

# *In-silico* Immunomodelling of 2019-nCoV

Amirhosein Maali (✉ [maali.amirhosein@gmail.com](mailto:maali.amirhosein@gmail.com))

Pasteur Institute of Iran, Tehran, Iran <https://orcid.org/0000-0002-0699-3010>

**Hossein Teimouri**

Golestan University of Medical Sciences, Golestan, Iran

**Shahin Amiri**

Pasteur Institute of Iran, Tehran, Iran

**Setare Adibzadeh**

Pasteur Institute of Iran, Tehran, Iran

**Mehdi Azad**

Qazvin University of Medical Sciences, Qazvin, Iran

---

## Research Article

**Keywords:** 2019-nCoV, Bioinformatics, Coronaviridae, Severe Acute Respiratory Syndrome

**Posted Date:** April 22nd, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-23657/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Background:** Novel Corona Virus 2019 (2019-nCoV) is a positive-sense single-strand RNA virus from coronaviridae family, responsible for corona virus infectious disease 2019 (COVID-19) with rapid transmission. The aim of this study is characterization of major viral proteins, prediction of antigen proteasomal cleavage pattern, MHC class I processing and presentation, B- and T-cell epitopes, and anti-inflammatory epitopes of 2019-nCoV, compared with SARS-CoV.

**Methods:** The aminoacid sequence of spike surface (S) glycoprotein, membrane (M) glycoprotein, envelop (E) protein and nucleocapsid (N) phosphoprotein were obtained from NCBI. The sequences were aligned by MEGA 7.0 and modeled by SWISS-MODEL. The proteasomal cleavage pattern, MHC class I processing and T-cells epitopes were predicted via IEDB analysis and EPISOFT. The B-cell epitopes were predicted by BepiPred 2.0. Also, prediction of anti-inflammatory epitopes was performed by AntiFlam.

**Results:** Two major antigen proteins, S glycoprotein and M glycoprotein of 2019-nCoV, respectively, have 26.57% and 20.59% less efficiency in proteasomal cleavage and presentation to MHC class I, comparing SARS-CoV. There are less B-cell predicted epitopes in 2019-nCoV, comparing SARS-CoV. The anti-inflammatory properties of 2019-nCoV S glycoprotein and N protein is higher than SARS-CoV.

**Discussion:** It seems that the evolution of 2019-nCoV is on the way of deficiency in antigen presenting to MHC class I and escaping from cellular immunity. Also, the predicted hotspot epitopes potentially can be used to induction of adaptive cellular immunity against 2019-nCoV. In addition, 2019-nCoV appears to be less immunopathogenic than SARS-CoV due to its higher anti-inflammatory proteins.

## Introduction

At the moment, pandemic of Novel Corona Virus 2019 (2019-nCoV) [responsible for coronavirus infectious disease of 2019 (COVID-19)] is a huge challenge facing to infectious medicine (1). This positive-sense single-strand RNA virus, a member of coronaviridae family, has high invasion power and rapid transmission rate via droplet and aerosol. 2019-nCoV has more than 186,000 confirmed cases and 4,500 mortality worldwide, as 22 February 2020 (2).

The primary role of the immune system in the face of viral infections is up to cellular immunity. Antigen-presenting cells (APCs) with antigen proteasomal cleavage presents digested peptides to T helper lymphocytes via MHC class I (3). Also, it has been shown that humoral immune system plays a role in adaptive immunity against viruses (4). Despite conventional coronaviruses, e.g. Severe Acute Respiratory Syndrome corona virus (SARS-CoV), 2019-nCoV shows more intensity in transmission and immune response failure. Bioinformatics tools have enabled us to simulate the evolution mechanism of 2019-nCoV in the way of incomplete immune response.

Sequencing of 2019-nCoV genome in the early days of COVID-19 pandemic has enabled humans to predict the probable protein and nucleoprotein structures of 2019-nCoV as a proteomic-based therapeutic and prophylaxis approaches. It is possible that 2019-nCoV has evolved in antigen presentation to adaptive immune system in the way of survival and escaping from immune system. This evolution may be responsible for the alteration of presented

proteomic homology, proteasomal cleavage pattern, Major histocompatibility complex (MHC) class I presenting, and B- and T-cells epitopes of SARS-CoV and 2019-nCoV with different bioinformatics tools, in four major viral proteins, including spike (S) surface glycoprotein, membrane (M) glycoprotein, envelop (E) protein and nucleocapsid (N) phosphoprotein.

