© 2016 IJSRST | Volume 2 | Issue 1 | Print ISSN: 2395-6011 | Online ISSN: 2395-602X
Themed Section: Science and Technology

Molecular Docking Studies of E-Bola Virus Protein VP30

***Uzma Paveen A. Shaikh, Yogesh N. Joshi**
Department of PG Studies and Research in Bioinformatics, Walchand Centre for Biotechnology,
Walchand College of Arts & Science, Solapur, Maharashtra, India

ABSTRACT

The Ebola virus (EBOV) genome encodes for several proteins that are necessary and sufficient for replication and transcription of the viral RNAs; NP, VP30, VP35, and L. VP30.VP30 binds to the RNA at the first gene start signal to initiate transcription. VP30 protein has playing important role in Ebola virus transcription and transcription reinitiation hence VP30 protein was targeted for the inhibition of Ebola virus. After target identification, Framycetin drug was taken from DrugBank database which is new lead and its derivatives were designed by bioinformatics virtual screening. Further, drug lead molecules were evaluated for their drug likeness using "Lipinski rule of five" and pharmacokinetic/ADMET properties. In molecular docking studies framycetin derivative shows the better binding energy with the target protein. This *in silico* approach can be appropriate to develop new drug lead molecules against Vp30 proteins Ebola virus infection.

Keywords: Ebola virus, VP30, framycetin, virtual screening, Molecular docking.

I. INTRODUCTION

Ebola viruses (EBOV) are non-segmented, negative RNA viruses, which together with Marburg virus constitute the family Filoviridae. Filoviruses cause severe and lethal hemorrhagic fevers in humans and nonhuman primates and, as such, are classified as biosafety level. Severe and often fatal hemorrhagic fever is the significant symptom of Ebola infection which happens in two stages; incubation period and late stage. Incubation period shows indications like joint pain, fever, weakness, sickness which can keep going for one week and late side effects incorporate sorrow, eye irritation, and hemorrhagic rash over the whole body [3]

A better understanding of the physical and functional interactions between Ebola virus proteins and cellular factors regulating the host innate immune response may reveal novel insights into the pathogenesis of Ebola virus and offer new strategies to inhibit Ebola virus replication[17]. The enveloped EBOV particle is composed of seven structural proteins, five of which form the helical nucleocapsid that represents the template for viral transcription and replication. The viral genome is encapsidated by the major nucleocapsid protein NP, and VP35, VP30, and VP24 interact with NP to form the mature nucleocapsid. In present study focused on VP30 protein as a potential drug target based on its function.

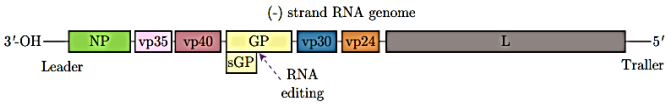

Figure 1 : Ebola Viral Genome Arrangement

Transcription is regulated by conserved transcription start and stop signal at the viral gene borders. The gene start signals are part of RNA secondary structures and it has been proposed that VP30 binds to the RNA at the first gene start signal to initiate transcription. In addition, VP30 was shown in Fig.1 which is important for transcription reinitiation of subsequent genes [6].

Phosphorylation of VP30 positively regulates and negatively regulates its transcriptional activity. In addition, enhancement of transcription by VP30 requires a putative RNA secondary structure located within

nucleotides (nt) 54 to 80 of the leader region. Deletion of the predicted RNA secondary structure permits VP30-independent transcription of viral messengers. These reported activities of VP30 suggest the possibility of a direct interaction of VP30 with EBOV RNA in its role in transcription. Recent publications with the mini genome system for EBOV suggest at least two possible mechanisms that VP30 may use in its transcriptional regulatory role. One possible mechanism could be that VP30 interacts with one or more of the other nucleocapsid proteins, polymerase, NP or VP35, and promotes increased stability of the transcriptional complex, VP30 may interact directly with viral RNA(s) to regulate transcription [14]. After entry into the host cell, the EBOV envelope fuses with host cell membranes to release the nucleocapsid into the cytoplasm where transcription and replication take place. Initial transcription of the newly entered encapsulated RNA genome is entirely accomplished by the nucleocapsid proteins that are associated with the intruding virus (primary transcription). Transcription is regulated by conserved transcription start and stop signals at the viral gene borders (Nadine Biedenkopfet al). VP30 represents an essential Ebola virus-specific transcription factor whose activity is regulated via its phosphorylation state. It has been hence conjectured that VP30 may help to beat this obstruction for transcriptional enactment, steady with its proposed part at a nearly phase of interpretation. VP30 due to its role in homooligomerization is considered as a potential target for antiviral treatment (Utkarsh Raj *et al).* Currently neither an approved vaccine nor antiviral therapy is available for humans.

