Journal of Pharmaceutical Chemistry
http://www.vensel.org/index.php/jphchem
VENSEL

High throughput virtual screening based discovery of dengue protease inhibitor

Sounak Bagchi, Uzma Alia, Faiz Mohammad, *Mohd Usman Mohd Siddique

Department of Pharmaceutical Sciences & Technology, Birla Institute of Technology, Mesra, Ranchi-835215, India

Abstract: High throughput virtual screening (HTVS) has been proved a successful tool for getting LEADs in drug design and discovery. In an attempt to design new Dengue protease inhibitors, we performed HTVS using Zinc13 database containing 13,195,609 drug-like molecules. ZINC42678127 was identified as potential HIT against Dengue protease. It's shape and electrostatic complimentary was found to be 0.608 and 0.078, respectively. Qikprop analysis of the compound complied with the Rule of Five (Ro5) and other drug-likeliness properties. Binding mode analysis of docked conformer of ZINC42678127, displayed favorable interaction with the active site residues of DENV protease. The identified HIT has a potential to become a LEAD against Dengue protease.

Key words: Dengue, Serine protease, NS2B/NS3, HTVS, Binding mode analysis

Introduction:

Dengue is a mosquito-borne viral infection triggering an acute flu like illness, sometimes initiating potentially lethal impediment known as severe dengue[1].It is a fast progressing pandemic-prone viral infection affecting millions of people worldwide. *Aedesa aegypti* is the prime vector that transmits the viruses and eventually leads to Dengue. It is caused through bites of an infected female *Aedes* mosquito, which primarily obtains the virus while sucking blood from an infected person. Over a term of 8-12 days, the virus infects the mid-gut of the mosquito and subsequently spreads to the salivary glands. The virus can be imparted to human beings via feeding, after the incubation period gets over[2]. Dengue virus belongs to the family Flaviviridae and has four sero-types, DENV1,[3] DENV2, DENV3[4] and DENV4[5].An estimated 50-100 million infection reported every year, Dengue alone presents an intimidating warning to 2.5 billion people round the globe[6-7]. Infection with Dengue virus lead to less severe Dengue Fever (DF) to lethal Dengue Hemorrhagic Fever (DHF) or Dengue Shock Syndrome (DSS). The virus consists of an enveloped (+ss)RNA of 10.7kb that encodes a precursor polyprotein[1-2, 8-9]. Polyprotein comprising of three structural (C, prM, M) and seven non-structural proteins (NS1,NS2A,NS2B,NS3,NS4A,NS4B and NS5) is further processed by host proteases and NS2B-NS3 viral protease, a major characteristic of the viral life cycle and replication performance. Consequently, the protease inhibition can be a successful way to fight against DENV infection[10-14]. Located in the N-terminal region, NS3 is a multifunctional protein that comprises of 180 amino acids of NS3 protease. The enzymatic functions of nucleoside triphosphate, helicase and RNA 5'-triphosphate are performed in the C-terminal region. The hydrophilic part of the integral membrane protein NS2B is required by the protease for the complete enzymatic activity[15]. Efforts in designing peptidomimetics to inhibit DENV protease has the major challenge due to the dibasic preface at P1 and P2 positions adjacent to the cleavage site. Hence, getting a potent peptidomimetic inhibitor with favorable pharmacokinetic profile is still on.[14]

Availability of X-ray crystallographic structure of DENV protease provided an opportunity to design the inhibitors through structure based drug design (SBDD) concept. High-Throughput Virtual Screening (HTVS) of small molecule libraries were reported against DENV NS2B-NS3 protease in identifying the LEAD compounds. Here we are reporting the HTVS based on shape and electrostatic similarity with the compound reported by Timiri *et al.*[16] as tool to enrich the huge database and subsequent molecular docking to identify the HITs. Binding mode analysis of the top-scoring molecule with DENV protease has also been carried out to establish the interaction at molecular level.

