57.342733 center_y = 18.398052 center_z = 48.695976 size_x = 70 size_y = 70 size_z = 70 exhaustiveness= 50). Among the 12 molecules gamma-glutamyl-S-allyl-L-cysteine had the highest docking score of -12.5 Kcal/mol (Figure 1) and was used for further experimentation.

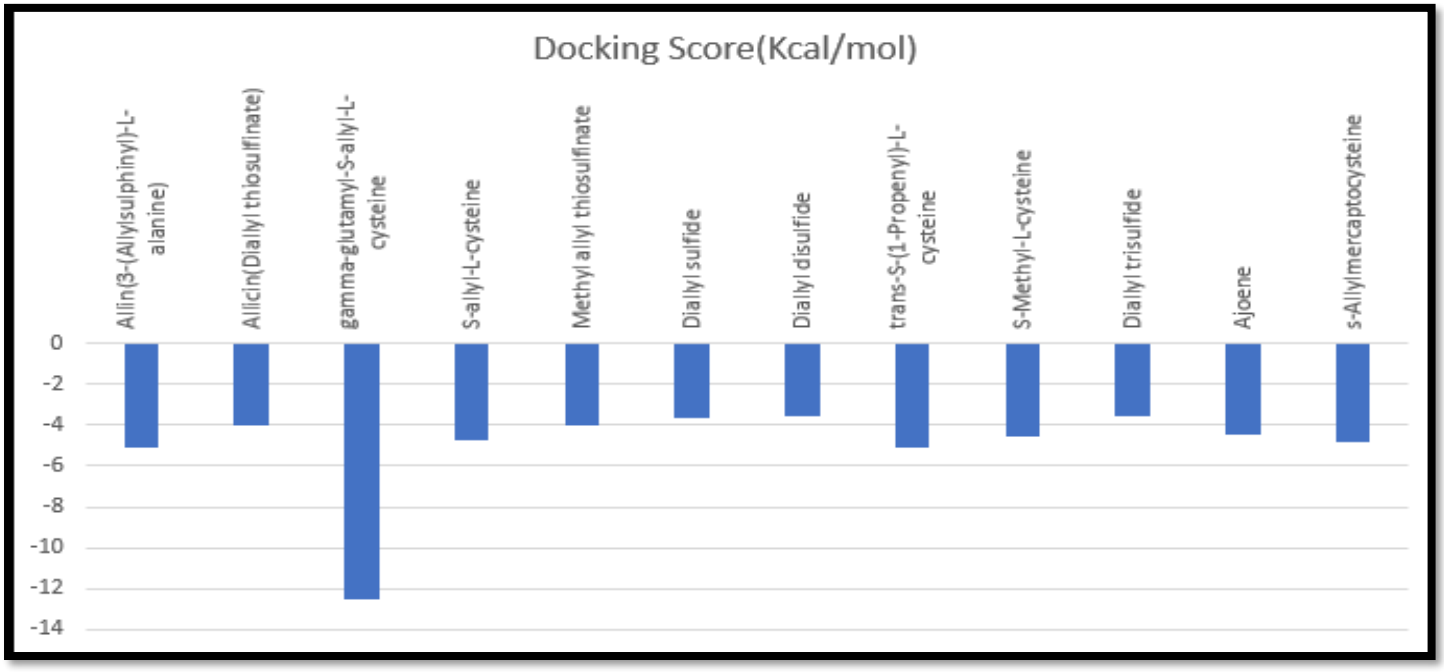

Figure 1. Docking scores of 12 organo-sulfur molecules of *Allium sativum*

4.2 Ligand-protein interaction observed using Ligplot+ v2.2.5

LigPlot is a software tool used for the visual analysis and representation of protein-ligand interactions. It allows the generation of 2D schematic diagrams that depict the interactions between a protein and its ligands, such as small molecules, peptides, or nucleic acids. LigPlot provides insights into the types of interactions occurring within the binding site, including hydrogen bonds, hydrophobic contacts, and other non-covalent interactions.

Interaction Diagrams: LigPlot generates interaction diagrams that illustrate the protein-ligand interactions in a clear and intuitive manner. The diagrams are created by mapping the ligand atoms and protein residues involved in the interactions onto a schematic representation of the protein structure. Interactions Analysis: LigPlot identifies and classifies different types of interactions between the ligand and protein. This includes hydrogen bonds, pi-pi stacking, salt bridges, hydrophobic contacts, and metal-ligand interactions. The diagrams highlight the specific residues involved in these interactions and provide a quantitative assessment of their strength. Ligand Orientation: LigPlot helps visualize the orientation of the ligand within the protein binding site. It shows how the ligand interacts with specific amino acid residues and how it fits within the active site or binding pocket. Water-Mediated Interactions:

LigPlot can also display water molecules involved in mediating interactions between the protein and ligand. These water-mediated interactions can play a significant role in ligand binding and stabilization.

The pdb file containg the ligand and protein with highest docking score orientation was analysed using Ligplot that helps visualize various interaction of ligand with the protein.

