ASIAN JOURNAL OF PHARMACEUTICAL AND CLINICAL RESEARCH

INNOVARE ACADEMIC SCIENCES
Knowledge to Innovation

Vol 9, Issue 3, 2016

Online - 2455-3891
Print - 0974-2441

Research Article

COMPUTATIONAL STUDIES OF THE NUCLEOCAPSID PROTEIN OF HUMAN IMMUNODEFICIENCY VIRUS-1 TO FIND A POTENT INHIBITOR

GOPALUNI SAI AKASH*

Department of Biotechnology, Sreenidhi Institute of Science and Technology, Yamnampet, Ghatkesar - 501 301, Telangana, India.
Email: akashk896@yahoo.co.in

Received: 07 March 2016, Revised and Accepted: 19 March 2016

ABSTRACT

Objective: Viruses are obligate intracellular parasites of the host cellular mechanism. Their life cycle is dependent on the host. Human immunodeficiency virus (HIV) is a retrovirus which is considered a threat to humanity and a potential bioweapon. These studies have been conducted to make a potent inhibitor of this virus, but the rate of success is a very low as the virus is prone to mutations. The objective of this study is to find a potent inhibitor molecule for the protein nucleocapsid (NC) of HIV.

Methods: Computational studies like docking play a major role in finding the protein-ligand interactions. In this study, the crystal structure of the protein NC was subjected to high-throughput screening (HTS) against ZINC database to find potential inhibitors of the protein. The selected ligands were then screened for toxicity using OSIRIS property calculator and Molinspiration. Docking was performed using AutoDock Vina.

Results: HTS analysis provided top 200 hits from the clean fragment subset of ZINC database, out of which best 10 were chosen. These molecules were then screened for toxicity and violations of Lipinski's rule of 5. A total of five molecules had no toxic effects and best drug scores. These five compounds were the selected for docking analysis and it was found that the ligand "benzene-1, 2-disulfonamide" had the best binding affinity and interacted with crucial residues of the active site.

Conclusion: From the AutoDock Vina studies, the best pose was obtained with least energy value from which it can be hypothesized that the compound "benzene-1, 2-disulfonamide" can be considered as a potential inhibitor of the protein NC. Furthermore, wet lab studies have to be performed to find its efficiency and off-target activity.

Keywords: Human immunodeficiency virus-1, Nucleocapsid protein, Psi region, Stem-loop 3, High-throughput screening.

INTRODUCTION

Human immunodeficiency virus (HIV) belongs to the family *Retroviridae*, subfamily Orthoretroviriane, genus *Lentivirus* [1]. HIV has two species HIV-1 and HIV-2, of which HIV-1 is considered to be widespread and more pathogenic [2]. Like all other virus, HIV-1 requires the host to replicate. It replicated its genome and packs into a new virion using the genome packaging called "Psi region" and protein "nucleocapsid (NC) protein." This protein recognizes the stem-loop 3 (SL3) of Psi region and packs into a virion. The interactions between SL3 and NC are vital for the survival of the protein and hence a target for therapeutics.

Like all the retroviruses, HIV-1 produces a gag polyprotein, in the late stages of infection; the gag polyprotein is converted into NC, matrix protein (MA), capsid protein (CA). The function of the protein NC is to encapsulate the viral genome into a new virion [3,4]. Like most of the others of its genus, HIV-1 has 2 zinc fingers/Knuckle domain with CCHC confirmation. Where CCHC accounts for Cys-X2-Cys-X4-His-X4-Cys and X = variable amino acid residues [5]. These zinc knuckle domains play a major role in the recognition of the genome packaging signal Psi and encapsidation of the genome into a virion [6]. The protein NC has two zinc knuckles F1 is from Val 13 to Ala 30, F2 is from Gly35 to Glu51, and the linker segment is from Pro31 to Lys34. The residues from Lys3 to Arg10 form a helical structure and interact with the major groove of the SL3 in psi and the zinc knuckles interact with the exposed bases of the G9 of the RNA SL3. Similarly, the F2 interacts with the G7 of the RNA SL3. The protein NC in HIV-1 has around 14 sites where Arg or Lys are conserved and around 10 of them play a major role in the interactions between NC and SL3 (Fig. 1) [7].

