COMPUTATIONAL DRUG DISCOVERY AGAINST HEPATITIS B VIRUS CORE PROTEIN

M.Sc. Thesis

2023

Submitted to

Central Department of Biotechnology

Tribhuvan University

Kirtipur, Kathmandu, Nepal

For partial fulfillment of the requirement for the

Master of Science in Biotechnology

By

Ramesh Raj Itani

Registration No: 5-2-37-903-2013

Supervisors

Assistant Professor Dr.Smita Shrestha

Visiting Faculty Dr. Pramod Aryal

Date :

RECOMMENDATION

This is to certify that the research work entiled "**COMPUTATIONAL DRUG DISCOVERY AGAINST HEPATITIS B VIRUS CORE PROTEIN"** has been carried out by **Mr.Ramesh Raj Itani** under my supervision.

This thesis work was performed for the partial fulfillment of the Master of Science in Biotechnology under the course code BT 621. The result presented here is his original findings. We, hereby, recommend this thesis for final evaluation.

…….

Supervisor

Asistant Professor Dr.Smita Shrestha

Central Department of Biotechnology

Tribhuvan University,

Kirtipur, Kathmandu, Nepal

…….

Supervisor

Dr. Pramod Aryal

Central Department of Biotechnology

Tribhuvan University,

Kirtipur, Kathmandu, Nepal

Date :

CERTIFICATE OF EVALUATION

This is to certify that this entitled"**COMPUTATIONAL DRUG DISCOVERY AGAINST HEPATITIS B VIRUS CORE PROTEIN"**presented to evaluation committed by **Mr.Ramesh Raj Itani** is found satisfactory for the partial fulfillment of Master of Science and Technology.

…….
Head of Department
Prof. Dr. Krishna Das Manandhar
Central Department of Biotechnology
Tribhuvan University,
Kirtipur, Kathmandu, Nepal

…….
Supervisor
Asistant Professor Dr.Smita Shrestha
Central Department of Biotechnology
Tribhuvan University,
Kirtipur, Kathmandu, Nepal

…….
Supervisor
Dr. Pramod Aryal
Central Department of Biotechnology
Tribhuvan University,
Kirtipur, Kathmandu, Nepal

…….
Internal Examiner
Alina Shri Sapkota
Assistant Professor
Central Department of Biotechnology
Tribhuvan University,
Kirtipur, Kathmandu, Nepal

……
External Examiner
Prof. Dr. Narayan Prasad Adhikari
Central Department of Physics
Tribhuvan University,
Kirtipur, Kathmandu, Nepal

ACKNOWLEDGEMENT

I would like to express my special thanks gratitude to my supervisor, Dr.Smita Shrestha as well as Dr.Pramod Aryal, for his continous support and guidance throughout the completion of my thesis. His encouragement and motivation during the difficulties I faced during my thesis period was invaluable.

I would also like to extend my special thanks to my friends Samiran Subedi and Ashish Pokhrel for providing immense support during my thesis work.

I am also thankful to Prof. Dr. Krishna Das Manandhar, the HOD of Central Department of Biotechnology at Tribhuvan University, for allowing me to complete my thesis.

Finally, last but not the least I would like to express my utmost gratitude to my parents, my batch mates, seniors for their ongoing support, and encouragement throughout and my thesis. Their unwavering faith in me has given me the strength and determination to complete my thesis. I am truly grateful for their love and support.

Thanks for all your support and encouragement!!!

Ramesh Raj Itani

LIST OF ABBREVATIONS

ADME/Tox	Absorption, Distribution, Metabolism, Excretion, Toxicity
CADD	Computer Aided Drug Design
DFT	Density function theory
FDA	Food and Drug Administration.
LBDD	ligand based drug design
RO5	Rule of Five
RCSB	Research Collaboratory for Structural Bioinformatics
SBDD	Structure based drug design
SBVS	structure based virtual screening
TOX	Toxicity
PDB	Protein data bank
CCC	covalently closed circular
DNA	Deoxy ribonucleic acid
PSA	polar surface area
WHO	World health organization

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Hepatitis B virus infection has been a major global health concern today.The present available drugs are not able to cure the hepatitis B completely.So the scientific community in world are in search of new drugs that can fulfill the void of existing non performing drugs. The problem of the low efficacy of present drugs and development of new resistant mutant virus needs to be solved immediately for the advancement in the cure of hepatitis B infected patients. Traditional drug discovery process is time and money consuming process so computer aided drug designing can be a good alternative for drug discovery process. In this study ligand library from different databases like asinex, zinc15, selleckchem are used.Ligands from these databases are screened for ADMET properties in order to ensure the drugability characters. The selected ligands are subjected to molecular docking purpose with the target protein i.e. hepatitis B core protein under the PDB name 5T2P .The ligands that have higher binding affinity with target protein was selected for further screening .hMAT1A screening was done with selected ligands to ensure that selected lead molecules donot harm liver.The lead compounds was selected after hMAT1A screening and subjected to density function theory to study its chemical properties computationally. Benzofuranonesshows the higher binding affinity with target protein and selected as lead molecules after hMAT1A screening.

Keywords: ADMET, CADD, Molecular Docking, Lead compound, FDA, hMAT1A

1. INTRODUCTION

1.1. Background

Hepatitis B virus is the major causative agent of the chronic and acute viral hepatitis as well as hepatocellular carcinoma(HCC) and liver cirrhosis.More than 260 million peoples are chronically infected world wide and about 6,80,000 people die every year due to the complications of HBV- related chronic liver diseases(WHO). Interferon alpha (IFN) and nucleos(t)ide analogues are currently the standard drugs against hepatitis B virus infections, although these antiviral agents have limited efficacy .The recently used FDA drugs cannot completely cure the hepatitis B infections due to their inability to eradicate covalently closed circular DNA (ccc) from nuclei of infected hepatocytes (Hyo gyeong na *et al*.,2020)**.**The drugs that can suppress the lifecycle or replication of viruses are the choices of drugs against hepatitis B virus.TAF (Tenofovir alafenamide fumarate), a prodrug of tenofovir has been recently approved by FDA for the treatment of hepatitis B virus. However, long-term use of nucleos(t)ide analogues may result in the development of drug resistant viruses .Covalently closed circular DNA (cccDNA) and chromosomally integrated HBC-DNA, which hide in the hosts' nucleus and evades current medicines, make it difficult to treat HBV. Therefore, the development of novel therapeutic approaches that disrupt various viral life cycle stages, such as viral entry inhibitors, polymerase inhibitors, capsid and assembly inhibitors, virus release blockers, and inhibitors of cccDNA formation with agents that enhance or activate human response considering disease stages, and can be combined with existing therapies, will result in a full or functional cure(Mohcine *et al*;2020).

Using insilico methods along with experimental works can reduce costs and time. Virtual screening (VS) can be one of the important computational methods for the identification of lead compunds during target based drug discovery. There has been several ongoing researches on drug discovery against hepatitis B virus core protein that has important role in maturation of virus or virion formation.

Computer aided drug designing (CADD) could be a promising tool for drug discovery against hepatitis B virus core protein. Experimental laboratory works is tedious job as it takes long time in trial and costly process also.CADD helps to select the drugs virtually from large databases and the suggeste drugs from virtual screenings can be used in trial so that time and money will be saved as well as the rate of sucesss of drugs efficiency will be increased.Recently drugs repurposing has been popular in scientific world since corona pandemic as various drugs that has been used previously to treat other disesases has been used to treat COVID.Principle of dug repurposing also could be a new dea to find out lead compounds that could work better in other diseases. Ligand from the various databases are used to create ligand library .Various ADMET screenings are done from the ligand databases in order to ensure the drugability properties of lead molecules.The target protein i.e. core proteins are used to dock with various ligand and the docking score are used to find out the lead molecules.Higher the binding affinity between ligand and protein higher the probality of the ligand to be a lead compound. hMAT1A screenings are done to make sure the drugs would not damage the liver.

1.2 Current Studies

In this research, we mainly focused on the computer aided drug designing (CADD) for the drug discovery against hepatitis B virus core protein. We intended to solve the time consuming and more tedious work of traditionaldrug discovery by doing computational work. For this research large sets of ligands are sorted to form ligand library after ADMET screening and molecular docking was carried out. The lead compound is selected on the basis of docking score and hMAT1A screening and various ligand protein interactions are visualized in different softwares.

1.3 Hypothesis

1.3.1 Null Hypothesis

The drug candidates to inhibit the hepatitis B virus core protein will not be identified through molecular docking techniques.

1.3.2 Alternative Hypothesis

The drug candidates to inhibit the hepatitis B virus core protein will be identified through molecular docking techniques.

1.4 Objectives

1.4.1 General Objective

- To discover novel lead compounds against hepatitis B virus core proteins.

1.4.2 Specific Objectives

- To identify potential drugs against the hepatitis B virus core protein from various ligand databases.
- To create ligand libraries from various databases by ADMET filtering.
- To perform Molecular docking against target protein and analyze protein ligand interactions.
- To identify the potential lead molecules after molecular docking with various screening parameters.
- To test the efficiency of lead compound to inhibit the mutant form of protein.

1.5 Rationale

Hepatitis B infection has been a global health concern today as more than 260 million people world wide are infected and 6,80,000 people are dying.The increasing emergence of drug resistant mutant of viruses and low efficacy of currentlty available FDA drugs to completely cure hepatitis infection demands the urgent need for the

development of novel antihepatitis B drugs. To address this challenge, the current study aims to develop a more efficient potential drug candidates through virtual screening of large sets of ligand databases.Computer aided drug designing could be a promising tool to solve this problem as it is more economical and time saving process than the traditional drug discovery process.CADD can increase the success rate of drugs in trial as it can recommends drugs with various screening of drugability parameters.

1.6 Scope of the study

The present study focuseson finding potential lead molecules with screening of natural products against hepatitis B core protein. This study involves the efficacy of lead compounds against the mutant protein of same viruses.

2. LITERATURE REVIEW

2.1. Literature review related to Hepatitis B virus

HBV is a member of the family Hepadnaviridae. It is an enclosed virus with a genome made of relaxed circular (RC) DNA, which is partially double stranded. It has eight genotypic class from A-H with distinct geographical distribution. It has a genome size of 3200bp and diameter 42 nm also called as dane particle that istranscribed to generate four known transcripts (3.5 kb, 2.4 kb, 2.1kb and 0.7 kb size) .One DNA strand is complete, except for a small nick (the minus strand), and the other is short and incomplete (the plus strand). The minus strand has four overlapping open reading frames (ORFs; Figure 1) that correspond to the following genes: (1) the preS/S gene, which codes for the large, middle, and small HBsAgs; (2) the P gene for DNA polymerase/reverse transcriptase (POL); (3) the X gene for the X protein, a crucial regulator during the natural infection process, which has transcriptional trans-activation activity and is necessary to start and maintain HBV replication ; 4) the precore/core gene, which produces the HBcAg or core protein that makes up the capsid as well as the HBeAg protein, which is produced on liver cells and secreted into the serum but is not part of the virus itself(Trevor Graham bell *et al*;2016).

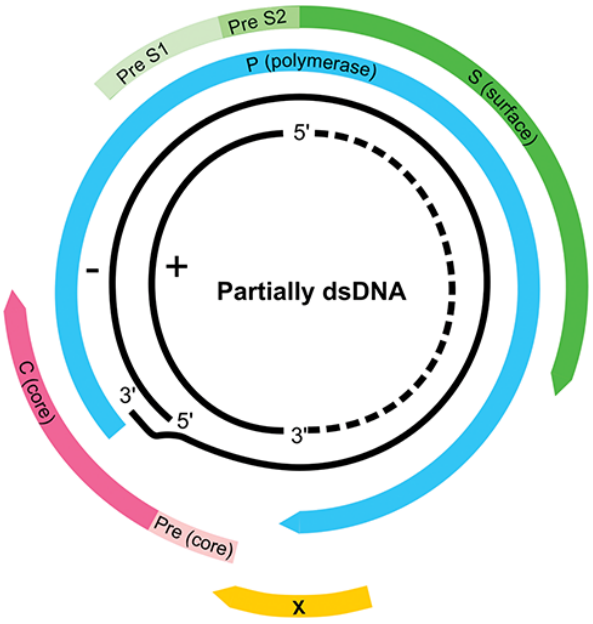


Figure 1 : HBV gene structure (intech open)

The genome of hepatitis B virus (HBV). The partially double-stranded DNA (dsDNA) with the complete minus (−) strand and the incomplete (+) strand. The four open reading frames (ORFs) are shown: precore/core (preC/C) that encodes the e antigen (HBeAg) and core protein (HBcAg); P for polymerase (reverse transcriptase), PreS1/PreS2/S for surface proteins [three forms of HBsAg, small (S), middle (M), and large (L)] and X for a transcriptional trans-activator protein.

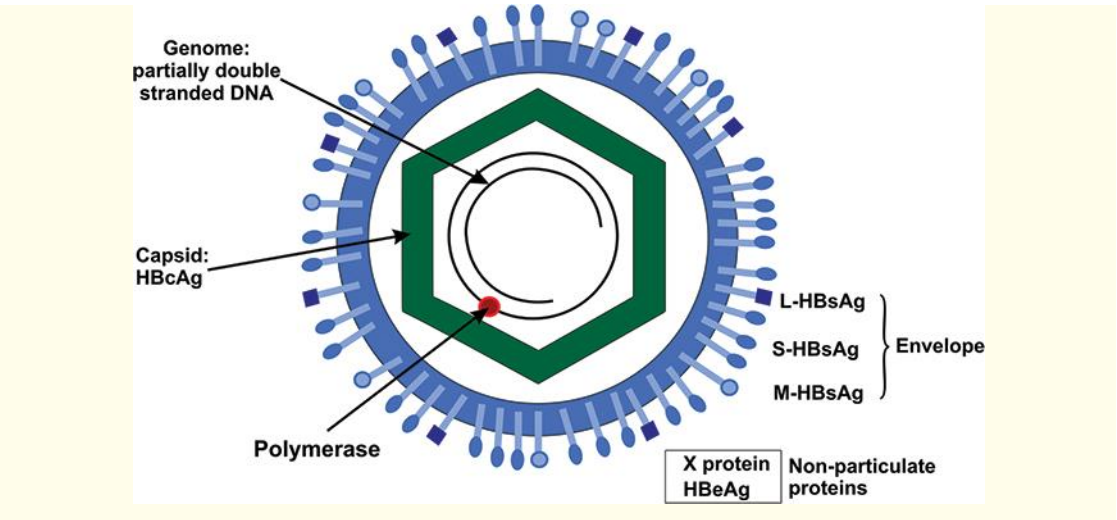


Figure 2 : Hepatitis B structure (intech open)

Schematic representation of hepatitis B virus (HBV), showing the structure of the virion, composed of a partially double-stranded DNA genome, enclosed by a capsid, comprised of HBcAg and surrounded by a lipid envelope containing large (L)-HBsAg, middle (M)-HBsAg, and small (S)-HBsAg. The virus also expresses two non-particulate proteins X protein and HBeAg.

2.1.1Regulatory elements of HBV

Every single nucleotide of the HBV genome is necessary for the translation of a protein and may also be part of one of the regulatory elements of HBV, which overlap with protein expressing regions. The regulatory elements include the S1 and S2 promoters, which overlap both the preS region and polymerase ORFs; the preC/pregenomic promoter, which includes the basic core promoter (BCP) and overlaps the X and preC ORF; and the X promoter. There are two enhancers (enhancer I and enhancer II) as well as cis-acting negative regulatory elements (URR: upper regulatory region, CURS:

core upstream regulatory sequence, NRE: negative regulatory element). These regulatory elements control transcription (Kramvis A. *et al;* 1999).

The hepatitis B surface antigen (HBsAg), which circulates in the blood in titres that can exceed 1012 HBsAg particles/mL, and the production of excess non-DNA containing viral envelope material are novel features of hepadnavirus replication. Normally, reversetranscription is only seen with retroviruses. Since the viral reverse transcriptase (rt) is incapable of proofreading, the virus generates a population of closely related variations known as a quasispecies when it replicates. Due to the presence of a resistant viral sub-population when the virus is threatened by the immune system or the introduction of an antiviral drug, this diversity protects the survival of HBV.

