Research article

The identification of synthetic peptide vaccine candidate against *SARS-CoV-2/COVID-19* through reverse vaccinology approach

S. Pushkala[1], Sudha Seshayyan[2], Tammanna Bhajantri[3]

[1]Professor,[3]Research Scholar, Department of Immunology, [2]The Vice Chancellor, The Tamil Nadu Dr MGR Medical University,69, Anna Salai Road, Guindy, Chennai,600032, Tamil Nadu, India

Corresponding author: **Tammanna Bhajantri.** Email: tm_1007@hotmail.com

ABSTRACT

Introduction and Aim: In emerging respiratory disease pathogens, the Corona viruses have become the main pathogens of respiratory viral disease outbreaks. SARS-Cov-2 is a new virus that has been identified in human and this very contagious novel corona virus has spread globally within the short period of time. The biological concept of the synthetic peptide vaccine is based on the induction of immune cells response depends on the immune cell receptor specificity to verify a presented peptide epitope. The identification of these epitopes by experimental procedures are expensive and time- consuming. Therefore, the approach of reverse vaccinology came into view. The approach of reverse vaccinology involves molecular docking, prediction of epitope tools, and desired immunogenic peptides analysis of population coverage in terms of design. The primary goal of this present study is to identify the antigenic determinant which might be a potent candidate vaccine against *SARS-CoV-2*.

Materials and Methods: The whole genome sequence of a contagious strain of *SARS-CoV-2* retrieved from genomic database. The whole genome screened to identify the protein sequence which is antigenic, and the antigen determinant peptide predicted with different databases and Accessible Surface Area (ASA) calculation. The selection of peptide depends on the prediction of identified epitope carried out according to their predictive scores by almost all bioinformatic tools.

Results: The identified antigenic determinant predicted to bind with MHC class I molecule using MHC binding prediction tools. In this study, the identified epitope is the best peptide having greater ASA value and binding with MHC class I molecule.

Conclusion: As this peptide is immunogenic epitope it might be a potent candidate vaccine against *COVID-19 or SARS-CoV-2* virus.

Keywords: Reverse vaccinology**;** synthetic peptide vaccine; *SARS-CoV-2; COVID-19;* epitope.

INTRODUCTION

Coronavirus disease 2019, which is the acronym of COVID-19, which has spread initially from China to several other countries around the world which is caused by *SARS-CoV-2* is a new virus accountable for a scourge of respiratory disease (1). In humans, the virus's epitope binds to the angiotensin-converting enzyme 2 (ACE2) receptor, which is found in abundance in the kidney, endothelium, lung, and heart. On December 31, 2019, the WHO Office in China received the first reports of outbreak cases with symptoms consistent with unidentified low respiratory infections found in Wuhan, the largest city in China's Hubei province (2).

There is no effective treatment available to treat COVID-19, which is caused by a new virus SARS-Cov-2 that has recently appeared in humans and in terms of therapeutics there is no known effective pharmaceutical agent to avoid infection from the virus SARS-Cov-2(3). A 6- to 18-month timetable was set for achieving the target of a safe and effective vaccine. As a result of this unprecedented effort, over 200 candidate vaccines are in various stages of growth, with over 50 in human clinical trials and 18 in efficacy testing (4).

Conventional vaccinology approach was developed in 1796 by Edward Jenner (5) and in 1880 the work of Louis Pasteur on chicken cholera opened the way to Conventional vaccinology approach in the laboratory (6). Since the time of vaccine developed by Louis Pasteur there have been 2 methods of vaccine development in Conventional vaccinology approach: attenuation and inactivation (7, 8).

The major drawback of this method; the outcome of an attempted attenuation unpredictable and based on the nature of the attenuating mutation attenuated pathogen may revert to virulence (9,10). Other disadvantages of this strategy include the difficulty in producing adequate titer preparations, their high cost per dose, and existing criteria for multiple vaccinations (11).

