# *IN SILICO* ANALYSIS OF THE ANTIGENIC PROPERTIES OF NOROVIRUS GII.4 SYDNEY [P16] VP1 PROTEIN

*A.-M.D. Zharova*[1*], *V.Yu. Talayev*[2], *A.D. Perenkov*[1], *I.Ye. Zaichenko*[2], *M.V. Svetlova*[2], *O.N. Babaykina*[2], *E.V. Voronina*[2], *V.A. Lapin*[1,2], *V.V. Novikov*[1,2]

[1] Institute of Biology and Biomedicine, National Research Lobachevsky State University of Nizhny Novgorod, 23 Gagarin Ave., Nizhny Novgorod, 603022, Russia;
[2] Academician I.N. Blokhina Nizhny Novgorod Scientific Research Institute of Epidemiology and Microbiology of Federal Service for Surveillance on Consumer Rights Protection and Human Wellbeing, 71 Malaya Yamskaya St., Nizhny Novgorod, 603950, Russia.

* Corresponding author: uglich_marie@mail.ru

**Abstract.** Norovirus infection is a leading cause of non-bacterial acute gastroenteritis. This study aimed to analyze the antigenic properties of the VP1 protein of norovirus GII.4 Sydney [P16] circulating in Russia. The analysis was conducted using *in silico* methods. VP1 amino acid sequence data was used to identify T-helper and T-killer epitopes, linear and conformational B-cell epitopes to assess the conservation of epitopes, and allergenicity of VP1. T cell epitopes with the highest estimated immunogenicity were identified at positions 207-223 and 378-394 in the S- and P-domains of the protein. The tertiary structure of VP1 was modeled, and 2 linear and 47 conformational B-cell epitopes were identified. In addition to the previously described epitopes, a new putative B-cell epitope was identified at position 307-316 of the P2 subdomain. *In silico* analysis of the primary and tertiary structure of the norovirus VP1 protein showed that it is not allergenic and has various immunogenic epitopes, potentially capable of inducing T- and B-cell immune responses.

**Keywords:** bioinformatics, antigen, epitope, norovirus, vaccine, virus-like particles.

## List of Abbreviations

HuNoV – human norovirus
$IC_{50}$ – half-maximal inhibitory concentration
mAbs – monoclonal antibodies
P domain – protruding domain
S domain – shell domain
VLP – virus-like particles
VP1 – norovirus capsid viral protein 1

## Introduction

Human Noroviruses (HuNoVs) are the leading cause of non-bacterial acute gastroenteritis worldwide and contribute to increased morbidity and mortality especially in children under 5 years of age, the elderly and immunocompromised individuals (Hall *et al.*, 2013). Despite this, there is currently no licensed vaccine against HuNoV. Three types of vaccines are being clinically tested, including virus-like particles (VLPs) constructed from VP1, a P-particle, and a vaccine based on recombinant adenoviruses and vesicular stomatitis virus containing the VP1 gene sequence. (Esposito & Principi, 2020; Ford-Siltz *et al.*, 2021).

HuNoV belongs to the genus Norovirus of the Caliciviridae family, which combines non-enveloped viruses with a linear positive polarity single-stranded RNA genome (+ssRNA). HuNoVs are classified into 10 genogroups (GI-GX) based on differences in the VP1 capsid protein, as well as several P-groups and P-types based on differences in RNA-dependent RNA polymerase (Chhabra *et al.*, 2019). HuNoV genogroups are divided into genotypes and variants resulting from mutations in the P2 subdomain of the VP1 protein and homologous recombination in the ORF1/ORF2 reading frame overlap region (Esposito & Principi, 2020). The most common and rapidly mutating genotype worldwide is GII.4, which, along with the GI and GIV genogroups, is responsible for 50-80% of cases of acute gastroenteritis in humans every year (Hassan & Baldridge, 2019; Epifanova *et al.*, 2022; Oparina *et al.*, 2023).

The capsid of the virus exhibits icosahedral symmetry (T = 3), has a diameter of 38 nm, and is composed of 180 monomers of the main structural protein VP1 (Hassan & Baldridge, 2019). The minor structural protein VP2 is present inside the capsid in quantities ranging from 1 to 8 copies. VP1 is the primary target of antibodies formed during infection due to its exter-

nal location (Campillay-Veliz *et al.*, 2020; Esposito & Principi, 2020; Ford-Siltz *et al.*, 2021). VP1 can spontaneously oligomerize and form virus-like particles (VLPs) with the morphology of native virions (Hassan & Baldridge, 2019). VP1 of the GII.4 genotype can also assemble into larger icosahedral particles with T = 4 triangulation number. VP1 has a molecular weight of 58 kDa and consists of a shell (S) and an overhanging (P) domain linked by a flexible linker region (Prasad *et al.*, 1999; Hardy, 2005). The S-domain contains the elements necessary for the assembly of the capsid. P-domains form protrusions on the outer surface of the virion. Dimerization of the P-domains appears to increase the stability of the capsid. The P-domain is divided into two subdomains, P1 and P2, with the P2 subdomain occupying the most distal position in the monomer and containing a hypervariable region. It is believed that P2, and especially the hypervariable region, plays an important role in immune reactivity, as well as in interaction with virus binding co-factors (Tan *et al.*, 2003; Tan *et al.*, 2005; Walker & Baldridge, 2019).

The aim of this work was an *in silico* analysis of the antigenic properties of the VP1 protein of norovirus GII.4 Sydney [P16] circulating in Russia.

## Materials and Methods

In this study, the amino acid sequence of VP1 (GenBank No. MZ958411) was utilized to identify epitopes and model the tertiary structure of the norovirus protein using homology modeling.

### *Identification of T-cell epitopes in silico*

To determine the epitopes of cytotoxic T-lymphocytes, a reference set of 27 frequently occurring alleles of HLA I molecules was used according to the international database of immunoepitopes IEDB (http://tools.iedb.org/mhci/). The search for alleles found in more than 1% of the population of the city Nizhny Novgorod was carried out in the international database AFND (http://www.allelefrequencies.net/hla6006a.asp/).