2.1.2 HBV core protein

The Pre-C/C ORF encodes the core protein, which is the major polypeptide of the nucleocapsid and expresses the HBV core antigen (HBcAg).The Hbv core protein length varies on viral genotype and can be either 183, 185 or 195 amino acids long. The HBc protein has two distinct domains.

The first one is necessary for the construction of the 32-nm nucleocapsid and spans amino acid residues 1-144. The second starts about position 140 and extends to the C-terminus. It is enriched in arginine clusters, forming a protamine-like domain that controls nucleic acid binding and is essential for viral encapsidation and DNA replication(Ganem and Schneider;2001). There is a potential nuclear localization sequence in this area. When expressed, the core protein's numerous hydrophilic and charged amino acids are phosphorylated (Kann and Gerlich; 2005). The nuclear localization of the core protein is negatively regulated by phosphorylation of serine 170-172 between arginine clusters 3 and 4. This blockage of nucleic acid binding may result from this. Pregenomic RNA is used to translate the HBc protein, which contains HBcAg epitopes that react with HBeAg epitopes.The HBc protein might actively participate in controlling viral transcription.

This theory was developed as a result of investigations indicating that HBcAg collects in various cellular compartments during various stages of infection. High HBV DNA levels are reported during the immunotolerant phase, which is linked to a mostly nuclear distribution of HBcAg. In contrast, HBcAg becomes predominately cytoplasmic during the immunological clearance phase, along with a reduction in viral load. Hence, during the immunological control phase, HBcAg is no longer detectable. Moreover, HBcAg shows preferential affinity for HBV cccDNA and has the ability to modify nucleosomal packing to influence viral transcription. These findings show that HBV proteins have several functions.

2.1.3 Replication cycle of HBV

The replication cycle of HBV and other members of the Hepadnaviridae family is peculiar. Pregenomic RNA (pgRNA), an intermediary RNA known for reverse transcription, is used by these DNA viruses to replicate themselves(Summers J, Mason W ;1982) .The sodium taurocholate cotransporting polypeptide (NTCP), a multiple transmembrane transporter that is mostly expressed in the liver, allows entry into the cell(Yan H, Zhong *et al*; 2012).

Upon entry, the virion is uncoated and its core particle is actively carried to the nucleus, where it releases a circular DNA molecule with a partly double strand relaxed structure. The host RNA polymerase II uses the covalently closed circular molecule of DNA (cccDNA) as a template for transcription after the viral polymerase closes the single-stranded (Kock J, Rosler C; 2010). The translation of the mRNAs takes place in the cytoplasm.

The encapsidation signal has a bipartite stem-loop structure made up of an apical loop, an upper and lower stem, and a bulge. In addition to encapsidation, performs a number of other tasks (Kramvis A *et al.,* 2007). It works in template restriction to prevent the encapsidation of any RNA, and it also participates in the activation of the viral polymerase to prevent random reverse transcription. Moreover, it starts the reverse transcription process. By attaching to the bulge of, the polymerase or reverse transcriptase functions as a primer for RNA-directed DNA synthesis. The bulge of DNA

synthesizes the first three nucleotides of the negative stand, which are then transported to an acceptor site.The HBV replication was described in brief below

- **Attachment and Penetration**

HBV enters the cell cytoplasm once the virion binds to its receptor on the surface membrane of the hepatocyte in the initial stage of infection. In order to enter cells, the AGL of HBsAg must first connect to hepatocyte-associated heparan sulphate proteoglycans (HSPG). The co-receptor stage that facilitates approximation is this one (Schulze *et al*., 2005).

The sodium taurocholate co-transporting polypeptide (NTCP), the entrance receptor for HBV, of the pre-S1 receptor-binding domains. The enterohepatic circulation is how the NTCP typically keeps bile acid balance in check. The human SLC10A1 gene, which is located on the long arm of chromosome 14, encodes NTCP. HBV's pre-S1 domain in its envelope.

In the N-terminal end of the protein, there are two areas within 75 amino acids that are in charge of NTCP binding (Ni *etal.,* 2014). To allow for virion infectivity, these sequences must be myristoylated. Following binding, viral nucleocapsids are transported into the cytoplasm via receptor-mediated endocytosis.

- **2. Genomic RC DNA Conversion into cccDNA and Transcriptional Activation of the Viral Minichromosome.**

The nuclear membrane is where the nucleocapsids are conveyed, where they uncoil and deposit their genetic cargo (Rabe *et al*; 2003). The viral minichromosome is created as a result of the host cell enzymes releasing the RC DNA and converting it into cccDNA.

Virulent persistence is caused by the transcription of all viral mRNA. Release of the viral polymerase protein from the 5’ end of the minus strand is the initial step in this conversion, and it is accomplished by using the host DNA repair enzyme tyrosyl-DNA-phosphodiesterase 2 (TDP2). There are still undetermined steps in the creation of cccDNA.

Topoisomerase I and II, as well as additional host domain modification enzymes, are likely involved but remain to be clarified. Host RNA polymerase II is used to create mRNA transcripts of both genomic and subgenomic length from the HBV minichromosome. All transcripts have a positive orientation, are heterogeneous, and are limited to both the 3’ end and the 5’ ends are polyadenylated. The subgenomic transcripts only serve as mRNAs for the translation of the X protein and envelope protein subunits. The precore, core, and polymerase proteins are encoded by the two genomic transcripts, which are longer than the genomic length. The pre-C/C mRNA is produced early in the transcription and translation process and aids in the translation of the precore protein, which is eventually processed and released as HBeAg. It doesn't participate in reverse transcription. The pregenomic RNA (pgRNA), in contrast, is multipurpose and acts as a template for both the translation of HBcAg and HBV Pol as well as reverse transcription into the negative DNA strand.The HBV minichromosome requires the synthesis of HBx to maintain transcriptional activity, primarily by preventing Smc5/6 from hyperchromatinizing it.

- **Genomic Replication via Reverse Transcription**

After the pgRNA and newly translated viral polymerase are packaged into subviral core particles and form replication complexes in the cytosol, HBV genomic replication begins. During reverse transcription, the HBV nucleocapsid is involved. The terminal protein domain of the polymerase specifically binds to the 'bulge' region of a distinctive RNA stem loop structure, known as the epsilon () loop, at the 5’ end of the pgRNA to prime reverse transcription while the polymerase is being translated off the same pgRNA molecule that it will eventually be packaged with(Bartenschlager andSchaller;1992). The encapsidation signal that cytoplasmic core protein dimers combine into nucleocapsids is the epsilon loop.

The polymerase undergoes a conformational shift when translation is finished, resulting in enzymatic activation, with the terminal protein domain now actively priming DNA synthsis (Zoulim and Seeger; 1994). The Pol-oligonucleotide (PolG-A-A) complex subsequently moves to the 3’ end of the pgRNA, where it complements the sequence of a direct repeat (DR-1) region. Then, until it reaches the 5’ end of the

pgRNA molecule, the creation of the negative DNA strand proceeds, generating a short terminal 8–9 nucleotide redundancy. While reverse transcription is taking place, the polymerase's RNaseH activity breaks down the pgRNA but retains the DR-1 sequence's 5'-capped terminal 18 nucleotides. This fragment contains a six-nucleotide homology to the direct repeat sequence, allowing the minus strand to be circularized. The 18 nucleotide-capped RNA structure is subsequently moved to a second DR sequence (DR-2) on the newly formed minus strand's 5' end, where it serves as a primer for plus strand synthesis using the minus strand as a template. The plus strand continues to be synthesized until it is between 50 and 70 percent of the length of the minus strand.

- **Virion Assembly and Release**

At the ER, the viral envelope and HBsAg subviral particles are created and assembled before budging into the lumen of the cell. Independent of the envelope protein, the HBc protein is created in the cytosol and assembled.

When the nucleocapsid is phosphorylated, which occurs in conjunction with the start of minus strand production, immature nucleocapsids harboring pgRNA-Pol can begin envelopment. Upon leaving the cell, the formed nucleocapsids with genomic RC DNA are then selectively wrapped. Genomes that have started plus strand synthesis and finished minus strand synthesis are more likely to be exported. Pre-S1 to S must be present at a key relative molar ratio for replicating cores to be enveloped; otherwise, aberrant virions develop, which hinder release. The endosomal sorting complex required for transport (ESCRT) network of proteins is a cellular network that aids in the formation of multivesicular bodies and is involved in daughter cell cytokinesis. HBV usurps this network, like many other enveloped viruses do. This system's various parts play crucial roles in the HBV replication cycle. Particularly, virion budding and egress require ESCRT-III and the VPS4 ATPase, while the production of mature nucleocapsids needs the ESCRT-II protein (Ganem and Schneider;2001).

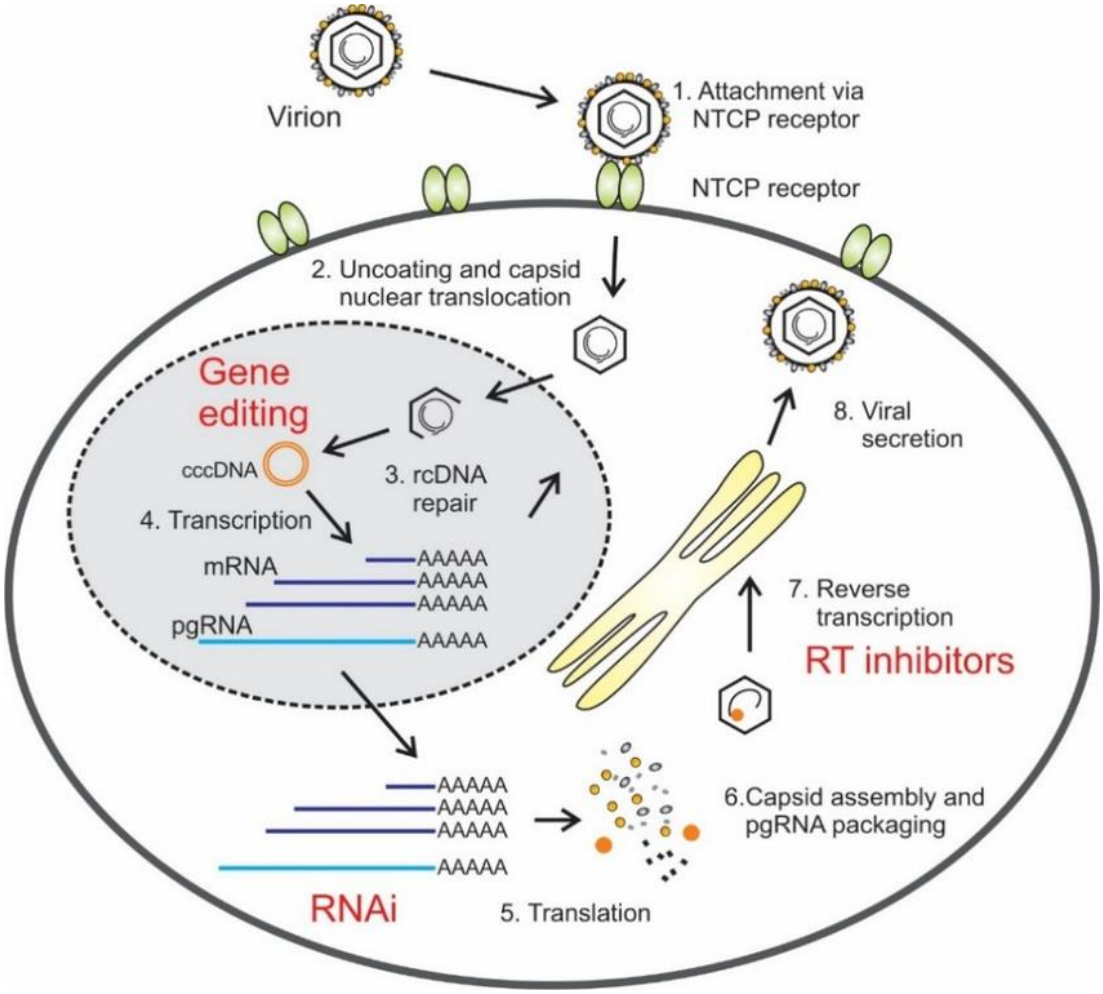


Figure 3 : Replication cycle of hepatitis B virus (Mohube maepa *et al*; 2015)

2.2 Literature review related to Computer Aided Drug Discovery (CADD)

2.2.1 Computer Aided Drug Discovery (CADD)

Using computer algorithms, computer-aided drug discovery is a process for developing drugs and analyzing their effectiveness against biological targets. Structure-based drug designing and ligand-based drug designing are two examples of computer-aided drug discovery that properly estimate the structure and makeup of a biological target, hence enhancing the drug development process.CADD approach saves time; it is fast and cost-effective .Four steps can be used using the CADD approach: To find hits or leads, one must first screen a small molecule library against the target using a virtual screening (VS) protocol. Next, one must examine the specificity of the hits from VS using molecular docking in the active site of other known

targets, and finally, one must predict the ADMET properties of the selected hit using in silico techniques. (4) Contributes to the leads' optimization by creating better compounds for synthesis and testing.

The 3D structure of the pharmacological target is necessary for SBDD. One can download the target structure from the protein databank (PDB) if the protein structure has been solved using crystallography and (2) utilize molecular modeling (a bioinformatics tool) to estimate the protein structure if it has not yet been determined.

According to Kapetanovic (2008), CADDD entails:

(1) The use of computing power to speed up the procedures involved in drug discovery and development.

2) Finding and refining novel medications by utilizing chemical and biological data regarding targets and/or ligands.

(3) Designing filters for undesired chemicals with poor activity and/or poor absorption, distribution, metabolism, excretion, and toxicity in silico, which makes it easier to choose the candidates with the best chances of success.

Benefits of CADD

The following are the key benefits of drug discovery using CADD

i) From huge compound libraries, a smaller group of compounds are chosen for experimental testing.

(ii) Lead compound optimization increases the drug metabolism and pharmacokinetics (DMPK) features including absorption, distribution, metabolism, excretion, and the potential for toxicity (ADMET).

(iii) Creating new chemotypes from pieces of existing ones or "growing" beginning molecules one functional group at a time are two ways to create novel chemicals (Veselovsky and Ivanov;2003).

(iv) CADD can replace conventional experimentation, which uses animal and human models, saving time and money (Mallipeddi *et al*., 2014).

v) Decreases the likelihood of drug resistance, which would encourage the development of lead medicines that would specifically address the underlying cause.

(vi) CADD facilitates the development of high-quality datasets and libraries that can be tailored for high molecular diversity or similarity (Ou-Yang *et al*., 2012).

