IJPSR (2024), Volume 15, Issue 6 (Research Article)

INTERNATIONAL JOURNAL OF PHARMACEUTICAL SCIENCES AND RESEARCH

Received on 25 November 2023; received in revised form, 02 February 2024; accepted, 05 April 2024; published 01 June 2024

IDENTIFICATION OF NOVEL NCP7 INHIBITORS FOR HIV DRUG DISCOVERY USING LIGAND-BASED COMPUTATIONAL APPROACHES

K. V. Lisina * and Shanmughavel Piramanayagam

DBT-Bioinformatics Centre, Computational Biology Laboratory, Department of Bioinformatics, Bharathiar University, Coimbatore - 641046, Tamil Nadu, India.

Keywords:

Pyridinioalkanoyl thioesters, Benzamide-based thiolcarbamates derivatives, Pharmacophore, 3D-QSAR, Glide

Correspondence to Author:
K. V. Lisina

Research Scholar,
DBT-Bioinformatics Centre,
Computational Biology Laboratory,
Department of Bioinformatics,
Bharathiar University, Coimbatore - 641046, Tamil Nadu, India.

E-mail: lisina.kv@gmail.com

ABSTRACT: Background: NCp7 is a nucleocapsid protein of HIV-1, a retrovirus causing AIDS in human beings. It is a small protein of 72 amino acids that plays an important role in being a chaperone protein involved in viral development. HIV has shown resistance to drugs engineered to treat its infections. **Objectives:** In this study, a Structure-Activity Relationship analysis of two distinct derivatives based on point pharmacophore was carried out to obtain an NCp7 inhibitor of better performance. **Materials and Methods:** A Pharmacophore and 3D-QSAR modeling of Pyridinioalkanoyl thioesters and Benzamide-based thiolcarbamates derivatives was performed. **Results:** A four-point hypothesis of HHRR pharmacophoric feature was determined. A 3D-QSAR model was built which was validated through regression analysis. Molecular docking often lead compounds showed that the compound ZINC65398698 possesses a higher affinity towards NCp7 protein and binds well whereas other leads possessed only moderate binding affinity. The ADME/T prediction of all the ten lead compounds showed accepted values of biophysical properties. As ZINC65398698 showed higher binding affinity, a molecular dynamics simulation was performed to determine the stability of the protein-ligand system and to study the atomistic detail of protein residue behavior after forming a complex with the ligand. The RMSD and RMSF analysis showed that the complex is highly stable throughout the production run. So the complex can be more stable within the same environment and act as potential inhibitors was determined in this procedure. **Conclusions:** The suggested Ligand can be then subjected to a model study after the synthesizing procedure.

INTRODUCTION: HIV (Human Immunodeficiency Virus) weakens the immune system of a patient by attacking T cells. HPV infection or cervical inflammation increases the risk of HIV infection in men and women, potentially fueling their mutual spread [1]. HIV also reduces the number of CD4 cells (T cells) in the body. Also, the person is more likely to get other infections.

QUICK RESPONSE CODE

DOI: 10.13040/IJPSR.0975-8232.15(6).1682-02

This article can be accessed online on www.ijpsr.com

DOI link: https://doi.org/10.13040/IJPSR.0975-8232.15(6).1682-02

HIV is a retroviral agent that causes acquired immune deficiency syndrome (AIDS) and has been linked to 4 million deaths in 2016 [2]. Sub-Saharan Africa remains the region with more than 80% of all infected individuals, while HIV-2 is restricted to West Africa along the Upper Guinean forests [3].

HIV-AIDS patients in India have a higher risk of various cancers, with a distinct pattern of malignancies based on age, sex, and CD4 counts at the time of diagnosis [4]. HIV-positive people in the Dharwad district face severe social, economic, and health challenges, with most living in rural areas, having no opportunistic infections, and living in nuclear families [5]. Small molecules that bind to the viral capsid protein can be potent inhibitors of HIV infection and exhibit high barriers to viral

resistance[6]. Systematic reviews regarding the mortality and rate of infections have been so prevalent that radical efforts are being undertaken for a curative emergence against the virus. Two types of HIV could be recognized genetically and antigenically. HIV-1 is the reason for the current overall pandemic, while HIV-2 is found in West Africa and yet infrequently somewhere else. Studies suggest that HIV-1 is more pandemic than HIV-2, which is primarily an epidemic in West Africa and better controlled by affected individuals. HIV-2 has not spread as a large pandemic and is well controlled by the majority of affected individuals, while HIV-1 has become more pandemic due to its low efficiency at productively infecting dendritic cells [7]. HIV-2 is less pathogenic and confined mainly to West Africa, while HIV-1 is pandemic and aggressive [8].

In the last four years, adult HIV occurrence has risen in some regions. Overall, India's HIV outbreak is slowing down. Between 2010 and 2017 new infections declined by 27% and AIDS-related deaths more than halved, falling by 56%. However, in 2017, new infections increased to 88,000 from 80,000 and AIDS-related deaths increased to 69,000 from 62,000. In 2017, 79% of people living with HIV were aware of their status, of which 56% were on antiretroviral treatment (ART) [9]. The HIV outbreak in India is driven by sexual transmission, which accounted for 86% of new infections in 2017/2018 [10]. The three states with the highest HIV incidence Manipur, Mizoram, and Nagaland are in the east of the country.

Protein Nucleocapsid (NCp7) is a nucleocapsid protein of HIV-1, a retrovirus that causes AIDS in humans [11]. It is a small protein of 72 amino acids that plays an important role in being a chaperone protein involved in viral development. The nucleocapsid protein plays a crucial role in compacting viral double-stranded DNA within the mature HIV-1 capsid, potentially preventing capsid uncoating and NC loss [12]. The nucleocapsid protein plays a crucial role in compacting viral double-stranded DNA within the mature HIV-1 capsid, potentially preventing capsid uncoating and NC loss. Multiple binding modes of the HIV-1 nucleocapsid protein compact double-stranded DNA into a conformation compatible with reverse transcription, regulating genomic pressure on the capsid and preventing premature uncoating. The nucleocapsid proteins of several retroviruses have been studied extensively, and high-resolution structural data are available for the NC proteins of HIV-1 [13]. Song *et al.,* 2002 and Goel *et al.,* 2002 reported Pyridinioalkanoyl Thioesters and Benzamide-based Thiolcarbamates [14, 15]. This work currently employs a computational measure to understand the effect of variants of pyridinioalkanoyl thioesters (PATEs) and benzisothiazolone derivatives (BITA) as drugs on the NCp7 domain present in the nucleocapsid of the HIV-1 protein. In this work, we describe the discovery of a new active combination of Carbamates and Thioesters ligands based on the NCp7 domain of HIV Type 1.

A pharmacophore model for HIV-1 Nucleocapsid P7 inhibitors was used for the QSAR study, it was compared using database screening and molecular docking to discover novel lead compounds. Pharmacomodels help define the mechanism of therapeutic compounds, enabling the development of novel drugs and improved understanding of pharmacology [16].

QSAR modeling is used during drug discovery for predicting the biological activity of drug candidates. QSAR models are used to determine the biological properties of chemical molecules based on their chemical structure [17]. Quantitative structure-activity relationship (QSAR) techniques have advanced in translational toxicology, reducing animal testing and promoting computational approaches as viable tools for reducing animal testing [18].

Further computational prediction of pharmacokinetic parameters like Absorption, Distribution, Metabolism, and Excretion (ADME) and toxicity studies have ended up progressively important in drug selection and advancement procedure and are guaranteeing tools for right-on-time screening of potential drug candidates[19]. Characterizing absorption, distribution, metabolism, and excretion (ADME) processes is crucial for drug-target and drug-body interactions, enabling reliable interspecies and human pharmacokinetic estimations [20]. Studies suggest that computer-aided drug design helps accelerate drug development by aiding in experiment design,

hit identification, optimization, and evaluation, and predicting pharmacokinetics and potential effects, while also benefiting fields like bioinformatics and biomedical engineering [21]. Characterizing absorption, distribution, metabolism, and excretion (ADME) processes is crucial for drug-target and drug-body interactions, enabling reliable interspecies and human pharmacokinetic estimations. De novo drug design generates novel molecular structures from atomic building blocks with no a priori relationships, using computational methods like structure-based and ligand-based design [22].

The application of Computer Aided Drug Design (CADD) and Quantitative Structure-activity Relationship (QSAR) study on developing novel drugs for these diseases is limited. Quantitative structure-activity relationship (QSAR) techniques have advanced in translational toxicology, reducing animal testing and promoting computational approaches as viable tools for reducing animal testing. 3D-QSAR/MD method combines 3D-quantitative structure-activity relationships (QSAR) with computational molecular dynamics (MD) studies for drug design, making it more cost-effective and less time-consuming [23]. 3D QSAR is a statistical model used in drug design and optimization, and can be taught and learned through web applications like 3d-qsar.com [24]. 3D-QSAR is a method for predicting the structural and functional properties of molecules, providing useful information for designing new and effective inhibitors [25].

Molecular dynamics (MD) is a computer simulation method for studying the physical movements of atoms and molecules. In this study, it is used to find out whether the best inhibitor obtained from the final screening is stable throughout the simulation period inside the nucleocapsid domain without any significant fluctuation within some period. Molecular dynamics simulation is a widely used approach for understanding complex systems on the atomistic scale, with applications in physics, chemistry, engineering, life, and medical science [26]. Molecular dynamics (MD) is a computer simulation method for studying the physical movements of atoms and molecules, providing detailed microscopic sampling on molecular scale [27].

Molecular dynamics techniques help characterize various conformations and molecular interactions, allowing for the search of suitable compounds for specific biochemical purposes and reducing the consumption of chemical reagents and time [28]. Molecular dynamics simulation is a widely used approach for understanding complex systems on the atomistic scale, with applications in physics, chemistry, engineering, life, and medical science. Molecular dynamics studies collect functional information on molecules at specific times and in specific environments, revealing properties and mechanisms of action. Molecular dynamics (MD) is a computer simulation method for studying physical movements of atoms and molecules, providing detailed microscopic sampling on molecular scale. Molecular dynamics simulation generates atomic trajectories of a system of particles by numerical integration of Newton's equation of motion, for a specific interatomic potential.

MATERIALS AND METHODS:

Biological Data Set: The selected target protein structure of the Nucleocapsid protein P7 domain of HIV type 1 virus was retrieved from Protein Databank with the PDB ID: 1ESK. Twenty-five Pyridinioalkanoyl Thioesters derivatives and eighteen Benzamide-based Thiolcarbamates derivatives were obtained from different references.

Preparing Ligands: The molecules were drawn using ACD chemdraw freeware. Studies suggest using ChemDraw freeware for drawing and optimizing chemical structures, improving chemistry skills and understanding, and predicting chemical shifts in organic compounds. ChemDraw software is used as a medium for learning hydrocarbon materials in drawing 2D and 3D chemical structures, improving teacher ability and facilitating students' understanding of abstract chemistry learning [29].

ChemDraw Professional 16.0 software was used to accurately draw the chemical structures of 1,3,4-Oxadiazole derivatives for molecular docking simulations targeting breast and lung cancers[30]. ChemDraw Ultra software was used to draw chemical structures of novel carboxamide series compounds for virtual molecular docking simulations on Mycobacterium tuberculosis target

DNA gyrase [31]. LigPrep2.7 was used for structure improvement and energy minimization [32]. Ligprep in Schrodinger is used for pre-processing of ligands before performing extra precision docking with target proteins. Studies suggest that LigPrep in Schrodinger is used for pre-processing and optimizing ligands for molecular interactions, *in-silico* screening, and virtual screening in drug discovery and repurposing. Ligprep in Schrodinger is used for pre-processing of ligands before performing extra precision docking with target proteins [33]. LigPrep in Schrodinger software optimizes ligands prepared from US-FDA drug-bank for molecular interactions and drug repurposing [34]. Ligprep in Schrodinger is a module that prepares approved drugs for virtual screening and molecular dynamics simulations [35]. Studies suggest that the ZINC database is a collection of commercially-available molecules for ligand and drug discovery, featuring tools for searching, predicting zinc binding sites, and estimating zinc deficiency prevalence. The ZINC database is a collection of commercially-available natural products [36]. Energy-minimized structures and their activity values were imported to PHASE [37] to generate a pharmacophore model. IC_{50} value of the reference structure converted into pIC_{50} value. The predicted IC_{50} values from the IC_{50} available in micromolar in the reference structure were calculated using the following equation; $pIC_{50} = (-\log IC_{50}+6)$. pIC_{50} values above 5.1 were considered active and below 3.9 were considered inactive and the rest were moderately active. The dataset was divided randomly into a training set and a test set by taking 80% of the total molecules in the training set and 20 % in the test set.