Progress in microbiology, genomics, synthetic biology, and biotechnology has provided a novel set of tools to approach modern-day vaccinology (12,

13). Since identifying epitopes experimentally is time-consuming and expensive, the reverse vaccinology approach was developed. One of the most important steps in vaccine development is antigenic epitope prediction using computational methods (14).

Pathogens have been dissected and the components useful for vaccine production identified using biochemical, serological, and microbiological approaches. While effective in many cases, this method is time-consuming and fails when pathogens cannot be cultured in vitro or when the most abundant antigens are sequenced in a variable manner. Without the need to grow the pathogen in vitro, Insilco genomic approaches now enable prediction of all antigens, regardless of their abundance or immunogenicity during infection, allowing vaccine production using non-conventional antigenic epitopes and exploiting non-conventional arms of the system. This novel process is named as Reverse Vaccinology since the method of vaccine discovery begins *in silico* using genetic information rather than the pathogen itself (15).

Reverse vaccinology employs epitope prediction techniques, molecular docking, and population coverage analysis to create desired immunogenic peptides. The use of bioinformatics techniques aids computational biology in the design of in silico vaccines. Prediction of the epitope is critical, since it reduces both the cost and the need for experimental results (16,17). We proposed the present study to identify the antigenic determinant that could be used as a potent vaccine candidate against the SARS-Cov-2 Virus.

MATERIALS AND METHODS

Selection and retrieval of sars-cov-2 genome

To identify the SARS-CoV-2 virus pathogenic protein sequence for their antigenic properties the bioinformatics tools are used. The publicly available comprehensive database GenBank contains amino-acid sequences for almost 260 000 formally described species (18). The complete genomic sequence of the virus SARS-CoV-2 was selected and isolated from GenBank (https://www.ncbi.nlm.nih.gov) with an Accession number MT012098.1 and the whole proteomic sequence of SARS-CoV-2 retrieved in the FASTA format.

Prediction of antigenic determinants

1. *Emboss antigenic*

The Antigenic predicts potentially antigenic regions of a protein sequence using Kolaskar and Tongaonkar's strategy. The hydrophobic residues Cys, Leu, and Val, if they occur on the surface of a protein, are more likely to be a part of antigenic sites, according to data analytically defined antigenic sites on proteins (19). the complete nucleotide sequences are retrieved using EMBOSS antigenic tool and located that the epitopes which having Accessible surface area (ASA) Value of over 30% (shown in Table. 2) are considered as antigenic epitopes using the mathematical formula. Accessibility surface area (ASA) represents the component of the peptide (in percentage) that is Accessible on the surface of the globular protein structure. for every antigenic peptide, ASA was Calculated by the formula.

ASA = [(L+C+V)/ Total no. of residues within the peptide] X 100

where,

L= number of Lysine residues in peptide

C= number of Cysteine residues inpeptide

V= number of Valine residues inpeptide

2. *Immunomedicine group*

This method helps to predict those determinants from within a gene sequence that are likely to be antigenic by eliciting a cell mediated immune response. Antigenic determinants are determined using the strategy of Kolaskar and Tongaonkar as this method gives 75% accuracy as compared to other available methods. These Predictions are supported a table by using IMMUNOMEDICINE group and, and that reflects the presence of amino acid residues in segmental epitopes that have been studied experimentally (19). The complete nucleotide sequences are retrieved one more time using IMMUNOMEDICINEgroup tool and located that the epitopes which having Accessible surface area (ASA) Value of over 30% are considered as antigenic epitopes using the formula. The retrieved nucleotide sequences which are antigenic epitopes having Accessible surface area (ASA) Value of over 30% are compared with the antigenic epitopes retrieved from EMBOSS antigenic tool (Table. 2).

3. *Prediction of MHC binding*

3.1 MHCPred is an additive approach for predicting MHC class I binding affinity, and allele-specific Quantitative Structure Activity Relationship (QSAR) models were created using partial least squares (PLS)(20-22). The MHCPred bioinformatic tool is used in this study to predict the possible antigenic epitopes presented on the cell surface by MHC I molecules to CD8 positive T lymphocytes using the MHCPred bioinformatic tool. It combines existing proteasomal cleavage predictions with peptide anchoring to MHC I molecules and the MHCPred results shown in table 3.