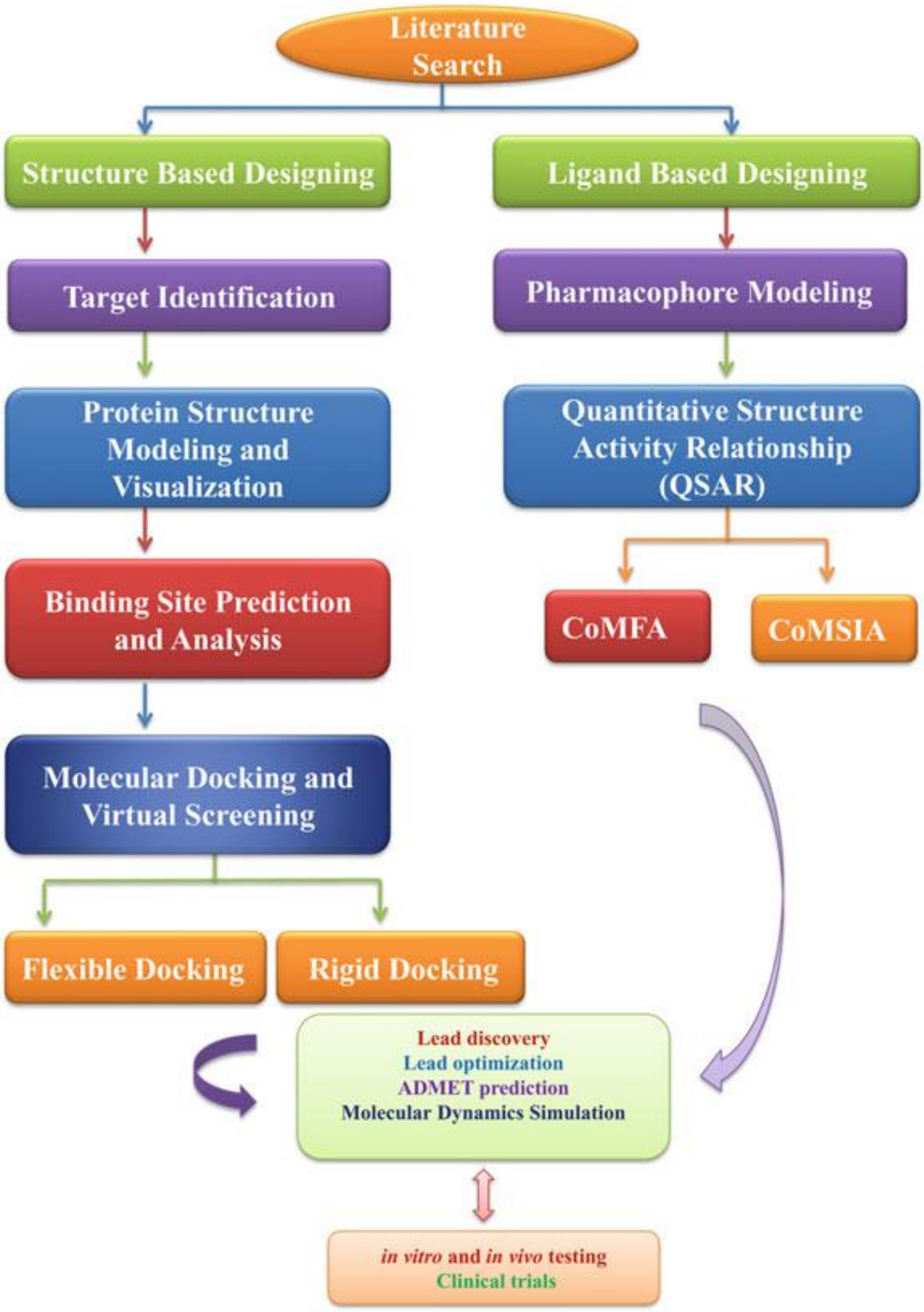


Figure 4: Application of molecular modeling approaches in drug discovery and design (Dev buksh singh; 2020)

There are two methods of CADD: Structure based drug design (SBDD) and ligand based drug design (LBDD).

2.2.2 Structure based drug design

In order to optimize the bound ligand or a group of congeneric molecules, this technique makes use of knowledge of the three-dimensional structure of a receptor complexed with a lead molecule (Lionta E., 2014).Understanding how receptors and ligands interact is crucial. X-ray crystallography, NMR, or homology modeling can be used to determine the structural information. A model with a specific structure can be used by a medicinal chemist to calculate the activity of a given molecule. De novo drug design and virtual screening are both steps in the process known as "structure-based drug design" (SBDD). These techniques offer a highly effective alternate strategy for discovering and creating new medication designs. Drug chemical compounds are computationally tested against known target structures via virtual scanning. Rational drug design is very expensive and effective in traditional, advanced, or legacy drug design and development.In order to create new biologically active molecules, the Structure-Based Drug Design (SBDD) method makes use of protein three-dimensional (3D) structural data of a target molecule obtained using cryo-electron microscopy (EM), NMR, X-ray crystallography, and computational methods like homology modeling and molecular dynamic (MD) simulation. Potential compounds are identified based on the ligand-receptor complex's affinity and whether they possess the requisite properties for the desired pharmacological and therapeutic effects. Hence, SBDD has been a promising approach for identifying lead compounds and medications. Structure-based virtual screening (SBVS), molecular docking, and molecular dynamics (MD) simulations are the most frequently utilized computational methods in SBDD. These methods leverage the target protein's 3D structure and knowledge of the disease at the molecular level to find drugs more quickly, efficiently, and specifically. A potential strategy for drug discovery and improvement has been SBDD.

Steps for structure based drug design

(i) Identification of the target protein

A critical first stage in the drug development process is the selection of a suitable drug target. The target should be capable of binding to tiny compounds and closely connected to human disorders. The first and most crucial phase in the drug discovery process is target identification. Proteins are selected as pharmacological targets in drug development because of their great specificity, potency, and low toxicity. Kinases, proteases, and peptides are the most often chosen protein types as therapeutic targets. A three-dimensional (3D) structure of the target protein is necessary in order to use a structure-based method for drug discovery. Cryo-electron microscopy, NMR spectroscopy, and X-ray crystallography can all be used to obtain this 3D structure, which is then saved in the Protein Data Bank (PDB).If the target protein's 3D structure cannot be determined through experimental means, there are a number of computational ways for doing it. These techniques include protein threading or fold recognition, homology modeling, ab initio techniques.

(ii) Identification of active site

Finding the binding pocket, also known as the active site, is essential after the target protein has been identified. In this situation, a ligand that can bind to the target protein and inhibit it is regarded as a lead molecule or lead contender. Computational techniques that rely on multiple geometric standards and algorithmic strategies are utilized to locate the binding pocket. The identification of the binding pockets of target proteins can be aided by a number of software programs, such as PyMOL.

(iii) Compound librarypreparation.

The SBDD procedure heavily relies on compound databases. A collection of chemical compounds that may be used as drug candidates are gathered and organized during the creation of these databases. An effective screening procedure and the selection of lead compounds for further research are made possible by a well-structured database.Ligand libraries are made by choosing compounds with suitable chemical

and physical characteristics. Ligand libraries are a subset of the compound database. These characteristics, sometimes referred to as physiochemical or drug-like features, are picked depending on the disease or biological target that is of interest. In order to improve the chances of discovering compounds that can successfully interact with the target and treat the disease, these libraries are being enriched.To find possible hits for drug discovery, the virtual screening technique makes use of multiple databases that include chemical compounds. These databases are available in a variety of formats, with commercial solutions like ZINC and PubChem (http://pubchem.ncbi.nlm.nih.gov) being popular choices. ZINC is a database of chemicals that have undergone processing to make them biologically relevant. It contains a sizable collection of 230 million chemicals that may be bought in 3D format.There is also PubChem, which gives users access to 112 million chemicals. For efforts to develop new drugs, these enormous libraries of chemicals provide an invaluable resource.

(iv) ADMET Filteration

ADMET offers pharmacokinetics (ADME, or Absorption, Distribution, Metabolism, and Excretion) and pharmacodynamics, or toxicity (T), of lead molecules before proceeding further experiments in a wet lab. It is one of the crucial elements in any drug development procedure. It lessens the chance of expensive, lengthy, and unsuccessful drug experiments. It has been widely accepted that computer-based ADMET prediction must be taken into account in the drug discovery program before moving on to in vitro and in vivo studies. Poor pharmacokinetics and pharmacodynamics are the primary causes of expensive and late-stage drug development failures. The ADMET parameters are based on the kinetics of drug exposure to tissues and the body's response to them, which affect the compound's functionality and pharmacological activity. As a result, this technique offers a critical understanding of how a medicinal molecule behaves inside an organism. This method assists in choosing compounds in the very early stages of drug development, which is essential for the discovery and development of new drugs. Because fewermedications are lost during pre-clinical and clinical phase trials at a later period, this strategy saves money and time.Lipinski's rule of five is given below.

The Lipinsinki rule of 5 states that the lead compounds i.e drugs should have the following properties:

Table 1 : Drugability parameters of a lead compound

Parameters	Minimum	Maximum
Molecular weight (MW) in Daltons	200	500
clogP	-3	6
ClogS	-4	-2
Hydrogen Bond Acceptors	0	10
Hydrogen Bond Donors	0	5
Topological Polar Surface Area (TPSA)	0	120
Rotatable Bonds	0	10

In addition to above parameters, to increase the druggability properties the polar surface area and rotatable bonds (Veber *et al.,* 2002 and Chagas *et al.,* 2018) has been added.

(v) Molecular docking and scoring

The process of inserting the ligand into the target's binding site cavity in three dimensions is known as molecular docking. It is regarded as a method for simulating molecular interactions using computing. In essence, the ligand is placed on the target's active site, and then, using a scoring function, the binding energy and interactions are visualized.In most cases, it accurately predicts both the complex structural conformation and the binding affinity of various chemicals that are directed at an active site of the target macromolecular structure. As a result, it has emerged as the most widely used and trustworthy technique in SBDD.

(vi) Post processing (Improving selection after docking)

Post processing of virtual screening data is needed to minimize the experimental errors so as to make data more reliable.there might be some errors like defective

bonding networks,intra-ligand steric clashes,perverse and twisted amides in high scoring compounds that makes molecules less attractive as lead compounds. . Additionally, molecular docking may not sample the entire conformational space of a ligand, which can result in missed binding opportunities.

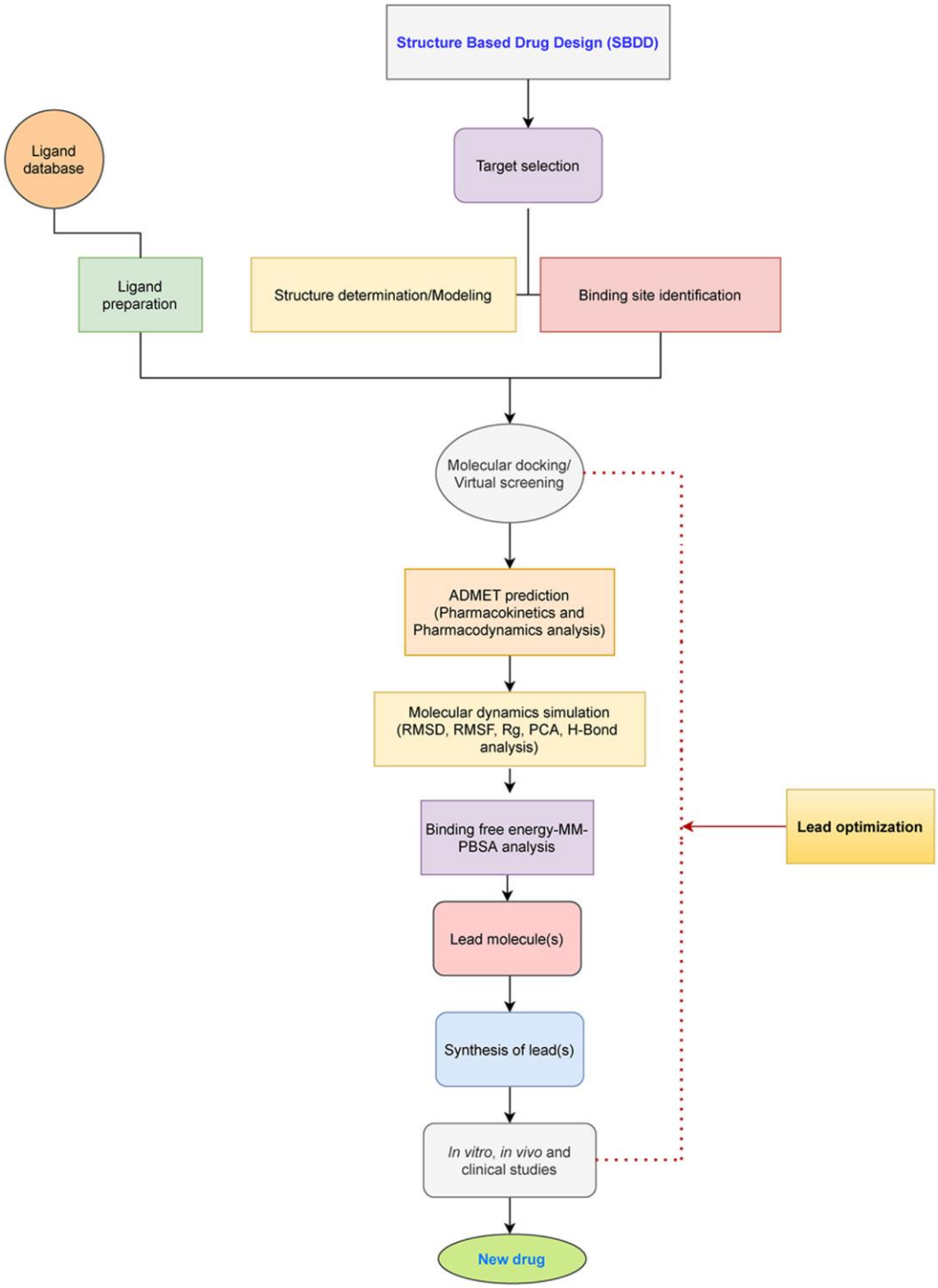


Figure 5 : Computational pipeline used in structure-based drug design for the identification of lead molecule(Dev buksh singh;2021)

2.2.3Ligand Based Drug Design (LBDD)

An strategy called ligand-based drug design, which relies on knowledge of compounds that bind to the desired biological target, is utilized in the lack of receptor 3D information. The most significant and often employed methods in ligand-based drug design are pharmacophore modeling and 3D quantitative structure activity relationships (3D QSAR). They can offer suitable predictive models for lead optimization and lead discovery.

To anticipate novel chemical entities that exhibit similar behavior in Ligand-based approaches, the current knowledge of active compounds against the target is utilised (Martin *et al*., 2002). A pharmacophore model can be created from a library of chemicals using a single known active molecule to specify the minimal structural requirements that a molecule must meet in order to bind to the desired target. The active molecule is compared to the library using a fingerprint-based similarity search because the molecules are provided here as bit strings that indicate whether certain preset structural descriptors are present or absent (Mishra and Siva-Prasad;2011).Structure-based techniques, in contrast, target structural information to assess the likelihood that a novel molecule would bind and interact with a receptor. This approach does not require prior knowledge of active ligands, which is a very important benefit (Kolb *et al*., 2009). New ligands that can trigger a therapeutic response from 3D structures can be created. Structure-based techniques have a significant impact on the discovery and optimization of the initial lead chemical, which is how novel medications are developed.

An strategy called ligand-based drug design, which relies on knowledge of compounds that bind to the desired biological target, is utilized in the lack of receptor 3D information.

a) The knowledge about compounds that bind to the desired target location is used in ligand-based drug design.

b) A pharmacophore model can be created using these compounds.

c) A molecule with the required structural capabilities to bind to a desired target location is referred to as a pharmacophore model.

d) After the pharmacophore is found, it is evaluated to see if the receptor will accept it; if not, the pharmacophore is further tweaked to create a possible medication.

2.2.3.1. Ligands based virtual screening (LBVS)

The "similarity principle," which states that similar compounds tend to have comparable biological properties, is the foundation of ligand-based virtual screening. The typical goal of employing LBVS is scaffold hopping, or the identification of iso-functional molecular structures with noticeably varied molecular backbones. Leapfrogging, scaffold searching, and "scaffold hopping" are other names for "scaffold hopping" (Kalliokoki; 2010). These techniques are typically beneficial in medication repurposing, which involves using existing therapeutic molecules to target new diseases and targets.

2.2.3.2. Molecular descriptors

The reference molecule or set of molecules are compared with a large library of compounds using one of the simplest methods possible. This is done by using descriptors of physicochemical properties, such as molecular weight, volume, geometry, surface areas, atom types, dipole moment, polarizability, molar refractivity, octanol-water partition coefficient (log P), planar structures, electronegativity, or solvation properties that are derived from Symbols are used to symbolize molecules in order to complete the assignment successfully (Prada-Graciaa *et al*., 2016).