2. Results and discussion

For similarity search based on the shape and electrostatic property of this reference molecule, clean Drug-like subset (comprising of 13,195,609 molecules) from the publicly accessible database, ZINC13 was considered. ROCS and EON module of OpenEye Tools were used to executed the shape and electrostatic based screening. Top 15,000 molecules with best tanimato-coefficient were considered for further molecular docking studies. The top-scoring HIT was then checked

Submitted on: Aug 07, 2017 | Accepted on: Aug 09, 2017
*corresponding author: MU Tel: +91-8809792075; E-mail: palladiumsalt@gmail.com

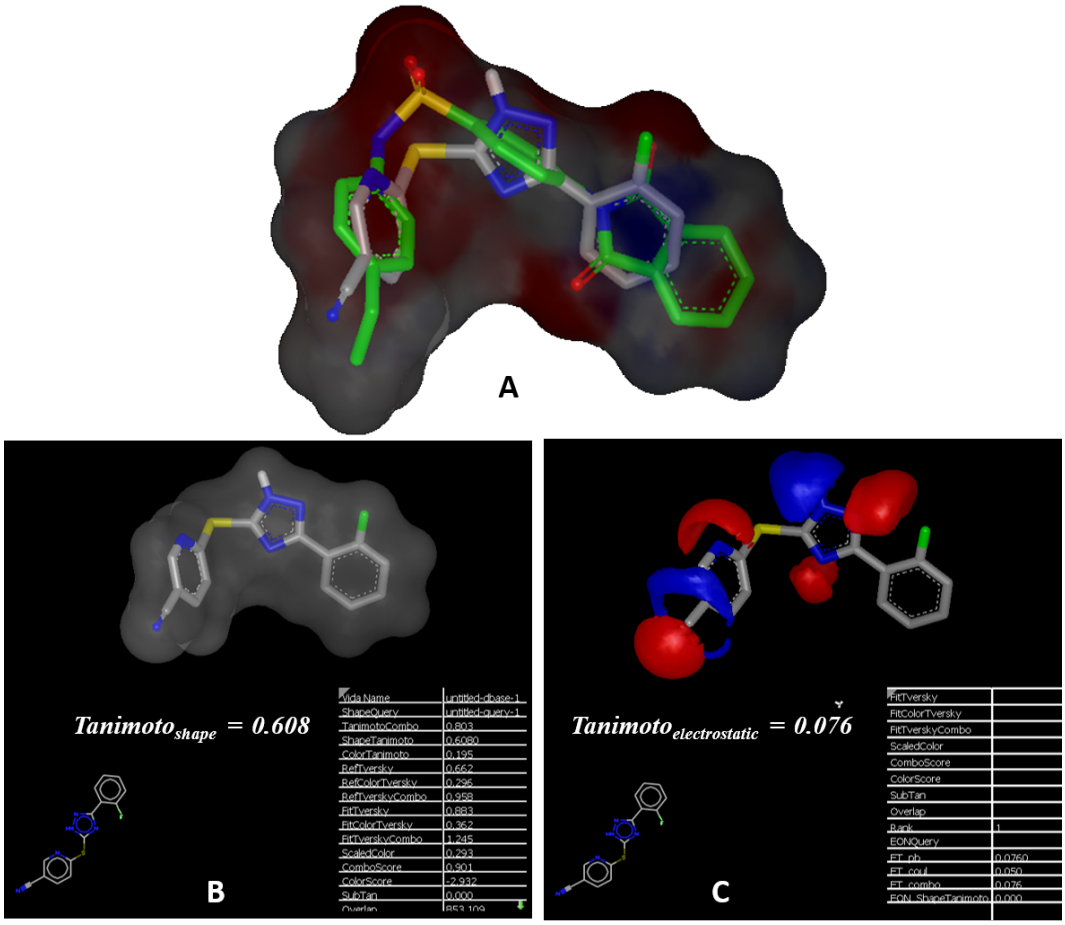

Figure 1. Overlapped structure of reference molecule (green) and ZINC42678127 (grey); **(B)** display of shape potential of ZINC42678127 with reference and **(C)** display of electrostatic potentials of ZINC42678127 with reference.

for ADME properties and subjected to Site Map studies.

