



REVERSE VACCINOLOGY APPROACH TOWARDS THE PREVENTION OF SARS-COV-2 INFECTIONS

Immunology

Mahadev A. Jadhav* Dept. of Biotechnology, Deogiri College, Aurangabad, Maharashtra, India. *Corresponding Author

Shivani V. Pawar Dept. of Biotechnology, Deogiri College, Aurangabad, Maharashtra, India.

Gargi S. Deshmukh Dept. of Biotechnology, Deogiri College, Aurangabad, Maharashtra, India.

ABSTRACT

The coronavirus family has been infecting the human population for the past two decades, but the ongoing coronavirus called SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2) has posed an enigmatic challenge to global public health security. Current available vaccines are less effective to prevent SARS-CoV-2 infections due to rapid rate mutations in his genome & it's also responsible for increased viral transmissibility leads to morbidity & mortality. The aim of the present study to design a peptide vaccine against SARS-Cov-2 by using epitope (B-cell & T-cell) based reverse vaccinology approach. The structural proteins Spike protein & Nucleoprotein of SARS-CoV-2 were used to identify potential epitope to construct effective peptide vaccine which prevents SARS-CoV-2 infections. Various reverse vaccinology resources & tools such as EXPASY, Vaxijen, IEDB, CTLPred, ALGPred, Pepfold and HPEPDOCK were used for immunoinformatics analysis. We investigated the parameters associated with antigenicity, toxicity, and allergenicity of selected proteins by evaluating their various properties. The result obtained from different steps of immunoinformatics analysis was satisfactory. The CTL epitopes that passed the antigenicity, allergenicity, toxigenicity, and immunogenicity assessment were 02 epitopes from the nucleoprotein becomes effective & safer vaccine to prevent SARS-CoV-2 infections but the experimental tests becomes essential to validate the efficacy of epitope based peptide vaccine.

KEYWORDS

SARS-CoV-2, reverse vaccinology, epitope, antigenicity, toxicity, allergenicity, peptide vaccine

INTRODUCTION

The immune system defends us against infection in a variety of ways. When the immune system isn't working properly; it can lead to diseases like autoimmune allergies. The immune system is a complex network of structures and processes that has evolved to keep us safe from disease. The immune system is made up of molecular and cellular components. These components' functions are divided into non-specific processes and responses that are unique to specific infections [1].

The adaptive immune system is slow, taking many days to activate two critical cell types, B-cells and T-cells. T-cells are divided into two groups: CD4++ and CD8+ cells. Cytokines are released by CD4+ and helper T lymphocytes. B-cells mature into plasma cells as a result of the cytokines. CD8+ cytotoxic T cells, on the other hand, actively kill diseased cells and create antibodies to neutralize the infections [2]. The virus that causes covid-19 is primarily spread through droplets produced by infected people, such as sneezes or exhalations. These droplets are heavy in the air and soon descend to the ground. If you are in close contact with someone who has covid-19, you can be infected by breathing in the virus or touching a contained surface and then touching your eyes or mouth. The majority of persons who get sick with covid-19 have mild to moderate symptoms and recover without any additional therapy. The vaccine that was previously produced had several serious adverse effects, such as fever, body aches, and so on [3].

Like SARS-CoV, SARS-CoV-2 binds to the receptor angiotensin-converting enzyme 2 (ACE2) on the host cell via the receptor binding domain (RBD) on the spike S protein of the virus [4]. The spike S protein of SARS-CoV-2 is a type I transmembrane glycoprotein with a predicted length of 1273 amino acids. Moreover, it comprises the major antigenic determinants that induce neutralizing antibodies [5, 6]. The vaccination process is significantly increased to develop a vaccine against the pandemic SARS-CoV-2, including the development of several RNA and DNA vaccines, recombinant protein vaccines, and cell-culture-based vaccines [7].

