



## Virtual Screening for Identification of Small Lead Compound Inhibitors of Nipah Virus Attachment Glycoprotein

Md Hazrat Ali<sup>1\*</sup>, Saeed Anwar<sup>1</sup>, Pradip Kumar Roy<sup>1</sup> and Md Ashrafuzzaman<sup>1,2</sup>

<sup>1</sup>Department of Genetic Engineering and Biotechnology, School of Life Science, Shahjalal University of Science and Technology, Sylhet 3114, Bangladesh

<sup>2</sup>Resource Conservation (INRES), University of Bonn, Karlobert-Kreiten-Str. 13, D-53115 Bonn, Germany

### Abstract

Nipah virus (NiV), a newly emergent zoonotic paramyxovirus, has caused several outbreaks in humans and associated with severe encephalitic diseases. Till these days, neither vaccines nor drugs with optimal appeasement against the virus are available. The attachment glycoprotein (NiV-G) on the surface of the virus is an important virulent factor and a promising antiviral target. To identify novel inhibitors of NiV-G using computer aided virtual screening of NCI diversity set 2 and 20,000 commercially available drug-like compounds in the ZINC database. Structure based molecular docking studies using the crystal structure of the NiV-G were performed to virtually screen for novel inhibitors of NiV-G and 4 potential compounds with potential ability to inhibit the NiV-G by competing with Ephrin binding site and prevent NiV encephalitis by blocking the Ephrin recognition zone at the peripheral site were found.

**Keywords:** Virtual screening; Nipah virus; Attachment glycoprotein; Encephalitis; Computer-aided drug design (CADD)

**Abbreviations:** NiV-G: Nipah Virus Attachment Glycoprotein; CADD: Computer Aided Drug Designing; CNS: Central Nervous System; BBB: Blood Brain Barrier.

### Introduction

Somewhere in the range of 20 years back, Nipah virus (NiV) emerged and was shown to be a previously obscured paramyxovirus, now being classified as one of the 5 members within the Henipavirus genus. The single-stranded negative-sense RNA genome of NiV contains only six genes corresponding to six structural genes; however, it can exert an unusual broad species tropism [1]. The viruses are highly pathogenic to several vertebrate animal groups including the humans and have been given biosecurity level 4 status [2]. The manifestation of pathogenicity of NiV infection appears to be essentially due to endothelial damage, multinucleated syncytia and vasculitis-induced thrombosis, ischemia and micro infraction in the central nervous system (CNS), allowing the virus to overcome the blood-brain-barrier (BBB) and subsequently causes infection of neurons and glial cells in the brain parenchyma [3,4].

Several fruit bats of the Pteropodidae family have been identified as natural reservoirs of NiV and pigs can serve as an amplifying host [5]. Furthermore, there are various other mammalian species that are vulnerable to NiV disease [6-8]. NiV has been isolated from the brain and spinal fluid of infected personnel and bat urine and partially-eaten fruit in Malaysia [9].

NiV outbreaks have occurred on a regular basis in Bangladesh and India, with human case fatality rates approaching 75% [10-12]. Although there is a lack of wholeness of data regarding the infection rate, till March, 2012 an aggregate of 209 human instances Bangladesh were accounted for; 161 (77%) of them died [11]. Since the first NiV outbreak in 2001, the virus caused outbreaks for at least 11 times in Bangladesh during the period 2001–2011 [11,12]. Contact with an infected cow, drinking fresh date.

Palm sap and men-to-men transmission were reported to be responsible for the outbreaks [11,13-15]. Viral isolates collected during the outbreaks in Bangladesh showed that diverse strains were involved

[16]. Viral infections caused by NiV which are occurring in Bangladesh potentially provide insight into broader clinical manifestations [17].

Till now, no effective treatment for Nipah virus infection has been reported. Chong et al., mentioned that, nausea, vomiting, and convulsions may get alleviated by Ribavirin [18]. All the existing treatment strategies for Nipah Virus infection are generally focused on managing fever and the neurological symptoms and for some severe cases, by providing ventilation support.

Novel antiviral strategy development against an emerging virus like the Nipah virus relies most on the better understanding of the molecular mechanism of their entry into the host cell [19]. The cell-cell adhesion mediated by glycoprotein is an important virulent factor for the virus to transmit directly from an infected cell to an uninfected cell in the vicinity and it significantly induces the cytological pathogenicity of a NiV infection [20]. The attachment glycoprotein (G) protein is composed of 602 amino acid residues and classified as a type II membrane glycoprotein which primarily functions as a receptor-binding protein [21]. This G protein is accountable for the viral attachment to sialic acid-containing host cell receptors and plays the most significant role in the replication process of the viral nucleic acid inside the host cell [22]. Hence, G protein possess a chief role in the process of replication of paramyxovirus, and of course can be taken as the primary target for neutralizing antibodies as well as potential inhibitory agents targeting the virus [23]. The G proteins of Henipavirus interact with ephrin-B2 and ephrin-B3, the only two members of ephrin family, in a very specific manner and the molecular basis of

**\*Corresponding author:** Md Hazrat Ali, Assistant Professor, Department of Genetic Engineering and Biotechnology, School of Life Science, Shahjalal University of Science and Technology, Sylhet 3114, Bangladesh, Phone: +880 17176 299 245; E-mail: [hazratali-geb@sust.edu](mailto:hazratali-geb@sust.edu)

Received April 25, 2018; Accepted May 15, 2018; Published May 25, 2018

**Citation:** Ali MH, Anwar S, Roy PK, Ashrafuzzaman MD (2018) Virtual Screening for Identification of Small Lead Compound Inhibitors of Nipah Virus Attachment Glycoprotein. J Pharmacogenomics Pharmacoproteomics 9: 180. doi: [10.4172/2153-0645.1000180](https://doi.org/10.4172/2153-0645.1000180)

**Copyright:** © 2018 Ali MH, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

that points out how they intervene in a unique fashion cell adhesion as well as the incitement of the fusion of the membrane during viral infection process [24]. The crystal structures of NiV-G have already been determined and suggestions that the interaction between NiV-G and ephrin can be effectively targeted to shatter the viral entry are available [25,26]. Essentially, this provides the base for designing the structure-based antiviral drug.