We observed 12 H-bonds(green dotted lines) and 1 non-bonding contact(spoked arc)(Figure 2)

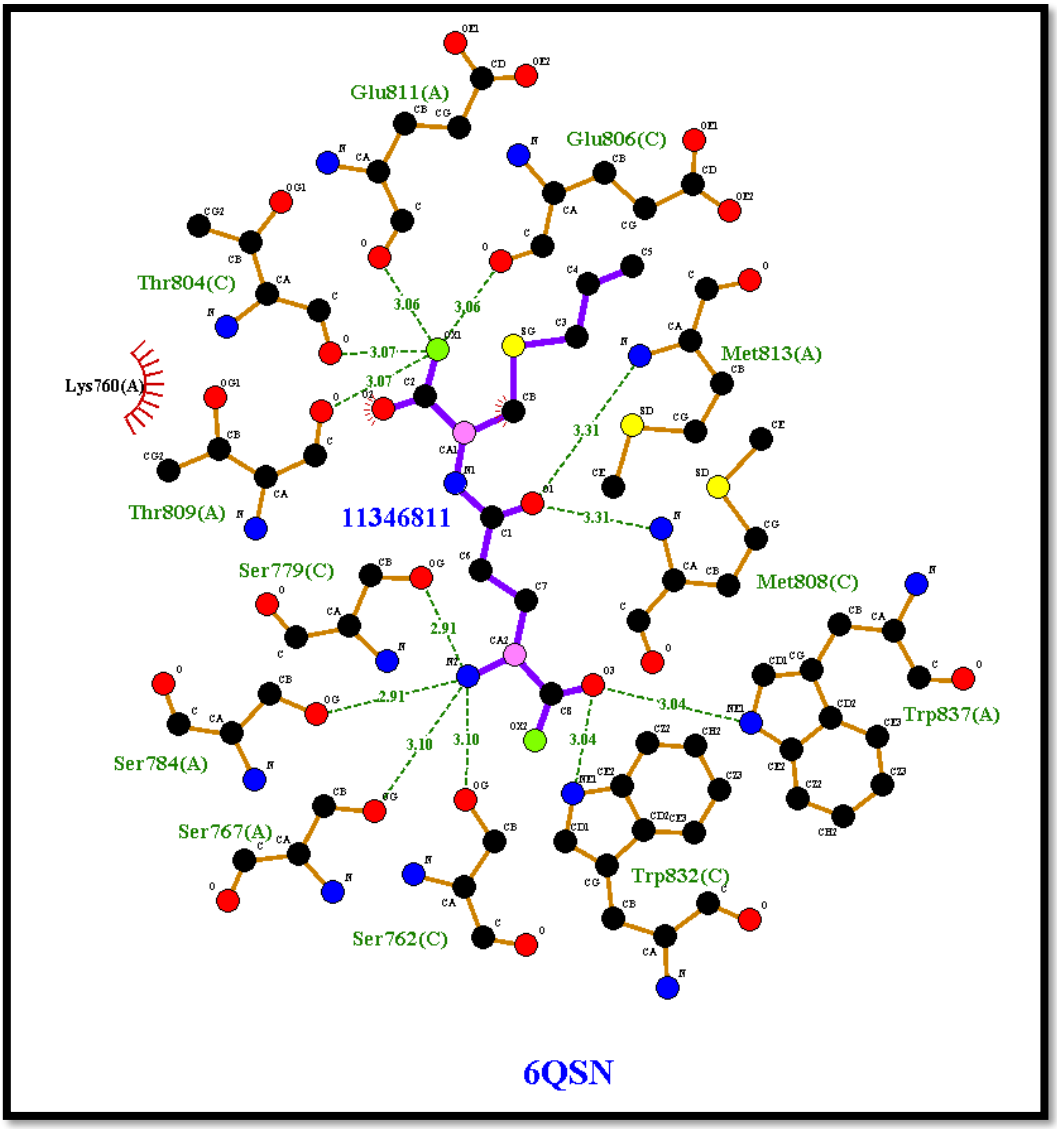

Figure 2. Protein-ligand interaction visualized using Ligplot, Protein used is NS5 of YFV (PDB id: 6QSN) with gamma-glutamyl-S-allyl-L-cysteine (Pubchem id: 11346811)

4.3 RMSD calculation

Root Mean Square Deviation (RMSD) is a measurement used to quantify the structural similarity or difference between two sets of atomic coordinates. It is commonly employed in structural biology and

computational chemistry to compare the similarity of protein structures, ligand poses, or other biomolecular conformations. RMSD calculates the average distance bet
en corresponding atoms in the two sets of coordinates, taking into account their individual displacements. Here are some key points about RMSD:

Calculation: The RMSD is calculated as follows:

RMSD = $\sqrt{(1/N)\sum_{i=1}^{N}.\delta_i^2}$

where N is the number of atoms being compared, and δ_i is the distance between the corresponding atoms in the two structures.

Units: The RMSD value is typically expressed in units of length, such as angstroms (Å) or nanometres (nm), depending on the coordinate system used.