The mRNA vaccines are a new type of vaccine to protect against infectious diseases. Recently Food and Drug Administration (FDA) has authorized the emergency use of the Pfizer-BioNTech COVID-19 Vaccine (BNT162b2) to prevent COVID-19 in individuals 16 years of age and older under an emergency use authorization given in two doses 3 weeks apart. However, this vaccine showed allergic reactions such as difficulty in breathing, swelling of the face and throat, fast heartbeat, skin rashes, dizziness, and weakness. Another vaccine by ModernaTX, Inc. (mRNA-1273) is recommended for people aged 18

years and older. But the vaccine also showed side effects that usually started within a day or two of getting the vaccine [8, 9]. The advances made in the field of immunoinformatics tools coincided with the knowledge of the host immune response leading to new disciplines in vaccine design against diseases via computer in silico epitope predictions. The epitopes-driven vaccine is a new concept that is being successfully applied in multiple studies, particularly to the development of vaccines targeting conserved epitopes in variable or rapidly mutating pathogens [10, 11, 12].

However, we designed a vaccination that has no side effects utilizing bioinformatics technologies in reverse vaccinology. Therefore, as the genome and proteome sequences of SARS-CoV-2 are swiftly made available [13, 14], this study aimed to use reverse vaccinology approach to design an epitope based vaccine against SARS-CoV-2 infection from the causative virus structural proteins.

MATERIALS AND METHODS

Data retrieval & Confirmation of antigenicity

The confirmation of antigenic nature of spike protein & nucleoprotein of SARS-CoV-2 was done by using Vaxijen server v2.0 (<http://www.ddg-pharmfac.net/vaxijen/vaxijen/vaxijen.html>), virus model was selected for analysis with threshold value 0.4 and its protein sequences was retrieved from Uniprot Database (www.uniprot.org) and their physicochemical properties were estimated by using ProtParam tool (<https://web.expasy.org/protparam/>) to define protein stability & charge.

B-cell Epitope Prediction

B-cell epitopes aid in the detection of viral infectious agents in the immune system. By using IEDB resource (<http://tools.iedb.org>) B cell epitopes from selected antigenic spike glycoprotein & nucleoprotein was predicted by using Kolaskar and Tongaonkar antigenicity scale method [15] window size 7 was selected for analysis, because it is appropriate for finding regions that may be potentially antigenic.

T-cell Epitope Prediction

T-cell monitor MHC bound ligands that trigger a T-cell immune response. T-cell epitopes in SARS CoV2's antigenic spike glycoprotein & nucleoprotein was predicted by using CTLpred tool (<http://crdd.osdd.net/raghava/ctlpred/>). Epitopes having a consensus score 1 or less than 1 were thought to be good binders and ANN prediction approach & ANN cutoff score 0.51 were chosen for analysis. Only 05 peptides from each antigenic protein were selected for further analysis.

Evaluation Of Allergenicity

The allergenicity of potential selected epitopes was checked to confirm their allergic and non allergic nature to host or host cells. It was predicted by using ALGpred server (<https://webs.iitd.edu.in/raghava/algpred2/>). Hybrid Machine Learning Techniques (RF+BLAST+MERCI) was used to develop model to predict allergenicity.

Secondary Structure Prediction

After a series of experiments and analyses of given sequences, 03 epitopes to reach a domain-containing sequence. The GOR4 (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_gor4.html) method of secondary structure prediction was used to predict the secondary structure of the designed vaccine (potential epitopes).

3d Structure Prediction Of Epitope

3D structure of selected 03 potential epitopes was modelled by using PEP-FOLD tool (<https://mobyle.rpbs.univ-paris-diderot.fr/cgi-bin/portal.py#forms::PEP-FOLD>) and it was further used to check antigen antibody interaction [16].

Antibody Selection & Its 3d Structure Retrieval

Recent study reveals that IgG3 antibody of human immune system responses binds to coronavirus spike proteins and forms an antigen-antibody complex and provokes immune response against SARS CoV2 infections [17]. 3D structure of IgG3 antibody was retrieved from Uniprot structure section ALPHAFOLD predicted model and visualized by using Discovery Studio Visualizer.

