In-silico docking studies of Plant Derived Natural Compounds to identify potential target for the treatment of Ebola Virus

To be submitted as Major Project Report in partial fulfillment of the requirement for the Degree of

Master of Technology

In

Bioinformatics

Submitted by

Kamlendra Gupta

(2K14/BIO/06)

Delhi Technological University, Delhi, India

Under the supervision of

Dr. Navneeta Bharadvaja

Assistant Professor

Department of Biotechnology
Delhi Technological University
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Bawana Road Delhi 110042

.

CERTIFICATE

This is to certify that the dissertation entitled ***In-silico*** **docking studies of Plant Derived Natural Compounds to identify potential target for the treatment of Ebola Virus (2k14/bio/06)** in the partial fulfillment of the requirements for the award of the degree of Masters of Engineering, Delhi Technological University (Formerly Delhi College of Engineering, University of Delhi), is an authentic record of the candidate's own work carried out by him under my guidance. The information and data enclosed in this thesis is original and has not been submitted elsewhere for honoring of any other degree.

Dr. Navneeta Bharadvaja
(Project Mentor)
Department of Bio-Technology
Delhi Technological University
Delhi-110042

Prof. D. Kumar
Head of Department
Department of Bio-Technology
Delhi Technological University
Delhi -110042

.

DECLARATION

I declare that my major project entitled **"*In-silico* docking studies of Plant Derived Natural Compounds to identify potential target for the treatment of Ebola Virus"**, submitted to Department of Biotechnology, Delhi Technological University as a result of the work carried out by me at "Plant Biotechnology Laboratory", Department of Biotechnology, as Major project.

Date: Kamlendra Gupta

.

ACKNOWLEDGEMENT

I am indebted to Prof. D. Kumar (Head of Department), Department of Biotechnology, Delhi Technological University, for providing me an opportunity to explore my intellect and giving me a chance to carry out my work. I thank him with all my heart for his guidance and valuable suggestions that have helped me throughout to put in my best efforts into the same.

I express my deepest sense of gratitude to Dr. Navneeta Bharadvaja, Assistant Professor, Department of Biotechnology, Delhi Technological University, who has been a wonderful mentor throughout the period of my project. I owe all the successful benchmarks of my entire work to her priceless counseling and thorough support.

I also acknowledge my work to all faculty members, technical and official staffs, for every kind of technical assistance and encouragement that have assisted me accomplish my targets in time.

Lastly, I would like to thank my class mates, as well as my seniors, who have helped me, directly or indirectly, during the course of my project and have aided in making it a success.

KAMLENDRA GUPTA

2K14/BIO/06

CONTENTS

.

LIST OF FIGURES

LIST OF TABLES

.

In-silico docking studies of Plant Derived Natural Compounds to identify potential target for the treatment of Ebola Virus

KAMLENDRA GUPTA

Delhi Technological University, Delhi, India
Email ID: kamlendra.gupta@gmail.com

ABSTRACT

Ebola virus is a negative sense, single stranded RNA virus that leads to acute hemorrhagic fever in humans and nonhuman primates. This virus is impervious to a big portion of the known antiviral drugs and there is no effective treatment as on date for disease caused by this pathogen. Looking into its capability to create a pandemic scenario across globe, there is an ultimate need for new drugs and therapy to combat this infection. The recent study deals with the assessment of the inhibitory activity of flavonoids contrary to the four selected Ebola virus receptor proteins, using *in-silico* studies. The viral protein VP35 was docked with small molecules obtained from flavonoid class and its derivatives, assessed on the basis of energetics, stereochemical considerations and pharmacokinetic properties to isolate potential lead compounds. The results exposed that the top ranking screened flavonoids i.e., and shown well docking charge and binding energies in all the EBOV receptors when matched with that of the described compound. All the screened flavonoids have known antiviral activity, suitable pharmacokinetic property and are being used on human, hence can be taken as anti-Ebola treatment without the time lag for clinical trial. Ligands were identified based on the active sites and docked subsequently to find out the best ligand, 7-hydroxyflavone. This ligand has better binding energy than the all other ligands. Further, Lipinski's filters, various others physicochemical properties and toxicity studies were also done to check the bioavailability and toxicity of the top ligands.

.

INTRODUCTION

The Ebola epidemic in 2014 is one of the largest viral occurrences in history and the first in West Africa. Even though currently it is distressing four countries in West Africa namely Guinea, Liberia, Nigeria, and Sierra Leone, conferring to the recent reports of the Centre for Disease Prevention and Control (http://www.cdc.gov/vhf/ebola/), it is spreading across globe as a impending pandemic scenario. Ebola virus is a fatal pathogenic virus which is categorized under filoviridea family. It was primarily reported in Africa in 1976 in democratic republic of Congo (Adrian *et al*.,2003; Towner *et al*., 2007). Severe and frequently fatal haemorrhagic fever is the major symptom of Ebola infection which occurs in two phases; incubation period and late stage. Incubation period displays indications like joint pain, fever, weakness, sickness which can retain going for one week and late side effects comprise sorrow, eye irritation, and hemorrhagic rash over the complete body (Dixon *et al*., 2002; International Commission report,1978). Ebola virus has cord like tubular filaments which are condensed with viral envelope (Hartlieb *et al*.,2003; Heinz *et al*., 2011). It has a negative sense stranded genome which comprises seven structural genes: out of which there are four structural proteins and three membrane associated proteins coded by Ebola virus genome (Hoenen *et al*., 2011; Saphire *et al*., 2008; Wei *et al*., 2002). No drug or vaccine is presently available for treatment of Ebola virus infection. Plant-derived flavonoids, also known as phytoestrogens, contribute a significant proportion to our daily diet including food supplements and neutraceuticals. Dietary flavonoids have been conveyed to significantly reduce incidence of breast cancer among women of Asian and Oriental origin (Adlercreutz, 1995; Fink et al., 2006; Jordan, Mittal, Gosden, Koch, & Lieberman, 1985; Peterson et al., 2003). Naturally occurring flavonoids and their byproducts have been reported to retain inhibitory effects on VP35 and therefore, have invited significant attention in development of these molecules as drugs against ebola virus disease. Thus, a complete quantitative structure–activity relationship analysis of these compounds would deliver insight into identification and optimization of the top molecules and in combinatorial library-based drug designing. Computer-

.

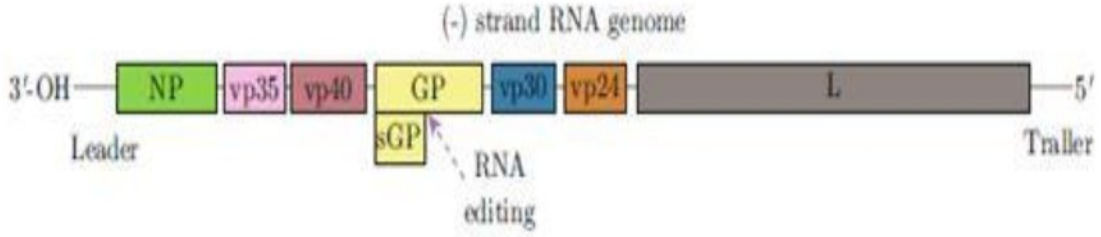


Fig. 1 Ebola viral genome arrangement (Hulo *et al.*, 2013)

aided drug discovery (CADD) comprises two main approaches, namely target-based and ligand based approaches. Target based drug discovery be determined by the structure of the target and its interfaces with the ligands. On the other hand, ligand based drug discovery is dependent on the structural statistics and molecular properties of known ligands only.

In-silico CADD methodologies can be considered as key to develop and screen the compounds or drugs to make effective leads for treatment of numerous diseases. Structure based drug designing methodologies involves the 3-D structure of protein on which docking studies of several specific small molecules have been approved in order to calculate their docking score and binding energy by utilizing a sequence of scoring functions as well as to discover the binding affinity of ligands and their cooperating residues.. The virtual screening & molecular docking of the drug candidates on mark protein could discover out the best lead like compounds with additional optimization of the compounds to finalize the lead (Rajamani et al., 2007).

OBJECTIVES

1. Novel drug target identification
2. Active site prediction of target protein
3. Finding ligands based on the target protein active sites
4. Checking lipinski's filters
5. Docking of different ligands with the target protein molecule
6. Checking toxicity of ligands molecules

.

REVIEW OF LITERATURE

Ebola virus codes seven polypeptides from its RNA genome of ca. 19.0 kb, comprising the glycoprotein (GP),nucleoprotein (NP), RNA-dependent RNA polymerase(L), Viral Protein 35, Viral Protein 30, Viral Protein 40, and Viral Protein 24 (Nan bo *et al*., 2013; Hulo *et al*., 2013). Among all proteins, we have concentrated on VP40, VP35, VP24 and VP30 proteins as the effective drug targets observing into the function of these proteins to cause defensive immune responses to Ebola virus. The most abundant protein found under the viral bilayer is VP40 and it is requisite to make the structural reliability of the viral particles. The association and budding development of the matrix protein VP40 occurs at the plasma membrane and needs lipid raft micro domains (Geisbert *et al*., 1995). During its replication, it also plays an significant role either in the RNA metabolism of viral or host cell (Gomis-R¨uth*et al*., 2003). The EBOV VP40 composed of the N terminal and C-terminal domains which are associated together by a flexible linker composed of residues ranging from 195 to 200. The amino-terminal domain is folded into a β-sandwich composed of six antiparallel strands organized in two β-sheets of three strands each. This laterally with the C terminal domain form the monomer of Viral Protein 40 receptor. The C-terminal domain of Viral Protein 40 act as a possible drug target due to its function in membrane association (Ruigrok *et al*., 2000), whereas the amino terminal domain is liable for the oligomerization of the protein. The carboxy-terminal domain consist of of a conserved proline-rich region in Viral Protein 40 EBOV receptor ranging from amino acids 205–219 which is liable for interaction with cellular Sec24C and also necessary for plasma membrane targeting and viral element release (Volchkov *et al*., 2011).The Ebola Viral Protein 35 protein is a key protein which acts as element of the viral RNA polymerase complex, viral assembly factor. It obstructs the host interferon (IFN) assembly hence is vital for virulence of EBOV. Additional the mutation of targeted residues within the caboxy-terminal of Viral Protein 35 damage its ds RNA-binding activity. It has been conveyed that the ds RNA binding cluster which is centered on Arg-312, is a highly preserved residue which is essential for EBOV virulence (Basler *et al*., 2003). Knockdown of Viral Protein 35 promotes reduced viral replication and reduced lethality in infected mice (Jin *et al*., 2010). Thus, Viral Protein 35 is a key drug target due to its role in viral replication and pathogenesis. Likewise the structural protein Viral Protein 24 of Ebola virus

.

(EBOV) has recognized to be antagonizing the host interferon function. It had also been proven that in a mice model this function could hang on the ability of VP24 to contrary the interferon system (Ebihara *et al*.,2006). There is similarly confirming that this protein can guidance the transcription and replication of an ebola virus mini genome (Ebihara *et al*., 2006; Watanabe *et al*., 2007) and, when rapidly conveyed together with Nucleo Protein and Viral Protein 35, is involved in the framing of viral nucleocapsids (Hoenen *et al*., 2010; Huang *et al*., 2002;Noda *et al*., 2006). Viral Protein 30 is important for the development of the viral mRNAs even though it has been notified that EBOV transcription could happen solely if the Nucleoprotein (NP) get changed; triggering the assimilation of the transcription start site (M¨uhlberger *et al*., 1999; Weik *et al*., 2002). It has been assumed that Viral Protein 30 may help to beat this obstacle for transcriptional enactment, steady with its mentioned part at an early phase of interpretation. Viral Protein 30 due to its function in homo oligomerization is mentioned as a potential target for antiviral treatment (Hartlieb *et al*., 2003).Configuration based drug designing approaches includes the 3-D structure of protein on which docking studies of several individual small molecules have been approved in order to compute their docking score and binding energy by using a series of scoring functions. The virtual screening & molecular docking of the drug molecule on target protein could find out the stop lead like compounds with additional optimization of the compounds to finalize the lead (Rajamani *et al*., 2007). According to modern reports, ZMapp, enhanced combination of drug contains monoclonal antibodies prepared from a tobacco-plant strain can act as antiviral remedy against Ebola infection (Lary Zeitlin *et al*., 2014). ZMapp is a mixture combining the best ingredients of two cures namely MB-003 (Mapp) and ZMAb (Defyrus/PHAC) and is formed in a laboratory by showing mice to fragments of the virus. But it doesn't get the FDA authorization yet. As of now, only two cases have been testified stating that ZMapp has been active in treating Ebola, while studies have revealed that only 43% of animals affected with the Ebola virus have been treated by ZMapp even though it is yet to be verified on human (Lary Zeitlin *et al*., 2014). Also another drug BCX4430 (Developed by BioCryst), a new artificial adenosine analogue, prevents infection of different filoviruses in human5cells (Sina *et al*., 2014). Biological and chemical, reporter-based and primer-extension tests specify that BCX4430 prevents viral RNA polymerase role, acting as a non-obligate RNA chain terminator. Exposure of intramuscular management of BCX4430 protects against Ebola

.

virus disease in rodent models but it is not verified in humans yet. Flavonoids have low toxic effects and are broadly available in plants, including edible. The chemistry of affinity of the flavonoids on Ebola virus receptor has not yet been defined. In this study, we chosen flavanoids as inhibitor for Viral Protein 35 receptors along with the reported inhibitors. Further we decided a comparative study of the efficiency of these flavonoids against all the stated drugs against these said above receptors and implication that the screened flavonoids could act as better inhibitors than the prevailing reported drugs.

Subtractive genome approach

Genomics are often used for the evaluation of possible targets quality. It includes two measures:

1. Essentiality
2. Selectivity

The encoded proteins from the essential gene are of utmost important for any organism to survive (Kobayashi *et al.*, 2003; Mushegian *et al.*, 1996; Itaya, 1995). Galperin proposed that previous drug targets finding among qualified proteins are essential as well as specific for any pathogen (Galperin *et al.*, 1999). There is increase in number of target identification employing genomics application. A remarkable method known as Subtractive genomics (also known as Differential genome display) is suggested for identification of possible targets (Huynen *et al.*, 1997). It is depending upon the concept that parasite has usually smaller genome and a lower number of proteins. And in addition to this, target shouldn't show any homolog within *Homo sapiens*. Therefore, those can be treated as the possible targets (Dutta *et al.*, 2006; Sakharkar *et al.*, 2004). This approach is used by various researchers to identify new drug targets in ebola virus.

.

Docking

Docking is a process by which the best configuration of binding molecules is determined. In this process, a complex structure is obtained having stable structure (Lengauer and Rarey, 1996). Knowledge of favored orientation in turn can be used in predicting strength of association between the two molecules and binding energy can be measured in terms of scoring function.

Docking is often used for predicting the binding of drug candidates to their target protein to predict the activity and affinity of the drug candidates. Thus, docking perform a very significant character in the rational drug design (Kitchen *et al.*, 2004).

Docking approaches

There are 2 approaches popular in molecular docking community. First approach employs a method of matching where protein and ligand molecular surfaces are reported as complementary to each other (Goldman and Wipke, 2000; Meng *et al.*, 2004; Morris *et al.*, 1998). In the next approach, process of docking takes place and the interaction energy of protein-ligand complex is determined (Feig *et al.*, 2004).

.

MATERIALS AND METHODOLOGY

Online Tools and Database:

1. NCBI

NCBI is one of the part of National Institutes of Health branch (United States National Library of Medicine). There are a number of databases available in NCBI which are useful for biomedicine and biotechnology. Major databases are: -GenBank: For DNA sequences, -PubMed: For biomedical literature, -Protein: For protein sequences, etc.