Alignment: Prior to calculating RMSD, the two sets of coordinates are usually aligned to optimize the overlap between the structures. This alignment step minimizes differences due to translation, rotation, or other global transformations. Interpretation: A lower RMSD value indicates a higher similarity or structural overlap between the two structures being compared. Conversely, a higher RMSD value suggests greater structural divergence or differences. Context and Comparison: The significance of an RMSD value depends on the context and purpose of the analysis. In protein structure comparison, RMSD values below 2 Å are often considered indicative of structurally similar conformations. Higher RMSD values, above 2 Å, indicate increasing structural deviation. Ligand Docking: In the context of ligand docking or virtual screening, RMSD is often used to assess the similarity between the predicted ligand pose (docked conformation) and the reference or experimental conformation. A low RMSD value implies a good agreement between the predicted and reference poses, suggesting a potentially accurate prediction. The variations caused during the simulation of the protein may be used to measure its stability relative to its conformation. The RMSD is 0 for identical structures and grows as the two structures become more dissimilar. When used to highly similar proteins, such as alternate conformations of the same protein, RMSD values are regarded as trustworthy markers of variability.

The RMSD value if below 2 Å or 0.2 nm for the protein backbone and even the ligand RMSD value becomes similar to protein backbone after 400ps(Figure 3).

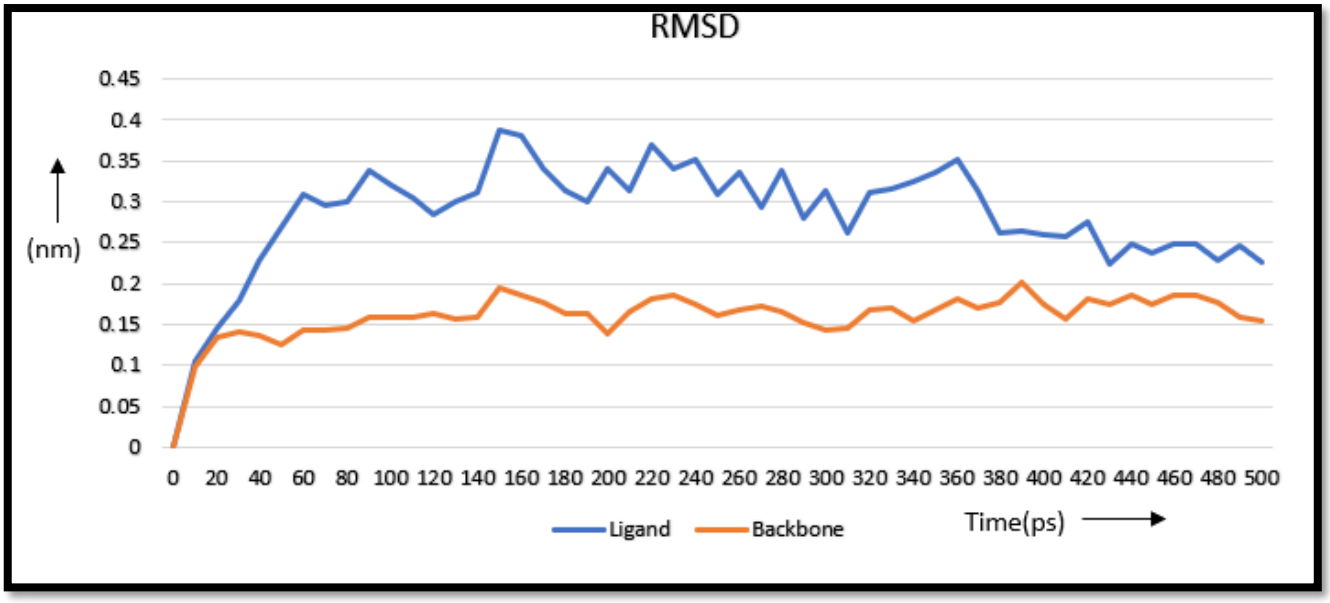

Figure 3. RMSD values of the protein backbone and the ligand molecule with respect to time(in picoseconds)

4.4 RMSF calculation

Root Mean Square Fluctuation (RMSF) is a measurement used to quantify the flexibility or fluctuation of individual atoms or groups of atoms in a biomolecular system, such as a protein or nucleic acid. It provides insights into the dynamic behaviour and mobility of the atoms within a structure. RMSF is calculated from molecular dynamics (MD) simulations or experimental data.

Calculation: RMSF is calculated by measuring the average displacement or fluctuation of each atom or group of atoms from their average position throughout a trajectory or ensemble of structures. The calculation involves determining the root mean square (RMS) value of the atomic fluctuations across the trajectory.

Units: The RMSF values are typically expressed in units of length, such as angstroms (Å) or nanometres (nm), reflecting the displacement or fluctuation of the atoms from their average position.

Interpretation: Higher RMSF values indicate greater flexibility or mobility of the atoms, while lower RMSF values suggest more rigid or stable regions. Flexible regions often correspond to surface loops, terminal regions, or regions involved in binding interactions, while more rigid regions are typically associated with secondary structural elements like alpha helices or beta sheets. Visualization: RMSF values can be visualized as a plot or heatmap, where each residue or atom is represented along the x-axis, and the corresponding RMSF values are plotted on the y-axis. This representation provides a profile of the flexibility or fluctuation along the protein sequence or structure.

Biological Implications: RMSF values can offer insights into the functional importance of certain regions

or residues. Highly fluctuating regions may indicate regions involved in binding, conformational changes, or allosteric regulation. Additionally, RMSF analysis can aid in identifying flexible regions suitable for targeting in drug discovery or protein engineering. Validation and Comparison: RMSF values can be compared between different systems or conditions, such as comparing the RMSF profiles of a protein in its apo (unbound) and holo (ligand-bound) states. Such comparisons can provide information on how ligand binding affects the flexibility of the protein and highlight key regions involved in the binding process. Prior atoms of ligand have less fluctuation as compared to later ones (Figure 4).

