High Throughput Screening for Identification of Potent Inhibitors Targeting Ebola Virus Major Matrix Protein VP40

J.V. Priya¹, A. Soma Prabha², S.V. Rajesh³, V. Aroulmoji⁴ and T.S. Gnanendra⁴*

¹Bioinformatics Division, OriGene Biosolutions, Salem-9, Tamilnadu, India.
²School of Biotechnology, Madurai Kamaraj University, Madurai-21, TamilNadu, India.
³Department of Botany, Vivekananda College of Arts and Sciences for Women, Elayampalyam, Tiruchengode, Namakkal, TamilNadu, India.
⁴Bioinformatics Division, Centre for Research and Development, Mahendra Educational Institutions, Mallasamudram, Tiruchengode, Namakkal, TamilNadu, India.

ABSTRACT: Ebola virus is a single-stranded, negative-sense RNA virus that causes severe hemorrhagic fever. The outbreak of Ebola infections in West Africa during 2014 has expanded exponentially with a doubling period of 34.8 days, resulting 83% mortality rate as it exhibited resistance to large number of anti-viral drugs. As there are no valid drugs to treat Ebola virus infections that cause severe hemorrhagic fever, the urgency in search and design of utmost anti-viral drug has propelled our research to screen the compounds to inhibit the Ebola infections. Thus in our present study, the membrane-associated major matrix protein (VP40) that plays a central role in the formation of filamentous virus particles and also binds to RNA tribonucleotide to form the viral matrix is chosen as potential drug target. Blocking the RNA site of interaction with VP40 may restrict the pathway of virus membrane formation. Hence the known anti-viral drugs were screened for their inhibitory activity by docking in the RNA binding site of VP40. The result of virtual screening revealed that all the compounds screened exhibited the encouraging docking scores. The highest docking score was observed for Nocodazole, which suggests that this compound might effectively occupy the RNA binding regions; the results of these binding interactions indicate that the most important amino acids in the binding site of VP40 are to be considered while designing the novel anti-viral drugs that can effectively inhibit the Ebola infections.

KEY WORDS: Ebola infections, Virtual Screening, Docking, Anti-viral drugs, matrix protein.

INTRODUCTION

The largest viral outbreak in the history of West Africa has occurred in 2014 by a thread like tubular filament virus namely Ebola [1]. This deadly pathogenic virus belongs to filoviridea family [2]. According to the Centre for Disease Prevention and Control, Ebola virus has majorly affected four countries in West Africa namely, Guinea, Liberia, Nigeria, and Sierra Leone and a scenario of potential pandemic spreading across globe is observed [3]. The first outbreak of Ebola was reported in Democratic Republic of Congo (Africa) during 1976 [4]. The Ebola outbreak in 2014 has expanded exponentially with a doubling period of 34.8 days, resulting 83% mortality rate [5]. The major source of these viral infections was observed as carcasses of gorilla and chimpanzees, during the outbreak of Ebola in 2001 and also bats were reported as a possible reservoir of Ebola virus infections [6].

The most significant symptoms of Ebola infection is severe and often fatal hemorrhagic fever that takes place in two phases; incubation phase shows symptoms like joint pain, fever, fatigue, nausea which can last for one week and second late phase symptoms include depression, eye inflammation, and hemorrhagic rash over the entire body [7]. The Ebola virus infections occur with binding to endothelial cells, mononuclear phagocytes and hepatocytes [8]. The virus escapes the immune system by interfering the neutrophiles signals and also by inhibiting the early steps of neutrophiles activation [9]. The neutrophiles signal inhibition is interfered by a diamic protein, secreted glycoprotein (sGP) [10].

The release of cytokines, signaling molecules for fever and inflammation is affected by the presence of viral particles and the budding of Ebola virus leads to cell damage [11]. The further synthesis of viral glycoproteins affects the cell adhesion to intracellular structure by reducing the specific integrins [12]. The large number of viral GP lead to coagulopathy due to loss of vascular integrity. This improper coagulation results in hypovolmic shock due to the leakage of blood through the vessels [13].