## Methods

### *2.1 Proteome Sequences and Alignment*

The amino acid sequences of S glycoprotein, M glycoprotein, E protein and N phosphoprotein in SARS-CoV (Accession Number NC\_004718) and 2019-nCoV (Accession Number MT106053) were obtained from “Nucleotide” database [National Center for biotechnology information (NCBI)]. The open reading frame of S glycoprotein, M glycoprotein, E protein and N phosphoprotein were aligned by MEGA ver.7.0 for investigation of row homology.

### *2.2 In-silico Characterization of Viral Proteins*

After approval of difference in SARS-CoV and 2019-nCoV proteomic sequence, the mentioned proteins were characterized via SWISS-MODEL (Swiss Institute of Bioinformatics, Biozentrum). (5-8) The comparative 3D-structures, table of comparison with Non-redundant set of PBD structure (adjusted by normalized QMEAN4 score and residual size), and Ramachandran plots were plotted to estimating the difference of the individual protein structure model between SARS-CoV and 2019-nCoV viral proteins.

### *2.3 Prediction of Proteasomal Cleavage Pattern and T Cell-Related Potential Epitope*

To prediction of proteasomal cleavage and MHC class I processing of viral antigens, The MHC class I binding predictions were made on 2/26/2020 using the IEDB analysis resource Consensus tool (9) which combines predictions from ANN aka NetMHC (4.0) (10-12), SMM (13) and Comblib (14). The results were categorized by 9-meric peptides, as the biological digestion pattern. For each 9-meric peptide, Proteasome Score, Torsion angle propensity (TAP) Score, MHC Score, Processing Score, Total Score and MHC- 50% inhibitory concentration (IC50) were reported. Then, all results were sorted by “Total Score.” Top-ten 9-meric peptides were inserted to EPISOFT. “EPISOFT predicts epitope HLA I (MHC class I) binding profiles and population protection computes (PPC). It also identifies minimal sets of epitopes that reach a target

PPC for 5 distinct user-selected ethnic groups,” manufacturer says (15). The EPISOFT results report PPC score and MHC class I binding profile. The statistical analysis was performed by EXCEL 2019.

#### ***2.4 Prediction of B Cell-Related Potential Epitope***

Sequential B-Cell Epitope Predictions were done by BepiPred ver.2.0. “The BepiPred-2.0 server predicts B-cell epitopes from a protein sequence, using a Random Forest algorithm trained on epitopes and non-epitope amino acids determined from crystal structures,” manufacturer says (16).

#### ***2.5 Prediction of Anti-inflammatory Peptides***

To predict the anti-inflammatory property of 9-meric epitopes on S glycoprotein, M glycoprotein, E protein and N phosphoprotein proteins in SARS-CoV and 2019-nCoV protein done via “Prediction of Anti-inflammatory Epitopes server (AntiFlam)” [metagenomics.iiserb.ac.in/antiinflam]. “It allows the users to predict the anti-inflammatory nature of the multiple variants of the query peptide (substitution of each amino acid of the peptide with other amino acids), and thus, is useful in assessing the position-specific effects of each amino acid in modulating the anti-inflammatory activity of the peptide”, manufacturer says.

## **Results**

### ***3.1 Minor Differences in Protein Homology of SARS-CoV and 2019-nCoV***

There are no major, but Minor differences within the structure of homologous S glycoprotein, M glycoprotein, E protein and N phosphoprotein of SARS-CoV and 2019-nCoV (*Figure 1*). The filled 3D-structures shows different exposure pattern, implying that sequences of 2019-nCoV viral proteins directly affect on 3D-structure and surface exposure pattern. This result establishes evolution of 2019-nCoV proteome being potentially able to lead to the different behavioral pattern, comparing to SARS-CoV. Minor differences in dots of Ramachandran plots confirm minor, but not major, structural homologous variation of SARS-CoV and 2019-nCoV.

Additionally, comparison of modeled proteins with Non-redundant set of PBD structure (figure 1-b) confirms the comparative results.

