Molecular determinants of antigenicity of two subtypes of the tick-borne flavivirus Omsk haemorrhagic fever virus

Li Li, Pierre E. Rollin, Stuart T. Nichol, Robert E. Shope, Alan D. T. Barrett and Michael R. Holbrook

1Department of Pathology and Center for Biodefense and Emerging Tropical Diseases, University of Texas Medical Branch, Galveston, TX 77555-0609, USA
2Special Pathogens Branch, Division of Viral and Rickettsial Diseases, Centers for Disease Control and Prevention, 1600 Clifton Road, Atlanta, GA 30333, USA

In 1964, D. H. Clarke defined two antigenic subtypes of Omsk haemorrhagic fever virus (OHFV) based on polyclonal antibody absorption and haemagglutination assays. The current report defines the molecular basis for these antigenic subtypes by comparison of the complete genomes of OHFV strains Kubrin (subtype I) and Bogoluvovska (subtype II). There were six nucleotide differences between these two strains throughout the entire genome and they encoded four amino acid changes including three in the viral envelope (E) protein. Two of these changes were in solvent-exposed regions of domain 3 of the E protein, one of which lies in a region that could easily function in virus–host cell or virus–antibody interactions. These results demonstrate the minimal changes that are required to significantly alter the antigenicity of flaviviruses and also demonstrate the tremendous genetic stability of the tick-borne flaviviruses.

The Kubrin strain of Omsk haemorrhagic fever virus (OHFV) was isolated from human blood in 1947, from the Omsk Oblast, USSR. This particular virus has been passaged at least 27 times through suckling mouse brains (SMBs). The stock came from pool 926 generated 26/6/1962 (day/month/year) by D. H. Clarke. The Bogoluvovska strain was isolated from Dermacentor marginatus ticks. The date and location of isolation of this virus are unknown. The sequenced virus stock is approximately SMB passage 24 and Vero E6 cell passage 2. D. H. Clarke stock was from 3/11/1961.

Omsk haemorrhagic fever virus (OHFV) is a member of the genus Flavivirus of the family Flaviviridae and is endemic to a localized region of Siberia, Russia. OHFV is transmitted by Dermacentor spp. ticks and is considered to be a potential biothreat agent. The flaviviruses are small positive-sense RNA viruses that are typically transmitted by either mosquitoes or ticks and include such major human pathogens as dengue virus, yellow fever virus, Japanese encephalitis virus, West Nile encephalitis virus and tick-borne encephalitis (TBE) virus. The tick-borne flavivirus serocomplex of viruses that affect animals (cf. birds) is effectively grouped into four major subtypes: TBE Western subtype, Siberian subtype and Far-eastern subtype and louping ill viruses (Fig. 1). Several other tick-borne flaviviruses make up this serocomplex including Powassan virus, Langat virus and three viruses that cause haemorrhagic disease, Alkhurma virus, Kyasanur Forest disease virus and OHFV.

OHFV was initially isolated in 1947 in Siberia from a man presenting with haemorrhagic fever (Burke & Monath, 2001). OHFV is endemic to a fairly localized region of Siberia within the Omsk and Novosibirsk Oblasts. OHFV is maintained in nature through circulation among ticks and rodents including water voles (Arvicola terrestris) and muskrats (Ondatra zibethica) (Kharitonova & Leonov, 1985). The primary tick vector in the native environment for this virus is Dermacentor reticulatus (Kharitonova & Leonov, 1985). The use of this vector for OHF transmission immediately differentiates this virus from other TBE viruses endemic to the same region, which are transmitted by Ixodes spp. of ticks. OHFV infection of humans is frequently associated with occupation, as muskrat trappers are common victims of OHF. OHFV causes a haemorrhagic disease infrequently associated with encephalitis and has a case fatality rate of 0.5–3%. Recently, members of this laboratory described the complete genome sequence of the OHFV strain Bogoluvovska (GenBank accession no. AY193805; Lin et al., 2003). This virus was found to be similar to other tick-borne flaviviruses at both the nucleotide and amino acid level. Thirty-five amino acid residues were identified throughout the genome of haemorrhagic tick-borne flaviviruses [OHFV and Alkhurma virus (ALKV)]
that were different from those in encephalitic tick-borne flaviviruses. These included four amino acid residues in the viral envelope (E) protein.