The Ebola virus has a negative stranded RNA genome of 19 kb (Figure.1) encoding seven polypeptides; out of which there are three membrane associated proteins: glycoprotein (GP), VP40, VP24 and four structural proteins: Nucleoprotein (NP), RNA-dependent RNA polymerase (L), VP35, VP30 (Figure.2). The glycoprotein (GP) plays a crucial role in release of cytokines and also in interfering the neutrophiles signals [14].

*Corresponding author.
+91-99421 57168
gnani.science@gmail.com
Received : 23.04.2015
Accepted : 15.05.2015
Published on: 20.05.2015

Figure.1: The negative ssRNA genome of Ebola virus (Adopted from Viral Zone)
The protein VP40 is the most abundant matrix protein located under the viral bilayer and plays a crucial role in structural integrity of the viral particles. Its assembly and budding process occurs at the plasma membrane. This protein plays an important role in the Viral RNA metabolism during Ebola replication [15].

The membrane-associated glycoprotein (GP) and matrix protein (VP40) plays a central role in the formation of filamentous virus particles. VP35 plays a major role in viral assembly by acting as a component of viral RNA polymerase complex. This protein reduces the host interferon production, a significant step in Ebola virulence [16]. Similarly the structural protein VP24 of Ebola virus plays a significant role in viral pathogenicity by antagonizing the host interferon function. This protein is also known to influence the transcription and replication of Ebola virus and in the framing of viral nucleocapsids [17]. The VP30 plays its significant role in the formation of the viral mRNAs and also in homooligomerization. The Nucleoprotein (NP), VP30, VP35, and RNA-dependent RNA polymerase (L), plays a role in the constitute of nucleocapsid a principal unit of transcription and replication of the viral genome [18]. Among these seven proteins, VP24, VP35, VP30 and VP40 were considered as the potential drug targets as they have crucial function and plays a significant role in viral replication and pathogenesis.

As such there are no drugs available for the treatment of Ebola virus infection [19]; the present study is focused on screening a group of known viral inhibitors targeting the Ebola virus major matrix protein VP40. Thus the compounds with the potential for treatment of Ebola virus infection may lead to the design and development of a novel drug for the treatment of patients with Ebola virus infection.

**METHODOLOGY**

**Potential Target**

The 3D structure of membrane-associated matrix protein (VP40) that plays a central role in the formation of filamentous virus particles is retrieved from PDB [20]. The structure of matrix protein VP40 of Ebola is determined by X-ray crystallography resolved at 1.60Å (1H2D) [21].

**Binding Site determination**

To determine the interactions between ligands and matrix protein VP40, the binding sites that favor the possible interactions with ligands were determined by submitting the co-ordinate file to DoGsite scorer [22]. The potential binding sites favoring the drug interactions were obtained as binding pockets. These binding pockets were used for further docking studies.

**Compound Library Preparation**

The Ebola VLP entry-blocking activity compounds reported in the study of Jennifer Kouznetsova et al. [23] and some anti-viral compounds were selected as ligands. The 3D structures of these compounds were retrieved in SDF format from Pubchem database [24]. All these compounds were retrieved as a single SDF for further virtual screening.

**High throughgput virtual screening**

The retrieved compounds in SDF file format from Pubchem database were docked with the amino acids in the predicted binding site of matrix protein VP40 using the following parameters [25] (i) default general docking information’s, (ii) base placement using triangle matching, (iii) scoring of full score contribution and threshold of 0.30 and No score contribution and threshold of 0.70, (iv) chemical parameters of clash handling values for protein-ligand clashes with maximum allowed overlap volume of 2.9 Å³ and intra-ligand clashes with clash factor of 0.6 and considering the hydrogen in internal clash tests, and (v) default docking details values of 200 for both the maximum number of solutions per iteration and maximum number of solutions per fragmentation.