Antigen-antibody Docking

To determine the binding affinity and energy of the antigen-antibody complex, the HPEPDOCK server (<http://huanglab.phys.hust.edu.cn/hpepdock/>) was used to perform peptide-peptide molecular docking to confirm antigen-antibody interaction.

RESULTS

Data Retrieval & Confirmation Of Antigenicity

The SARS CoV2 spike & nucleoprotein sequences were retrieved from the Uniprot database for development of a good peptide vaccine candidate (Table 1). Both proteins were discovered to be antigenic. By using Vaxijen server we were able to estimate the antigenicity of viral proteins.

Table1: Details Of Sars Cov2 Antigenic Protein Selected For Reverse Vaccinology

Protein Identifier	Name of the protein	Source Organism	Length (AA)
P0DTC2 · SPIKE_SARS2	Spike glycoprotein	Human coronavirus 229E (HCoV-229E)	1273
A0A6B9VLF5_SARS2	Nucleoprotein	Severe acute respiratory syndrome coronavirus 2 (2019-nCoV) (SARS-CoV-2)	419

To increase the specificity, the cutoff threshold value 0.4 was selected to check full length protein antigenicity of spike and nucleoprotein. It was found that spike protein has 0.4646 & nucleoprotein has 0.5059 antigenicity score to become a good potential antigen (Table 2).

Table 2. Antigenicity Of Sars Cov2 Spike And Nucleoprotein Predicted By Vaxijen V2.0

Protein Identifier	Name of the protein	Model Selected & threshold	Overall Protective Antigen Prediction Score	Antigen Prediction
P0DTC2 · SPIKE_SARS2	Spike glycoprotein	Virus 0.4	0.4646	Antigen
A0A6B9VLF5_SARS2	Nucleo-protein	Virus 0.4	0.5059	Antigen

These two proteins were then further investigated by ProtParam tool to calculate physiochemical properties. It found that spike glycoprotein has a pI: 6.24 which indicate the positive charge of protein and it have a low instability index (II) 33.01 & 84.67 aliphatic index to confirm its thermo stability. Whereas nucleoprotein has a pI: 10.4, 55.09 instability index & 52.53 aliphatic index. It indicates slightly unstable

nature of nucleoprotein.

B-cell Epitope Prediction

Kolaskar and Tongaonkar method was used for the predictions of known amino acid antigenicity. The maximum antigenicity propensity of spike protein B-Cell epitope prediction was 1.261 while minimum antigenicity was 0.866 and threshold value was set 1.041. For nucleoprotein maximum antigenicity propensity was 1.197, while minimum propensity was 0.874 and threshold value was 0.988 (Fig.1 & Fig. 2).

Average: 1.041 Minimum: 0.866 Maximum: 1.261

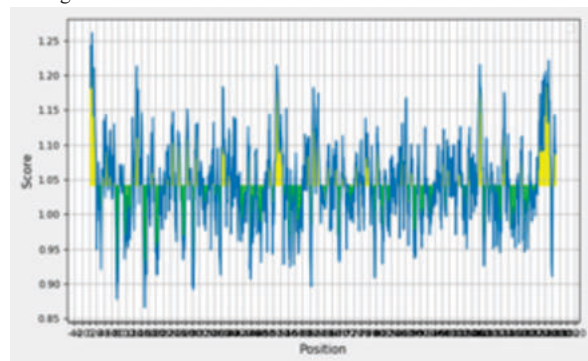


Fig.1. Antigenicity Score Of Sars Cov-2's Spike Protein Predicted By Kolaskar-tongaonkar Method For Detection Of B-cell Epitopes.