2.2.3.3. Quantitative structure–activity relationship (QSAR)

The quantitative structure-activity relationship (QSAR) method is a tool used in drug discovery to forecast the biological activity of chemical compounds. To determine how thecompound's structural relationships relate to one another, statistical and mathematical methods are used.and the associated biological actions.. As a result, the QSAR model is created utilizing structural parameters to forecast a drug's biological characteristics. The 2-D QSAR (2D-QSAR) interprets the two-dimensional structural

features of descriptors such steric, electrostatic, hydrophobicity, and geometric behavior to interpret themolecular biological activity using multiple regression analysis.

2.2.3.4. Pharmacophore modeling

Using different conformations of a variety of ligands rather than just one ligand structure allows for the extraction of more meaningful information. With a large enough selection of ligands, a pharmacophore model of the receptor site can be created. There have been promising results from pharmacophore modeling of smaller, non-peptide compounds that may be more stable and bioavailable than their peptide counterparts (Nielsen *et al.,* 1999).

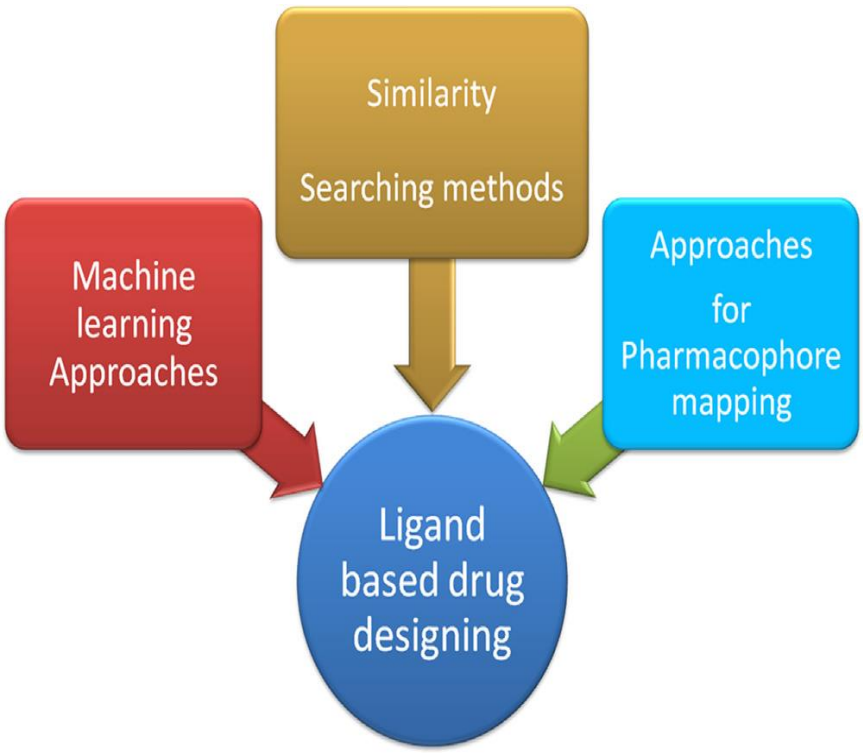


Figure 6 : Ligand based drug design process (Dev buksh singh; 2021)

3. MATERIALS AND METHODOLOGY

3.1 Obtaining the 3D crystal structures of the target proteins

The 3D crystal structures of target protein was obtained from RCSB Protein data bank (PDB) under the name 5T2P (Z Zhou*et al*., 2017).

3.2 Ligand library preparation

The ligand library for the docking purpose was prepared from different databases available in websites such as zinc 15 databases, asinex, selleckchem,drug bank,pubchem etc.The ligand library from Zinc databases includes the categories like fda drugs,natural products,and from asinex antiviral and nucleotide mimeticsetc. All the selected ligands were then processed for ADME/Tox screening for docking.

3.3 ADME/Tox Screening

Drug discovery and development process are time-consuming processes. Manypromising lead compounds fail to reach the clinical stage of drug research due to poor pharmacokinetic properties and toxicity issues. Hence, in order to improve absorption, distribution, metabolism, and excretion, the ligands was screened for drugability based on Lipinski's rule of five and LogS. Using the OSIRIS data warrior program ,the druglikeness properties of ligands such as-molecular weight, cLogP and cLogS, druglikeness, and toxicities like mutagenicity, tumorigenicity, reproductive effects, and irritating effectswere calculated.The following parameters are set for ADMET Process

ADME/Tox

Total Molecular weight: 200-500 Daltons

cLogP: -3 to +6

cLogS: -4 to -2

Hydrogen bond acceptor: 0 to 10

Hydrogen bond donors: 0 to 5

Topological Polar surface area: 0 to 120

Drug likeliness: positive value (as default)

LE/TOx/S criteria

Mutagenesis: None

Reproductiveness: None

Tumorogenicity: None

Irritant: None

Rotable bonds: 0 to 10

3.4 Identification of active site of target protein

It is crucial to locate the protein's ligand binding site, also known as the active site, before performing molecular docking between a target protein and a ligand. PyMol, a program for visualizing and analyzing protein structures, can be used to accomplish this. The procedure began by loading the native ligand and protein structure into the PyMol window. After that, the active site was located by determining the specific interactions between the ligand and protein by looking at the amino acid residues that were five amino acids away from each other. Polar interactions were also assessed in addition. The data was then displayed with Ligplot+ and Biovia Discovery Studio and saved as a PDB file.

3.5 Molecular docking and virtual screening

Using the AutoDock Vina Wizard in the PyRx 0.9.8 platform, molecular docking experiments were carried out to examine the binding energy of5t2P with a variety of natural product ligands. The center grid box values used for the docking trials were x= 28.347 , y= 11.962 ,z= -43.702,and dimensions value x= 28.021 y= 32.705 and z= 21.667, as well as some parameters like the number of modes used, which was 16, and the number of exhaustiveness, which was 8. The ligands were then sorted by

binding energy, and those with binding energies greater than those of the natural ligand were chosen for additional examination.

3.6 Screening with hMAT1A Proteins and Preference index

The ability of the chosen ligands to interact with the human hMAT1A protein, which is involved in the synthesis of SAM, was examined after a screening procedure. It was determined through additional testing that the selected medication candidate would not prevent these proteins, both of which are located in the liver and have significant effects on human health, from producing SAM. In order to determine whether the ligands could potentially impair the function of the hMAT1A protein, they were docked with it.

A set of largely selected ligands and the hMAT1A protein were used for molecular docking. The hMAT1A protein's center grid box and dimension values were, respectively, x: 31.096, y: -0.571, z: 24.983 and x: 27.625, y: 57.151, z: 32.622. The lead molecule was ultimately chosen because of its high preference index, low binding energy against the human hMAT1A protein, which is involved in the synthesis of SAM, and high binding energy against the target protein. Preference index, which was generated using the hydrogen bond acceptor and donor capacities and the number of rotatable bonds of the pharmacophore, was added to further refine the selection of the lead molecule among the screened ligands.

The formula for calculating the preference index was as follows:

{(H-acceptor + H-donor + Rotatable bonds count) * 5}/ 25.

3.7 Visualization of interaction of Protein-ligand

Ligplot and PyMOL were used to visualize the protein-ligand interactions. By analyzing and visualizing the binding of proteins and ligands, these techniques were able to shed light on the nature of the interactions and the spatial organization of the molecules. The binding energy in Kcal/mol between each docked ligand and the target protein was also computed. This number measures how strongly the ligand interacts with the protein and can be used to assess the ligand's potential as a therapeutic candidate.

(a) Using Ligplot

For the purpose of viewing 2D interactions, the protein and ligand pdb file (created for visualizing active sites) was opened in ligplot+. This method was used to identify the hydrophobic and hydrogen bonds that were causing the greater binding energy. The specific interactions between the protein and the ligand that contribute to their binding were better understood as a result of these interactions.

(b) Using Biovia Discovery Studio

The BioVia Discovery Studio, a computational tool that enables the detection of various types of bonds between atoms of a protein and a ligand, was used to open the protein and ligand pdb file. Other interactions, such as alkyl bonds, pi-alkyl bonds, pi-pi t-shaped links, pi-sigma bonds, carbon-hydrogen bonds, etc., as well as the residues of the protein and ligand that are involved in those interactions, were also identified using this method. The bond distances, which reveal details on how powerful the interactions are, were also computed. The specificity and affinity of the ligand's binding to the protein were further understood thanks to this knowledge.

3.8 Mutagenesis study and lead compound efficiency to mutant protein

The wild type hepatitis B virus core protein is subjected to different mutation form by using PyMol software.The purpose of this study is to check the efficiency of the lead compound to in+hibit the different mutant form.The single mutant Y132 (Z Zhou *et al*;2017) is prepared by converting the tyrosine residue of the aminoacid chain at 132 number to alanine.similarly the double mutant form P5T,I97L is prepared by converting proline at residue 5 to threonine whereas isolecuine to lysine at residue 97(Christina R Bourne*et al*.,2009).The different lead compound and mutant protein interactions were visualized in biovia discovery studio after the molecular docking of the lead compound and mutant protein.

4. RESULTS AND DISCUSSION

4.1 Target protein selection

Protein target identification is one of the most important and primary step in the drug discovery process.Hepatitis B core protein was our target protein in this docking study.The target protein 3D crystal structure was searched in the RCSB websites.The PDB ID for the protein was 5T2P.The 3D crystal structure of 5T2P wild was given below.

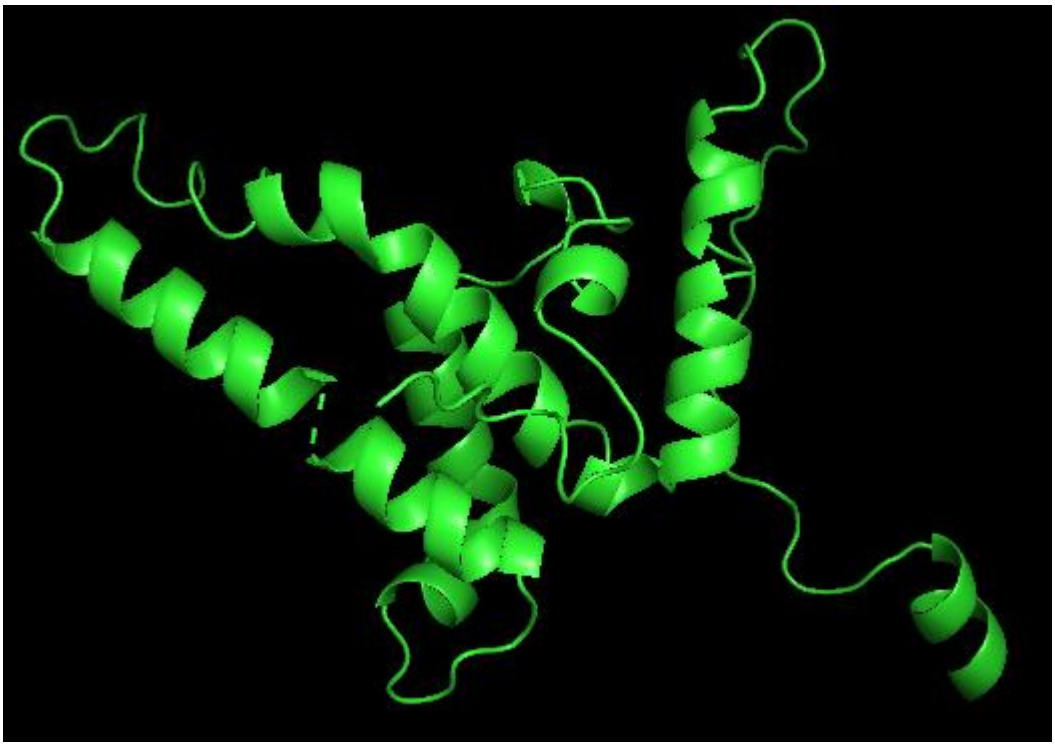


Figure 7 : 3D structure of 5T2P wild protein observed in PyMol

This protein has six chains, each chain with 155 aminoacids sequence in length.we have considered bchain in molecular docking and the protein in PDB has mutation in the 132 aminoacid chain changing the aminoacid from tyrosine to alanine.we have revert this mutant to original wild sequence and docked this protein structure.

4.1.1 Ligand LibraryPreparation

Ligands were prepared from various ligand library such as ZINC15which consists of different categories such as- FDA, in Trial,natural products ,investigation,world not

fdaetc and ASINEX for nucleoside mimeticsand antivirals as well as selleckchem for bioactive and antiviral compounds.

Ligands sets from the above databases were downloaded in sdf format and was subjected to OSIRIS datawarrior for ADMET filtration where the parameters of the druggability were set from the ADMET filteration .About 35000 Natural products were selected for molecular docking obtaining from Zinc 15 databases. Ligands were prepared with a view to find lead compounds from natural products category.

4.1.2Active sites identification and molecular docking

The active sites are the binding sites in protein structure where the ligands binds.The active sites in the protein given below are within5Å of the ligand ssociated with protein.

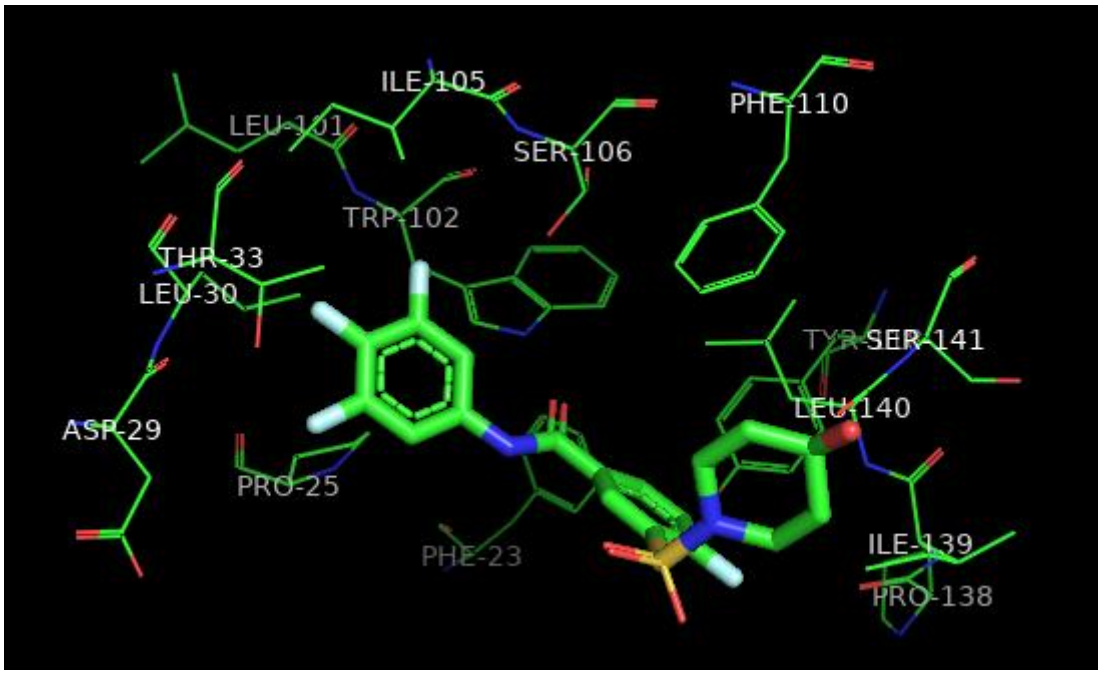


Figure 8: Binding sites identification within 5Å of ligand with HBV core protein

The active sites of Hepatitis B core protein which were used for docking are mentioned in the table 2 below

Table 2 : Active binding sites of HBV core protein obtained by PyMol

Residues	**Amino Acid**
23	PHE
25	PRO
29	ASP
30	LEU
33	THR
102	TRP
105	ILE
106	SER
118	TYR
138	PRO
140	LEU
141	SER

The molecular docking of taget protein with the ligand are done with the active sites mentioned above by using the PyRx software.The center grid box values were set as x= 28.347, y= 11.962, z=- 43.702 and dimension values x=28.021, y= 32.705 and z= 21.667 as well as certain parameters such as numbers of exhaustiveness of 8 and number of modes of 16 were set up.Final docking were done at 32-32 parameters.