The NetCTLpan 1.1 online program was used to identify $CD8^+$ T-lymphocyte epitopes (https://services.healthtech.dtu.dk/service.php?NetCTLpan-1.1/). The analysis was performed at the default parameters. Over the course of the analysis, the binding affinity index for MHC I was calculated for the studied peptides using the following formula: $1\text{-}\log_2(A)/\log(50000)$, where A is the half-maximal inhibitory concentration $IC_{50}$ (nM) (Nielsen *et al.*, 2003). The efficiency of peptide transport by TAP proteins was expressed in terms of $IC_{50}$ values. Predictive assessment of proteasomal cleavage of ligands of histocompatibility molecules was expressed as a score. The selection of epitope peptides from the total number of identified peptides was performed using R Studio 4.2.2.

A reference set of 27 frequently occurring alleles of MHC II molecules according to IEDB (http://tools.iedb.org/mhcii/) and alleles found in over 1% of people in the city Nizhny Novgorod were used to identify helper T-lymphocyte epitopes. Epitopes of $CD4^+$ T-lymphocytes were predicted using the NetMHCIIpan-4.0 (https://services.healthtech.dtu.dk/service.php?NetMHCIIpan-4.0/). Sequence analysis was also performed in the default web server mode. The analysis included predictions of eluted ligands (EL) and binding affinity (BA) of peptides for class II histocompatibility molecules expressed as percentile rank for comparison with a reference set of 100,000 peptides.

Analysis of the immunogenicity of the detected epitopes was performed using the VaxiJen 2.0 program at a threshold of 0.4 for viral peptides and proteins (http://www.ddg-pharmfac.net/vaxijen/). Mapping of the detected epitopes was performed using the Unipro Ugene software (version 33). The allergenicity of the studied protein was analyzed using the AllerTOP 2.0 online server (http://www.ddg-pharmfac.net/AllerTOP/).

### *Identification of B-cell epitopes in silico*

Linear B-cell epitopes were determined using the BepiPred-3.0 (https://services.healthtech.dtu.dk/services/BepiPred-3.0/) and ABCPred servers (http://crdd.osdd.net/raghava/abcpred/), which were analyzed at the recommended thresholds. The immunogenicity of the identified peptides was assessed using VaxiJen 2.0.

Using the IEDB tools, the Emini solvent accessibility of the protein surface and Karplus & Schultz flexibility (http://tools.iedb.org/bcell/) were further analyzed. The analysis was carried out at automatic threshold values equal to the average value of the estimated score. The assessment of the conservatism of the identified epitopes was performed using the IEDB tool (http://tools.iedb.org/conservancy/) with a sequence identity threshold for calculation of at least 100%. Conservation of the predicted epitopes was assessed with respect to 5 complete amino acid sequences of the VP1 protein of some norovirus genotypes occurring in the Nizhny Novgorod region (GII.3 [P12], GII.2 [P16] and GII.6 [P7]), 3 complete amino acid sequences of the VP1 protein genotype GII.4 [P16], identified in 2022 in the Nizhny Novgorod region according to GenBank data, and 1 VP1 protein sequence of norovirus GII.4 Sydney 2012 [P16]. The search and selection of sequences was carried out according to the GenBank database and literature sources (On the state of sanitary and epidemiological well-being…, 2022; Epifanova *et al.*, 2022).

The search for conformational epitopes using a homologous model of protein tertiary structure was performed in BepiPred-3.0 at a threshold of 0.1512 and DiscoTope 2.0 (https://services.healthtech.dtu.dk/services/DiscoTope-2.0/). The measurement in DiscoTope 2.0 was carried out at the default threshold.

*Homological modeling of protein tertiary structure*

Homological modeling of protein tertiary structure to determine conformational B-cell epitopes was performed using the SWISS-MODEL server (https://swissmodel.expasy.org/). The crystal structure of the human norovirus GII.4 Houston virus-like particle (PDB ID: 7k6v.1.B) was used for modeling. The protein model was visualized using the PyMol 2.5.4 program. The stereochemical quality and accuracy of model building were analyzed using the Ramachandran plot, GMQE, MolProbity score and clash score. Additionally, the model was protonated using the Reduce program, the inversion of Asn, Gln, His side chains and invalid overlaps of atoms were eliminated. Corrections to the constructed model were made using the MolProbity 4.5.2 resource (http://molprobity.biochem.duke.edu/).

The study was carried out as part of the work approved by the Local Ethical Committee of Academician I.N. Blokhina Nizhny Novgorod Scientific Research Institute of Epidemiology and Microbiology, protocol No. 3 of 24.03.20. Experiments with humans and animals were not carried out, personal data was not used.

## Results

*Analysis of T-cell epitopes*

To determine the epitopes of cytotoxic T-lymphocytes, 27 frequently occurring alleles of HLA I molecules were selected according to IEDB data. The search for alleles necessary for analysis was also carried out in AFND (Table 1).

The analysis revealed 14 peptides with a length of 9 amino acids, capable of interacting with HLA I, among which 6 peptides of the VP1 protein had the maximum immunogenicity values (Table 2). The identified peptides are optimal in terms of $IC_{50}$ ($50 < IC_{50} < 500$) for binding to MHC I and TAP, have a high probability of proteasomal cleavage, which was the criteria for evaluating peptides as potential epitopes.

A search was made for epitopes of helper T-lymphocytes capable of interacting with HLA II. For this, 27 frequently occurring alleles of MHC II molecules were selected according to IEDB data. An allele search was also carried out using the AFND database (Table 1).

The search revealed 20 peptides with a length of 15 amino acids capable of cross-linking with at least 6 analyzed alleles of histocompatibility molecules. 6 peptides localized in the S- and P-domains of the VP1 protein had the highest indices of expected immunogenicity (Table 3).

The selected immunogenic sequences had the best MHC II binding affinity ($IC_{50} <$ < 1000 nM) and a percentile rank of eluted ligands ≥ 1%, which made it possible to identify the sequences as epitopes.