Computer-aided drug design (CADD) has become a very important part of the rational drug design process now a day. This involves extensive computer modeling methods to reduce the costs and speed up the drug developing process [26]. Designing a drug is not a talk of overnight- it is a very sensitive, time-consuming, costly, sophisticated, and inefficient process. Collier et al. estimated that, the average cost on a new drug development is about US\$ 1.3-1.7 billion and it takes about eight to ten years [27]. CADD helps to find out potential drug by its speed and efficiency. High-throughput robotic screening methods accelerate this process, but still, this is a time-consuming process as a great number of compounds must be trialed [28]. By the use structure-based drug designing process, it is hoped that at least a smaller number of compounds will be found active against the target and that very small number of compounds are then taken for trial. The availability of the 3D structure of the targeted macromolecule, usually by X-ray crystallography or nuclear magnetic resonance (NMR) or in few instances, homology models, enhances the procedure [29]. In general, the more precise the 3D structure and its information are the more accurate predictive results will be found in the context of drug discovery, importance of high-throughput docking has getting elevation every day [30,31]. Though there are technical challenges in reliably predicting the mode of binding of a molecule to a target and making comparison among the binding affinities to other compounds, molecular docking campaigns have produced significant upliftment of hit rate compared to random screening in a commencing number of cases [32-35]. In this particular study, we have incorporated structure based virtual screening method to search out the potential drug like compounds for the treatment against NiV-encephalitis targeting the NiV attachment glycoproteins (Niv-G).

## Materials and Methods

### Preparation of G-glycoprotein structure

The 3D crystal structure of G-glycoprotein of NiV (PDB ID: 3D11) in complex with N-glycosylation site was retrieved from the Protein Data Bank [25]. All the water molecules and iodine ions were removed to make the structure of G-glycoprotein of NiV prepared for structure based virtual screening as well as molecular docking processes.

### Active site prediction

Binding site of the G-glycoprotein was determined by the aid of CASTp protein prediction server [36]. As determined by CASTp using a 1.4 Å radius probe, the volume of the internal cavity surface of the ligand binding sites was computed.

### Ligand preparation for virtual screening

Compound input libraries of 20,000 drug like small compounds from about 3,50,000 drug-like compounds, based on Lipinski rule of five from ZINC database and 1592 compounds with a rich structural as well as pharmacophore diversity, were chosen as the molecules in demand for virtual screening.

The NCI diversity set-II is a compilation of about 1592 compounds

that are structurally delegate of a wide range of pharmacophores. The "Diversity set" was attained from almost 140,000 compounds accessible on plates and from these 71,756 compounds were then reduced to the final set using the Chem-X program [37]. The eliminated compounds were reduced as at least 1 gr of the material was not available. The pharmacophores for the current structure were weighted to all other pharmacophores found in structures already accepted in to the diverse subset. If the current structure had more than a predetermined number of new pharmacophores, it was kept to the diverse subset and by this process we could select 1592 compounds in the diversity data set. As the selection procedure was order dependent, the order was randomized.

### Virtual screening with Auto dock vina

For this particular study, we used Auto Dock Vina as the primary docking program for virtual screening [38]. We prepared the 3D11 pdbqt file and determined grid box sizes using Auto dock Tools version 4.2 (The Scripps Research Institute, La Jolla, USA). Auto dock Vina is reported for its accuracy and speed, which is far faster than its ancestor program, Auto dock4 [39]. However, we prepared the input pdbqt file for 3D11 and set the size and the center of the grid box by employing Auto dock Tools. The 3D11 structure was included with the Kollman charges and polar hydrogen atoms. Over and above, we set the grid box sizes at 60, 60, and 60 Å and the spacing between grid points at 0.375 angstroms. The predicted binding strength (kcal/mol), which fingers at how strongly a ligand binds to the receptor, is computed depending on the scoring function used in Auto Dock Vina. The more the negative value of binding affinity is, the stronger the affinity is.

Pursuant that, the scoring function of the software Auto Dock Vina is divided into 2 portions:

- i) Conformation-dependent portion, which can be regarded as a sum of intramolecular and intermolecular contributions, and
- ii) Conformation-independent portion, which relies on the number of rotatory bonds between heavy atoms in the ligand.

Contributions of each of the portions are given an unidentical weighted value in the scoring function of the Auto Dock Vina [39]. The contributions include the steric, hydrophobic, hydrogen bonding and number of rotatory bonds

### ADMET properties prediction

The physico-chemical properties such as log P value, H-bond donors, H-bond acceptors, molecular weight and rotational bonds, for ligands obtained from NCI diversity set-II are calculated using the ADME/Tox filter server and presented in Table 1 [40].

The top hits from each screen were initially filtered for drug likeness by their adherence to Lipinski's "rule of fives" [41]. It is recommended that compounds should conform to two or more of these rules and we only considered compounds conforming all rules in this study, i.e., having no more than five hydrogen bond donors, no more than 10 hydrogen bond acceptors, a molecular weight under 500 gr/mol,

HBD	0-5
HBA	0-10
Log P	0-5
Molecular weight (M.wt)	160-500
TSPA	20-140 Å <sup>2</sup>
Number of rotatable bond	0-10

Table 1: Physicochemical parameter used for filtering ZINC database.

and a partition coefficient (log P) under 5. Compounds were filtered after rather than prior to docking to prevent premature exclusion of compounds that do not meet this criterion. The remaining hits were further filtered based on clustering criteria in Auto Dock: we only consider a compound among the top hits if the most populated cluster included at least 25% of all docked conformations. Receptor ligand interactions, including hydrogen-bonding interactions and molecular surfaces, were calculated in Auto Dock Tools version 4.2 [42,43].

