Acta Scientific Biotechnology

Volume 2 Issue 5 August 2021

Research Article

MHC Binding Potential of RBD Domain in Spike Protein of SARS CoV2

Shiny Pravallikan Karukutla and Krupanidhi Sreerama*

Department of Biotechnology, Vignan's Foundation for Science, Technology and Research, India

***Corresponding Author:** Krupanidhi Sreerama, Department of Biotechnology, Vignan's Foundation for Science, Technology and Research, India.

Received: July 09, 2021

Published: July 29, 2021

Abstract

SARS-CoV-2 infection became pandemic and public concern due to the outbreak of COVID-19. Since its appearance in Wuhan, China in December 2019 the infection spread across continents. Substantial numbers of global population were disrupted both in their health and family structure. The spike protein is found to be the key protein of coronavirus as it anchors to human alveolar epithelial cells. Hence, the present study is focused on immune informatics analysis of spike glycoprotein (S-protein) of SARS-CoV-2 to explore MHC binding potential of its epitopes. Out of 5058 peptides, each with nine residues obtained through NetMHCpan version 4.1b, six peptides were predicted as epitopes with strong binding potential to HLA-A*02:01, whose distribution is ~33.8% among Indian Asians. Among these six peptides, the amino acid sequence starting from 417th position namely KIADYNYKL was found to be highly probable antigen with a value of 1.6639 derived through VaxiJen V2.0. Peptides predicted in the receptor binding region also revealed that the aforementioned peptide ranked as the strong binder to HLA-A*02:01. The secondary structure prediction of RBD performed through Phyre2 revealed that these peptide residues spanned to stretch in beta strand followed by random coil and reported as immunodominant region using Immunome Browser tool. This study would pave the path to understand the prevalent Delta Plus variant of CoV2 (K417N) and facilitates to design a recombinant vaccine.

Keywords: HLA-A02:01; Spike Protein; SARS CoV2; Delta Plus Variant

Introduction

The 'self-molecule' namely MHC was introduced by British immunologist Peter Gorer in 1936 [1]. George Snell [2] unfolded the role of MHC in the organ transplantation. Its participation in antigen presentation was elucidated and further described its genetic complex system including their haplotypes prevailing in human population by Jean Dausset and Jan van Rood who brought the jargon namely human leukocyte antigen into the literature [3,4].

MHCs are transmembrane cell surface glycoprotein molecules present on all nucleated cells with varying distribution for MHC class I and MHC class II. The attribute of antigen presentation bestowed on MHC molecule an important role in immune elicitation in the host organism. Hence, the peptide residues of antigen have to anchor to the residues of MHC cleft and the same would be predicted through bioinformatics tools. An antigen after having been processed, its fragments anchor to MHC of dendritic cells which further accelerate a series of steps involved in immune elicitation. These investigations have attracted several scientists across the globe to explore *in silico* design of potential epitopes of S-protein [5-10].

In the present article an attempt is made to explore immunoinformatics tools to evaluate the potential epitopes of RBD region of S-protein in SARS CoV2 to understand the Delta variant (K417N) prevailing in India during the second wave of Covid-19.

Methods

The spike protein sequence of Wuhan coronavirus was retrieved from NC_045512 in GeneBank. The protein id of the chosen sequence in the present study is YP_009724390.1 and its related gene id is 43740568. The S-protein sequence of Wuhan isolate comprises of 1284 amino acids. The receptor binding protein ranges from ~401-541 residues of S-protein viz., "VIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNF". The highlighted residues were prone for mutation which led to Delta Plus variant (K417N) and Kurnool variant (N440K) in India.

NetMHCpan version 4.1b open software tool [11] was employed to predict residues of S-protein to bind to HLA-A*02:01 and tabulated the results in table 1. Also, the prediction was done for peptides of RBD region of S-Protein as 417th position is prone for Dalta Plus variant. VaxiJen V2.0 online tool was used to predict the antigenicity of peptide binding to HLA-A*02:01 [12]. The secondary structure of RBD region was determined through Phyre2 software tool to evaluate whether the peptide at 417th position was either in beta strand/helix/random coil [13]. Further, to confirm and visualize the 417th position as an immunodominant region, the online Immunome Browser tool was employed [14].