*Table 1*

**Alleles HLA I and HLA II used for epitope analysis**

| Database | HLA I | HLA II |
|---|---|---|
| IEDB | A*01:01, A*02:01, A*02:03, A*02:06, A*03:01, A*11:01, A*23:01, A*24:02, A*26:01, A*30:01, A*30:02, A*31:01, A*32:01, A*33:01, A*68:01, A*68:02, B*07:02, B*08:01, B*15:01, B*35:01, B*40:01, B*44:02, B*44:03, B*51:01, B*53:01, B*57:01, B*58:01 | DRB1*01:01, DRB1*03:01, DRB1*04:01, DRB1*04:05, DRB1*07:01, DRB1*08:02, DRB1*09:01, DRB1*11:01, DRB1*12:01, DRB1*13:02, DRB1*15:01, DRB3*01:01, DRB3*02:02, DRB4*01:01, DRB5*01:01, DQA1*05:01/DQB1*02:01, DQA1*05:01/DQB1*03:01, DQA1*03:01/DQB1*03:02, DQA1*04:01/DQB1*04:02, DQA1*01:01/DQB1*05:01, DQA1*01:02/DQB1*06:02, DPA1*02:01/DPB1*01:01, DPA1*01:03/DPB1*02:01, DPA1*01:03/DPB1*04:01, DPA1*03:01/DPB1*04:02, DPA1*02:01/DPB1*05:01, DPA1*02:01/DPB1*14:01 |
| AFND | A*25:01, B*13:02, B*14:02, B*18:01, HLA-B*27:02, B*27:05, B*35:02, B*35:03, B*37:01, B*38:01, B*39:01, B*40:02, B*41:02, B*44:27, B*49:01, B*50:01, B*52:01 | DRB1*01:02, DRB1*04:02, DRB1*04:04, DRB1*08:01, DRB1*10:01, DRB1*11:04, DRB1*13:01, DRB1*13:03, DRB1*14:54, DRB1*15:02, DRB1*16:01 |

*Table 2*

**Norovirus VP1 epitopes interacting with HLA I**

| Peptide | Allele | Immunogenicity |
|---|---|---|
| **TVSCRVLTR*** (197..205) | HLA-A*31:01, HLA-A*33:01, HLA-A*68:01 | 1.18 |
| **RPSPDFDFI** (205..213) | HLA-B*07:02, HLA-B*51:01, HLA-B*53:01 | 1.21 |
| **SPDFDFIFL** (207..215) | HLA-B*35:02, HLA-B*35:03, HLA-B*51:01, HLA-B*53:01 | 1.96 |
| RTKPFSVPV (223..231) | HLA-A*02:06, HLA-A*30:01, HLA-A*31:01, HLA-A*32:01 | 1.04 |
| NQNTKFTPV (378..386) | HLA-A*02:03, HLA-A*02:06, HLA-B*08:01 | 1.64 |
| YPNMDLDCL (444..453) | HLA-B*35:02, HLA-B*35:03, HLA-B*53:01 | 1.43 |

* *Note:* S-domain peptides are shown in bold.

*Table 3*

**Norovirus VP1 epitopes interacting with HLA II**

| Peptide | Allele | Immunogenicity |
|---|---|---|
| **MFPHIIVDVRQLEPV*** (140..154) | DRB1*03:01, DRB1*11:04, DRB1*13:01, DRB1*13:03, DRB1*13:02, DRB1*14:54, DRB3*01:01 | 0.98 |
| **IVDVRQLEPVLIPLP** (145..159) | DRB1*01:01, DRB1*01:02, DRB1*07:01, DRB1*09:01, DRB1*10:01, DRB1*12:01, DQA1*05:01/DQB1*02:01, DQA1*01:01/DQB1*05:01, DPA1*02:01/DPB1*14:01 | 1.19 |
| **DFDFIFLVPPTVESR** (209..223) | DRB1*01:01, DRB1*01:02, DRB1*04:05, DRB1*07:01, DRB1*08:02, DRB1*09:01, DRB1*10:01, DRB1*11:01, DRB1*12:01, DRB1*15:02, DRB1*16:01, DRB3*02:02, DRB5*01:01, DQA1*01:01/DQB1*05:01, DPA1*03:01/DPB1*04:02 | 1.82 |
| DHDFEANQNTKFTPV (372..386) | DRB1*04:01, DRB1*07:01, DRB1*13:02, DRB1*13:03, DRB1*14:54, DRB1*15:02, DRB3*01:01, DRB3*02:02 | 1.11 |
| NTKFTPVGVIQDGST (380..394) | DRB1*04:01, DRB1*04:05, DRB1*07:01, DRB1*08:01, DRB1*08:02, DRB1*09:01, DRB1*10:01, DRB1*11:01, DRB1*13:03, DRB3*02:02, DQA1*03:01/DQB1*03:02, DQA1*04:01/DQB1*04: 02 | 1.13 |
| KSGYVTVAHTGQHDL (493..507) | DRB1*04:01, DRB1*08:01, DRB1*08:02, DRB1*10:01, DRB1*11:01, DRB1*11:04, DRB1*13:03, DRB1*16:01, DRB5*01:01, DQA1*05:01/DQB1*02:01, DQA1*04:01/DQB1*04: 02 | 0.78 |

* *Note:* S-domain peptides are shown in bold.

After epitope identification and calculation of immunogenicity values, the location of the MHC I and MHC II ligands was mapped to the primary sequence of the VP1 protein (Fig. 1).

Potential immunogenic ligand peptides for HLA I and HLA II were found to be located in both the P- and S-domains. The ligand of HLA I and HLA II with the highest expected immunogenicity was the sequence 207-223, corresponding to the end of the S-domain and the VP1 hinge region, and the region 378-394 in the P2 subdomain. In addition, we found that the VP1 protein, as well as its S- and P-domains, are not potential allergens. This proves the possibility of using the protein or individual VP1 domains in VLP vaccines.

*Identification and assessment of the conservatism of linear B-cell epitopes*

BepiPred-3.0 and ABCPred predicted 13 linear B-cell epitopes. After that, the theoretical solvent accessibility of the VP1 protein surface was further analyzed using the IEDB database tools (Fig. 2).