According to latest reports, ZMapp, optimized combination of drug contains monoclonal antibodies made from a tobacco-plant strain can act as antiviral therapy against Ebola infection[1]. ZMapp is a cocktail combining the best components of two treatments namely MB-003 (Mapp) and ZMAb(Defyrus/PHAC) and is produced in a laboratory by exposing mice to fragments of the virus. Another drug BCX4430 (Developed byBioCryst), a new synthetic adenosine analogue, inhibits infection of different filoviruses in human cells[13]. Biochemical, reporter-based and primer-extension assays indicate that BCX4430 inhibits viral RNA polymerase function, acting as a non-obligate RNA chain terminator. Post-exposure intramuscular administration of BCX4430 protects against Ebola virus and Marburg virus disease in rodent models but it is not tested in humans yet.

In present study, we used *in silico* approach or discovering drug lead candidate against Ebola virus infection. The use of bioinformatics methods allows, using all aspects of drug discovery, forming core of structure based drug design and has advantage of delivering drug more quickly and at economic cost. Structure based drug designing approaches involves the 3-D structure of protein on which docking studies of various individual small molecules have been carried in order to calculate their docking score and binding energy by utilizing a series of scoring functions. The virtual screening & molecular docking of the drug candidates on target protein could find out the best lead like compounds with further optimization of the compounds to designing the lead (Rajamaniet al., 2007).

II. METHODS AND MATERIAL

1. Homology Modeling and Validation:

The sequence of VP30 protein was retrieved from the universal protein resource database (Uniprot) (UniprotKb ID: P35258). Secondary structure prediction was performed by using SOPMA (Geourjon and Deleage).The template for sequence alignment was identified through searching VP30 against PDB using the BLASTp[16].The 3D Crystal structure of the C-terminal domain of Ebola virus VP30 was downloaded from PDB (PDB ID: 3V7O) as the template structure. The homology model of VP30 was built by Swiss Model server. After modeling, the quality and validation of the model was evaluated by Ramachandran plot analysis using PDBsum server. The active site prediction of protein was predicted by using CASTp (Computed Atlas of Surface Topography of proteins) server [4].

2. Designing of Drug library :

Ligand molecule was downloaded from DrugBank database. Before docking ligand molecule was cleaned in 2D and 3D and saved it in mol2 chemical structure file format. After selecting the drug compound the drug library of compounds were generated using Marvin sketch software (version 5.8) and 2D structure of the all

compounds were subsequently converted in to 3D structures and saved in mol2 file format.

3. Virtual Screening by Lipinski's Rule of Five:

Molecular descriptors and drug likeliness properties of compounds were analyzed using the Molinspiration server with based on Lipinski's Rules . Molinspiration server supports for calculation of important molecular properties such as LogP, polar surface area, number of hydrogen bond donors and acceptors, as well as prediction of bioactivity score for the most important drug targets GPCR ligands, kinase inhibitors, ion channel modulators, enzymes and nuclear receptors[4].

4. Virtual Screening by ADMET Properties:

The pharmacokinetic properties such as Absorption, Distribution, Metabolism, Excretion and Toxicity of the compounds were predicted using admetSAR database. In admetSAR is a web based query tools incorporating a molecular build-in interface enable the database to be queried by SMILES and structural similarity search [4].

5. Molecular Docking Studies:

Based on the drug likeliness properties and pharmacokinetic properties 10 compounds were selected for docking studies. To validate drug- target association, the molecular docking was performed on active compound with screened compounds by Vlife MDS (Molecular Dynamics and Simulation) software (version 4.3) and we choose GA (Genetic Algorithm). All compounds and target protein were saved in mol2 file format before subjecting to software for docking [15].