2. PDB

The Protein Data Bank is a repository for 3D biological molecules (nucleotides and proteins) structural data. The structural data found experimentally by NMR spectroscopy or X-ray crystallography and put in by the biochemists or biologists and are accessible freely on Internet. Users can search in this database by PDB ID, macromolecule, author, sequence or ligands and download the required files in pdb format.

3. ZINC Database

The ZINC database comprises of commercially available chemical compounds. ZINC database is mainly used for virtual screening. ZINC is used by various scholars in research field as well as by the investigators in pharmaceutical companies or biotech companies. Users can search in this database by IDs, SMILES, etc. The ZINC database finds the compounds based on similarity to the query compound. The output result of query can be downloaded in the mol2, sdf, SMILES, ddb (flexibase) format. Other uses of the ZINC database: -obtaining a compound for purchasing, -obtaining compounds which can be used as a drug molecule, etc.

.

4. SCF BIO-IITD

Supercomputing facility at IIT Delhi has various online softwares which can be used for ligand screening and ligand optimization. Active site prediction, Lipinski filter etc are some of the few online resources which are accessible online on this site.

5. RASPD

RASPD is used to excluding the ligand molecules in the beginning based on physicochemical properties of ligands and active site of the target protein molecule. This tool searches based on various physicochemical properties like chemical formula, H-bond donors as well as acceptors, number of rings, etc. for every molecules.

6. Toxicity checker

As the name suggests, Toxicity Checker aims to identify whether any toxic substructure of the query compound is available or not and it also calculates the different properties of the compounds. It is available freely to the users. It helps the scholars, companies and research institutes by allowing them to use the available tools online. In this tool, users have two options to check for the toxic substructure in the compound. They can check it either by drawing the molecule or by providing molecule ID, SMILES, InChI, InChIKey.

Software:

1. AutoDock 4.2.5

Auto Dock 4.2.5 is a software used for the purpose of molecular docking of ligand to macromolecules like DNA, proteins, etc. There are two main programs in Auto Dock: (a) Auto Grid program for the identification of pre-computing grids, and (b) Auto Dock program for docking ligand molecule to a number of grids of the target protein. Binding energy calculated is the combination of intermolecular and torsional energies.

.

2. Osiris Data Warrior

OSIRIS Data Warrior is data analysis and visualization software. OSIRIS data warrior is helpful in predicting various physico-chemical properties and toxicity risk indication that must be optimized while designing pharmaceutically active compounds.

3. Chemsketch

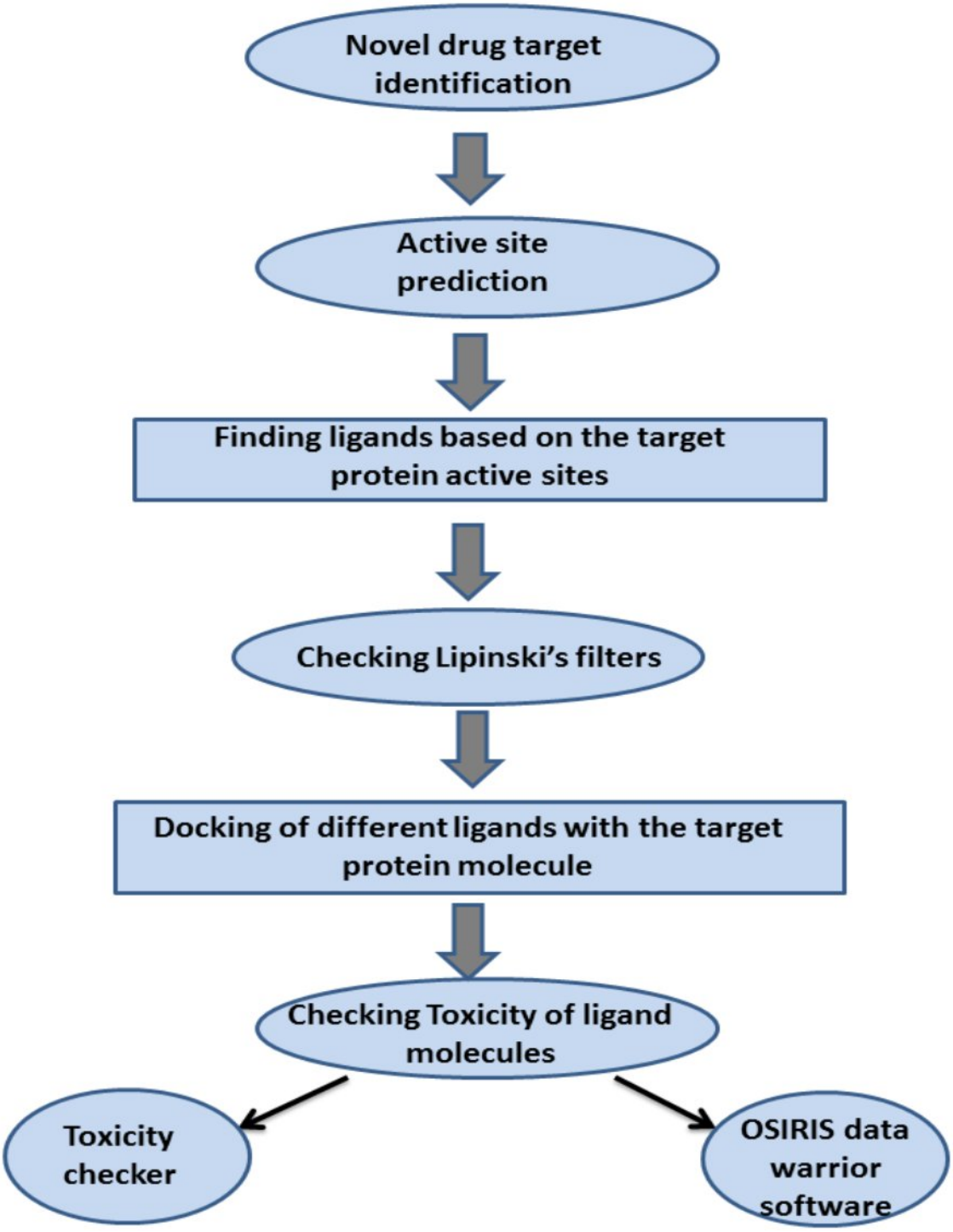
Chemsketch is a software which is used to draw chemical structure and create a file in desired format. This was used to draw chemical structure for molecules whose mol2 format was not available online.

4. Pymol

Pymol is a protein molecule visualization software. This software can further be used to edit protein structure by removing unnecessary ligands which are attached to the protein molecule file.

.

STEPS:


Novel drug target identification
Active site prediction
Finding ligands based on the target protein active sites
Checking Lipinski's filters
Docking of different ligands with the target protein molecule
Checking Toxicity of ligand molecules
Toxicity checker
OSIRIS data warrior software

.

1. Novel Drug Target Identification:

Ebola virus shows seven polypeptides from its RNA genome of ca. 19.0 kb, comprising the glycoprotein, nucleoprotein, RNA-dependent RNA polymerase (L), VP35, VP30, VP40, and VP24. Out of these proteins, we have concentrated on VP35 proteins as the key drug targets looking into the role of these proteins to cause defensive immune responses to Ebola virus. The most ample protein located under the viral bilayer is VP40 and it is required to make the structural reliability of the viral particles. The Ebola Viral Protein 35 protein is a key protein which acts as constituent of the viral ribonuclease polymerase complex, viral assembly factor. It obstructs the host interferon (IFN) assembly hence is vital for virulence of ebola virus. Further the mutation of designated residues within the carboxy terminal of Viral Protein 35 impair its dsRNA-binding activity. Therefore, Viral Protein35 is a key drug target due to its role in viral replication and pathogenesis.

2.Active site Prediction:

The Active site prediction server designed by IIT delhi can be used to determine the active sites. The server computes the cavities in a given protein using the uploaded pdb file. The active sites of VP35 were identified using SCFBIO IIT Delhi, The most essential property produced by SCFBIO site is an overall Site score, which has demonstrated to be successful at differentiating identified binding sites in co-crystallized centers. Active sites with best site scores were employed as prerequisite for receptor grid generation with the Viral Protein 35.

.

Cavities	
cavity_1_VPTWRDIQSHLKAGN	cavity_2_KDTLRNQGVSIHAPYF
cavity_3_FGPIRDKQYVWLTNHA	cavity_4_KPLRWNTVDSQIGHA
cavity_5_DKIVQPRAELSTGN	cavity_6_RLKPDWSNTVQIGA
cavity_7_QVDTKRPLIWHYCGFSA	cavity_8_PKNWTVDSGLIRQAE
cavity_9_KRQGAPSFTLCHIWVD	cavity_10_RFGAPTCSQHIDYNVK
cavity_11_VAHKLPQIDTSRGNW	cavity_12_KRASLCTQFIPWGVD
cavity_13_QDIRKPSTVLWCFA	cavity_14_AKRQIVNHYFDPGTC
cavity_15_DQKLVPAIRHTGS	cavity_16_VPSKQRTGWLIA
cavity_17_VPTRISQLKADW	cavity_18_HPDYLANFIQEKS
cavity_19_IFCKQGVWRPTDSAEL	cavity_20_VERAKDQTLIGCPFS
cavity_21_SPDGCLIHQAEKFRV	cavity_22_IFCQGDKVRSTEAP
cavity_23_DSIPAQVERK	cavity_24_NIQDKHRAESLP
cavity_25_EQIHYALDGCPFST	cavity_26_PQSHDGYLANFI
cavity_27_KDNSQGLITARP	cavity_28_GPYIVQDHNFKRAL
cavity_29_RDKVTSIPLGAQC	cavity_30_KQRASLCPIWTV
cavity_31_AIHQENYLDP	cavity_32_DSKILAVRHQN

Fig2. Active site predicted cavities of VP35

3. Docking of ligand with the target protein molecule

From Protein Data Bank, the natural ligand of our target protein molecule was found. And the ligand molecule was downloaded in mol2 format from ZINC database.

Then, we had used the docking software, AutoDock 4.2.5 for docking the natural ligand with the target protein molecule and binding energy had been recorded. Now we have to identify the ligands having lesser binding energy.

4. Finding ligands based on the target protein active sites

Online available tool RASPD had been used to identify various ligands. This method is useful when only mark protein molecule is available. RASPD determines ligands depending upon the active grooves present on the target molecule. So, we obtained the library of ligands with their

.

ZINC ID, IUPAC name and the 3D coordinates of the atoms involved. Then, all the ligands were downloaded from the ZINC database and thus we had generated the virtual library of ligands.

5. Docking of different ligands with the target protein molecule

Each of the ligands as well as drug molecules was docked one by one with the target molecule using AutoDock 4.2.5 software. All the values for binding energies are displayed in Table 1.

6. Checking Lipinski's filters

Lipinski's rule of five had been checked for top 10 ligand molecules, according to binding energy. These filters are obtained from one of the drug design tools at scfbio-IIT Delhi. The results are displayed in Table 2.

7. Cheking Toxicity of ligand molecules

a) Using Toxicity checker

Using SMILES sequence of the ligand, molecule structures was drawn and was checked for the toxic substructure. It is done for top 10 ligands and is shown in the result section.

b) Using Osiris data warrior software

Using OSIRIS data warrior software, toxicity as well as physicochemical properties was obtained for top 10 ligand molecules and compared with the available drug molecules and summarized in Table 3.

.

RESULTS AND DISCUSSION

1)3D structure view of the protein with PDB ID 3FKE.

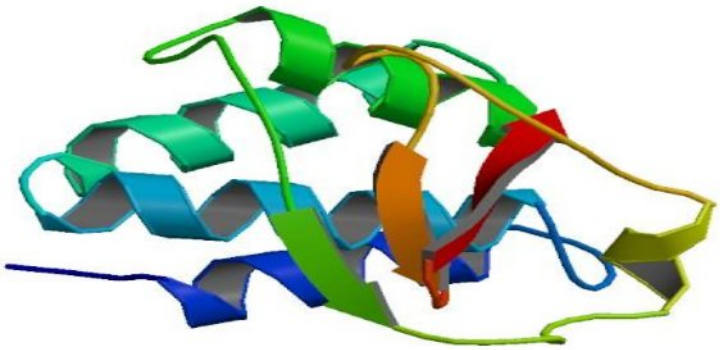


Fig3.3D structure of VP35

2.Docking of ligand with the target protein molecule

The ligand identified for protein with PDB ID 3FKE 7-hydroxyflavone, 7,8-dihydroxyflavone, 7-methoxyflavone,Flavan-4-ol,Apigenin,Chrysin,3-hydroxyflavone,Flavone,3,4Dimethoxyflavone, Flavanone, 3,4-Dimethoxyflavone, Flavanone, Naringenin, 4-hydroxyflavanone, 5-Hydroxyflavone, Isoflavanone, 5,7-dihydroxyflavone, 4-o-methylequol, Quercetin, Eriodictyol, Luteolin, Taxifolin. Now we have to identify the ligands having lesser binding energy.

.

AutoDock result of first 20 ligands:

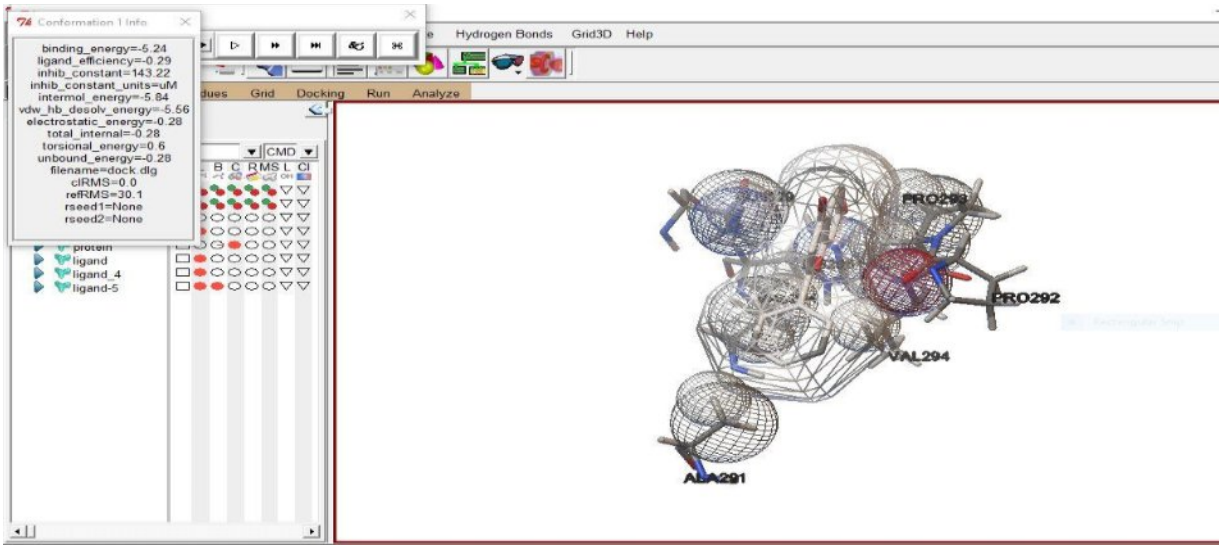


Fig1. Docking result of 7-hydroxyflavone with the target protein 3FKE

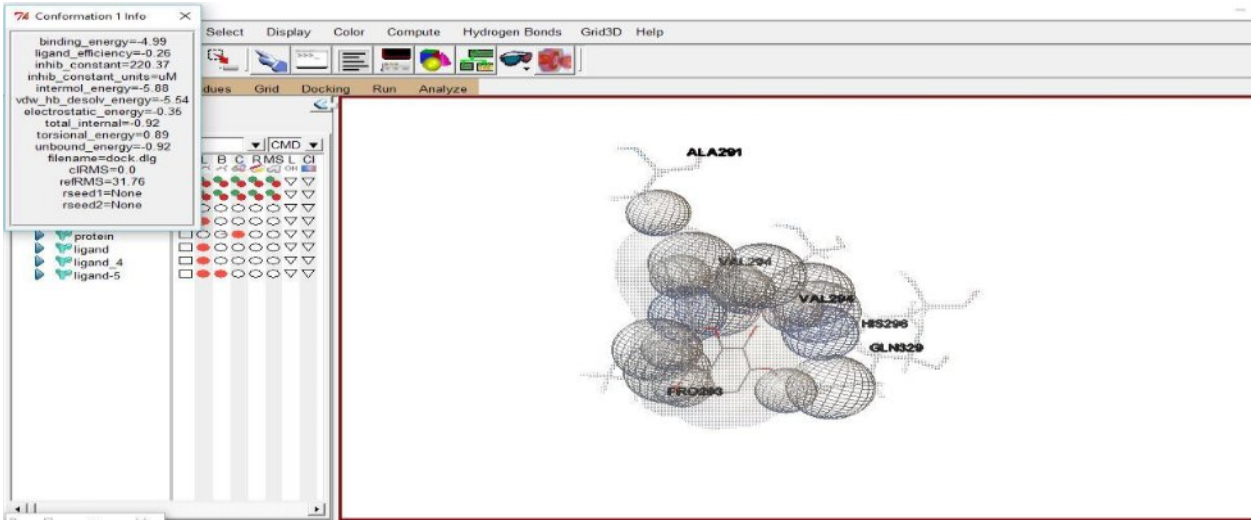


Fig2. Docking result of 7,8-dihydroxyflavone with the target protein 3FKE

.