With a threshold equal to 1, only 4 sites can be theoretically accessible and simultaneously predicted by the BepiPred 3.0 and ABCPred programs, presented in Table 4 and highlighted in red in Figure 2.

All peptides identified by us belonged to the P-domain of the VP1 protein. The conservatism of

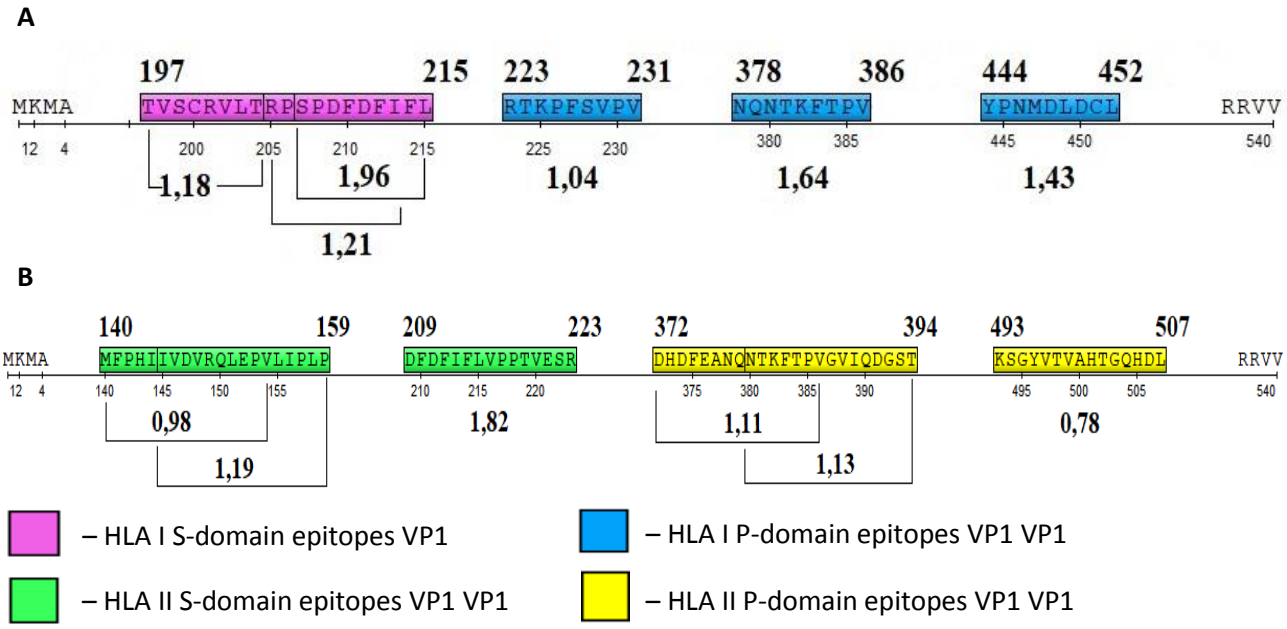


**Fig. 1.** Distribution of HLA I (A) and HLA II (B) cross-linking epitopes of the VP1 protein. The values below the sequences correspond to the VaxiJen assessment of the immunogenicity of the peptides

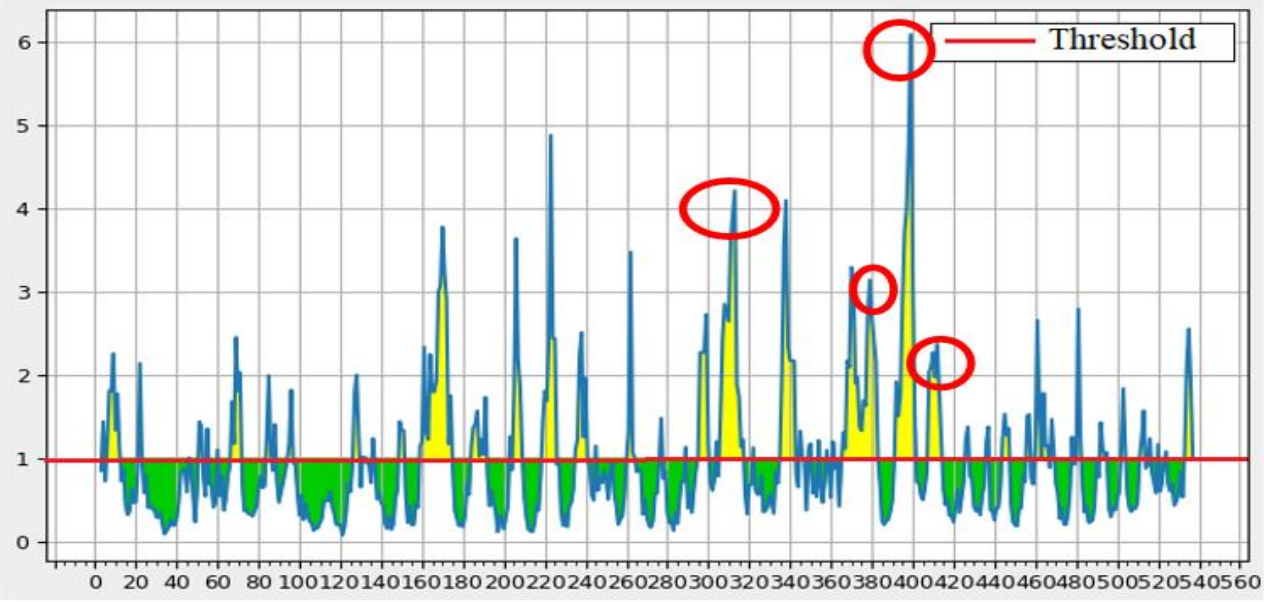


**Fig. 2.** Predicted amino acid surface arrangement of the norovirus VP1 protein for the Emini solvent. The yellow and green areas correspond to positive and negative forecasts, respectively

linear epitopes was then assessed. For this, the analyzed sequence was compared with VP1 of various noroviruses. The identified epitopes are not 100% conserved among the GII.3 [P12], GII.2 [P16] and GII.6 [P7] norovirus genotypes. The DGSTTHRNEPQ sequence is 100% conserved with the 2022 VP1 sequences, the YSGRNTHNV peptide is 66.67% conserved, and TMNLASQNWNNYDPTEEIPAPLGTPDFV, DHDFEANQNTK are 33.33% identical. The single peptide YSGRNTHNV showed 100% identity with GII.4 Sydney 2012 [P16].