2.1. Shape and electrostatic studies

Comparative shape and electrostatic study provided good correlation between reference molecule (reported by Timiri *et al.*)[16] and ZINC42678127. The molecule displayed *Tanimoto*$_{shape}$ *of* 0.608 and *Tanimoto*$_{electrostatic}$ *of* 0.076, inspite of different chemical scaffold. Figure 1A. depicted the complete overlapping of reference molecule and ZINC42678127, whereas Figure 1B. and Figure 1C. denotes the shape & electrostatic co-efficient, respectively with reference molecule. Top 15000 molecules from similarity based enrichment were then selected for docking studies and top hit was considered for assessment of their binding with the protein.

2.2. Molecular docking studies

Molecule that are having structure complementary to the catalytic site of the receptor will be identified through molecular docking simulation. The enriched database was subjected to Glide docking with extra precision (XP) mode aginst DENV NS2B-NS3 protease. Molecule with Zinc ID ZINC42678127, the one has exhibited the best similarity score, showcased the best docking score. Figure 2. represents the 3D interactive diagram of reference and ZINC42678127 with the active site of DENV NS2B-NS3 Protease.

Binding mode analysis:

The 2D interactive diagram of reference ligand Figure 3A. displayed one backbone H-bond interactions with Asn152. While identified HIT (Figure 3B.) formed polar H- bonding with Phe130 and Tyr150. The first H-bond appeared as backbone and later was side chain. Moreover the ZINC42678127 molecule also displayed hydrophobic interaction with Hip51 in similar fashion as reference molecule. Both the molecules displayed the hydrophobic interactions with amino acids Phe130, Tyr150, Pro132 and Leu128. Similarly, polar interactions with Ser131 and Ser135 residues present in the active site of DENV NS2B-NS3 protease.

Site map: Recognizing ligand binding site, hydrophobic, H-Bond donor and H-bond acceptor site was aided by active site mapping of reference as well as ZINC42678127 in the receptor as demonstrated in Figure 4. These factors were used to find out the druggability of the ligand molecules. The scoring functions of site, site map properties are mentioned in Table 1. It was found that ZINC42678127 had site score as well as D-score very similar to our reference molecule. Thus, for the treatment of Dengue disease, it can be regarded as druggable agent for inhibition of viral serine protease. This similarity in scores is attributed to the degree of resemblance in shape and electrostatic properties of the two, which allowed added advantage to virtual-screening method.

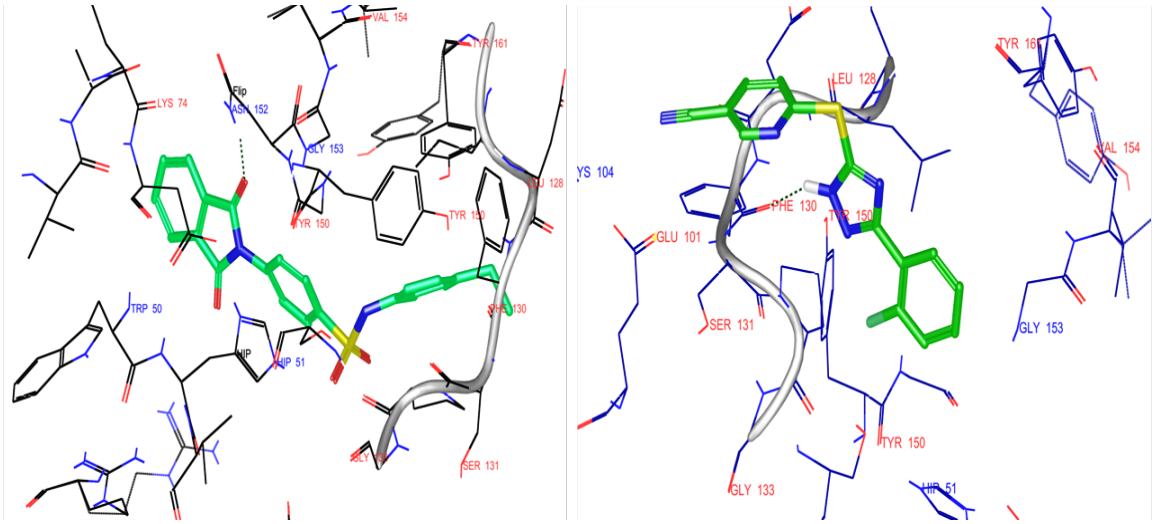

Figure 2. (A) Interaction of reference molecule with active site residues of DENV NS2B-NS3 protease (B) interactions ZINC42678127 with active site residues of DENV NS2B-NS3 protease (PDB ID 2FOM).