### Molecular docking

To carry out the further docking simulation of the top-ranking compound resulted by Auto Dock Vina, we used the Auto Dock 4.2 suite as molecular-docking tool [39]. It is suitable software for performing automated docking of ligands to their macromolecular receptors. Typically, the ligands are substrates or drug candidates and the macromolecule is a protein of known three dimensional structures. In this docking simulation, we used rigid docking protocols. In which the target protein G-glycoprotein and the ligand molecule both were kept as rigid. The Graphical User Interface program "Auto dock Tools" was used to prepare, run, and analyze the docking simulations. Kollman united atom charges, solvation parameters and polar hydrogens were added into the receptor PDB file for the preparation of protein in docking simulation. Auto dock requires pre-calculated grid maps, one for each atom type; present in the ligand being docked and its stores the potential energy arising from the interaction with macromolecule. This grid must surround the region of interest in the macromolecule. In the present study, the binding site was selected based on the active site component and also the Ephrin binding site of the G-glycoprotein. Both the active site and Ephrin binding site residues are covered as grid box. The grid box size was set at 60, 60, and 60 Å (x, y, and z), though it was changed depending on the ligand size. Auto Grid 4.2 Program, supplied with Auto dock 4.2 was used to produce grid maps. The spacing between grid points was 0.375 angstroms. PyMol (De Lano Scientific LLC, USA) and DS Visualizer (Accelrys, Inc., USA) were employed to visualize and modify the receptor and ligand structures.

## Results

### Binding pocket of NiV-G

The crystal structure (PDB ID: 3D11) of the Nipah G Attachment Glycoprotein (NiV-G) of nipah virus complexed with ephrin-B3 at a resolution of 2.2 Å was retrieved from PDB database, which was then employed to search for potential inhibitors. As the binding pocket was calculated by CASTp the amino acid residues surrounding the binding pockets are ASP219, PRO220, PRO276, ASN277, VAL279, TYR280, HIS281, CYS282, TYR351, GLY352, PRO353, PRO448, PHE458, GLY489, GLN490, GLY506, VAL507, TYR508, LYS560, GLY559. Thus, all these residues are predicted as important residues for docking and later used to create grid during the docking study. The predicted binding pocket has a volume of 807.2 Å and area of 1617.5. The residues constitute the binding pocket of NiV-G are shown in Figure 1.

### Virtual screening for potential ligands

Prior to the docking study, all considered ligands were subjected to virtual screening to identify promising ligands based on their affinity score obtained from Auto dock Vina. A total of 21592 compounds of NCI diversity set II and ZINC database were screened for this purpose. We selected top 64 and 10 compounds out of 1592 and 20000 compounds from NCI diversity set II and ZINC database, respectively according to their binding affinity of the molecule to the G-glycoprotein of Nipah virus measured by Auto dock Vina. All selected ligands were

then checked through ADMET/Tox filter to ensure their compatibility as a drug considering several structural features. All those criterions we used to select druglike molecules are having no more than five hydrogen bond donors, no more than 10 hydrogen bond acceptors, a molecular weight under 500 g/mol, and a partition coefficient (log P value) under 5 which are also known as Lipinski's rule of five. Only 19 compounds of 64 NCI diversity set II were passed through the ADMET test. A brief analysis of top binders, however, revealed that most compounds that failed the Lipinski rules were unlikely to be drug candidates. Although some compounds that failed the Lipinski rules also had good docking scores, this cut off criterion was imposed to focus the scope of our search to the more promising compounds. The top 19 and top 10 compounds of NCI diversity set II and ZINC database are given to the Tables 2 and 3 respectively.

### Selection of positive control and validation of the docking

In molecular docking, the size and center of the coordinates of the grid box need to be validated to ensure that ligands bind to the binding pocket in the correct conformation. Our aim of this work is to block the binding of the G-glycoprotein to the Ephrin receptor molecule. So, the Ephrin binding site was utilized as the grid box center for molecular docking studies of ligands in search of potential inhibitors. As there is no any natural ligand bound to the active site of the G-glycoprotein, we considered the Rivabirin as control molecule. Heng-Thay et al. suggested that Rivabirin can be a potential drug for the treatment of acute Nipah encephalitis [14]. We docked the Rivabirin into the active site of the protein to analyze the binding pattern of the Rivabirin. The detailed view of docking of Rivabirin to the active site is shown in Figure 2A and B which exhibited molecular docking with  $\Delta G$  value of -6.36 (kcal/mol). Rivabirin forms 7 hydrogen bonds and several hydrophobic interactions with residues inside the active site. Rivairin has been found to form seven hydrogen bonds among oxygen O18 of Rivabirin and the backbone hydrogen of Gly352 (bond length is 2.25 Å), the H28 and the backbone oxygen of Cys282 (2.08 Å), the O3 of and the backbone hydrogen of Tyr508 (2.18 Å), the H27 and the backbone oxygen of Pro220 (1.80 Å), the O21 and the side chain hydrogen of

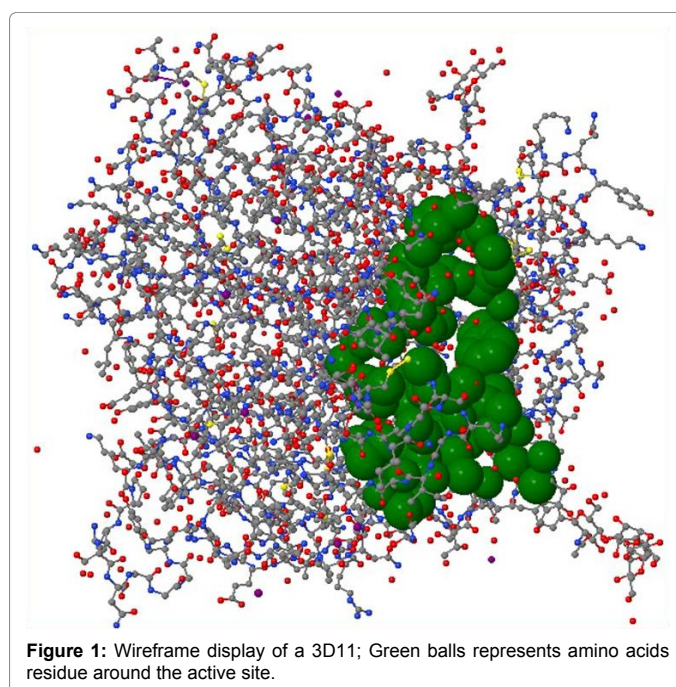


Figure 1: Wireframe display of a 3D11; Green balls represents amino acids residue around the active site.

