Virion Protein 24 of Ebola Virus as a Potential Drug Target

Vivek Darapaneni1*

1 Department of virology and computational biochemistry, Saket Institute for Biomolecular Research, Visakhapatnam, India

Abstract
The virion protein 24 of ebola Virus is multifunctional protein. The virion protein of ebola Virus plays imperative roles in the virus life cycle. The objective of the present study was to identify the residue conservation in the virion protein 24 of various ebola viruses. The study was based on 173 sequences from Bundibugyo, Sudan, Tai Forest, Zaire and Reston ebola viruses. The virion protein 24 showed higher level of sequence conservation. On the whole, this study exposed residues which are universally conserved among different viral species. These universally conserved residues might be involved in either structure stabilizing or protein-protein interactions. The conserved residues identified in the present study in conjunction with structural analysis of virion protein 24 identified novel binding sites: Ile157, Ile161, Leu164; Leu164, Leu167; Leu150, Ile161; Ala99, Ile161; Asp205, Ser207; Phe76, His78 and Glu46, Asp48. The present study could form basis for designing universal anti-ebola drugs which are resistant to mutations arising in the future.

Keywords: Conservation; Universally; Drugs; Mutation; Resistance; Anti-influenza

Peer Reviewer: Sihua Peng, Department of Biology, Shanghai Ocean University, China

Received: February 14, 2015; Accepted: April 20, 2015; Published: June 5, 2015

Competing Interests: The authors have declared that no competing interests exist.

Copyright: 2015 Darapaneni V 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.

*Correspondence to: Vivek Darapaneni, Department of virology and computational biochemistry, Saket Institute for Biomolecular Research, Visakhapatnam, India; Email: darapanenivivek.sibr@gmail.com
Introduction

The Ebola virus (EBOV) causes an acute hemorrhagic fever in humans and non-human primates, with casualty rates as high as 90% [1]. Previously EBOV was responsible for an epidemic in the year 1976 in the remote areas of central Africa with 318 cases [2]. Recently EBOV is responsible for prolonged and complex epidemic in the year 2014 with 13,676 registered infected cases and a death toll of 4,910 (as on 29th October 2014) [3]. EBOV belongs to the family Filoviridae and is classified in the genus Ebolavirus, which include Bundibugyo EBOV; Sudan EBOV; Tai Forest EBOV; Zaire EBOV and Reston EBOV [4, 5]. EBOV is a filamentous shaped virus with a negative strand RNA genome which is non-segmented [4, 6]. EBOV genome encodes seven proteins namely, glycoprotein, nucleoprotein, RNA-dependent RNA polymerase, virion protein 35, virion protein 30, virion protein 40 and virion protein 24 [4].

The virion protein (VP) 24 is one of the proteins encoded by the EBOV genome, which contributes to the pathology of the infection and might be a key factor in the EBOV virulence [7]. The VP24 is 251 amino acid residues long and has a molecular weight of 29 kDa [4]. The VP24 is a multifunctional protein with several functions, which include: (i) nucleocapsid formation [8, 9] (ii) regulation of viral replication [8, 9] (iii) disrupt interferon signalling [10] (iv) inhibition of IFN-α/β-induced and IFN-γ-induced gene expression [11] (v) involvement in mimicking the host transporter/cargo interactions [12, 13] (vi) averts phosphorylation of p38 mitogen-activated protein kinase [14] (vii) prevents the relocation of nuclear ribo-nuclear protein complex C1/C2 [15]. Residues Leu115, Leu121, Asp124, Trp125, Thr128-Thr131, Asn135, Arg137, Thr138, Arg140, Lys142-Ser146, Gln184-His186 and Leu201-Ser207 of VP24 are involved in KPNA5 binding [16, 17]. A functional late domain consensus sequence (YXXXL) is located between residues Tyr172 to Leu175 [18]. VP24 was shown to oligomerize and preferentially form tetramers with N-terminal region playing important role in oligomerization [18]. The regions Val96-Leu98 and Leu106-Leu112 of VP24 are shown to be involved in VP24-STAT1 interaction [19]. When STAT1 binds to the regions Met71-Leu79 and Ile181-Leu198 of VP24, then there would be an increased flexibility in these regions [19].

