



Scholars Research Library

Der Pharmacia Lettre, 2014, 6 (3):255-264  
(<http://scholarsresearchlibrary.com/archive.html>)



## ***In silico* structural analysis of Hantaan virus glycoprotein G2 and conserved epitope prediction for vaccine development**

**Monzilur Rahman and Md. Masud Parvege**

*Department of Genetic Engineering and Biotechnology, University of Dhaka, Dhaka, Bangladesh*

### **ABSTRACT**

*Hantaan virus (HNTV), a prototypic member of Hantavirus genus, is an etiological agent of potentially fatal hemorrhagic fever with renal syndrome (HFRS). The virus infects a large number of patients annually, with a mortality rate more than 10%. However, no treatment option or vaccine is available against the virus. The virus expresses two envelope proteins having role in viral attachment to host. Among them HNTV glycoprotein G2 has higher antigenicity making it a better target for vaccine development. However, 3-D structure of the protein is not available which is important for identifying epitopes that are essential for vaccine design. Therefore, this study was designed to predict a structural model of glycoprotein G2 and to predict peptide sequences for vaccine development containing conserved epitopes within the structure. Many of the physio-chemical and structural properties including secondary structure and di-sulfide linkage of the protein were predicted using a number of computational tools. The 3D structure of the protein was modeled using I-TASSER online tool. The quality of the predicted models was evaluated with Ramachandran plot and Z-score. The structural and sequence information was used to predict B-cell and T-cell epitopes on glycoprotein G2. A model with reliable quality was generated. Using various bio-informatics and immuno-informatics tools, a total of 9 continuous B-cell and 22 T-cell epitopes were predicted having significant antigenicity. These antigenic epitopes were further analyzed for conservation and a total of 4 B-cells and 8 T-cell epitopes were found to be highly conserved in sequences from diverse origins. These epitopes revealed by the current study are recognized by immune system to protect host from HNTV infection can be potential targets for vaccine development.*

**Keywords** vaccine, hantaan, virus, glycoprotein G2

### **INTRODUCTION**

*Hantaan virus (HNTV) belongs to the Hantavirus genus within the Bunyaviridae family. This virus can cause hemorrhagic fever with renal syndrome (HFRS). The disease, HFRS, can be fatal during outbreaks resulting in a mortality rate of more than 10% [1-3]. About 60,000 – 100,000 cases are reported annually worldwide, and the infection is predominant in the East Eurasian continent, including China, Korea and far East Russia [4]. This disease has taken a heavy toll on the Chinese people health. Approximately, 1.2 million symptomatic infections and 44,300 deaths have been reported [5]. Therefore, development of remedy against the virus e.g. treatment option for the disease and immunization by vaccine is necessary. Unfortunately, no known antiviral treatment option against the virus is currently reported [6,7]. However, several attempts have been made to develop vaccine against the virus.*

Strategies that have already been carried out to develop vaccine against HNTV are mainly limited to live attenuated form of the virus and DNA vaccine. DNA vaccine is currently under trial [8,9]. However, no attempt has been made to predict or design peptide vaccine against the virus. To our best knowledge, there is a lack of information about the structure of proteins of the virus and any computational analysis has not yet been done that can help develop effective inhibitors against the virus. With the advent of *in silico* methods, it is possible to predict the structure of proteins using homology based modeling and inhibitors against the virus.

The virus contains negative sense single stranded RNA genome. The genome is segmented into three parts: a small (S) segment encoding capsid proteins, a medium (M) segment encoding the envelope glycoproteins G1 and G2, and a large (L) segment encoding viral RNA dependent RNA polymerase [10-12]. G1 and G2 are first produced as a single polypeptide chain and then the polypeptide is cleaved to form separate G1 and G2 [12]. The two glycoproteins undergo maturation to be associated with each other [13]. Most of portions of G1 glycoprotein reside in transmembrane region while G2 is more exposed to the outer environment [14].

The sequence and structure of protein may provide important insight for designing effective vaccine against the virus. Therefore, the aim of this study was to predict the structure of G2 glycoprotein and vaccine target within the sequence of G2 glycoprotein. G1 glycoprotein was not considered because it is embedded in the envelope while G2 is more exposed to the outer environment for immune reaction [14]. Conserved B-cell and T-cell epitopes were predicted that can be used further to design effective vaccine against the virus.

## MATERIALS AND METHODS

### Protein retrieval and homology modeling

G2 glycoprotein sequence was retrieved from ref\_seq database in NCBI (<http://www.ncbi.nlm.nih.gov>). The accession of the sequence was NP\_941988.1. It was checked if any 3-D structure was available for the protein in Protein Data Bank (PDB) [15]. As there was no available structure, the three dimensional structure of the protein was predicted using I-TASSER online server [16]. This tool generates five structural models for the protein and assigns confidence score (C-score) for each model. The model with maximum C-score is the best model. Primary structure analysis was performed by ProtParam online tool [17]. Different parameters computed using the tool were molecular weight, theoretical pI, amino acid composition, extinction coefficient, estimated half-life, instability index, aliphatic index, and grand average of hydropathicity (GRAVY). Protein secondary structure was analyzed using SOPMA [18]. DiANNA tool was used to check disulfide connectivity of the protein [19].