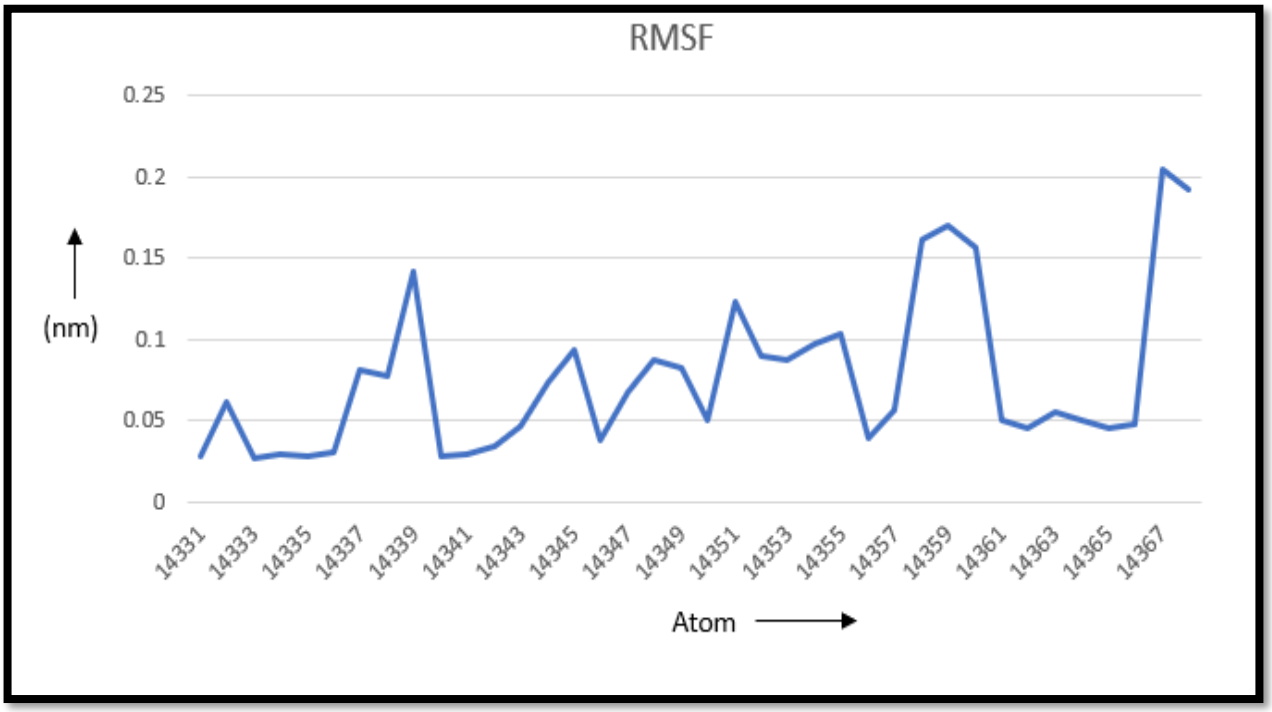

Figure 4. RMSF values of the ligand molecule's atoms

4.5 Hydrogen bonds formed during molecular docking simulation

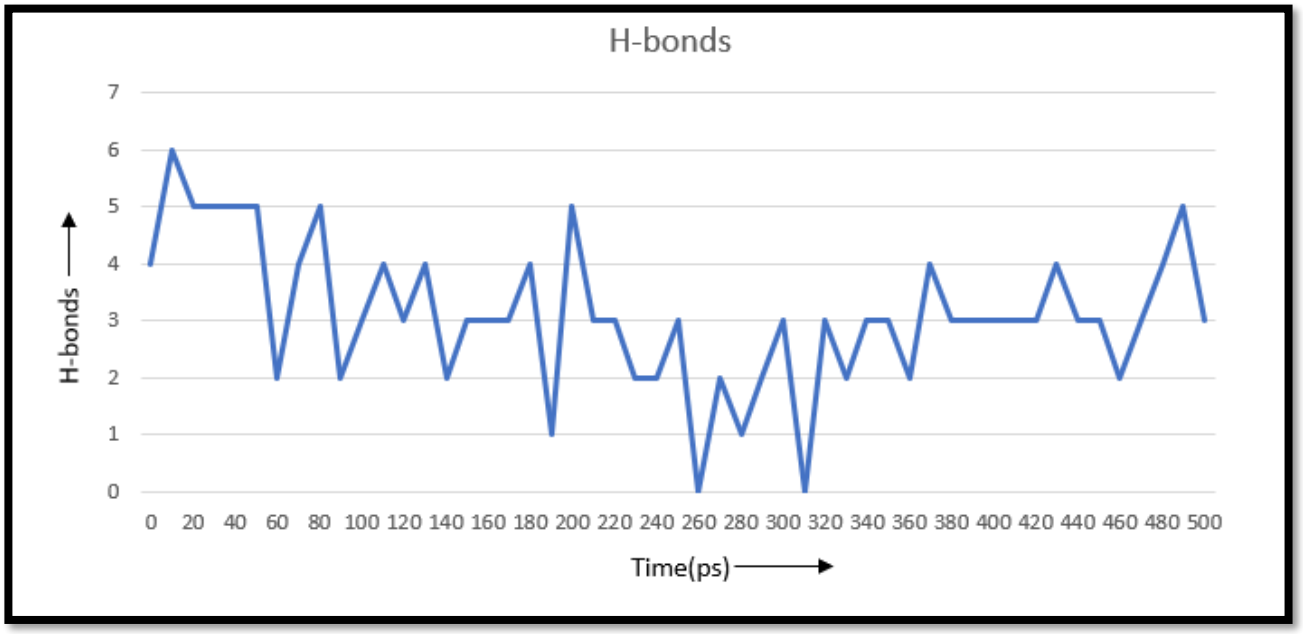

Figure 5. H-bonds formed between protein and ligand with respect to time (in picoseconds)