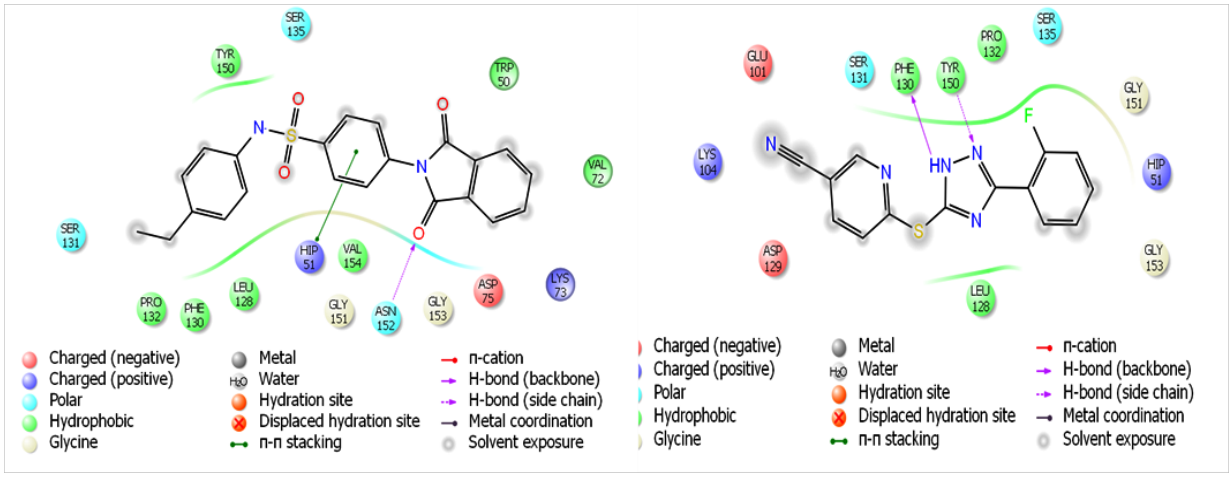

Figure 3. 2D interactive diagram (A) Interaction of reference molecule with active site residues of DENV NS2B-NS3 protease (B) interactions ZINC42678127 with active site residues of DENV NS2B-NS3 protease (PDB ID 2FOM)