3.2 nHLAPred bioinformatic tool is employed to predict MHC I binders and are filtered to potential CTL epitopes by refining through Proteasomal matrices (23). The entire protein sequence of SARS-CoV-2 was retrieved using nHLAPred, which allowed all predicted binders for unique MHC alleles to be displayed in a single line by simply colouring the predicted binders.

3.3 RANKPEP -The conserved epitopes predicted using the RANKPEP web server, which has a variability masking function, as a predictor of MHC-peptide binding from a collection of aligned peptides known to bind to a given MHC molecule (24). The complete Protein sequence SARS-CoV-2 analyzed with the help of RANKPEP web server to find out individually binding epitopes to a particular MHC class I molecule based on the predetermined threshold value.

3.4 VAXIGN is a vaccine design software that used microbial genomic sequence as input data and epitope binding to MHC class I molecule to predict potential vaccine targets based on various vaccine design criteria (25). The complete proteomic sequence of SARS-Cov-2 analyzed using Vaxign. In Vaxign-Vaxitop method used to predict SARS-Cov-2peptide candidates for vaccine development.

RESULTS

Identification of epitope (peptide)

The Accessible Surface area (ASA) value of the EMBOSS antigenic determinants were calculated using mathematical formula. The antigenic determinants with more than 30 % ASA value are selected as epitopes and we found 27 epitopes having ASA value more than 30% (Table 1). The EMBOSS results are compared with the IMMUNOMEDICINE group results to find out common epitopes. The antigenic determinant with 50 % highest ASA value considered as a potential epitope (Table. 2).

ASA= (L+V+C)/Total number of residues X 100

Table 1: EMBOSS results

#Sequence	Score	Max_score_pos	ASA Value
GCTACCCTCT	1.219 lengths 10 at residues 836->845	840	50%

Table 2: 10 Epitopes with more than 30 % ASA Value and sequence number 4 is a potential Epitope (peptide) with 50 % ASA value.

Sequence Number	EMBOSS Antigenic Determinant	IMMUNOMEDICINE Antigenic Determinant	Score Emboss	Accessible Surface area (ASA) Emboss/ Immunomed (%)
1	GTGCCACTAC	GGTGCCACTA	1.241	40
2	GCTCGAACTGCACCTCA	TGCTCGAACTGCACCTC	1.241	41.17
3	GTCCTTGTCCCTCA	TGTCCTTGTCCCTC	1.219	42.85
4	**GCTACCCTCT**	**GGCTACCCTC**	**1.219**	**50**
5	CATACACTCGCT	GCATACACTCGC	1.214	41.66
6	CGGCCCAAA	TCGGCCCAA	1.209	44.44
7	TACACCAG	GTACACCA	1.191	37.5
8	CAGCCGATCATCAGCACATCT	GCAGCCGATCATCAGCACATC	1.186	38
9	CGTCCGG	TCGTCCG	1.120	42.85
10	CTGCTCG	GCTGCTC	1.115	42.85

Table 3: MHCPredresult shows that GCTACCCTC epitope (peptide) binds with MHC Class I allele HLA-A*0201 and '- 'represents non-binders.

Amino acid groups	Predicted - logIC_{50} (M)	Predicted IC_{50} Value (nM)	Confidence of prediction (Max = 1)
ACATCTAGG	7.026	94.19	0.78
CCTTCTAGC	6.75	177.83	0.89
GCTGGTAGC	6.691	203.70	0.89
GCATGCTTA	6.423	377.57	0.89
AAAGATGGC	6.344	452.90	1.00
ATGTGTTCA	6.277	528.45	1.00
GCTGGTGGC	6.199	632.41	0.89
CCTGATGGC	6.134	734.51	0.89
CATTAAAGA	6.059	872.97	1.00
AAATACCAG	6.014	968.28	0.89
ACGGTAATA	5.966	1081.43	0.89
GTTGACAGG	5.806	1563.15	0.89
AACTTTCGA	5.804	1570.36	1.00
CAGTACGGT	5.803	1573.98	1.00
GCAAGGTTC	5.803	1573.98	0.89
TCATGTTAT	5.802	1577.61	0.89
AGATGGAGA	5.802	1577.61	1.00