Average: 0.988 Minimum: 0.874 Maximum: 1.197

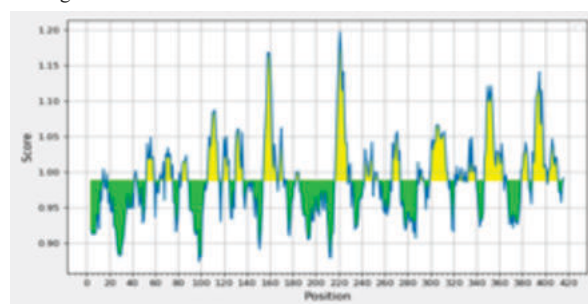


Fig.2. Antigenicity Score Of Sars Cov-2's Nucleoprotein Predicted By Kolaskar-tongaonkar Method For Detection Of B-cell Epitopes.

For further analysis of B-cell epitopes were separated based on antigenicity & allergenicity (Table 3, Table 4 & Table 8).

Table 3: B-cell Epitopes In Sars Cov-2's Spike Protein Predicted By Kolaskar-tongaonkar Method

SN	Posi	Residue	Start	End	Peptide	Score	Prediction
1	4	F	1	7	MFVFLVL	1.182	Epitope
2	5	L	2	8	FVFLVLL	1.243	Epitope
3	6	V	3	9	VFLVLLP	1.239	Epitope
4	7	L	4	10	FLVLLPL	1.22	Epitope
5	8	L	5	11	LVLLPLV	1.261	Epitope

Table 4: B-cell Epitopes In Sars Cov-2's Nucleoprotein Predicted By Kolaskar-tongaonkar Method

Sr. No	Position	Residue	Start	End	Peptide	Score	Prediction
1	156	A	153	159	NNAAIVL	1.066	Epitope
2	157	I	154	160	NAAIVLQ	1.101	Epitope
3	158	V	155	161	AAIVLQL	1.168	Epitope
4	159	L	156	162	AIVLQLP	1.168	Epitope
5	160	Q	157	163	IVLQLPQ	1.161	Epitope

Peptide sequence from 1 to 7 (spike glycoprotein) and 154 to 160 (nucleoprotein) have ability to initiate B-cell immune response.

T-cell Epitope Prediction

T-cell epitopes of SARS CoV2 spike glycoprotein & nucleoprotein were predicted with the help of CTLpred. T-cell epitopes plays crucial role in sub unit vaccine candidate design. T-cell epitope peptide score

equal to 1 or less than 1 is potential to become vaccine. Peptides are ranked as per score. O5 peptides from each protein are capable to become epitopes (Table 5 & Table 6).

Table 5: T-cell Epitopes In Sars Cov2 Spike Glycoprotein Predicted By CTLpred

Peptide Rank	Start Position	Sequence	Score	Prediction
1	318	FRVQPTESI	1.000	Epitope
2	448	NYNYLYRLF	1.000	Epitope
3	783	AQVKQIYKT	1.000	Epitope
4	28	YNSFTRGV	0.990	Epitope
5	236	TRFQTLLAL	0.990	Epitope

Table 6: T-cell Epitopes In Sars Cov2 Nucleoprotein Predicted By CTLpred

Peptide Rank	Start Position	Sequence	Score	Prediction
1	115	TGPEAGLPY	1.000	Epitope
2	222	LLDRLNQL	1.000	Epitope
3	339	LDDKDPNFK	1.000	Epitope
5	377	DETQALPQR	1.000	Epitope

The nucleoprotein peptide DETQALPQR which starting position was 377 has a score 1.00 and it is non allergen.

Evaluation Of Allergenicity

Allergenicity of SARS CoV2 entire spike glycoprotein & nucleoprotein as well as selected B-cell & T-cell epitope was predicted with the help of ALGPred. After allergenicity prediction It is found that both spike glycoprotein and nucleoprotein are Non-Allergen (Table 7).

Table 7: Allergenicity Of Sars Cov2 Entire Spike Glycoprotein & Nucleoprotein Predicted By Algpred 2.0.