III. RESULT AND DISCUSSION

1. Homology Modeling and Validation:

PDB id 3V7O (Crystal structure of the C terminus domain of Ebola virus) was selected as template with 37.50% sequence identity to query sequence. The quality and validation of the model was evaluated by Ramachandran plot analysis using PDBsum server. Ramachandran plot analysis showed that 2.7% in additional allowed region and 96.9% in favored region, indicating that the models were of reliable and good quality tools which shown in Fig. 1. The predicted model was visualized in Jmol visulazation tools which shown in Fig. 2. The predicted protein was subjected to CASTp server for pocket prediction. 53 pockets were predicted and docked all pockets which found that most repeating interacting residues were present in pocket ID 53 on the volume and area of this pocket is large as compare to other predicted pockets.

PROCHECK statistics

1. Ramachandran Plot statistics

		No. of residues	%-tage
Most favoured regions	[A,B,L]	219	96.9%
Additional allowed regions	[a,b,l,p]	6	2.7%
Generously allowed regions	[~a,~b,~l,~p]	1	0.4%
Disallowed regions	[XX]	0	0.0%
Non-glycine and non-proline residues		226	100.0%
End-residues (excl. Gly and Pro)		4	
Glycine residues		16	
Proline residues		4	
Total number of residues		250	

Figure 1: Procheck Analysis

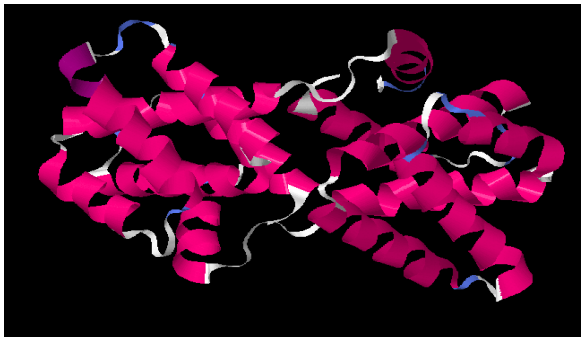

Figure 2: Visualization of VP30 in Jmol

2. Designing of Drug library :

Framycetin compound was selected from DrugBank as inhibitor for VP30 protein. Drug library of selected framycetin derivatives compound was designed in Marvin sketch software (version 5.8).Total 31 derivative compounds were generated.

3. Virtual Screening by Lipinski's rule of five:

Drug likeliness properties of compounds were predicted by Molinspiration server. The CLogP (octanol /water partition coefficient) was calculated by the methodology developed by Molinspiration as a sum of fragment based

contributions and correlation factors. The molecular descriptors of five compounds given in Table 1 were tested to Lipinski's rule of five. In that top ten compounds which showed drug likeness properties were selected for further analysis.

The ADMET (Absorption, Distribution, Metabolism, Excretion and Toxicity) properties of the target compounds were calculated using admetSAR. Blood-Brain Barrier (BBB) penetration, HIA (Human Intestinal Absorption), and AMES toxicity were calculated. The predicted ADMET data were summarized in Table 2. The cytochromeP450 super family plays an important role in drug metabolism and clearance in the liver.

4. **Virtual Screening by ADMET properties:**

Sr. NO	SMILE	Log P	Mol. Wt	H boud accp	H bound donar	Rot bound
1	NC4C(O)C(O)C(CN)OC4OC3C(N)CC(N)C(O)C3OC2OC(CO)C(OC1OC(CN)C(O)C(O)C1N)C2C#N	-6.0	623.66	18	19	9
2	NC4C(O)C(O)C(CN)OC4OC3C(N)CC(N)C(C#N)C3OC2OC(CO)C(OC1OC(CN)C(O)C(O)C1N)C2O	-6.0	623.66	18	19	9
3	NC4C(O)C(O)C(CN)OC4OC3C(N)CC(N)C(O)C3OC2OC(CC#N)C(OC1OC(CN)C(O)C(O)C1N)C2O	-6.1	623.66	18	19	9
4	NC4C(O)C(O)C(CN)OC4OC3C(N)CC(N)C(O)C3OC2OC(CO)C(OC1OC(CN)C(C#N)C(O)C1N)C2O	-6.03	623.66	18	19	9
5	NC4C(O)C(O)C(CN)OC4OC3C(N)CC(N)C(O)C3OC2OC(CO)C(OC1OC(CN)C(O)C(C#N)C1N)C2O	-6.03	623.66	18	19	9
6	NC4C(C#N)C(O)C(CN)OC4OC3C(N)CC(N)C(O)C3OC2OC(CO)C(OC1OC(CN)C(O)C(O)C1N)C2O	-6.03	623.66	18	19	9
7	NC4C(O)C(C#N)C(CN)OC4OC3C(N)CC(N)C(O)C3OC2OC(CO)C(OC1OC(CN)C(O)C(O)C1N)C2O	-6.03	623.66	18	19	9
8	[H+].NC4C(O)C(O)C(CN)OC4OC3C(N)CC(N)C(O)C3OC2OC(CO)C(OC1OC(CN)C(O)C(=[N+]=O)\C1N)C2O	-6.49	623.66	20	19	9
9	NC4C(O)C(O)C(CN)OC4OC3C(N)CC(N)C(O)C3OC2OC(CO)C(OC1OC(CN)C(O)C(CO)C1N)C2O	-6.09	628.68	19	19	9
10	NC4C(O)C(O)C(CN)OC4OC3C(N)CC(N)C(O)C3OC2OC(CO)C(OC1OC(CN)C(CO)C(O)C1N)C2O	-6.09	628.68	19	19	9
Framycetin	NC4C(O)C(O)C(CN)OC4OC3C(N)CC(N)C(O)C3OC2OC(CO)C(OC1OC(CN)C(O)C(O)C1N)C2O	-6.11	614.65	19	19	9