CTTTGTCCG	5.802	1577.61	0.89
AAAGTCATT	5.801	1581.25	1.00
AACTCGAAG	5.801	1581.25	0.89
TTATGAAGA	5.697	2009.09	1.00
TCATTTGAC	5.507	3111.72	0.89
TACTGTCGT	5.397	4008.67	1.00
GCTACCCTC	**5.324**	**4742.42**	**0.89**
CACGTCAAC	5.303	4977.37	1.00
ATGGTTGAG	-	-	-
TAAGAACGG	-	-	-
TTGTCCCTG	-	-	-

The results are shown in Table 3. with three columns. The peptide sequences shown in the first column, the predicted IC_{50} and IC_{50} values shown in the second and the third column respectively. The peptide sequence is sorted according to their IC_{50} values. Non-binders are listed at the bottom of the table, and peptides with lower IC_{50} values (or higher expected IC_{50} values) are listed first. If the IC_{50} value is greater than 5000, the peptide will not bind to MHC molecules.

Table 4: nHLA Pred result shows that predicted GCTACCCTC epitope (peptide) is a potential MHCbinder which binds with MHC Class I alleles HLA-A*0202 and HLA-A*0203

HLA-A*0202 [ANNs]	ACCTTCCCAGGTAACAAACCAACCAACTTTCGATCTCTTGTAGATCTGTTCTCTAAACGAACTTTAAAAT CTGTGTGGCTGTCACTCGGCTGCATGCTTAGTGCACTCACGCAGTATAATTAATAACTAATTACTGTCGT TGACAGGACACGAGTAACTCGTCTATCTTCTGCAGGCTGCTTACGGTTTCGTCCGTGTTGCAGCCGATCA TCAGCACATCTAGGTTTCGTCCGGGTGTGACCGAAAGGTAAGATGGAGAGCCTTGTCCCTGGTTTCAACG AGAAAACACACGTCCAACTCAGTTTGCCTGTTTTACAGGTTCGCGACGTGCTCGTACGTGGCTTTGGAGA CTCCGTGGAGGAGGTCTTATCAGAGGCACGTCAACATCTTAAAGATGGCACTTGTGGCTTAGTAGAAGTT GAAAAAGGCGTTTTGCCTCAACTTGAACAGCCCTATGTGTTCATCAAACGTTCGGATGCTCGAACTGCAC CTCATGGTCATGTTATGGTTGAGCTGGTAGCAGAACTCGAAGGCATTCAGTACGGTCGTAGTGGTGAGAC ACTTGGTGTCCTTGTCCCTCATGTGGGCGAAATACCAGTGGCTTACCGCAAGGTTCTTCTTCGTAAGAAC GGTAATAAAGGAGCTGGTGGCCATAGTTACGGCGCCGATCTAAAGTCATTTGACTTAGGCGACGAGCTTG GCACTGATCCTTATGAAGATTTTCAAGAAAACTGGAACACTAAACATAGCAGTGGTGTTACCCGTGAACT CATGCGTGAGCTTAACGGAGGGGCATACACTCGCTATGTCGATAACAACTTCTGTGGCCCTGATGGCTAC CCTCTTGAGTGCATTAAAGACCTTCTAGCACGTGCTGGTAAAGCTTCATGCACTTTGTCCGAACAACTGG ACTTTATTGACACTAAGAGGGGTGTATACTGCTGCCGTGAACATGAGCATGAAATTGCTTGGTACACGGA ACGTTCTGAAAAGAGCTATGAATTGCAGACACCTTTTGAAATTAAATTGGCAAAGAAATTTGACACCTTC AATGGGGAATGTCCAAATTTTGTATTTCCCTTAAATTCCATAATCAAGACTATTCAACCAAGGGTTGAAA