### ***3.3 Difference in Proteasomal Cleavage Pattern and MHC Class I Presentation of T-cells Epitopes of 2019-nCoV and SARS-CoV***

The IEDB analysis showed that the proteasomal cleavage pattern of 9-meric peptides of 2019-nCoV and SARS-CoV are different. The top-ten 9-mer peptides of each viral protein were compared via “total score” index. The S glycoprotein of 2019-nCoV has 26.57% less efficiency in antigen proteasomal cleavage and presentation to MHC class I, compared to SARS-CoV S glycoprotein. Also, the other giant surface-exposed protein, M protein of 2019-nCoV, has 20.59% less efficiency in proteasomal cleavage and presentation to MHC class I compared to SARS-CoV M glycoprotein. (*Table 1*)

Despite above proteins, the E protein and N phosphoprotein of 2019-nCoV and SARS-CoV show the different pattern in antigen proteasomal cleavage and presentation to MHC class I. There is 0.08% more efficiency in the cleavage of E protein of SARS-CoV and its presentation to MHC class I, compared to 2019-nCoV. Also, proteasomal cleavage and MHC class I presentation of N phosphoprotein of SARS-CoV done 15.19% more efficient, compared to 2019-nCoV.

As our results show, “LTDEMIAQY” (total score: 1.71, PPC: 0.0423), “CVADYSVLY” (total score: 1.16, PPC: 0.2654) and “TSNQVAVLY” (total score: 0.94, PPC: 0.2724) are top-three hotspot 9-meric presented peptides in S glycoprotein of 2019-nCoV. In addition, “ATSRTLSTYY” (total score: 0.92, PPC: 0.0159) and “YANRNRFLY” (total score: 0.51, PPC: 0.0140) are top-two hotspot 9-meric presented peptides in M glycoprotein of 2019-nCoV.

### ***3.2 Difference in B-cells Epitopes of 2019-nCoV and SARS-CoV***

The results of prediction of potential B-cells epitopes via BepiPred V2.0 show that S glycoprotein and M glycoprotein of 2019-nCoV contain higher-potent B-cell epitopes than SARS-CoV, while E protein and N phosphoprotein of SARS-CoV contain higher-potent B-cell epitopes than 2019-nCoV. (*Figure 2*) Also, SWISS-MODEL results of structure prediction approved that the predicted epitopes have surface exposure.

### ***3.3 Differences in Anti-inflammatory Properties of 2019-nCoV and SARS-CoV Epitopes***

The Protein scan of four major viral proteins of 2019-nCoV and SARS-CoV, *e.g.* S glycoprotein, M glycoprotein, E protein and N phosphoprotein proteins, show that the 9-meric epitopes of 2019-nCoV have higher anti-inflammatory properties in S glycoprotein (264 9-meric epitopes in 2019-nCoV [Top-ten mean score =  $4.50 \pm 0.10$ ], comparing 77 9-meric epitopes in SARS-CoV [Top-ten mean score =  $4.34 \pm 0.36$ ]) and N phosphoprotein (224 9-meric epitopes in 2019-nCoV, comparing 54 9-meric epitopes in SARS-CoV). Despite two above proteins, analysis of the anti-inflammatory property of the 9-meric M glycoprotein epitopes in 2019-nCoV and SARS-CoV shows that this protein has the slightly more anti-inflammatory property SARS-CoV (29 9-meric anti-inflammatory epitopes), comparing 2019-nCoV (23 9-meric anti-inflammatory epitopes). The analysis of E protein in these two viruses also showed the same result (7 9-meric anti-inflammatory epitopes). (*Table 2*)

## Discussion

2019 novel corona virus (2019-nCoV), which is responsible for corona virus infectious disease 2019 (COVID-19), evolved to higher transmission potential comparing to Severe Acute Respiratory Syndrome corona virus (SARS-CoV), which introduced this virus as a global concern. Bioinformatic tools equipped us to realize the molecular evolution pattern of 2019-nCoV.

