




Check for updates

RESEARCH ARTICLE

Design of SARS-CoV-2 protein S peptides recognized by the most frequent HLA alleles in the Moroccan population using an immunoinformatics approach

[version 1; peer review: 1 not approved]

Meryem Fakhkhari ¹, Bouabid Badaoui², Hicham Oumzil³, Khalid Sadki ¹¹Research Laboratory in Oral Biology and Biotechnology, Faculty of Dentistry, Mohammed V University in Rabat, Rabat, 10112, Morocco²Department of Biology, Faculty of Sciences, Mohammed V University in Rabat, Rabat, 10106, Morocco³Medical Biotechnology Laboratory, Faculty of Medicine and Pharmacy, Mohammed V University in Rabat, Rabat, 10112, Morocco

V1 First published: 20 May 2024, 13:514
<https://doi.org/10.12688/f1000research.150049.1>
Latest published: 20 May 2024, 13:514
<https://doi.org/10.12688/f1000research.150049.1>

Abstract

Background

The coronavirus disease 2019 (COVID-19) is an infectious disease, caused by the new coronavirus known as Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), and exhibits diverse clinical outcomes and symptoms in infected individuals, emphasizing the need to investigate how human genetic diversity influences the virus's impact. This study aims to employ in silico methods to identify epitopes capable of eliciting an immune response, focusing on the most prevalent HLA-I and HLA-II alleles in the Moroccan population.

Methods

Our research consisted in predicting peptide-binding affinities between the most prevalent HLA Class I and Class II alleles in the Moroccan population and SARS-CoV-2 spike glycoprotein (S protein) peptides of variants isolated from strains of Moroccan patients. We performed the same analyses for SARS-CoV-2 wild type S protein to assess the ability of these HLA alleles to interact with peptides in the presence or absence of SARS-CoV-2 mutations.

Results

In a broader sense, 12 distinct HLA Class I and Class II alleles in the Moroccan population have been identified as possibly interacting with

Open Peer Review

Approval Status 


1

version 1

20 May 2024



view

1. **Gustavo Fioravanti Vieira** , Federal University of Rio Grande do Sul, Universidade La Salle, Porto Alegre, Canoas, Brazil

Any reports and responses or comments on the article can be found at the end of the article.

19 epitopes in the SARS-CoV-2 S protein. Findings of this study must be validated in both in vitro and in vivo models.

Conclusions

These data may help clarify the issue of host cell susceptibility and the outcome of SARS-CoV-2 infection, and may guide further research to uncover potential targets for the vaccination strategy.

Keywords

SARS-CoV-2, S protein, peptides, Variants, HLA alleles Moroccan population



This article is included in the **Bioinformatics** gateway.

Corresponding author: Khalid Sadki (ksadki1@yahoo.fr)

Author roles: **Fakhkhari M:** Formal Analysis, Investigation, Methodology, Resources, Visualization, Writing – Original Draft Preparation; **Badaoui B:** Data Curation, Resources, Writing – Review & Editing; **Oumzil H:** Investigation, Resources, Writing – Review & Editing; **Sadki K:** Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Supervision, Validation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: The author(s) declared that no grants were involved in supporting this work.

Copyright: © 2024 Fakhkhari M *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Fakhkhari M, Badaoui B, Oumzil H and Sadki K. **Design of SARS-CoV-2 protein S peptides recognized by the most frequent HLA alleles in the Moroccan population using an immunoinformatics approach [version 1; peer review: 1 not approved]** F1000Research 2024, 13:514 <https://doi.org/10.12688/f1000research.150049.1>

First published: 20 May 2024, 13:514 <https://doi.org/10.12688/f1000research.150049.1>

List of abbreviations

COVID-19: Coronavirus disease
 E: Envelope protein
 HLA: Human leukocyte antigen
 M: Membrane glycoprotein
 MHC: Major histocompatibility complex
 N: Nucleocapsid phosphoprotein
 ORF: Open reading frame
 RBD: Receptor-binding domain
 S: Spike glycoprotein
 SARS-CoV-2: Severe acute respiratory syndrome 2
 WHO: World Health Organization

Introduction

The COVID-19 pandemic has caused unparalleled economic and social disruption across the world. COVID-19 is a respiratory illness that results from an infection with the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The virus was initially in December 2019 in Wuhan, China, and has quickly spread throughout the world.^{1,2} By December 2022, the World Health Organization (WHO) had reported over 651 million cases of COVID-19, with more than 6 million deaths attributed to the disease. The first instance of the COVID-19 virus was registered in the Kingdom of Morocco on March 02, 2020, in Casablanca city. The Moroccan patient had acute pneumonia, and the infection was imported from Europe.³ Between January 3, 2020, and August 16, 2023, the Ministry of Health in Morocco reported a total of 1 275 320 confirmed cases of COVID-19 with 16 297 deaths in within the country. The mortality rate is stated at 1.3% (Ministry of Health, Morocco, CNOUSP report, April 2023).