**Prediction of ligand-receptor interactions**

The interactions of 42 compounds with matrix protein VP40 in the docked complex were analyzed by the pose-view of LeadIT [26].

**RESULT AND DISCUSSION**

Ebola infections in 2014 were the largest outbreak in West Africa. The genome of this virus is made up of single negative RNA strand of 19 kb that encodes for seven polypeptides namely glycoprotein (GP), VP40, VP24 Nucleoprotein (NP), RNA-dependent RNA polymerase (L), VP35, VP30 [27]. Among these proteins, the most abundant matrix protein VP40 located under the viral bilayer that plays a crucial role in structural integrity of the viral particles is considered as the potential target. This protein plays an important role in the Viral RNA metabolism during Ebola replication [28].

Thus considering the matrix protein VP40 and also taking in to account that there are no drugs available for the treatment of Ebola virus infection, virtual screening of known viral inhibitors targeting the Ebola virus major matrix protein VP40 was carried out. As the 3D structure of the matrix protein VP40 is already available in the form of X-ray crystal structure, it was obtained from the PDB database (Figure.3).

The N and C-terminal domains of VP40 are linked together by a flexible linker made up of residues ranging from 195 to 200. The VP40 is classified as Apha+beta as its N terminal domain is folded into a β-sandwich consisting of six anti-
parallel strands arranged in two β-sheets of three strands each [29]. This along with the C terminal domain forms the monomer of VP40 receptor. The C-terminal domain comprises of a conserved proline-rich region in VP40 EBOV receptor ranging from 205–219aa responsible for interaction with cellular Sec24C and also essential for plasma membrane targeting and viral particle release is considered as a potential drug target due to its role in membrane association. The N-terminal domain is responsible for the oligomerization of the protein [30].

A total of 43 compounds were screened for their binding affinities to signify their inhibition activity against the Ebola major membrane associated matrix protein VP40. The compounds considered in this study also include the Ebola VLP entry blocking activity compounds reported in the study of Jennifer Kouznetsova et al. 2014 [23]. The 2D structures of all these 42 compounds were shown in Figure 4. All these compounds were docked in to the predicted binding site of VP40 using the FlexX module of LeadIT. The binding efficiency of each compounds were determined in the terms of binding score. The five best docked compounds with their binding score and the binding site amino acids favoring the interactions were summarized in Table 1. The best docking compounds was selected based on the dock score (KJ/mol), that was used to rank the poses of the ligands with possible H-bond interactions and non-bonded interactions and excess internal energy of generated ligand conformation for flexible docking. The studies of Raj and Varadwaj et al [31], reported that Gossypetin and Taxifolin compound isolated from the flowers and the calyx of Hibiscus sabdariffa (roselle) and Taxus chinensis respectively, possess natural antiviral activity and also Ebola VP40 inhibition activity. Similarly, the studies of Palamthodi et al [32], reported four compounds (i) 2-(1,3-benzothiazol-2-ylsulfanyl)acetate, (ii) 2-(1,8-dihydroxy-9-oxo-10h-anthracen-2yl)acetic acid (iii) 1-[(2s,4s,5r)-4-hydroxy-5-methylxolan-2-yl]-5-methylpyrimidine,2,4 dione and (iv) 1-[(2r, 4s, 5s)-5-(hydroxymethyl)-4-methylxolan-2-yl]-1,2,4-triazole-3-carboxamide screened form ZINC database to possess inhibition activity against Ebola matrix protein VP40. Similarly, the high throughput screening studies conducted by Tamilvanan & Hopper [33] reported the compounds ASN03576800, ASN06396768,ASN05439185, ASN08735135, ASN08745583 from ASINEX database and 693, 234 from TCM database ligands occupy the RNA binding region of VP40.