Table 1: Virtual screening of Framycetin derivatives by Lipinski's rule.

Sr. NO	SMILE	BBB	Carcinogenic	LD-50	CYP 450	AMES toxicity
1	NC4C(O)C(O)C(CN)OC4OC3C(N)CC(N)C(O)C3OC2OC(CO)C(OC1OC(CN)C(O)C(O)C1N)C2C#N	0.9708	Non carcinogenic	1.7003	0.6398	Non toxic
2	NC4C(O)C(O)C(CN)OC4OC3C(N)CC(N)C(C#N)C3OC2OC(CO)C(OC1OC(CN)C(O)C(O)C1N)C2O	0.9708	Non carcinogenic	1.7003	0.6398	Non toxic
3	NC4C(O)C(O)C(CN)OC4OC3C(N)CC(N)C(O)C3OC2OC(CC#N)C(OC1OC(CN)C(O)C(O)C1N)C2O	0.6953	Non carcinogenic	1.8101	0.6311	Non toxic
4	NC4C(O)C(O)C(CN)OC4OC3C(N)CC(N)C(O)C3OC2OC(CO)C(OC1OC(CN)C(C#N)C(O)C1N)C2O	0.9708	Non carcinogenic	1.7003	0.6390	Non toxic
5	NC4C(O)C(O)C(CN)OC4OC3C(N)CC(N)C(O)C3OC2OC(CO)C(OC1OC(CN)C(O)C(C#N)C1N)C2O	0.9708	Non carcinogenic	1.7003	0.6398	Non toxic
6	NC4C(C#N)C(O)C(CN)OC4OC3C(N)CC(N)C(O)C3OC2OC(CO)C(OC1OC(CN)C(O)C(O)C1N)C2O	0.9708	Non carcinogenic	1.7003	0.6398	Non toxic
7	NC4C(O)C(C#N)C(CN)OC4OC3C(N)CC(N)C(O)C3OC2OC(CO)C(OC1OC(CN)C(O)C(O)C1N)C2O	0.9708	Non carcinogenic	1.7003	0.6398	Non toxic
8	[H+].NC4C(O)C(O)C(CN)OC4OC3C(N)CC(N)C(O)C3OC2OC(CO)C(OC1OC(CN)C(O)C(=[N+]=O)\C1N)C2O	0.9920	Non carcinogenic	2.2020	0.5701	Non toxic

9	NC4C(O)C(O)C(CN)OC4OC3C(N)CC(N)C(O)C3OC2OC(CO)C(OC1OC(CN)C(O)C(CO)C1N)C2O	0.968	Non carcinogenic	1.5675	0.6458	Non toxic
10	NC4C(O)C(O)C(CN)OC4OC3C(N)CC(N)C(O)C3OC2OC(CO)C(OC1OC(CN)C(CO)C(O)C1N)C2O	0.9632	Non carcinogenic	1.7059	0.6393	Non toxic
Framycetin	NC4C(O)C(O)C(CN)OC4OC3C(N)CC(N)C(O)C3OC2OC(CO)C(OC1OC(CN)C(O)C(O)C1N)C2O	0.9659	Non carcinogenic	1.4850	0.6473	Non toxic

Table 2: Virtual screening by ADMET.