The genome packaging signal of the HIV-1 virus is around 120 nucleotides long and is present between the 5' long translated region and start codon of the gag polyprotein [8]. Nucleus accessibility mapping and computational studies have shown that the Psi region has four SLs (SL1, SL2, SL3 and SL4, respectively). It has been observed that all the four SLs have a major role in the encapsidation of the viral genome. The SL1 plays a major role in the formation of the kissing-loop dimer which allows the virus to put two unspliced positive ssRNA into a virion. Whereas the functions of SL2 and SL4 is not that clearly understood [9]. HIV infection is considered a global threat but no effective drugs are discovered because of its high rates of mutations. Although inhibitors are made for protein, the mutation rate of the virus changes the protein structure, making the drug ineffective. Hence, inhibitors should be designed for conserved domains of the viral proteins. As some of the regions of the HIV-1 protein NC are conserved, and the same regions have a vital role in the encapsidation of the virus, these sites can be targeted with drugs that inhibit the interactions between NC and SL3.

The protein NC can be screened for a large number of ligands through high-throughput screening (HTS) of ZINC database. The ligand molecules obtained from this can be screened for their toxicity and docked against the protein NC using docking tools like AutoDock Vina. This study can find a potent inhibitor molecule that can stop the viral lifecycle and can be used as therapeutic against HIV-1 infections.

METHODS

Tools and materials

In this study, the crystal structure of the protein NC was retrieved from PDB databank, which has crystal structures of proteins [10]. Active site analysis of the protein was done using CASTp; this online server finds the possible active site pockets [11]. The server Dock Blaster was used to perform HTS against ZINC database [12]. Toxicity analysis of the ligands can be done through OSIRIS property calculator, which

reveals the toxic effects of the ligand such as mutagenic, tumorigenic, carcinogenic, and reproductive effects [13]. Molinspiration is an online tool used for QSAR analysis [14]. Docking studies can be done using AutoDock Vina [15]. The results of AutoDock Vina can be visualized using Accelrys Discovery Studio [16].

Preparation of protein

The crystal structure of the protein NC, 2M3Z was retrieved from RCSB Protein Databank. The structure of the protein was retrieved in the form of "Pdb." This structure was added with polar hydrogen atoms using Accelrys Discovery Studio and was converted into "Pdbqt" format for further docking analysis.

Active site analysis

The crystal structure of the protein was submitted to CASTp. CASTp finds pockets that can be possible active sites of the protein. This server also gives information about the active site residues.

HTS

The PDB structure of the protein was submitted to Dock Blaster, which screens the ZINC database for the best possible ligands for the query protein. It can also be used to screen a particular subset of the database like the natural products or the drug like subset. The clean-fragment database was chosen, which had 1611889 ligand molecules in it.

Toxicity analysis

The PDB structures of the top 10 hits of HTS were subjected to toxicity analysis using OSIRIS property calculator. It is an online server that reveals various properties of ligand molecules such as mutagenic, tumorigenic, irritant, reproductive effective, Clogp value, solubility, molecular weight, drug-likeness, and finally the drug score. Molinspiration is another online tool that gives information about the number of acceptor and donor atoms and violations from the Lipinski rule of 5. Based on these parameters ligands were chosen for docking studies.

Docking analysis

Non-toxic ligands with competitive drug scores were chosen for docking studies. The docking studies were performed by AutoDock Vina.

RESULTS AND DISCUSSION

Preparation of the protein and active site analysis

The crystal structure of the protein was retrieved from RCSB Protein DataBank. Polar hydrogen atoms were added to the protein, and this file was submitted to CASTp. This online server revealed 4 pockets in the protein. But only one of the pockets had residues interact with SL3, hence that pocket was selected as the active site for further experimentation (Fig. 2).

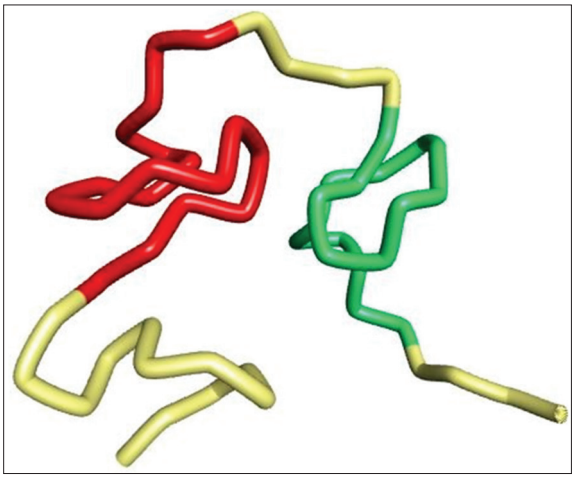

Fig. 1: The protein structure visualized using Accelrys Discovery Studio. F1 region colored red and F2 region colored green

HTS

The protein was then submitted to dock blaster for HTS against ZINC database. It predicted 2 active sites and one of the active sites predicted by this server was the same as the active site predicted by CASTp. The subset of ZINC database chosen for HTS was clean-fragments. This subset of the database contained 1611889 molecules. After HTS was performed, 200 hits of ligands were generated in the order of descending binding affinity with the protein. The top 10 ligands were chosen, for toxicity studies (Fig. 3).