Pharmacophore Modeling: The PHASE module of the Schrodinger suite is used for modeling pharmacophore. Pharmacophore sites were created from pharmacophore features. Common Pharmacophore hypotheses were generated from the 28-member training set molecules by using a grid spacing of 1.0 Å. Then the developed pharmacophore model was validated by calculating the activity of test set molecules. The association between the experimental and predicted activities of the molecules of training and test sets is shown in **Table 1.**

TABLE 1: ACTUAL AND PREDICTED ACTIVITIES FOR TRAINING AND TEST SET COMPOUNDS ALONG WITH CALCULATED PROPERTIES

Compound ID	R_1	R_2	R_3	X	IC_{50} (µM)	Experimental $logPIC_{50}$	Predicted $logIC_{50}$	Residual	Pharm set	QSAR set
1	NH_2	H	H	Cl	6.8	5.167	5.27	-0.103	Active	Training
2	NH_2	H	H	Cl	5.4	5.267	5.45	-0.183	Active	Training
6	NH_2	H	H	Br	201	3.698	4.77	-1.072	Inactive	Training
7	NH_2	H	CH_3	Br	101	3.995	4.58	-0.585	Moderate	Training
8	NH_2	H	CH_2CH_3	Br	200	3.698	3.6	0.098	Inactive	Training
10	NH_2	H	$CH_2CH(CH_3)_2$	Br	5.2	5.283	5.45	-0.167	Active	Training
12	NH_2	H	$CH_2C_6H_5$	Br	4.5	5.346	4.97	0.376	Active	Training
13	NH_2	H	CH_2CONH_2	Br	317	3.498	3.82	-0.322	Inactive	Training
14	NH_2	H	$CH_2CH_2CONH_2$	Br	317	3.498	3.9	-0.402	Inactive	Training
15	NH_2	CH_3	CH_3	Br	23	4.46	4.35	0.11	Moderate	Training
17	NH_2	H	CH_3	Br	132	3.879	3.88	-0.001	Inactive	Training
20	$NHCH_3$	H	H	Br	21	4.677	4.54	0.137	Moderate	Training
21	OH	H	H	Br	317	3.498	3.46	0.038	Inactive	Training
22	OH	H	H	Br	10	5	4.74	0.26	Moderate	Training
23	OC_2H_5	H	H	Br	317	3.498	3.61	-0.112	Inactive	Training
24	OC_2H_5	H	H	Br	1.4	5.853	5.35	0.503	Active	Training
25	OC_2H_5	H	H	Br	0.9	6.045	6.11	-0.065	Active	Training
26	OCH_3	H	CH_3	Br	5.9	5.229	5.36	-0.131	Active	Training
5a	H	CH_3	-	-	41.2	4.385	4.16	0.225	Moderate	Training
5c	CH_3	CH_3	-	-	38.3	4.416	4.37	0.046	Moderate	Training
5d	H	CH_3	-	-	40.2	4.395	4.44	-0.045	Moderate	Training
5h	H	$C(CH_3)_3$	-	-	142	3.847	-	-	Inactive	Training
5k	CH_3	4-FC_6H_4	-	-	130	3.886	-	-	Inactive	Training
5m	H	4-$CF_3C_6H_4$	-	-	128	3.892	3.71	0.182	Inactive	Training
5n	H	Br	-	-	4.9	5.309	-	-	Active	Training
5o	CH_3	Br	-	-	8	5.096	-	-	Moderate	Training
6b	H	Py+Br-	-	-	153	3.815	3.9	-0.085	Inactive	Training
6c	CH_3	Py+Br-	-	-	200	3.698	3.88	-0.182	Inactive	Training

4	NH_2	H	CH_3	Br	63	4.2	4.99	-0.79	Moderate	Test
5	NH_2	H	$CH(CH_3)(C_2H_5)$	Cl	10	5	4.18	0.82	Moderate	Test
9	NH_2	H	$CH(CH_3)_2$	Br	9.9	5.004	5.11	-0.106	Moderate	Test
11	NH_2	H	$CH(CH_3)(C_2H_5)$	Br	20.1	4.696	4.52	0.176	Moderate	Test
16	NH_2	H	H	Br	110	3.958	4.21	-0.252	Moderate	Test
18	NH_2	H	H	-	12	4.92	4.8	0.12	Moderate	Test
19	NH_2	H	H	Br	9	5.045	4.58	0.465	Moderate	Test
5b	H	CH_3	-	-	39.4	4.404	4.02	0.384	Moderate	Test
5e	CH_3	$CH(CH_3)_2$	-	-	76	4.119	-	-	Moderate	Test
5f	H	$CH(CH_3)_2$	-	-	55.1	4.258	-	-	Moderate	Test
5g	CH_3	$C(CH_3)_3$	-	-	125	3.903	-	-	Moderate	Test
5i	H	Cyc-Pentyl	-	-	17.3	4.761	-	-	Moderate	Test
5j	H	Cyc-Hexyl	-	-	11.9	4.924	-	-	Moderate	Test
5l	CH_3	4-$NO_2C_6H_4$	-	-	23.1	4.636	-	-	Moderate	Test
6a	H	Py+Br-	-	-	70	4.154	4.49	-0.336	Moderate	Test

Tabular representation of predicted and calculated properties of reference ligands selected as training and test set compounds. The activity above 5.1 was arbitrarily taken as active and activity less than 3.9 as inactive during model generation. A total of 7 were active in the training set and compound 25 was selected as a highly active compound in the training set. Compound 6c was selected as the least active in the test set. The scaffolds of molecules are categorized into R1, R2, and R3 for the functional group and X for the Ionic group but the compound from two different sources does not have the same feature set so the hyphen denoted in the tables do not have any R3 group and X for the corresponding compound. In the Pharmacophore hypothesis scoring function, out of 43 molecules 5h, 5k, 5n, 5o, 5e, 5f, 5g, 5i, 5j, and 5l molecules are not scored with Predicted IC_{50} value because of a lack of a functional group to score the defined hypothesis correspondingly.

3D-QSAR Studies: The 3D-QSAR studies were carried out using PHASE version 3.0 implemented in the Maestro 9.3 molecular modeling package from Schrodinger Molecular Modeling Interface. In QSAR studies, the appropriate conformation of the compound is required for the accurate calculation of 3D descriptors. All the molecules were divided into a training set and a test set to maintain the structure and activity diversity in both the sets for the QSAR model and pharmacophore generation and validation. The prepared ligands were used for generating common pharmacophore and QSAR model building. QSAR modeling was carried out using the selected hypothesis by dividing the datasets into a training set (80%) and a test set (20%) in a random manner.

Phase presents two options for the alignment of the 3D structure of molecules the pharmacophore-based alignment and the atom-based alignment. In this study, an atom-based QSAR model was explained the Structure Activity Relationship (SAR). In atom-based QSAR, a molecule is treated as a set of overlapping van der Waals spheres. Each atom is placed into one of six categories according to a simple set of rules: hydrogens attached to polar atoms are classified as hydrogen bond donors (D); carbons, halogens and hydrogens are classified as hydrophobic/non-polar (H); atoms with an explicit negative ionic charge are classified as negative ionic (N); atoms with an explicit positive ionic charge are classified as positive ionic (P); non-ionic nitrogen and oxygen are classified as electron-withdrawing (W); and all other types of atoms are classified as miscellaneous (X). Atom-based QSAR models were generated for two sets of point pharmacophores. Here 8 active compounds, 12 inactive compounds, and the rest moderately active compounds were chosen to perform a point pharmacophore. Grid spacing was set up to 1.0 Å.

Pharmacophore Screening against Chemical Databases with Predicted Pharmacophore Model: Around 7,75,000 property-predicted ligands from both ZINC and NCI databases [38] were allowed to interact with the pharmacophore model to find a potent molecule. Based on a structure target-based search against ZINC and NCI database, around 42,72,818 compounds were obtained in 2D SDF (Standard Dimension Format) format. Virtual screening is one of the fastest and most accurate methods to identify the potential ligand with desired pharmacophore kinetics properties. The well-validated HHRR hypothesis was used to search a 3D database for structures that match the pharmacophoric features of the model. Virtual screening was carried out using Zinc Pharmer [39] which uses the pharmacophore to efficiently search the ZINC database of fixed conformers for pharmacophore matches.

To accomplish the best 3D similarity search, the constraints used also included a maximum of 0.7 Å RMSD, obeying 10 rotatable bonds cut off and molecular weight range of under 500 Dalton, hydrogen bond acceptor not more than 10, hydrogen bond donor not more than 5, and partition coefficient log *p* less than 5. A molecule that fits well with the pharmacophoric features of the HHRR hypothesis was retrieved as a hit.

Molecular Docking: In the current study, we carried out docking of the top fitted ligand in the 3D-QSAR model i.e., compound ZINC 65398698 into the active site of the NCp7 (PDB ID: 1ESK). For performing molecular docking between the protein structure and ligand structure, both of them have to be prepared (i.e.) proper atom charges and bond orders to be added, necessary Hydrogen atoms were added and optimized. Glide ligand docking jobs require a set of previously calculated receptor grids and ligand structures. The ligand was docked with the target protein using the Glide module of the Schrödinger suite. The best-fitted ligand interacts with the NCP7 domain. Extra precision docking profile and high throughput screening methods were followed.

Molecular Dynamic Simulation of the Docked Complex: GROMACS[40] is computational software used to perform Molecular Dynamics Simulations and Energy Minimization. Molecular dynamics simulation studies provide the dynamical behavior of the system in real-time. The molecular dynamic simulation was performed with GROMOS96 43a1 force fields using the GROMACS 4.5.5-1 package. Before starting the simulations, all the models were solvated with the explicit simple point charge (SPC216) water in a cubic box with the Periodic Boundary Conditions (PBC). The system was neutralized with 7 chlorine ions by replacing the solvent atoms. The energy minimization was performed for the system concerned by using the steepest descent method. The topology files and charges for the ligand atoms were generated by the PRODRG2 Server [41]. Then, 20 ns for the complex having the best G Score and 10 ns for four consecutive best score fits of MD simulations were carried out with a timestep of 1 fs.

Pharmacodynamic and Kinetic Study of the Docked Ligand: Pharmacodynamics is the investigation of a drug's influence on an organism while pharmacokinetics is the investigation of how an organism affects the drug. Pharmacodynamics and pharmacokinetics influence the measurement of dosage and its effects. Using the Qikprop module, [42] the biophysical, biochemical, and physicochemical properties of the molecule were measured. These properties aid in the identification and development of drug candidates.

RESULTS AND DISCUSSION

3-Dimensional Ligand Structure Generation: Based on the substitution of the R group from the literature source, R group atoms were added to the basic structure of Pyridinioalkanoyl Thioesters and Benzamide-based Thiolcarbamates **Fig. 1**, and they were generated using ChemDraw freeware. The structures of 43 derivatives are shown in **Fig. 2.**

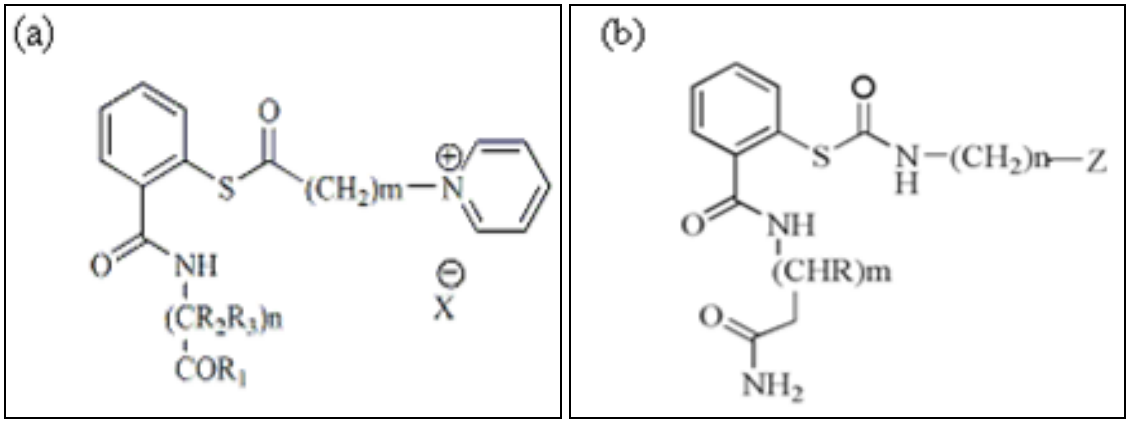
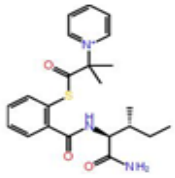
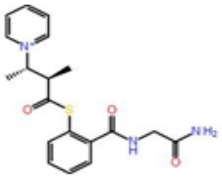