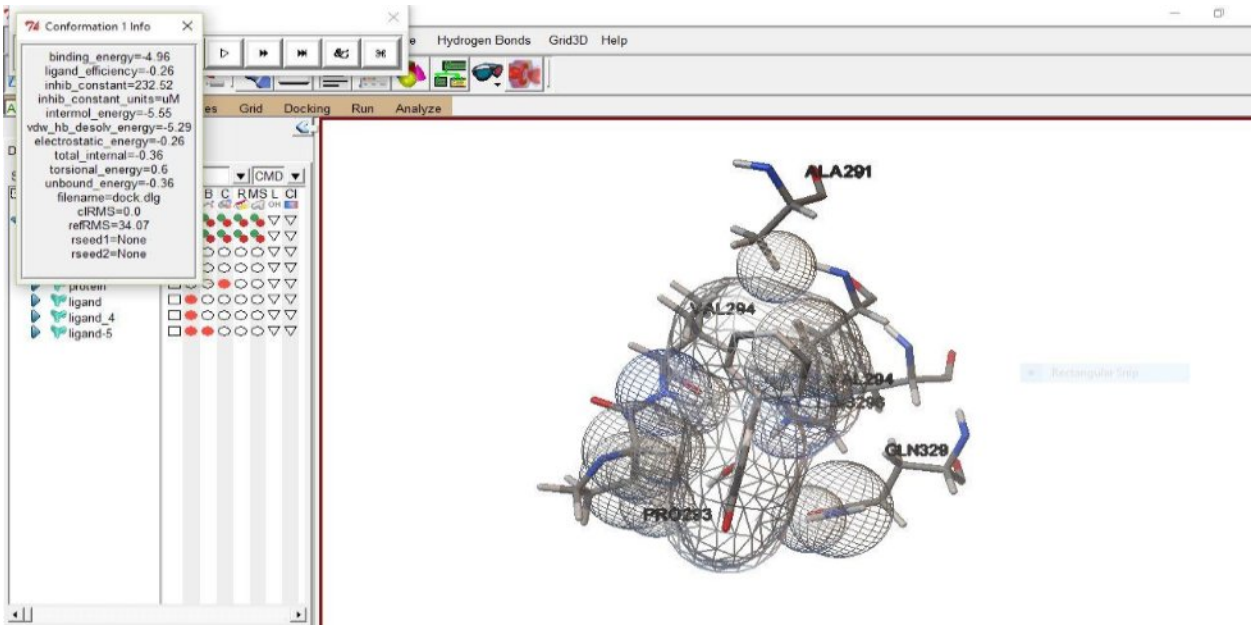


Fig3. Docking result of 7- methoxyflavone with the target protein 3FKE

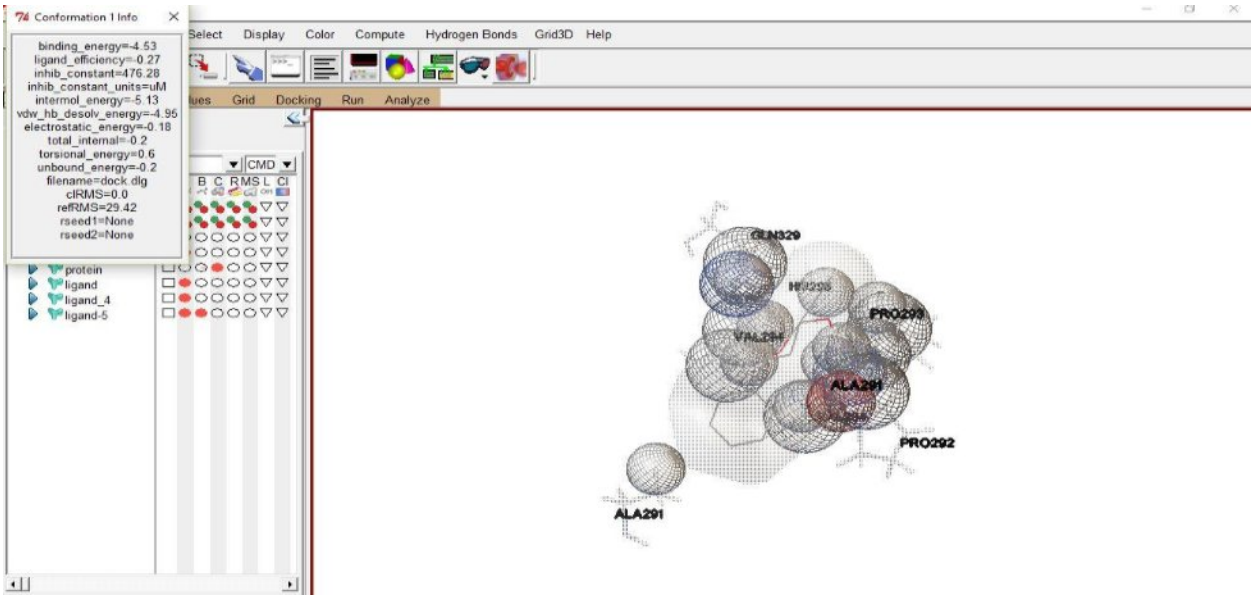


Fig4. Docking result of flavan-4-ol with the target protein 3FKE

.

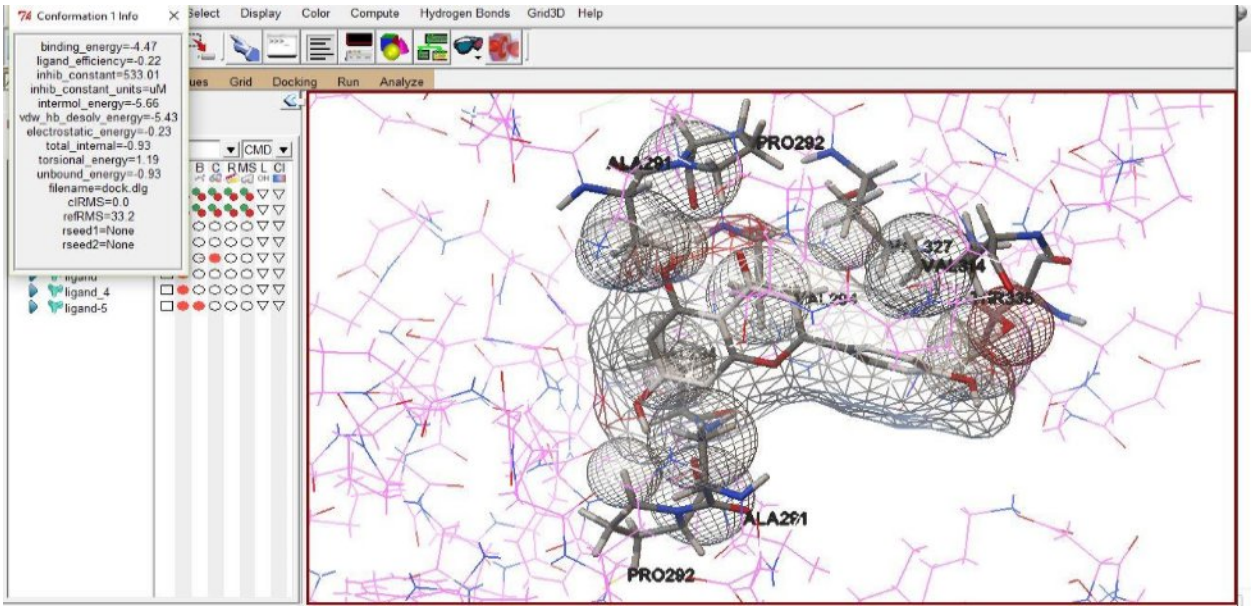


Fig5. Docking result of apigenin with the target protein 3FKE

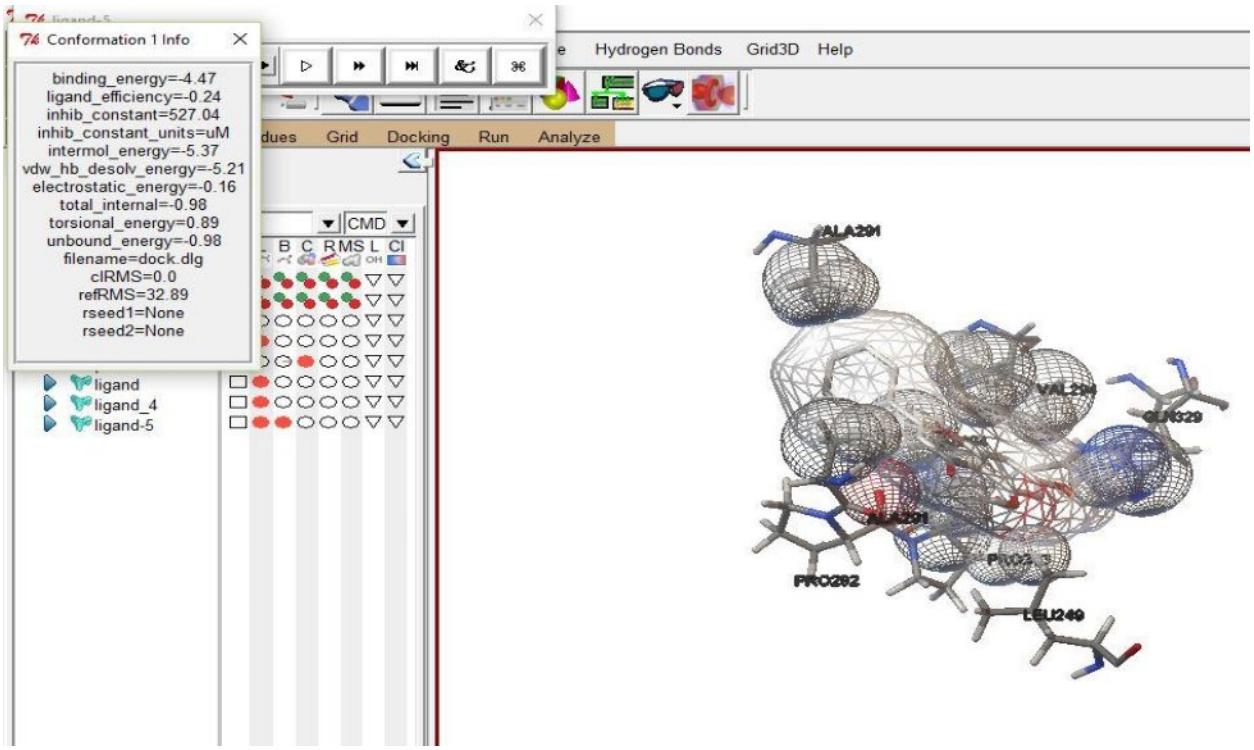


Fig6. Docking result of chrysin with the target protein 3FKE

.

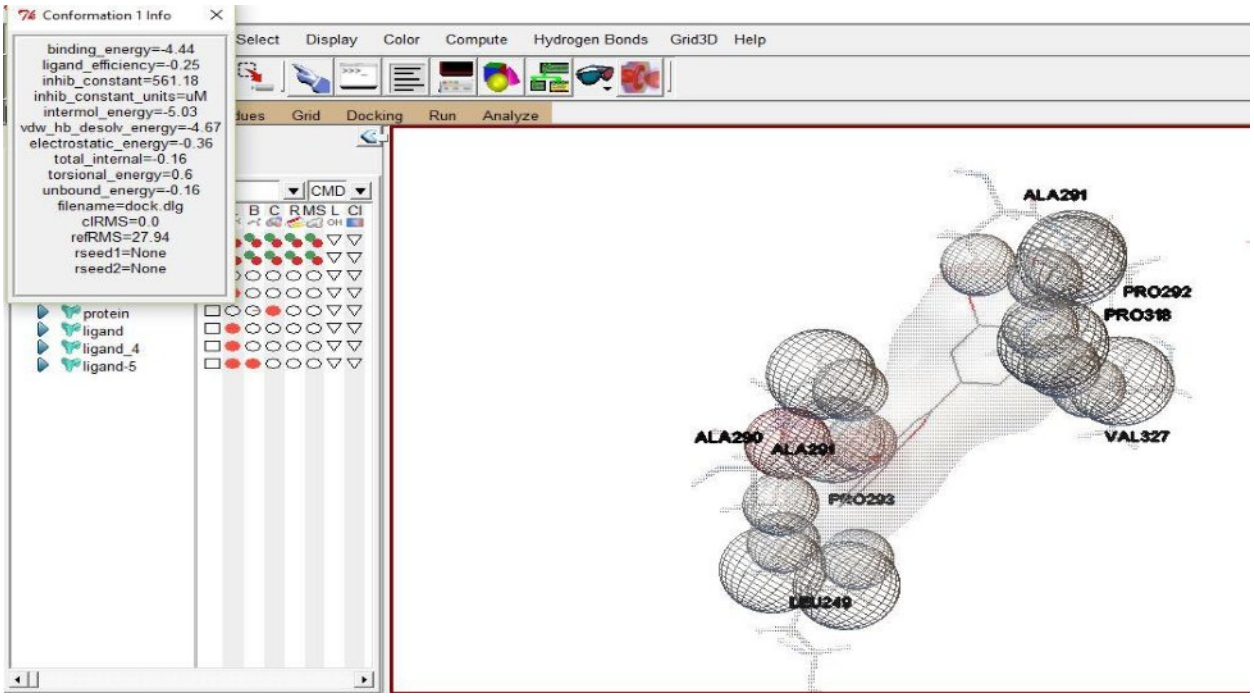


Fig7. Docking result of 3-hydroxyflavanone with the target protein 3FKE

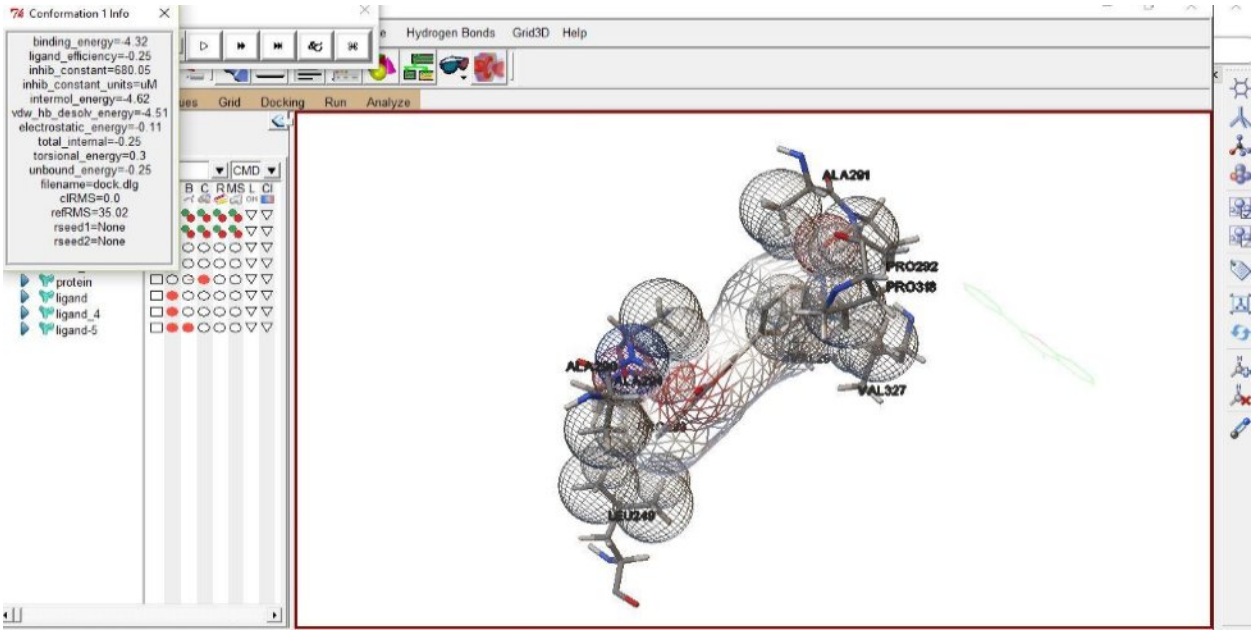


Fig8. Docking result of flavone with the target protein 3FKE

.