After the virus enters the host body, the antigen presenting cells (APCs) capture them and start proteasomal cleavage process of viral proteins. After antigen digestion, 9-meric peptides present to T-helper lymphocytes in the cellular immunity system, via major histocompatibility complex (MHC) class I. Now, cellular immunity leads to eliminate infectious cells from host body. Different pattern of viral protein digestion leads to different MHC class I presenting, and finally, different immune responses to viruses. Spike (S) surface glycoprotein, membrane (M) glycoprotein, envelop (E) protein and nucleocapsid (N) phosphoprotein are major structural proteins in 2019-nCoV. Our results authenticate that there is different in proteasomal cleavage pattern of S glycoprotein, M glycoprotein, E protein and N phosphoprotein in 2019-nCoV comparing to SARS-CoV. The results show that two major antigen proteins of 2019-nCoV, S glycoprotein and M glycoprotein, have 26.57% and 20.59% fewer efficiency in antigen proteasomal cleavage and presentation to MHC class I respectively, comparing to SARS-CoV. Thus, cellular immune system is more powerless in the elimination of 2019-nCoV, in comparison with SARS-CoV.

Our results show that there is minor difference homology in row alignment of S glycoprotein, M glycoprotein, E protein and N phosphoprotein. Also, SWISS-MODEL results reveal the minor difference in row homology leads to minor changes in 3D-structure of these viral proteins. In another hand, various researches established the role of humoral immunity against viral infections. Thus, it is possible that 2019-nCoV has evolved to escape from humoral immune system. The exposed B-cell epitopes of 2019-nCoV are predicted by BepiPred, and were compared to SARS-CoV. The results showed that S glycoprotein and M glycoprotein (most exposed proteins) of SARS-CoV contain higher potent epitopes. So, it seems that probably, humoral immune system is more involved in response to SARS-CoV compared to 2019-nCoV.

IEDB results predict "LTDEMIAQY" 9-mer as higher scored cleaved peptide (total score: 1.71). "LTDEMIAQY" binds to different MHC class I alleles (A0207, B1508, B1516, B3801, B5702 and B5801) CPP 0423. Thus, we suggest that "LTDEMIAQY" linear peptide potentially can be applied as peptide Antigen vaccine, to induction cellular immunity prophylaxis. Subsequent research can be conducted on *in-vitro* stimulation of T-helper lymphocytes by "LTDEMIAQY" peptide or generation of induced "LTDEMIAQY" presenting cells.

Control of inflammatory response of the immune system to these two viruses is of great importance in controlling the complications, since the basis of the pathogenesis in these two viruses is immunopathogenic. According to the analysis results of the anti-inflammatory properties of 9-meric epitopes on the four main proteins of 2019-nCoV and SARS-CoV, 2019-nCoV appears to be less pathogenic than SARS-CoV due to its higher anti-inflammatory proteins and better control of the

inflammatory response of the immune system. The less mortality of 2019-nCoV confirms this. The immune response should be controlled to the extent that it prevents both the proliferation of the virus and does not lead to inflammation and its complications. Therefore, it is better to prioritize the design of the drug and vaccine to achieve this goal.

## Declarations

### Acknowledgment

We declare that there is no Acknowledgment.

### Conflict of interest

We declare that there is no conflict of interest.

## References

1. World Health Organization. Novel Coronavirus (2019-nCoV): situation report 52. 2020. Available from: <https://www.who.int/docs/default-source/coronaviruse/20200312-sitrep-52-covid-19.pdf>
2. World Health Organization. Coronavirus disease 2019 (COVID-19): Situation Report 33. 2020. Available from: <https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200222-sitrep-33-covid-19.pdf>
3. Buseyne F, Le Gall S, Boccaccio C, Abastado J-P, Lifson JD, Arthur LO, et al. MHC-I-restricted presentation of HIV-1 virion antigens without viral replication. *Nature medicine*. 2001;7(3):344-9.
4. Oxenius A, Bachmann MF, Zinkernagel RM, Hengartner HJEjoi. Virus-specific major MHC class II-restricted TCR-transgenic mice: effects on humoral and cellular immune responses after viral infection. *European journal of Immunology*. 1998;28(1):390-400.
5. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer F.T., de Beer T.A.P., Rempfer C, Bordoli L, Lepore R, Schwede T. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res*. 46(W1), W296-W303 (2018).
6. Guex N, Peitsch M.C., Schwede, T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis* 30, S162-S173 (2009).
7. vBenkert P, Biasini M., Schwede, T. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* 27, 343-350 (2011).
8. Bertoni M, Kiefer F., Biasini M., Bordoli L., Schwede, T. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. *Scientific Reports* 7 (2017).
9. Kim Y, Ponomarenko J, Zhu Z, Tamang D, Wang P, Greenbaum J, Lundegaard C, Sette A, Lund O, Bourne PE, Nielsen M, Peters B. 2012. Immune epitope database analysis resource. *NAR*.