### Stereochemical analysis and model evaluation

The models were further evaluated through Psi/Phi ramachandran plot and their Z-score. Overall quality of a model is evaluated by Z-score which was determined using PROSA web tool [20]. To construct ramachandran plot PROCHECK software was used through PDBsum online tool [21]. Ramachandran plot assesses the stereochemical quality of 3-D structure analyzing residue-by-residue geometry and overall structure geometry. Finally, in combination with all the evaluation and validation methods a structure was predicted for the G2 glycoprotein of HNTV.

### Prediction of B-cell and T-cell epitope

B-cells and T-cells are the main players in provoking immune response against an antigen. The antigenicity of G2 glycoprotein was evaluated using VaxiJen online antigen prediction server [22]. Transmembrane topology was checked using TMHMM online tool [23]. Using hidden Markov model this tool predicts the locations of transmembrane, intracellular and extracellular regions. The transmembrane and exo-membrane regions of the protein were analyzed for the presence of T-cell epitopes. Propred-1 and propred predicts epitopes for 47 MHC class-I alleles and 51 MHC class-II alleles respectively [24,25]. A 5% threshold was set for MHC class-I alleles to filter proteasome and immune proteasome. Proteosomal cleavage site of MHC binder at C-terminal indicates greater chances for being T-cell epitopes [26]. BCPred online server was used to predict B-cell epitopes taking 75% specific criteria for epitope prediction [27]. For this purpose only protein segment exposed on the surface of the membrane was taken into account. Again, the antigenicity was determined using VaxiJen online server [22]. Discontinuous B-cell epitopes were predicted using Disco Tope 2.0 server using 3-D structure of the G2 glycoprotein [28].

### Conservancy analysis of the predicted epitopes

A total of 91 HNTV envelope glycoprotein sequences having diverse origins were downloaded from the database and the G2 glycoprotein portion was truncated from each of the given sequence. Altogether, these 91 sequences were subjected to multiple sequence alignment using MEGA software [29]. This result was further analyzed and conservation of selected epitopes was evaluated using IEDB conservation analysis tool [30].

## RESULTS

### Analysis of the predicted protein structure

The protein sequence was downloaded from the NCBI protein database using accession NP\_941988.1. The primary structure analyzed by ProtParam online tool showed that the G2 protein is 478 amino acids long having a molecular weight of 52904.2 Daltons. The computed theoretical isoelectric point (pI) was 5.65 indicating a negatively charged protein. The computed instability index (II) was 37.76 referring the protein as stable. The negative Grand average of hydropathicity (GRAVY) of -0.116 indicated that protein is hydrophilic in nature. Glycine (9.6%), Serine (8.2%), Leucine (7.9%), Threonine (7.3%) and Isoleucine (6.7%) are most predominant. SOPMA secondary structure prediction results showed that it has 19.04% alpha helices, 27.62% extended strand, 10.25% beta turns and 43.10% coils. Total 27 cysteines are present in the sequence. Disulfide bonds predicted by DiANNA are shown in the Table 1. Disulfide connectivity was predicted to be in between amino acids 1-4, 2-15, 3-26, 5-10, 6-16, 7-8, 9-12, 11-14, 13-19, 17-21, 18-25, 20-22, and 23-24. This knowledge might provide an important insight in understanding the protein secondary structure because disulfide bridges play important roles in protein fold stabilization. Protein 3D structure is pivotal in understanding of its function, interaction with other proteins, localization and many other aspects of it. Therefore, 3D structure of the protein was predicted using I-TASSER online server. The tool predicted five possible structural models and assigned confidence score (C-score) for each of the models. C-score for five predicted structures of HNTV glycoprotein G2 were 2.30, 2.64, 2.47, 1.91 and 1.64. Model with the maximum C-score (2.64) was selected as the best predicted structure for HNTV G2 glycoprotein. The structure was visualized using UCSF Chimera. The structure was further evaluated using Ramachandran plot and Z-score. The result of the Ramachandran plot showed 72.5% residues in the favorable region. The Z-score, indicative of overall model quality, is used to check whether the input structure is within the range of scores typically found for native proteins of similar size. The Z-score of the protein was -4.91 (Figure 1). The Ramachandran plot and Z-score results ascertained the quality of the homology model of G2 protein.

Table 1 Predicted disulfide bonds by DiANNA tool

Predicted bonds	
87 - 117 KTSFHICYGACT - ETSWGCNPSDC	184 - 237RHKVVCIIIGTV - DKGFLCPEFPG
91 - 248 HCYGACTKYEY - SFRKKCNFATT	219 - 352WCTSTCQFGDP - VSLTECPFTLT
103 - 441WHTAKCHYERD - QCGIKCWFVKS	322 - 367LGENPCKIGLQ - CDKAICYGAES
122 - 175CNPSDCPGVGT - IDMNDCFVSRH	345 - 437GFTLTCLVSLT - DGAPQCGIKCW
129 - 256GVGTGCTACGL - ATTPICEYDGN	362 - 397TSIKACDKAIC - GSTFRCCHGED
132 - 158TGCTACGLYLD - YSRRVCVQFGE	398 - 403STFRCCHGEDC - CHGEDCSQIGL
167 - 215GEENLCKIIDM - IFKHWCTSTCQ	

### Epitope Prediction

Overall antigenicity of the protein predicted by VaxiJenv2.0 online server was 0.5977 indicating it as a probable antigen. TMHMM online tool was used to check the transmembrane protein topology and results showed that residues 1 - 452 are located outside of the envelope while residues 453 - 475 are within the transmembrane region and residues 476 - 478 are found to be presented inside the envelope.