A notable characteristic of COVID-19 that continually surprises us is the extensive range of clinical symptoms that patients exhibit across different populations. These symptoms range from mild to severe, with severe cases potentially leading to pneumonia, respiratory failure, multi-organ failure, and death.⁴ These differences highlight the significance of studying and understanding human genetic variation during infections. Previous studies have linked the susceptibility and outcomes of multiple infectious diseases to the genetic background of the host. The Human Leukocyte Antigen system is believed to be among the components that could explain differences in virus susceptibility and severity, and it has been recommended as a potential genetic factor that affects a person's immune response to SARS-CoV-2.^{5,6}

The human leukocyte antigen (HLA) system, which is a major component of the adaptive immune system, is in charge of recognizing and binding both endogenous and exogenous antigens.⁵ HLA molecules are classified into two classes: HLA class I and HLA class II. While HLA class II (DR, DQ, DP) molecules are primarily responsible for displaying peptides from external pathogens, HLA class I (A, B, C) molecules are crucial for immune protection against intracellular pathogens. HLA polymorphism is the highest degree of variability found in the genetic code of the proteins expressed by this system. This variability is considered an important factor in determining a person's resistance to, or susceptibility to specific infectious illnesses.^{6,7}

The SARS-CoV-2 genome encodes a number of structural proteins, including the spike protein (S), envelope protein (E), membrane glycoprotein (M), and nucleocapsid phosphoprotein (N), which is in accordance with other coronaviruses. Additionally, it encodes nonstructural proteins, such as open reading frame 1ab (ORF1ab), ORF3a, ORF6, ORF7a, ORF8, and ORF10.⁸ This study focused on the spike protein, which is the primary antigen present on the surface of the virus, facilitating SARS-CoV-2 entrance into human host cells,^{9,10} and additionally because of its high mutation rate, which enables it to alter its shape and escape host immune responses.¹¹ The Spike protein, also known as S protein, is made up of two subunits, S1 and S2, which are connected by a furin cleavage site. The S1 subunit comprises the receptor-binding domain (RBD), which is responsible for the virus's ability to adhere to the host cell membrane and initiate infection. The S2 subunit has a hydrophobic fusion loop which facilitates membrane fusion. Therapeutic developments against SARS-CoV-2 are often targeted at RBD.¹⁰ The spike is frequently studied for the development of neutralizing antibodies and vaccines, and is widely considered as a successful target for detection purposes.^{12–17}

Due to their effectiveness and up speed, computational techniques are far superior to laboratory tests in the drug development process because they can anticipate the antigenic epitopes of specific viral proteins.¹⁸

In Morocco, like in other countries, the COVID-19 pandemic had a noticeable impact on the health system. Although many risk factors of COVID-19 severity have been described, data from North Africa are limited. This study used an immunoinformatics approach to predict the peptide-binding affinity between the most frequent HLA Class I and Class II

alleles in the Moroccan population and SARS-CoV-2 S protein peptides of variants isolated from strains of Moroccan patients. The same analysis was also performed for the SARS-CoV-2 wild-type S protein to assess the ability of these HLA alleles to interact with peptides in the presence or absence of the SARS-CoV-2 mutation, and thus predict which epitopes would be most effective at acting as potent immunogens.

Methods

Selection of Major Histocompatibility Complex (MHC) alleles

The most frequent HLA class I and class II alleles in the Moroccan population were obtained from the Allele Frequency Net Database (<http://www.allelefrequencies.net/pop6001a.asp>), which compiled data from studies conducted in different regions of Morocco. The average allelic frequency of each allele was then computed using data from multiple regions.

Retrieval of Mutated sequences of SARS-COV-2 S protein

To evaluate the HLA-peptide-binding affinity predictions, we obtained the mutated sequences of the SARS-COV-2 S protein for each variant of concern, including Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1), Kappa (B.1.617.1), Delta (B.1.617.2), and Omicron (B.1.1.529), from strains of Moroccan patients available on the GISAID databank.