At present, there are no licensed drugs which target the VP24 of EBOV. The aims of the present study are to identify degree of conservation of VP24 among all the EBOV’s of the genus Ebolavirus in order to identify universally conserved regions and map the conservation on to the protein structure. In conjunction with binding site analysis, prospective binding sites were identified, which may delineate novel in vivo protein-protein interactions sites. Furthermore, this study will provide basis for designing new anti-ebola drugs which will be effective universally.

Methods

Sequence analysis and protein structure retrieval

The protein sequences of VP24 of ebola virus were obtained from virus variation resource [20]. Full-length sequences from all hosts and species were chosen. For alignment of the obtained sequences of VP24, MUSCLE version 3.8 [21] was used with default parameters. Multiple refinements of the obtained alignment were carried out resulting in 4-5 iterations, until no further improvement was attained. The
experimental structure of the VP24 (residues 16-231) was obtained from Protein Data Bank (PDB) [22] with the identifier 4U2X [17].

Conservation analysis

By providing sequence alignment and protein structure files as an input; conserved regions on VP24 were identified and mapped onto the experimental protein structures using ConSurf server (http://consurf.tau.ac.il/) [23-26]. By taking evolutionary relationships among protein sequences into account, ConSurf algorithm produces consequential conservation scores. ConSurf algorithm gives more emphasis to those protein sequences which are evolutionarily distant, thus producing conservation scores which are significant [23-26]. The ensuing conservation scores are criterion scores with an average of 0 and a standard deviation of 1. The residues with score < 0 denote higher conservation and those with score > 0 are variable residues [23-26]. The Bayesian algorithm is used to evaluate the confidence intervals of calculated conservation scores [23-26]. The conservation score given by ConSurf server is divided into scale of nine grades which are given for the purpose of visualization [23-26]. Most variable positions are placed in grade one (turquoise), intermediately conserved positions are placed in grade five (white), and most conserved positions are placed in grade nine (maroon) [23-26].

Binding site analysis

Ligand binding sites on VP24 protein structure were identified using COFACTOR (http://zhanglab.ccmb.med.umich.edu/COFACTOR/), detects the binding sites using both global and local with templates from PDB and match local motifs of the identified template with that of query structure [27, 28, 29]; TM-site (http://zhanglab.ccmb.med.umich.edu/COACH/), detects the binding sites using intermediary approach if both local and global alignments [29, 30]; S-site (http://zhanglab.ccmb.med.umich.edu/COACH/), detects the binding sites by explicitly comparing binding site specific sequence profiles [29, 30] and SiteHound (http://scbx.mssm.edu/sitehound/sitehound-web/Input.html) which uses energy based method to find regions with high potential for ligand interactions. The structure of the protein is characterized by different probes for identification of binding sites of different types [31].

Results

Multiple alignment of VP24 protein sequences

For the VP24, 173 sequences were obtained from the database, which were mainly from Zaire EBOV (149 sequences), Bundibugyo EBOV (5 sequences), Sudan EBOV (10 sequences), Tai Forest EBOV (2 sequence), and Reston EBOV (8 sequences). The sequence alignment of the VP24, revealed the fact that albeit the sequences belong to different lineages of EBOV and were from different hosts, they had a high degree of similarity. N-terminus of the VP24 showed high degree of conservation, while the C-terminal end was found to be less conserved.
Table 1 Conserved and variable residues of VP24 identified by the ConSurf server [23-26].

<table>
<thead>
<tr>
<th>Residues</th>
<th>VP24</th>
</tr>
</thead>
</table>

Conserved and variable residues

The variable and conserved residues in the VP24 were identified using ConSurf server [23-26] and are illustrated in figure 1. The highly variable residues of grades 1-3 and highly conserved residues of grades 7-9 are clustered together in Table 1. The conservation scores were projected onto the backbone of the protein structures and are shown in Figure 1.