HLA-A*0203	ACCTTCCCAGGTAACAAACCAACCAACTTTCGATCTCTTGTAGATCTGTTCTCTAAACGAACTTTAAAAT CTGTGTGGCTGTCACTCGGCTGCATGCTTAGTGCACTCACGCAGTATAATTAATAACTAATTACTGTCGT TGACAGGACACGAGTAACTCGTCTATCTTCTGCAGGCTGCTTACGGTTTCGTCCGTGTTGCAGCCGATCA TCAGCACATCTAGGTTTCGTCCGGGTGTGACCGAAAGGTAAGATGGAGAGCCTTGTCCCTGGTTTCAACG AGAAAACACACGTCCAACTCAGTTTGCCTGTTTTACAGGTTCGCGACGTGCTCGTACGTGGCTTTGGAGA CTCCGTGGAGGAGGTCTTATCAGAGGCACGTCAACATCTTAAAGATGGCACTTGTGGCTTAGTAGAAGTT GAAAAAGGCGTTTTGCCTCAACTTGAACAGCCCTATGTGTTCATCAAACGTTCGGATGCTCGAACTGCAC CTCATGGTCATGTTATGGTTGAGCTGGTAGCAGAACTCGAAGGCATTCAGTACGGTCGTAGTGGTGAGAC ACTTGGTGTCCTTGTCCCTCATGTGGGCGAAATACCAGTGGCTTACCGCAAGGTTCTTCTTCGTAAGAAC GGTAATAAAGGAGCTGGTGGCCATAGTTACGGCGCCGATCTAAAGTCATTTGACTTAGGCGACGAGCTTG GCACTGATCCTTATGAAGATTTTCAAGAAAACTGGAACACTAAACATAGCAGTGGTGTTACCCGTGAACT CATGCGTGAGCTTAACGGAGGGGCATACACTCGCTATGTCGATAACAACTTCTGTGGCCCTGATGGCTAC CCTCTTGAGTGCATTAAAGACCTTCTAGCACGTGCTGGTAAAGCTTCATGCACTTTGTCCGAACAACTGG ACTTTATTGACACTAAGAGGGGTGTATACTGCTGCCGTGAACATGAGCATGAAATTGCTTGGTACACGGA ACGTTCTGAAAAGAGCTATGAATTGCAGACACCTTTTGAAATTAAATTGGCAAAGAAATTTGACACCTTC AATGGGGAATGTCCAAATTTTGTATTTCCCTTAAATTCCATAATCAAGACTATTCAACCAAGGGTTGAAA

Table 4 shows the nHLAPred result, which displays all predicted binders for unique MHC alleles in a single line by simply colouring the expected or predicted binders. Every predicted binder's starting residue is shown in red, while the remaining residues are shown in blue. The option is very helpful in detecting the promiscuous MHC binder in the series.

Table 5: RANKPEP Result: Epitope (Peptide) highlighted in colour represents predicted binder to MHC class I allele HLA-A*0207

RANK	POS.	N	SEQUENCE	C	MW (Da)	SCORE
1	808	TAT	GTCGATAAC	AAC	735.82	10.756
2	2085	GCA	GTGGCTAAC	TAA	721.79	6.365
3	1428	TGG	CTTGAAAAC	CAT	749.85	4.34
4	1603	CTT	CTTGAAATA	CTC	747.81	4.054
5	**833**	**CTG**	ATG**GCTACC**	**CTC**	767.88	2.681

The peptides predicted out of a protein sequence according to absolute scores and specify peptides as binders or non-binders based on predetermined thresholds were the product of the RANKPEP.The peptides predicted out of a protein sequence according to absolute scores and specify peptides as binders or non-binders based on predetermined thresholds were the product of the RANKPEP.

The peptides predicted out of a protein sequence according to absolute scores and specify peptides as binders or non-binders based on predetermined thresholds were the result of the RANKPEP.