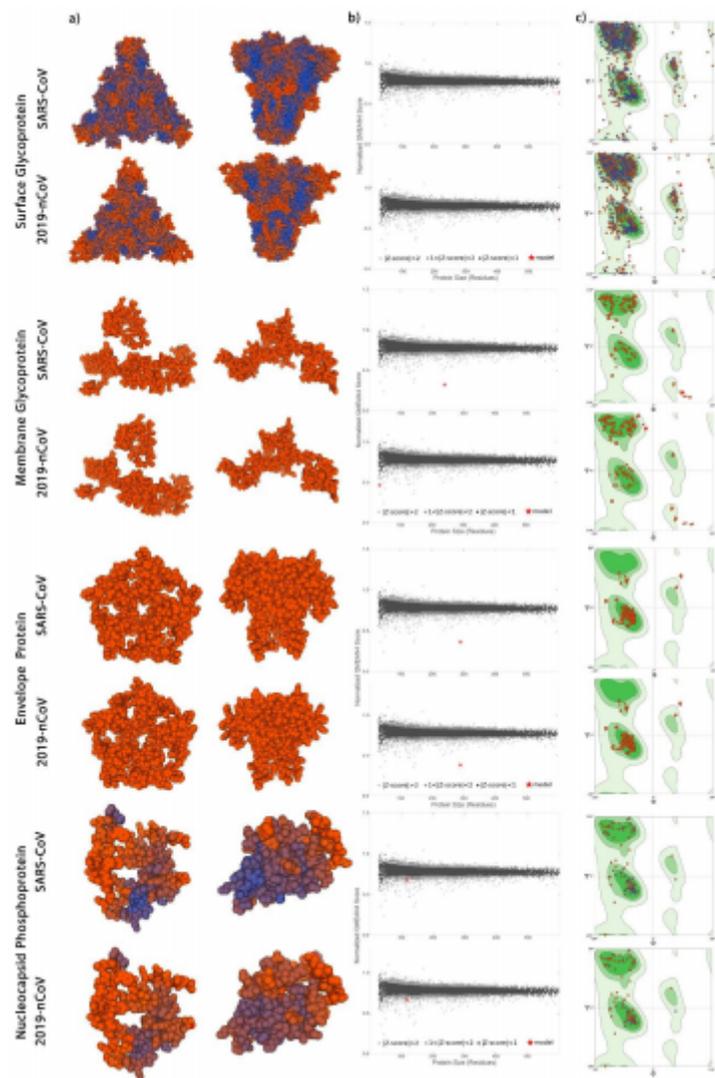
10. Nielsen M, Lundegaard C, Worning P, Lauemøller SL, Lamberth K, Buus S, Brunak S, Lund O. 2003. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein Sci* 12:1007-1017.
11. Lundegaard C, Lamberth K, Harndahl M, Buus S, Lund O, and Nielsen M. 2008. NetMHC-3.0: Accurate web accessible predictions of Human, Mouse, and Monkey MHC class I affinities for peptides of length 8-11. *NAR* 36:W509-512.
12. Andreatta M. and Nielsen M. 2016. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinformatics* 32:511-7.
13. Peters B, Sette A. 2005. Generating quantitative models describing the sequence specificity of biological processes with the stabilized matrix method. *BMC Bioinformatics* 6:132.
14. Sidney J, Assarsson E, Moore C, Ngo S, Pinilla C, Sette A, Peters B. 2008. Quantitative peptide binding motifs for 19 human and mouse MHC class I molecules derived using positional scanning combinatorial peptide libraries. *Immunome Res* 4:2.
15. Molero-Abraham M, Lafuente EM1, Flower DR, Reche PA (2013) Selection of conserved epitopes from hepatitis C virus for pan-population stimulation of T-cell responses. *Clin Dev Immunol.* 2013:601943. doi: 10.1155/2013/601943. Epub 2013 Nov 21
16. Jespersen MC PB, Nielsen M, Marcatili P. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res* 2017 (Web Server issue). doi: 10.1093/nar/gkx352

## Tables

Please see the supplementary files section to view the tables.