FIG. 1: GENERAL R-GROUP SCAFFOLD OF (A) PYRIDINIOALKANOYL THIOESTERS AND (B) BENZAMIDE-BASED THIOLCARBAMATES HYPOTHESIS GENERATION. IT SHOWS THE GENERAL STRUCTURES OF PYRIDINIOALKANOYL THIOESTERS AND BENZAMIDE-BASED THIOLCARBAMATES USED FOR PHARMACOPHORE GENERATION DRAWN USING CHEMDRAW FREEWARE



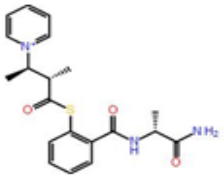
Title: 1

Ref #: 1



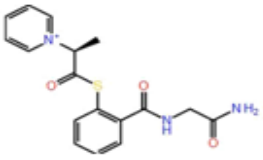
Title: 2

Ref #: 2



Title: 3

Ref #: 4

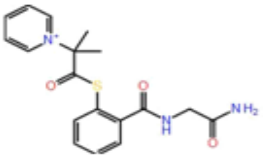


IC_{50}: 6.8µM

Title: 4

Ref #: 5

IC_{50}: 10 µM

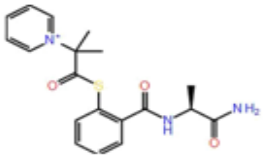


IC_{50}: 5.4 µM

Title: 5

Ref #: 6

IC_{50}: 201 µM

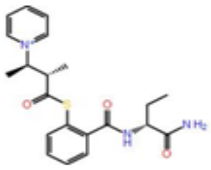


IC_{50} : 63 µM

Title: 6

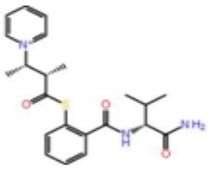
Ref #: 7

IC_{50} : 101 µM



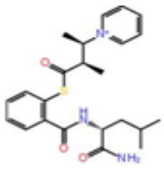
Title: 7

Ref #: 8



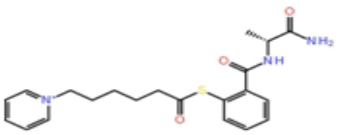
Title: 8

Ref #: 9



Title : 9

Ref #: 10

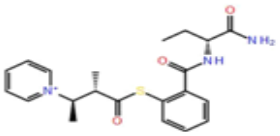


IC_{50}: 200 µM

Title: 10

Ref #: 11

IC_{50}: 20.1 µM

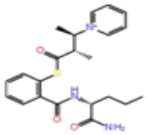


IC_{50}: 9.9 µM

Title: 11

Ref #: 12

IC_{50}: 4.5 µM

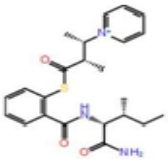


IC_{50}: 5.2 µM

Title: 12

Ref #: 13

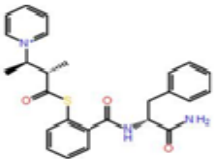
IC_{50}: 317 µM



Title: 13

Ref #: 14

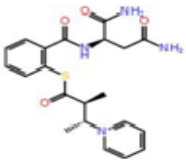
IC_{50}: 317 µM



Title: 14

Ref #: 15

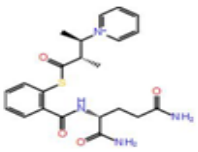
IC_{50}: 23 µM



Title: 15

Ref #: 16

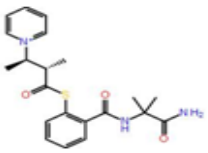
IC_{50}: 110 µM



Title: 16

Ref #: 17

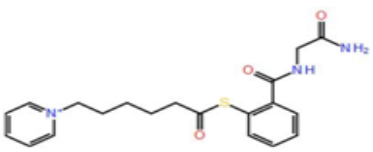
IC_{50}: 132 µM



Title: 17

Ref #: 18

IC_{50} : 12 µM

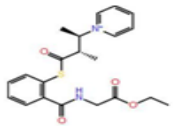


Title: 18

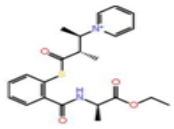
Ref #: 19

IC_{50} : 9 µM

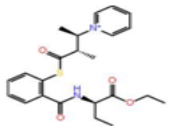
Title: 19
Ref #: 20
IC_{50}: 21 µM



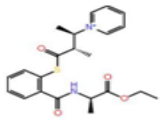
Title: 20
Ref #: 21
IC_{50}: 317 µM



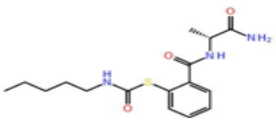
Title: 21
Ref #: 22
IC_{50}: 10 µM



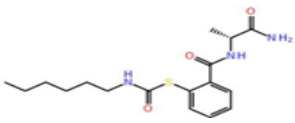
Title: 22
Ref #: 23
IC_{50}: 317 µM



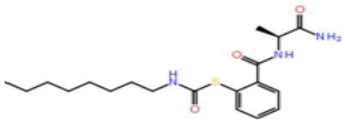
Title: 23
Ref #: 24
IC_{50}: 1.4 µM



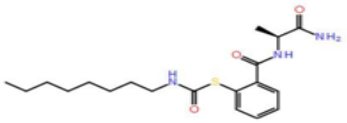
Title: 24
Ref #: 25
IC_{50}: 0.9 µM



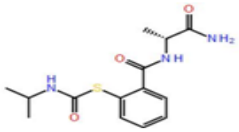
Title: 25
Ref #: 26
IC_{50}: 5.9 µM



Title: 26
Ref #: 5a
IC_{50}: 41.2 µM



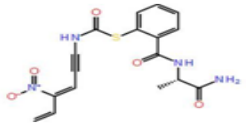
Title: 27
Ref #: 5b
IC_{50}: 39.4 µM



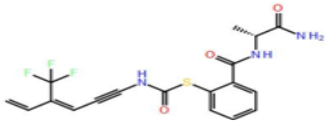
Title: 28
Ref #: 5c
IC_{50}: 38.3 µM

Title: 29
Ref #: 5d
IC_{50}: 40.2 µM

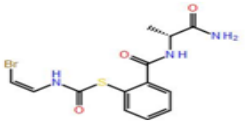
Title: 30
Ref #: 5e
IC_{50}: 76 µM



Title: 31
Ref #: 5f
IC_{50}: 55.1 µM



Title: 32
Ref #: 5g
IC_{50}: 125 µM

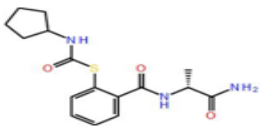


Title: 33
Ref #: 5h
IC_{50}: 142 µM

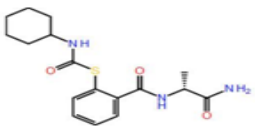
Title: 34
Ref #: 5i
IC_{50}: 17.3 µM

Title: 35
Ref #: 5j
IC_{50}: 11.9 µM

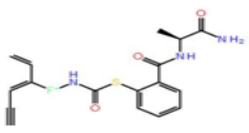
Title: 36
Ref #: 5k
IC_{50} : 130 µM



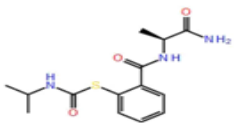
Title: 37
Ref #: 5l
IC_{50}: 23.1 µM

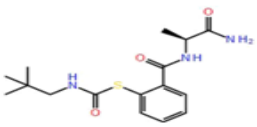


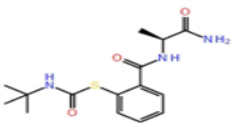
Title: 38
Ref #: 5m
IC_{50}: 128 µM



Title: 39
Ref #: 5n
IC_{50} : 4.9 µM







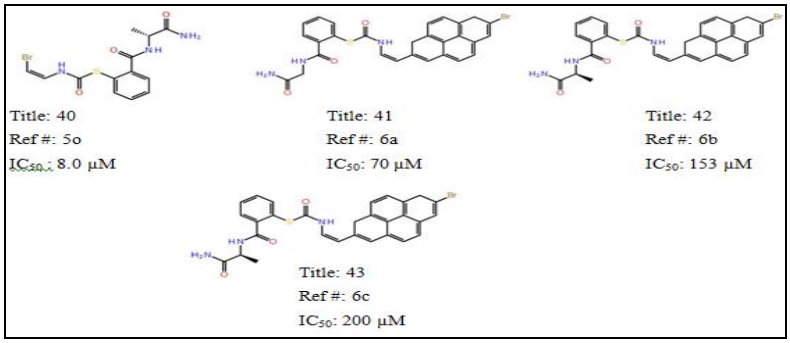


FIG. 2: THE 2D STRUCTURES OF 43 REFERENCE LIGANDS FOR PYRIDINIOALKANOYL THIOESTERS AND BENZAMIDE-BASED THIOLCARBAMATES DERIVATIVES DRAWN USING CHEMDRAW FREEWARE FROM LITERATURE SOURCE

The figure shows 2D structures of reference ligands for Pyridinioalkanoyl Thioesters and Benzamide-based Thiolcarbamates derivatives drawn using ChemDraw freeware from literature sources.

Ligand Set Taken for the Study: The study was carried out using a training set of 25 compounds to determine the factors required for the anti-HIV activities, which was measured by pIC_{50}.

Pharmacophore Model Generation and its Dataset: For the generation of an atom-based 3D-QSAR model, forty-three molecules were randomly divided into training (25) and test set (18) molecules. Eight most active molecules, twelve moderately active molecules, and less active molecules were included to spread out to assess the predictive accuracy of the model, and a set of 9 molecules was set as a test set. The training set molecules were aligned on the common pharmacophore hypotheses and analyzed by PLS with three factors. The correlation between predicted and observed activities of training and test set compounds is shown in **Fig. 3**. The top pharmacophore model was found to be associated with the four-point pharmacophore hypothesis containing two hydrophobic groups (H) and two aromatic rings (R) was denoted as HHRR. The distance and angles between the pharmacophoric features and their sites are reported in **Fig. 4.** Statistically significant 3D-QSAR models were predicted using the PLS factor based on the training set compounds by validating with test set compounds. The training set correlation is characterized by PLS factors (R^2= 0.901, SD= 0.6056, F= 18.2, P=0.0003115). The test set correlation is characterized by PLS factors (Q^2= 0.8563, RMSE= 0.3212, Pearson-R= 0.7033). The statistical results of PLS factors of atom-based 3D-QSAR are shown in **Table 2.**

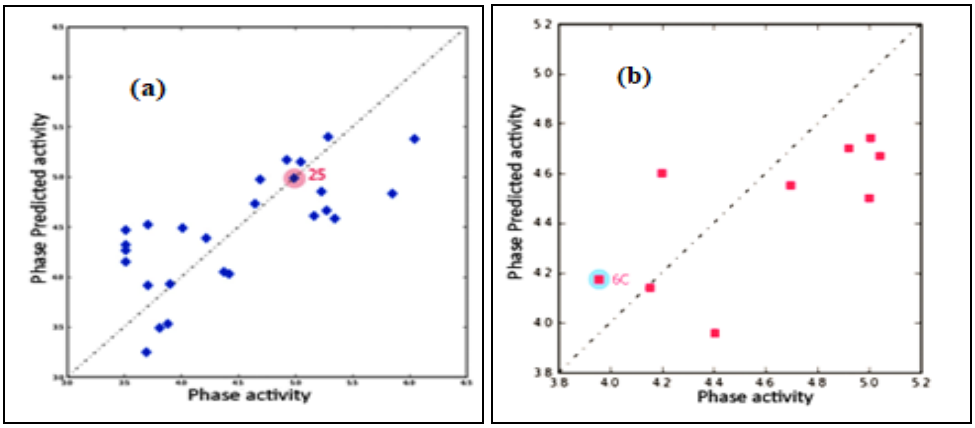


FIG. 3: GRAPH OF ACTUAL VERSUS PREDICTED IC_{50} OF TRAINING SET (A), AND TEST SET (B) USING AN ATOM-BASED 3D-QSAR MODEL FROM PHASE. THE ABOVE FIGURE SHOWS THE GRAPHICAL REPRESENTATION OF IC_{50} VERSUS THE PREDICTED IC_{50} OF THE TRAINING SET AND TEST SET USING AN ATOM-BASED 3D-QSAR MODEL FROM PHASE. AMONG THE 511 GENERATED HYPOTHESES ONLY 26 DIFFERENT HYPOTHESES WERE SELECTED BASED ON SURVIVAL SCORE AND BASED ON THE PHARMACOPHORE-BASED ALIGNMENT FOR THE TRAINING SET (A), AND 9 DIFFERENT HYPOTHESES FOR THE TEST SET (B)