Table 6: VAXIGN RESULT : the vaccine epitope GCTACCCTCT binds with MHC CLASS I Allele HLA-A*0202

MHC I Binding Prediction							
Index	**Epitope**	**Epitope Length**	**MHC Allele**	**P value**	**Matching from**	**Matching to**	**Location**
1	ACAGCCCTA	9	HLA-A*02:02	0.0132	447	455	
2	ATACTCCAA	9	HLA-A*02:02	0.0173	1609	1617	
3	CTAAACTTA	9	HLA-A*02:02	0.0223	2345	2353	
4	**CTACCCTCT**	**9**	**HLA-A*02:02**	**0.0235**	**837**	**845**	
5	ACAACCATT	9	HLA-A*02:02	0.0258	2541	2549	
6	CTTCCCACA	9	HLA-A*02:02	0.0296	2491	2499	
7	AAAGCCCCA	9	HLA-A*02:02	0.0296	2452	2460	
8	ATCGCCATT	9	HLA-A*02:02	0.0354	1669	1677	

The Vaxign result shown in Table 6 using Vaxitop epitope prediction method show that selected epitope (peptide)after the initial screening predicted to bind with MHC class I allele HLA-A*02:02, and it relies on statistical P-value.We identified 8 peptide candidates in which CTACCCTCT epitope is a potential peptide candidate for SARS-Cov-2 vaccine development.

Epitope (peptide) binding with MHC class I molecule

From the result of EMBOSS antigenic and IMMUNOMEDICINE group the selected epitope GCTACCCTCT binds with MHC Class I molecule and it is confirmed with MHC binding tools such as MHCPrednHLAPred, RANKPEP and web server vaccine design program VAXIGN.

DISCUSSION

Corona Virus disease 2019 is an acronym for COVID-19 is a communicable disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) novel virus emerged in humans in December 2019 from Wuhan, China (26). COVID-19 has a wide range of clinical manifestations, ranging from asymptomatic and mild flu-like symptoms to acute respiratory distress syndrome and death. In COVID-19 cases, long-term neurological, pulmonary, and cardiological complications have also been identified (27). SARS-CoV-2 vaccines of various types have been produced and tested in preclinical studies. Although only a few of them advanced to clinical trials, the creation and mass production of effective vaccines is a costly and time-consuming operation. To reduce the time needed to produce a potent SARS-CoV-2 vaccine candidate, researchers used immuno-informatics and reverse vaccinology methods to classify and design a peptide vaccine. In recent time, there are many well defined bioinformatics approaches available to design a successful new generation vaccine which are safe in humans (28). To develop a vaccine in short time of period due to the pandemic situation like a COVID-19, the new technologies like computational immunology and immuno-informatic tools are used to develop the vaccine candidates or vaccine by understanding the human immune response against a pathogen (29).

The epitope's proper binding to the MHC I antigen-binding cleft is needed for the induction of the desired immune response (30).

The main aim of this research is to find an antigenic epitope that could be used as a vaccine candidate against the SARS-Cov-2 Virus. The complete genome sequence of a contagious SARS-CoV-2 strain was obtained in FASTA format from the GenBank database. The whole genome screened to identify the protein sequence which is antigenic, and the antigen determinant peptide predicted with different databases like EMBOSS antigenic, IMMUNOMEDICINE group based on the Accessible Surface area (ASA) calculation. The peptide is chosen only if an epitope has been found. Almost all bioinformatic systems forecast it based on their predictive ratings. The identified antigenic determinant predicted to bind with MHC class I molecule using MHC binding prediction

tools like MHCPred, nHLAPred, RANKPEP and web-based vaccine design program VAXIGN.

CONCLUSION

We proposed in this study to identify an antigenic determinant that could be used as a possible vaccine candidate against the SARS-Cov-2 Virus. An infectious SARS-CoV-2 strain's complete genome sequence was obtained from the GenBank database. The analysis from this study provides information about antigenic determinants and their ASA value analysis in the predicted model. We identified the protein sequence which is potential epitope predicted with different databases and Accessible Surface area (ASA) value. The identified epitope is predicted to bind with MHC class I molecule using MHC binding prediction tools.