Residue Position	Peptide	Score_EL	%Rank_EL	Bind Level	Overall Prediction for the Protective Antigen
269	YLQPRTFLL	0.9711980	0.013	<= SB	0.4532 (Probable ANTIGEN)
976	VLNDILSRL	0.9384980	0.028	<= SB	-0.8524 (Probable NON-ANTIGEN)
109	TLDSKTQSL	0.9149980	0.041	<= SB	1.0685 (Probable ANTIGEN)
1000	RLQSLQTYV	0.8737600	0.062	<= SB	-0.2167 (Probable NON-ANTIGEN)
417	KIADYNYKL	0.8646110	0.067	<= SB	1.6639 (Probable ANTIGEN)
983	RLDKVEAEV	0.8250450	0.090	<= SB	0.0765 (Probable NON-ANTIGEN)

Table 1: Peptides predicted in the spike protein of SARS CoV2 (YP_009724390.1) as strong binders to HLA-A*02:01 using NetMHCpan version 4.1b (http://www.cbs.dtu.dk/services/NetMHCpan/) [11].
Overall prediction of the protective antigen was evaluated through VaxiJen V2.0 [12].

Results and Discussion

It is interesting that within a period of 20 months since December 2019, SARS CoV2 has undergone several mutations and resulted in Delta, Alpha, Beta, Gamma and Variants of Interest clades [15]. A total of 2,038,763 full genomes of SARS CoV2 were sequenced by the team from GISAID Initiative [15] and reported that Delta variant was first noticed in India in October, 2020. 57% and 89% of samples collected from India in the months of April and May 2021 comprised of Delta variants (gisaid.org/variants). These variants became highly infective and found resistant to immune surveillance in the convalescent human host, thus spreading and causing increased risk. Hence, it is imperative to unfold features of the Variants of Interest and Variants of Concern among emerging CoV2s.

The S-protein being a characteristic attribute of coronavirus is attracting the attention of researchers to explore its binding potential to the anchoring residues located in the cleft of MHC class I. The obtained S-protein sequence residues from GenBank was submitted to NetMHCpan version 4.1b open software tool [11] and observed the appearance of 5058 peptides. Of which, six peptides shown in Table 1 ranked the highest score ranging from 82% to 97% with strong binding potential to HLA-A*02 which is prevalent in Indian population with a frequency of 33.8% [16]. The VaxiJen tool predicted these six peptides for their antigenicty, of which the peptide at 417th position was predominantly reported to be highly antigenic with a score of 1.6639, the highest among the three peptides identified as probable antigens (Table 1) [12]. An attempt is

also made to choose the peptide sequence of delta plus variant of S-protein by replacing "K" with "N" residue at 417th position (NIADYNYKL) and submitted to VaxiJen V2.0 which yielded a value viz., 1.5485 comparable to the wild strain, indicating as probable antigen. Further, the RBD region of the S-protein was subjected for the evaluation of the MHC binding potential using NetMHCpan version 4.1b. Interestingly, it was shown that 17th (417th position in S-protein) residue position of RBD bearing the peptide KIADYNYKL was found to be predicted as strong binder with 86% score. Thus, the Delta variants as compiled and reported by GISAID [15] are persisting as infective agents. These results were further authenticated through Immunome Browser tool [14] which confirmed that 417th residue bearing peptide as an immunodominant region to bind to MHC I.

Residue Position	Peptide	Score_EL	%Rank_EL	Bind Level
17	KIADYNYKL	0.8646110	0.067	<= SB
112	VLSFELLHA	0.3507340	0.627	<= WB
116	ELLHAPATV	0.3364190	0.664	<= WB
24	KLPDDFTGC	0.3349030	0.668	<= WB
195	YQPYRVVVL	0.1890140	1.193	<= WB

Table 2: Peptides predicted in the RBD region of spike protein of SARS CoV2 (YP_009724390.1) as strong binders to HLA-A*02:01 using NetMHCpan version 4.1b (http://www.cbs.dtu.dk/services/NetMHCpan/) [11].

Score_EL: The raw prediction score.

% Rank_EL: Rank of the predicted binding score compared to a set of random natural peptides.

SB: Strong binders are defined as % rank <0.5.

Conclusion

The Delta variant of SARS CoV2 was first noticed in India in October, 2020. These variants became highly infective. It is analysed that six peptides of S-protein scored 82% to 97% with strong binding potential to HLA-A*02 which is prevalent in Indian population. The Delta plus variant of S-protein is found comparable to wild strain, as a potential infective antigen. The peptide at 417th position bearing the sequence KIADYNYKL was found to be immunodominant region to bind to MHC I molecule

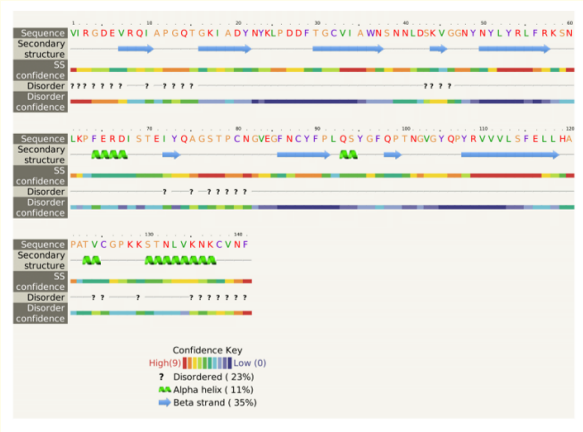

Figure 1: Secondary structure prediction of RBD region of the S-protein of SARS CoV2 (YP_009724390.1) through Phyre2 [13]. The strong binder residues from 17th position namely KIADYNYKL was found in the beta strand followed by random coil.