4.6 Average Bond Energy between Protein and ligand

The bond energy between a protein and a ligand refers to the strength of the covalent and noncovalent interactions that hold the ligand within the binding site of the protein. These interactions contribute to the stability of the protein-ligand complex and play a crucial role in determining the binding affinity and specificity. It's important to note that the bond energy between a protein and a ligand is typically not a single covalent bond but a combination of various covalent and noncovalent interactions.

Hydrogen Bonds: Hydrogen bonds occur when a hydrogen atom is covalently bonded to an electronegative atom (e.g., nitrogen or oxygen) and interacts with another electronegative atom in the protein or ligand. Hydrogen bonding is a significant contributor to the overall bond energy and can involve multiple hydrogen bonds within the complex.

Electrostatic Interactions: Electrostatic interactions arise from the attraction or repulsion between charged particles. Protein-ligand complexes may involve interactions between positively charged residues (e.g., arginine or lysine) and negatively charged ligand atoms, or vice versa. These electrostatic interactions can significantly contribute to the overall bond energy.

Van der Waals Interactions: Van der Waals interactions include attractive forces between atoms or groups due to induced dipoles or temporary fluctuations in electron density. These interactions, including London dispersion forces and dipole-dipole interactions, play a crucial role in stabilizing protein-ligand

complexes.

Pi-Stacking Interactions: Pi-stacking occurs when aromatic rings in the protein and ligand align in a parallel or near-parallel fashion, resulting in attractive interactions. Pi-stacking interactions can contribute to the overall bond energy in complexes involving aromatic ligands or aromatic residues in the protein.

Hydrophobic Interactions: Hydrophobic interactions refer to the tendency of nonpolar molecules or groups to cluster together in the presence of water. In protein-ligand complexes, hydrophobic interactions can occur between nonpolar regions of the ligand and hydrophobic residues in the protein. These interactions help drive the binding process and contribute to the bond energy.

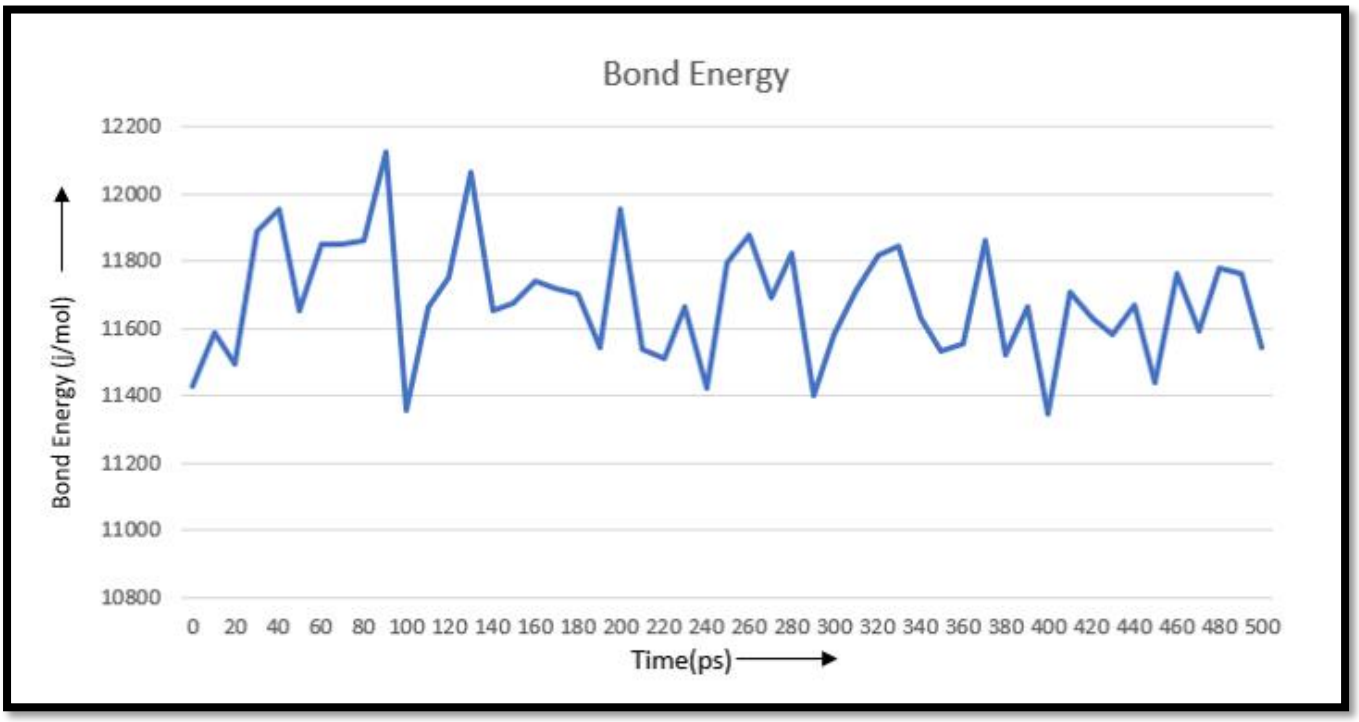

Figure 6. Average Bond Energy between protein and ligand with respect to time (in picoseconds)