Toxicity studies

The top 10 hits of HTS were submitted to OSIRIS property calculator and it was found that one of the 10 compounds had mild reproductive effects as shown in Table 1. The other nine compounds were screened based on drug score as shown in the Table 2 and top 5 molecules were selected for further analysis. Molinspiration revealed that none of the compounds violated Lipinski's rule of 5 as shown in Table 3, and hence, all the molecules were further considered for docking analysis.

Docking studies

Docking was performed between the crystal structure of the protein NC and the five ligands. The protein and ligand files were converted to "pdbqt" format, and grid parameters were set as per the result provided by CASTp. AutoDock Vina produces 9 poses of the ligand interacting with the protein, which has different binding affinities. The docked

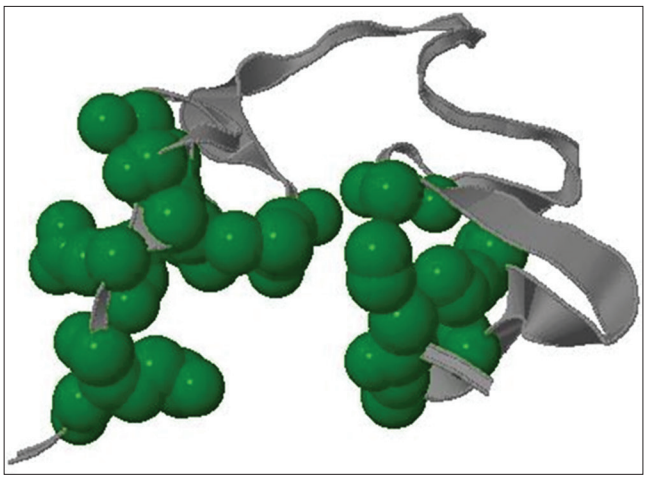

Fig. 2: A possible active site pocket predicted by CASTp. The pocket is colored green

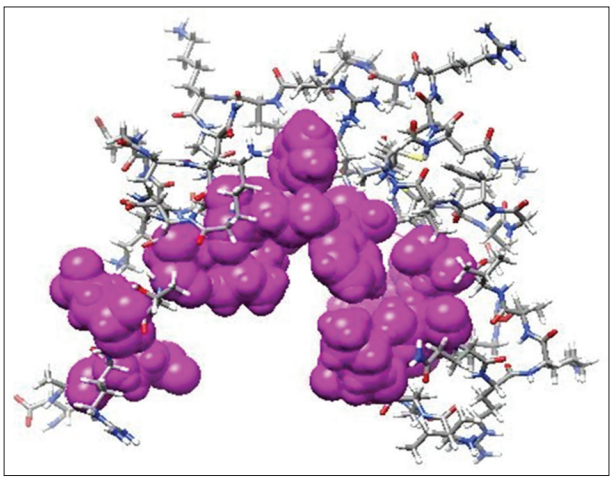

Fig. 3: One of the two active sites predicted by Dock Blaster. The same active site was considered for high-throughput screening

Table 1: Toxicity analysis of the top 10 ligands produced by Dock Blaster after HTS against clean-fragment subset of ZINC database

Ligand	Mutagenic	Tumorigenic	Irritant	Reproductive effect
N-(2-aminoethyl)pyrrolidine-1-sulfonamide	Normal	Normal	Normal	Normal
1-Dimethylsulfamoyl-pyrrolidine-2-carboxylic acid	Normal	Normal	Normal	Normal
N-[2-(methanesulfonamido)-1,1-dimethyl-ethyl]methanesulfonamide	Normal	Normal	Normal	Normal
1,2-bis(sulfamoylamino)ethane	Normal	Normal	Normal	Normal
3-sulfamoylpyridine-4-sulfonic acid	Normal	Normal	Normal	Normal
N-[2-[(5-cyano-2-pyridyl)amino]ethyl]methanesulfonamide	Normal	Normal	Normal	Normal
Benzene-1,2-disulfonamide	Normal	Normal	Normal	Normal
2-(N-benzylmethylsulfoamido)acetic acid	Normal	Normal	Normal	Slightly toxic
2-(N-benzylmethylsulfoamido)acetic acid	Normal	Normal	Normal	Normal
N-[2-(3,5-dimethylphenoxy)ethyl]methanesulfonamide	Normal	Normal	Normal	Normal