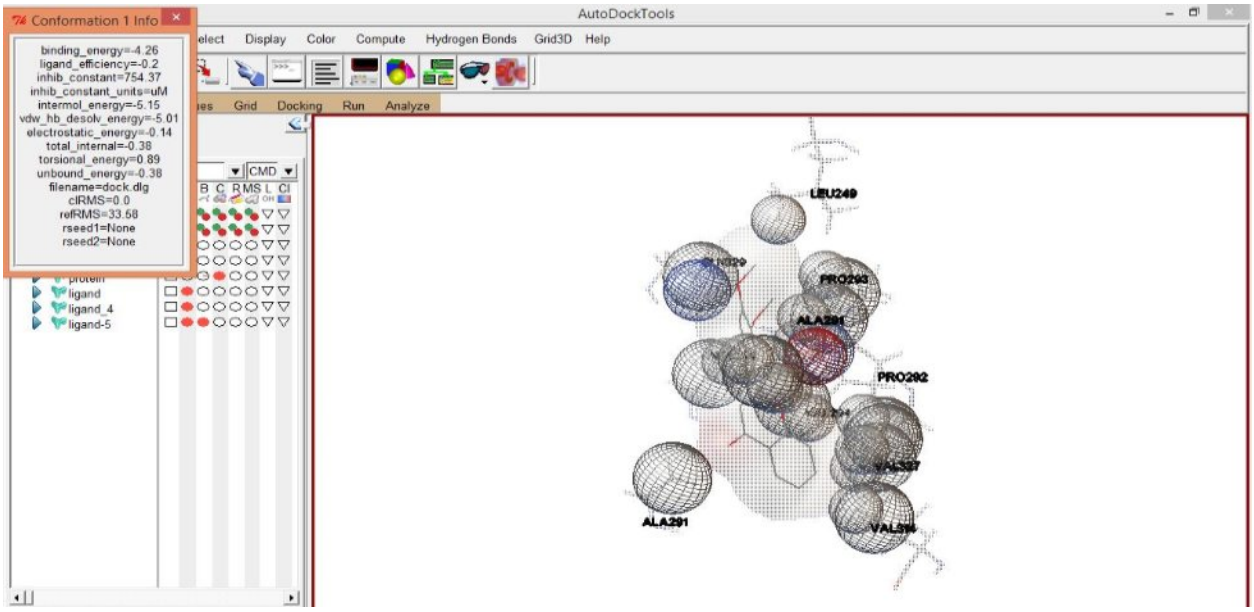


Fig9. Docking result of 3,4-Dimethoxyflavone with the target protein 3FKE

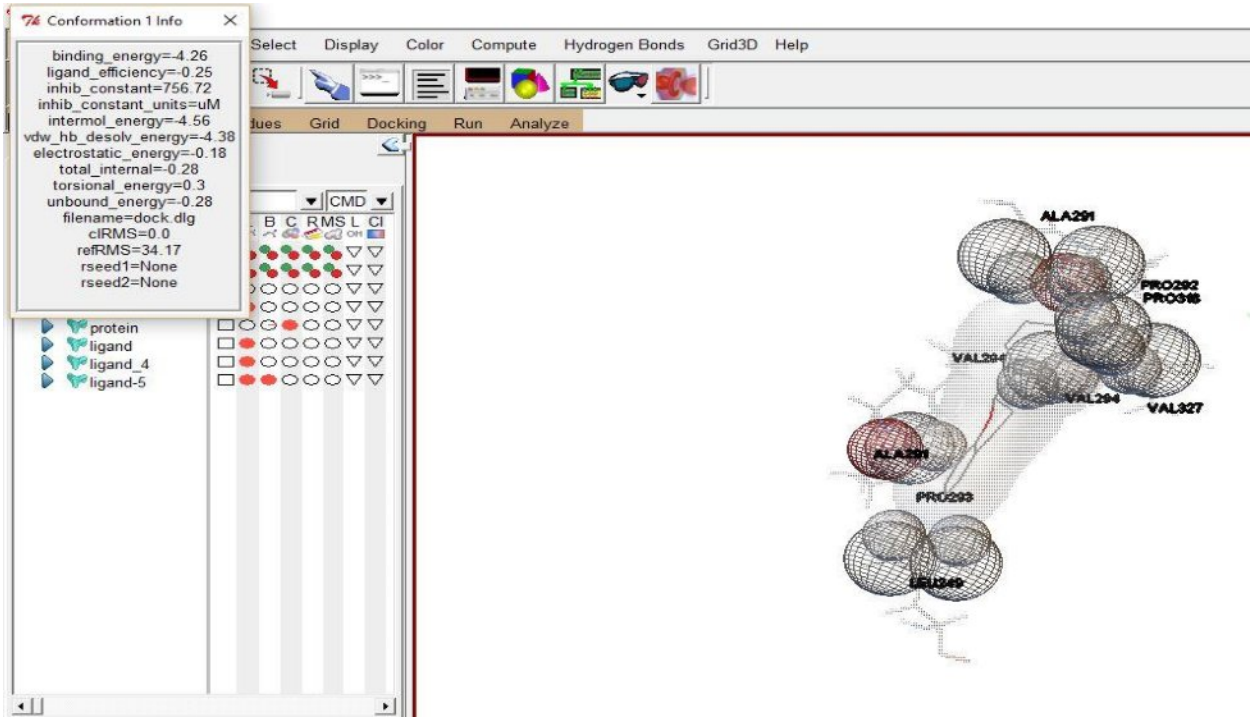


Fig10. Docking result of Flavanone with the target protein 3FKE

.

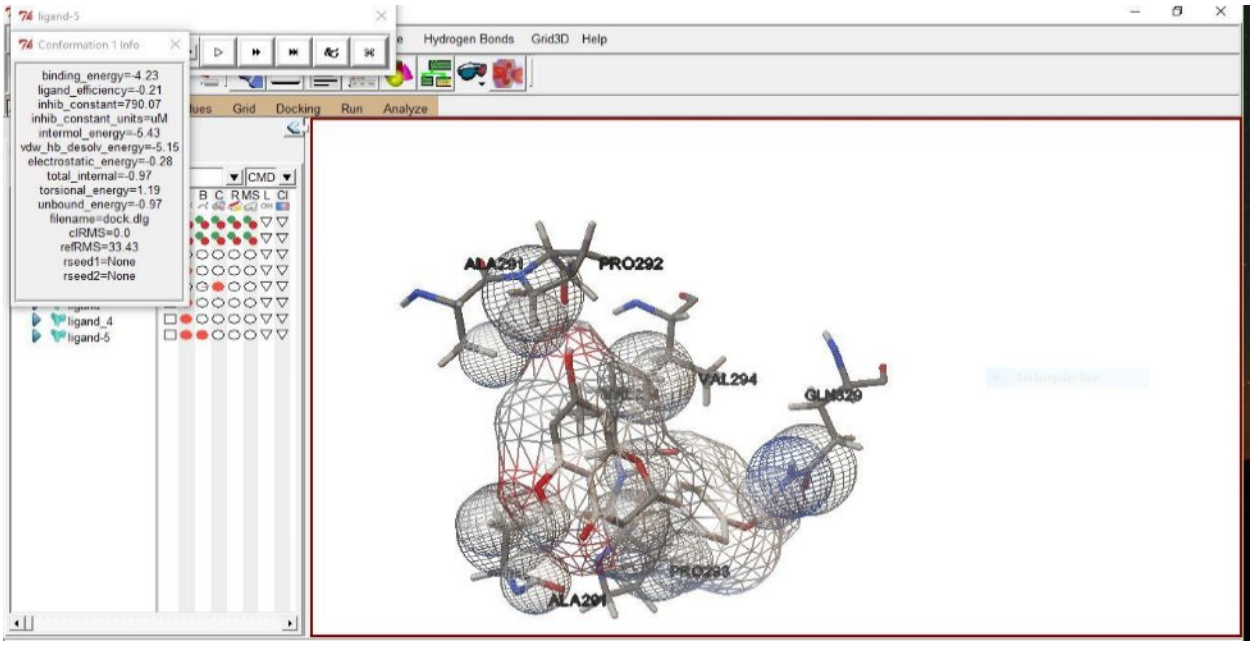


Fig11. Docking result of naringenin with the target protein 3FKE

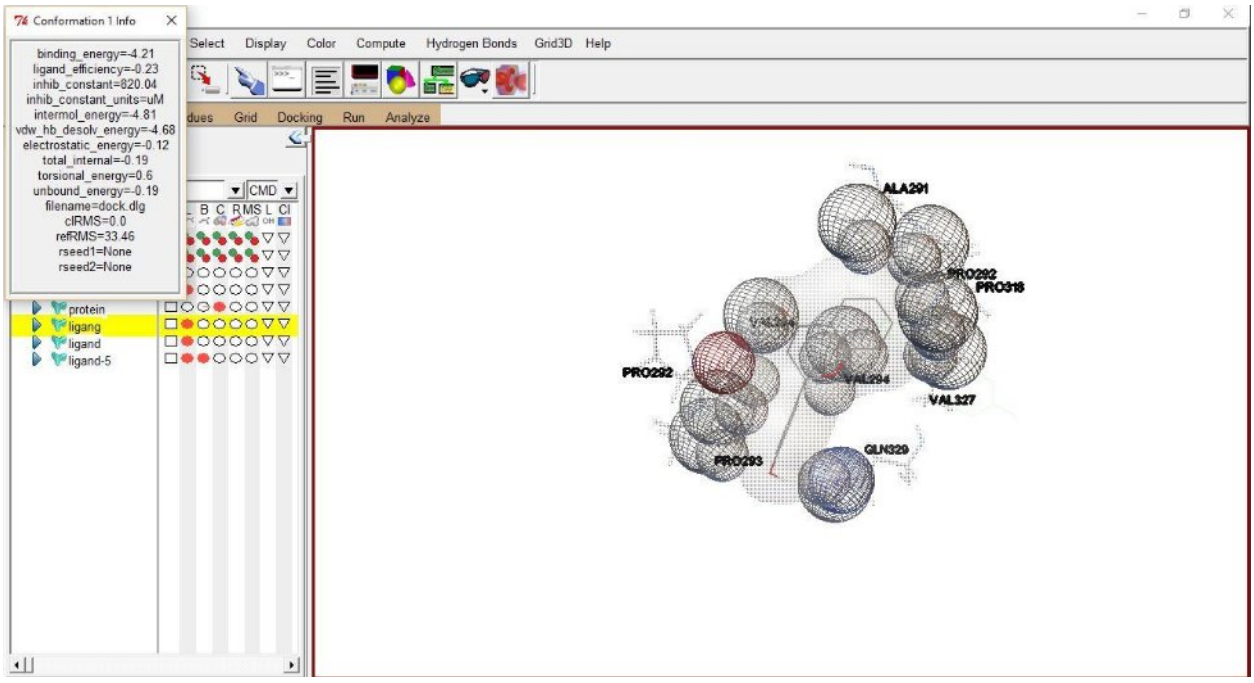


Fig12. Docking result of 4-hydroxyflavanone with the target protein 3FKE

.