The flavivirus genome encodes a single open reading frame (ORF) that is co- and post-translationally cleaved to generate three structural and seven non-structural proteins. The structural proteins include the core protein, the pre-membrane/membrane (M) protein and the viral E protein. The M and E proteins are found on the surface of mature virions. The flavivirus E protein is the major viral surface protein and the primary antigenic domain for these viruses. Many monoclonal antibodies have been mapped to the E protein. The E protein has been shown in cryo-electron microscopy studies to exist on the virus surface as a dimer (Kuhn et al., 2002; Zhang et al., 2003). The E protein crystal structure for Western subtype TBE strain Neudorfl has been solved (Rey et al., 1995). This structure contained the first 400 of the 500 amino acids within the E protein, lacking the C-terminal 100 residues that compose the stem–anchor region. The crystal structure clearly differentiated the E protein into three distinct domains that corresponded to previously defined antigenic domains (Heinz et al., 1991; Mandl et al., 1989). Domain I was the central domain and contained potential glycosylation sites and domain II was the dimerization domain, named due to apparent extensive interactions with the second E protein monomer in the dimer pair. Domain III (D3) was considered the putative receptor-binding domain due to its Ig-like fold and an RGD integrin-binding motif in some mosquito-borne flaviviruses. In addition, recombinant E protein D3 binds directly to host cells (M. R. Holbrook, unpublished results).

The description of two antigenic subtypes of OHFV in 1964 was based initially on gel-diffusion assays, which suggested a variation between the strains but could not clearly differentiate them (Clarke, 1964). More conclusive assays using serum absorption of viral antigen followed by haemagglutination inhibition (HI) assays to test residual HI activity clearly separated the viruses into two antigenic subtypes. Subtype I includes the Kubrin (isolated from human blood) and Balangul (isolated from D. marginatus) strains, while subtype II includes Bogoluvovska (from D. marginatus) and Guriev (from human blood) strains. Differences in pathogenicity between the subtypes have not yet been examined. The current study extends those findings by using molecular techniques to examine viruses from each antigenic subtype in an effort to determine the molecular determinants for antigenic specificity.

![Fig. 1. Phylogenetic maximum-likelihood tree of the TBE serocomplex E protein gene. The serocomplex is divided into three major subtypes: Western subtype, Siberian subtype and Far-eastern subtype. The tree was rooted with yellow fever type strain Asibi.](image-url)
In this study, the complete genome of the Kubrin strain of OHFV (antigenic subtype I) was sequenced as described previously (Lin et al., 2003) for the Bogoluvovska strain (antigenic subtype II). Viral RNA from the Kubrin strain was isolated from suckling mouse brain 5 days after intracerebral inoculation. This virus was derived from a pool generated by D. H. Clarke in 1962. The Kubrin strain was found to be 10,787 nt in length with a 132 nt 5’ UTR and a 413 nt 3’ UTR. The resulting ORF was 10,242 nt long encoding a 3,414 aa ORF (Table 1). The Kubrin strain of OHFV was precisely the same length as the previously sequenced Bogoluvovska strain. To determine potential molecular determinants of antigenicity, a comparison was made between the nucleotide and amino acid sequences for the two viruses. This comparison identified a total of six nucleotide changes throughout the virus ORF, four of which encoded amino acid changes, and no changes within the 5’ or 3’ non-coding regions (Table 2). Three of the amino acid changes were in the viral E protein and two of these fell in E protein D3, the putative receptor-binding domain of the flaviviruses. The E364 change lies on the upper surface of D3 (Fig. 2) in a region that has previously been identified as a binding epitope for West Nile D3-specific monoclonal antibodies (Beasley & Barrett, 2002), and thus is clearly antigenic and may potentially function as a receptor-binding epitope for these viruses based on its physical location. The E349 change of Ser→Ala lies at the outer bottom edge of D3. This part of D3 could interact with other E proteins on the surface of the virion but also appears to be solvent-exposed in those D3s that lie within the fivefold axis of symmetry based on cryo-EM reconstructions of dengue and yellow fever virus (Kuhn et al., 2002; Zhang et al., 2003). This residue is unlikely to be able to interact with antibodies given its position, but might be able to affect the positioning or orientation of D3, as well as the stability of interactions with the second E protein monomer.