Peptides with 100% conservatism were then analyzed for flexibility. The DGSTTHRNEPQ and YSGRNTHNV sequences (at mean thresholds of 1.061 and 1.064) have the flexible properties of linear B-cell epitopes (Fig. 3).

*Table 4*

**Linear epitopes of B-lymphocytes of norovirus VP1 capsid protein**

| Peptide | Length | Immunogenicity |
|---|---|---|
| TMNLASQNWNNYDPTEEIPAPLGTPDFV (300..327) | 28 | 0.7026 |
| DHDFEANQNTK (372..382) | 11 | 0.5604 |
| DGSTTHRNEPQ (391..401) | 9 | 0.6991 |
| YSGRNTHNV (408..416) | 9 | 1.3185 |

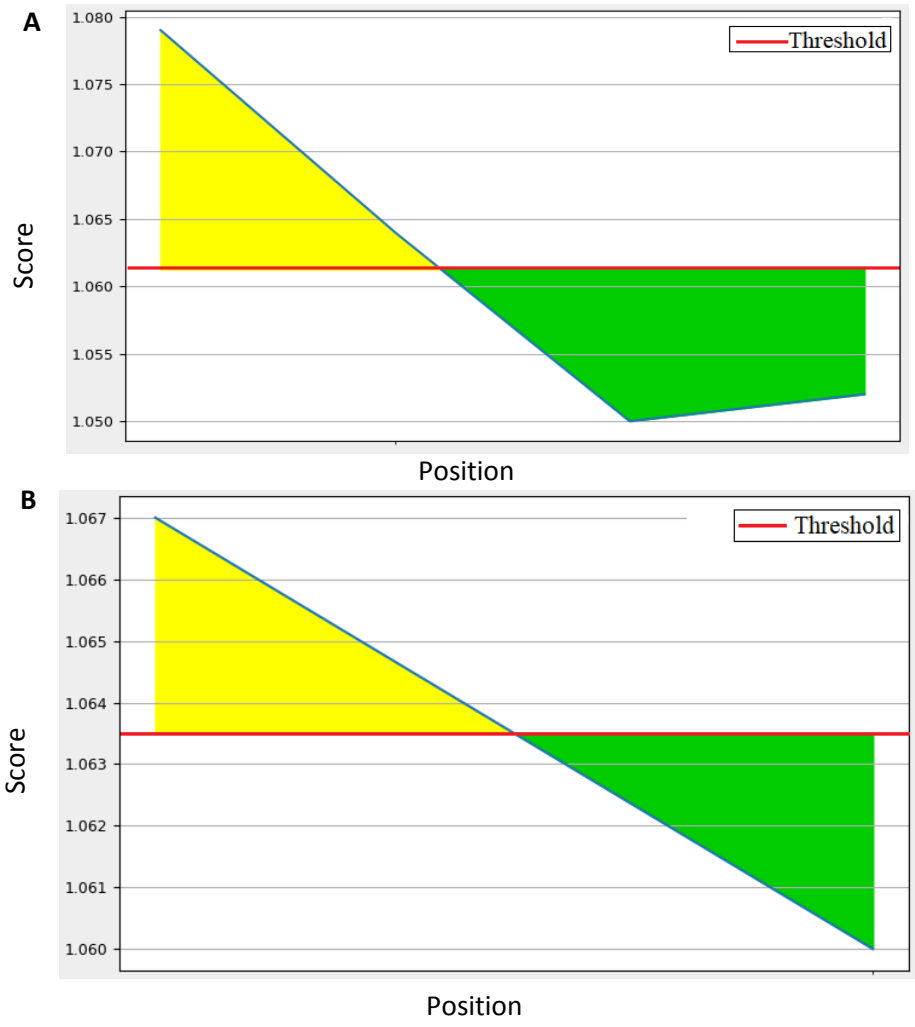


**Fig. 3.** Flexibility of epitopes «DGSTTHRNEPQ» (A) and «YSGRNTHNV» (B) of the norovirus VP1 protein according to Karplus and Schultz. The yellow and green areas correspond to positive and negative forecasts, respectively

*Homological modeling of protein tertiary structure and analysis of conformational B-cell epitopes*

3D modeling of norovirus VP1 was performed using homology modeling (Fig. 4). The crystal structure of the virus-like particle of human norovirus Houston GII.4 (PDB ID: 7k6v.1.B) with a resolution of 3 Å was chosen as a suitable template for building the model. The consensus sequence of this template had a high identity with the test sequence (92.78%) and coverage from 10 to 532 amino acids. The positions at which amino acid substitutions were found are shown in Figure 4 in red (except for positions 10, 534, 539 and 540, which were not included in the model due to the lack of corresponding sections in the template). The model makes it possible to clearly differentiate the S- and P-domains of the VP1 protein connected by a linker sequence. The model is rich in β-structures characteristic of HuNoV VP1. In the S-domain, β-folding of 8 antiparallel β-chains was noted (one of the chains is not shown in Figure 4a). The P1 subdomain consists of 9 β-strands and one short α-helix. The structure of the P1 subdomain contains the P2 subdomain of 7 β-strands.

To assess the quality of the obtained model and to identify prohibited positions of amino acids, a Ramachandran map was constructed, on which the most favorable regions correspond to dark green, neutral to light green, and unfavorable to white areas (Fig. 5).