4.1.3hMAT1A protein screening and Preference Index set up

The production of SAM in the liver and extra-hepatic organs is carried out by these enzymes. In order to identify new therapeutic candidates that could avoid inhibiting

human hMAT1A and hMAT2A while targeting the target protein precisely. Secondly, the best candidates for additional examination were chosen from among the ligands that had a high affinity for the target protein, HBV core After molecular docking the docking score was obrtained.The compounds that have higher binding affinity with target protein was sorted and subjected to HMAT1A screening to ensure that the lead compound donot damage the liver.The enzyme methionine adenosyltransferase catalyzes the catabolism of methionine, which results in the production of S-adenosylmethionine (SAM), a crucial chemical for human health (MAT). The MAT protein is encoded by two distinct genes, MAT1A and MAT2A, which result in the production of the MAT enzymes, hMAT1A and hMAT2A, respectivelyprotein.These ligands were then docked with, and those that exhibited a reduced affinity for binding to human enzymes were chosen. Due to their greater affinity for the HBV core protein and lesser affinity for the human SAM biosynthesizing enzymes, ligands with higher binding affinity for HBV core proteins were therefore explored as prospective therapeutic targets.This is important since the newly identified medication targets need to be safe for human use and shouldn't harm human health. The potential as drug candidates of the ligands that passed this screening was then further investigated.

The active sites of hMAT1A (PDB: 6SW5) binding with SAM are mentioned in the table 3

Table 3 : Active sites of hMAT1A (PDB: 6SW5) binding with SAM obtained from PyMol

Residues	**Amino acid**
55	ALA
70	GLU
112	GLN
113	GLN

114	SER
133	GLY
134	ASP
289	LYS
291	ASP

The center grid box values in x= 31.096 y= -0.571 z= 24.983 and dimension values x= 27.625 y= 57.151 and z= 32.622 for hMAT1A (PDB: 6SW5) was set up.

Anempirical formula was set up to make screening process easier. The formula was based on the number of H-bond acceptors, H-bond donors, and rotatable bond counts in the ligands.

{(H-acceptor + H-donor + Rotatable bonds count) * 5}/ 25.

This formula is used because a medicine must be able to pass across biological membranes in order to reach its target protein in order to be effective. The partitioning of biologically active substances and their interactions with target proteins depend heavily on hydrogen bonding. Atoms or groups of atoms that can accept a hydrogen bond are known as H-bond acceptors, whereas those that can donate a hydrogen bond are known as H-bond donors. Drug molecules must be able to break hydrogen bonds in the aqueous environment of the body in order to pass biological membranes and get to their target proteins. . The number of H-bond acceptors and donors in a ligand can affect its ability to partition across membranes and interact with its target protein.

Chemical interactions known as rotatable bonds enable a molecule to rotate along a certain axis, which may impact the flexibility and capacity of the molecule to bind to a target protein. An increased binding affinity may result from a molecule's greater flexibility and ability to bind to a target protein in many directions.

4.1.4 Protein- Ligand Interaction Result

The protein ligand interactions are viewed in different software after having a molecular docking with target protein in PyRx software.

4.1.5 Analyzing virtual screening of natural products

A virtual screening of natural products against the HBV core proteinof PDB ID 5T2P was performed using the PyRx software. Among the natural products database available in the ZINC15 software after ADMET filteration, 35000 igands wereselected.The molecular docking of these compounds with target proteins results that 500 compounds show higher binding affinity with target protein.The reference ligand of 5T2P protein shows the binding energy of -8.1Kcal/mol.500 compounds shows the higher binding affinity than these reference ligands.The ligands that shows higher binding affinity with target protein are docked with hmat1A protein.The reference ligand of hmat1A protein SAM shows the binding energy of -7.9 Kcal/mol.Only one compound shows the less binding affinity with the hMAT1A protein. the compound named **(3R, 3aR, 4aR, 8aR, 9aR)-3-[[[(1S,2S)-1-hydroxy-1-phenylpropan-2-yl]-methylamino]methyl]-8a-methyl-5-methylidene-3a, 4, 4a, 6, 7, 8, 9, 9a-octahydro-3H-benzo[f][1]benzofuran-2-one** shows binding affinity of -8.5Kcal/mol with targetprotein and -7.8 Kcal/mol with hmat1A protein.

The result analysis are done with the binding affinity of the ligands to the HBV core protein, as well as comparing their binding affinity with hMAT1A, a human protein. The ligands that showed higher binding affinity with HBV core protein and lower affinity with hMAT1A were chosen for further analysis.

The binding energy and preference index of the high scoring selected 5 compounds were analyzed, and it was found that the two compounds,ZINC4260269 and ZINC 85878547 showed the the highest binding energy with HBV core protein at -9.5 kcal/mol by bothand the preference index of 1.8 by both of these compound. The reference ligands associated with protein showed the docking score of -8.1Kcal/mol. The top five compounds with docking score better than reference ligand are given below. ZINC4260269 and ZINC 85878547 had a higher binding energy of -9.5 kcal/mol,

but it was not selected as it had a greater binding affinity with hMAT protein as it can create problem to liver.

The compounds ZINC 12660727 showed the higher binding affinity with HBV core protein of -8.5Kcal/mol as compared to its reference ligand -8.1Kcal/mol and lower affinity with hMAT1Aof -7.8 Kcal/mol as compared to its reference ligand SAM of -7.9Kcal/mol.so the compound of ZINC id 12660727 i.e. benzofuranone derivatives is considered as potential lead compound for the treatment of hepatitis B.

Table 4 : Summary of top hits Natural Product compounds after molecular docking against HBV core protein.

Natural Products	Molecular Characters						Binding energy (kcal/mol)		
Database ID	cLogP	cLogS	H-Acceptors	H-Donors	Polar Surface Area	Drug likeness	HBV core protein	hMAT1A	Preference Index
ZINC12660727	**3.1688**	**-3.871**	**4**	**1**	**49.77**	**1.0664**	**-8.5**	**-7.8**	**1.6**
ZINC4260269	3.0902	-3.734	6	1	69.72	6.2688	-9.5	-8.3	1.8
ZINC85878547	1.1323	-3.319	8	1	89.61	5.0898	-9.5	-9.8	1.8
ZINC247722440	0.7038	-3.062	9	1	94.55	6.8164	-9.4	-9.9	1.6
ZINC604405342	2.4134	-3.684	8	0	81.24	0.057699	-9.3	-9.7	2

4.1.6 Analysis of protein ligand interactions

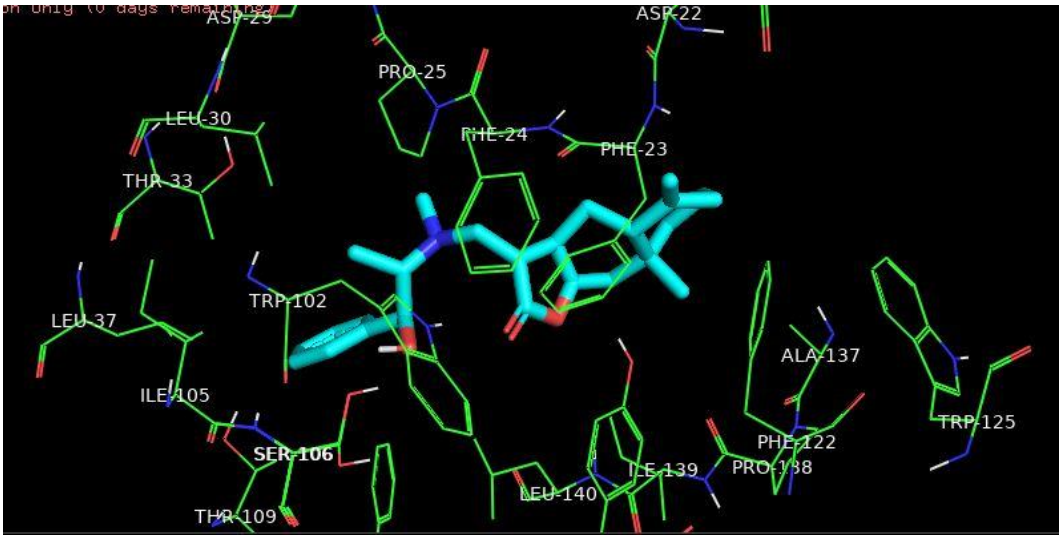
The lead compounds that have been identified as potential drug candidates based on their binding affinity to a target protein were screened for their protein-ligand

interactions using PyMol, Discovery studio. The interactions between the lead compounds and the protein were then analyzed in detail. Different amino acid residues in the protein can be involved in these interactions and they can interact via different types of bonding, such as hydrogen bonding, hydrophobic interactions, and electrostatic interaction. Understanding how the lead drug attaches to the protein can help us better understand the molecule's mechanism of action and increase the binding affinity. This can be done by identifying the precise amino acid residues involved and the kinds of interactions they make.

a) Using Pymol

PyMol is molecular visualization software that can be used to create detailed 3D models of protein-ligand complexes. Here we can visualize the interactions between the lead compound and the protein, including hydrogen bonding interactions, hydrophobic interactions and electrostatic interactions which is useful for identifying the specific residues and groups involved in the interaction and for understanding the structural basis of the binding.

The lead compound benzofuranones derivatives and HBV core protein interaction were visualized in PyMol software.The aminoacids within 5Å of benzofuranones derivatives are shown in the figure below.



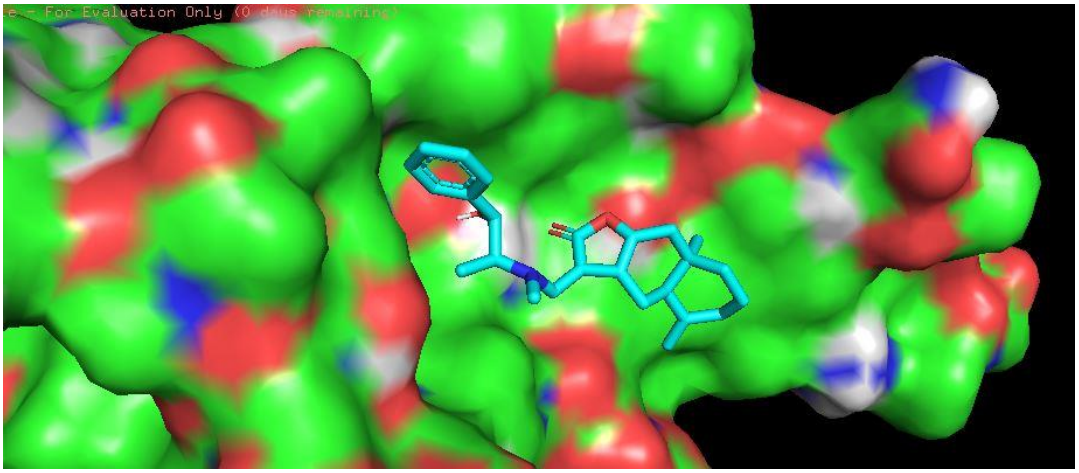


Figure 9 : Visualization of Protein ligand interaction A) interaction of ligand (benzofuranones derivatives) with the HBV core protein.figure showing the active sites of protein within 5Å of ligand. B) Surface visualization of protein with ligand in PyMol

b)Using Discovery studio

Discovery Studio is a software package that can be used to generate 2D and 3D interaction maps of protein-ligand interactions. These maps provide detailed information on the specific protein residues and groups involved in the interaction, as well as the types of bonds that are formed and the distance between the interacting groups. It also allows to identify the residues that interact more often in a protein-ligand interaction. The orientation and stereochemistry of the individual amino acids and groups can also be visualized, which can provide insight into the specific mechanisms of the interaction. This can be useful for understanding the structural basis of protein-ligand interactions and for drug design. The different types of bonds are formed during protein-ligand interactions. The bonds can be both covalent and non-covalent bonds. Non-covalent interactions include pi-alkyl, pi-pi T shaped, and pi-sulphur interactions. In a pi-alkyl interaction, the pi electron cloud of an aromatic group interacts with the electron group of an alkyl group. In a pi-pi T shaped interaction, the pi electron cloud of two aromatic groups interact in a T-shaped manner. In a pi-sulphur interaction, the pi electron cloud of an aromatic ring interacts with the lone pair of electrons of a sulphur atom. These interactions play an important role in the binding of the ligand to the protein and affect the overall binding energy.

The 2D map of ligand -protein interactions shows that the TRP-102, PHE-110,TRP-118and PHE-122 involves in bonded interactions. The TRP-102 involve in the conventional H- bonding with the oxo group of bond length 2.47Å ,the PHE-110 involves the Pi-Pi T shaped bonded interactions with the benzene ring of bond length of 4.82 Å. TRP-118 involves in the conventional H bonding with the oxygen of the furan of bond length of 2.78 Å and PHE-122 involves in the Pi alkyl bonding with cyclohexane of bond length of 5.40 Å.

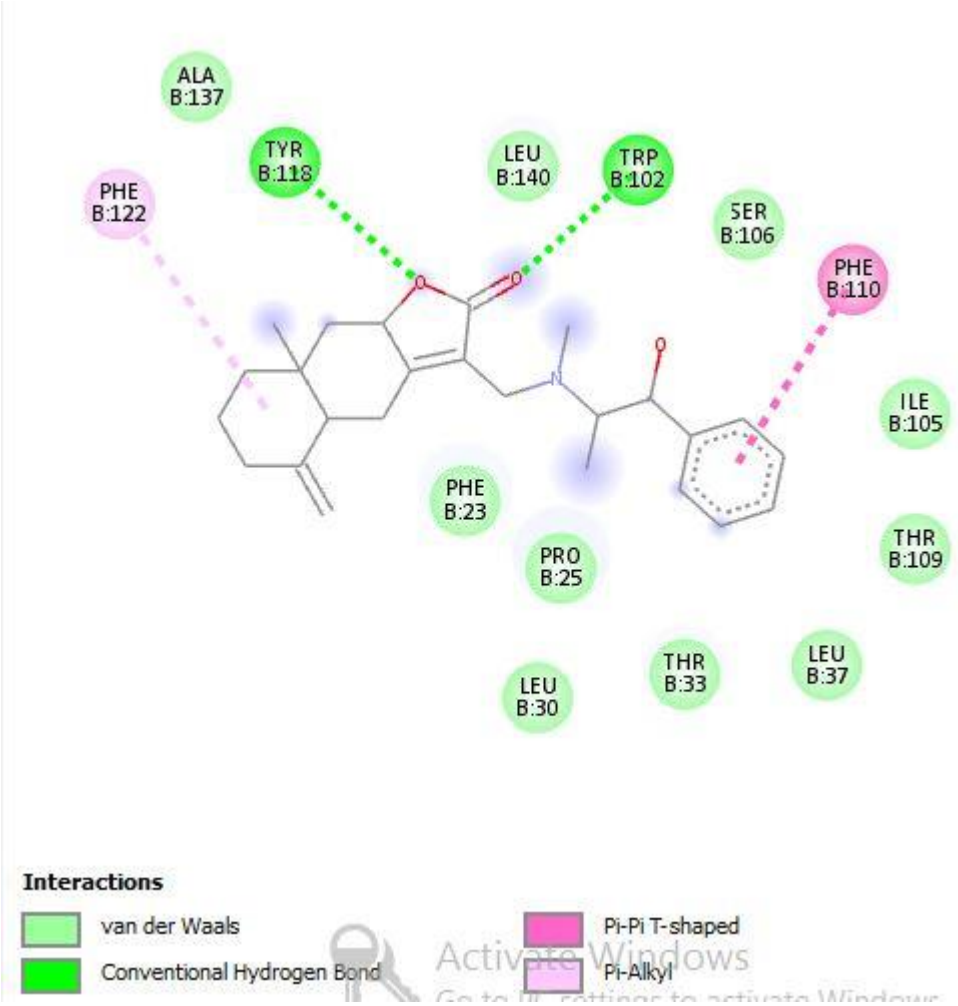


Figure 10 : 2D Visualization of Protein ligand interaction between HBV core protein and ligand in Discovery Studio

Table 5 : Summary of Protein-ligand interaction between Benzofuranones and HBV core protein