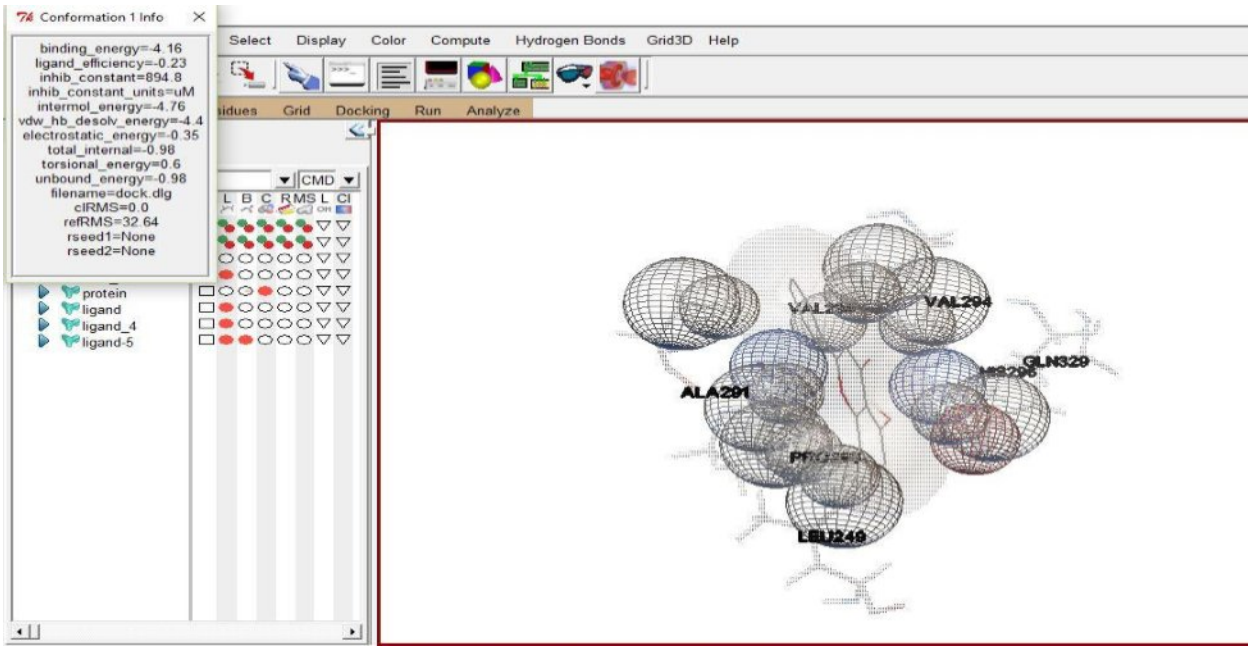


Fig13. Docking result of 5-Hydroxyflavone with the target protein 3FKE

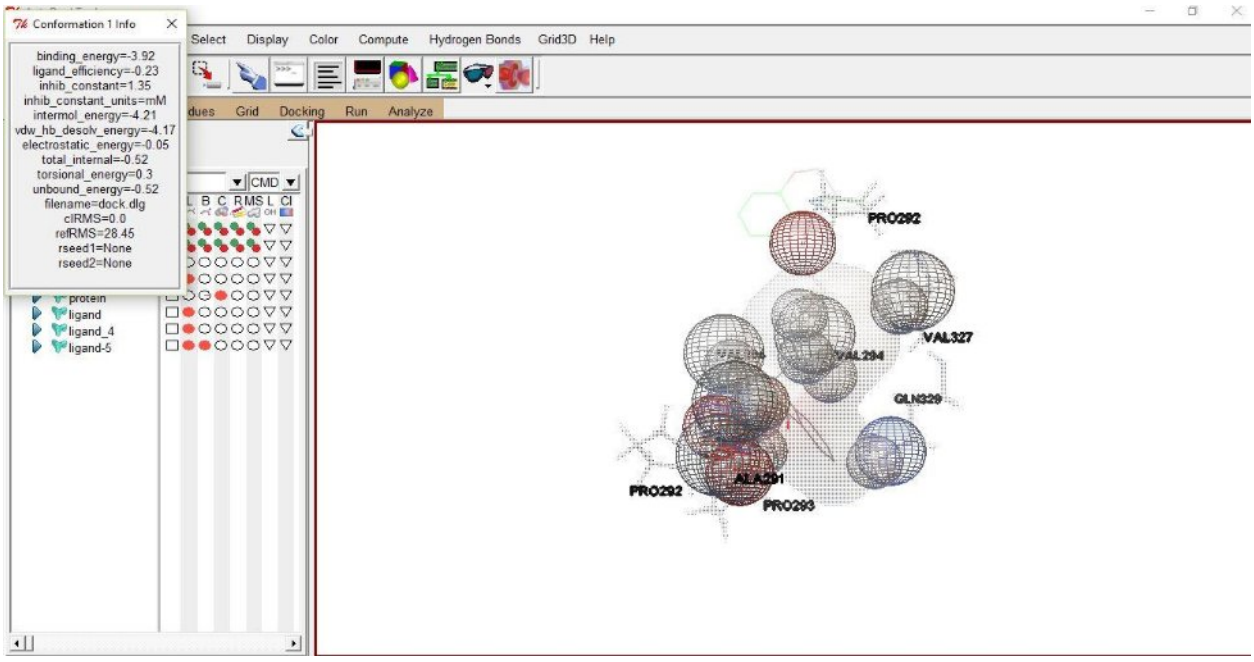


Fig14. Docking result of isoflavanone with the target protein 3FKE

.

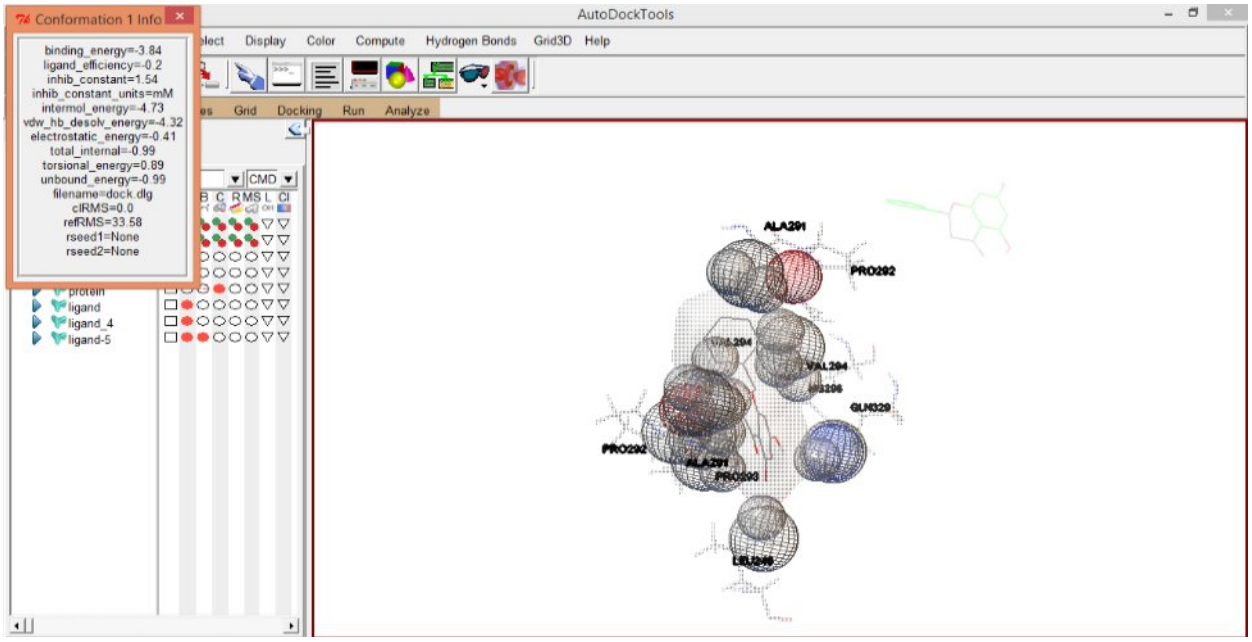


Fig15. Docking result of 5,7-dihydroxyflavone with the target protein 3FKE

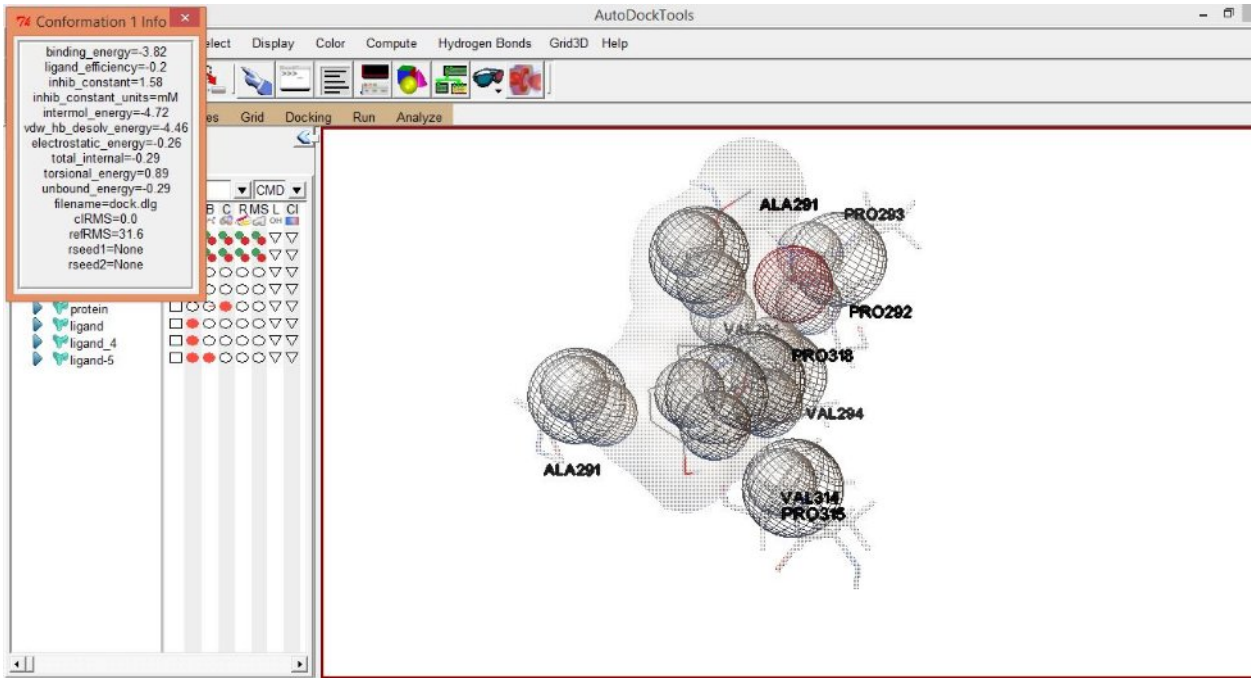


Fig16. Docking result of 4-o-methylequol with the target protein 3FKE

.

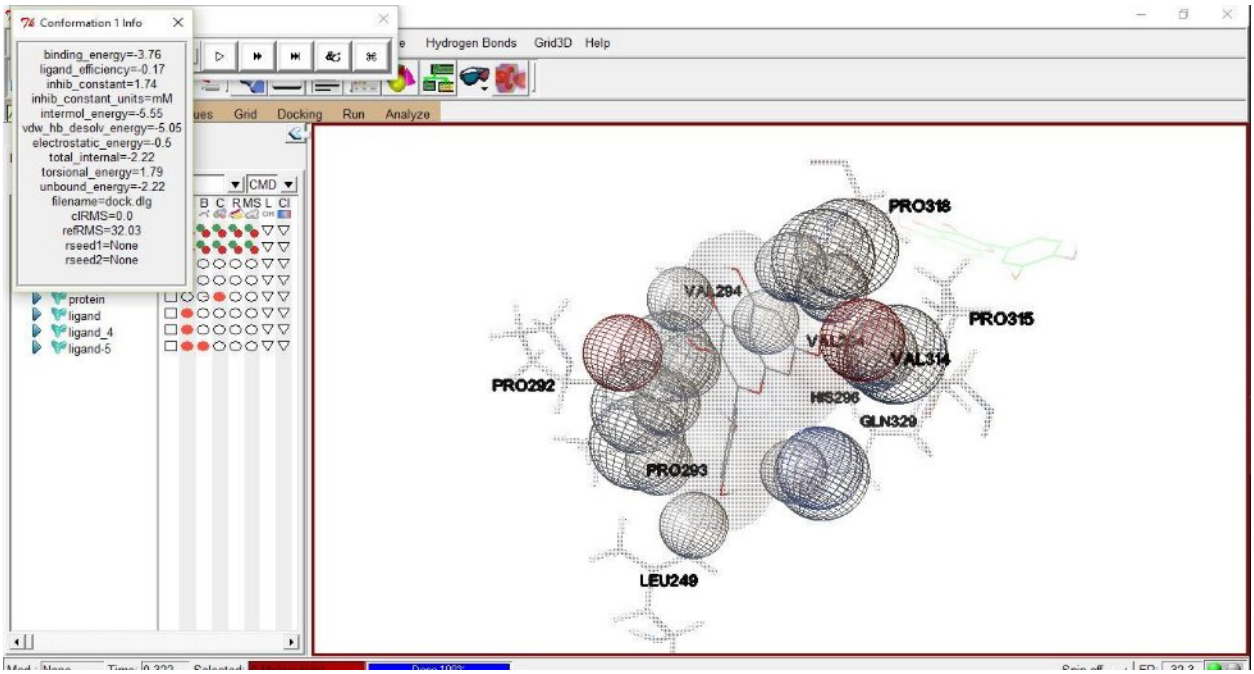


Fig17. Docking result of quercetin with the target protein 3FKE

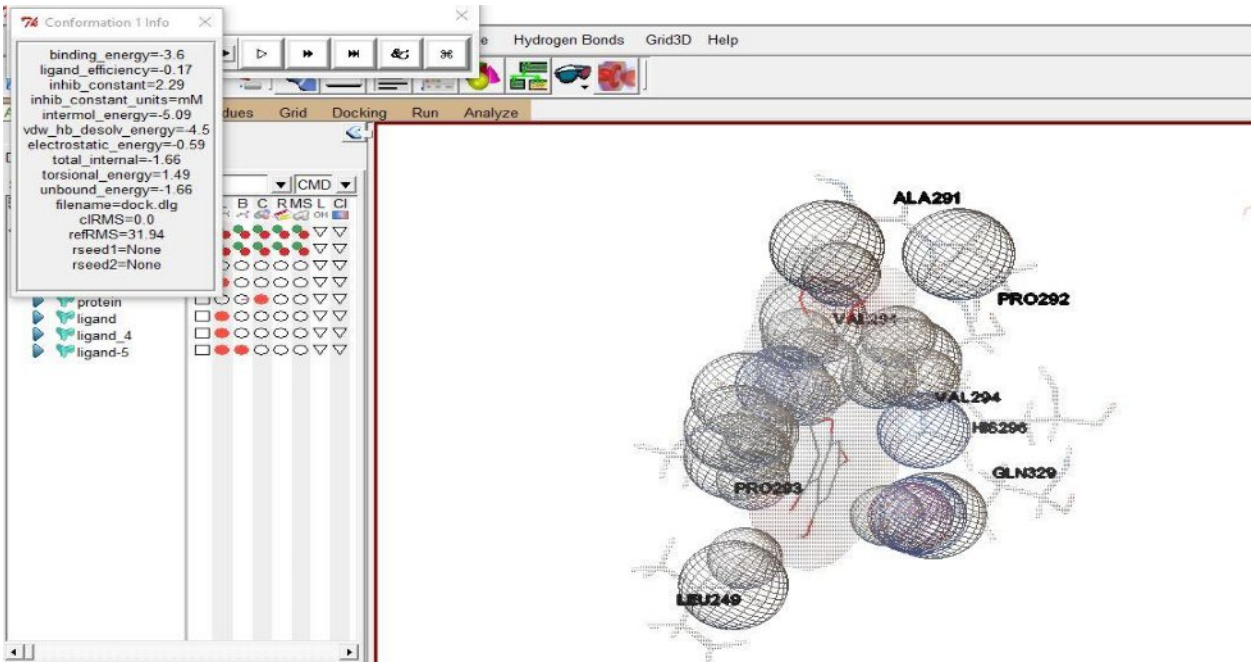


Fig18. Docking result of eriodictyol with the target protein 3FKE

.