Similarly, the compounds from this study also exhibited encouraging binding and docking energy (Table 1). Among these, the compounds CID-4122, CID-159269, CID-73115, CID-60786 and CID-21704 exhibited better docking with VP40 of docking scores -19.5138 kJ/mol, -16.6071 kJ/mol, -16.3261 kJ/mol, -14.7275 kJ/mol and -13.8140 kJ/mol respectively. The docking scores of all the compounds ranged from -19.5138 to -13.3496 kJ/mol.

Table 1: Amino acids interactions among the docked compounds and their docking score

<table>
<thead>
<tr>
<th>Protein</th>
<th>CID-4122</th>
<th>CID-159269</th>
<th>CID-73115</th>
<th>CID-60786</th>
<th>CID-21704</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>*Asn62</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>*Pro63</td>
<td>*Pro63</td>
<td>-</td>
<td>*Pro63</td>
</tr>
<tr>
<td>Pro63</td>
<td>Pro63</td>
<td>Pro63</td>
<td>-</td>
<td>Pro63</td>
<td></td>
</tr>
<tr>
<td>*Gln91</td>
<td>*Gln91</td>
<td>*Gln91</td>
<td>*Gln91</td>
<td>*Gln91</td>
<td></td>
</tr>
<tr>
<td>Gln91</td>
<td>Gln91</td>
<td>Gln91</td>
<td>-</td>
<td>Gln91</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>*Glu92</td>
<td>*Glu92</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Glu92</td>
<td>Glu92</td>
<td>-</td>
<td>-</td>
<td>Glu92</td>
<td></td>
</tr>
<tr>
<td>*Leu95</td>
<td>*Leu95</td>
<td>*Leu95</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Leu95</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>*Pro96</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>*Pro96</td>
</tr>
<tr>
<td>Pro96</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Pro96</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pro97</td>
<td>Pro97</td>
<td>-</td>
<td>Pro97</td>
<td></td>
</tr>
<tr>
<td>Gln99</td>
<td>-</td>
<td>-</td>
<td>Gln99</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>*Leu100</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>*Leu100</td>
<td></td>
</tr>
<tr>
<td>Leu100</td>
<td>-</td>
<td>-</td>
<td>Leu100</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>Pro101</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>Gln102</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

* Amino acids forming H-bond interactions

Docking interactions of CID-4122 with matrix protein VP40

The compound Nocodazole (CID-4122) showed the highest docking score of -19.5138 kJ/mol with the matrix protein VP40. It is observed that the compound occupied the RNA binding region of VP40 by favoring the H-bond interactions and Non-bonded interactions. The amino acids Gln91, Leu95, Pro96 and Leu100 favored the H-bond interactions and the non-bonded interactions are supported by Pro63, Gln91, Glu92, Leu95, Pro96, Pro97, Gln99 and Leu100. The docking complex and docking interactions of Nocodazole within the active site of matrix protein VP40 is shown in figure 5.
Docking interactions of CID-159269 with matrix protein VP40

The compound Telbivudine (CID-15969) showed the second highest docking score of -16.6071 kJ/mol with the matrix protein VP40. It is observed that the compound occupied the RNA binding region of VP40 by favoring the H-bond interactions and Non-bonded interactions. The amino acids Pro63, Gln91, Glu92 and Leu95 favored the H-bond interactions and the non-bonded interactions are supported by Pro63, Gln91, Glu92 and Pro97. The docking complex and docking interactions of Telbivudine within the active site of matrix protein VP40 is shown in figure 6.
Docking interactions of CID-73115 with matrix protein VP40

The compound Clevudine (CID-73115) showed the docking score of -16.3261 kJ/mol with the matrix protein VP40. It is observed that the compound occupied the RNA binding region of VP40 by favoring the H-bond interactions and Non-bonded interactions. Similar to the Telbivudine binding interactions, the compound Clevudine also bonded to the amino acids Pro63, Gln91, Glu92 and Leu95 by favoring H-bond interactions and the non-bonded interactions by Pro63, Gln91, Glu92 and Pro97. The docking complex and docking interactions of Telbivudine within the active site of matrix protein VP40 is shown in figure.7.