The portion of an antigen which binds to an antibody molecule is called an epitope or antigenic determinant. When an antigen is a protein molecule, epitope may be either a short peptide sequence from the protein or a patch of atoms on the protein surface in the three-dimensional space. Both continuous and discontinuous epitopes are extremely important for immunity against viral infection. B-cell continuous epitope prediction was carried out using BCPred server. Epitope length was set as 10 amino acids with 75% classifier specificity while the overlapping filter was on. A total of 21 B-cell epitopes were predicted by BCPred server. All the epitopes were found either located in the transmembrane region or exposed outside. Antigenic score for each of the epitopes was calculated using VaxiJen online server and it was found that only 9 out of 21 had probable antigenic properties and all the 9 epitopes were exposed outside (Table 2). Moreover, Disco Tope 2.0 server used to predict discontinuous B-Cell epitopes from the protein 3D structure showed that a total 29 epitopic locations were found (Table 3, Figure 2).

B- Cell epitope prediction

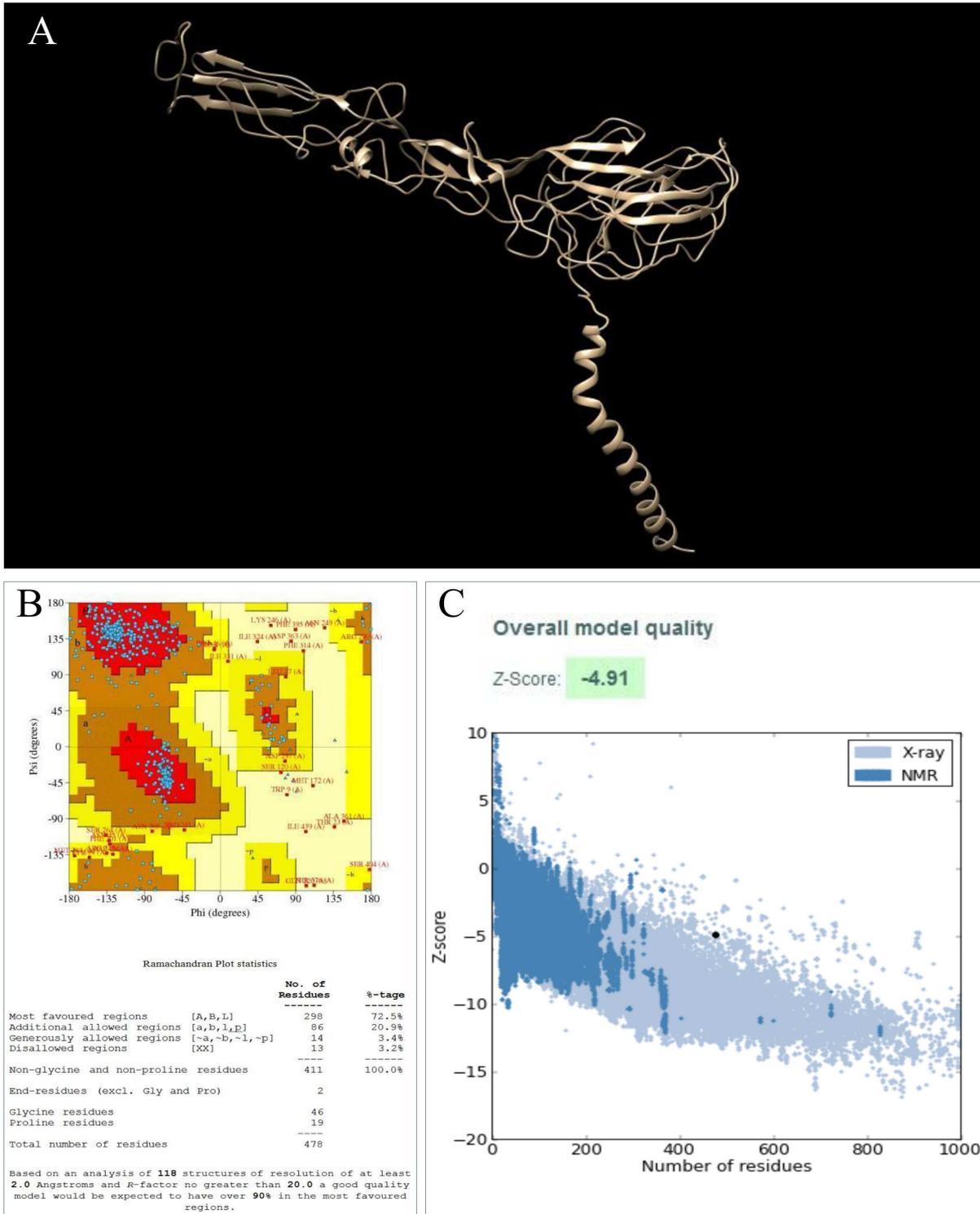


Figure 1 3D structure of hantaan virus glycoprotein G2. A. Predicted 3D structure of the glycoprotein G2 using homology modeling B. Ramachandran plot showing residues in the most favorable and disallowed regions. C. Z-score showing the quality of 3D structure.