Chapter-5

Conclusion

Docking Score: The highest docking score of -12.5 kcal/mol indicates a strong binding affinity between gamma-glutamyl-S-allyl-L-cysteine and the NS5 protein. A lower docking score suggests a more favourable interaction and suggests that the compound has a high potential to act as an inhibitor.

H-Bond Formation: Analysis using Ligplot reveals that 12 hydrogen bonds are formed between the ligand and the protein. This indicates a significant level of interaction between gamma-glutamyl-S-allyl-L-cysteine and the NS5 protein, further supporting its potential as an inhibitor.

RMSD Stabilization: Molecular dynamics simulations conducted using GROMACS show that the system stabilizes after 400ps. The reduced difference in RMSD values between the ligand and the protein backbone suggests a decreased likelihood of the ligand slipping out of the active site. This stability is crucial for the compound to maintain its inhibitory effect.

RMSF Analysis: The Root Mean Square Fluctuation (RMSF) analysis reveals that the exposed portion of the ligand, not fully covered by the protein and exposed to solvent (water), exhibits higher fluctuations compared to the rest of the ligand molecule. This insight provides valuable information about the dynamic behaviour of the ligand within the system.

Molecular Dynamics Simulation Realism: Molecular dynamics simulations are performed under constraints that resemble real-life scenarios, making them more realistic compared to molecular docking. The simulations demonstrate an average of 3.09 hydrogen bonds formed between the ligand and the protein over a time period of 500 picoseconds. These hydrogen bonds contribute to the stability of the interaction. The average energy of the hydrogen bonds is calculated to be 11.68 kJ/mol over the 500ps time period.

Inhibition Potential: The cumulative data strongly suggests that gamma-glutamyl-S-allyl-L-cysteine is a potent inhibitor of the NS5 protein of YFV. However, further in vivo assessment is necessary to validate its inhibitory effects and evaluate its potential as an antiviral agent.

Chapter 6

Future Prospects

The future prospects of the study on gamma-glutamyl-S-allyl-L-cysteine as an inhibitor of the NS5 protein of the Yellow Fever Virus (YFV) and its potential application against other related viruses are quite promising. Here are some potential future directions and prospects for this research:

In Vivo Assessment: Conducting in vivo studies to evaluate the inhibitory effects of gamma-glutamyl-S-allyl-L-cysteine is a crucial next step. These studies involve testing the compound in living organisms, such as animal models, to determine its efficacy, safety, and pharmacokinetic properties. In vivo assessments will provide more realistic and reliable data to support the compound's potential as an antiviral agent. Structure-Activity Relationship (SAR) Studies: Further investigations can be conducted to explore the structure-activity relationship of gamma-glutamyl-S-allyl-L-cysteine and its analogs. Modifying the compound's structure and systematically evaluating its impact on inhibitory activity can help identify key functional groups and optimize its potency as an inhibitor. Mechanistic Studies: Detailed mechanistic studies can provide insights into the specific interactions between gamma-glutamyl-S-allyl-L-cysteine and the NS5 protein or other target proteins. Understanding the precise binding sites, the nature of interactions, and the impact on viral replication processes can guide the design of more effective inhibitors and contribute to the development of targeted antiviral therapies. Drug Delivery Systems: Investigating and developing suitable drug delivery systems can enhance the compound's bioavailability, stability, and targeted delivery to infected cells or tissues. Formulations such as nanoparticles, liposomes, or prodrugs can improve the compound's pharmacokinetic properties and facilitate its effective delivery to the target site. Clinical Trials: If the compound demonstrates promising results in preclinical studies, progressing to clinical trials is the next crucial step. Clinical trials involve testing the compound's safety, efficacy, and dosage regimens in human subjects. This stage is essential for determining its clinical effectiveness and potential as a therapeutic option for viral infections. Evaluation against Other Viruses: Since the NS5 protein of YFV exhibits similarity with the polymerases of Zika and Dengue viruses, exploring the inhibitory effects of gamma-glutamyl-S-allyl-L-cysteine against these viruses is a worthwhile pursuit. It may open up new avenues for combating other mosquito-borne viral infections and contribute to the development of broad-spectrum antiviral strategies. Combination Therapy: Investigating the potential of combining gamma-glutamyl-S-allyl-L-cysteine with other antiviral agents or therapies could enhance its effectiveness. Combination therapies can target different stages of viral replication and increase the likelihood of inhibiting viral replication and reducing the emergence of drug resistance.