Retrieval of the Wild-Type SARS-CoV-2 sequence

The reference genome (WuhanHu-1 strain) was obtained from the NCBI Refseq database under GenBank accession number NC_045512.2. The amino acid sequence of the S protein from the reference genome was retrieved in FASTA format and used as a reference sequence for comparison with the sequences of the variants.

Prediction of SARS-CoV-2-derived HLA-binding peptides

The NetMHCpan - v4.0 and NetMHCIIpan v. 3.2 programs were used to predict HLA peptide-binding affinity for HLA class I and class II alleles, respectively. The FASTA sequences from both the GISAID databank and the NCBI database were imported and analyzed. Using the aforementioned sequences, we predicted the binding affinity of each HLA allele to all potential 10-mer and 15-mer overlapping peptides for HLA class I and class II, respectively.

Antigenicity prediction

We used Vaxijen V.2.0 (<http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html>), to assess the potential antigenicity of the predicted peptides by setting a threshold score of less than 0.5 and employing a virus model to eliminate non-antigenic peptides.

The antigenicity response assesses the capability of the proposed epitopes to prompt an immune response.

Toxicity prediction

Toxicity of the predicted peptides was set using the ToxinPred server (https://webs.iitd.edu.in/raghava/toxinpred/multi_submit.php). This server enables the identification of highly toxic or non-toxic peptides from a large pool of peptides submitted by analyzing their key physico-chemical properties such as hydrophobicity, hydropathicity, amphipathicity, molecular weight, and pI charge.

Binding affinity assessment

The HLA binding affinity frequencies of peptides from the mutated S protein were compared with those from the reference S protein (Wild-Type) to identify which HLA alleles had varying binding affinities in the presence of mutations.

This analysis aimed to identify mutated peptides that bind well the most frequent HLA alleles in the Moroccan population.

Results

Table 1 lists the most frequent HLA class I and class II alleles in the Moroccan population, according to the HLA Allele Frequency Net Database. The data reveals a prevalence of HLA class I alleles over class II alleles, with a total of 10 HLA-A, 6 HLA-B, 7 HLA-C, and 7 HLA-DRB1 alleles identified.

In **Table 2**, mutations identified in the SARS-CoV-2 S protein isolated from strains of Moroccan patients are presented, sourced from the GISAID databank. A total of 23 mutations were identified on the spike protein across the different variants of SARS-CoV-2. The mutations highlighted in bold are the common ones found among all six variants, while each variant also includes unique mutations not observed in other variants of concern.

Table 1. Averages frequency of most common HLA class I and II alleles in the Moroccan population.

HLA Class I					
HLA-A allele	Frequency	HLA-B allele	Frequency	HLA-C allele	Frequency
A*01:01	0.131	B*08:01	0.0703	C*02:02	0.065
A*02:01	0.1673	B*18:01	0.0443	C*04:01	0.134
A*03:01	0.043	B*35:01	0.037	C*05:01	0.055
A*23:01	0.08	B*44:03	0.084	C*06:02	0.22
A*24:02	0.048	B*45:01	0.06	C*07:01	0.093
A*29:02	0.059	B*50:01	0.10	C*08:02	0.0613
A*30:01	0.05			C*16:01	0.06
A*30:02	0.057				
A*33:01	0.032				
A*68:02	0.0473				
HLA Class II					
HLA-DRB1			Frequency		
DRB1*01:02			0.0355		
DRB1*03:01			0.152		
DRB1*04:02			0.059		
DRB1*04:05			0.045		
DRB1*07:01			0.163		
DRB1*13:02			0.071		
DRB1*15:01			0.105		

Table 2. SARS-CoV-2 S protein mutations isolated from strains of Moroccan patients (n=23).

SARS-CoV-2 variants	S protein mutations	SARS-CoV-2 variants	S protein mutations
Alpha	V6F	Delta	L452R
	H69-V70		E484K
	Y144		D614G
	N501Y		P681R
	D614G		T19R
	P681H		T478K
Bêta	K417N	Kappa	E484Q
	E484K		D614G
	N501Y	Omicron	E484A
	D614G		Q498R
Gamma	K417T		Q954H
	E484K		Q493R
	N501Y		S704L
	D614G		T547K
	V1176F		T376A
			V213G

Table 3. SARS-CoV-2 S peptides for Class I and II HLA alleles of wild and mutated types.