Comparison of the D3 amino acid sequences of several tick-borne flaviviruses demonstrated the high degree of homology within this region of the viral E protein (Fig. 3). The two amino acid changes at E349 and E364 were at positions that are fairly well conserved, but not completely so. Position E349 is predominately serine residues, though amphipathic valine residues are tolerated in the genetically distinct LGT and POW viruses and a glutamic acid residue is found in Alkhurma and Kyasanur Forest disease viruses, two haemorrhagic fever viruses. Amino acid residues at position E364 are either isoleucine or methionine residues so it is curious how this subtle change at this position might affect the antigenic differences observed previously between the Bogoluvovska and Kubrin strains of OHFV. However, Ryman et al. (1997) have shown that a conservative amino acid substitution at E240 (Ala→Val) can alter the mouse neurovirulence phenotype of yellow fever virus.
Fig. 3. Amino acid alignment of the E protein D3 from residue 300 of several TBE serocomplex viruses including both Kubrin (OHFV-K) and Bogoluvovska (OHFV-B) strains of OHFV. Also included are the haemorrhagic fever virus KFDV, the BSL-2 virus LGTV, BSL-3 virus POWV and the TBE viruses Sofjin (SOF Far-easter subtype), Vasilchenko (VAS Siberian subtype) and the Western subtype virus strain Hypr (HYPR), Kumlinge (KUM) and Neudorfl (NEUD). The genetic relatedness of these viruses is shown in Fig. 1. Numbering is based on the CEE E protein sequence from Rey et al. (1995). Asterisks indicate that residues are associated with nucleotide differences in the genome are shown as space-filling models of Van der Waals interactions. Modelling was performed using the Swiss-model program (http://us.expasy.org/spdbv/mainpage.html).

<table>
<thead>
<tr>
<th>Residue</th>
<th>OHFV-B</th>
<th>OHFV-K</th>
<th>KFDV</th>
<th>POWV</th>
<th>LGTV</th>
<th>SOF</th>
<th>VAS</th>
<th>HYPR</th>
<th>KUM</th>
<th>NEUD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>KGLTYTMCNK</td>
<td>KGLTYTMCNK</td>
<td>KGMYTCNG</td>
<td>KGTTYSMCNK</td>
<td>KGLTYTMCNK</td>
<td>KGLTYTMCNK</td>
<td>KGLTYTMCNK</td>
<td>KGLTYTMCNK</td>
<td>KGLTYTMCNK</td>
<td>KGLTYTMCNK</td>
</tr>
<tr>
<td>300</td>
<td>AKFTWKRAPT</td>
<td>AKFTWKRAPT</td>
<td>SKFANAKRPT</td>
<td>AKEKWKVRPV</td>
<td>TKFTWKRIPR</td>
<td>TKFTWKRIPR</td>
<td>TKFTWKRIPR</td>
<td>TKFTWKRIPR</td>
<td>TKFTWKRIPR</td>
<td>TKFTWKRIPR</td>
</tr>
<tr>
<td>310</td>
<td>DSGHDTVVME</td>
<td>DSGHDTVVME</td>
<td>DSGHDTVVME</td>
<td>DSGHDTVVME</td>
<td>DSGHDTVVME</td>
<td>DSGHDTVVME</td>
<td>DSGHDTVVME</td>
<td>DSGHDTVVME</td>
<td>DSGHDTVVME</td>
<td>DSGHDTVVME</td>
</tr>
<tr>
<td>320</td>
<td>VAFSGT-KPC</td>
<td>VAFSGT-KPC</td>
<td>VTYTG-KPC</td>
<td>VSYTG-SPK</td>
<td>VAFSGT-KPC</td>
<td>VAFSGT-KPC</td>
<td>VAFSGT-KPC</td>
<td>VAFSGT-KPC</td>
<td>VAFSGT-KPC</td>
<td>VAFSGT-KPC</td>
</tr>
<tr>
<td>330</td>
<td>RIPVRVAGH</td>
<td>RIPVRVAGH</td>
<td>RIPVRVAGH</td>
<td>RIPVRVAGH</td>
<td>RIPVRVAGH</td>
<td>RIPVRVAGH</td>
<td>RIPVRVAGH</td>
<td>RIPVRVAGH</td>
<td>RIPVRVAGH</td>
<td>RIPVRVAGH</td>
</tr>
<tr>
<td>340</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 2. Molecular models of E protein D3 from OHFV strains Bogoluvovska and Kubrin. Amino acid residues associated with nucleotide changes in the genome are shown as space-filling models of Van der Waals interactions. Modelling was performed using the Swiss-model program (http://us.expasy.org/spdbv/mainpage.html).
One coding change was also found in the stem–anchor region of the E protein at residue E437, which lies within the predicted helix 2 of the stem–anchor region (Allison et al., 1999). Helix 2 has been proposed to function in stabilization of interactions with the viral pre-membrane/membrane protein (Allison et al., 1999). A mutation here could theoretically affect the orientation of the E protein enough to affect the antigenicity, but this scenario seems unlikely given the essentially conservative change at E437 (Leu→Val).