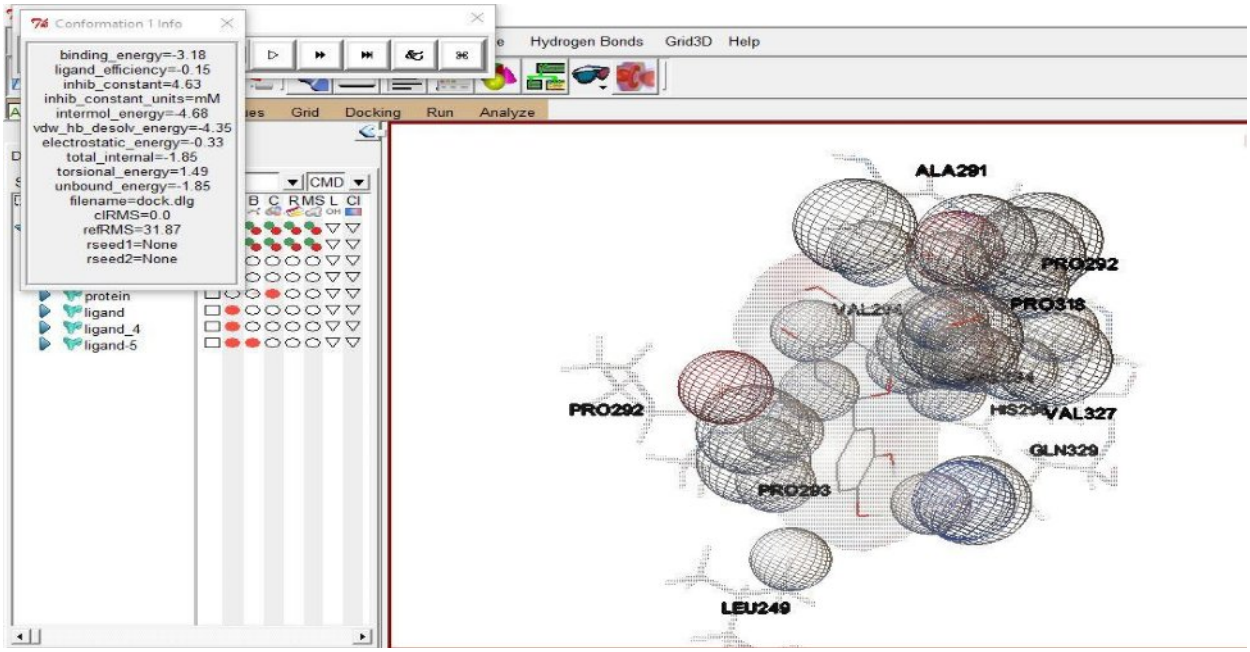


Fig19. Docking result of luteolin with the target protein 3FKE

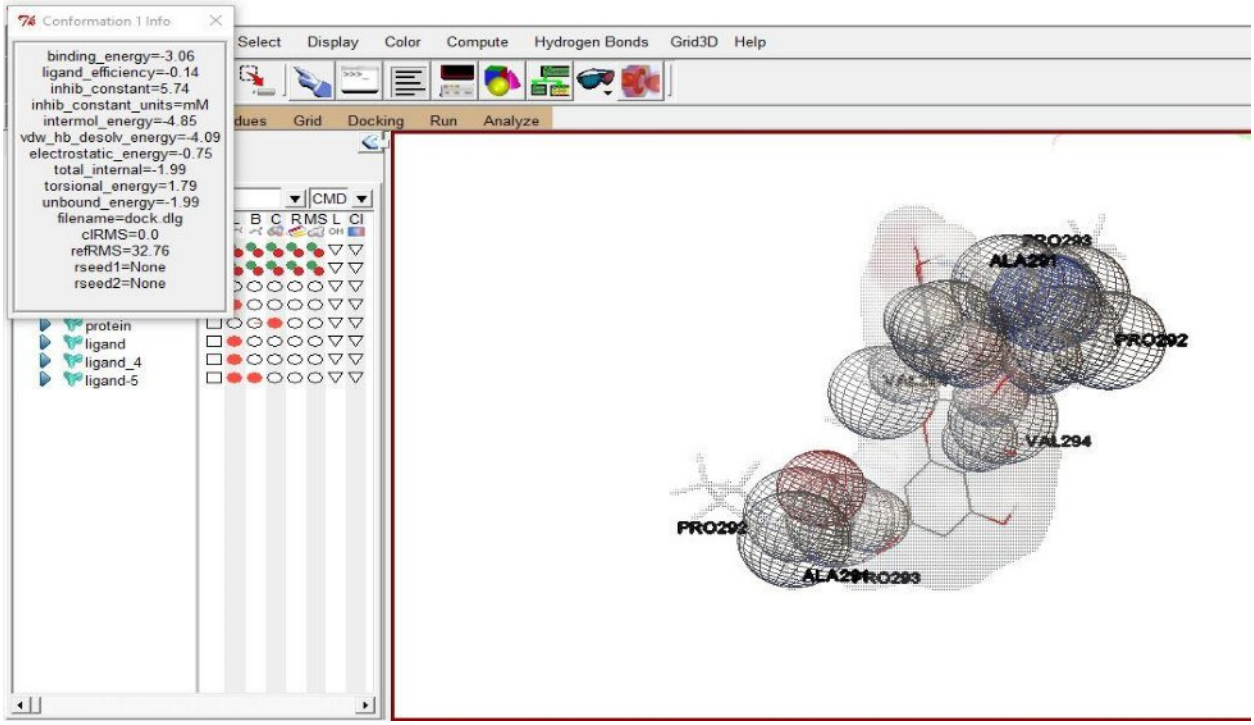


Fig20. Docking result of taxifolin with the target protein 3FKE

S.No.	Name	ZINC ID	Binding energy
1	7-hydroxyflavone	zinc_5934541	-5.24
2	7,8-dihydroxyflavone	zinc_57657	-4.99
3	7- methoxyflavone	zinc_18056	-4.96
4	Flavan-4-ol	zinc_85342	-4.53
5	Apigenin	zinc_3871576	-4.47
6	Chrysin	zinc _3872070	-4.47
7	3-hydroxyflavone	zinc_57678	-4.44
8	Flavone	zinc_57674	-4.32
9	3,4-Dimethoxyflavone	zinc_57672	-4.26
10	Flavanone	zinc_58113	-4.26
11	Naringenin	zinc_156701	-4.23
12	4-hydroxyflavanone	zinc_57922.mol2	-4.21
13	5-Hydroxyflavone	zinc_57676	-4.16
14	Isoflavanone	zinc_134619	-3.92
15	5,7-dihydroxyflavone	zinc_4935	-3.84
16	4-o-methylequol	zinc_2558138	-3.82
17	Quercetin	zinc_3869685	-3.76
18	Eriodictyol	zinc_58117	-3.6
19	Luteolin	zinc_18185774	-3.18
20	Taxifolin	zinc_105086	-3.06

Table1. Binding energy of different ligands with the target protein

.

3. Checking Lipinski's filters

Lipinski's rule of five had been checked for top 20 ligand molecules selected based on the binding energy.

S.No.	Ligands (ZINC ID)	Molecular Mass	LogP	Hydrogen Bond Donors	Hydrogen Bond Acceptors
1	Zinc _5934541	238	3.00	1	3
2	Zinc_57657	254	2.71	2	4
3	Zinc_18056	252	3.31	0	3
4	Zinc_85342	226	3.24	1	2
5	Zinc_3871576	270	2.41	3	5
6	Zinc _3872070	254	2.71	2	4
7	Zinc_57678	238	3.00	1	3
8	Zinc_57674	222	3.30	0	2
9	Zinc_57672	282	3.31	0	4
10	Zinc_156701	272	2.50	3	5
11	Zinc_57676	238	3.00	1	3
12	Zinc_134619	224	3.04	0	2
13	Zinc_4935	256	2.80	2	4
14	Zinc_2558138	256	3.11	1	3
15	Zinc_3869685	302	2.01	5	7
16	Zinc_58117	288	2.21	4	6
17	Zinc_18185774	286	2.12	4	6
18	Zinc_105086	304	1.18	5	7
19	Zinc_58113	224	3.39	0	2
20	Zinc_57922	240	3.09	1	3

Table2. Lipinski's filters for top 20 ligands

.

4. Cheking Toxicity of ligand molecules

a)Using Toxicity checker

Presence of any toxic substructure was checked by the online tool Toxicity checker by providing the SMILES sequence of the ligand. It is done for the top 10 ligands and the result is shown below.

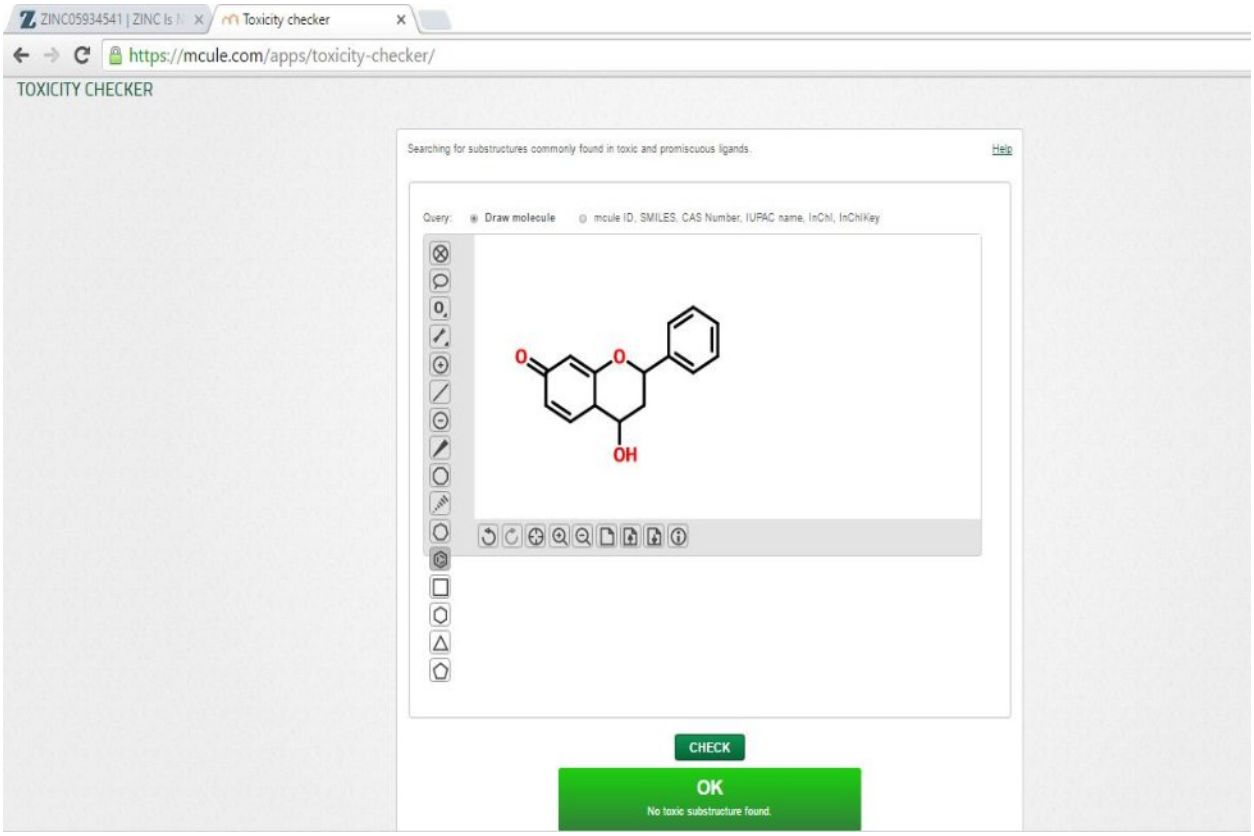


Fig23a.Toxicity checker result of ligand 1

.

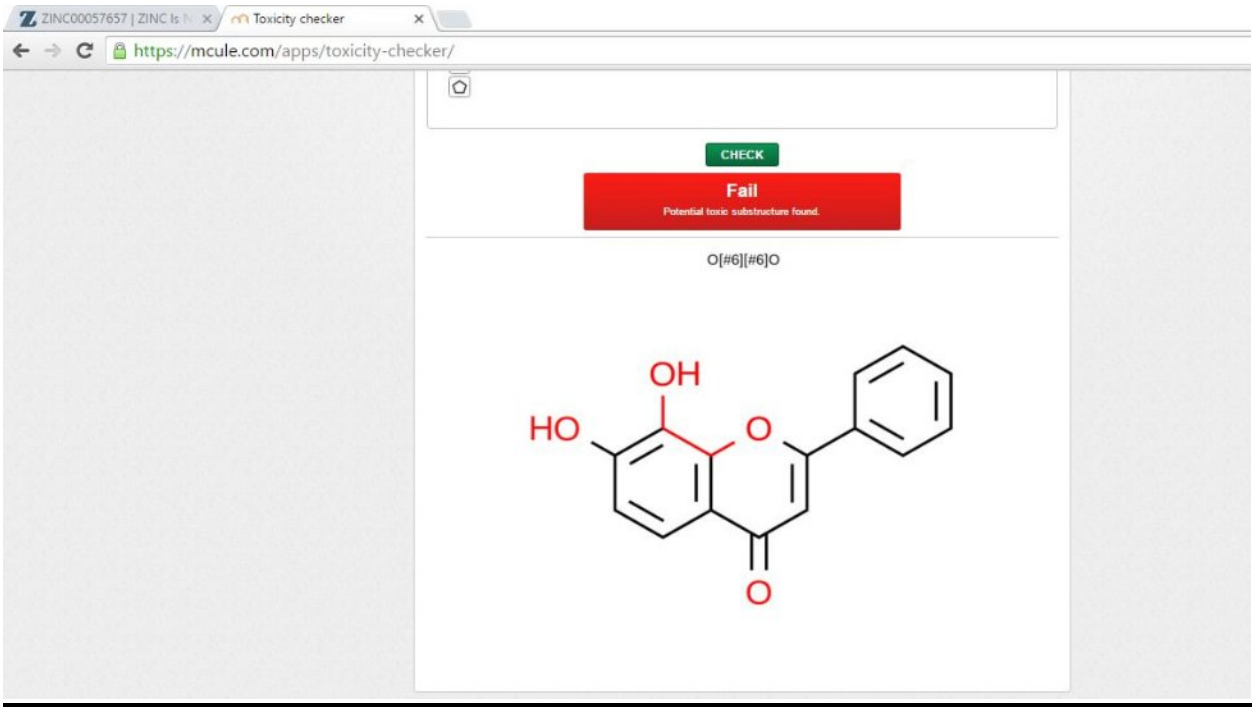


Fig23b.Toxicity checker result of ligand 2

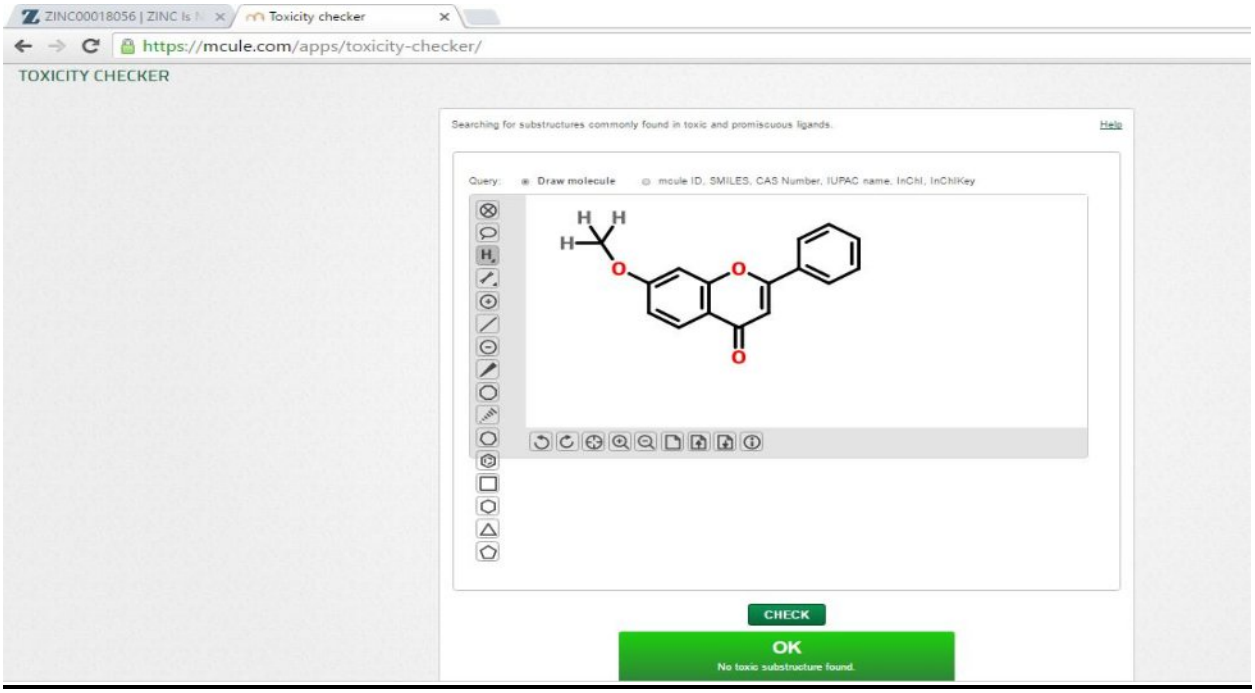


Fig23c.Toxicity checker result of ligand 3

.