Amino acid	Bond type	Bonded with	Bond length
TRP-102	Conventional H-bond	Oxo	2.47 Å
PHE-110	Pi-Pi T shaped	Benzene	4.82 Å
TRP-118	Conventional H-bond	oxygen of furan	2.78 Å
PHE-122	Pi-Alkyl	Cyclohexane	5.40 Å

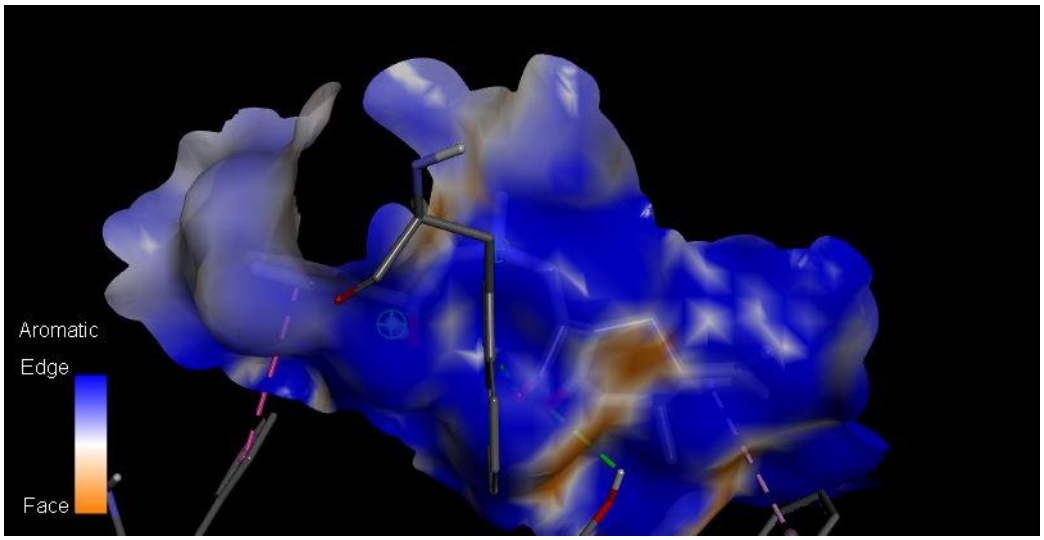


Figure 11 : Showing aromaticity

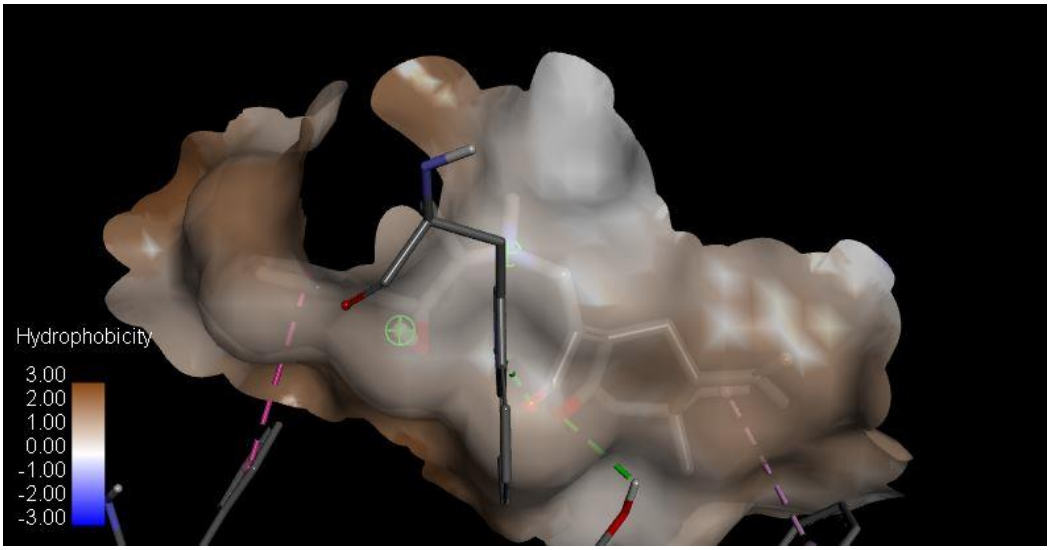


Figure 12 : Showing hydrophobicity

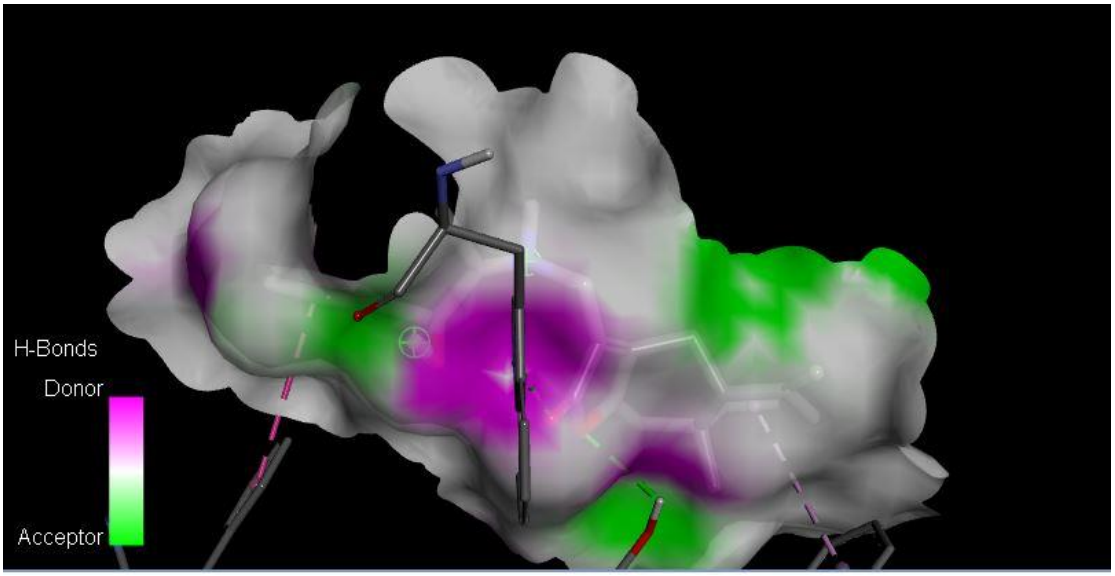


Figure 13 : Showing H bonds

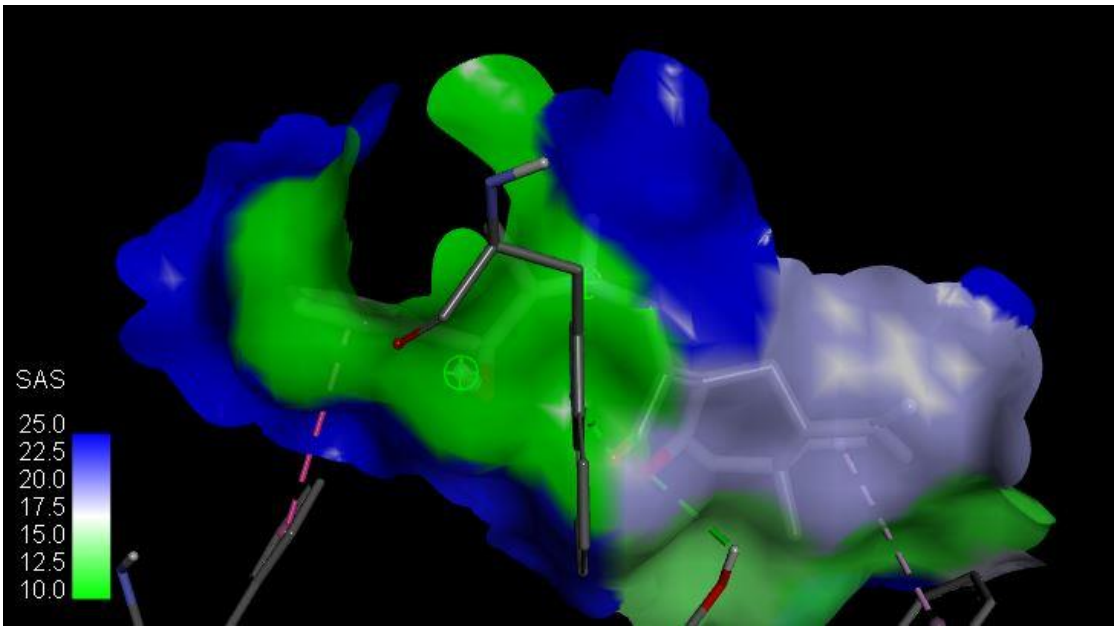


Figure 14 : Showing solvent accessible surface(SAS)

4.2Mutational Analysis and Efficiency of Screened Drug Candidate

The lead compound Benzofuranones derivatives obtained after molecular docking with hepatitis B virus core protein is used to study the mutgenesis of the target protein and the efficiency of lead compound antiviral properties to the mutant protein.

The Hepatitis B virus core protein different mutant form i.e. Y132A single mutant(Z Zhou *et al* ;2017) and P5T,I97L double mutant(chua *et al* ;2003) are used to dock with lead compound Benzofuranones.The docking score obtained after molecular docking of lead compound Benzofuranones with the mutant form Y132A is -8.5Kcal/mol whereas with the double mutant form P5Tand I97L is -8.4Kcal/mol.The docking score of the reference ligand Sulfamoyl benzamide is -8.1Kcal/mol.and with the wild type of protein is -8.5Kcal/mol.

From the above result, the benzofuranones have similar antihepatitis B properties with both wild and single mutant protein where as a less more efficiency with double mutant form.

4.2.1 Ligand-mutant protein interaction analysis

The ligand- mutant protein analysis was studied using discovery studio.

4.2.2Single mutant Y132A

The 2D interaction map as visualized in discovery studio shows that the PHE110 involves a Pi -Pi T shaped interaction with benzene group of benzofuranones compound with bond length of 4.82 Å.similarly SER106 makes conventional H bonding with oxygen atom of the chain of bond length of 2.49 Å.TRP102 makes a conventional H bonding with oxygen atom of chain and oxo group of furan of bond length 2.15 Å and 2.48 Å respectively.TYR-118 makes conventional H bonding with oxygen atom of furan of bond length 2.77 Å .PHE-122 makes Pi alkyl bond with cyclohexane of the chain of bond length 5.38Å.

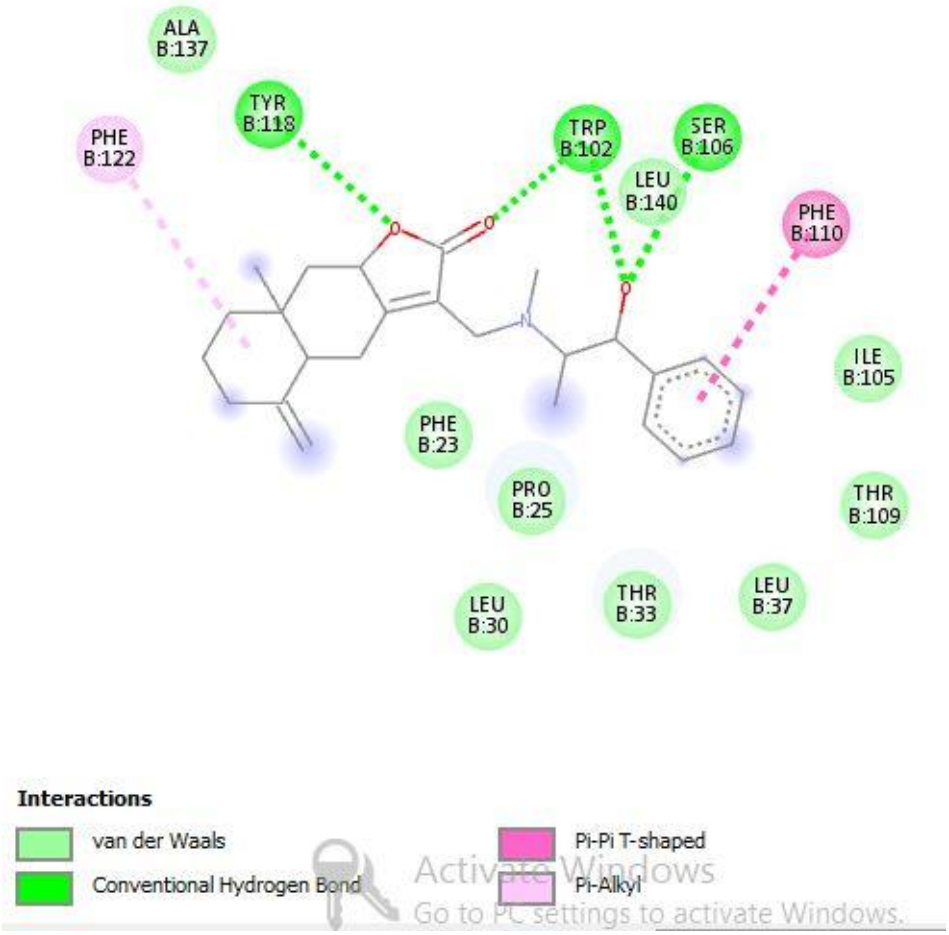


Figure 15: 2D interaction map of mutant protein Y132A and benzofuranones viewed from discovery studio.

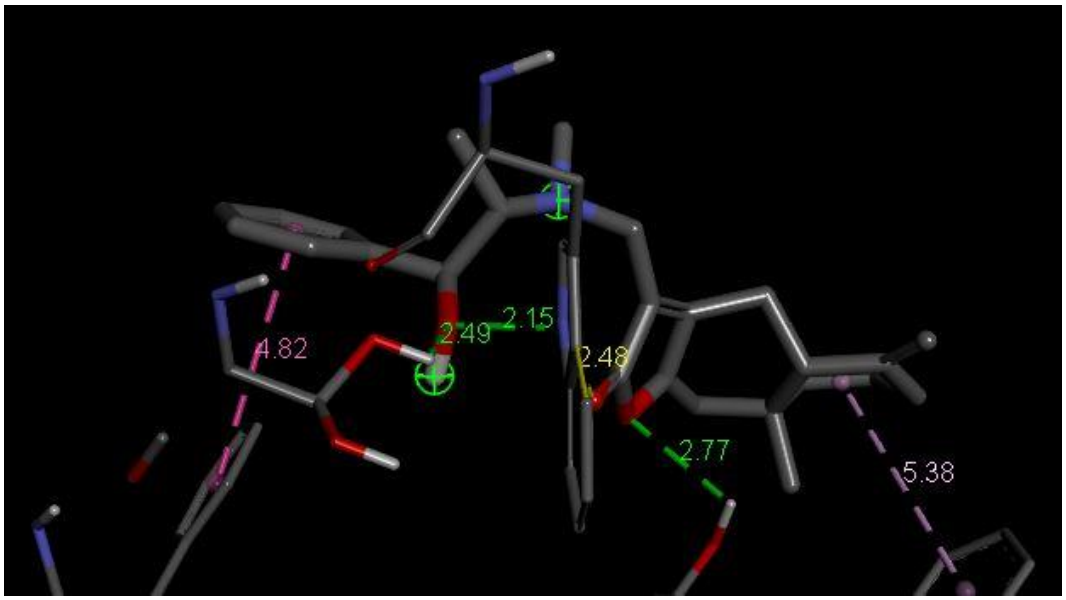


Figure 16: Bond distances of mutant protein Y132A and benzofuranones viewed from discovery studio.

4.2.3 Double mutant PT5, I97L

The 2D interaction map as visualized in discovery studio shows that the PHE110 involves a Pi -Pi T shaped interaction with benzene group of benzofuranones compound with a bond length of 4.84 Å.similarly SER106 makes conventional H bonding with oxygen atom of the chain of bond length of 2.49 Å.TRP102 makes a conventional H bonding with oxygen atom of chain and oxo group of furan of bond length 2.16 Å and 2.45Årespectively.TYR-118 makes conventional H bonding with oxygen atom of furan of bond length 2.79 Å.PHE-122 makes Pi alkyl bond with cyclohexane of the chain of bond length 5.41 Å.