Figure 5 shows that most of the amino acid residues occupied the upper left quadrant of the map, which is typical for β-structures and is consistent with the obtained spatial structure. The number of amino acid residues in favorable areas of the map according to the values of torsion angles φ and ψ is 96.74%, and in unfavorable areas - 0.19% (Ser494). Additionally, the quality of the model was evaluated by GMQE equal to 0.89, which corresponds to high quality. The Molprobity score is 0.88, which is lower than the resolution of the original template structure of 3 Å, which confirms the quality of the constructed model. The number of collisions of atoms is 0.5 per 1000 atoms, which is the 99th percentile. Consequently, the constructed model has almost the best quality among models with a resolution of 3 ± 0.25Å. The analysis performed proved the absence of conformational stresses and confirmed the quality of the model and the possibility of further use of the model.

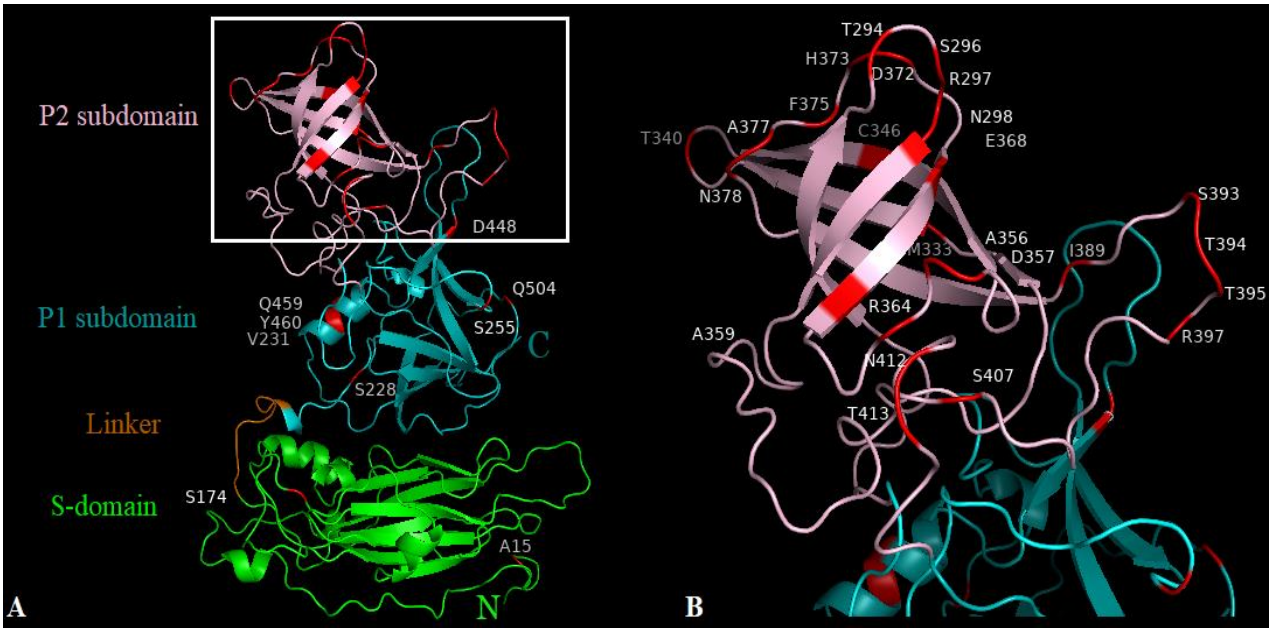


**Fig. 4.** Model of the tertiary structure of the VP1 protein (A) and scaled image of the P2 VP1 subdomain (B) of human norovirus Sydney GII.4 [P16]. The hypervariable subdomain P2 of the VP1 protein (B) is indicated by a white rectangle

The construction of the VP1 protein monomer made it possible to identify B-cell conformational epitopes, the position of which was marked on the molecular surface of the P2 subdomain of the VP1 protein, stained in pink (Fig. 6).

Detected epitopes are marked in yellow in Figures 6. These epitopes are located in the P2 subdomain of the VP1 protein, except for positions 442, 444, and 445 of the P1 subdomain. Positions within conformational epitopes 287, 297, 342-345, 370, 371, 379, 390, and 397 are out of view in the background of the protein structure in fig. 6. Also, *in silico* analysis re vealed for the first time a probable B-cell epitope motif having residue position 307–316 (red area).

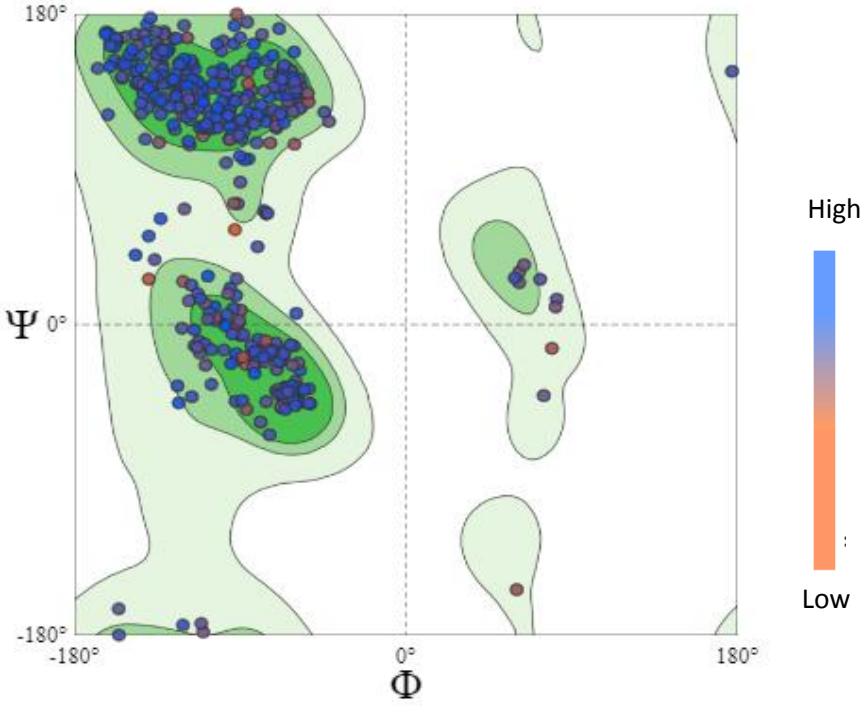


**Fig. 5.** Ramachandran map of the VP1 protein of human norovirus Sydney GII.4 [P16]. The dots show amino acid residues, the color reflects the level of prediction confidence