No	NSC	Compound name	Molecular Weight (g/mol)	Auto dock Vina Score (Kcal/mol)	Predicted Ki (μM)
1	308835	9-dioxo-8-azaspiro[4.4]nonan-8-yl)-3-methylphenyl]-2-methylphenyl]-8-azaspiro[4.4]nonane-7,9-dione	484.6	-10.2	0.033
2	37553	1-N,4-N-bis(3-phenylphenyl)piperazine-1,4-dicarboxamide	476.6	-10.1	0.039
3	400770	(4'-hydroxy[1,1'-biphenyl]-4-yl)(phenyl)methanone	274.3	-9.8	0.065
4	5157	N-(4,7-dihydroxy-8-methyl-2-oxo-2H-chromen-3-yl)-2,2-dimethyl-3,4-dihydro-2H-1-benzopyran-6-carboxamide	395.4	-9.5	1.08
5	97920	dispiro[2,4]diazatricyclo[7.3.1.0 <sup>5,13</sup> ]tridecane-3,1'-cyclohexane-4',3"-[2,4] diazatricyclo[7.3.1.0 <sup>5,13</sup> ]tridecane]-1(12),1"(12"),5,5",7,7",9(13),9"(13"),10,10"-decaene	392.5	-9.5	1.08
6	80997	1-(1,11-dihydroxy-2,5,10a,12a-tetramethyl-7-phenyl-1,2,3,3a,3b,7,10,10a,10b,11,12,12a-dodecahydrocyclopenta [5,6]naphtho[1,2-f]indazol-1-yl) ethanone	472.6	-9.4	1.2
7	670283	2,2'-spiro[3,6,7,8-tetrahydro-1H-cyclopenta[g]naphthalene]-5,5'-dione	356.5	-9.4	1.2
8	345845	3-(4,5-dimethylbenzo[h][1,6]naphthylidin-2-yl)-2-methyl-4-quinolinylamine	361.9	-9.4	1.2
9	127133	2'-(((6-hydroxy-3-phenanthridinyl)amino) carbonyl)[1,1'- biphenyl]-2-carboxylic acid	433.4	-9.3	1.52
10	332670	2-(1H-phenanthro[9,10-d]imidazol-2-yl)phenol	310.4	-9.3	1.52
11	67436	4-chloro-1-N,3-N-bis[4-(4,5-dihydro-1H-imidazol-2-yl)phenyl]benzene-1,3 dicarboxamide	489	-9.3	1.52
12	324623	3-(2-pyridinyl)-1-(3-pyridinylcarbonyl)-4H-quinolizin-4-one	327.3	-9.2	1.8
13	37641	4b-fluoro-6b-glycoloyl-5-hydroxy-4a,6a,8-trimethyl-8-phenyl-4a,4b,5,6,6a,6b,9a,10,10a,10b,11,12-dodecahydro-2H-naphtho[2,1':4,5] indeno[1,2-d][1,3]dioxol-2-one	496.6	-9.2	1.8
14	60339	2-chloro-N1,N4-bis(4-(4,5-dihydro-1H-imidazol-2-yl)phenyl)terephthalamide	487	-9	2.5
15	117268	4-(2-(4-chlorophenyl)-4,6-diimino-1,3,5-triazinan-1-yl)benzenesulfonamide	378.8	-9	
16	59620	4a,6a,9-trimethyl-6,7-dioxo-2,3,4,4a,4b,5,6,6a,6b,7,10,10a,11,11a,11b,12,13,13a-octadecahydro-1H-indeno[2,1-a] phenanthren-2-yl acetate	412.6	-8.9	2.9
17	26258	2-isopropenyl-8,9-dimethoxy-1,2,12,12a-tetrahydrochromeno[3,4-b]furo[2,3-h] chromen-6(6aH)-one	394.9	-8.9	2.9
18	26692	4,5-bis(4-methoxyphenyl)-1,5-dihydro-2H-imidazol-2-one	296.3	-8.9	
19	379651	2-(2-furyl)-5-phenylpyrazolo[1,5-a]pyrimidin-7(4H)-one	277.3	-8.8	3.5

**Table 2:** NSC Database ID number, IUPAC name, molecular weight, binding energy and Ki value of the top scored ligands with the G-glycoprotein of Nipah Virus using docking software AutoDock Vina.

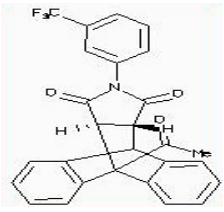
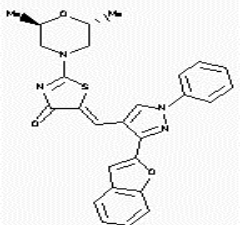
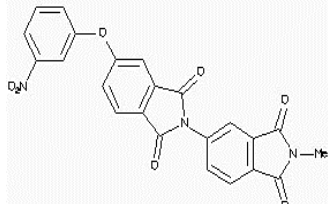
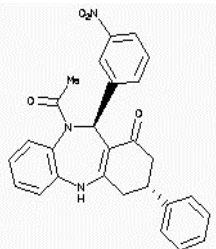
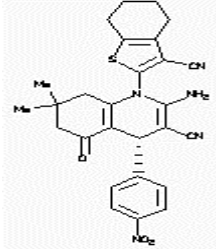
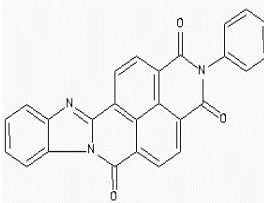
His281 (2.14 Å) and O and O15 and the side chain hydrogen of Lys560 (2.04 Å and 2.14 Å). In addition to hydrogen bonds, Val507, Tyr351, Phe458, Pro441, Pro353 and Ile562 were involved in the van der Waals interactions which further stabilize the Rivabirin binding strength.