Drug Optimization and Development: The findings from this study can serve as a foundation for further optimization and development of gamma-glutamyl-S-allyl-L-cysteine or its analogs as a lead compound for antiviral drug development. This could involve chemical modifications, lead optimization, and formulation development to enhance its pharmaceutical properties and potential for clinical use. Overall, the study on gamma-glutamyl-S-allyl-L-cysteine as an inhibitor of the NS5 protein of YFV holds significant future prospects. Continued research in these areas can contribute to the development of effective antiviral therapies, expanding our arsenal against viral infections and potentially improving public health outcomes.

Chapter-6

References

(1) Gardner, C. L.; Ryman, K. D. Yellow Fever: A Reemerging Threat. (Special Issue: Emerging Pathogens.). *Clin. Lab. Med.* **2010**, *30* (1), 237–260. https://doi.org/10.1016/j.cll.2010.01.001.Yellow.

(2) WHO. *Yellow Fever*; 2019. https://www.who.int/news-room/fact-sheets/detail/yellow-fever.

(3) Douam, F.; Ploss, A. Yellow Fever Virus: Knowledge Gaps Impeding the Fight Against an Old Foe. *Trends Microbiol.* **2018**, *26* (11), 913–928. https://doi.org/10.1016/j.tim.2018.05.012.

(4) Monath, T. P.; Vasconcelos, P. F. C. Yellow Fever. *J. Clin. Virol.* **2015**, *64*, 160–173. https://doi.org/10.1016/J.JCV.2014.08.030.

(5) Waggoner, J. J.; Rojas, A.; Pinsky, B. A. Yellow Fever Virus: Diagnostics for a Persistent Arboviral Threat. *J. Clin. Microbiol.* **2018**, *56* (10), 1–13. https://doi.org/10.1128/JCM.00827-18.

(6) Mukhtar, M.; Arshad, M.; Ahmad, M.; Pomerantz, R. J.; Wigdahl, B.; Parveen, Z. Antiviral Potentials of Medicinal Plants. *Virus Res.* **2008**, *131* (2), 111–120. https://doi.org/10.1016/j.virusres.2007.09.008.

(7) Bookchin, D.; Schumacher, J. The Virus and the Vaccine. *Atl.* **2000**, *10* (5), 94–98.

(8) Selisko, B.; Papageorgiou, N.; Ferron, F.; Canard, B. Structural and Functional Basis of the Fidelity of Nucleotide Selection by Flavivirus RNA-Dependent RNA Polymerases. *Viruses* **2018**, *10* (2). https://doi.org/10.3390/v10020059.

(9) Coutard, B.; Barral, K.; Lichière, J.; Selisko, B.; Martin, B.; Aouadi, W.; Lombardia, M. O.; Debart, F.; Vasseur, J.-J.; Guillemot, J. C.; Canard, B.; Decroly, E. Zika Virus Methyltransferase: Structure and Functions for Drug Design Perspectives. *J. Virol.* **2017**, *91* (5), 1–15. https://doi.org/10.1128/jvi.02202-16.

(10) Dong, H.; Chang, D. C.; Xie, X.; Toh, Y. X.; Chung, K. Y.; Zou, G.; Lescar, J.; Lim, S. P.; Shi, P. Y. Biochemical and Genetic Characterization of Dengue Virus Methyltransferase. *Virology* **2010**, *405* (2), 568–578. https://doi.org/10.1016/J.VIROL.2010.06.039.

(11) Hyde, J. L.; Diamond, M. S. Innate Immune Restriction and Antagonism of Viral RNA Lacking 2'-O Methylation. **2015**, *479–480* (January), 66–74.

(12) Lim, S. P.; Noble, C. G.; Shi, P. Y. The Dengue Virus NS5 Protein as a Target for Drug Discovery. *Antiviral Res.* **2015**, *119*, 57–67. https://doi.org/10.1016/J.ANTIVIRAL.2015.04.010.

(13) Rai, S. K.; Sharma, M.; Tiwari, M. Inhibitory Effect of Novel Diallyldisulfide Analogs on HMG-CoA Reductase Expression in Hypercholesterolemic Rats: CREB as a Potential Upstream Target. *Life Sci.* **2009**, *85* (5–6), 211–219. https://doi.org/10.1016/J.LFS.2009.05.020.

(14) Reinhart, K. M.; Coleman, C. I.; Teevan, C.; Vachhani, P.; White, C. M. Effects of Garlic on

Blood Pressure in Patients with and without Systolic Hypertension: A Meta-Analysis. *Ann. Pharmacother.* **2008**, *42* (12), 1766–1771. https://doi.org/10.1345/aph.1L319.

(15) Arroll, B. Non-Antibiotic Treatments for Upper-Respiratory Tract Infections (Common Cold). *Respir. Med.* **2005**, *99* (12), 1477–1484. https://doi.org/10.1016/j.rmed.2005.09.039.

(16) Ebrahimi, M.; Hassan, Z. M.; Mostafaie, A.; Mehrjardi, N. Z.; Ghazanfari, T. Purified Protein Fraction of Garlic Extract Modulates Cellular Immune Response against Breast Transplanted Tumors in BALB/c Mice Model. *Cell J.* **2013**, *15* (1), 65–74.

(17) Casella, S.; Leonardi, M.; Melai, B.; Fratini, F.; Pistelli, L. The Role of Diallyl Sulfides and Dipropyl Sulfides in the in Vitro Antimicrobial Activity of the Essential Oil of Garlic, Allium Sativum L., and Leek, Allium Porrum L. *Phyther. Res.* **2013**, *27* (3), 380–383. https://doi.org/10.1002/ptr.4725.