The molecule that is rejected due to toxic affects has its toxic effects in italic. HTS: High-throughput screening

Table 2: Drug score and other vital parameters analysis of the nine non-toxic ligand molecules

Ligand	cLogP	Solubility	Molecular weight	TPSA	Drug likeliness	Drug score
N-(2-aminoethyl)pyrrolidine-1-sulfonamide	−0.88	−0.51	193	83.81	0.11	0.75
1-dimethylsulfamoyl-pyrrolidine-2-carboxylic acid	−0.82	−0.08	222	86.3	0.88	0.83
N-[2-(methanesulfonamido)-1,1-dimethyl-ethyl] methanesulfonamide	−0.36	−1.95	244	109.1	3.02	0.93
1,2-bis(sulfamoylamino)ethane	−0.63	−3.6	210	92.86	0.55	0.72
3-sulfamoylpyridine-4-sulfonic acid	−2.81	0.67	238	144.1	−1.18	0.6
N-[2-[(5-cyano-2-pyridyl)amino]ethyl] methanesulfonamide	0.02	−2.62	226	71.35	1.1	0.82
Benzene-1,2-disulfonamide	−0.81	−1.41	236	137	0.23	0.75
2-(N-benzylmethylsulfoamido)acetic acid	−0.93	−0.26	201	88.28	−1.07	0.62
N-[2-(3,5-dimethylphenoxy)ethyl]methanesulfonamide	1.65	−2.86	243	63.78	−2.93	0.48

Molecules selected for docking have their parameter values in italic

Table 3: Molinspiration property values of compounds with best drug scores

Ligand	H-acceptors	H-donors	Violations
N-(2-aminoethyl) pyrrolidine-1-sulfonamide	5	4	0
1-dimethylsulfamoyl-pyrrolidine-2-carboxylic acid	6	1	0
N-[2-(methanesulfonamido)-1,1-dimethyl-ethyl]methanesulfonamide	6	2	0
N-[2-[(5-cyano-2-pyridyl) amino]ethyl]methanesulfonamide	6	2	0
Benzene-1,2-disulfonamide	6	4	0

Table 4: Docking results of the protein NC and the 5 ligands using AutoDock Vina

Ligand	Binding affinity (Kcal/mol)
N-(2-aminoethyl)pyrrolidine-1-sulfonamide	−4.3
1-dimethylsulfamoyl-pyrrolidine-2-carboxylic acid	−4.3
N-[2-(methanesulfonamido)-1,1-dimethyl-ethyl] methanesulfonamide	−3.9
N-[2-[(5-cyano-2-pyridyl)amino]ethyl] methanesulfonamide	−5.0
Benzene-1,2-disulfonamide	−5.3

NC: Nucleocapsid

ligands were again screened based on the best binding affinity to find the best inhibitor of the protein. "Benzene-1, 2-disulfonamide" is the ligand that has the best binding affinity of −5.3 Kcal/mol affinity of −5.3 Kcal/mol as shown in Table 4.

The ligand was found to be interacting with 4 of the active site residues Phe 6, Lys 14, Cys 15, and Phe 16. The ligand formed a classical hydrogen bond with Cys 15, Pi hydrophobic bond with Phe 6 and Phe 16, mixed pi/alkyl hydrophobic bond with Lys 14 and sulfur of the ligand interacted with Phe 6 by sulfur interactions (Fig. 4). Phe 6 belongs to the helix that interacts with the major groove of SL3. Lys 14, Cys 15, and Phe 16 belong to zinc knuckle 1 (F1) that interact with G9 of SL3. These interactions show that the ligand is binding to functional part of

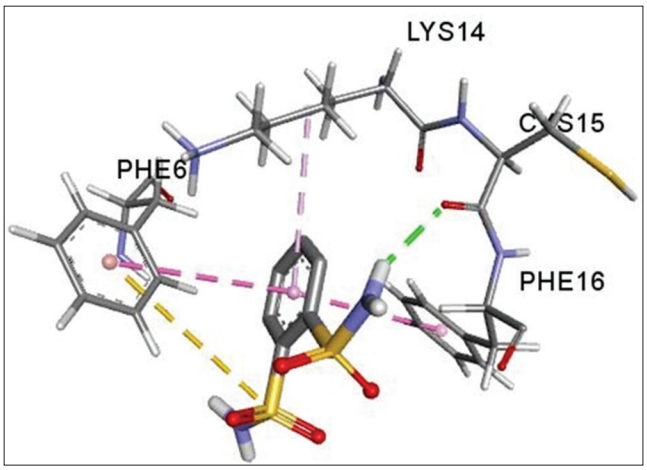

Fig. 4: Interactions between the active site residues of the protein nucleocapsid and the ligand "benzene-1, 2-disulfonamide"

the protein and can inhibit the efficient binding of the NC to SL3, which will lead to the genome of HIV-1 unpacked into a new virion. This can help decrease the number of viral budding and propagation of the virus.