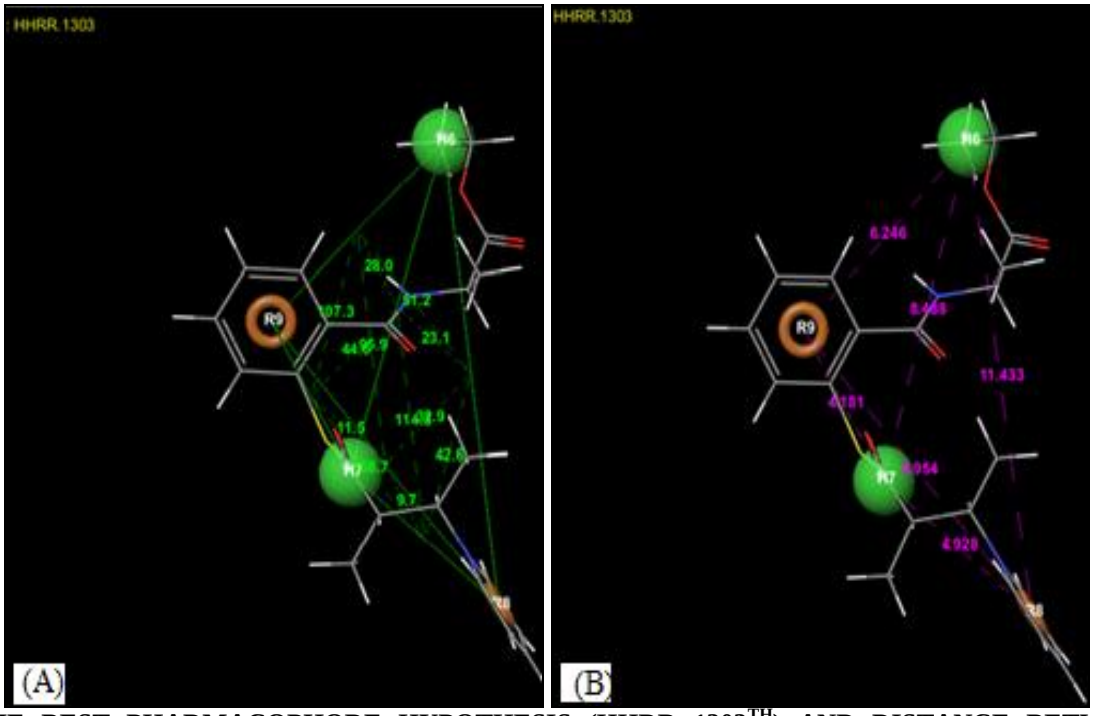


FIG. 4: THE BEST PHARMACOPHORE HYPOTHESIS (HHRR 1303TH) AND DISTANCE BETWEEN THE PHARMACOPHORIC SITES (A) AND ANGLES BETWEEN THE PHARMACOPHORIC SITES ARE IN Å UNIT (B). COLOR CODES FOR THE PHARMACOPHORIC FEATURES - GREEN SPHERE FOR HYDROPHOBIC (H); ORANGE RING FOR AROMATIC RING (R)

TABLE 2: PLS STATISTICAL PARAMETERS OF THE SELECTED QSAR MODEL

S. no.	Hypothesis	SD	R^2	F	P	RMSE	Q^2	Pearson-R
1	HHRR.1303	0.6056	0.901	18.2	0.0003115	0.3212	0.8563	0.7033
2	HHRR.1313	0.5361	0.5715	29.3	1.931e-005	0.3422	0.2695	0.5807
3	HHRR.1304	0.5904	0.4804	20.3	0.0001737	0.3693	0.1491	0.5827
4	HHRR.1517	0.4963	0.6328	37.9	3.379e-006	0.4392	-0.2034	0.1405
5	HHRR.1312	0.6318	0.4049	15	0.0008303	0.4686	-0.3705	0.1933

SD = standard deviation of the regression; R2 = correlation coefficient; F= variance ratio; P = significant level of variance ratio; RMSE= root-mean-square error; Q2= for the predictive activities; Pearson-R = correlation between the predicted and observed activity for the test set. Tabular representation of statistical results of PLS features of atom-based 3D-QSAR. For the best hypothesis HHRR, the result at PLS factor 3 was R^2 = 0.901, Q^2 = 0.8563, Pearson R = 0.7033, RSME = 0.3212 for statistically best model and this model can add an edge to the development of potential anti-HIV agents.

3D-QSAR Analysis: Additional insights into the inhibitory activity were gained by visualizing the 3D-QSAR model in the context of one or more ligands in the series with diverse activity.

The identification of important features is necessary for the interaction between ligands and the target protein.

The atom-based 3D-QSAR with the combined effect of hydrogen bond donor, hydrophobic/nonpolar, positive ionizable features, electron-withdrawing, and other features are depicted in **Fig. 5** and **6**. **Fig. 5** and **6** represents the cubes generated with the most active compound (compound 25) and least active compound (compound 6c) respectively.

The blue-colored region indicates the favorable features contributing to the ligand interactions with the target enzyme, while the red-colored region indicates the unfavorable region of the activity.

The blue cubes around the substitution of the R-group suggest that the R-group is significant in enhancing the activity, while some unfavorable region is indicated by the reference ligand, this red cube indicates very little activity when compared to the least active compound 6c.

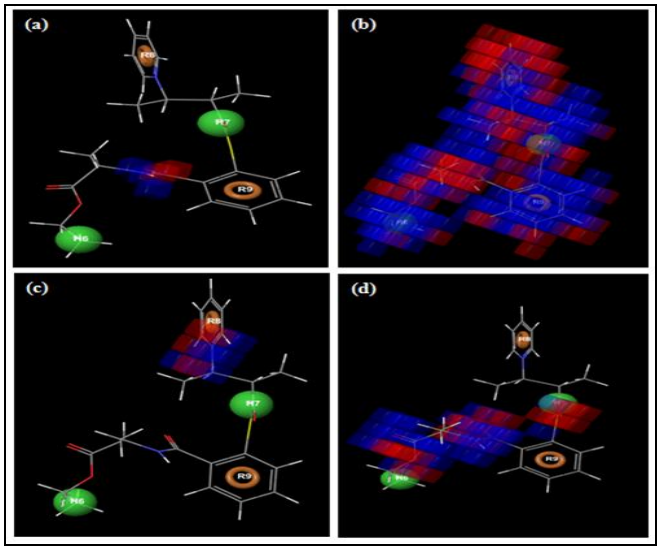


FIG. 5: PICTORIAL REPRESENTATION OF THE HYDROPHOBIC AND HYDROPHILIC REGIONS GENERATED FOR THE MOST ACTIVE MOLECULE FROM THE QSAR MODEL. BLUE CUBES DEPICTED FAVORABLE REGIONS, WHILE RED CUBES INDICATE UNFAVORABLE REGIONS FOR THE ACTIVITY OF THE MOLECULE. ATOM-BASED 3D-QSAR MODEL VISUALIZED IN THE CONTEXT OF HYDROGEN BOND DONOR FEATURES FOR MOST ACTIVE COMPOUND 25 (A) QSAR MODEL VISUALIZED IN THE CONTEXT OF THE FAVORABLE, AND UNFAVORABLE REGIONS OF THE HYDROPHOBIC FEATURES OF THE MOST ACTIVE COMPOUND 25 (B) QSAR MODEL VISUALIZED IN THE CONTEXT OF FAVORABLE AND UNFAVORABLE REGION POSITIVE IONIZABLE FEATURES OF THE MOST ACTIVE COMPOUND 25 (C) QSAR MODEL VISUALIZED IN THE CONTEXT OF FAVORABLE AND UNFAVORABLE REGION ELECTRON WITHDRAWING FEATURES OF THE MOST ACTIVE COMPOUND 25 (D)

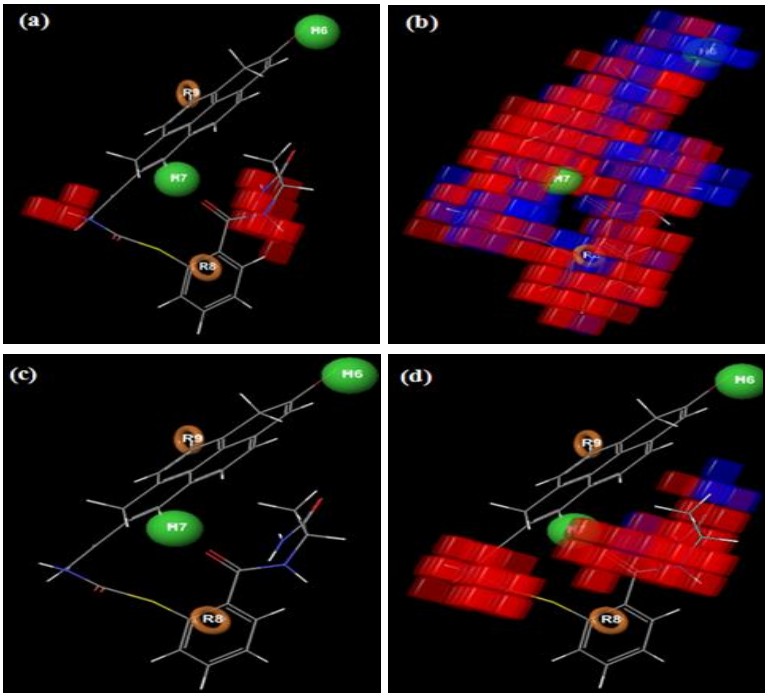


FIG. 6: PICTORIAL REPRESENTATION OF THE HYDROPHOBIC AND HYDROPHILIC REGIONS GENERATED FOR THE LEAST ACTIVE MOLECULE FROM THE QSAR MODEL. BLUE CUBES DEPICTED FAVORABLEREGIONS, WHILE RED CUBES INDICATE UNFAVORABLEREGIONS FOR THE ACTIVITY OF THE MOLECULE. ATOM-BASED 3D-QSAR MODEL VISUALIZED IN THE CONTEXT OF HYDROGEN BOND DONOR FEATURES FOR LEAST ACTIVE COMPOUND 6C (A) QSAR MODEL VISUALIZED IN THE CONTEXT OF FAVORABLE, AND UNFAVORABLEREGIONS OF THE HYDROPHOBIC FEATURES OF THE LEAST ACTIVE COMPOUND 6C (B) QSAR MODEL VISUALIZED IN THE CONTEXT OF FAVORABLE AND UNFAVORABLE REGION POSITIVE IONIZABLE FEATURES OF THE LEAST ACTIVE COMPOUND 6C (C) QSAR MODEL VISUALIZED IN THE CONTEXT OF FAVORABLE AND UNFAVORABLE REGION ELECTRON WITHDRAWING FEATURES OF THE LEAST ACTIVE COMPOUND 6C (D)

Docking Profile of Pharmacophore Screened Molecule: The compounds obtained from the ZINC and NCI database as a control drug interacted with the nucleocapsid protein domain p7 to compare the interaction profile **Table 3.** The 2D structure of ten lead compounds with their corresponding database ID is shown in **Fig. 7.**