(18) Altonsy, M. O.; Andrews, S. C. Diallyl Disulphide, a Beneficial Component of Garlic Oil, Causes a Redistribution of Cell-Cycle Growth Phases, Induces Apoptosis, and Enhances Butyrate-Induced Apoptosis in Colorectal Adenocarcinoma Cells (HT-29). *Nutr. Cancer* **2011**, *63* (7), 1104–1113. https://doi.org/10.1080/01635581.2011.601846.

(19) Karmakar, S.; Roy Choudhury, S.; L. Banik, N.; K. Ray, S. Molecular Mechanisms of Anti-Cancer Action of Garlic Compounds in Neuroblastoma. *Anticancer. Agents Med. Chem.* **2012**, *11* (4), 398–407. https://doi.org/10.2174/187152011795677553.

(20) Huang, Y. S.; Xie, N.; Su, Q.; Su, J.; Huang, C.; Liao, Q. J. Diallyl Disulfide Inhibits the Proliferation of HT-29 Human Colon Cancer Cells by Inducing Differentially Expressed Genes. *Mol. Med. Rep.* **2011**, *4* (3), 553–559. https://doi.org/10.3892/mmr.2011.453.

(21) Aggarwal, B.; Prasad, S.; Sung, B.; Krishnan, S.; Guha, S. *Prevention and Treatment of Colorectal Cancer by Natural Agents from Mother Nature*; 2013; Vol. 9. https://doi.org/10.1007/s11888-012-0154-1.

(22) Kodera, Y.; Ushijima, M.; Amano, H.; Suzuki, J. I.; Matsutomo, T. Chemical and Biological Properties of S-1-Propenyl-l-Cysteine in Aged Garlic Extract. *Molecules* **2017**, *22* (4), 1–18. https://doi.org/10.3390/molecules22040570.

(23) Yoo, D. Y.; Kim, W.; Nam, S. M.; Yoo, M.; Lee, S.; Yoon, Y. S.; Won, M. H.; Hwang, I. K.; Choi, J. H. Neuroprotective Effects of Z-Ajoene, an Organosulfur Compound Derived from Oil-Macerated Garlic, in the Gerbil Hippocampal CA1 Region after Transient Forebrain Ischemia. *Food Chem. Toxicol.* **2014**, *72*, 1–7. https://doi.org/10.1016/J.FCT.2014.06.023.

(24) Dubankova, A.; Boura, E. Structure of the Yellow Fever NS5 Protein Reveals Conserved Drug Targets Shared among Flaviviruses. *Antiviral Res.* **2019**, *169*, 104536. https://doi.org/10.1016/J.ANTIVIRAL.2019.104536.

Plagiarism Report

Shalesh Thesis

ORIGINALITY REPORT

7%
SIMILARITY INDEX

PRIMARY SOURCES

1	shodhbhagirathi.iitr.ac.in:8081 Internet	72 words — 1%
2	www.xiahepublishing.com Internet	38 words — 1%
3	Iqbal Azad, Tahmeena Khan, Naseem Ahmad, Abdul Rahman Khan, Yusuf Akhter. "Updates on drug designing approach through computational strategies: a review", Future Science OA, 2023 Crossref	24 words — < 1%
4	coek.info Internet	24 words — < 1%
5	www.indianjournals.com Internet	24 words — < 1%
6	www.ehow.com Internet	22 words — < 1%
7	Suchita Chaudhry, Sathyamangalam Swaminatha, Navin Khanna. "The Murine Charged Multivesicular Body Protein 2A, CHMP2A Interacts with the 5â€™ and 3â€™ Terminal Regions of Dengue Virus Complementary Minus-strand RNA", American Journal of	20 words — < 1%

8	github.com Internet	20 words — < 1%
9	Younis Ahmad Hajam, Premlata, Raksha Rani, Riya Sharma, Preeti Sharma, Rajesh Kumar. "Immunogenetic mechanism driving neurological and neurodegenerative disorders", Elsevier BV, 2022 Crossref	19 words — < 1%
10	dspace.dtu.ac.in:8080 Internet	17 words — < 1%
11	repository.up.ac.za Internet	16 words — < 1%
12	link.springer.com Internet	14 words — < 1%
13	www.biorxiv.org Internet	13 words — < 1%
14	www.sanofi.com Internet	12 words — < 1%
15	kupdf.net Internet	11 words — < 1%
16	scholarworks.uno.edu Internet	11 words — < 1%
17	Vinícius de Almeida Paiva, Isabela de Souza Gomes, Cleiton Rodrigues Monteiro, Murillo	10 words — < 1%

	Internet	10 words — < 1%
19	dr.ntu.edu.sg Internet	10 words — < 1%
20	researchonline.ljmu.ac.uk Internet	10 words — < 1%
21	vital.seals.ac.za:8080 Internet	10 words — < 1%
22	www.researchgate.net Internet	10 words — < 1%
23	www.uochb.cz Internet	10 words — < 1%

EXCLUDE QUOTES	ON	EXCLUDE SOURCES	OFF
EXCLUDE BIBLIOGRAPHY	ON	EXCLUDE MATCHES	< 10 WORDS