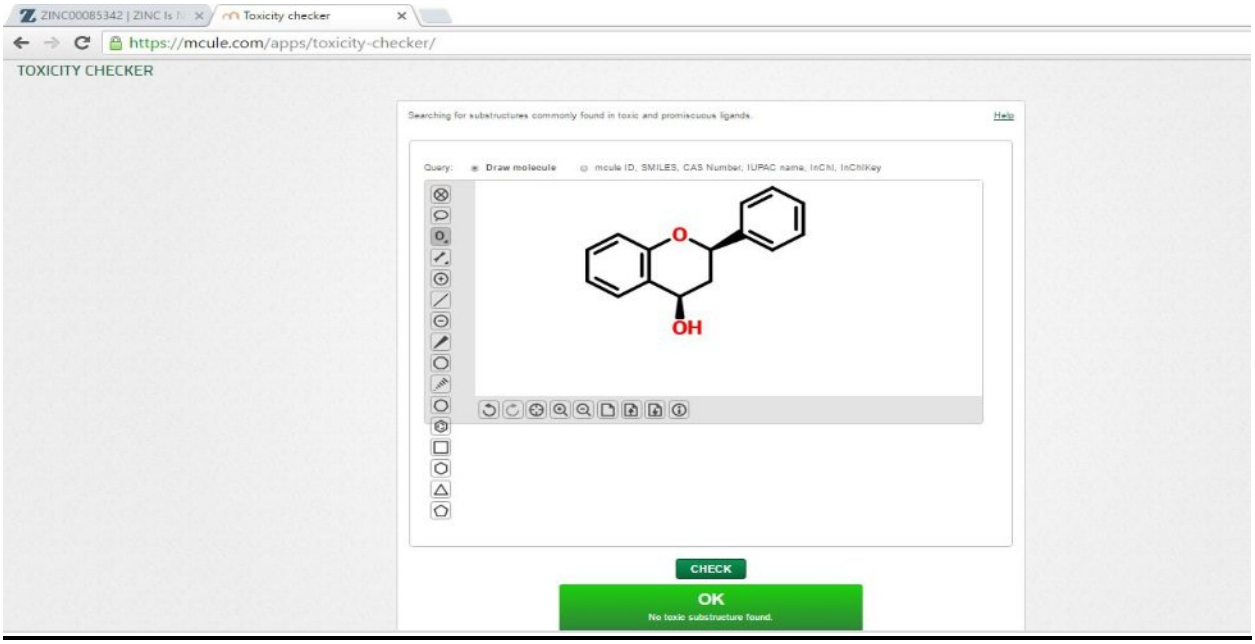


Fig23d.Toxicity checker result of ligand 4

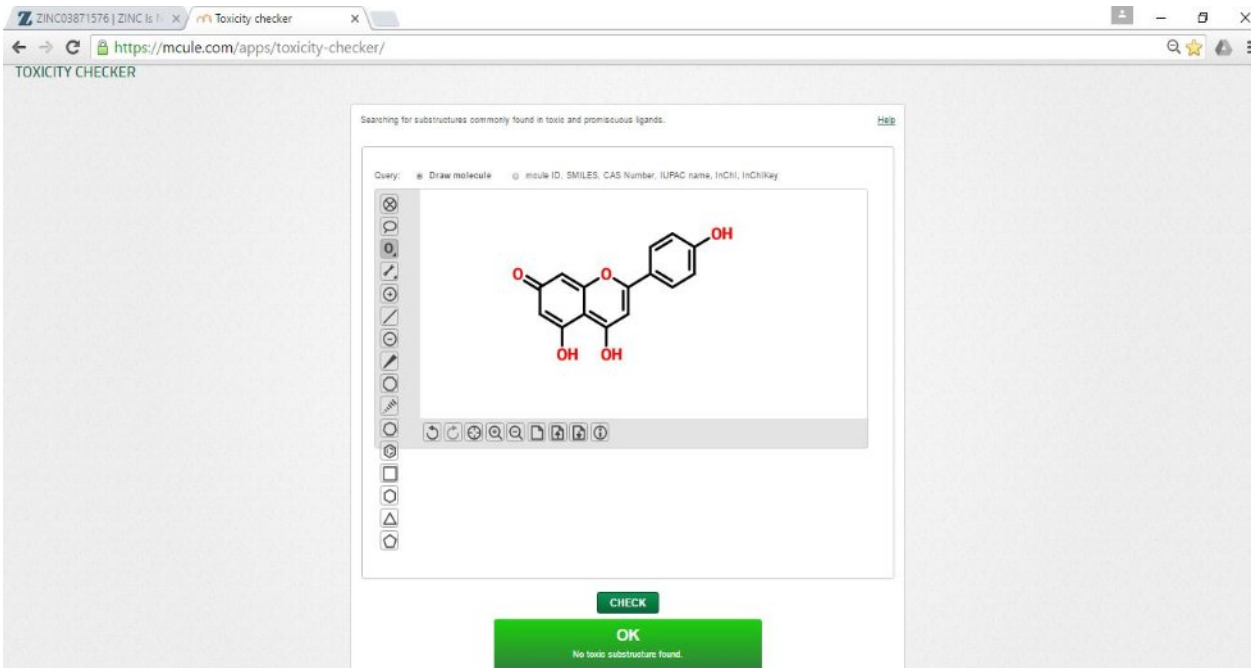


Fig23e.Toxicity checker result of ligand 5

.

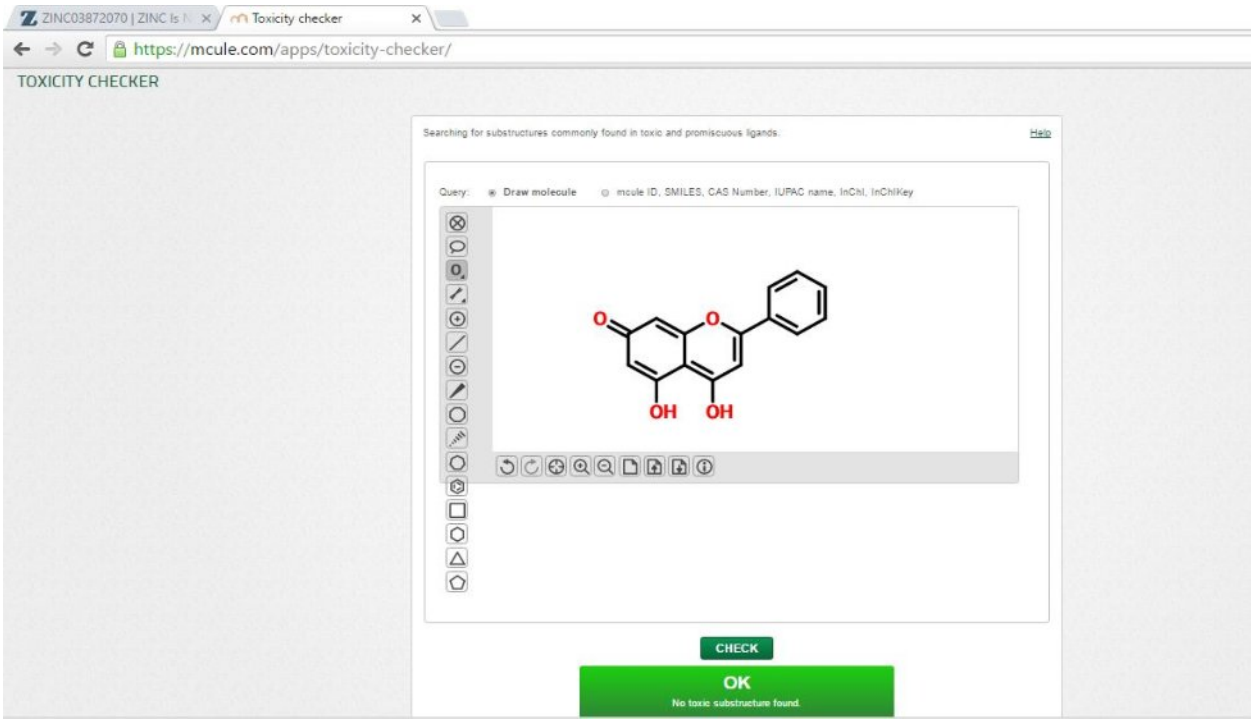


Fig23f.Toxicity checker result of ligand 6

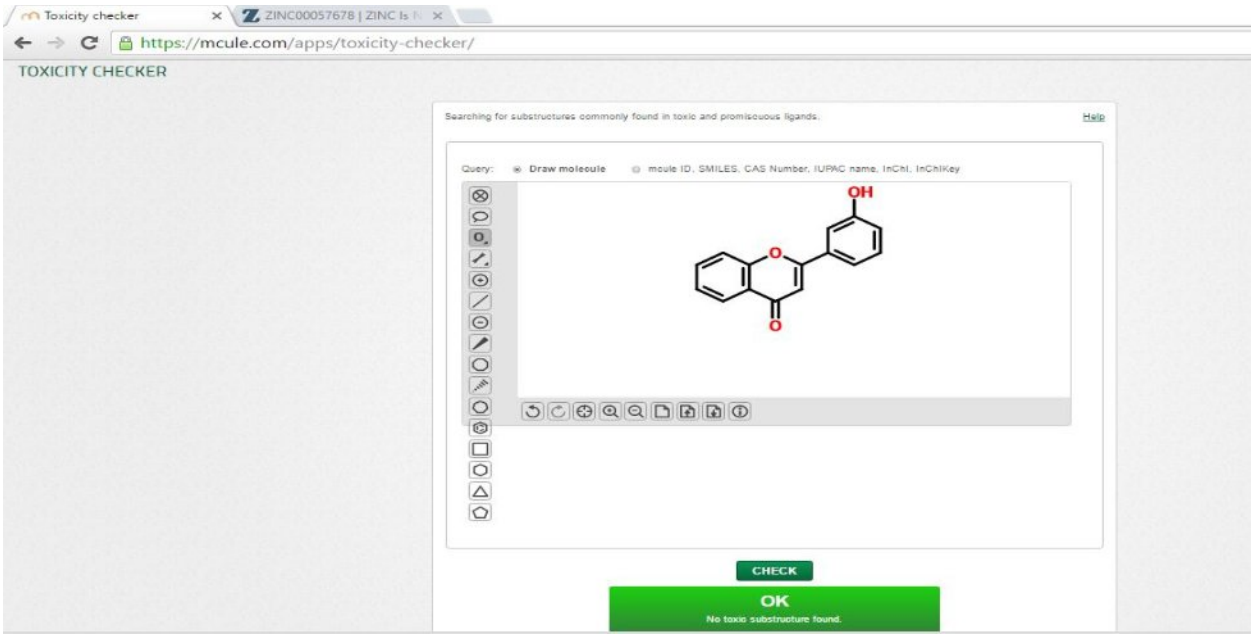


Fig23g.Toxicity checker result of ligand 7

.

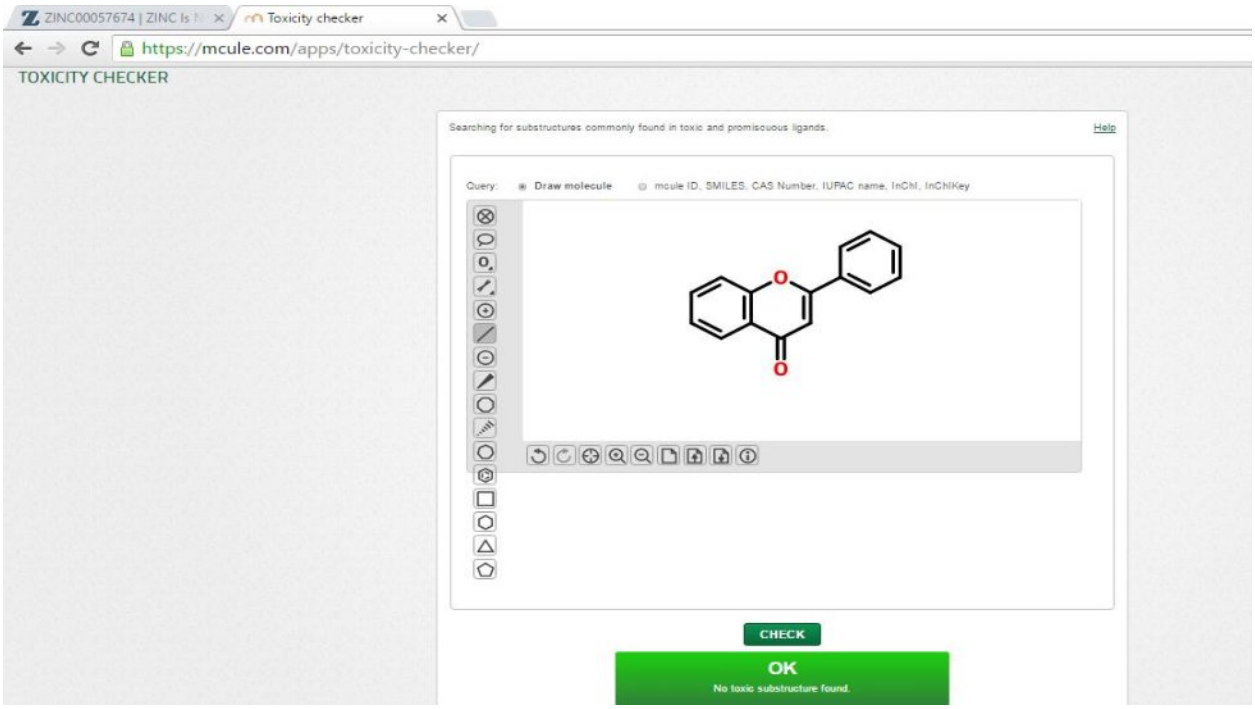
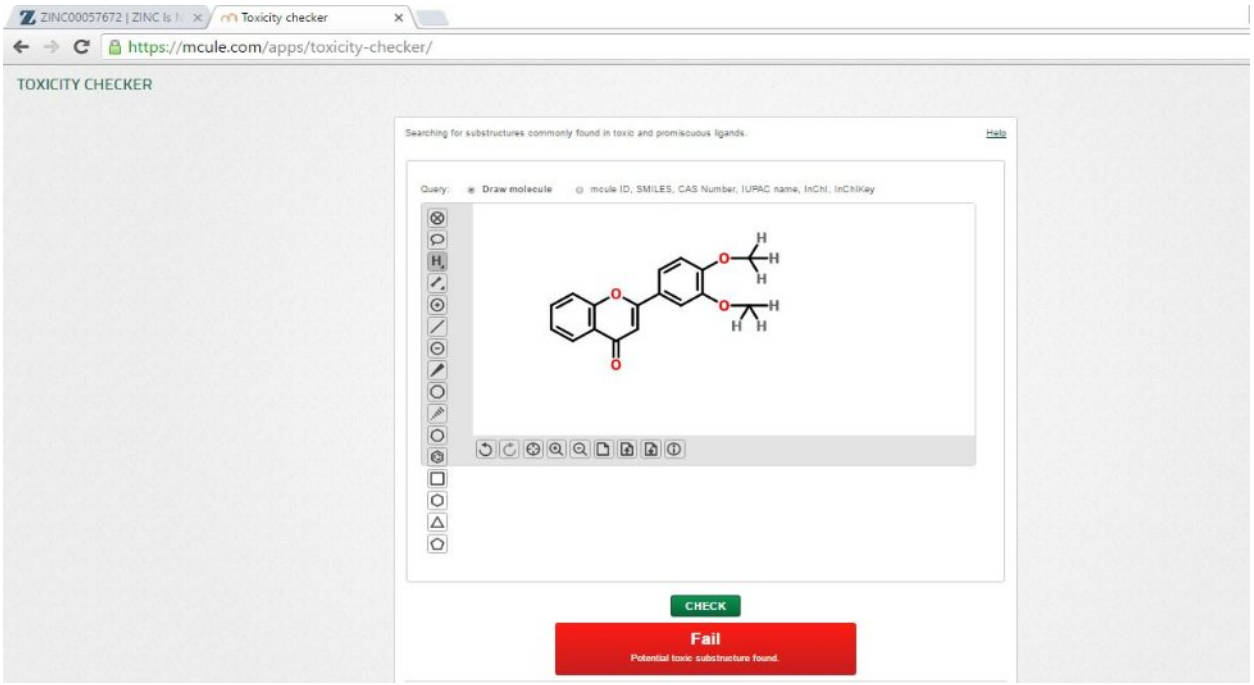


Fig23h.Toxicity checker result of ligand 8



.