Table 2 Linear B-Cell epitopes with their antigenic score

Position	Epitope	VaxiJen score	Comments	Position	Epitope	VaxiJen score	Comments
201	FGPLEGGGLI	0.0723	Non-antigen	189	VSKFSQGDTL	0.0795	Non-antigen
43	LTNPLEEAQS	0.2558	Non-antigen	19	VPMHTDLELD	0.9948	Antigen
113	TSWGCNPSDC	0.1504	Non-antigen	257	EYDGNMVSGY	0.1221	Non-antigen
1	SETPLTPVWN	0.3557	Non-antigen	313	DFDNLGENPC	0.1107	Non-antigen
55	LHIEIEEQTI	1.132	Antigen	83	TSFHCYGACT	0.9770	Antigen
135	YLDQLKPVGS	0.1589	Non-antigen	399	HGEDCSQIGL	1.3138	Antigen
372	SVTLTRGQNT	1.1341	Antigen	294	KDPDGMLRDH	-0.284	Non-antigen
431	DDGAPQCGIK	-0.164	Non-antigen	443	FVKSGEWISG	0.6316	Antigen
219	CQFGDPGDIM	0.7637	Antigen	385	SGKGGHSGST	1.2906	Antigen
234	GFLCPEFPGS	0.1712	Non-antigen	94	YEYPWHTAKC	0.4521	Antigen
345	CLVSLTECPT	0.0444	Non-antigen				

Table 3 Amino acid residues in the discontinuous epitopes predicted from glycoprotein G2 structure

Positions	Residues	Contact number	Propensity score	DiscoTope score	Positions	Residues	Contact number	Propensity score	DiscoTope score
67	ASP	3	-2.964	-2.968	290	ARG	13	-1.815	-3.101
114	SER	0	-2.396	-2.120	295	ASP	11	-2.534	-3.508
115	TRP	5	-3.439	-3.618	297	ASP	13	-2.172	-3.417
116	GLY	0	-3.845	-3.403	299	MET	13	1.543	-0.129
149	ILE	3	-3.399	-3.353	300	LEU	12	0.941	-0.547
260	GLY	5	-3.445	-3.624	301	ARG	14	1.407	-0.364
261	ASN	2	-3.794	-3.588	302	ASP	9	-0.677	-1.635
278	SER	11	-2.729	-3.680	420	GLY	6	-1.488	-2.007
282	SER	9	0.086	-0.958	421	ILE	3	-1.164	-1.375
283	THR	26	1.433	-1.722	422	SER	1	0.263	0.117
284	MET	18	0.542	-1.591	423	GLU	10	0.785	-0.455
285	HIS	6	-0.257	-0.917	424	ILE	1	1.284	1.022
287	THR	12	-2.610	-3.690	425	GLU	8	1.102	0.056
288	ASP	14	-1.555	-2.987	426	ASN	17	-1.688	-3.449
289	GLU	1	-0.489	-0.548					

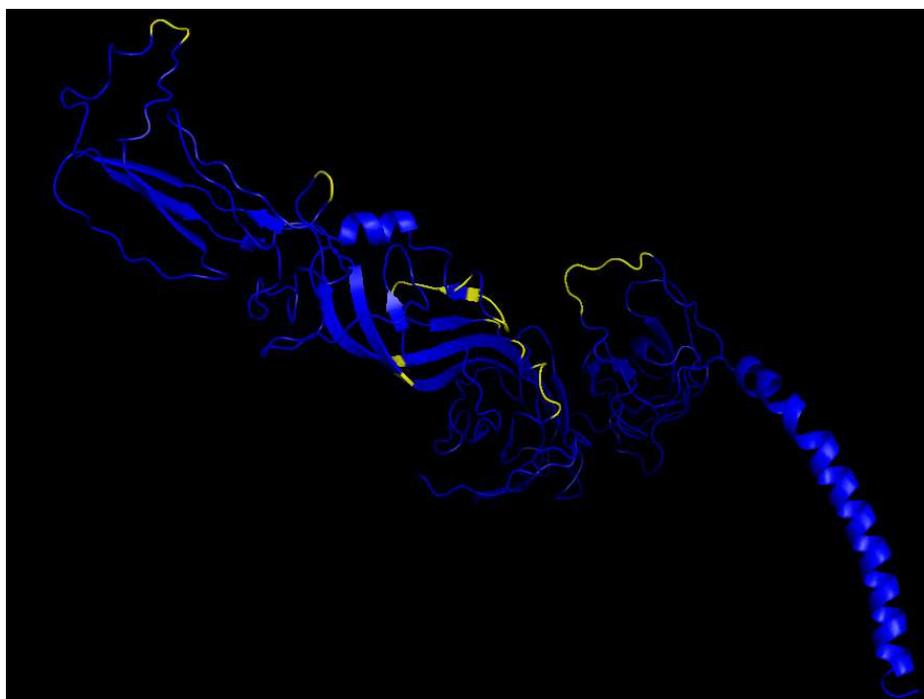


Figure 2 Discontinuous B-cell epitopes of the glycoprotein G2 shown in yellow.