Molecule	Docking Score
Parent molecule	-2.4194
NC4C(O)C(O)C(CN)OC4OC3C(N)CC(N)C(O)C3OC2OC(CO)C(OC1OC(CN)C(O)C(CO)C1N)C2O	-2.4932

Table 3: Docking score.

5. Molecular Docking :

Molecular docking analysis were carried out using Biopredicta module of Vlife MDS 4.3.The receptor(VP30) was kept rigid, while ligands were set flexible to rotate and explore most probable binding poses. Software predicted cavities in the protein structure, first cavity was selected as binding pocket as it shows the similar residues as predicted in CASTp server pocket ID 53. The molecular docking complex which shows minimum binding energy was selected as shown in Table 3. The docking score or binding energy for the complex of VP30 and compound no 9((2R,3R,4S,5S,6S)-5-amino-2-(aminomethyl)-6-{[(1R,2R,3S,4R,6S)-4,6-diamino-2-{[(2R,3S,4S,5S)-4-{[(2R,3R,4S,5R,6R)-3-amino-6-(aminomethyl)-5-hydroxy-4-(hydroxymethyl)oxan-2-yl]oxy}-3-hydroxy-5-(hydroxymethyl)oxolan-2-yl]oxy}-3-hydroxycyclohexyl]oxy}oxane-3,4-diol) which shown in Fig. 3 was found out to be -1.3561 and the interacting residues Leu-151, Gyl-155, His-156, Ser-159, Gly-184 and Lys-187 were shown in Fig. 4.

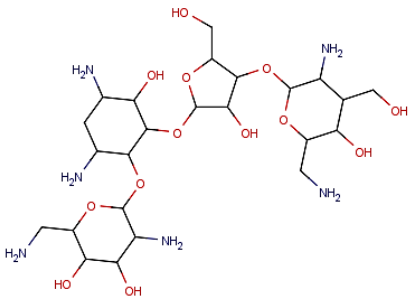

Figure 3: Chemical structure of compound 9 (2D representation).

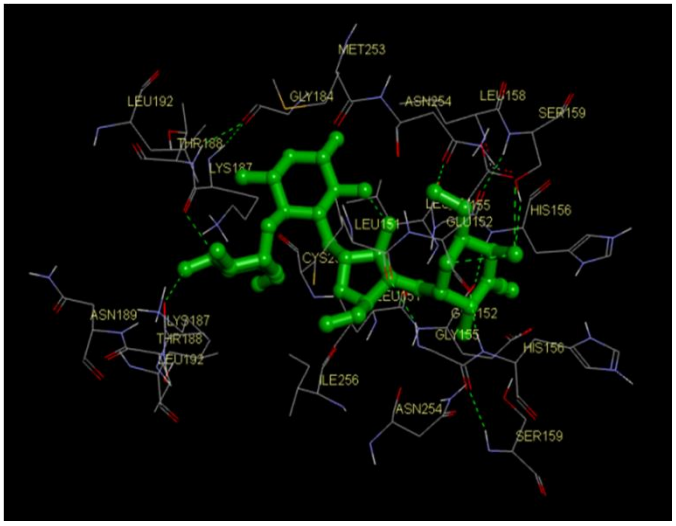

Figure 4: Molecular docking between the VP30 and compound 9 :(Binders are represented by green colored sticks and protein was represented by lines and colored according atom types.)

IV. CONCLUSION

The Ebola virus is most lethal disease. At present there is no potent drug available to treat Ebola virus. Molecular docking is one of the powerful technique for identifying biological significance and exploring new drugs by screening millions of compounds. The results of the current study clearly demonstrated that screened compounds of Framycetin are better inhibitors for viral protein (VP30)target as they interact with VP30 protein. The findings of this study are important as there is need for new drug to inhibit Ebola virus. The lead found out, could possibly inhibit the infection. However, the lead should undergo various preclinical analysis and optimization process before going into clinical trials.