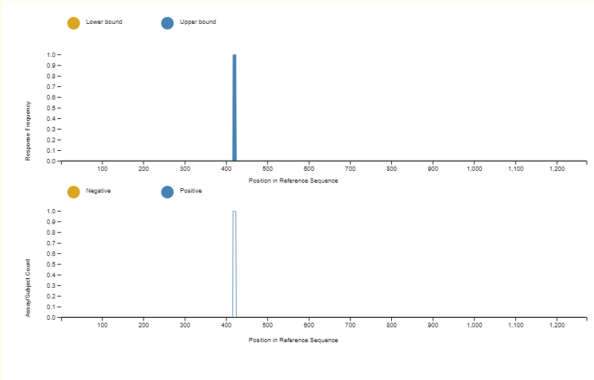

Figure 2: The immunodominant region of S-protein from 417th position is authenticated through Immunome Browser tool http://tools.iedb.org/immunomebrowser/ [14].

Acknowledgements

Authors acknowledge DST, Govt. of India, New Delhi (FIST Project No: LSI-576/2013), Net-working facility and Centre of Excellence, Department of Biotechnology, Vignan`s Foundation for Science, Technology and Research for providing facilities.

Bibliography

1. Thorsby E. "A short history of HLA". *Tissue Antigens* 74.2 (2009): 101-116.

2. Snell GD. "Studies in histocompatibility". *Science* 213.4504 (1981): 172-178.

3. Dausset J and Soulillou JP. "An interview with Jean Dausset". *American Journal of Transplantation: Official Journal of the American Society of Transplantation and the American Society of Transplant Surgeons* 4.1 (2004): 4-7.

4. Cruz-Tapias P., *et al*. "Major histocompatibility complex: antigen processing and presentation". InAutoimmunity: From Bench to Bedside. El Rosario University Press (2013).

5. Kolla HB., *et al*. "An immunoinformatics approach for the design of a multi-epitope vaccine targeting super antigen TSST-1 of Staphylococcus aureus". *Journal of Genetic Engineering and Biotechnology* 19.1 (2021): 1-4.

6. Peele KA., *et al*. "Insilico drug repurposing using FDA approved drugs against Membrane protein of SARS-CoV-2". *Journal of Pharmaceutical Sciences* 110.6 (2021): 2346-54.

7. Karlapudi AP., *et al*. "Design of CRISPR-based Targets for the Development of a Diagnostic Method for SARS-CoV-2: An in Silico Approach". *Eurasian Journal of Medicine and Oncology* 4.4 (2020): 304-308.

8. Abraham Peele K., *et al*. "Design of multi-epitope vaccine candidate against SARS-CoV-2: A in-silico study". *Journal of Biomolecular Structure and Dynamics* 27 (2020): 1-9.

9. Ayyagari VS., *et al*. "Design of a multi-epitope-based vaccine targeting M-protein of SARS-CoV2: an immunoinformatics approach". *Journal of Biomolecular Structure and Dynamics* 16 (2020): 1-5.

10. Krupanidhi S. "In silico design of multiepitope vaccine candidate against SARS CoV and CoV2". 7.2 (2020): 61-73.

11. Reynisson B., *et al*. "NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data". *Nucleic Acids Research* 48 (2020): W449-454.

12. Doytchinova IA and Flower DR. "VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines". *BMC Bioinformatics* 8.1 (2007): 1-7.

13. Kelley LA., *et al*. "The Phyre2 web portal for protein modeling, prediction and analysis". *Nature Protocols* 10.6 (2015): 845-858.

14. Dhanda SK., *et al*. "ImmunomeBrowser: a tool to aggregate and visualize complex and heterogeneous epitopes in reference proteins". *Bioinformatics* 34.22 (2018): 3931-3933.

15. https://www.gisaid.org/

16. Mehra NK., *et al*. "Molecular diversity of HLA-A* 02 in Asian Indians: predominance of A* 0211". *Tissue Antigens* 57.6 (2001): 502-507.

Volume 2 Issue 5 August 2021