**T – Cell epitope prediction**

T lymphocytes play a pivotal role in recognition and subsequent elimination of intracellular pathogen. T cell epitopes that bind to MHC Class I molecules were predicted using Propred I online server along with proteasome and immunoproteasome filter on at 5 thresholds. 47 MHC Class I alleles were considered when G2 protein sequence was analysed by Propred I. Only the top 4 epitopes for each of the alleles were taken into account and a total of 13 epitopes were predicted. Each of the epitopes were analysed for antigenic properties using VaxiJen v2.0 online server and out of 13 epitopes, 8 had probable antigenic properties (Table 4). The protein sequence was also used to predict MHC Class II binding regions using the Propred online server. All the predicted epitopes were then checked for their antigenicity (Table 5). Epitope LELDFSLTS at position 24 was found to have highest antigenicity ensuring maximum binding affinity.

**Table 4 MHC class I binding peptides on the basis of antigenicity**

Starting position	Peptide	MHC class I allele	Antigenic score
47	LEEAQSIDL	HLA-A1, HLA-B*2702, HLA-B*2705, HLA-B*3701, HLA-B40, HLA-B*4403, HLA-B60, HLA-B61, MHC-Kk	1.3540
128	GCTACGLYL	HLA-A1, HLA-A*1101, HLA-B*3901, HLA-B*3902, HLA-B7, HLA-B8, MHC-Ld	0.5195
23	TDLELDFSL	HLA-A2, HLA-B*3701, HLA-B40, HLA-B*4403, HLA-B*5401, HLA-B60, HLA-B61, HLA-Cw*0301, HLA-Cw*0602, HLA-Cw*0702, MHC-Db, MHC-Kk, MHC-Ld	2.0861
335	WGSVGVFTL	HLA-A2, HLA-A*0205, HLA-A2.1, HLA-B*3801, HLA-B*3901, HLA-B*3902, HLA-B*5101, HLA-B*5102, HLA-B*5103, HLA-B*5201, HLA-B*5301, HLA-B*5401, HLA-B*5401, HLA-B*51, HLA-B*5801, HLA-B60, HLA-B62, HLA-B7, HLA-B*0702, HLA-Cw*0702, MHC-Db revised, MHC-Dd, MHC-Kb, MHC-Ld	0.6654
464	LCVFLFSL	HLA-A*0205, HLA-A*1101, HLA-A24, HLA-A*3101, HLA-A68.1., HLA-B14, HLA-B*3901, HLA-B*3902, HLAB40, HLA-B*5102, HLA-B*5301, HLA-B*5401, HLA-B*51, HLA-B*5801, HLA-B8, HLA-Cw*0301, HLA-Cw*0401, HLA-Cw*0602, HLA-Cw*0702, MHC-Db, MHC-Db revised, MHC-Kb, MHC-Kd	0.8415
367	CYGAESVTL	HLA-A24, HLA-A*3302, HLA-B*3801, HLA-Cw*0401, MHC-Kb, MHC-Kd	0.8999
318	GENPCKIGL	HLA-B*2702, HLA-B*2705, HLA-B*3701, HLA-B40, HLA-B*4403, HLA-B60, HLA-B61, HLA-B*0702, HLA-Cw*0301, MHC-Kk, MHC-Ld	1.1630
411	AAPHLDKVN	HLA-B*5101, HLA-B*5102, HLA-B*5103, HLA-B*5801	0.4827

**Table 5 MHC class II binding peptides on the basis of antigenicity**

Starting position	Peptide	MHC class II allele	Antigenic score
457	WIVLIVLCV	DRB1_0101, DRB1_0421, DRB1_1101, DRB1_1120, DRB1_1128, DRB1_1302, DRB1_1305, DRB1_1307, DRB1_1321	0.497
465	VFLFSLVL	DRB1_0101, DRB1_0102, DRB1_0410, DRB1_1501, DRB1_1502, DRB1_1506	0.8378
150	IRYSRRVCV	DRB1_0301, DRB1_0309, DRB1_1107	0.8257
467	LLFSLVLLS	DRB1_0301, DRB1_0305, DRB1_0306, DRB1_0307, DRB1_0308, DRB1_0309, DRB1_0311, DRB1_0401, DRB1_0426, DRB1_1101, DRB1_1102, DRB1_1104, DRB1_1106, DRB1_1107, DRB1_1121, DRB1_1311, DRB1_1322	0.7849
24	LELDFSLTS	DRB1_0305, DRB1_0306, DRB1_0307, DRB1_0308, DRB1_0311, DRB1_0401, DRB1_0421, DRB1_0426, DRB1_1107	2.4905
148	ITIRYSRRV	DRB1_0402	1.1647
182	VCIIGTVSK	DRB1_0404, DRB1_0423	1.2303
323	IGLQTSSIE	DRB1_0410	0.772
340	FTLTCLVSL	DRB1_0701, DRB1_0703	0.795
175	FVSRHVKVC	DRB1_0801, DRB1_0802, DRB1_0804, DRB1_0806, DRB1_0813, DRB1_0817, DRB1_1114, DRB1_1323	0.9935
78	LNLKTSFHC	DRB1_0804	1.8984