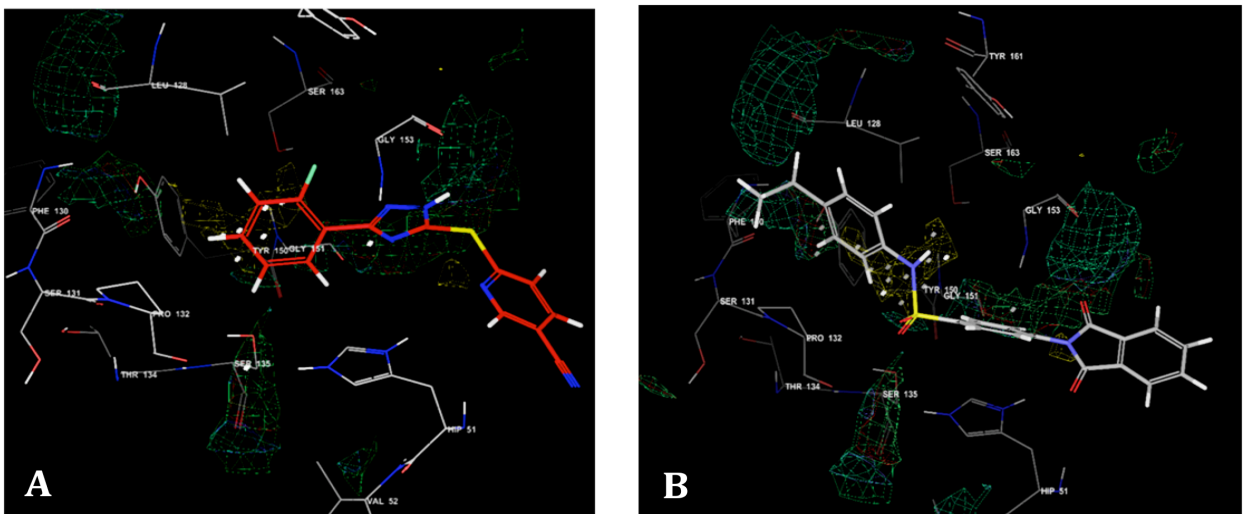

Figure 4. Binding mode analysis of reference molecule (A); identified hit ZINC42678127 (B). The various crucial interactions are blue color shows H-Bond donor, red color denotes H Bond acceptor, green color represents hydrophilic interactionand Yellow color indicates hydrophobic interaction

Table1. SiteMap Property Values and Dscore Ranks reference molecule and identified potential hit ZINC42678127

Entry	DScore	SiteScore	Size	Enclosure	Exposure	hydrophilic	hydrophobic
Reference	0.522	0.545	14	0.540	0.827	0.472	1.071
ZINC42678127	0.440	0.492	8	0.578	0.871	0.530	1.028

2.3 ADME calculations

To check the drug likeness properties of molecule, In-silico ADME calculation was performed. The percentage of oral absorption, dipole moment, and Plog BB were inspected. For anymolecule to be orally active, (i) themolecular weight (MW) less than 500D, (ii) hydrogen bond acceptor (HBA) not exceeding 10, (iii) hydrogen bond donor (HBD) groups less than 5 and (iv) oil/water partition coefficient (logP o/w) below 5 has been suggested by Lipinski's rule-of-five (Ro5) .Qikprop module in Maestro-8.5 (Schrodinger LLC) has been used to calculate these parameters for the reference &HIT. The results obtained were tabulated in Table 2. Both molecules (reference and top hit) were found to complywith the Ro5 as described by Lipinski. The CNS toxicity was found out by blood-brain coefficient (Plog BB), and the outcomes indicate that an identified molecule have values within the range and proves that ZINC42678127 cannot cross the blood brain barrier to produce any CNS toxicity.

3. Experimental

Material and methods: Dengue protease inhibitor reported from our laboratory, 4-(1,3-dioxoisoindoin-2-yl)-N-(4-ethylphenyl)-benzenesulfonamide[16], has been selected as a reference molecule and ZINC13 database (comprising of 13,195,609 molecules)[17] was selected for similarity search. Shape and electrostatic evaluation was performed using VROCS (3.1.2) and EON (2.1.0) of OpenEye toolkit (OEChem)[18]. Ligand preparation, protein preparation, receptor grid generation, molecular docking and site map analysis were performed using appropriate modules from Maestro-8.5 (Schrodinger LLC),[19,20] Qikprop module in Maestro-8.5 has been used to check the Ro5 violations. All the simulations were carried out in DELL Precision T3400 machine (n-series, Intel core 2 Quad processor, 8GB RAM, 500GB and running on RHEL5 operating system.

3.1. Virtual Screening

We used the novel compound reported by Timiri *et al.*[16] for HTVS screening, exhibiting an IC_{50} value of 48.2 µM. On the basis of shape and electrostatic studies using VROCS (3.1.2) and EON (2.1.0) interface of OpenEye toolkit, respectively and via ZINC13 database along with default parameters and molecules were filtered and High Throughput Virtual Screening (HTVS) was carried out from compound libraries. QikProp utility for drug likeliness properties of these molecules was used for further screening of ADME properties and Lipinski Rule of Five.

3.1.1. Shape and Electrostatic Study

Standard and significant virtual screening protocol for ligand based computer aided drug design demands for shape and electrostatic based screening. The logic in contrasting shape properties is that it takes into consideration molecules having shape similarity by rapid flexible superimposition of several similar molecules so that they can fit well and bind to the active pocket of protein or enzyme more efficiently.[21] Formation of non-covalent bonds between oppositely charged moieties is enabled by force involving electrostatic interactions. Relationship between polarity of the ligand and its relative biological activity is examined by electrostatic parameter.[22] Electrostatic parameters as well as shape give an idea regarding physiochemical properties of recognition of active molecules. ROCS algorithm (Rapid Overlay of Chemical Structures) and EON of OpenEye Toolkit was applied for this function. In ROCS analysis molecules were filtered on the basis of their shape agreement with that of reported DENV NS2B-NS3 protease inhibitor, output from the ROCS was then subjected to electrostatic correlations using EON. The results were inspected on the basis of $Tanimoto_{shape}$ and $Tanimoto_{electrostatic}$ coefficient