![Figure 1](image-url)
Conservation scores for the VP24 were obtained between -0.767 (maximum conservation) and 3.934 (maximum variability) by the ConSurf server [23-26]. In general the VP24 is conserved with 61.5% of the residues belong to grades 7-9 (conserved), while 27.5% of the residues belong to grades 1-3 (variable). Altogether, ninety two residue positions (36.7% of total residues) were found to be highly conserved (grade 9). Out of these highly conserved residues, only 57.5% of the residues were found to be exposed at the surface of the protein whereas the remaining 42.5% are buried residues. The residue positions 85, 110, 146, 151, 155, 177, 178, 207, 220, 225 and 247 showed highest conservation among all the sequences analyzed. In total, forty five residues (17.9% of total residues) were found to be highly variable (grade 1). Out of these highly variable residues, 55.5% of the residues were found to be exposed at the surface of the protein whereas the remaining 44.5% are buried residues. The residue positions 184, 211, 233, 234 and 237 showed highest variations among all the sequences analyzed.

Small ligand binding potential

Putative ligand binding sites were identified using COFACTOR (detects LBS by global and local alignments with template structures in PDB), TM-site (detects LBS by transitional approach balancing global and local alignments), S-site (detects LBS by comparing sequence profiles which are binding site specific) and SiteHound (detects LBS by positive interaction between a chemical probe and protein structure) algorithms [27-31]. By amalgamating the binding site results with that of evolutionary analysis, binding sites which are in close proximity to the regions of conservation were identified and are shown in Table 2 and Table 3.

**Table 2** Potential ligand binding sites of VP24 identified by COFACTOR, TM-site and S-site algorithms [27-30].

<table>
<thead>
<tr>
<th>Site</th>
<th>Residue position in the ligand binding site of VP24</th>
<th>Ligands</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>200, 201, 202, 219</td>
<td>2-(Hydroxymethyl) phenol [SA9]</td>
</tr>
<tr>
<td>2</td>
<td>18, 21, 25</td>
<td>L-Threitol [DTL]</td>
</tr>
<tr>
<td>3</td>
<td>222, 224, 227</td>
<td>Glycerol [GOL]</td>
</tr>
<tr>
<td>4</td>
<td>73, 74, 77, 81</td>
<td>L-Threitol [DTL]</td>
</tr>
<tr>
<td>5</td>
<td>95, 102, 123, 126</td>
<td>2'-Deoxythymidine triphosphate [TTP]</td>
</tr>
<tr>
<td>6</td>
<td>45, 96, 101, 102, 153, 154, 157</td>
<td>Adenosine Monophosphate [AMP]</td>
</tr>
<tr>
<td>7</td>
<td>100, 101, 102, 180, 181</td>
<td>Sodium [NA]</td>
</tr>
<tr>
<td>8</td>
<td>157, 161, 164</td>
<td>Tetra ethylene glycol monoocyl ether [C8E]; Retinol [RTL]; Decyl-beta-D-Maltopyranoside [DMU]</td>
</tr>
<tr>
<td>9</td>
<td>109, 156</td>
<td>Zinc [ZN]</td>
</tr>
<tr>
<td>10</td>
<td>164, 167</td>
<td>Dodecyl-beta-D-Maltoside [LMT]</td>
</tr>
<tr>
<td>11</td>
<td>150, 161</td>
<td>Bacteriochlorophyll A [BCL]</td>
</tr>
<tr>
<td>12</td>
<td>99, 161</td>
<td>Tetra ethylene glycol monooctyl ether [C8E]</td>
</tr>
<tr>
<td>13</td>
<td>58, 104, 168, 186, 193</td>
<td>Magnesium [MG]; Zinc [ZN]</td>
</tr>
<tr>
<td>14</td>
<td>24, 180</td>
<td>Magnesium [MG]; Manganese [MN]</td>
</tr>
<tr>
<td>15</td>
<td>205, 207</td>
<td>Magnesium [MG]</td>
</tr>
<tr>
<td>16</td>
<td>76, 78</td>
<td>SUD</td>
</tr>
<tr>
<td>17</td>
<td>46, 48</td>
<td>Zinc [ZN]</td>
</tr>
</tbody>
</table>
Table 3 Predicted ligand binding sites of VP24 identified using SiteHound algorithm [31].