V. REFERENCES

[1] Lary Zeitlin 2014 Reversion of advanced Ebolavirus disease in nonhuman primates with ZMapp, Nature 514, 47-53

[2] Uzma Parveen A Shaikh, Yogesh Joshi and Vinod P. S 2015 Homology modeling and binding site prediction of vp30 protein involved in Ebola virus, International Journal of Current Research, vol.7, issue, 12, pp.23947-23950.

[3] Utkarsh Raj and Pritish Kumar Varadwaj 2015 Flavonoids as Multi-target Inhibitors for Proteins Associated with Ebola Virus: insilico Discovery Using Virtual Screening and Molecular Docking Studies, Interdiscip SciComput Life Sci (2015) 7: 1–10.

[4] Sathish Kumar Paramashivam, KalaivaniElayaperumal, Boopala bhagavan Natarajan, Manjuladevi Ramamoorthy, Suganthana Balasubramanian & Kannan Narayanan Dhiraviam 2015 Insilico pharmacokinetic and molecular docking studies of small molecules derived from IndigoferaaspalathoidesVahl targeting receptor tyrosine kinases in Bioinformation 11(2): 073-084.

[5] Strahinja Z Kovacevic, Lidija R Jevric, Sanja O PodunavacKuzmanović and Eva S Lončar: 2014 Prediction of InsilicoADME Properties of 1,2-O-Isopropylidene Aldohexose Derivatives, Iranian Journal of Pharmaceutical Research, 13 (3): 899-907.

[6] Hartlieb B, Modrof J, M¨ uhlberger E, KlenkHD, Becker S 2003 Crystal structure of the C-termina ldomain of Ebola virus VP30 reveals a role in transcription and nucleocapsid association, JBiolChem 278,41830-41836.

[7] 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, 7260-7264 .

[8] Hoenen T, Jung S, Herwig A, Groseth A, Becker S 2010 Both matrix proteins of Ebola virus contribute to the regulation of viral genome replication and transcription, Virology 403, 56-66.

[9] Weik M, Modrof J, Klenk HD, Becker S, MuhlbergerE 2002 Crystal structure of the C-terminal domain ofEbola virus VP30 reveals a role in transcription andnucleocapsid association, J Virol 76, 8532-8539.

[10] Nadine Biedenkopf, BettinaHartlieb 2013 Phosphorylation of Ebola Virus VP30 Influences the Composition of the Viral Nucleocapsid Complex, the Journal of Biological Chemistry vol. 288, NO. 16, pp. 11165–11174.

[11] Miguel J. Martinez, Nadine Biedenkopf, Valentina Volchkova, Bettina Hartlieb, Nathalie Alazard-Dany, Olivier Reynard, Stephan Becker, and Viktor Volchkov 2008 Role of Ebola Virus VP30 in Transcription Reinitiation, in Journal Of virology, p. 12569–12573.

[12] Shanooba Palamthodi, DhirajPatil, Ashwini Sankpal, SehalZarekar, YatinPatil: 2012 Identification of Drug Lead Molecules against Ebola Virus: an Insilico Approach, J.Comput. Methods Mol., 2 (2):76-84.

[13] Sina Bavariet al. 2014 Protection against filovirus diseases by a novel broad-spectrum nucleoside analogueBCX4430, Nature 508, 402-405.

[14] John, S. P., T. Wang, S. Steffen, S. Longhi, C. S. Schmal john, and C. B.Jonsson 2007 Ebola virus VP30 is an RNA binding protein, J. Virol.81:8967–8976.

[15] Sujit Arun Desai, Santosh Sahedeo Kumbhar, Vishal Shankar Katti, Prafulla Balkrishna Choudhari, Manish Sudesh Bhatia 2013 3D QSAR and Pharmacophore Modelling on Chalcones as Anti leishmanial Agents potential Trypanothione reductase Inhibitors, Journal of Applied Pharmaceutical Science Vol. 3 (12), pp. 099-102.

[16] Jie Shen, Wenqian Zhang, Hong Fang, Roger Perkins, Weida Tong, Huixiao Hong 2013 Homology modeling, molecular docking, and molecular dynamics simulations elucidated-fetoprotein binding modes, Shenet al. BMC Bioinformatics,14(Suppl 14):S6.

[17] Atsushi Okumura, Paula M. Pitha, Akihiko Yoshimura, and Ronald N. Harty 2010 Interaction between ebola virus glycoprotein and host toll-like Receptor 4 leads to induction of proinflammatory Cytokines and socs1, journal of virology, p. 27–33