Classical hydrogen bonds are shown in green, Pi hydrophobic bonds shown in pink, mixed pi/alkyl hydrophobic bond is shown in light pink and sulfur interactions shown in golden color.

CONCLUSION

The aim of the study was to find an inhibitor for the protein NC, which would interfere with the recognition of the genome packaging signal of HIV-1 viruses. HTS was done against ZINC database to find possible inhibitors of this protein. HTS produced 200 hits of which best 10 were selected for toxicity analysis using OSIRIS property calculator and Molinspiration. This analysis revealed five ligand molecules with no toxic effects, competitive drug score and obeyed Lipinski's rule of 5. These 5 molecules were selected for docking analysis using AutoDock Vina.

Based on the docking studies using AutoDock Vina, it was found that the ligand "benzene-1, 2-disulfonamide" is the best inhibitor of the protein NC with a binding affinity of −5.3 KCal/mol. Upon visualizing the results through Accelrys Discovery Studio, it was found that the ligand was binding to 4 residues Phe 6, Lys 14, Cys15 and Phe 16, all of which belong to domains that play a major role in the packaging of the viral genome. From these results, if can be concluded that the ligand "benzene-1, 2-disulfonamide" is a potent inhibitor of the protein NC. However, further wet lab studies have to be performed to find the efficiency and off-target interactions of the ligand.

ACKNOWLEDGMENT

I thank the faculty of Department Biotechnology, SNIST for their support and encouragement.

REFERENCES

1. International Committee for Taxonomy of Viruses. Available from: http://www.ictvonline.org/virusTaxonomy.asp. [Last accessed on 2016 Mar 07].
2. Gilbert PB, McKeague IW, Eisen G, Mullins C, Guéye-NDiaye A, Mboup S, *et al.* Comparison of HIV-1 and HIV-2 infectivity from a prospective cohort study in Senegal. Stat Med 2003;22(4):573-93.
3. Dickson C, Eisenman R, Fan R, Hunter E, Teich N. In: RNA Tumor Viruses. 2nd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 1984.
4. Mervis RJ, Ahmad N, Lillehoj EP, Raum MG, Salazar FH, Chan HW, *et al.* The gag gene products of human immunodeficiency virus Type 1: Alignment within the gag open reading frame, identification of posttranslational modifications, and evidence for alternative gag precursors. J Virol 1988;62(11):3993-4002.
5. Berg JM. Potential metal-binding domains in nucleic acid binding proteins. Science 1986;232(4749):485-7.
6. Gorelick RJ, Henderson LE, Hanser JP, Rein A. Point mutants of Moloney murine leukemia virus that fail to package viral RNA: Evidence for specific RNA recognition by a "zinc finger-like" protein sequence. Proc Natl Acad Sci U S A 1988;85(22):8420-4.
7. De Guzman RN, Wu ZR, Stalling CC, Pappalardo L, Borer PN, Summers MF. Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element. Science 1998;279(5349):384-8.
8. Lever AM, Gottlinger HG, Haseltine WA, Sodroski JG. Identification of a sequence required for efficient packaging of human immunodeficiency virus type 1 RNA into virions. J Virol 1989;63:4085-7.
9. Moore MD, Hu WS. HIV-1 RNA dimerization: It takes two to tango. AIDS Rev 2009;11:91-102.
10. Berman HM. The Protein Data Bank: A historical perspective. Acta Crystallogr A 2008;64:88-95.
11. Dundas J, Ouyang Z, Tseng J, Binkowski A, Turpaz Y, Liang J. CASTp: Computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. Nucleic Acids Res 2006;34:W116-8.
12. Irwin JJ, Shoichet BK, Mysinger MM, Huang N, Colizzi F, Wassam P, *et al.* Automated docking screens: A feasibility study. J Med Chem 2009;52(18):5712-20.
13. OSIRIS Property Explorer. Available from: http://www.openmolecules.org/propertyexplorer/index.html. [Last accessed on 2016 Mar 07].
14. Molinspiration Property Calculation Services. Available from: http://www.molinspiration.com/cgi-bin/properties. [Last accessed on 2016 Mar 07].
15. Trott O, Olson AJ. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem 2010;31(2):455-61.
16. Ref. Dassault Systèmes BIOVIA, Discovery Studio Modeling Environment, Release 4.5. San Diego: Dassault Systèmes; 2015.