Subject	ML Score	MERCI Score	BLAST Score	Hybrid Score	Prediction
sp P0DTC2 SPIKE_SARS2	0.21	0.5	-0.5	0.21	Non-Allergen
tr A0A6B9VLF5 A0A6B9VLF5_SARS2	0.11	0.5	-0.5	0.11	Non-Allergen

As well as out of 10 B-cell epitopes 02 epitopes MFVFLVL & NAAIVLQ are Non-Allergen as per ML & Hybrid score while 01 T-cell epitope DETQALPQR was found to be Non-Allergen. Rest of the selected B-Cell & T-Cell epitopes from both antigenic protein found to be allergen and it was omitted from further analysis (Table 8).

Table 8: Allergenicity Of Sars Cov2 Spike Glycoprotein & Nucleoprotein's Predicted B-cell & T-cell Epitope Predicted By Algpred 2.0.

Nature of Epitope	Epitope peptide	Antigenic Protein	ML Score	MERCI Score	BLAST Score	Hybrid Score	Prediction
B-cell	MFVFLVL	Spike Glycoprotein	0.26	0	0	0.26	Non-Allergen
B-Cell	FVFLVLL	Spike Glycoprotein	0.36	0	0	0.36	Allergen
B-Cell	VFLVLLP	Spike Glycoprotein	0.45	0	0	0.45	Allergen
B-Cell	FLVLLPL	Spike Glycoprotein	0.46	0	0	0.46	Allergen
B-Cell	LVLLPLV	Spike Glycoprotein	0.48	0	0	0.48	Allergen
B-cell	NNAIVL	Nucleoprotein	0.28	0	0	0.28	Allergen
B-Cell	NAAIVLQ	Nucleoprotein	0.29	0	0	0.29	Non-Allergen
B-Cell	AAIVLQL	Nucleoprotein	0.31	0	0	0.31	Allergen
B-Cell	AIVLQLP	Nucleoprotein	0.41	0	0	0.41	Allergen

B-Cell	IVLQLPQ	Nucleoprotein	0.46	0	0	0.46	Allergen
T-Cell	FRVQPTESI	Spike Glycoprotein	0.4	0	0	0.4	Allergen
T-Cell	NYNYLYRLF	Spike Glycoprotein	0.31	0	0	0.31	Allergen
T-Cell	AQVKQIYKT	Spike Glycoprotein	0.37	0	0	0.37	Allergen
T-Cell	YTNSFTRGV	Spike Glycoprotein	0.31	0	0	0.31	Allergen
T-Cell	TRFQTLLAL	Spike Glycoprotein	0.36	0	0	0.36	Allergen
T-Cell	TGPEAGLPY	Nucleoprotein	0.4	0	0	0.4	Allergen
T-Cell	LLDRLNQL	Nucleoprotein	0.33	0	0	0.33	Allergen
T-Cell	LDDKDPNFK	Nucleoprotein	0.33	0	0	0.33	Allergen
T-Cell	DETQALPQR	Nucleoprotein	0.29	0	0	0.29	Non-Allergen
T-Cell	TGPEAGLPY	Nucleoprotein	0.35	0	0	0.35	Allergen

Secondary Structure Prediction

GOR4 is an online tool used to predict secondary structure of 03 non-allergen B-cell & T-cell epitopes. B-Cell epitopes consists high percentage of extended strand than T-cell epitope. T-cell epitope DETQALPQR consists 77.78% Random coil (Table 9).

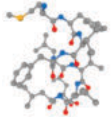
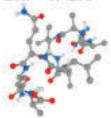
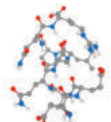
Table 9: Secondary Structure Prediction Of Selected Peptide By GOR4

Nature of Epitope	Epitope peptide	Alpha Helix	Extended strand	Random Coil
B-cell	MFVFLVL	0.00%	71.43%	28.57%
B-Cell	NAAIVLQ	0.00%	57.14%	42.86%
T-Cell	DETQALPQR	0.00%	22.22%	77.78%

3D Structure Prediction Of Epitope

3D structure of selected non-allergen 03 potential epitopes predicted by pepfold based on hidden markov model. The conformation of selected epitopes was refined by greedy algorithm & coarse-grained force field structures visualized by using discovery studio visualizer and further used for molecular docking between antibody & antigen (Table 10).