Mutation	Peptides for Class I (10mers)		Peptides for Class II (15mers)	
	Wild	Mutated	Wild	Mutated
V6F	FVFLVLLPLV	FVFLFLLPLV	-	--
H69-V70	NVTWFHAIHV	NVTWFHAISG	FSNVTWFHAIHVSGT	FSNVTWFHAISGTNG
Y144-145	CNDPFLGVYY	CNDPFLGVYH	QFCNDPFLGVYYHKN	QFCNDPFLGVYHKNN
K417N	GQTGKIADYN	GQTGNIADYN	VRQIAPGQTGKIADY	VRQIAPGQTGNIADY
K417T	GQTGKIADYN	GQTGTIADYN	VRQIAPGQTGKIADY	VRQIAPGQTGTIADY
L452R	GGNYNYLYRL	GGNYNYRYRL	DSKVGGNYNYLYRLF	DSKVGGNYNYRYRLF
E484A	GVEGFNCYFP	GVAGFNCYFP	QAGSTPCNGVEGFNC	QAGSTPCNGVAGFNC
E484K	GVEGFNCYFP	GVKGFNCYFP	PCNGVEGFNCYFPLQ	PCNGVKGFNCYFPLQ
E484Q	GVEGFNCYFP	GVQGFNCYFP	PCNGVEGFNCYFPLQ	PCNGVQGFNCYFPLQ
Q498R	YGFQPTNGVG	YGFRPTNGVG	LQSYGFQPTNGVGYQ	LQSYGFRPTNGVGYQ
N501Y	QSYGFQPTNG	QSYGFQPTYG	QSYGFQPTNGVGYQP	QSYGFQPTYGVGYQP
D614G	YQDVNCTEVP	YQGVNCTEVP	NQVAVLYQDVNCTEV	NQVAVLYQGVNCTEV
P681H	NSPRRARSVA	NSHRRARSVA	YQTQTNSPRRARSVA	YQTQTNSHRRARSVA
P681R	TNSPRRARSV	TNSRRRARSV	ASYQTQTNSPRRARS	ASYQTQTNSRRRARS
V1176F	ISGINASVNN	ISGINASFVN	DISGINASVNNIQKE	DISGINASFVNIQKE
T19R	QCVNLTTRTQ	QCVNLRTRTQ	SSQCVNLTTRTQLPP	SSQCVNLRTRTQLPP
T478K	---	---	---	---
Q954H	GKLQDVVNQN	---	GKLQDVVNQNAQALN	GKLQDVVNHNAQALN
Q493R	NCYFPLQSYG	NCYFPLRSYG	NCYFPLQSYGFQPTN	NCYFPLRSYGFQPTN
S704L	MSLGAENSVA	MSLGAENLVA	SLGAENSVAVSNNIS	SLGAENLVAVSNNIS
T547K	NFNFNGLTGT	NFNFNGLKGT	KCVNFNFNGLTGTGV	KCVNFNFNGLKGTGV
T376A	SFSTFKCYGV	SFSAFKCYGV	SASFSTFKCYGVSPT	SASFSAFKCYGVSPT
V213G	KHTPINLVRD	KHTPINLGRD	PINLVRDLPQGFSAL	PINLGRDLPQGFSAL

-No peptide contains V in its core; --No peptide contains the mutations in its core; ---All predicted peptides were probable non-antigen on Vaxijen.

Our dataset for peptide-binding affinity prediction comprises only peptides with a mutation in their core. Two hundred and twenty peptides were identified based on the results from NetMHCpan servers. Non-immunogenic or toxic peptides were discarded from further analysis. Consequently, eighty-five immunogenic peptides in the SARS-CoV-2 S protein were retained for the assessment of binding affinities (as shown in Table 3).

The comparison of HLA class I binding repertoires between wild type and mutated S protein peptides revealed that 10 mutations in the S protein had the potential to bind to HLA-A*02:01, HLA-A*29:02, HLA-A*68:02, HLA-B*45:01, and HLA-C*16:01.

Among these HLA alleles, HLA-A*02:01 and HLA-C*16:01 had one good binder each with T547K and T376A, respectively. HLA-A*29:02 had three good binders with H69-V70, P681R, and D614G. For HLA-A*68:02, four good binders were found with H69-V70, Y144-145, L252R, and T547K. Similarly, HLA-B*45:01 had four good binders with E484K, E484Q, E484A, and D614G.