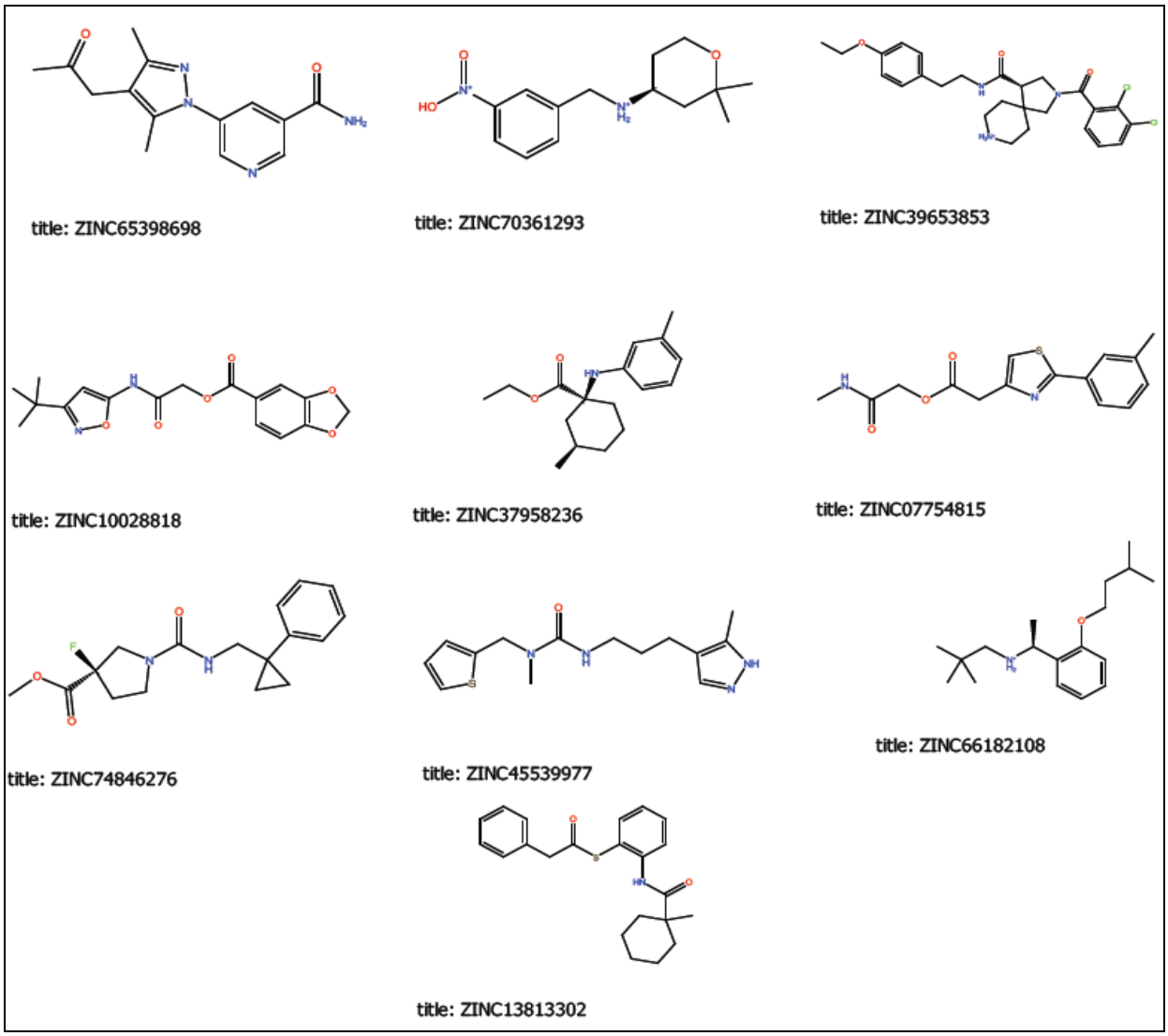


FIG. 7: THE 2D STRUCTURE OF TEN LEAD COMPOUNDS WITH THEIR CORRESPONDING DATABASE ID REPRESENTATION OF TEN LEAD COMPOUNDS WITH THEIR CORRESPONDING DATABASE ID. FINALLY, TEN LIGANDS WERE IDENTIFIED BASED ON THE GLIDE SCORE, WHICH INTERACTED WITH THE ACTIVE SITE OF THE NCP7 PROTEIN.

TABLE 3: GLIDE EXTRA PRECISION (XP) RESULTS FOR TEN LEAD MOLECULES

S. no.	Compound ID	Glide score	No. of Hydrogen Bonds	Interacting Residues	Distance (Å)	Hydrogen bond acceptor	Hydrogen bond donor
1	ZINC65398698	-11.51	4	LYS 34	2.45	A: LYS 34: (H) H	Ligand: (O) O2
				GLY 35	2.09	A: GLY 35: (H) H	Ligand: (O) O2
				TRP 37	2.03	A: TRP 37: (O) O	Ligand: (H) H8
				CYS 36	2.11	A: CYS 36: (O) O	Ligand: (H) H7
2	ZINC70361293	-8.42	2	ASN 17	2.11	A: ASN 17: (O) O	Ligand: (H) H11
				LYS 38	1.91	A: LYS 38 :(O) O	Ligand: (H) HO3
3	ZINC39653853	-8.34	3	CYS 18	2.21	A: CYS 18: (O) O	Ligand: (H) H15
				LYS 34	2.46	A: LYS 34: (H) H	Ligand: (O) O3
				GLY 35	2.09	A: GLY 35: (H) H	Ligand: (O) O3
4	ZINC10028818	-8.21	1	ASN 17	2.00	A: ASN 17: (O) O	Ligand: (H) H11

5	ZINC37958236	-7.24	2	CYS 18	2.23	A: CYS 18: (O) O	Ligand: (H) H14
				CYS 36	2.35	A: CYS 36: (O) O	Ligand: (H) H8
6	ZINC07754815	-7.16	2	ASN 17	2.09	A: ASN 17: (O) O	Ligand: (H) H13
				PHE 16	2.35	A: PHE 16: (O) O	Ligand: (H) H13
7	ZINC74846276	-7.06	2	LYS 14	2.17	A: LYS 14: (H) 3HZ	Ligand: (O) O2
				CYS 36	1.86	A:CYS 36: (O) O	Ligand: (H) H10
8	ZINC45539977	-6.91	1	PHE 16	2.10	A: PHE 16: (O) O	Ligand: (H) H12
9	ZINC66182108	-6.76	2	PHE 16	2.14	A: PHE 16: (O) O	Ligand: (H) H32
				CYS 36	1.79	A: CYS 36: (O) O	Ligand: (H) H20
10	ZINC13813302	-6.31	3	GLY 19	2.38	A: GLY 19: (O) O	Ligand: (H) H21
				PHE 16	2.37	A: PHE 16: (O) O	Ligand: (H) H15
				LYS 38	2.70	A: LYS 38: (O) O	Ligand: (H) H3

*Glide score (Kcal/mol); Number of hydrogen bond interactions; Interacting residues, Distance between the protein and ligand (Å); Hydrogen bond donor; Hydrogen bond acceptor. Tabular representation of Glide Score and number of Hydrogen bonds of ten NCp7 docked inhibitors out of 209 NCp7 Inhibitors. ZINC65398698 has having highest Glide score, selected as hit.

Interaction of Screened Ligands with NCP7 Domain: Totally 1000 compounds were subjected to interact with the NCP7 domain by incorporating the high throughput virtual screening (HTVS) method and Glide Docking panel with parameterization of Extra Precision value and Rigid Docking Algorithm. The resulting docking profile shows 10 feasible compounds that can act as to inhibit the function of Nucleocapsid Protein Domain 7. The Glide score of 10 compounds shows a minimum value of -6.31 and a maximum value of -11.51.

Among the docking results, below are the best complex structures with better G Score along with Protein and Ligand atom bonding **Fig. 8, 9, 10.**

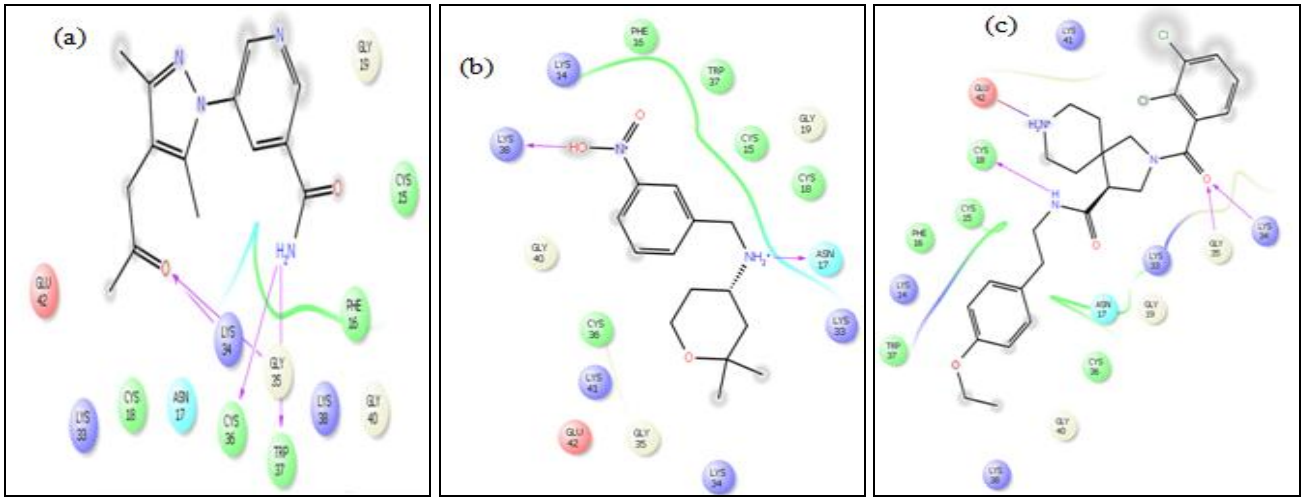


FIG. 8: BINDING POSES OF THE THREE LEAD COMPOUNDS WITH THEIR CORRESPONDING DATABASE ID

(A) Docked pose of NCp7 receptor with new molecule ZINC65398698. Ligand interaction diagram with pink arrows representing electrostatic interactions and green linesrepresenting π-π interactions. In this **Fig. 4** electrostatic interactions are observed. The docking simulation of ZINC65398698 into the binding site of the NCp7 was analyzed. The Glide score (-11.51 kcal/mol) was calculated as the highest interacted compound. A totalof four hydrogen bond interactions were formed. The oxygen atom of the ZINC65398698 well interacted with the backbone hydrogen atom of the LYS 34 (Hydrophilic), GLY 35 (Hydrophobic), and the oxygen atom of TRP 37 and CYS 36. The hydrogen bond distances were calculated (2.45Å, 2.09Å, 2.03Å, 2.11Å).

(B) Docked pose of NCp7 receptor with new molecule ZINC 70361293. The ligand interaction diagram with pink arrows represents the electrostatic interactions and the green line represents π-π interactions. In this **Fig. 2** electrostatic interactions were observed. The docking simulation of ZINC70361293within the active site of the NCp7 was analyzed with a glide score of -8.42 Kcal/mol.

Upon the examination of docking features between Hit ZINC 70361293 and NCp7, it was found to establish close contacts with NCp7 which included hydrogen bonding formed between the hit ZINC 70361293 compound and the active site region of NCp7. Totally two hydrogen bond interactions were formed between ZINC70361293 into the active site of the NCp7. The hydrogen atom of the hit molecule ZINC 70361293 strongly interacted with the backbone oxygen atom of the ASN 17, LYS 38 with bond length 2.11Å, 1.91Å.

(C) Docked pose of NCp7 receptor with new molecule ZINC 39653853. The ligand interaction diagram with pink arrows represents electrostatic interactions and the green line represents π-π interactions. In this figure, three electrostatic interactions are observed. The docking of ZINC 39653853 had a glide score of -8.34 kcal/mol. Upon examination of docking features between ZINC 39653853 and NCp7, it was found only three hydrogen bonds were formed between the ZINC 39653853 compound and the active site region of NCp7. The first one is the hydrogen atom of the ligand molecule well interacted with the backbone oxygen atom of the negatively charged residues CYS 18, and the second one is the oxygen atom of the ligand molecule was nicely bonded with the backbone hydrogen atom of the positively charged residue of LYS 34 and the third one is the oxygen atom of the ligand molecule were tightly interacted with the backbone hydrogen atom of the positively charged residue of GLY 35. The hydrogen bond length was noted (2.21Å, 2.46 Å, 2.09Å).

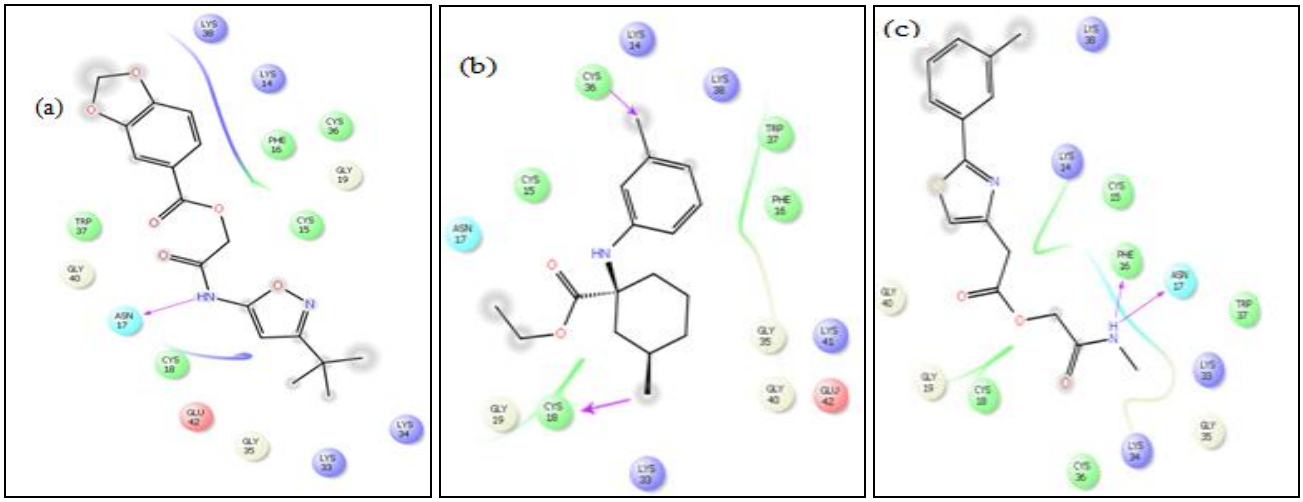


FIG. 9: BINDING POSES OF THE THREE LEAD COMPOUNDS WITH THEIR CORRESPONDING DATABASE ID

(A) Docked pose of NCp7 receptor with new molecule ZINC10028818. The ligand interaction diagram with pink arrows represents electrostatic interactions and the green line represents π-π interactions. In this figure, one electrostatic interaction was observed. The binding conformation of ZINC10028818 within the active site of NCp7 had a Glide score of -8.21 kcal/mol. Hit ZINC10028818 formed one hydrogen bond with the active site of the NCp7 residues like ASN 17 (2.00Å). The hydrogen atom of the ligand molecule ZINC10028818 nicely interacted with the backbone oxygen atom of the ASN 17. The hydrogen atom of the ligand molecule acts as a hydrogen bond donor, and the oxygen atom of the protein molecule NCp7 acts as a hydrogen bond acceptor. (B) Docked pose of NCp7 receptor with new molecule ZINC37958236. The ligand interaction diagram with pink arrows represents electrostatic interactions and the green line represents π-π interactions. In this figure, two electrostatic interactions are observed. Hit ZINC37958236 had a Glide score of -7.24 kcal/mol. Upon the examination of docking features between Hit ZINC37958236 and NCp7, it was found only two hydrogen bond interactions were formed between the hit compound and the active site region of NCp7. The first one is the hydrogen atom of the hit compound ZINC37958236well interacted with the backbone oxygen atom of the negatively charged residues of the CYS 18, second interaction was the side chain oxygen atom of the negatively charged residue of CYS 36 strongly interacted with hydrogen atom of the hit compound ZINC37958236. The hydrogen atom of the hit compound acts as a hydrogen bond

donor, and the oxygen atom of the NCp7 protein molecule acts as a hydrogen bond acceptor. The bond lengths were observed (2.23Å and 2.35Å).