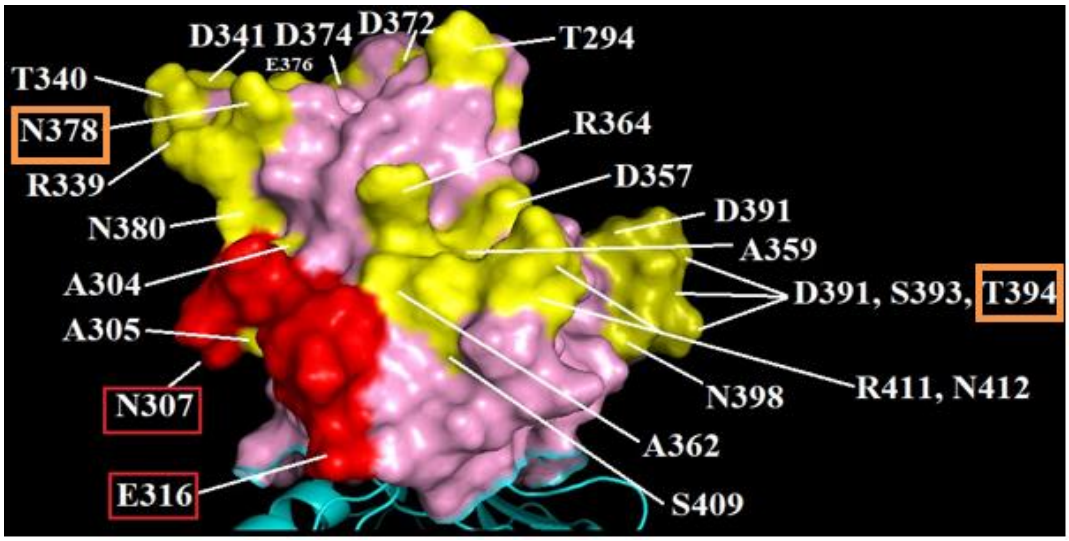


**Fig. 6.** Position of conformational epitopes on the molecular surface of the P2 subdomain of the norovirus VP1 protein

**Discussion**

The presence of immunogenic epitopes in the structure of a foreign antigen is essential for the effective initiation of T- and B-cell immune responses. Therefore, analyzing immunological properties is crucial for selecting antigens in developing new vaccines. In this study, we investigated the structural protein of norovirus VP1, which is circulating in the central part of Russia (Epifanova *et al.*, 2022; Oparina *et al.*, 2023). The development of a vaccine is promising because norovirus is the leading cause of gastroenteritis worldwide, and no effective vaccines have been developed against this infection. Using *in silico* methods, we identified potential immunogenic epitopes of the VP1 protein for helper and cytotoxic T-lymphocytes. For identifying T-cell epitopes, we chose HLA alleles that are frequently present in the world population and in the population of Nizhny Novgorod city. Additionally, we determined that the VP1 protein is non-allergenic and can be used as a basis for creating virus-like particle (VLP) vaccines.

It is known that pathogens have evolved a variety of immune escape mechanisms, including the variability of surface capsid proteins. The norovirus VP1 protein is no exception. Its evolutionary rate is $5.94*10^{-3}$ substitutions per position per year, and only the S-domain of this protein, unlike the P-domain, is conservative (Cotten *et al.*, 2014; Ford-Siltz *et al.*, 2021). The main role of the P-domain is to regulate the size and stability of the HuNoV capsid by establishing intermolecular contacts between VP1 dimeric subunits, as well as interacting with blood group antigens (HBGA). When using virus-like particles in the form of vaccines, it is important that the peptides in such structures, despite mutations, induce protective reactions against several genogroups and genotypes of norovirus. We determined the conservatism of the identified B-cell linear epitopes. The identified peptides were found to be less than 100% conserved among the GII.3 [P12], GII.2 [P16] and GII.6 [P7] norovirus genotypes. 100% identity with GII.4 Sydney 2012 [P16] and among the 2022 VP1 sequences was found for only one peptide $YSGRNTHNV_{408\text{-}416}$. The analysis performed confirms the high intensity of mutation processes in this protein, and this process affects potential protective epitopes (Ford-Siltz *et al.*, 2021).

In addition, the production of protective antibodies and the escape of the virus from immune surveillance are affected by the dynamics of the capsid. The spatial alteration of VP1 parts allows the immunogenic sequences to be hidden from direct recognition by receptors. Due to the linker site between the two domains, the P-domain can be in an open «floating» conformation, rising at a distance of ~15 Å above the S-domain, or in a «closed» state, which contributes to the preservation of the viability of the virus. The researchers suggest that, probably, in the open state, protease-sensitive sites open at the P-domain, as a result of which soluble forms of the P-domain are formed, allowing the virus to escape immunity. However, on the other hand, the open state of the virion makes the virus vulnerable to neutralizing antibodies that recognize buried epitopes of S- and P-domains (Smith H., Smith T., 2019).

Using *in silico* methods, we have identified B-cell linear and conformational epitopes of this protein. The bioinformatically determined position of the identified epitopes is consistent with the literature, according to which the accessible sites for B-cell epitopes are flat, mobile protein loops (Karplus & Schulz, 1985). Previously, the topology of several epitopes was also established using mAbs: epitopes A (294-298, 368, 372-373), C (339-341, 375-378), D (393-397), E (407, 411-414) and G (352, 355-357, 359, 364) using GII.4 norovirus as an example (Debbink *et al.*, 2012; Kendra *et al.*, 2021; Tohma *et al.*, 2022; Zheng *et al.*, 2022). In this work, we have identified epitopes previously identified in norovirus VP1 and identified a new putative B-cell epitope at position 307-316 of the P2 subdomain. The neutralizing capacity of this antigenic site should be investigated by in vitro methods in the future.