Regarding HLA Class II, 9 mutated peptides were found able to interact with HLA-DRB1*01:02, HLA-DRB1*03:01, HLA-DRB1*04:02, HLA-DRB1*04:05, HLA-DRB1*07:01, HLA-DRB1*13:02, and HLA-DRB1*15:01.

Among these, HLA-DRB1*01:02, DRB1*04:02, and HLA-DRB1*04:05 had a common good binder with V213G. HLA-DRB1*03:01 and HLA-DRB1*15:01 had four good binders each, which were (H69-V70, K417N, K417T, and D614G) and (Y144-145, K417N, K417T, and V213G) respectively. On the other hand, HLA-DRB1*07:01 and

Table 4. Sequences of the S protein good binders with their HLA alleles.

	HLA alleles	Good binders' sequences
Class I	HLA-A*02:01	NFNFNGLKGT
	HLA-A*29:02	NVTWFHAISG, TNSRRRARSV, and YQGVNCTEVP
	HLA-A*68:02	NVTWFHAISG, CNDPFLGVYH, GGNYNRYRL and NFNFNGLKGT
	HLA-B*45:01	GVKGFNCYFP, GVQGFNCYFP, GVAGFNCYF, and YQGVNCTEVP
	HLA-C*16:01	SFSAFKCYGV
Class II	HLA-DRB1*01:02	PINLGRDLPQGFSAL
	HLA-DRB1*03:01	FSNVTWFHAISGTNG, VRQIAPGQTGNIADY, VRQIAPGQTGTIADY, and NQVAVLYQGVNCTEV
	HLA-DRB1*04:02	PINLGRDLPQGFSAL
	HLA-DRB1*04:05	PINLGRDLPQGFSAL
	HLA-DRB1*07:01	VRQIAPGQTGNIADY, VRQIAPGQTGTIADY, DSKVGGNRYRL, SSQCVNLRTRTLPP, and PINLGRDLPQGFSAL
	HLA-DRB1*13:02	QSYGFQPTYGVGYQP, FSNVTWFHAISGTNG, VRQIAPGQTGTIADY, SSQCVNLRTRTLPP, and PINLGRDLPQGFSAL
	HLA-DRB1*15:01	QFCNDPFLGVYHKN, VRQIAPGQTGNIADY, VRQIAPGQTGTIADY, and PINLGRDLPQGFSAL

HLA-DRB1*13:02 had five good binders each with (K417N, K417T, L252R, T19R, and V213G) and (N501Y, H69-V70, K417T, T19R, and V213G) respectively. Table 4 provides more details on the peptide sequences.

The HLA-DRB1*15:02 allele was the only one to exhibit good binding affinity to the N501Y mutation, which emerged in the Alpha, Beta and Gamma lineages. Conversely, the H69-V70, Y144-145, L452R, and D614G mutations were presented by both HLA class I and class II alleles, while the K417N, K417T, and V213G mutations were only presented in Class II.

NetMHCIIpan v. 3.2 did not predict any peptides for the most frequent HLA class II alleles in the Moroccan population for both the Spike V6F mutation and the wild-type spike sequence. All predicted peptides for the T478K mutation in the receptor binding domain (RBD) of the Delta variant were deemed non-antigenic by the Vaxijen tool. This mutation is believed to help evade recognition by the immune system,¹⁹ particularly with regards to antibodies and neutralization, or in impairing the interaction between the RBD and drugs.²⁰ Tables 5 and 6 present the antigenicity scores of the selected binders.

Table 5. Vaxijen scores (antigenicity) of the predicted MHC class I allele binding peptides.

Variant	Peptide	Vaxijen score	Final decision
Alpha	NVTWFHAISG	0.6203	-
Delta	TNSRRRARSV	1.1716	*
Alpha/Beta/Gamma/Delta	YQGVNCTEVP	1.0639	*
Alpha	CNDPFLGVYH	0.4600	-
Delta	GGNRYRL	1.2747	*
Beta/Gamma/Delta	GVKGFNCYFP	0.7415	-
Kappa	GVQGFNCYFP	0.7874	-
Omicron	GVAGFNCYFP	1.2211	*
Omicron	NFNFNGLKGT	1.3565	*
Omicron	SFSAFKCYGV	0.5764	-

*Selected good binders.

Table 6. Vaxijen scores (antigenicity) of the predicted MHC class II allele binding peptides.