(C) Docked pose of NCp7 receptor with new molecule ZINC07754815. The ligand interaction diagram with pink arrows represents electrostatic interactions and the green line represents π-π interactions. In this figure, two electrostatic interactions are observed. The binding mode of the hit molecule ZINC07754815 within the binding pocket of the NCp7 was analyzedby Glide score (-7.16 kcal/mol). Upon the examination of docking features between Hit ZINC07754815and the binding site of the NCp7, it was found hydrogen bonds were formed between the hit compound and the active site region of NCp7.

First interactions were formed with the positively charged residue of ASN 17 and hydrogenatom of the hit molecule. The side chain oxygen atom of the positively charged residue of PHE 16 were well interacted with the hydrogen atom of the ZINC07754815. The hydrogen bond distance between the protein-ligand was observed (2.09Å, 2.35Å).

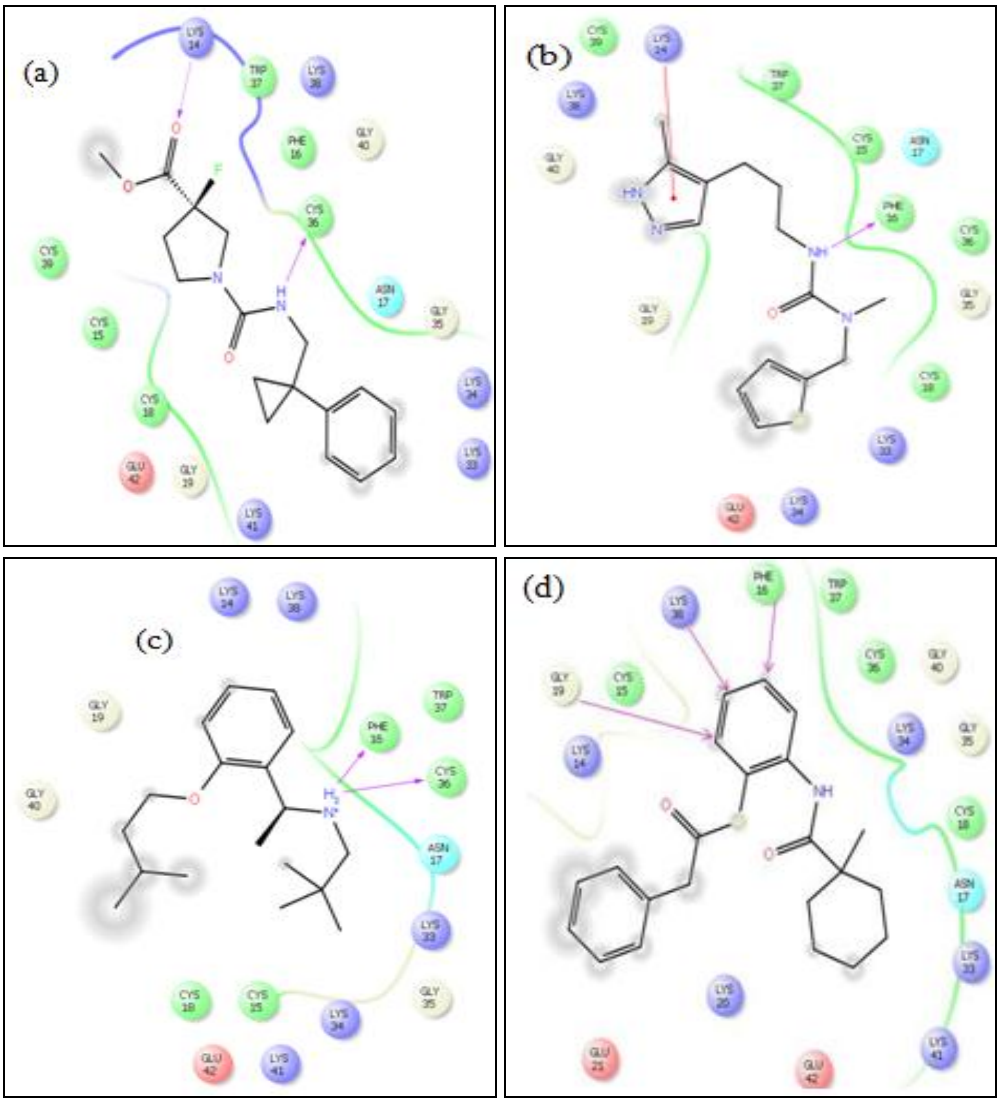


FIG. 10: BINDING POSES OF THE FOUR LEAD COMPOUNDS WITH THEIR CORRESPONDING DATABASE ID

(A) Docked pose of NCp7 receptor with new molecule ZINC74846276. The ligand interaction diagram with pink arrows represents electrostatic interactions and the green line represents π-π interactions. In this figure, two electrostatic interactions were observed. The binding mode of the ZINC74846276 within the active site of the NCp7 was visualized. Upon the examination of docking features between ZINC74846276 and NCp7, it was found hydrogen bonding interactions were formed between the hit compound and the active site region of NCp7. The backbone hydrogen

atom of the positively charged residues of LYS 14 was well interacted with the oxygen atom of the ZINC74846276, and then the hydrogen atom of the ZINC74846276tightly interacted with the side chain oxygen atom of the hydrophobic residue of CYS 36 with bond distance (2.17Å, 1.86Å).

(B) Docked pose of NCp7 receptor with new molecule ZINC45539977. The ligand interaction diagram with pink arrows represents electrostatic interactions and the green line represents π-π interactions. In this figure, one electrostatic interaction was observed. The docking simulation of ZINC45539977 had a glide score of -6.91 kcal/mol and Upon the examination of docking features between ZINC45539977 and NCp7, it was found only one hydrogen bond was formed between the ZINC45539977 compound and active site region of NCp7. The hydrogen atom of the ligand molecule was nicely bonded with the backbone oxygen atom of the negatively charged residue of PHE 16. The hydrogen bond length was noted as 2.10Å.

(C) Docked pose of NCp7 receptor with new molecule ZINC66182108. The ligand interaction diagram with pink arrows represents electrostatic interactions and the green line represents π-π interactions. In this figure, two electrostatic interactions were observed. The binding mode of the ZINC66182108 within the active site of the NCp7 was visualized. There are two hydrogen bond interactions formed between ZINC66182108 and NCp7. The hydrogen atom of the ligand molecule tightly interacted with the backbone oxygen atom of the negatively charged residues of PHE 16, and then the hydrogen atom of the hit molecule interacted with the backbone oxygen atom of the hydrophobic residues of CYS 36. The bond lengths were noted as 2.14Å, 1.79Å correspondingly.

(D) Docked pose of NCp7 receptor with new molecule ZINC13813302. The ligand interaction diagram with pink arrows represents electrostatic interactions and the green line represents π-π interactions. In this figure, three electrostatic interactions were observed. The binding mode of the hit molecule ZINC13813302 within the binding pocket of the NCp7 wasanalyzed with, a Glide score (-6.31 kcal/mol). Upon the examination of docking features between ZINC13813302 and NCp7, hydrogen bonding interactions were observed between the hit compound and the active site region of NCp7. It was found that the hydrogen atom of the ligand formed hydrogen bonds with the oxygen atoms of GLY 19, PHE 16, and LYS 38 respectively with hydrogen bond distances of 2.38Å, 2.37Å, and 2.70Å between protein residues and ligand atoms.

Pharmacodynamic and Kinetic Study on the Identified Lead Molecules: The ADME properties of the ten newly identified lead molecules were assessed using the QikProp tool of the Schrödinger suite. The above-mentioned ten lead molecules satisfied drug-like properties based on Lipinski's rule of five. The molecular weight of the lead molecules was found to be less than 500Da, the number of hydrogen bond donors less than 5, and hydrogen bond acceptors are less than 10, and the predicted octanol/water partition coefficient (QPlogPo/w) is less than 5. Then these lead compounds were further evaluated for their drug-like behavior through analysis of pharmacokinetic parameters required for absorption, distribution, metabolism, excretion, and toxicity (ADMET) by use of QikProp. For the ten lead molecules, the aqueous solubility (QPlogS) critical for estimation of absorption and distribution of drug within the body ranged between -5.010 to -0.485 respectively. The predicted IC_{50} value for the blockage of HERG K+ channels is in the acceptable range below -5. The predicted value of binding to human serum albumin (QPksha) lies in the acceptable range of ~ -0.397 to -2.661. The predicted Brain/Blood barriers are under the acceptable range of ~ -5.050 to -1.791. All the pharmacokinetics parameters fit well with the acceptable range defined for use in humans. The results of the Pharmacodynamic and kinetic study on the identified lead molecules are listed in **Table 4.**

TABLE 4: ADME PROPERTIES OF THE LEAD MOLECULES AS VERIFIED BY USING QIKPROP

Ligand ID	QPlogS	QPlogKhsa	LogBB	MW	HBD	HBA	QPlog(o/w)	HERG
ZINC65398698	-3.240	-0.313	-1.316	272.306	2.000	7.000	0.880	-4.470
ZINC70361293	-2.810	0.259	-0.258	264.324	1.000	3.250	2.417	-5.277

ZINC39653853	-5.780	0.471	-0.225	504.455	2.000	7.750	4.186	-5.804
ZINC10028818	-3.852	-0.181	-1.030	346.339	1.000	7.500	2.072	-5.165
ZINC37958236	-4.389	0.473	0.037	275.390	1.000	3.000	3.897	-4.125
ZINC07754815	-3.809	-0.264	-0.768	304.363	1.000	6.000	2.176	-4.396
ZINC74846276	-4.846	0.319	-0.378	320.363	1.000	4.000	3.471	-3.900
ZINC45539977	-4.460	0.102	-0.756	292.398	2.000	3.000	3.128	-4.038
ZINC66182108	-4.322	0.825	0.469	277.449	1.000	2.250	4.822	-5.782
ZINC13813302	-5.464	0.682	-0.028	367.505	1.000	4.500	5.014	-5.696

Ligand IDs are of the zinc database. Predicted aqueous solubility; S in mol/L (acceptable range; -6.5 to 0.5); Prediction of binding to human serum albumin (acceptable range; -1.5 to 1.5); Prediction of brain/blood (acceptable range; -3.0 to 1.2); Molecular weight (<500Da); Hydrogen bond donor (<5); Hydrogen bond acceptor (<10); Predicted octanol /water partition coefficient log p (acceptable range; -2.0 to 6.5); Predicted IC_{50} value for blockage of HERG K+ channels (acceptable range: below. -5.0). Tabular representation of ten newly identified leads assessed by the use of the Qikprop tool of the Schrödinger suite. The above-mentioned ten lead molecules satisfy drug-like properties based on Lipinski's rule of five.

Molecular Dynamic Simulation of Best G-score Docked Complex: Molecular dynamics of the protein-ligand complex were performed by applying Berenson's Temperature (301K) coupling method and the Parrinello-Rahman Pressure coupling method. Five complexes simulated in the water environment have the G Score of -11.51, -8.42, -8.34, -8.21, and -7.24 respectively. A best-fit score of -11.51 was simulated for 20 ns whereas other G-scored complexes were done for a 10 ns equilibration run. The following observables were determined as the density of the system was constant at around 1000(kg/m^3) which shows the system is well stabilized inside a solvated cubic box. The temperature of the system was stable throughout the equilibration run at 310K with a thermal fluctuation of ±5K. The kinetic energy of the system was found to be stable at 28 J/mol with a mean fluctuation of ± 5 J/mol. The applied pressure to the system was observed to be fluctuating around ±500 bar however on average is stable.