<table>
<thead>
<tr>
<th>Site</th>
<th>Residue positions in the ligand binding site of VP24</th>
<th>Site volume ($\text{Å}^3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>63, 64, 65, 66, 67, 68</td>
<td>28</td>
</tr>
<tr>
<td>2</td>
<td>34, 36, 38, 48, 49, 50, 52, 225</td>
<td>9</td>
</tr>
<tr>
<td>3</td>
<td>61, 80, 81, 82, 83</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>46, 47, 48, 229, 230</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>56, 60, 69, 70, 71, 72, 75, 76, 169, 170, 172, 192, 193, 194, 195</td>
<td>141</td>
</tr>
<tr>
<td>6</td>
<td>110, 111, 112, 115, 144, 146, 148, 149, 152</td>
<td>89</td>
</tr>
<tr>
<td>7</td>
<td>113, 114, 117, 118, 121, 137, 138, 139, 140, 141</td>
<td>85</td>
</tr>
<tr>
<td>8</td>
<td>76, 78, 170, 171, 175, 176, 177, 178, 180, 189, 191, 200, 209</td>
<td>37</td>
</tr>
</tbody>
</table>

One of the important objectives of the present study was to detect potential drug binding sites which are spatially in close proximity to the conserved regions. Site 8, site 10–site 12, site 15–site 17 identified by the COFACTOR, TM-site and S-site algorithms and site 8 detected by the SiteHound algorithm are found to be completely conserved. While the remaining ligand binding sites detected on the VP24 and shown in Table 2 and Table 3 were found to be located in close proximity to the conserved regions of the protein.

**Discussion**

**Sequence conservation**

The objective of the present study was to determine the degree of conservation of the VP24 among all the of ebola viruses belonging to the genus *Ebolavirus* and from different hosts. The VP24 from all ebola viruses belonging to the genus *Ebolavirus* were analyzed together to facilitate the identification of universally conserved residues of potential epidemic ebola viruses that might arise in future due to either a event of mutation. The conserved residues detected on the VP24 may have either functional importance or structural importance [32]. On the contrary, variable sites arise as a result of either adaptation or evolutionary pressure to evade the host immune system.

In the regions responsible for the VP24-KPNA5 interactions, residues Leu115, Leu121, Asp124, Trp125, Thr128, Thr129, Arg137, Thr138, Lys142, Gln144-Ser146 and Gln202- Ser207 were found to be conserved while the remaining residues were found to be variable residues. In the functional late domain consensus sequence, residues Tyr172, Gly174 and Leu175 were found to be conserved among all the viruses analysed. In the VP24-STAT1 interaction region, residues Val96-Leu98, Leu106, Asp108, Ser110, Leu111, Gln144-Leu115, Gln117-Leu119, Leu121, Thr72- Leu79, Ile181-Thr183, Ile188-Gly196 and Leu198 were found to be conserved among all the viruses. The residues Met1-Leu10, Pro13, Lys14, Glu18, and Gly20 play a pivotal role in the oligomerization of VP24.
Novel binding sites

Most of the predicted LBS identified on the VP24 are surrounded by conserved residues. Targeting these LBS using small molecule ligands will most probably result in disruption of VP24 function. Among the predicted binding sites identified by different methods, few novel binding sites were detected. These novel sites may be either involved in new protein-protein interactions or sites of known protein-protein interactions. The novel binding sites identified are Site 8, site 10–site 12, site 15–site 17 identified by the COFACTOR, TM-site and S-site algorithms and site 8 detected by the SiteHound algorithm. These novel sites are completely conserved and the functions of these sites are not known from previous studies. Targeting these sites using small ligand molecule would inhibit the functions of the VP24 protein. Apart from the novel sites, conserved sites were also identified. These sites might be functionally important which can be explained by the degree of conservation found in these sites and the functions of these sites are not elucidated yet.

Conclusion

In conclusion the study of VP24 protein sequences divulge an elevated level of sequence conservation pattern that intersects with prospective ligand binding site analysis rendering the VP24 protein an exceptional drug target. Targeting the conserved binding sites identified in this study using small drug molecule will effectively diminish the activity of the VP24 protein. Moreover, anti-VP24 drugs targeting these conserved sites are less likely to become ineffective due to drug resistance in the future.

Acknowledgement

The author received no funding for this work.

References