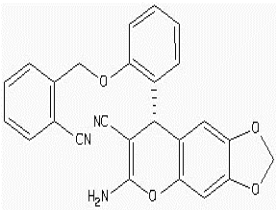
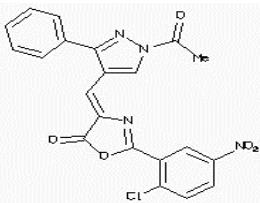
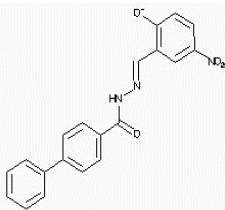
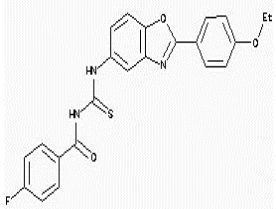
All the amino acid residues which involved in molecular interaction are shown in wire frame drawing and colored by atom types in which hydrogen is colored gray, carbon green, oxygen red, nitrogen blue, and sulfur yellow. Residues with hydrophobic contacts with the ligands are leveled in black. (B) Binding of Ribavirin into the active site pocket of NiV-G.

### Docking of the top four discovered NSC ligands

After the ADMET test we found that the top four compounds which have the strongest affinities for the active site of the G-glycoprotein were NSC308835, NSC37553, NSC227186 and NSC5157 each with a binding affinity of -10.6 kcal/mol, -10.1 kcal/mol, -10.0 kcal/mol and -9.5 kcal/mol respectively. We performed further analysis of these top four compounds on target protein to evaluate these compounds as drug. Binding of 9-dioxo-8-azaspiro[4.4]nonan-8-yl)-3-methylphenyl]-2-methylphenyl]-8-azaspiro[4.4]nonane-7,9-dione (NSC308835) and G-glycoprotein is shown in Figure 3A. NSC308835 formed three hydrogen bonds with in the active site of the of the target protein. It formed one hydrogen bond at O<sub>2</sub> from side chain hydrogen of Lys560 (bond length is 2.13 Å), one hydrogen bond at O1 from side chain oxygen of Asp219 (bond length is 2.8 Å) and another hydrogen bond at

O4 from side chain oxygen of Gly506 (bond length is 2.5 Å). In addition, several hydrophobic interactions are formed between NSC308835 with the hydrophobic side chain of Tyr351, Pro441, Phe458, His281, Asp219, Gln559 and Ala532. The predicted Ki value for NSC308835 is 0.033 (Table 1) which shows the lowest Ki value. Binding of 1-N, 4-N-bis(3-phenylphenyl) piperazine-1,4-dicarboxamide (NSC37553) and G-glycoprotein is shown on Figure 3B. NSC37553 formed two hydrogen bonds within the active site residues of the target protein. It forms one hydrogen bond at O1 from side chain hydrogen of Lys560 (bond length is 2.12 Å) and one hydrogen bond at O<sub>2</sub> from backbone oxygen of Gly506 (bond length is 3.4 Å). Hydrophobic interactions are formed between NSC37553 with the hydrophobic side chain of Gln490, Val507, Tyr508, Phe458, Pro441, Asp219, Gln559 and His281. Binding of (4'-hydroxy[1,1'-biphenyl]-4-yl)(phenyl)methanone (NSC400770) and G-glycoprotein is shown on Figure 3C. NSC400770 formed two hydrogen bonds, in which one hydrogen bond at O1 from the side chain hydrogen of His281 (bond length is 2.20 Å) and another hydrogen bond at side chain H14 of Gly5.6 from O of the ligand molecule (bond length is 3.00 Å). And the binding of N-(4,7-dihydroxy-8-methyl-2-oxo-2H-chromen-3-yl)-2,2-dimethyl-3,4-dihydro-2H-1-benzopyran-6-carboxamide (NSC5157) with G-glycoprotein is shown in Figure 3D and E. NSC5157 formed two hydrogen bonds with the active site of the target protein. It forms one hydrogen bond at O5 from the side chain hydrogen of His281 (bond length is 2.22 Å) and one hydrogen bond at O4 from backbone nitrogen of Tyr508 (bond length is 2.30 Å). In this study, we also determined the binding affinity between protein and its

No	ZINC ID	Compound Structure	Compound Name	Molecular weight (g/mol)	AutoDock Vina Score (Kcal/mol)
1.	ZINC00707572		1-acetyl-17-[3-(trifluoromethyl)phenyl]-17-azapentacyclo[6.6.5.0~2,7~.0~9,14~.0~15,19~]nonadeca-2,4,6,9,11,13-hexaene-16,18-dione	461.439	-10.9
2.	ZINC01019934		(5Z)-5-[[3-(1-benzofuran-2-yl)-1-phenylpyrazol-4-yl]methylidene]-2-[(2R,6R)-2,6-dimethylmorpholin-4-yl]-1,3-thiazol-4-one	484.581	-10.8
3.	ZINC00827822		2-methyl-5-[5-(3-nitrophenoxy)-1,3-dioxoisindol-2-yl]isindole-1,3-dione	443.371	-10.6
4.	ZINC00654610		(6S,9R)-5-acetyl-6-(3-nitrophenyl)-9-phenyl-8,9,10,11-tetrahydro-6H-benzob[1,4]benzodiazepin-7-one	453.498	-10.6
5.	ZINC02573756		(4S)-2-amino-1-(3-cyano-4,5,6,7-tetrahydro-1-benzothiophen-2-yl)-7,7-dimethyl-4-(4-nitrophenyl)-5-oxo-6,8-dihydro-4H-quinoline-3-carbonitrile	499.596	-10.4
6.	ZINC00988780		2-phenylbenzimidazo[2,1-b]benzo[lmn][3,8]phenanthroline-1,3,6 (2H)-trione	415.408	-10.3

7.	ZINC00663998		6-amino-8-[2-[(2-cyanophenyl) methoxy] phenyl]-8H[1,3]dioxolo [4,5-g]chromene-7-carbonitrile	423.428	-10.2
8.	ZINC00726045		4-[(1-acetyl-3-phenylpyrazol-4-yl) methylidene]-2-(2-chloro-5-nitrophenyl)-1,3-oxazol-5-one	436.811	-10.2
9.	ZINC01922868		N'-(2-hydroxy-5-nitrobenzylidene)-4-biphenylcarbohydrazide	360.349	-10.0
10.	ZINC00724735		N-[[2-(4-ethoxyphenyl)-1,3-benzoxazol-5-yl]carbamoithoyl]-4-fluorobenzamide	435.48	-9.9

**Table 3:** ZINC Database ID number, chemical structure, IUPAC name, molecular weight and binding energy of the top scored ligands with the G-glycoprotein of Nipah Virus using docking software AutoDock Vina.

inhibitor shown in Table 1 by calculating the inhibitory constant ( $K_i$ ) of the protein inhibitor complex, which is correlated to the concentration at which 50% of the protein is inhibited ( $IC_{50}$ ). NSC308835 showed the lowest  $K_i$  value by which we can measure how much effective this compound as drug.