NCp7 Protein-ligand Complex Stability: Molecular dynamics simulation for 20 ns, of NCp7 protein-ligand complex ZINC65398698 (5-[3,5-dimethyl-4-(2-oxopropyl) -1H-pyrazol-1-yl] pyridine-3-carboxamide) having high binding affinity was performed [43]. To make comparative values of the stability further consequent G Scored complexes were simulated. The stability of the NCp7 protein after it binds to the ligand was checked by analyzing the trajectories obtained after the Molecular Dynamics Simulation.

Root Mean Squared Deviation (RMSD): The complex stability was determined by calculating root mean squared deviation (RMSD), and root mean squared fluctuation (RMSF) of the complex. The RMSD analysis predicts the stability of protein and its structural variation while evolving with time. The RMSD of the protein-ligand complex for the trajectories written for a 20 ns production run was analyzed to identify the stability of the system at each time interval. It was observed that the complex is stable within the early 10 ns production run **Fig. 11**. Between 1.5 to 12.5 ns, the complex was observed to fluctuate from equilibrium with minor deviation of 0.1 nm, thereafter i.e. from 14.1 ns to 20 ns the system again attained equilibrium. On average, the system was observed to be stable with an RMSD of 0.3 Å. The system achieves equilibrium in the early 100-150 ps and remains stable thereafter for a 20 ns simulation period. The complex stability shows that there is no major structural variation in protein after binding with ligand whereas ZINC10028818, ZINC37958236, ZINC70361293, and ZINC39653853 **Fig. 12** ID with protein complexes are showing simulation within 0.1 to 0.25 Å in 10 ns itself so the best-fit score of ZINC65398698 complex are showing best complex.

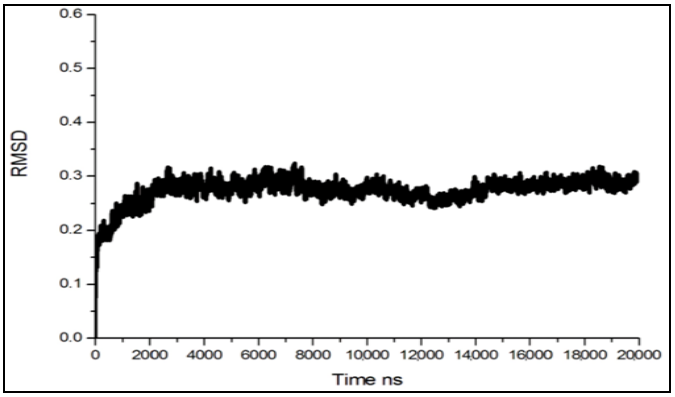


FIG. 11: RMSD PLOT OF PROTEIN-LIGAND COMPLEX IN 20 NS MD SIMULATION RUN

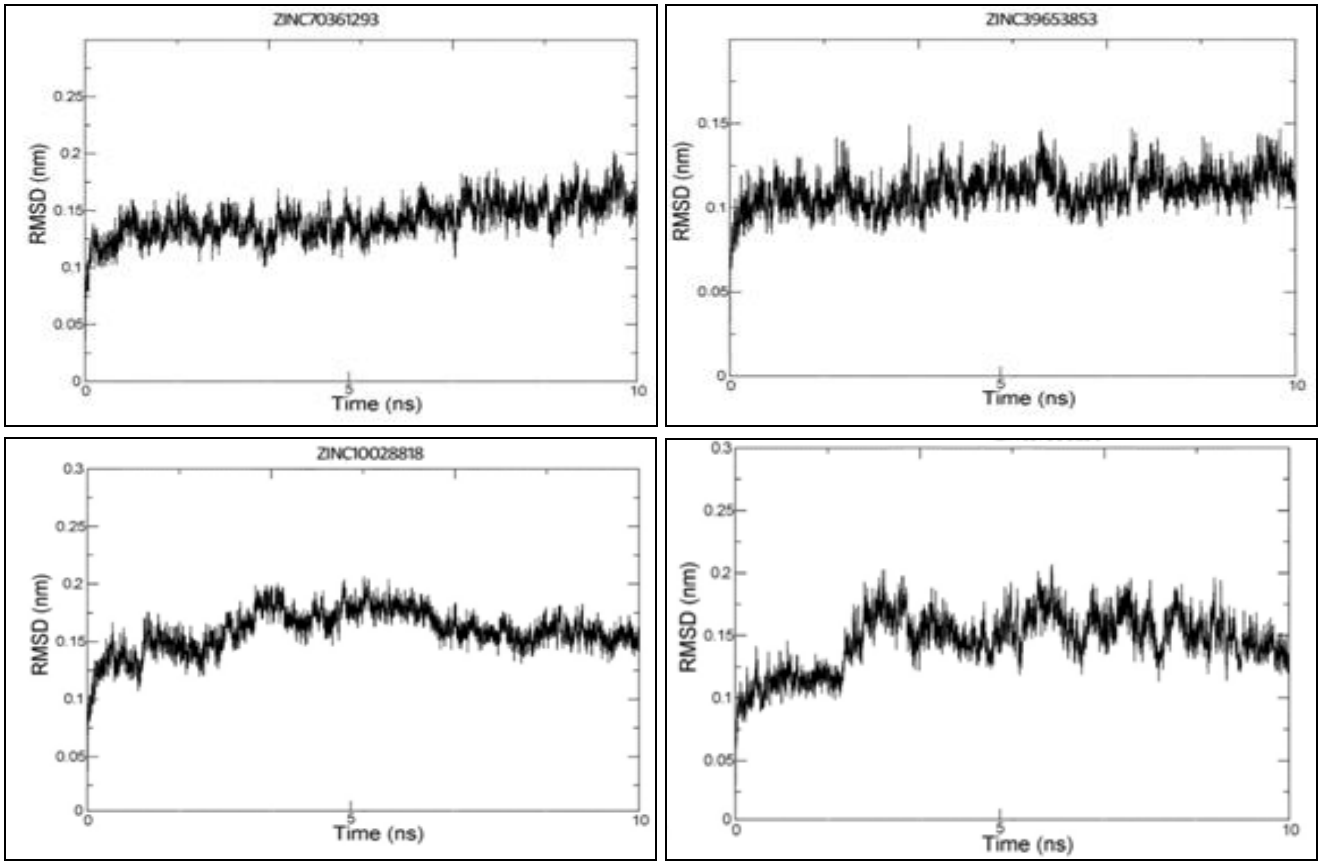


FIG. 12: RMSD PLOT OF PROTEIN-LIGAND COMPLEX IN 10 NS MD SIMULATION RUN

Root Mean Square Fluctuation (RMSF): The RMSF plot depicted in **Fig. 13** presents the protein backbone flexibility after forming a complex with ligand, ZINC65398698 (5-[3,5-dimethyl-4-(2-oxopropyl) - 1H – pyrazol – 1 - yl] pyridine-3-carboxamide). The RMSF plot showed to what extent each residue of NCp7 protein fluctuated during the production run. It was observed that the protein residues, Asp (N), Ala (A), Lys (K), Met (M), and Glu (Q) were found to be more flexible with RMSF of 0.5 nm, 0.1 nm, 0.12 nm, 0.14 nm, and 0.15 nm respectively which implies that mentioned residues had greater movement from its native position during dynamics simulation. The remaining residues were observed to be flexible with an average RMSF of 0.5-0.17 nm which shows that these residues had constrained flexibility from their mean position and, hence were rigid and possessed limited movement during dynamics simulation whereas the other four complexes showed fluctuation in between 0.5 to 0.35 which has maximum simulation in intermolecular interaction **Fig. 14.**

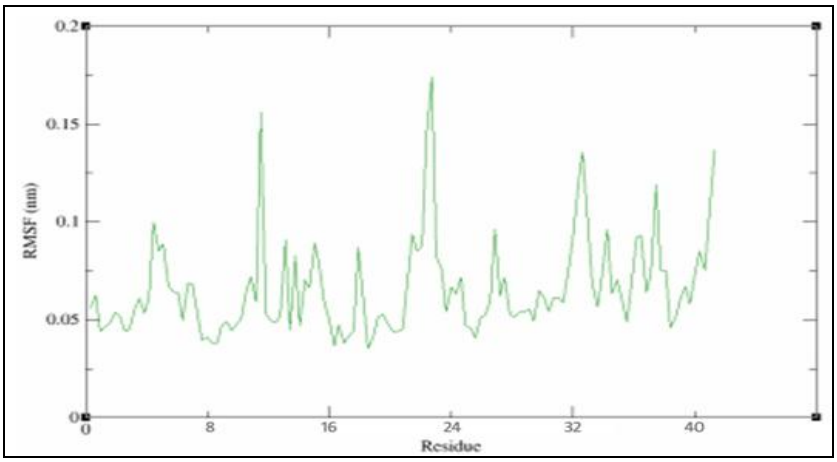


FIG. 13: RMSF PLOT OF PROTEIN-LIGAND COMPLEX IN 20 NS MD SIMULATION RUN

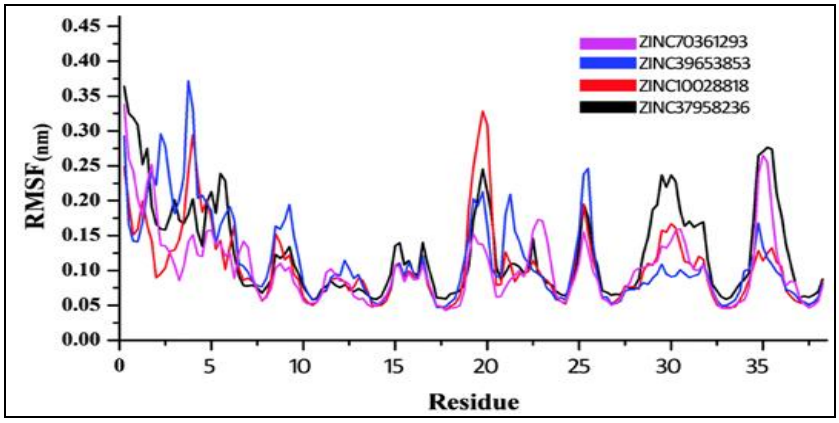


FIG. 14: RMSF PLOT OF PROTEIN-LIGAND COMPLEX IN 10 NS MD SIMULATION RUN

Hydrogen Bond Interaction: ZINC65398698 complex **Fig. 15** has having total number of hydrogen bonds within the complex is 24. Whereas the other four **Fig. 16** complexes have a maximum of 12 hydrogen bonds with a complex system of MD simulation for 20 ns and 10 ns equilibrium productivity run.

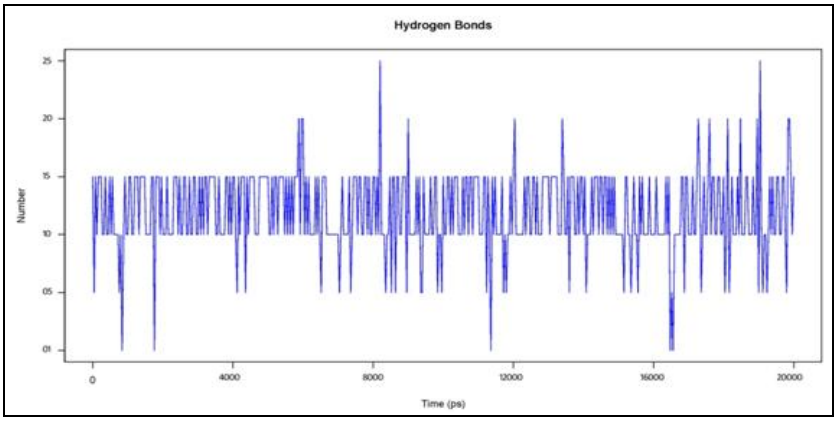


FIG. 15: HYDROGEN BOND INTERACTIONS OF THE PROTEIN-LIGAND COMPLEX IN 20 NS MD SIMULATION RUN

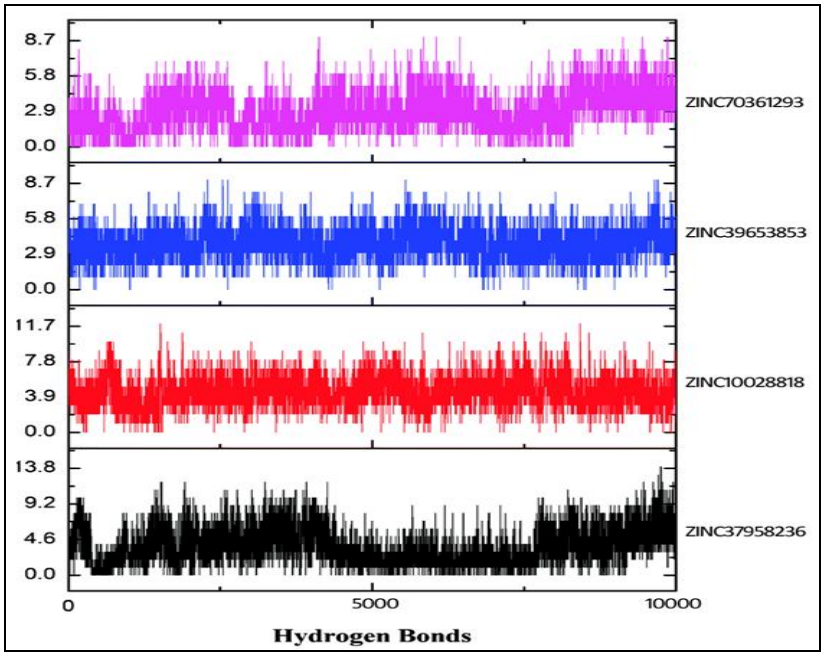


FIG. 16: HYDROGEN BOND INTERACTIONS OF THE PROTEIN-LIGAND COMPLEX IN 10 NS MD SIMULATION RUN

The overall objective of this work is to identify potential drug targets for the HIV-I sub-type which is infecting millions of human beings worldwide. We have adopted Nucleocapsid Protein Domain p7 as a target as it has a critical role in HIV-I replication that can inactivate the virus infectious process.