Table 10: 3D Structures Of Selected Epitopes

Nature of Epitope	Epitope peptide	Antigenic Protein	Prediction	Structure of Epitope peptide (In Balls & Sticks model)
B-cell	MFVFLVL	Spike Glycoprotein	Non-Allergen	
B-Cell	NAAIVLQ	Nucleoprotein	Non-Allergen	
T-Cell	DETQALPQR	Nucleoprotein	Non-Allergen	

Antibody Selection & Its 3D Structure Retrieval

The refined 3D structure of human IgG3 antibody was retrieved from Uniprot structure section. It was an AlphaFold predicted structure AF-P01860-F1 (1-377). This model was analyzed and is based on high

coverage value.

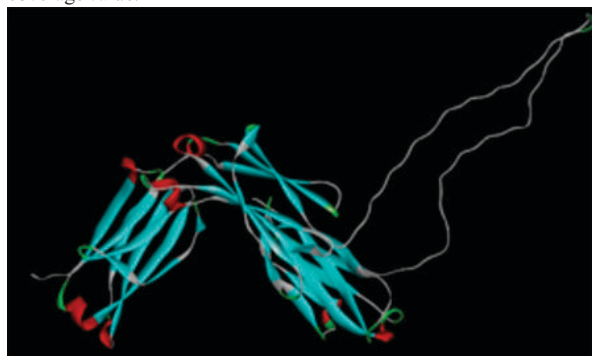


Fig. 3: 3d Structure Of Igg3 Antibody (Discovery Studio Visualizer)

Antigen-Antibody docking

Hpepdock is an online server used for molecular docking of antibody and antigen to estimate vaccine model interaction by ligand binding domain of immune receptor. 3D structures of 02 epitope peptides DETQALPQR & NAAIVLQ was selected for docking studies. The DETQALPQR peptide shows docking energy score -160.412 kcal/mol & NAAIVLQ shows -160.657 kcal/mol docking energy score (Fig.3). The peptide perfectly binds to the binding pose of antibody and it provokes desired immune response to prevent SARS CoV2 infections.

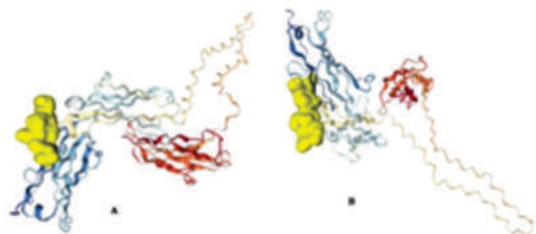


Fig. 4: B-cell Epitope Peptide Naaivlq (yellow) Of Nucleoprotein Perfectly Binds To Igg3 Antibody (a), While T-cell Epitope Peptide Detqalpqr (yellow) Of Nucleoprotein Perfectly Binds To Igg3 Antibody (b).

CONCLUSION

In the present study, several reverse vaccinology approaches were used to construct effective vaccine against SARS CoV2 infections. IEDB shows antigenic nature of protein to construct vaccine candidate. ALGPred suggests allergic and non-allergic nature of peptide vaccine candidates. HPEPDOCK shows a higher binding energy to confirm antigen antibody interaction. After various in silico trials our peptide vaccine shows good immune response, but in vivo and in vitro experiments are required for validation. The proposed vaccine met all the required parameters. Collectively, our designed vaccine construct is safer & effective; however preclinical validation must be performed before clinical trials.