Docking interactions of CID-60786 with matrix protein VP40

The compound Lobucavir (CID-60786) showed the docking score of -14.7275 kJ/mol with the matrix protein VP40. It is observed that the compound occupied the RNA binding region of VP40 by favoring the H-bond interactions and Non-bonded interactions. This compound H-bond interaction is favored by Gln91 and Leu100 whereas the non-bonded interactions are favored by Gln99, Leu100, Pro101 and Gln102. The docking complex and docking interactions of Lobucavir within the active site of matrix protein VP40 is shown in figure.8.

Docking interactions of CID-21704 with matrix protein VP40

The compound Vidarabine (CID-21704) showed the docking score of -13.8140 kJ/mol with the matrix protein VP40. It is observed that the compound occupied the RNA binding region of VP40 by favoring the H-bond interactions and Non-bonded interactions. This compound H-bond interaction is supported by Asn62, Pro63, Gln91 and Pro96 whereas the non-bonded interactions are favored by Pro63, Gln91, Glu92, Pro96 and Pro97. The docking complex and docking interactions of Lobucavir within the active site of matrix protein VP40 is shown in figure.9.
In many enveloped virus, the matrix protein and RNA interactions have been reported to involve interactions with viral RNA as a crucial for the M-NP association [34]. In addition, some studies implicated that the matrix proteins interferes with the cellular RNA metabolism. The dimer interface is stabilized by the interaction with ssRNA segment [35]. It is reported the octameric VP40 target protein binds to RNA tribonucleotide at the inner pore surface of each anti parallel homodimer to form the viral matrix; hence blocking these RNA binding site may restrict this pathway of virus membrane formation [36]. The virtual screening of 42 compounds revealed that the compounds, Nocodazole, Telbivudine, Clevudine, Lobucavir and Vidarabine effectively occupied the RNA binding regions resulting in restricting the virus membrane formation. The docking studies also implies that H-bond interactions of these anti-viral inhibitors implies that the C=O (keto group) present in the compounds and NH (amino group) on the amino acids namely Glutamine, Glutamic acid and Proline observed in the active site of VP40 favors the H-bond interactions. Thus suggesting that the compound Nocodazole (CID-4122) could be a potent inhibitor for Ebola virus matrix protein VP40 and also these binding interactions can imply the design of novel compounds with Ebola inhibition activity by restricting the membrane formation pathway of Ebola virus.

**CONCLUSION**

As there are no valid drugs to treat Ebola virus infections that cause severe hemorrhagic fever, the urgency in search and design of utmost anti-viral drug has propelled our research to screen the compounds to inhibit the Ebola infections. The virtual screening of 42 anti-viral drugs were evaluated by their binding efficacy in occupying the RNA binding regions in the membrane-associated major matrix protein (VP40) was performed. The VP40 protein plays a central role in the formation of filamentous virus particles and also binds to RNA tribonucleotide to form the viral matrix was chosen as the potential drug target. The molecular docking studies suggested Nocodazole, Telbivudine, Clevudine, Lobucavir and Vidarabine were effective in occupying the RNA binding regions and can result in restricting the virus membrane formation. Further, the binding interactions of these anti-viral inhibitors implies that the C=O (keto group) present in the compounds and NH (amino group) on the amino acids namely Glutamine, Glutamic acid and Proline observed in the active site of VP40 favors the H-bond interactions. Thus suggesting that the compound Nocodazole (CID-4122) could be a potent inhibitor for Ebola virus matrix protein VP40 and also these binding interactions can imply the design of novel compounds with Ebola inhibition activity by restricting the membrane formation pathway of Ebola virus.

**REFERENCE**


Priya al
International Journal of Advanced Science and Engineering
www.mahendrapublications.com