On the basis of the findings, it could be concluded that from the whole analysis, the *in silico* identified epitope **GCTACCCTCT** is the potent vaccine candidate against *SARS-CoV-2* virus or *COVID-19* and further pre-clinical study is needed to validate the results.

ACKNOWLEDGEMENT

First of all, my thanks to almighty and I sincerely thank Dr. Sudha Seshyyan, Vice Chancellor, The Tamil Nadu Dr MGR Medical University, Chennai, who was the main person to spear head this study.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

1. Hu, L., Chen, S., Fu, Y., Gao, Z., Long, H., Ren, H. W., *et al.,* Risk factors associated with clinical outcomes in 323 Coronavirus disease 2019 (COVID-19) hospitalized patients in Wuhan, China. Clinical Infectious Diseases. 2020; 71(16): 2089-2098. https://doi.org/10.1093/cid/ciaa539
2. Zhou, P., Yang, X. L., Wang, X. G., Hu, B., Zhang, L., Zhang, W., *et al.,* A pneumonia outbreak associated with a new coronavirus of probable bat origin. nature. 2020 Mar; 579(7798): 270-3. https://doi.org/10. 1038/s41586-020-2012-7.
3. Fisher D, Heymann D. Q&A: The novel coronavirus outbreak causing COVID-19. BMC medicine. 2020 Dec; 18(1): 1-3. https://doi.org/10.1186/s12916-020- 01533-w.
4. Kim, J. H., Marks, F., Clemens, J. D. Looking beyond COVID-19 vaccine phase 3 trials. *Nat Med.* 2021; 27: 205-211. https://doi.org/10.1038/s41591-021-01230-y.
5. Bragazzi, N. L., Gianfredi, V., Villarini, M., Rosselli, R., Nasr, A., Hussein, A., *et al.,* Vaccines meet big data: State-of-the-Art and future prospects. From the classical 3Is ("Isolate–Inactivate–Inject") vaccinology 1.0 to vaccinology 3.0, vaccinomics, and beyond: a historical overview. Frontiers in public health. 2018 Mar 5; 6: 62. DOI: 10.3389/fpubh.2018. 00062.
6. Plotkin, S. A. Vaccines, vaccination, and vaccinology. The Journal of Infectious Diseases. 2003; 187(9): 1349–1359. https://doi.org/10.1086/374419
7. Plotkin, S. L., Plotkin, S. A., Orenstein, W. O. The history of vaccination Vaccines. 3rd*ed*, 1999 Philadelphia WB Saunders (pg.1-12).
8. Smith, K. A. Louis pasteur, the father of immunology? Frontiers in Immunology. 2012; 3: 68. DOI: 10.3389/fimmu.2012.00068.
9. Badgett, M. R., Auer, A., Carmichael, L. E., Parrish, C. R., Bull J. J. Evolutionary dynamics of viral attenuation. Journal of Virology. 2002; 76(20): 10524-10529. doi:10.1128/JVI. 76.20.10524-10529.2002.
10. Pulendran, B., Ahmed, R. Immunological mechanisms of vaccination. Nat Immunol. 2011; 12(6): 509-517. doi: 10.1038/ni.2039.
11. Barrett, A. D. T., Stanberry, L. R. Vaccines for biodefense and emerging and neglected diseases. Elsevier Inc., 2009. Pages 285, 287-324.
12. Nabel, G. J. Designing tomorrow's vaccines. N Engl J Med. 2013; 368(6): 551-560. doi: 10.1056/NEJMra1204186.
13. Votter, V., Denizer, G., Friedland, L. R., Krishnan, J., Shapiro, M. Understanding modern-day vaccines: what you need to know. Annals of Medicine. 2018; 50(2): 110-120. DOI: 10.1080/07853890.2017.1407035
14. Dubey, K. K., Luke, G. A., Knox, C., Kumar, P., Pletschke B. I., Singh, P. K., *et al.,* Vaccine and antibody production in plants: developments and computational tools. Briefings in Functional Genomics. 2018; 17(5): 295-307. https://doi.org/10.1093/bfgp/ely020