Variant	Peptide	Vaxijen score	Final decision
Alpha	FSNVTWFHAISGTNG	0.8214	-
Beta	VRQIAPGQTGNIADY	1.1378	*
Gamma	VRQIAPGQTGTIADY	1.1046	*
Alpha/Beta/Gamma/Delta	NQVAVLYQG ^V NCTEV	0.7970	-
Delta	DSKVGGN ^Y NYR ^Y RLF	1.0304	*
Alpha/Beta/Gamma	QSYGFQPT ^Y GVGYQP	0.7291	-
Alpha	QFCNDPFLGVYHKNN	0.7581	-
Delta	SSQCVNLR ^T RTQLPP	1.2742	*
Omicron	PINLGRDLPQGF ^S AL	0.9873	*

*Selected good binders.

Discussion

Due to their efficiency and speed, computational approaches have consequently arisen as potent alternatives or choices for the development of diagnostic tools for infectious diseases, as well as new immunotherapies and vaccines.¹⁰ Predictive computational for identifying antigenic epitopes in viral or bacterial proteins are critical and valuable in the development of new drugs.^{18–21} It has been recommended to employ these tools before performing laboratory experiments because they are more efficient and faster to use.²² Several studies have utilized immunoinformatic approaches on different SARS-CoV-2 proteins to design potential epitope vaccine candidates against SARS-CoV-2.^{8,23–31} In this study, we designed SARS-CoV-2 S protein peptides that bind to the most frequent HLA class I and class II alleles in the Moroccan population. The aim was to evaluate the potential of these peptides to elicit an immune response using in silico methods. The SARS-CoV-2 S protein was chosen as the target in this study because it is the most mutated structural protein of SARS-CoV-2 and is frequently studied as it helps in target recognition and cellular entry. This protein promotes viral infection and is essential for the development of neutralizing antibodies and vaccines.^{10,12} Additionally, the S protein is commonly acknowledged as a suitable target for detection purposes.¹⁶

In this study, we selected twenty-three mutations from five SARS-CoV-2 variants (Alpha, Beta, Gamma, Delta, and Omicron) as they are considered key mutations associated with higher transmission and reinfection rates.^{32,33} We evaluated the ability of class I and class II HLA molecules to present the mutated peptides of the SARS-CoV-2 spike protein. Our findings showed that HLA class I molecules had a higher proportion of good binders compared to HLA class II alleles, with 10 versus 9, respectively.

The peptides SSQCVNLR^TRTQLPP, VRQIAPGQTGNIADY, VRQIAPGQTGTIADY, DSKVGGN^YNYR^YRLF, and PINLGRDLPQGF^SAL were selected as immunogens for HLA class II based on their antigenic score (1.2742; 1.1378; 1.1046; 1.0304 and 0.9873 respectively), non-allergenicity, and lack of toxicity. For HLA class I, the peptides NFNFNGLKGT, GGN^YNYR^YRL, GVAGFNCYFP, TNSRRRARSV, and YQGVNCTEVP exhibited the best antigenic scores of 1.3565; 1.2747; 1.2211; 1.1716; and 1.0639, respectively. The scores therefore indicate the stimulation of the immune system in response to the proposed epitopes. Given their binding affinity with the most frequent HLA alleles in the Moroccan population and antigenicity response, these epitopes may be promising candidates for vaccine development.

Potential immunogenic peptides have been identified as prospective COVID-19 vaccine targets using in silico studies. Previous research has indicated that HLA-A*02:01, among other alleles, exhibited the strongest binding to COVID-19 epitopes.^{34,35} Our study revealed that the “NFNFNGLKGT” peptide of the **T547K** mutation in the Omicron variant showed a higher antigenicity score and had the highest affinity to this particular allele. Consequently, this epitope could be a promising candidate for the development of a COVID-19 vaccine.