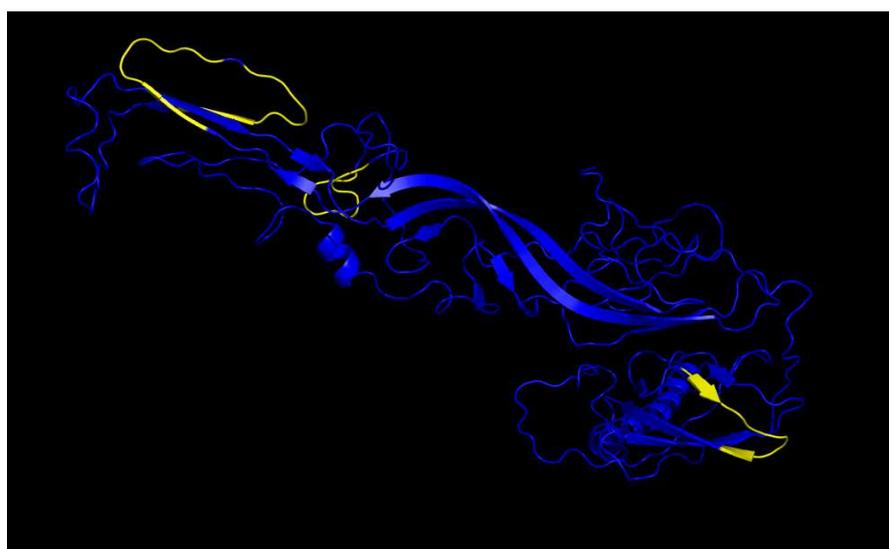
463	LCVFLLFSL	DRB1_1501, DRB1_1506	0.8415
466	FLLFSLVLL	DRB1_1502	0.8332
138	LKPVGSAYK	DRB5_0101, DRB5_0105	0.5253

**Conservation of epitopes**

Conservancy of effective antigenic 9 B-cell epitopes (Table 2) and 22 T-cell epitopes (Table 4 and 5) were checked against 91 HNTV glycoprotein G2 sequences retrieved from the NCBI protein database. B-cell epitope TSFH CYGACT at position 83 with antigenic score 0.977 was found to be conserved for all sequences with 100% identity. Other B-cell eptiopes showing significant conservancy with more than 94% matches at 100% identity were: CQFGD PGDIM at postion 219, YEYPWHTAKC at position 94, and SVTLTRGQNT at position 372. A total of 8 T-cell epitopes were found highly conserved showing more than 95% matches at 100% identity. They were WSGVGFTL at position 335, GCTACGLYL at position 128, VCIIGTVSK at position 182, TDLELDFSL at position 23, CYGAESVTL at position 367, AAPHLDKVN at position 411, LNLKTSFHC at position 78, and LELEDFSLTS at position 24 (Table 6). Conserved B-cell and T-cell epitopes are shown in Figure 3 and 4.

**Table 6** Conservancy analysis of B-cell and T-cell epitopes

B-cell/ T-cell	Epitope sequence	Epitope length	Percent of protein sequence matches at 100% identity	Minimum identity	Maximum identity
B-cell	TSFH CYGACT	10	100.00% (91/91)	100.00%	100.00%
B-cell	CQFGD PGDIM	10	98.90% (90/91)	70.00%	100.00%
B-cell	YEYPWHTAKC	10	98.90% (90/91)	90.00%	100.00%
T-cell	WSGVGFTL	9	98.90% (90/91)	88.89%	100.00%
T-cell	GCTACGLYL	9	97.80% (89/91)	88.89%	100.00%
T-cell	VCIIGTVSK	9	97.80% (89/91)	88.89%	100.00%
T-cell	TDLELDFSL	9	96.70% (88/91)	88.89%	100.00%
T-cell	CYGAESVTL	9	96.70% (88/91)	88.89%	100.00%
T-cell	AAPHLDKVN	9	96.70% (88/91)	88.89%	100.00%
T-cell	LNLKTSFHC	9	96.70% (88/91)	88.89%	100.00%
T-cell	LELEDFSLTS	9	95.60% (87/91)	88.89%	100.00%
B-cell	SVTLTRGQNT	10	94.51% (86/91)	90.00%	100.00%