15. Rappuoli, R. Reverse vaccinology. CurrOpinMicrobiol. 2000; 3(5): 445-450. doi: 10.1016/s1369-5274(00)00119-3.
16. Florea, L., Halldorsson, B., Kohlbacher, O., Schwartz, R., Hoffman, S., Istrail, S., *et al.,* Epitope prediction algorithms for peptide-based vaccine design. Proc IEEE Comput Soc Bioinform Conf 2. 2003; 17-26.
17. María, R. R., Arturo, C. J., Alicia, J. A., Paulina, M. G., Gerardo, A. O. The impact of bioinformatics on vaccine design and development. Vaccines. 2017 Sep 6; 2: 3-6.
18. Benson, D. A., Cavanaugh, M., Clark, K., Ilene K. M., Lipman, D. J., Ostell, J., *et al.,* GenBank Nucleic Acids Res 41 (D1). D36–D42. 2013. https://doi.org/10.1093/nar/gks1195
19. Kolaskar, A. S., Tongaonkar, P. C. A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Letters 1990; 276(1-2): 172-174. doi: 10.1016/0014-5793(90)80535-q.
20. Guan, P., Doytchinova, I. A., Zygouri, C., Flower, D. R. MHCPred: bringing a quantitative dimension to the online prediction of MHC binding. Applied Bioinformatics. 2003; 2(1): 63-66.
21. Guan, P., Doytchinova, I. A., Zygouri, C., Flower, D. R. MHCPred: A server for quantitative prediction of peptide-MHC binding. Nucleic Acids Res. 2003; 31(13): 3621-3624. doi:10.1093/nar/gkg510.
22. Hattotuwagama, C. K., Guan, P., Doytchinova, I. A., Zygouri, C., Flower, D. R. Quantitative online prediction of peptide binding to the major histocompatibility complex. J Mol Graph Model. 2004; 22(3): 195-207. doi: 10.1016/S1093 -3263(03)00160-8.
23. Bhasin, M., Raghava, G. P. A hybrid approach for predicting promiscuous MHC class I restricted T cell epitopes. J Biosci. 2007; 32(1): 31-42. doi: 10.1007/s12038-007-0004-5.
24. Reche, P. A., Glutting, J. P., Zhang, H., Reinherz, E. L. Enhancement to the RANKPEP resource for the prediction of peptide binding to MHC molecules using profiles. *Immunogenetics*. 2004; 56(6): 405-419. doi: 10.1007/s00251-004-0709-7.
25. Xiang, Z., He, Y. Vaxign: a web-based vaccine target design program for reverse vaccinology. Procedia in Vaccinology. 2009*;* 1(1): 23-29.
26. Han, H. J., Nwagwu. C., Anyim, O., Ekweremadu, C., Kim, S. COVID-19 and cancer: From basic mechanisms to vaccine development using nanotechnology. Int Immunopharmacol. 2021; 90: 107247. doi: 10.1016/j.intimp.2020.107247.
27. Li, Y. D., Chi, W. Y., Su, J. H., Ferrall, L., Hung, C. F., Wu, T. C., *et al.,* Corona virus vaccine development: from SARS and MERS to COVID-19. J Biomed Sci. 2020; 27: 104 https://doi.org/10.1186/s12929-020-00695-2
28. Seib, K. L., Zhao, X., Rappuoli, R. Developing vaccines in the era of genomics: a decade of reverse vaccinology. Clin

Microbiol Infect. 2012; 18(5): 109-116. doi: 10.1111/j.1469-0691.2012.03939.x.
29. De Groot, A. S., Rappuoli, R. Genome-derived vaccines. Expert Rev Vaccines. 2004; 3(1): 59-76. doi: 10.1586 /14760584.3.1.59.
30. Stern, L. J., Wiley, D. C. Antigenic peptide binding by class I and class II histocompatibility proteins. Structure. 1994; 2(4): 245-251. doi: 10.1016/s0969-2126(00)00026-5.