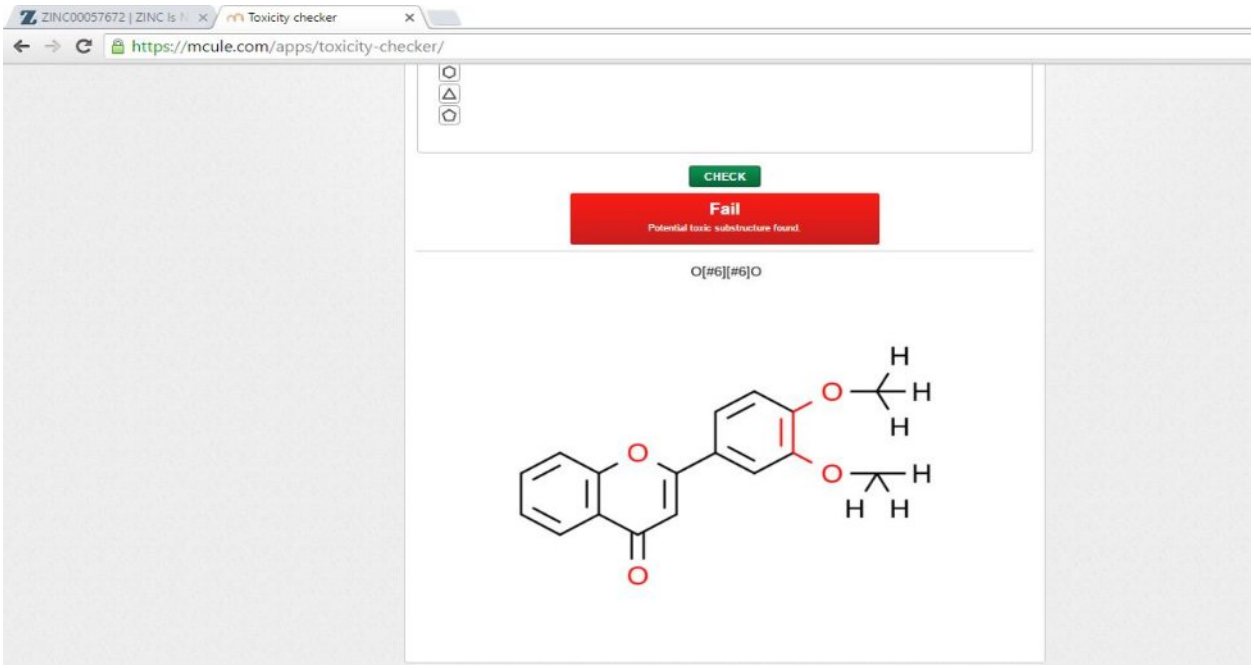


Fig23i.Toxicity checker result of ligand 9

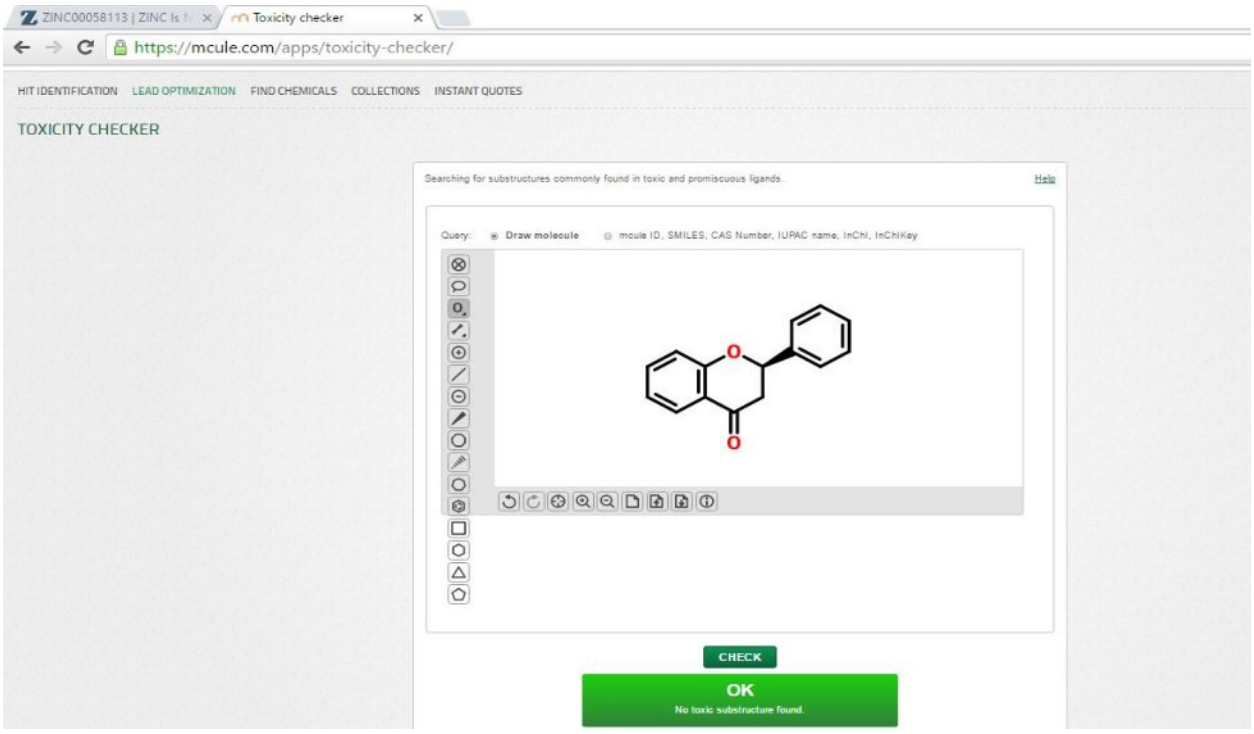


Fig23j.Toxicity checker result of ligand 10

.

b) Using OSIRIS data warrior software

Toxicity as well as various physicochemical properties was again checked by OSIRIS data warrior software for top 10 ligand molecules. These results are summarized in Table 3.

Ligand	zinc_59345 41	zinc_ 57657	zinc_ 18056	zinc_8 5342	zinc_38 71576	zinc_38 72070	zinc_57 678	zinc_57 674	zinc_57 672	zinc_58 113
Total mol. Wt.	238.24 1	254.24 0	252.26 8	226.27 4	270.239	254.240	238.241	222.242	282.294	224.258
CLogP	3.0271	2.6814	3.3028	3.0176	2.3357	2.6814	3.0271	3.3728	3.2328	3.1928
CLogS	-3.448	-3.152	-3.762	-2.799	-2.856	-3.152	-3.448	-3.744	-3.78	-3.528
H-Acceptors	3	4	3	2	5	4	3	2	4	2
H- Donars	1	2	0	1	3	2	1	0	0	0
Polar Surface Area	46.53	66.76	35.53	29.46	86.99	66.76	46.53	26.3	44.76	26.3
Drug likeness	0.2819 4	0.2819 4	0.3599 7	-0.0720 24	0.28194	0.28194	0.28194	0.24192	0.35997	-0.2709
Mutagenic	None	high	None	None	high	none	None	High	None	none
Tumorigenic	None	none	None	None	none	none	None	None	None	none
Reproductive Effect	None	none	None	None	none	none	None	None	None	none
Irritant	None	none	None	None	none	none	None	None	None	none
Non C/H Atoms	3	4	3	2	5	4	3	2	4	2
Stereo centers	0	0	0	2	0	0	0	0	0	1
Rotatable bonds	1	1	2	1	1	1	1	1	3	1
Rings	3	3	3	3	2	3	3	3	3	3
Aromatic Rings	2	2	2	2	3	2	2	2	2	2

Table3.Physicochemical properties and toxicity results for top 10 ligands

.

The knowledge of full genome and the application of subtractive approach have been very useful for drug target identification against many viral pathogens. The present work is an attempt to identify potential drug targets against the viral pathogen. Also the screened inhibitors are effective in contrast to multiple targets of Ebola cure, hence useful in blocking several signaling pathways simultaneously. Likewise VP35 protein obstructs post interferon initiation by association with TBK1-IKBKEDDX3 complex; hence hampering the active site of this protein by the screened flavonoids may obstruct the duplication and virulence factor of the EBOV. Since the multi-target binding affinity of offered ligands, these may be resolved as active inhibitors for Ebola Virus. The ligand is determined from PDB as hydroxy ion and docking of natural ligand with the target protein gives a cut-off value of binding energy for any ligand to be the drug molecule. Then based on the active sites present on our target protein, 100 ligands were identified as suitable ligands. Thus, virtual library of ligands were generated. And each ligand was separately docked with the target protein to determine top 10 ligands having better binding energy. And the best shows better binding energy than the natural ligand as well as available drugs. Then, various physicochemical properties and toxicity of the selected ligands were determined. Finally, we come up with ligand, 7-hydroxyflavone (ZINC ID 05934541) as the best ligand molecule and have the potential to be a drug molecule. So, in future one can go for clinical trials with the best ligand and study their different properties like pharmacodynamic, pharmacokinetics, solubility etc.

.

CONCLUSION

The whole genome analysis of the Ebola virus and use of database and tools like PDB, ZINC DATABASE, AUTODOCK TOOLS etc. can be helpful in the prediction of possible protein targets. And then use of docking software like AutoDock may lead to novel or optimized drugs against the Ebola virus. Ebola virus is a single-stranded, negative-sense RNA virus that leads to the acute hemorrhagic fever in humans and nonhuman primates. Here we present a computer aided prediction of target protein against the Ebola virus and then docked 100 ligands, identified based on active site of protein, to identify the best ligand which can be the possible drug molecule. Further, we check different physicochemical properties like Lipinski filters, etc. and toxicity was also checked to find out whether any harmful substructure is present or not in our ligand molecules. In this study, we come up top 20 ligands (based on their binding energy against target protein) which are having better binding energy than the available ligand as well as drug molecules, and also no toxic substructure is there in them. 7-hydroxyflavone (ZINC ID 05934541) is identified as the best ligand molecule in our study.

.

REFERENCES

Accelrys Software Inc.2013. Discovery Studio Modeling Environment, Release 4.0, San Diego: Accelrys Software Inc.

Adrian M et al. 2003. A current review of Ebola virus: pathogenesis, clinical presentation, and diagnostic assessment. Biol Res Nurs 4, 268

Adlercreutz, H.1995. Phytoestrogens: Epidemiology and a possible role in cancer protection. Environmental Health Perspectives, 103, 103–112.

Basler CF, Mikulasova A, Martinez-Sobrido L, Paragas J, M¨uhlberger E, Bray M, Klenk HD, Palese P, Garc´ıaSastre A. 2003. The Ebola Virus VP35 Protein Inhibits Activation of Interferon Regulatory Factor 3. J Virol 77, 7945-7956.

Costanzi, S., Tikhonova, I. G., Harden, T. K., & Jacobson, K. A. 2009. Ligand and structure-based methodologies for the prediction of the activity of G protein-coupled receptor ligands. Journal of Computer-Aided Molecular Design, 23, 747–754.

Cushnie TP, Lamb AJ. 2011. Recent advances in understanding the antibacterial properties of flavonoids. Int J Antimicrob Agents 38(2), 99-107.

Dixon MP, Pau RN, Howlett GJ, Dunstan DE, Sawyer WH, Davidson BE. 2002. Crystal structure of the Cterminal domain of Ebola virus VP30 reveals a role in transcription and nucleocapsid association. J Biol Chem 277, 23186-23192.

Dutta A, Singh SK, Ghosh P, Mukherjee R, Mitter S, Bandyopadhyay D 2006. *In silico* identification of potential therapeutic targets in the human pathogen *Helicobacter pylori*. In Silico Biol. 6:43–47.

Ebihara H, Takada A, Kobasa D, et al. 2006. Molecular determinants of Ebola virus virulence in mice. PLoS Pathog 2, e73.

Feig M, Onufriev A, Lee MS, Im W, Case DA, Brooks CL 2004. "Performance comparison of generalized born and Poisson methods in the calculation of electrostatic solvation energies for

.

protein structures". Journal of Computational Chemistry 25 (2): 265-84.

Geisbert TW et al. 1995. Differentiation of filoviruses by electron microscopy. Virus Res 39, 129.

Gomis R¨uth, Xavier F, et al. 2003. The matrix protein VP40 from Ebola virus octamerizes into pore like structures with specific RNA binding properties. Structure 11, 423.

Gonz´alez R, Ballester I, et.al. 2011. Effects of flavonoids and other polyphenols on inflammation. Crit Rev Food Sci Nutr 51(4), 331-62.

Goldman BB, Wipke WT 2000. "QSD quadratic shape descriptors. 2. Molecular docking using quadratic shape descriptors (QSDock)". Proteins 38(1): 79-94.

Hartlieb B, Modrof J, M¨uhlberger E, Klenk HD, Becker S. 2003. Crystal structure of the C-terminal domain of Ebola virus VP30 reveals a role in transcription and nucleocapsid association. J BiolChem 278, 41830-41836.

Hoenen T, Groseth A, Kolesnikova L, Theriault S, Ebihara H, Hartlieb B, Bamberg S, Feldmann H, Stroher U, Becker S. 2011. Crystal structure of the C-terminal domain of Ebola virus VP30 reveals a role in transcription and nucleocapsid association. J Virol 80, 72607264.

Huynen M, Diaz-Lazcoz Y, Bork P 1997. Differential genome display. Trend Genet.13:389–390.

Kitchen DB, Decornez H, Furr JR, Bajorath J 2004. "Docking and scoring in virtual screening for drug discovery: methods and applications". Nature reviews. Drug discovery 3(11): 935-49.

Lary Zeitlin et al. 2014. Reversion of advanced Ebola virus disease in nonhuman primates with ZMapp. Nature 514, 47-53.

.

M¨uhlberger E, Weik M, Volchkov VE, Klenk H-D, Becker S. 1999. Comparison of the transcription and replication strategies of Marburg virus and Ebola virus by using artificial replication systems. J Virol 73, 23332342.

Nanbo, Asuka, Watanabe, Shinji; Halfmann, Peter; Kawaoka, Yoshihiro. 2013. The spatio-temporal distribution dynamics of Ebola virus proteins and RNA in infected cells. Nature 10, 1038.

Rajamani R, Good AC. 2007. Ranking poses in structure-based lead discovery and optimization: current trends in scoring function development. Curr Opin Drug Discov Devel 10, 308-315.

Ruigrok RW et al. 2000. Structural characterization and membrane binding properties of the matrix protein VP40 of Ebola virus. J Mol Biol 300, 103.

Sina Bavari et al. 2014. Protection against filovirus diseases by a novel broad-spectrum nucleoside analogue BCX4430. Nature 508, 402-405.

Vilar, S., Cozza, G., & Moro, S. 2008. Medicinal chemistry and the molecular operating environment (MOE): Application of QSAR and molecular docking to drug discovery. Current Topics in Medicinal Chemistry, 8, 1555–1572.

Volchkov Victor E, et al. 2011. Conserved Proline-Rich Region of Ebola Virus Matrix Protein VP40 Is Essential for Plasma Membrane Targeting and Virus-Like Particle Release. J Infect Dis 204 (suppl 3), 884-891.