**Figure 3** Conserved B-cell epitopes of the glycoprotein G2 shown in yellow.

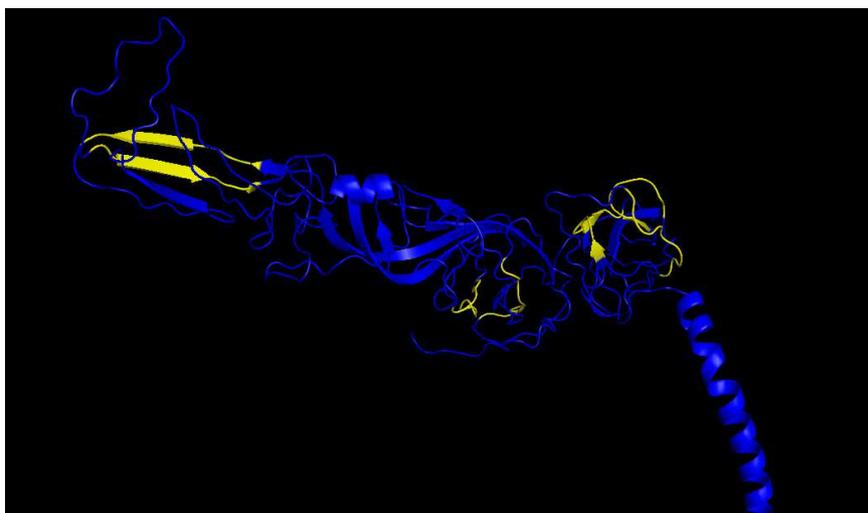


Figure 4 Conserved T-cell epitopes of the glycoprotein G2 shown in yellow.

## DISCUSSION

The current study was focused on prediction of potential vaccine against HNTV. This approach of vaccine development is based only on in-silico analysis. The protein sequences of antigens present in the virus are analyzed to find out regions in the protein that can be used as a vaccine. This kind of vaccine development has already been successful in case of Hepatitis B virus, Hepatitis C virus and HIV where synthetic peptide vaccines were developed using gene sequence of the viruses. Those vaccines showed efficacy in clinical trials. With the increasing availability of genomic data this approach is being used more often to develop vaccines against virus now-a-days. Based on the principle of these studies we analyzed HNTV glycoprotein G2 to propose potential peptide sequence to be used as peptide vaccine.

Currently no high-resolution structure is available for HNTV glycoprotein G2. Also, there is no evidence to develop peptide vaccine against the virus. *In silico* prediction of structure of the protein facilitates vaccine development and a deeper understanding about the mechanism of viral entry and immune response. For structure determination homology modeling approach was used. In the predicted structure there is only a single helical region which lies in the transmembrane region of the protein. The rest of the protein is located in the exomembrane region and comprises of strands, turns and coils. The antigenicity of the protein was evaluated and it was found to be a potential antigen that elicits immune response. The antigenic determinants or epitopes on the protein were determined. These epitopes are regions of protein structure interact directly with the immune system and are recognized by antibodies, T-cell receptors and B-cell receptors, consequently, the virus particle is either neutralized or subjected to various effector mechanisms. Conservancy analysis of the epitopes showed that B-cell epitope TSFHICYGACT (Figure 5) was conserved in all possible genotypes at 100% identity. Eleven other B-cell and T-cell epitopes were also found to be highly conserved showing variation in only one amino acid position for some genotypes. Among four B-cell epitopes SVTLTRGQNT at position 138 have the highest antigenic score of 1.1341. Among eight conserved T-cell epitopes LELDFSLTS at position 24 has the highest antigenic score of 2.4905. A higher value of antigenic score increases binding affinity to the immune system. It is notable that most of the epitopes are located in two regions of the protein: a domain close to the transmembrane helical portion and a region of the protein that lies away from the transmembrane region. Peptides containing the sequence of these epitopes along with the predicted discontinuous epitopes can be used for vaccine development. The information about the epitopes can also be used for inhibitor development against the virus because they play indispensable role in viral attachment to host cell. Furthermore, the structure of the glycoprotein G2 along with microscopy data can provide further insight about attachment and entry.

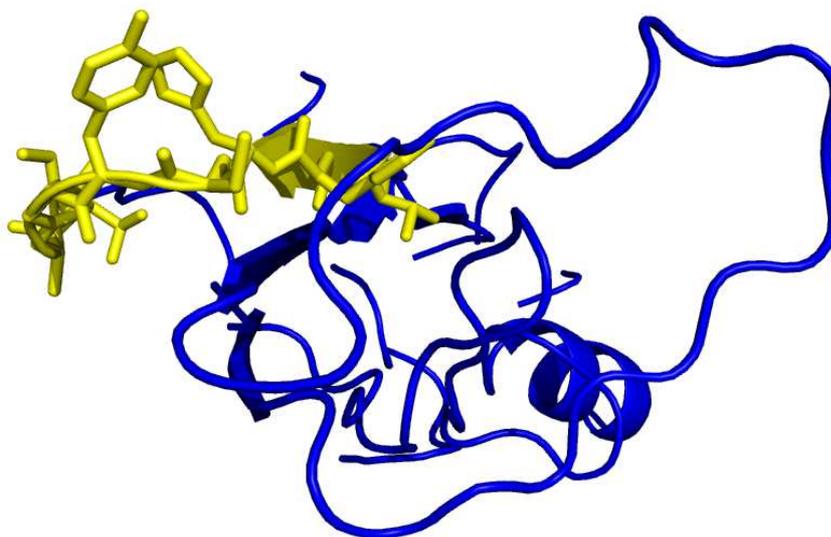


Figure 5 A B-cell epitope in the HNTV glycoprotein G2 model. The epitope (TSFH CYGACT) structure is colored in yellow.