3.2. Protein Preparation

Crystal structure of DENV NS2B-NS3 protease (PDB ID: 2FOM) was retrieved from RCSB protein data bank (http://www.rcsb.org). To evade unphysical interaction, the raw protein was prepared via Protein preparation wizard. Chain A was selected for simulation. Hydrogens were added, water molecules were removed, bond order and formal charges were assigned for protein, ligands and co-factors. Hetero state for co-crystallised ligand was generated at pH 7 ± 0.3 through Epik. Hydrogen bonding network is optimised through ProtAssign. Finally the protein was minimized using OPLS2005 Force field.

3.3. Receptor Grid Generation

Receptor Grid Generation wizard of Glide module facilitated the generation of receptor grid along with default parameters: scaling factor (1.0) and partial charge cut-off (0.25) without any force. Grid box having the dimension of 20x20x20 Å has been constructed by picking ligand. Centre of the box defined as centroid of the co-crystalline ligand.

Table 2: In silico calculation of ADME properties of reference molecule and identified potential hit ZINC42678127

Compound ID	Mol weight	Glide score	DonorHB	AcceptorHB	Log P	Human Oral Absorption	PSA
Reference	406.455	-5.516	1	7.5	3.292	90.039	104.65
ZINC42678127	297.309	-6.516	1	5	2.59	88.882	80.02

3.4. Docking

LigPrep utility has been used to prepare ligands for docking with default parameters. A maximum of 32 tautomers and stereo isomers were produced and minimization was done through OPLS 2005 Force field and ligands were desalted. By utilizing GLIDE software,the prepared ligands were docked into the active site of the receptor in XP mode. The vdW scaling factor of 0.8 and the potential charge cut-off of 0.15 were applied respectively but the core and constraints were not applied. The ligand-docking process was followed by the execution of post dock minimization. Under grid space of 12.0 Å, the command was given to include 5 poses per ligands and to note interactions of residue and writes per residue interaction scores for residues within 12.0 Å of the grid centre with the number of poses per ligand being 5. The result was declared with best-scoring poses first and followed by examining with Maestro's Pose Viewer utility and XP Visualizer. The docking protocol was validated by obtaining the internal ligand and re-docking into the active site of the receptor.

3.5. Site map analysis[23]

For the purpose of identification of single binding site region prediction of druggability of the ligand, Site Map a Schrödinger tool is used. Site Map takes into account the following parameters like site point, hydrophilicity, hydrophobicity, H-bond donor, H-bond acceptor, enclosure and exposure to calculate the Site score (SiteScore) and Druggability Score (D-score) of ligand molecule.

MAP_{POD} is the parameter used to access druggability. It can be calculated as:

$$\Delta \mathrm{GMAP}_{POD} = 300\text{Å}^2 f_{non-polar}\gamma(\alpha)/(1 - 1.4/r)$$

Where,
$f_{non-polar}$ = Non polar part of SAS
(SAS: solvent accessible surface area)
$\gamma(\alpha)$= 45 cal/mol/Å^2
r = radius of curvature determined for the site

It is model for accessing the maximum affinity that can be achieved by passively absorbed oral drug for a given target.

Following parameters are considered while scoring the site.

3.5.1. Number of Site Point

Site point can be explained as the collection of points in a 1 Å grid box, either continuous or in short gaps present is the exposed region of the receptor. Therefore, 2-3 site points are recommended for each atom of bonded ligand involving hydrogen.

3.5.2. Hydrophobic and hydrophilic site

The information regarding Shape, nature of hydrophobic and hydrophilic region is provided by a site map assessment. Regions that are neither hydrophobic nor hydrophilic are also noticed as it points out areas where it might be possible to modify the ligand to increase it's physical properties. (E.g. By modifying its solubility without compromising its binding affinity). The average value of hydrophobic and hydrophilic score for submicromolar site is 1.0. A potent drug molecule should have reasonably higher hydrophobicity and lower hydrophilicity, bringing the average ratio to 1.6.