The magenta dot lines denote the hydrogen bonds. Residues with hydrophobic contracts with the ligands are leveled in black. All the amino acid residues which involved in molecular interaction are shown in wire frame drawing and colored by atom types in which hydrogen is colored gray, carbon green, oxygen red, nitrogen blue, and sulfur yellow. Ligands are shown in stick drawing. (E) Superimposition of Ribavirin and four discovered compounds. Crystallized conformation of Ribavirin is shown in red, NSC308835 is shown in red green, NSC37553 is shown blue color, NSC400770 is shown in yellow color and NSC5157 is shown in cyan color.

## Discussion

Virtual Screening, an *in silico* tool for drug discovery, has emerged as an important tool in our quest access towards the novel drug like compounds. Nipah virus encephalitis remains a pharmacologically under estimated disease despite the high mortality rate, especially in the developing countries like Bangladesh, where access to medication facilities is limited. To perform virtual screening NCI Diversity Set-11 containing 1592 diverse compounds tested in NCI National Center Institute) that facilitated by not to scorn the whole NCI database were

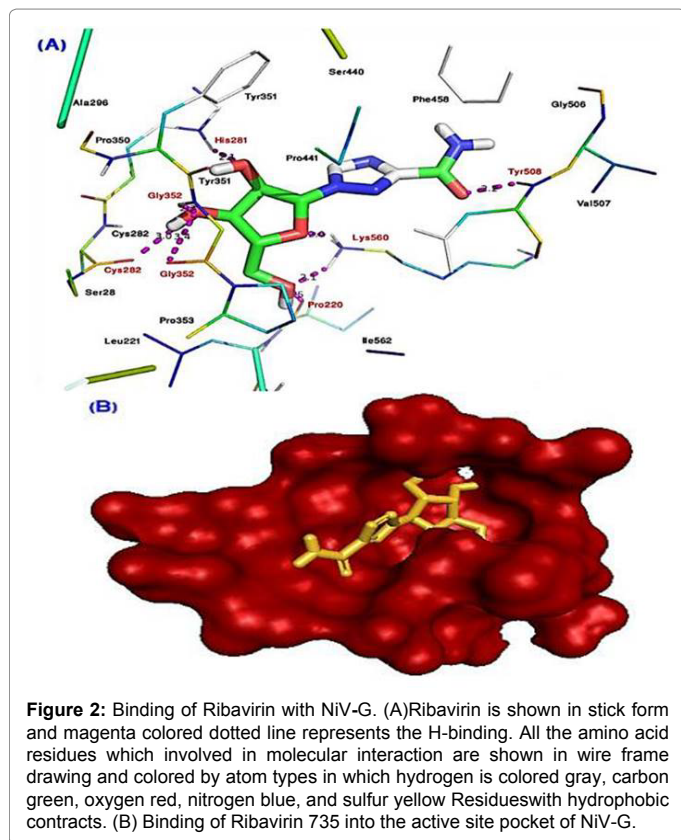
chosen. A more 20,000 drug-like compounds from ZINC database were also screened. Alongside, the fact which motivated us to use Auto dock Vina as virtual screening platform for molecular docking as it is not commercial and can be run on a single personal computer.

## Binding site of NiV-G

The binding pocket of NiV-G determined by CASTp server manifests that a number of residues are engaged in the active site of NiV-G protein. The most active residues in binding pocket are used for grid box selection. As the Ephrin binding site area is closely placed we selected the whole area as Grid box for performing virtual screening. Among the active site residue Asp219, Lys560 and His281 play an important role in small ligand binding.

## Docking of NSC compounds

Computer aided simulation study using molecular docking techniques performed on the NCI Diversity Set- 11 resulted NSC308835 and NSC37553 are bounded at the minimal energy to the NiV-G. In the docking study of Ribavirin, we found that it binds with the active site residues such as with Lys560, His281, Pro220, Gly352 and Cys282 which are revealed as active site in the study. The structure activity relationships of the best hit, NSC308835, against NiV-G observed VIS docking interaction exposes that the oxygen and hydrogen functionalities have a strong hydrogen bond interaction with Lys560, Asp219 and Gly506 amino acid present in the active site



The combination of these interactions is also found in the docking analysis of Ribavirin which is suggested as the potential drug for the treatment of NiV encephalitis [18] and this prove that NSC308835 and NSC37553 is meaningful for the test compound. In addition although all compounds are able to bind the active site of the gorge, not all of them are able to interact with all the important residues at the binding sites. Ligand size may be one of the reasons for some of the activities being low.

### Docking of ZINC database compounds

We have performed a virtual screening study enumerating the ZINC drug-like database for potential inhibitor against NiV-G. About twenty thousand drug-like compounds have been virtually screened for the potent drug-like molecule against NiV-G. The drug-like compounds were filtered by Lipinski's "Rule of five" that sets the criteria for drug-like properties. According to this rule, poor absorption is expected if MW>500, logP>5, hydrogen bond donors>5, and hydrogen bond acceptors>10 [41]. Virtual screening of ZINC database revealed that ZINC00707572 and ZINC01019034 manifests the lowest binding affinity after docking simulation. ZINC00707572 showed the lowest binding free energy confirming the best docking feature in the active site of NiV-G. The docking study of ZINC00707572 against NiV-G resulted that fluorine and oxygen functionalities have strong hydrogen interactions with Gly506 and Tyr508 amino acids present in the active site of NiV-G and therefore these groups are essential for activity. In the active site, two benzyl rings form two  $\pi-\pi$  interactions within the Lys560. Another benzyl ring forms another  $\pi-\pi$  interaction within the Trp504.