It should be noted that along with the use of VP1 in vaccines based on VLPs, it is possible to use individual peptides as part of epitope vaccines. In our study, it was found that the region from residues 378 to 394 is a potential source for the formation of highly immunogenic ligands for HLA I and HLA II, as well as an anti-

body recognition site. In our opinion, this peptide is promising for use in peptide vaccines. Overall, we have provided an in silico characterization of the antigenic properties of the VP1 protein of the norovirus GII.4 Sydney [P16] to further test the in vitro immune properties of the vaccine candidate protein.

## Conclusion

The bioinformatic analysis of the primary and tertiary structure of the norovirus VP1 protein showed that the studied protein is potentially capable of inducing T- and B-cell immune responses due to the presence of immunogenic epitopes.

**References**


CAMPILLAY-VELIZ C.P., CARVAJAL J.J., AVELLANEDA A.M., ESCOBAR D., COVIÁN C., KALERGIS A.M. & Lay M.K. (2020): Human Norovirus Proteins: Implications in the Replicative Cycle, Pathogenesis, and the Host Immune Response. *Frontiers in Immunology* **11**, 1–14.

CHHABRA P., DE GRAAF M., PARRA G.I., CHAN M.C., GREEN K., MARTELLA V., WANG Q., WHITE P.A., KATAYAMA K., VENNEMA H., KOOPMANS M.P.G. & VINJÉ J. (2019): Updated classification of norovirus genogroups and genotypes. *Journal of General Virology* **100**, 1393–1406.

COTTEN M., PETROVA V., PHAN M.V., RABAA M.A., WATSON S.J., ONG S.H., KELLAM P. & BAKER S. (2014): Deep sequencing of norovirus genomes defines evolutionary patterns in an urban tropical setting. *Journal of Virology* **88**, 11056–11069.

DEBBINK K., DONALDSON E.F., LINDESMITH L.C. & BARIC R.S. (2012): Genetic mapping of a highly variable norovirus GII.4 blockade epitope: potential role in escape from human herd immunity. *Journal of Virology* **86**(2), 1214–1226.

EPIFANOVA N.V., SASHINA T.A., MOROZOVA O.V., OPARINA S.V., NOVIKOVA N.A. (2022): An increase in prevalence of recombinant GII.3 [P12] norovirus in sporadic acute diarrhea in children in Nizhny Novgorod, Russia, 2018-2021. *Virus Genes* **58**, 467–472.

ESPOSITO S. & PRINCIPI N. (2020): Norovirus Vaccine: Priorities for Future Research and Development. *Frontiers in Immunology* **11**, 1–9.

FORD-SILTZ L.A., TOHMA K. & PARRA G.I. (2021): Understanding the relationship between norovirus diversity and immunity. *Gut Microbes* **13**(1), 1–13.

HALL A.J., LOPMAN B.A., PAYNE D.C., PATEL M.M., GASTAÑADUY P.A., VINJÉ J. & PARASHAR U.D. (2013): Norovirus disease in the United States. *Emerging Infectious Diseases* **19**, 1198–1205.

HARDY M.E. (2005): Norovirus protein structure and function. *FEMS Microbiology Letters* **253**(1), 1–8.

HASSAN E. & BALDRIDGE M.T. (2019): Norovirus encounters in the gut: multifaceted interactions and disease outcomes. *Mucosal Immunology* **12**, 1259–1267.

KARPLUS P.A. & Schulz G.E. (1985): Prediction of chain flexibility in proteins. *Naturwissenschaften* **72**, 212–213.

KENDRA J.A., TOHMA K., FORD-SILTZ L.A., LEPORE C.J. & PARRA G.I. (2021): Antigenic cartography reveals complexities of genetic determinants that lead to antigenic differences among pandemic GII.4 noroviruses. *Proceedings of the National Academy of Sciences of the United States of America* **16**, 1–12.

NIELSEN M., LUNDEGAARD C., WORNING P., LAUEMØLLER S.L., LAMBERTH K., BUUS S., BRUNAK S. & LUND O. (2003): Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein Science* **12**, 1007–1017.

OPARINA S.V., EPIFANOVA N.V., NOVIKOVA N.A. (2023): Features of the circulation of the epidemic variant of norovirus GII.4 Sydney in the Nizhny Novgorod region. In *Molecular diagnostics and biosafety – 2023*; Congress with international participation: Abstracts book (Ed. Akimkin V.G.), Moscow, Russia, 27–28 April 2023, 127-128, Bath: Central Research Institute for Epidemiology.

On the state of sanitary and epidemiological well-being of the population in the Russian Federation in 2021 (2022). State Report, 191–194.

PRASAD B.V., HARDY M.E., DOKLAND T., BELLA J., ROSSMANN M.G. & ESTES M.K. (1999): X-ray crystallographic structure of the Norwalk virus capsid. *Science* **286**, 287–290.

SMITH H.Q. & SMITH T.J. (2019): The Dynamic Capsid Structures of the Noroviruses. *Viruses* **11**(3), 1–18.

TAN M. & JIANG X. (2005): The P-domain of norovirus capsid protein forms a subviral particle that binds to histo-blood group antigen receptors. *Journal of Virology* **79**, 14017–14030.

TAN M., HUANG P., MELLER J., ZHONG W., FARKAS T. & JIANG X. (2003): Mutations within the P2 domain of norovirus capsid affect binding to human histo-blood group antigens: evidence for a binding pocket. *Journal of Virology* **77**, 12562–12571.

TOHMA K., FORD-SILTZ L.A., KENDRA J.A. & PARRA G.I. (2022): Dynamic immunodominance hierarchy of neutralizing antibody responses to evolving GII.4 noroviruses. *Cell reports* **39**(2), 1–23.

WALKER F.C. & BALDRIDGE M.T. (2019): Interactions between noroviruses, the host, and the microbiota. *Current opinion in virology* **37**, 1–9.

ZHENG G.L., ZHU Z.X., CUI J.L. & YU J.M. (2022): Evolutionary analyses of emerging GII.2 [P16] and GII.4 Sydney [P16] noroviruses. *Virus Evolution* **8**(1), 1–11.