3.5.3. Exposure and Enclosure

The measure of the degree of exposure of the site to the solvent is described by these properties. For calculation of the exposure property, "extension" sites points are included to 1- Å site-point grid. Open site allows collection of additional extension point, whereas deep and well closed site, contributes to higher exposure score. Lower exposure score and higher enclosure score of the site is recommended. Regular value of exposure and enclosure scor for submicromolar site is 0.52 and 0.76, respectively.

3.5.4. Donor/Acceptor

As reproduced from the size and strengths of donor and acceptor Site Map Regions, this factor calculates the capability of well-structured ligand to donate, rather than to accept hydrogen bonds. The normal value for submicromolar site is 0.76.

3.5.5. Contacts

This factor inspects the strength of vdW non-bonded interaction of average site points with the receptor. The normal value for contact score is 1.0.

3.5.6. Site score

To determine the druggability and assess binding sites, Site score is generally used. It depends on factors like hydrophilic score, number of site point and enclosure score. It supports in determination of target site.

$$\text{Site Score} = 0.0733\mathrm{n}^{1/2} + 0.6688e^{-0.020}p$$

Where,
n = square-root of no. of site points (up to 100)
e = enclosure score and
p = hydrophilic score (up to 1.0)

The average value of Site Score for submicromolar site is 1.01.

3.5.7. Druggability score (D score)

It involves provisions that supports ligand bindings, adequate size, and isolation from solvent but forestalls them with a term that forbids the increase of hydrophilicity as shown in the equation:

$$\text{D Score} = 0.094\mathrm{n}^{1/2} + 0.60e^{-0.324}p$$

Where,
n = number of site points capped at 100
e = degree of enclosure
p = hydrophilic score

Druggabililty of a ligand is directly proportional to its *Dscore*; and explained as druggable, undruggable and challenging sites having high, low and intermediate *Dscore*, respectively.

Undruggable sites form covalent bonds and are minor in size, shallow and highly hydrophilic with minuscule hydrophobicity. To be administered as pro-drug, challenging sites are appropriately hydrophilic and less

hydrophobic, while druggable site are of acceptable size, enclosure properties and have exceptional hydrophobicity.

3.5.8. Site Volume

It is important to determine the volume of site surrounded by protein surface, devoid of contact with solvent. The "Shrink-wrap" volume is determined in this step.

Conclusion

In summary, shape and electrostatic similarity based HTVS has been applied to enrich the huge database like ZINC13. Enriched database has been further subjected to molecular docking simulation to identify the potential HITs. The current study identified a HIT that has the potential to become the LEAD. Further *in vitro* screening of ZINC42678127 for its antiviral activity against DENV has to be carried out to establish the effectiveness of the virtual screening protocol employed in identifying the LEADs quickly.

Acknowledgment

Mohd Usman acknowledges UGC for award of MANF JRF fellowship (201516-MANF-2015-17-MAH-60712).