#### CONCLUSION

HNTV glycoprotein G2 can be used to develop vaccine against the virus. Structural analysis of the protein may facilitate development of effective inhibitor. Antigenic determinants present on the protein can be important target for vaccine design and development. The conserved regions of the protein can also give useful information for development of diagnostic method of the virus in human body.

#### REFERENCES

- [1] CB Jonsson; LT Figueiredo; O Vapalahti. *Clin Microbiol Rev*, **2010**, 23, 2, 412-441.
- [2] CJ Peters; AS Khan. *Clin Infect Dis*, **2002**, 34, 9, 1224-1231.
- [3] C Schmaljohn; B Hjelle. *Emerg Infect Dis*, **1997**, 3, 2, 95-104.
- [4] KM Johnson. Manual of hemorrhagic fever with renal syndrome and hantavirus pulmonary syndrome. Seoul (South Korea): WHO Collaborating Center for Virus Reference and Research (Hantaviruses), Asian Institute for Life Sciences, **1999**; 1-6.
- [5] YZ Zhang; Y Zou; ZF Fu; A Plyusnin. *Emerg Infect Dis*, **2010**, 16, 8, 1195-1209.
- [6] CS Schmaljohn. *Expert Rev Vaccines*, **2012**, 11, 5, 511-513.
- [7] H Kariwa; K Lokugamage; N Lokugamage; H Miyamoto; K Yoshii; M Nakauchi; K Yoshimatsu; J Arikawa; LI Ivanov; T Iwasaki; I Takashima. *Jpn J Vet Res*, **2007**, 54, 4, 145-161.
- [8] HW Cho; CR Howard; HW Lee. *Intervirology*, **2002**, 45, 4-6, 328-333.
- [9] JW Hooper; JE Moon; KM Paolino; R Newcomer; DE McLain; M Josleyn; D Hannaman; C Schmaljohn. *Clin Microbiol Infect*, **2014**.
- [10] W Severson; L Partin; CS Schmaljohn; CB Jonsson. *J Biol Chem*, **1999**, 274, 47, 33732-33739.
- [11] CB Jonsson; CS Schmaljohn. *Curr Top Microbiol Immunol*, **2001**, 256, 15-32.
- [12] C Lober; B Anheier; S Lindow; HD Klenk; H Feldmann. *Virology*, **2001**, 289, 2, 224-229.
- [13] CF Spiropoulou. *Curr Top Microbiol Immunol*, **2001**, 256, 33-46.
- [14] AJ Battisti; YK Chu; PR Chipman; B Kaufmann; CB Jonsson; MG Rossmann. *J Virol*, **2011**, 85, 2, 835-841.
- [15] HM Berman; J Westbrook; Z Feng; G Gilliland; TN Bhat; H Weissig; IN Shindyalov; PE Bourne. *Nucleic Acids Res*, **2000**, 28, 1, 235-242.
- [16] Y Zhang. *BMC Bioinformatics*, **2008**, 9, 40.

- 
- [17] E Gasteiger; C Hoogland; A Gattiker; S Duvaud; MR Wilkins; RD Appel; A Bairoch. The Proteomics Protocols Handbook. Humana Press, **2005**; 571-607
- [18] C Geourjon; G Deleage. *Comput Appl Biosci*, **1995**, 11, 6, 681-684.
- [19] F Ferre; P Clote. *Nucleic Acids Res*, **2006**, 34, *Web Server issue*, W182-185.
- [20] M Wiederstein; MJ Sippl. *Nucleic Acids Res*, **2007**, 35, *Web Server issue*, W407-410.
- [21] RA Laskowski; MW MacArthur; DS Moss; JM Thornton. *J App Cryst*, **1993**, 11, 2, 283-291.
- [22] IA Doytchinova; DR Flower. *BMC Bioinformatics*, **2007**, 8, 4.
- [23] A Krogh; B Larsson; G von Heijne; EL Sonnhammer. *J Mol Biol*, **2001**, 305, 3, 567-580.
- [24] H Singh; GP Raghava. *Bioinformatics*, **2001**, 17, 12, 1236-1237.
- [25] H Singh; GP Raghava. *Bioinformatics*, **2003**, 19, 8, 1009-1014.
- [26] P Somvanshi; V Singh; PK Seth. *OMICS*, **2008**, 12, 1, 61-69.
- [27] Y El-Manzalawy; D Dobbs; V Honavar. *J Mol Recognit*, **2008**, 21, 4, 243-255.
- [28] JV Kringelum; C Lundegaard; O Lund; M Nielsen. *PLoS Comput Biol*, **2012**, 8, 12, e1002829.
- [29] K Tamura; D Peterson; N Peterson; G Stecher; M Nei; S Kumar. *Mol Biol Evol*, **2011**, 28, 10, 2731-2739.
- [30] R Vita; L Zarebski; JA Greenbaum; H Emami; I Hoof; N Salimi; R Damle; A Sette; B Peters. *Nucleic Acids Res*, **2010**, 38, Database issue, D854-862.