REFERENCES

- [1] National Library of Medicine. Immune System and Disorders | Autoimmune Disease | MedlinePlus, medlineplus.gov, 2020 Aug 17, <https://medlineplus.gov/immunesystemanddisorders.html>.
- [2] Alberts, Bruce, et al. Helper T Cells and Lymphocyte Activation - Molecular Biology of the Cell - NCBI Bookshelf, www.ncbi.nlm.nih.gov, 2002 Jan.1, <https://www.ncbi.nlm.nih.gov/books/NBK26827/>.
- [3] <https://www.who.int/news-room/commentaries/detail/modes-of-transmission-of-virus-causing-covid-19-implications-for-ipc-precaution-recommendations>.
- [4] Zhou P, Yang XL, Wang XG, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin [published correction appears in Nature. 2020 Dec; 588(7836):E6]. Nature. 2020; 579(7798):270-273. doi:10.1038/s41586-020-2012-7
- [5] He Y, Lu H, Siddiqui P, Zhou Y, Jiang S. Receptor-binding domain of severe acute respiratory syndrome coronavirus spike protein contains multiple conformation-dependent epitopes that induce highly potent neutralizing antibodies. J Immunol. 2005; 174(8):4908-4915. doi:10.4049/jimmunol.174.8.4908.
- [6] Saif LJ. Coronavirus immunogens. Vet Microbiol. 1993;37(3-4):285-297. doi:10.1016/0378-1135(93)90030-b
- [7] Amanat F, Krammer F. SARS-CoV-2 Vaccines: Status Report. Immunity. 2020;52(4):583-589. doi:10.1016/j.immuni.2020.03.007
- [8] Polack FP, Thomas SJ, Kitchin N, et al. Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. N Engl J Med. 2020;383(27):2603-2615. <https://www.cdc.gov/coronavirus/2019-ncov/vaccines/different-vaccines.html>. Updated Dec. 28, 2020
- [10] Gershoni JM, Roitburd-Berman A, Siman-Tov DD, Tarnovitski Freund N, Weiss Y. Epitope mapping: the first step in developing epitope-based vaccines. BioDrugs. 2007;21(3):145-156. doi:10.2165/00063030-200721030-00002
- [11] Iurescia S, Fioretti D, Fazio VM, Rinaldi M. Epitope-driven DNA vaccine design

- employing immunoinformatics against B-cell lymphoma: a biotech's challenge. Biotechnol Adv. 2012;30(1):372-383. doi:10.1016/j.biotechadv.2011.06.020
- [12] Abu-haraz AH, Abd-elrahman KA, Ibrahim MS, Hussien WH, Mohammed MS, et al. Multi epitope peptide vaccine prediction against Sudan Ebola virus using immunoinformatics approaches. Adv Tech Biol Med. 2017;5:203.
 - [13] Wu F, Zhao S, Yu B, et al. A new coronavirus associated with human respiratory disease in China [published correction appears in Nature. 2020 Apr;580(7803):E7]. Nature. 2020;579(7798):265-269. doi:10.1038/s41586-020-2008-3
 - [14] Zhu N, Zhang D, Wang W, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. N Engl J Med. 2020;382(8):727-733. doi:10.1056/NEJMoa2001017
 - [15] Kolaskar AS, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Lett. 1990 Dec 10;276(1-2):172-4.
 - [16] Shen Y, Maupetit J, Derreumaux P, Tufféry P. Improved PEP-FOLD approach for peptide and miniprotein structure prediction J. Chem. Theor. Comput. 2014; 10:4745-4758
 - [17] Iles, J. K., Zmuidinaite, R., Sadee, C., Gardiner, A., Lacey, J., Harding, S., Wallis, G., Patel, R., Roblett, D., Heeney, J., Baxendale, H., & Iles, R. K. (2022). Determination of IgG1 and IgG3 SARS-CoV-2 Spike Protein and Nucleocapsid Binding-Who Is Binding Who and Why?. International journal of molecular sciences, 23(11), 6050. <https://doi.org/10.3390/ijms23116050>.