Docking of ZINC01019934 exposed that hydrogen and nitrogen functionalities have strong interactions with Tyr508, Gly506 and Asp219 amino acids which are suggested as active site residues predicted by CASTp active site prediction server. One  $\pi-\pi$  interaction is formed between one benzyl ring of active site and His281. Additionally, H-bond and  $\pi-\pi$  interaction, several other interactions are also analyzed which played important role in ligand binding. In both cases Lys560 and His281 play a major role in molecular docking of small ligand into the active site of Nipah virus attachment glycoprotein. The binding pattern of this two ZINC compound is relevant to the control ligand Ribavirin. That is, ZINC00707572 and ZINC01019934 would be used as test compounds for curing NiV encephalitis.

### Conclusion

In conclusion, the previously mentioned  $\pi-\pi$  interactions, hydrogen bonds, and strong hydrophobic interactions formed between the inhibitors and the nearby NiV-G side chains serve dual roles:

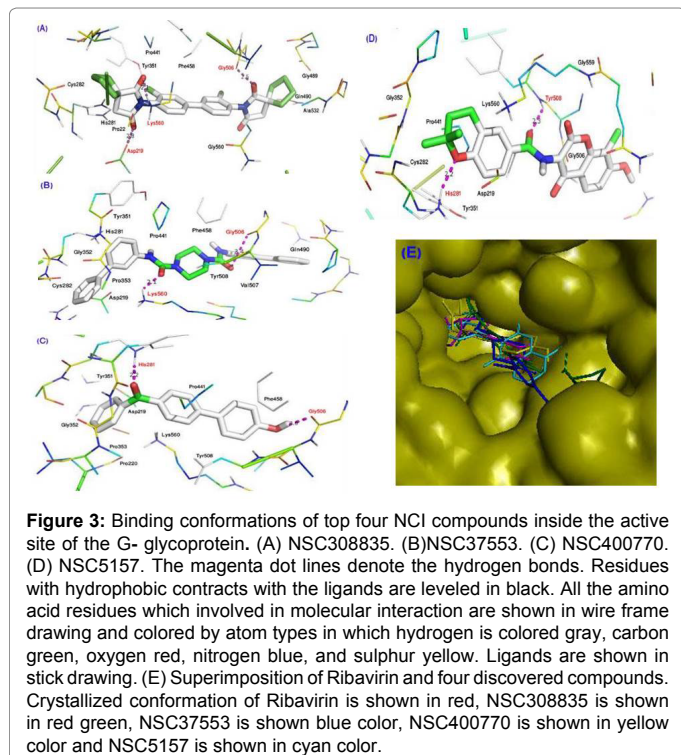
- 1) To inhibit the catalytic activity of NiV-G by competing with Ephrin binding site, and
- 2) To prevent NiV encephalitis by blocking the Ephrin recognition zone at the peripheral site.

### Acknowledgement

The authors gratefully acknowledge the co-operation of the Department of Genetic Engineering and Biotechnology, Shahjalal University of Science and Technology, Sylhet-3114, Bangladesh.

### Authors' Contributions

MHA conceived and designed the experiments. PKR carried out computational work. PKR wrote the draft of the manuscript. SA read and edited the draft. MA and AI read the manuscript and have given important intellectual advices. SA revised the draft critically and edited accordingly. MHA and SA read, reedited and approved the final manuscript.



of NiV-G I and thus these groups are essential for activity. Docking studies of the NSC37553 compounds with NiV-G. revealed that oxygen and hydrogen functionalities are making hydrogen interacting with the active site containing Lys560 and Gly506 amino acid of active site.