References

1. Gubler, D. J. In *Dengue/dengue haemorrhagic fever: history and current status*, Novartis foundation symposium, Chichester; New York; John Wiley; 1999: 2006; p 3.
2. WHO report on global surveillance of epidemic-prone infectious diseases. 2000.
3. Austin, S. K.; Dowd, K. A.; Shrestha, B.; Nelson, C. A.; Edeling, M. A.; Johnson, S.; Pierson, T. C.; Diamond, M. S.; Fremont, D. H., Structural basis of differential neutralization of DENV-1 genotypes by an antibody that recognizes a cryptic epitope. PLoS Pathog 2012**,** 8 (10), e1002930.
4. Lok, S.-M.; Kostyuchenko, V.; Nybakken, G. E.; Holdaway, H. A.; Battisti, A. J.; Sukupolvi-Petty, S.; Sedlak, D.; Fremont, D. H.; Chipman, P. R.; Roehrig, J. T., Binding of a neutralizing antibody to dengue virus alters the arrangement of surface glycoproteins. Nature Struct Mol Biol 2008, 15 (3), 312-317.
5. Alcaraz-Estrada, S. L.; del Angel, R.; Padmanabhan, R., Construction of self-replicating subgenomic dengue virus 4 (DENV4) replicon. Dengue: Methods and Protocols 2014, 131-150.
6. Bhatt, S.; Gething, P. W.; Brady, O. J.; Messina, J. P.; Farlow, A. W.; Moyes, C. L.; Drake, J. M.; Brownstein, J. S.; Hoen, A. G.; Sankoh, O., The global distribution and burden of dengue. Nature 2013, 496 (7446), 504.
7. Murray, N. E. A.; Quam, M. B.; Wilder-Smith, A., Epidemiology of dengue: past, present and future prospects. Clin Epidemiol 2013**,** 5, 299.
8. Fields, B.; Knipe, D.; Howley, P.; Griffin, D., Fields virology, Lippincott Williams & Wilkins. Philadelphia, Pa 2001.
9. Yang, C.-C.; Hsieh, Y.-C.; Lee, S.-J.; Wu, S.-H.; Liao, C.-L.; Tsao, C.-H.; Chao, Y.-S.; Chern, J.-H.; Wu, C.-P.; Yueh, A., Novel dengue virus-specific NS2B/NS3 protease inhibitor, BP2109, discovered by a high-throughput screening assay. Antimicrob Agents Chemother 2011, 55 (1), 229-238.
10. Nitsche, C.; Behnam, M. A.; Steuer, C.; Klein, C. D., Retro peptide-hybrids as selective inhibitors of the Dengue virus NS2B-NS3 protease. Antiviral Res 2012, 94 (1), 72-79.
11. Rodenhuis-Zybert, I. A.; Wilschut, J.; Smit, J. M., Dengue virus life cycle: viral and host factors modulating infectivity. Cell Mol Life Sci 2010**,** 67 (16), 2773-2786.
12. Hanley, K. A.; Weaver, S. C., *Frontiers in Dengue Virus Research*. Horizon Scientific Press: 2010.
13. Tomlinson, S. M.; Watowich, S. J., Anthracene-based inhibitors of dengue virus NS2B–NS3 protease. Antiviral Res 2011, 89 (2), 127-135.
14. Gao, Y.; Cui, T.; Lam, Y., Synthesis and disulfide bond connectivity–activity studies of a kalata B1-inspired cyclopeptide against dengue NS2B–NS3 protease. Bioorg Med Chem 2010, 18 (3), 1331-1336.
15. Steuer, C.; Gege, C.; Fischl, W.; Heinonen, K. H.; Bartenschlager, R.; Klein, C. D., Synthesis and biological evaluation of α-ketoamides as inhibitors of the Dengue virus protease with antiviral activity in cell-culture. Bioorg Med Chem 2011, 19 (13), 4067-4074.
16. Timiri, A. K.; Selvarasu, S.; Kesherwani, M.; Vijayan, V.; Sinha, B. N.; Devadasan, V.; Jayaprakash, V., Synthesis and molecular modelling studies of novel sulphonamide derivatives as dengue virus 2 protease inhibitors. Bioorg Chem 2015, 62, 74-82.
17. Irwin, J. J.; Shoichet, B. K., ZINC− a free database of commercially available compounds for virtual screening. J Chem Inf Model 2005, 45 (1), 177-182.
18. OEChem, T., OpenEye Scientific Software. Inc., Santa Fe, NM, USA 2012.
19. Maestro, S., Version 9.2. Schrodinger LLC, New York 2011.
20. Schrodinger Software Suite. Schrödinger LLC, New York 2011.
21. Sastry, G. M.; Dixon, S. L.; Sherman, W., Rapid shape-based ligand alignment and virtual screening method based on atom/feature-pair similarities and volume overlap scoring. J Chem Inf Model 2011, 51 (10), 2455-2466.
22. Radhakrishnan, M. L.; Tidor, B., Specificity in molecular design: a physical framework for probing the determinants of binding specificity and promiscuity in a biological environment. J Phys Chemy B 2007, 111 (47), 13419-13435.
23. Halgren, T., New Method for Fast and Accurate Binding-site Identification and Analysis. Chem Biol Drug Des 2007, 69 (2), 146-148.