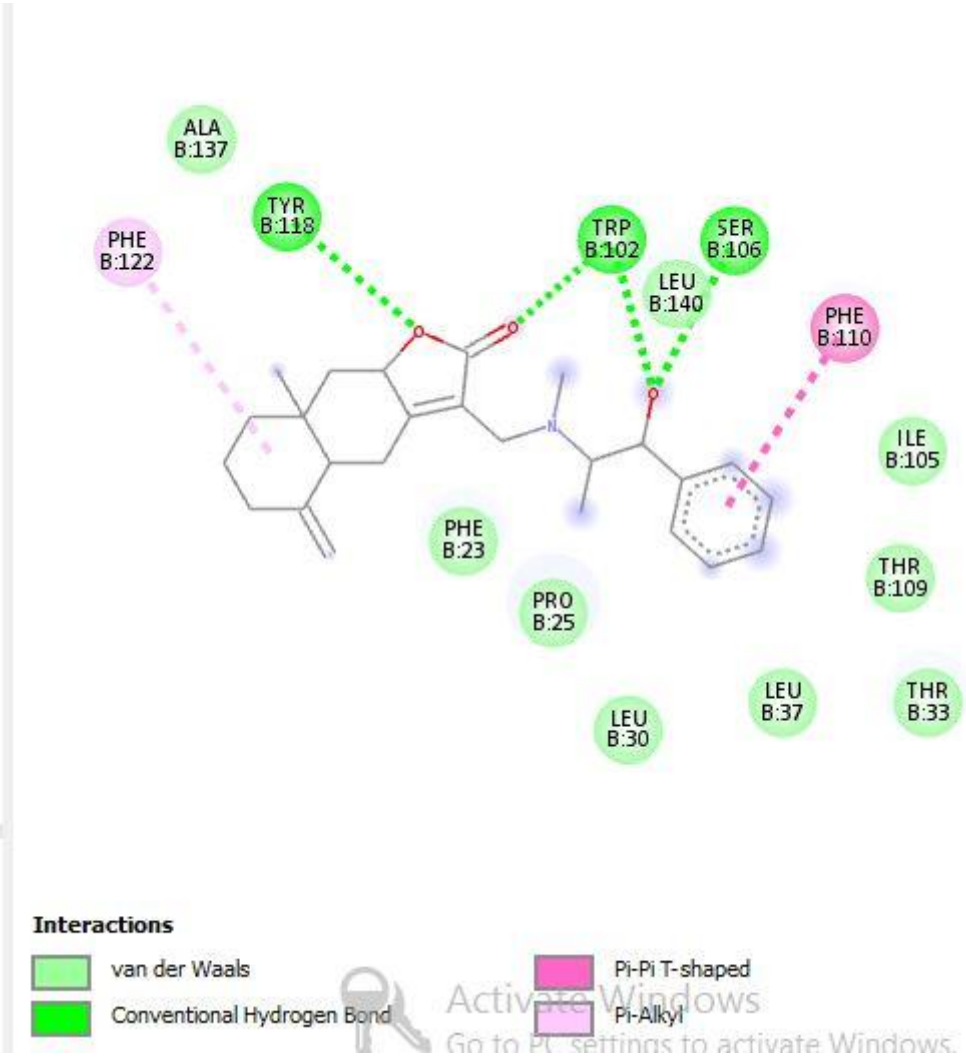


Figure 17: 2D interaction map of mutant protein P5T,I97L with benzofuranones viewed from discovery studio.

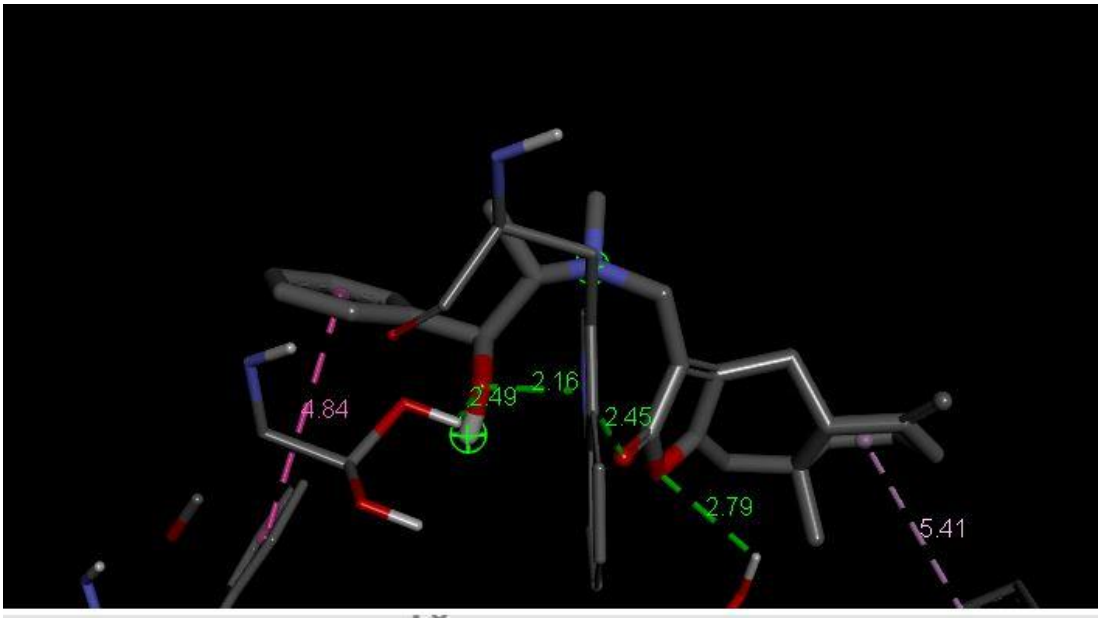


Figure 18: Bond distance of mutant protein P5T, I97L with the benzofuranones visualized in discovery studio.

4.3 Prospects of using benzofuranone derivatives as antihepatitis B drugs.

The compound named **(3R, 3aR, 4aR, 8aR, 9aR)-3-[[[(1S,2S)-1-hydroxy-1-phenylpropan-2-yl]-methylamino]methyl]-8a-methyl-5-methylidene-3a, 4, 4a, 6, 7, 8, 9, 9a-octahydro-3H-benzo[f][1]benzofuran-2-one.i.e benzofuranone derivatives** is selected as lead compound from molecular docking between hepatitis B core protein and natural products ligands.These compound shows the docking score of -8.5Kcal/mol which has greater binding affinity with this protein than the reference ligand sulfamoyl benzamide i.e.-8.1Kcal/mol. as well as shows lesser binding affinity with hMAT protein i.e.-7.8Kcal/mol than its reference ligand (SAM) -7.9Kcal/mol.hMAT screening shows that benzofuranone derivative is safer to use as drugs as it shows lesser binding affinity to hMAT protein.This means it creates lesser hepatotoxicity.Greater docking score than reference ligand means benzofuranones derivative binds more to the active sites of hepatitis B core protein to inhibits its viral properties.Various literature review have also suggested benzofuranone compound as antiviral drugs(Mahesh & Murugan;2016).The pubchem ID for this compound is **25426069**.The molecular formula is **$C_{25}H_{35}NO_3$.**

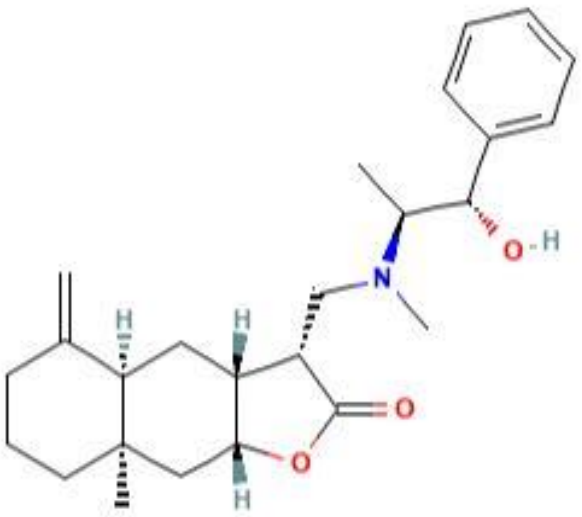


Figure 19: Structure of benzofuranone derivatives (pubchem).

6. SUMMARY

The currently available FDA drugs are not able to cure hepatitis B functionally due to development of mutant viruses as well as the low efficacy of present drugs. So the scientific community in the world are in search of new drugs that can combat hepatitis B virus and fulfill the new drugs development void. Computer aided drug designing can be a good option to tackle this problem .In this study we create ligand library from different popular databases like ZINC,Asinex etc .Ligands from natural products are screened to about 35000 ligands from ADMET filteration. the filterated ligand are subjected to molecular docking where 500 compounds shows the higher binding affinity with the HBV core protein.We have done the hMAT1A filteration technique to sort down these ligands. Ligands of ZINC ID12660727 theBenzofuranones derivatives shows the higher binding affinity towards the HBV core protein and lesser binding affinity to hMAT1A protein as compared to their respective native ligands.Several literature review have also said that Benzofuranones derivatives have shown greater antiviral activity.So the Benzofuranones derivatives is taken as potential lead compound against the hepatitis B virus core proteins.The efficiency of the lead compound to inhibit the mutant protein form was also studied and the result shows the benzofurnanone can inhibit the mutant form with good efficiency as with wild form.

7. CONCLUSION

The Benzofuranones derivative compound is selected as lead compounds against the hepatitis B virus core protein. The several thousands of ligands are subjected to ADMET filteration and docking score obtained after molecular docking between ligand protein as well hMAT1A screenings leads to Benzofuranones derivatives compounds as anti-hepatitis B drugs.Computer aided drug designing could be a promising tool for drug discovery process.

8. RECOMMENDATION

Thus, it is recommended that the interaction between the lead compound Benzofuranones and the target HBV protein can be further verified by Molecular dynamics simulations. Cell lines work can be conducted in laboratory for verifying its antiviral activity by wet lab.Various animal models trial testing ,toxicity testing,enzyme inhibition kinetic testing can be carried out further to verify its as potential lead molecules against hepatitis B virus .

REFERENCES

Acharya, C., Coop, A., Polli, J. E., & Mackerell, A. D., Jr (2011). Recent advances in ligand-based drug design: relevance and utility of the conformationally sampled pharmacophore approach. *Current computer-aided drug design*, *7*(1), 10–22.

Allen, W. J., Balius, T. E., Mukherjee, S., Brozell, S. R., Moustakas, D. T., Lang, P. T., Case, D. A., Kuntz, I. D., & Rizzo, R. C. (2015). DOCK 6: Impact of new features and current docking performance. *Journal of computational chemistry*, *36*(15), 1132–1156.

Ashburn, T. T., & Thor, K. B. (2004). Drug repositioning: identifying and developing new uses for existing drugs. *Nature reviews. Drug discovery*, *3*(8), 673–683.

Baig, M. H., Ahmad, K., Roy, S., Ashraf, J. M., Adil, M., Siddiqui, M. H., Khan, S., Kamal, M. A., Provazník, I., & Choi, I. (2016). Computer Aided Drug Design: Success and Limitations. *Current pharmaceutical design*, *22*(5), 572–581.

Bartuzi, D., Kaczor, A., Targowska-Duda, K., & Matosiuk, D. (2017). Recent advances and applications of molecular docking to G protein-coupled receptors. *Molecules*, *22*(2), 340.

Beaumont, C., Young, G. C., Cavalier, T., & Young, M. A. (2014). Human absorption, distribution, metabolism and excretion properties of drug molecules: A plethora of approaches. *British Journal of Clinical Pharmacology*, *78*(6), 1185–1200.

Brady, G. P., Jr, & Stouten, P. F. (2000). Fast prediction and visualization of protein binding pockets with PASS. *Journal of computer-aided molecular design*, *14*(4), 383–401.

Böttcher, B., & Nassal, M. (2018). Structure of mutant hepatitis B core protein capsids with premature secretion phenotype. *Journal of molecular biology*, *430*(24), 4941-4954.

Bourne, C. R., Katen, S. P., Fulz, M. R., Packianathan, C., & Zlotnick, A. (2009). A mutant hepatitis B virus core protein mimics inhibitors of icosahedral capsid self-assembly. *Biochemistry*, *48*(8), 1736-1742.

Chagas, C. M., Moss, S., & Alisaraie, L. (2018). Drug metabolites and their effects on the development of adverse reactions: Revisiting Lipinski's Rule of Five. *International journal of pharmaceutics*, *549*(1-2), 133–149.

Chen, L., Morrow, J. K., Tran, H. T., Phatak, S. S., Du-Cuny, L., & Zhang, S. (2012). From laptop to benchtop to bedside: structure-based drug design on protein targets. *Current pharmaceutical design*, *18*(9), 1217–1239.

Chen, L., Liu, W. G., Xiong, F., Ma, C., Sun, C., Zhu, Y. R., ... & Wang, Z. H. (2021). 3D-QSAR, molecular docking and molecular dynamics simulations analyses of a series of heteroaryldihydropyrimidine derivatives as hepatitis B virus capsid assembly inhibitors. *New Journal of Chemistry*, *45*(47), 22062-22076.

Chua, P. K., Wen, Y. M., & Shih, C. (2003). Coexistence of two distinct secretion mutations (P5T and I97L) in hepatitis B virus core produces a wild-type pattern of secretion. *Journal of virology*, *77*(13), 7673-7676.

Colovos, C., & Yeates, T. O. (1993). Verification of protein structures: Patterns of nonbonded atomic interactions. *Protein Science*, *2*(9), 1511–1519.

Corbeil, C. R., Williams, C. I., & Labute, P. (2012). Variability in docking success rates due to dataset preparation. *Journal of computer-aided molecular design*, *26*(6), 775–786.

Cosconati, S., Forli, S., Perryman, A. L., Harris, R., Goodsell, D. S., & Olson, A. J. (2010). Virtual Screening with AutoDock: Theory and Practice. *Expert opinion on drug discovery*, *5*(6), 597–607.

da Rocha, A. B., Lopes, R. M., & Schwartsmann, G. (2001). Natural products in anticancer therapy. *Current opinion in pharmacology*, *1*(4), 364–369.

Dickson, M., & Gagnon, J. P. (2004). Key factors in the rising cost of new drug discovery and development. *Nature reviews. Drug discovery*, *3*(5), 417–429.

Dudley, J. T., Deshpande, T., & Butte, A. J. (2011). Exploiting drug-disease relationships for computational drug repositioning. *Briefings in bioinformatics*, *12*(4), 303–311.

Durán-Iturbide, N. A., Díaz-Eufracio, B. I., & Medina-Franco, J. L. (2020). *in silico* ADME/tox profiling of natural products: A focus on BIOFACQUIM. *ACS Omega*, *5*(26), 16076–16084.

Ferreira, L. G., Dos Santos, R. N., Oliva, G., & Andricopulo, A. D. (2015). Molecular docking and structure-based drug design strategies. *Molecules (Basel, Switzerland)*, *20*(7), 13384–13421.

Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., Repasky, M. P., Knoll, E. H., Shelley, M., Perry, J. K., Shaw, D. E., Francis, P., & Shenkin, P. S. (2004). Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of medicinal chemistry*, *47*(7), 1739–1749.

Gu, Q., Wang, R. R., Zhang, X. M., Wang, Y. H., Zheng, Y. T., Zhou, J., & Chen, J. J. (2007). A new benzofuranone and anti-HIV constituents from the stems of Rhus chinensis. *Planta medica*, *73*(03), 279-282

Goh, B. C., Hadden, J. A., Bernardi, R. C., Singharoy, A., McGreevy, R., Rudack, T., Cassidy, C. K., & Schulten, K. (2016). Computational methodologies for real-space structural refinement of large macromolecular complexes. *Annual Review of Biophysics*, *45*(1), 253–278.

Govindarajan, M., Karabacak, M., Suvitha, A., & Periandy, S. (2012). FT-IR, FT-Raman, ab initio, HF and DFT studies, NBO, HOMO-LUMO and electronic structure calculations on 4-chloro-3-nitrotoluene. *Spectrochimica acta. Part A, Molecular and biomolecular spectroscopy*, *89*, 137–148.