## References

1. Wang LF, Harcourt BH, Yu M, Tamin A, Rota PA, et al. (2001) Molecular biology of Hendra and Nipah viruses. *Microbes and Infection* 3: 279-287.
2. Eaton BT, Broder CC, Middleton D, Wang LF (2006) Hendra and Nipah viruses: different and dangerous. *Nature Reviews Microbiology* 4: 23.
3. Chua KB, Lam SK, Tan CT, Hooi PS, Goh K, et al. (2000) High mortality in Nipah encephalitis is associated with presence of virus in cerebrospinal fluid. *Annals of Neurology* 48: 802-805.
4. Wong KT, Shieh WJ, Kumar S, Norain K, Abdullah W, et al. (2002) Nipah virus infection: pathology and pathogenesis of an emerging paramyxoviral zoonosis. *The American Journal of Pathology* 161: 2153-2167.
5. Yob JM, Field H, Rashdi AM, Morrissy C, van der Heide B, et al. (2001) Nipah virus infection in bats (order Chiroptera) in peninsular Malaysia. *Emerging Infectious Diseases* 7: 439.
6. Chua KB, Bellini WJ, Rota PA, Harcourt BH, Tamin A, et al. (2000) Nipah virus: a recently emergent deadly paramyxovirus. *Science* 288: 1432-1435.
7. Geisbert TW, Daddario-DiCaprio KM, Hickey AC, Smith MA, Chan YP, et al. (2010) Development of an acute and highly pathogenic nonhuman primate model of Nipah virus infection. *PLoS one* 5: e10690.
8. Chua KB, Koh CL, Hooi PS, Wee KF, Khong JH, et al. (2002) Isolation of Nipah virus from Malaysian Island flying-foxes. *Microbes and Infection* 4: 145-151.
9. Chadha MS, Comer JA, Lowe L, Rota PA, Rollin PE, et al. (2006) Nipah virus-associated encephalitis outbreak, Siliguri, India. *Emerging Infectious Diseases* 12: 235.
10. Gurley ES, Montgomery JM, Hossain MJ, Bell M, Azad AK, et al. (2007) Person-to-person transmission of Nipah virus in a Bangladeshi community. *Emerging Infectious Diseases* 13: 1031.
11. World Health Organization. Nipah virus outbreaks in the WHO South-East Asia Region. *Surveillance and Outbreak Alert*.
12. Rahman M, Chakraborty A (2012) Nipah virus outbreaks in Bangladesh: a deadly infectious disease. *WHO South-East Asia Journal of Public Health* 1: 208.
13. Hsu VP, Hossain MJ, Parashar UD, Ali MM, Ksiazek TG, et al. (2004) Nipah virus encephalitis reemergence, Bangladesh. *Emerging Infectious Diseases* 10: 2082.
14. Luby SP, Rahman M, Hossain MJ, Blum LS, Husain MM, et al. (2006) Foodborne transmission of Nipah virus, Bangladesh. *Emerging Infectious Diseases* 12: 1888.
15. Hossain MJ, Gurley ES, Montgomery JM, Bell M, Carroll DS, et al. (2008) Clinical presentation of Nipah virus infection in Bangladesh. *Clinical Infectious Diseases* 46: 977-984.
16. Bellini WJ, Harcourt BH, Bowden N, Rota PA (2005) Nipah virus: an emergent paramyxovirus causing severe encephalitis in humans. *Journal of Neurovirology* 11: 481-487.
17. Mire CE, Satterfield BA, Geisbert JB, Agans KN, Borisevich V, et al. (2016) Pathogenic differences between Nipah virus Bangladesh and Malaysia strains in primates: implications for antibody therapy. *Scientific Reports* 6: 30916.
18. Chong HT, Kamarulzaman A, Tan CT, Goh KJ, Thayaparan T, et al. (2001) Treatment of acute Nipah encephalitis with ribavirin. *Annals of Neurology* 49: 810-813.
19. Guillaume V, Aslan H, Ainouze M, Guerbois M, Wild TF, et al. (2006) Evidence of a potential receptor-binding site on the Nipah virus G protein (NiV-G): identification of globular head residues with a role in fusion promotion and their localization on an NiV-G structural model. *Journal of Virology* 80: 7546-7554.
20. Moll M, Diederich S, Klenk HD, Czub M, Maisner A (2004) Ubiquitous activation of the Nipah virus fusion protein does not require a basic amino acid at the cleavage site. *Journal of Virology* 78: 9705-9712.
21. Harcourt BH, Tamin A, Ksiazek TG, Rollin PE, Anderson LJ, et al. (2000). Molecular characterization of Nipah virus, a newly emergent paramyxovirus. *Virology* 271: 334-349.
22. Bossart KN, Wang LF, Eaton BT, Broder CC (2001) Functional expression and membrane fusion tropism of the envelope glycoproteins of Hendra virus. *Virology* 290: 121-135.
23. Eshaghi M, Tan WS, Chin WK, Yusoff K (2005) Purification of the extra-cellular domain of Nipah virus glycoprotein produced in *Escherichia coli* and possible application in diagnosis. *Journal of Biotechnology* 116: 221-226.
24. Xu K, Rajashankar KR, Chan YP, Himanen JP, Broder CC, et al. (2008) Host cell recognition by the henipaviruses: crystal structures of the Nipah G attachment glycoprotein and its complex with ephrin-B3. *Proceedings of the National Academy of Sciences* 105: 9953-9958.
25. Bowden TA, Crispin M, Harvey DJ, Aricescu AR, Grimes JM et al. (2008) Crystal structure and carbohydrate analysis of Nipah virus attachment glycoprotein: a template for antiviral and vaccine design. *Journal of Virology* 82: 11628-11636.
26. Liljefors T, Krosgaard-Larsen P, Madsen U (2002) *Textbook of drug design and discovery*. CRC Press.
27. Collier R (2009) Drug development cost estimates hard to swallow. *CMAJ* 180: 279-280.
28. Houston JG, Banks M (1997) The chemical-biological interface: developments in automated and miniaturised screening technology. *Current Opinion in Biotechnology* 8: 734-740.
29. Ring CS, Sun E, McKerrow JH, Lee GK, Rosenthal PJ, et al. (1993) Structure-based inhibitor design by using protein models for the development of antiparasitic agents. *Proceedings of the National Academy of Sciences* 90: 3583-3587.
30. Schneider G, Böhm HJ (2002) Virtual screening and fast automated docking methods. *Drug Discovery Today* 7: 64-70.
31. Waszkowycz B (2002) Structure-based approaches to drug design and virtual screening. *Current Opinion in Drug Discovery & Development* 5: 407-413.
32. Verkhivker GM, Bouzida D, Gehlhaar DK, Rejto PA, Arthurs S, et al. (2000) Deciphering common failures in molecular docking of ligand-protein complexes. *J Computer-Aided Molecular Design* 14: 731-751.
33. Stahl M, Rarey M (2001) Detailed analysis of scoring functions for virtual screening. *Journal of Medicinal Chemistry* 44: 1035-1042.
34. Doman TN, McGovern SL, Witherbee BJ, Kasten TP, Kurumbail R, et al. (2002) Molecular docking and high-throughput screening for novel inhibitors of protein tyrosine phosphatase-1B. *Journal of Medicinal Chemistry* 45: 2213-2221.
35. Trosset JY, Dalvit C, Knapp S, Fasolini M, Veronesi M, et al. (2006) Inhibition of protein-protein interactions: The discovery of druglike  $\beta$ -catenin inhibitors by combining virtual and biophysical screening. *Proteins: Structure, Function, and Bioinformatics* 64: 60-67.
36. Dundas J, Ouyang Z, Tseng J, Binkowski A, Turpaz Y, et al. (2006). CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic acids research* 34: W116-W118.
37. Nicklaus MC, Milne GW, Zaharevitz D (1993) Chem-X and CAMBRIDGE. Comparison of computer generated chemical structures with X-ray crystallographic data. *Journal of Chemical Information and Computer Sciences* 33: 639-646.
38. Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Computational Chemistry* 31: 455-461.
39. Morris GM, Huey R, Olson AJ (2008) Using autodock for ligand receptor docking. *Current protocols in Bioinformatics* 2008: 8-14.
40. Lagorce D, Sperandio O, Galons H, Miteva MA, Villoutreix BO (2008) FAF-Drugs2: free ADME/tox filtering tool to assist drug discovery and chemical biology projects. *BMC Bioinformatics* 9: 396.
41. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (2012) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews* 64: 4-17.
42. Sanner MF (1999) Python: A programming language for software integration and development. *J Mol Graph Model* 17: 57-61.
43. Feng BY, Simeonov A, Jadhav A, Babaoglu K, Inglesse J, et al. (2007) A high-throughput screen for aggregation-based inhibition in a large compound library. *J Medicinal Chem* 50: 2385-2390.