## Figures

**Figure 1**



**Figure 1**

SWISS-MODEL characterization of surface glycoprotein, membrane glycoprotein, envelop protein and nucleocapsid phosphoprotein of 2019-nCoV and SARS-CoV. a) The top-view and side-view of filled 3D-structures shows different exposure pattern, implying that different row sequences of 2019-nCoV viral proteins directly effect on 3D-structure and surface exposure pattern, but not majorly. b) table of comparison with Non-redundant set of PBD structure (adjusted by normalized QMEAN4 score and residual size) and c) Ramachandran plots, approves minor differences of homology of surface glycoprotein, membrane glycoprotein, envelop protein and nucleocapsid phosphoprotein of 2019-nCoV and SARS-CoV.

Figure 2

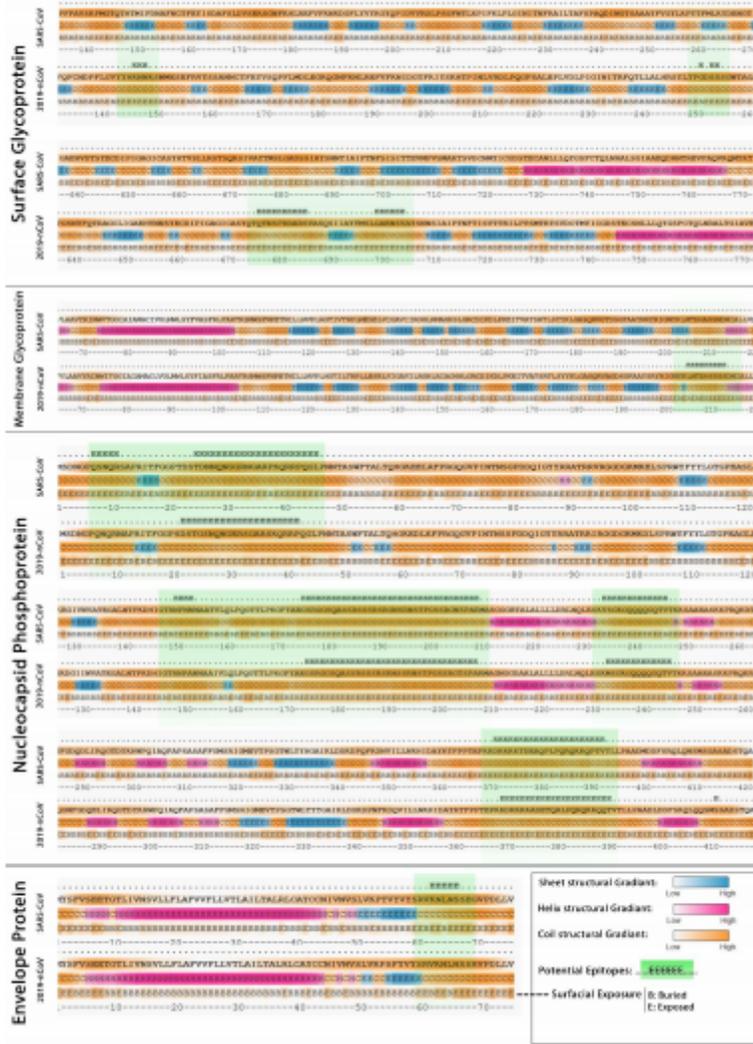


Figure 2

Difference in B-cells Epitopes of 2019-nCoV and SARS-CoV in surface glycoprotein, membrane glycoprotein, envelop protein and nucleocapsid phosphoprotein via BepiPred. The results of prediction of potential B-cells epitopes via BepiPred show that surface glycoprotein, membrane glycoprotein of 2019-nCoV contain higher-potent B-cell epitopes than SARS-CoV, while envelop protein and nucleocapsid phosphoprotein of SARS-CoV contain higher-potent B-cell epitopes than 2019-nCoV. The epitope threshold is taken 0.64. All predicted epitopes contain coiled structure and exposed surface.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Table2.PNG](#)
- [Table1Part2.PNG](#)

- [Table1Part1.PNG](#)
- [Table1Part3.PNG](#)
- [ARVID20002842.pdf](#)
- [TableLegends.docx](#)