Gromiha, M. M., Nagarajan, R., & Selvaraj, S. (2019). Protein structural bioinformatics: An overview. *Encyclopedia of Bioinformatics and Computational Biology,* 445-459.

Guedes, I. A., Pereira, F. S. S., & Dardenne, L. E. (2018). Empirical Scoring Functions for Structure-Based Virtual Screening: Applications, Critical Aspects, and Challenges. Frontiers in pharmacology, 9, 1089.

Hendlich, M., Rippmann, F., & Barnickel, G. (1997). LIGSITE: automatic and efficient detection of potential small molecule-binding sites in proteins. Journal of molecular graphics & modelling, 15(6), 359–389.

Hopkins, A. L., & Groom, C. R. (2002). The druggable genome. *Nature Reviews Drug Discovery*, *1*(9), 727–730.

Hou, S., Yang, X., Tong, Y., Yang, Y., Chen, Q., Wan, B., Wei, R., Wang, Y., Zhang, Y., Kong, B., Huang, J., Chen, Y., Lu, T., Hu, Q., & Du, D. (2021). Structure-based discovery of 1H-indole-2-carboxamide derivatives as potent ASK1 inhibitors for potential treatment of ulcerative colitis. *European journal of medicinal chemistry*, *211*, 113114.

Huang, S. Y., Grinter, S. Z., & Zou, X. (2010). Scoring functions and their evaluation methods for protein-ligand docking: recent advances and future directions. Physical chemistry chemical physics : PCCP, 12(40), 12899–12908.

Irwin J. J. (2008). Using ZINC to acquire a virtual screening library. Current protocols in bioinformatics, Chapter14,14.6.1–14.6.23.

Jain A. N. (2003). Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine. *Journal of medicinal chemistry*, *46*(4), 499–511.

Jones, G., Willett, P., Glen, R. C., Leach, A. R., & Taylor, R. (1997). Development and validation of a genetic algorithm for flexible docking. *Journal of molecular biology*, *267*(3), 727–748.

Jorgensen, W. L. (2009). Efficient drug lead discovery and Optimization. *Accounts of Chemical Research*, *42*(6), 724–733.

Kao, Jia Horng. (2018).Hepatitis B virus and liver disease, Springer publication .

Kapetanovic I. M. (2008). Computer-aided drug discovery and development (CADDD): in silico-chemico-biological approach. Chemico-biological interactions, 171(2), 165–176.

Karabacak, M., Bilgili, S., & Atac, A. (2015). Molecular structure, spectroscopic characterization, HOMO and LUMO analysis of 3, 3'-diaminobenzidine with DFT quantum chemical calculations. *Spectrochimica acta. Part A, Molecular and biomolecular spectroscopy*, *150*, 83–93.

Karayiannis, P. (2017). Hepatitis B virus: virology, molecular biology, life cycle and intrahepatic spread. *Hepatology international*, *11*, 500-508.

Kramvis A, Kew MC. The core promoter of hepatitis B virus. J Viral Hepat. 1999; 6 (6):415–27. PubMed PMID: 10607259. Epub 1999/12/22. eng.

Kitchen, D. B., Decornez, H., Furr, J.R., & Bajorath, J. (2004). Docking and scoring in virtual screening for drug discovery: methods and applications. Nature reviews. Drug discovery, 3(11), 935–949.

Klumpp, K., & Crépin, T. (2014). Capsid proteins of enveloped viruses as antiviral drug targets. *Current opinion in virology*, *5*, 63-71.

Klumpp, K., Lam, A. M., Lukacs, C., Vogel, R., Ren, S., Espiritu, C., ... & Flores, O. A. (2015). High-resolution crystal structure of a hepatitis B virus replication inhibitor bound to the viral core protein. *Proceedings of the National Academy of Sciences*, *112*(49), 15196-15201.

Kosar, B., & Albayrak, C. (2011). Spectroscopic investigations and quantum chemical computational study of (E)-4-methoxy-2-[(p-tolylimino) methyl] phenol. *Spectrochimica acta. Part A, Molecular and biomolecular spectroscopy*, *78*(1), 160–167.

Kosar, B., & Albayrak, C. (2011). Spectroscopic investigations and quantum chemical computational study of (E)-4-methoxy-2-[(p-tolylimino) methyl] phenol. *Spectrochimica acta. Part A, Molecular and biomolecular spectroscopy*, *78*(1), 160–167.

Kubinyi, H. (n.d.). Hydrogen bonding: The last mystery in drug design? *Pharmacokinetic Optimization in Drug Research*, 513–524.

Küçükgüzel, Ş. G., & Şenkardeş, S. (2015). Recent advances in bioactive pyrazoles. *European Journal of Medicinal Chemistry*, *97*, 786–815.

Lane, H. Y., Lin, C. H., Green, M. F., Hellemann, G., Huang, C. C., Chen, P. W., Tun, R., Chang, Y. C., & Tsai, G. E. (2013). Add-on treatment of benzoate for schizophrenia: a randomized, double-blind, placebo-controlled trial of D-amino acid oxidase inhibitor. *JAMA psychiatry*, *70*(12), 1267–1275.

Laskowski R. A. (1995). SURFNET: a program for visualizing molecular surfaces, cavities, and intermolecular interactions. Journal of molecular graphics, 13(5), 323–308.

Laskowski, R. A., & Swindells, M. B. (2011). LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. Journal of chemical information and modeling, 51(10), 2778–2786.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). Procheck: A program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, *26*(2), 283–291.

Levitt, D. G., & Banaszak, L. J. (1992). POCKET: a computer graphics method for identifying and displaying protein cavities and their surrounding amino acids. Journal of molecular graphics, 10(4), 229–234.

Li, H., Peng, J., Leung, Y., Leung, K. S., Wong, M. H., Lu, G., & Ballester, P. J. (2018). The Impact of Protein Structure and Sequence Similarity on the Accuracy of

Machine-Learning Scoring Functions for Binding Affinity Prediction. Biomolecules, 8(1), 12.

Liang, J., Edelsbrunner, H., & Woodward, C. (1998). Anatomy of protein pockets and cavities: measurement of binding site geometry and implications for ligand design. Protein science: a publication of the Protein Society, 7(9), 1884–1897.

Lin, X., Li, X., & Lin, X. (2020). A Review on Applications of Computational Methods in Drug Screening and Design. *Molecules (Basel, Switzerland)*, *25*(6), 1375.

Lionta, E., Spyrou, G., Vassilatis, D., & Cournia, Z. (2014). Structure-based virtual screening for Drug Discovery: Principles, applications and recent advances. *Current Topics in Medicinal Chemistry*, *14*(16), 1923–1938.

Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Advanced drug delivery reviews, 46(1-3), 3–26.

Liu, Z., Tang, L., Zhu, H., Xu, T., Qiu, C., Zheng, S., Gu, Y., Feng, J., Zhang, Y., & Liang, G. (2016). Design, synthesis, and structure–activity relationship study of novel indole-2-carboxamide derivatives as anti-inflammatory agents for the treatment of sepsis. *Journal of Medicinal Chemistry*, *59*(10), 4637–4650.

López-López, E., Naveja, J. J., & Medina-Franco, J. L. (2019). DataWarrior: an evaluation of the open-source drug discovery tool. Expert opinion on drug discovery, 14(4), 335–341.

Lounnas, V., Ritschel, T., Kelder, J., McGuire, R., Bywater, R. P., & Foloppe, N. (2013). Current progress in Structure-Based Rational Drug Design marks a new mindset in drug discovery. *Computational and structural biotechnology journal*, *5*, e201302011.

Lupia, T., Corcione, S., Mornese Pinna, S., & De Rosa, F. G. (2020). New cephalosporins for the treatment of pneumonia in internal medicine wards. Journal of thoracic disease, 12(7), 3747–3763.

Ma, C., Liu, W. G., Liu, W. D., Xi, C. C., Xiong, F., & Zhang, S. P. (2022). Molecular Docking and 3D-QSAR Studies on a Series of Benzenesulfonamide Derivatives as a Hepatitis B Virus Capsid Assembly Inhibitor. *Polycyclic Aromatic Compounds*, *42*(6), 3622-3637.

Mahesh, A. R., & Murugan, V. (2016). SYNTHESIS OF A BIOLOGICALLY PERTINENT MOLECULE, AURONES–AN EXTENSIVE REVIEW. *European Journal of Biomedical*, *3*(1), 112–116.

Na, H. G., Imran, A., Kim, K., Han, H. S., Lee, Y. J., Kim, M.-J., Yun, C.-S., Jung, Y.-S., Lee, J.-Y., & Han, S. B. (2020). Discovery of a new sulfonamide hepatitis B capsid assembly modulator. *ACS Medicinal Chemistry Letters*, *11*(2), 166–171.

Maia, R. T., & Amador, V. C. (2018). Molecular docking for detoxifying enzyme studies. *Molecular Docking*.

Mallipeddi, P. L., Kumar, G., White, S. W., & Webb, T. R. (2014). Recent advances in computer-aided drug design as applied to anti-influenza drug discovery. *Current topics in medicinal chemistry*, *14*(16), 1875–1889.

McGann, M. R., Almond, H. R., Nicholls, A., Grant, J. A., & Brown, F. K. (2003). Gaussian docking functions. *Biopolymers*, *68*(1), 76–90.

Meng, X. Y., Zhang, H. X., Mezei, M., & Cui, M. (2011). Molecular docking: a powerful approach for structure-based drug discovery. Current computer-aided drug design, 7(2), 146–157.

Miller, D. C., Lunn, G., Jones, P., Sabnis, Y., Davies, N. L., & Driscoll, P. (2012). Investigation of the effect of molecular properties on the binding kinetics of a ligand to its biological target. *MedChemComm*, *3*(4), 449.

Morris, A. L., MacArthur, M. W., Hutchinson, E. G., & Thornton, J. M. (1992). Stereochemical quality of protein structure coordinates. *Proteins*, *12*(4), 345–364.

Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. Journal of computational chemistry, 30(16), 2785–2791.

Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. Journal of computational chemistry, 30(16), 2785–2791.

O'Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., & Hutchison, G. R. (2011). Open Babel: An open chemical toolbox. Journal of cheminformatics, 3, 33.

Ou-Yang, S. S., Lu, J. Y., Kong, X. Q., Liang, Z. J., Luo, C., & Jiang, H. (2012). Computational drug discovery. *Acta pharmacologica Sinica*, *33*(9), 1131–1140.

Pang, X., & Zhou, H. X. (2017). Rate Constants and Mechanisms of Protein-Ligand Binding. *Annual review of biophysics*, *46*, 105–130

Petit, J., Meurice, N., Kaiser, C., & Maggiora, G. (2012). Softening the Rule of Five--where to draw the line?. *Bioorganic & medicinal chemistry*, *20*(18), 5343–5351.

Pina, A. S., Hussain, A., & Roque, A. C. (2009). An historical overview of drug discovery. *Methods in molecular biology (Clifton, N.J.)*, *572*, 3–12.

Prentis, R. A., Lis, Y., & Walker, S. R. (1988) Pharmaceutical innovation by the seven UK-owned pharmaceutical companies (1964-1985). British journal of clinical pharmacology, 25(3), 387–396.

Pushpakom, S., Iorio, F., Eyers, P. A., Escott, K. J., Hopper, S., Wells, A., Doig, A., Guilliams, T., Latimer, J., McNamee, C., Norris, A., Sanseau, P., Cavalla, D., & Pirmohamed, M. (2019). Drug repurposing: progress, challenges and recommendations. *Nature reviews. Drug discovery*, *18*(1), 41–58.

Rajan, S., Baek, K., & Yoon, H. S. (2013). C-H…O hydrogen bonds in FK506-binding protein-ligand interactions. *Journal of molecular recognition: JMR*, *26*(11), 550–555.

Ramachandran, G. N., & Sasisekharan, V. (1968). Conformation of polypeptides and proteins. *Advances in protein chemistry*, *23*, 283–438.

Rarey, M., Kramer, B., Lengauer, T., & Klebe, G. (1996). A fast flexible docking method using an incremental construction algorithm. *Journal of molecular biology*, *261*(3), 470–489.

Revelle, L. K., D'Avignon, D. A., & Wilson, J. A. (1996). 3-[(Phenylacetyl) amino]-2, 6-piperidinedione hydrolysis studies with improved synthesis and characterization of hydrolysates. *Journal of pharmaceutical sciences*, *85*(10), 1049–1052.

Ruiz-Carmona, S., Alvarez-Garcia, D., Foloppe, N., Garmendia-Doval, A. B., Juhos, S., Schmidtke, P., Barril, X., Hubbard, R. E., & Morley, S. D. (2014). rDock: a fast, versatile and open source program for docking ligands to proteins and nucleic acids. *PLoS computational biology*, *10*(4), e1003571.

Sander, T., Freyss, J., von Korff, M., & Rufener, C. (2015). DataWarrior: an open-source program for chemistry aware data visualization and analysis. Journal of chemical information and modeling, 55(2), 460–473.

Schreiber, G., Haran, G., & Zhou, H. X. (2009). Fundamental aspects of protein-protein association kinetics. *Chemical reviews*, *109*(3), 839–860.

Seeliger, D., & de Groot, B. L. (2010). Ligand docking and binding site analysis with PyMOL and Autodock/Vina. Journal of computer-aided molecular design, 24(5), 417–422.

Singh, Dev buksh; 2020 Computer aided drug designing, springer publications.

Sliwoski, G., Kothiwale, S., Meiler, J., & Lowe, E. W. (2013). Computational methods in drug discovery. *Pharmacological Reviews, 66*(1), 334-395.

Sousa, S. F., Fernandes, P. A., & Ramos, M. J. (2006). Protein-ligand docking: current status and future challenges. Proteins, 65(1), 15–26.

Tan, J. J., Cong, X. J., Hu, L. M., Wang, C. X., Jia, L., & Liang, X. J. (2010). Therapeutic strategies underpinning the development of novel techniques for the treatment of HIV infection. *Drug discovery today*, *15*(5-6), 186–197.

Trott, O., & Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*, *31*(2), 455–461.

Unni, S., Aouti, S., Thiyagarajan, S., & Padmanabhan, B. (2020). Identification of a repurposed drug as an inhibitor of Spike protein of human coronavirus SARS-CoV-2 by computational methods. *Journal of Biosciences*, *45*, 1-20.

Venkatachalam, C. M., Jiang, X., Oldfield, T., & Waldman, M. (2003). LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites. *Journal of molecular graphics & modelling*, *21*(4), 289–307.

Wallace, A. C., Laskowski, R. A., & Thornton, J. M. (1995). LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein engineering*, *8*(2), 127–134.

Wang, X., Song, K., Li, L., & Chen, L. (2018). Structure-based drug design strategies and challenges. *Current Topics in Medicinal Chemistry*, *18*(12), 998–1006.

Waring M. J. (2010). Lipophilicity in drug discovery. *Expert opinion on drug discovery*, *5*(3), 235–248.

Yuan, S., Chan, H. C. S., & Hu, Z. (2017). Using pymol as a platform for computational drug design. *WIREs Computational Molecular Science*, *7*(2).

Yuan, T. T. T., & Shih, C. (2000). A frequent, naturally occurring mutation (P130T) of human hepatitis B virus core antigen is compensatory for immature secretion phenotype of another frequent variant (I97L). *Journal of virology*, *74*(10), 4929-4932.

Zhou, P., Huang, J., & Tian, F. (2012). Specific noncovalent interactions at protein-ligand interface: implications for rational drug design. *Current medicinal chemistry*, *19*(2), 226–238.

Zhou, Z., Hu, T., Zhou, X., Wildum, S., Garcia-Alcalde, F., Xu, Z., ... & Qin, N. (2017). Heteroaryldihydropyrimidine (HAP) and sulfamoylbenzamide (SBA) inhibit hepatitis B virus replication by different molecular mechanisms. *Scientific reports*, *7*(1), 1-12.