A challenge in structure-based virtual screening is the assessment and ranking of predicted ligand conformations because the scoring functions implemented in the Maestro v9.3 include assumptions simplifying the complexity of molecular recognition. The core structure of the nucleocapsid p7 domain was docked with the compounds obtained from ZINC and NCI. We generated pharmacophore models for the HIV-1 NCp7 and screened commercial databases to find possible lead compounds for our drug discovery project.

We identified compounds that interacted best at the active binding site of NCp7 in terms of pharmacophore mapping and docking using two different scoring functions, showing very promising inhibitory activity. Altogether, we have shown that pharmacophore modeling, database screening, and prediction of the compounds with docking resulted in truly promising compounds for the drug-designing process in the future.

CONCLUSION: In conclusion, a ligand-based computational approach was used to identify molecular structural features required for an effective NCp7 inhibitor to discover drugs to prevent and cure varieties of HIV. Based on 25 training set compounds, a highly predictive pharmacophore model was generated which consists of two hydrophobic groups (H) and two aromatic rings (R). The pharmacophore model generated was used for three-dimensional queries in database searches for new potent molecules and to design new molecules and forecast their inhibitory activity for NCp7 quantitatively before undertaking any synthesis. Thus, our pharmacophore model helps identify novel lead compounds with improved NCp7 inhibitory activity through 3D database searches.

ACKNOWLEDGEMENTS: The Authors thank the DBT Bioinformatics Centre, Bharathiar University for providing facilities.

CONFLICTS OF INTEREST: Nil

REFERENCES:

1. Zayats R, Murooka T and McKinnon L: HPV and the Risk of HIV Acquisition in Women. Frontiers in Cellular and Infection Microbiology 2022; 12: 814948.
2. Mei J and Zhao J: Prediction of HIV-1 and HIV-2 proteins by using Chou's pseudo amino acid compositions and different classifiers. Scientific Reports 2018; 8: 2359.
3. Zautner A, Herchenroder O, Moussi A, Schwarz, N, Wiemer D, Groß U & Frickmann H: Pharmaceutical interactions between antiretroviral and antimalarial drugs used in chemoprophylaxis. Actatropica 2018; 179: 25-35.
4. Haleshappa R, Deepak Koppaka, Lakshmaiah K, Govind K, Dasappa L, Premalata C, Jacob L, Suresh B, Sumathi B, Lokesh K, Rajeev L, Saldanha S, Abhishek A, Mulchandani N and Vikas Asati: Patterns of malignancies in patients with HIV-AIDS: a single institution observational study. The Journal of Community and Supportive Oncology 2018; 16: 188-192.
5. Ravindra and Patil S: A Study of Socio-Demography and Health Condition of HIV Positive People in Dharwad District. IAHRW International Journal of Social Sciences Review 2019; 7: 259-264.
6. Carnes S, Sheehan J & Aiken C: Inhibitors of the HIV-1 capsid, a target of opportunity. Current Opinion in HIV and AIDS, 2018; 13: 359-365.
7. Dalod M and Brooks A: Broadening our knowledge of the differences between HIV-2 and HIV-1 innate sensing: HIV-2 infected cells selectively express B7-H6 and engage NKp30 on natural killer cells. AIDS 2019; 33: 153-154.
8. Esbjornsson J, Jansson M, Jespersen S, Mansson F, Honge B, Lindman J, Medina C, Silva Z, Norrgren H, Medstrand P, Rowland-Jones S and Wejse C: HIV-2 as a model to identify a functional HIV cure. AIDS Research and Therapy 2019; 16: 24.
9. UNAIDS 'AIDSinfo'. Accessed September 2018.
10. NACO 'Annual Report 2017-2018'.
11. Guo F, Saadatmand J, Niu M, Kleiman L: Roles of Gag and NCp7 in Facilitating Annealing to Viral RNA in Human Immunodeficiency Virus Type 1. Journal of Virology 2009; 83: 8099-8107.
12. Jiang K, Humbert N, KK S, Rouzina I, Mely Y & Westerlund F: The HIV-1 nucleocapsid chaperone protein forms locally compacted globules on long double-stranded DNA. Nucleic Acids Research 2021; 49: 4550-4563.
13. Summers MF, Henderson LE, Chance MR, Bess JW, South TL and Blake PR: Nucleocapsid zinc fingers detected in retroviruses: EXAFS studies of intact viruses and the solution structures of the nucleocapsid protein from HIV-1. Protein Science 1992; 1: 563-574.
14. Song Y, Goel A, Basrur V, Roberts PEA, Mikovits JA and Inman JK: Bioorganic & Medicinal Chemistry 2002; 10: 1263-1273.
15. Goel A, Sharlyn JM, Rasem JF, Tracy LH, Turpin JA and Huang M: Benzamide-Based Thiolcarbamates: A New Class of HIV-1 NCp7 Inhibitors. Bioorganic & Medicinal Chemistry Letters 2002; 12: 767-770.
16. Meier J & Schnermann M: Pharmacology by Chemical Biology. Molecular Pharmaceutics 2018; 5: 703-704.
17. Kausar S and Falcao A: An automated framework for QSAR model building. Journal of Cheminformatics 2018; 10: 1.
18. Slavov S and Beger R: Quantitative structure–toxicity relationships in translational toxicology. Current Opinion in Tolxicology 2020; 46-49.

19. Moroy G, Martiny VY, Vayer P, Villoutreix BO and Miteva MA: Toward *in-silico* structure-based ADMET prediction in drug discovery. Drug Discovery Today 2012; 17: 44-55.
20. Cordes H and Rapp H: Gene expression databases for physiologically based pharmacokinetic modeling of humans and animal species. CPT: Pharmacometrics & Systems Pharmacology 2023; 12: 311-319.
21. Basith S, Cui M, Macalino S & Choi S: Expediting the Design, Discovery and Development of Anticancer Drugs using Computational Approaches. Current medicinal chemistry 2017; 24: 4753-4778.
22. Mouchlis V, Afantitis A, Serra A, Fratello M, Papadiamantis A, Aidinis V, Lynch I, Greco D and Melagraki G: Advances in De Novo Drug Design: From Conventional to Machine Learning Methods. International Journal of Molecular Sciences 2021; 22: 1676.
23. Haghshenas, H, Kaviani B, Firouzeh M & Tavakol H: Developing a variation of 3D-QSAR/MD method in drug design. Journal of Computational Chemistry 2021; 42: 917- 929.
24. Ragno R, Esposito V, Mario M, Masiello S, Viscovo M & Cramer R. Teaching and Learning Computational Drug Design: Student Investigations of 3D Quantitative Structure–Activity Relationships through Web Applications. Jouranl of Chemical Education 2020; 97: 1922-1930.
25. Gu C, Zhang Y, XieX, Wang X, Zhao B & Zhu Y: Pharmacophore modeling, atom-based 3D-QSAR and molecular docking studies on N-benzylpyrimidin-4-amine derivatives as VCP/p97 inhibitors. Medicinal Chemistry Research 2020; 29: 727-737.
26. Wan S, Sinclair R and Coveney P: Uncertainty quantification in classical molecular dynamics. Philosophical transactions. Series A, Mathematical, Physical, and Engineering Sciences 2020; 379: 20200082.
27. Peng S, Zhang X, Su W, Dong D, Lu Y, Liao X, Lu K, Yang C, Liu J, Zhu W and Wei D: High-Scalable Collaborated Parallel Framework for Large-Scale Molecular Dynamic Simulation on Tianhe-2 Supercomputer. IEEE/ACM Transactions on Computational Biology and Bioinformatics 2020; 17: 804-816.
28. Stefaniu A: Introductory Chapter: Molecular Docking and Molecular Dynamics Techniques to Achieve Rational Drug Design. Molecular Docking and Molecular Dynamics 2019.
29. Novitasari D: Pelatihan Pemanfaatan Software ChemDrawsebagai Media Pembelajaran Materi Hidrokarbon Bagi Guru di SMK N 1 BP. BR OKU Timur. Jurnal Indonesia Mengabdi 2022; 3: 10-15.
30. Ghazi Y, Mahdi M and Dawood A: Theoretical drug design, molecular docking and adme study of new 1,3,4-oxadiazole derivatives: promising anticancer agents against both breast and lung cancers. Egyptian Journal of Chemistry 2021; 64: 6269-6283.
31. Abdullahi M, Uzairu A, Shallangwa G, Arthur D, Umar B and Ibrahim M: Virtual molecular docking study of some novel carboxamide series as new anti-tubercular agents. European Journal of Chemistry 2020; 11: 30-36.
32. Ligprep, Version 2.7, Schrodinger, LLC, New York, NY, 2013.
33. Ramakrishnan P, Kumar T, Saraswathy G and Sujatha S: *In-silico* evaluation of drugs used in treatment oral lichen planus. Journal of oral pathology & medicine: Official publication of the International Association of Oral Pathologists and the American Academy of Oral Pathology 2020; 49: 926-932.
34. Krishnaprasad B, Maity S, Mehta C, Suresh A, Nayak U and Nayak Y: *In-silico* Drug Repurposing of Penicillins to Target Main Protease Mpro of SARS-CoV-2. Pharmaceutical Sciences 2020; 26: 52-62.
35. Prajapat M, Sarma P, Shekhar N, Chauhan A, Kaur G, Bhattacharyya A, Avti P, Choudhary G, Bansal S, Sharma S, Kaur H, Kumar S, Mann H, Raja A, Singh A, Singh R, Sharma A, Prakash A and Medhi B: Virtual screening and molecular dynamics simulation study of approved drugs as a binder to the linoleic acid binding site on spike protein of SARS-CoV-2 and double mutant (E484Q and L452R). Indian Journal of Pharmacology 2022; 54: 431 -442.
36. Sorokina M and Steinbeck C: Review on natural products databases: where to find data in 2020. Journal of Cheminfomatics 2020; 12: 20.
37. Phase, version 3.6. Schrödinger, New York, NY, 2013.
38. Voigt JH, Bienfait B, Wang S, Nicklaus MC: Comparison of the NCI open database with seven large chemical structural databases. Journal of Chemical Information and Computer Sciences 2001; 41: 702-712.
39. David Ryan Koes, Carlos J. Camacho: ZINCPharmer: pharmacophore search of the ZINC database. Nucleic Acids Research 2012; 40: 409-414.
40. Hess B, Kutzner C, van der Spoel D and Lindahl E: GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. Journal of Chemical Theory and Computation 2008; 4: 435-447.
41. Schuttelkopf AW, van Aalten DMF: "PRODRG: a tool for high-throughput crystallography of protein-ligand complexes". Acta Crystallographica 2004; D60:1355-1363.
42. Protein preparation Wizard Maestro. New York: Schrodinger LLC. Qikprop, Version 3.7, LLC, New York, NY, 2013.
43. Hayik SA, Dunbrack R and Merz KM: A Mixed QM/ MM Scoring Function to Predict Protein-Ligand Binding Affinity. Journal of Chemical Theory and Computation 2010; 6: 3079-3091.

How to cite this article:
Lisina KV and Piramanayagam S: Identification of novel NCP7 inhibitors for HIV drug discovery using ligand-based computational approaches. Int J Pharm Sci & Res 2024; 15(6): 1682-02. doi: 10.13040/IJPSR.0975-8232.15(6).1682-02.