The final coding change in NS4B is an Ile→Val change at NS4B179. As the structure and function of NS4B are unknown it is hard to predict whether such a change would affect the function of the protein. Residue NS4B179 lies immediately adjacent to the fourth of four predicted transmembrane domains. The location of this change, much like that at E437, could theoretically affect the structure of NS4B, but again, given that the Ile→Val change is chemically very conservative, it is unlikely that this mutation would have a dramatic effect on NS4B structure or virus function.

When sequencing the ends of the genome, a procedure incorporating decapping and RNA ligation steps (Mandl et al., 1991) was used and sequences were compared from several clones to be certain of the integrity of the 5' and 3' termini. When sequencing the Bogoluvovska strain, all clones that were sequenced had the complete and correct 5' and 3' ends. When sequencing the Kubrin strain for this study, approximately half of the clones were lacking the 3'-terminal 4–6 nt. Based on many studies on several flaviviruses and a model presented by Khromykh et al. (2003), these nucleotides are absolutely required for virus viability. Therefore, it seems likely that the loss of these nucleotides is due to RNA degradation rather than having a functional role in regulation of virus replication. However, the fact that a large percentage of clones were lacking 4–6 nt suggests that these viruses might not have an absolute requirement for these nucleotides or, more likely, that they are not associated with secondary structures and can be degraded easily by cellular RNases. This is an interesting finding that will be examined more closely.

The fact that only six nucleotide differences were identified within the approximately 11 kb genomes between these two viruses indicates the stability of the tick-borne flaviviruses. The OHFV strains examined here were isolated from either a human sample (Kubrin) or a tick (Bogoluvovska) and were passaged a number of times (>20) through SMB prior to sequencing (see first paragraph). Given the large number of SMB passages, it is possible that the viruses had adapted to the mouse. However, the few differences between these viruses also suggest that the source of the isolate does not affect the genetic stability of the virus. Tick-borne flaviviruses exist in a significantly different ecosystem to mosquito-borne flaviviruses. Tick-borne flaviviruses typically spend a considerable amount of time in the arthropod vector as the life cycle is considerably longer (~3 years) compared with that of a mosquito (~3 weeks). Subsequently, tick-borne flaviviruses are not required to adapt to different hosts as frequently, thus reducing selective pressure due to host/vector variability. This decrease in selective pressure would provide a much more stable, or less variable, viral genome.

To examine this more extensively, alternative passage through animals and ticks would need to be performed. This is a technically challenging set of experiments, given the safety restrictions required for OHFV.

This study has identified molecular determinants for antigenicity between the two antigenic subtypes of OHFV, as defined by D. H. Clarke in 1964. Only two amino acid changes in the E protein D3 and one in the E protein stem–anchor region were associated with the antigenic subtypes. In addition, the incredibly high identity between the two viruses, which were isolated from an arthropod and a human source, demonstrate the tremendous stability of OHFV within its natural environment.

**Acknowledgements**

The authors would like to thank the UTMB Protein Chemistry Core facility for their help with DNA sequencing. This work was supported by a T32 Postdoctoral Fellowship in Emerging Infectious Diseases (M. R. H.) and the Centers for Disease Control and Prevention (R. E. S., A. D. T. B.).

**References**