Regarding the HLA-C allele, a peptide belonging to the Omicron subvariant peptide with the sequence SFS^AFKCYGV was predicted to have high binding affinity with HLA-C*16:02. However, it was not selected due to its Vaxijen score of 0.5764. This gene has been previously reported to have a less distinctive peptide repertoire when compared to HLA-A and HLA-B.³⁶ HLA-C*16:01 was found to be more prevalent among individuals who had a mild form of COVID-19 compared with those with severe or critical forms of the disease in a cohort of Spanish Mediterranean Caucasians.³⁷

HLA-A*68:02 was the MHC-I molecule that bound the immune epitopes of the S protein **L452R** mutation of the Delta variant (GGNYNYRYRL). A study conducted in Tapachula Chapas, found that although the frequency of HLA-A*68 was lower in COVID-19 patients who were ill, the allele provided 3.3 times more protection against a fatal outcome from SARS-CoV-2 infection in mestizo individuals.³⁸ Both HLA-A*68:01 and HLA-A*68:02 have the ability to bind to a large number of SARS-CoV-2 peptides with varying degrees of affinity. Furthermore, a systematic review and meta-analysis found that HLA-A*68:02, along with other HLA class I and class II alleles, were associated with COVID-19 severity.³⁹

HLA-A, HLA-B, HLA-C, and HLA-DRB1 may serve as potential indicators of the severity and likelihood of death from COVID-19. However, further research on a larger scale are required to confirm this hypothesis.

A literature review has revealed that HLA class I alleles may be deemed as determinants of either resistance or susceptibility to COVID-19. This is due to the fact that these alleles have the ability to bind to SARS-CoV-2 peptides, leading to the modulation of the immune response against the virus.⁴⁰

Several studies have employed a similar approach to design peptides for the protein S of SARS-CoV-2. *Baruah et al.* found five CD8+ T cell epitopes YLQPRTFLL, GVFYFASTEK, EPVLKGVKL, VVNQNAQAL, and WTA-GAAAYY, along with eight B cell epitopes that are more likely to bind MHC class I commonly found in China.⁴¹ In a second study, *Bhattacharya et al.* identified thirteen potential MHC-I antigenic peptides (SQCVNLITR, GVYYHKNNK, GKQGNFKNL, GIYQTSNFR, VSPTKLNDL, KIADYNYKL, KVGGNYYNL, EGFNCYFPL, GPKKSTNLV, SPRRARSVA, LGAENSVAY, FKNHTSPDV, and DEDDSEPV) and three potential MHC-II antigenic peptides (IHVSGTNGT, VYYHKNNKS, and FKNHTSPDV).⁴²

Joshi et al. proposed the MHC-I ITLCFTLKR peptide as a potential vaccine candidate,⁴³ while another study conducted on the Brazilian population found 24 epitopes that bind to 17 different MHC-I alleles.⁴⁴ Other studies have also identified B-cell epitopes on spike protein for developing a protective vaccine against SARS-CoV-2.^{23,24,45}

The current study's findings are novel and have not been previously published. These findings are valuable for the development of broadly accessible vaccine epitopes targeting SARS-CoV-2, and can also offer valuable insights for investigating T-cell responses.

Nevertheless, to confirm their immunogenicity against SARS-CoV-2, further in-vitro experimental validation or in vivo studies are necessary.

Conclusions and perspectives

This is the first Moroccan in silico study to assess potential immunogenic peptides within the S protein of various SARS-CoV-2 variants according to the most frequent HLA alleles in the Moroccan population, using an immunoinformatic approach. The findings of the current study have not been published previously. To sum up, we identified 19 epitopes in the SARS-CoV-2 S protein that can bind to 12 distinct HLA Class I and Class II alleles among the Moroccan population, as they were characterized by a probability of triggering an immune response. However, in order to validate their immunogenicity against SARS-CoV-2, additional in-vitro experimental validation or in vivo studies are essential.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Author's contributions

MF: Methodology, Data curation, Formal analysis, Writing – original draft. **BB**: Data curation, Writing – Review & Editing. **HO**: Visualization, Writing – review & editing, Investigation. **KS**: Conceptualization, Methodology, Formal analysis, Supervision, Writing – review & editing.

Data availability statement

Underlying data

Figshare: Dataset - Design of SARS-CoV-2 protein S peptides recognized by the most frequent HLA alleles in the Moroccan population using an immunoinformatics approach. <https://doi.org/10.6084/m9.figshare.25737534.v1>.⁴⁶

The dataset contains the following data:

- Data.docx: COVID-19 Wild Type Sequence and Selected Mutations from Various Variants
- Data.xlsx: Peptide-HLA Class I Binding Affinity Assessment Wild Type (WT) and Mutated (MT) Peptide Binding Scores
- Data.xlsx: Peptide-HLA Class II Binding Affinity Assessment Wild Type (WT) and Mutated (MT) Peptide Binding Scores
- Data.docx: SARS-CoV-2 Sequences of different variants retrieved from GISAID Databank
- Data.xlsx: Table 1. Averages frequency of most common HLA class I and II alleles in the Moroccan population
- Data.xlsx: Table 2. SARS-CoV-2 S protein mutations isolated from strains of Moroccan patients (n=23)
- Data.xlsx: Table 3. SARS-CoV-2 S peptides for Class I and II HLA alleles of wild and mutated types
- Data.xlsx: Table 4. Sequences of the S protein good binders with their HLA alleles
- Data.xlsx: Table 5. Vaxijen scores (antigenicity) of the predicted MHC class I allele binding peptides.
- Data.xlsx: Table 6. Vaxijen scores (antigenicity) of the predicted MHC class II allele binding peptides.

Data are available under the terms of the [Creative Commons Attribution 4.0 International license](#) (CC-BY 4.0).

Extended data

Extended data is not applicable in this instance. All relevant data and materials utilized in our study have been comprehensively outlined in the 'Materials and Methods' section and provided in the Data Availability section. This includes the data uploaded to the designated repository, as specified in the Data Availability section, which encompasses all materials pertinent to our research. We encourage readers to refer to the repository for access to the complete set of data and materials used in this study. Additionally, all servers, databases and methods utilized in this research are outlined in detail within the manuscript. For further inquiries regarding the data or materials, please contact the corresponding author.

Acknowledgements

Not applicable.

References

1. Iturrieta-Zuazo I, Rita CG, García-Soldán A, *et al.*: **Possible role of HLA class-I genotype in SARS-CoV-2 infection and progression: A pilot study in a cohort of Covid-19 Spanish patients.** *Clin. Immunol. Orlando Fla.* 2020; **219**: 108572.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
2. Toyoshima Y, Nemoto K, Matsumoto S, *et al.*: **SARS-CoV-2 genomic variations associated with mortality rate of COVID-19.** *J. Hum. Genet.* 2020; **65**(12): 1075–1082.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
3. Hammoui A, Qesmi R: **Impact assessment of containment measure against COVID-19 spread in Morocco.** *Chaos. Solitons Fractals.* 2020; **140**: 110231.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
4. Anastassopoulou C, Gkizarioti Z, Patrinos GP, *et al.*: **Human genetic factors associated with susceptibility to SARS-CoV-2 infection and COVID-19 disease severity.** *Hum. Genomics.* 2020; **14**(1): 40.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
5. Blackwell JM, Jamieson SE, Burgner DHLA, *et al.*: **HLA and infectious diseases.** *Clin. Microbiol. Rev.* 2009; **22**(2): 370–385, Table of Contents.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
6. Medhasi S, Chantratita N: **Human Leukocyte Antigen (HLA) System: Genetics and Association with Bacterial and Viral Infections.** *J. Immunol. Res.* 2022; **2022**: 9710315–9710376.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
7. Sanchez-Mazas A: **A review of HLA allele and SNP associations with highly prevalent infectious diseases in human populations.** *Swiss Med. Wkly.* 2020; **150**(1516): w20214.
[PubMed Abstract](#) | [Publisher Full Text](#)
8. Kiyotani K, Toyoshima Y, Nemoto K, *et al.*: **Bioinformatic prediction of potential T cell epitopes for SARS-Cov-2.** *J. Hum. Genet.* 2020; **65**(7): 569–575.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
9. Hadj Hassine I: **Covid-19 vaccines and variants of concern: A review.** *Rev. Med. Virol.* 2022; **32**(4): e2313.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
10. Hajikarimlou M, Hooshyar M, Moutaoufik MT, *et al.*: **A computational approach to rapidly design peptides that detect SARS-CoV-2 surface protein S.** *NAR Genomics Bioinforma.*

- 2022; 4(3): lqac058.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
11. Mengist HM, Kombe AJK, Mekonnen D, *et al.*: **Mutations of SARS-CoV-2 spike protein: Implications on immune evasion and vaccine-induced immunity.** *Semin. Immunol.* 2021; 55: 101533.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 12. Jia Z, Gong W: **Will Mutations in the Spike Protein of SARS-CoV-2 Lead to the Failure of COVID-19 Vaccines?** *J. Korean Med. Sci.* 2021; 36(18): e124.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 13. Tian F, Tong B, Sun L, *et al.*: **N501Y mutation of spike protein in SARS-CoV-2 strengthens its binding to receptor ACE2.** *elife.* 2021; 10: e69091.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 14. Hoffmann M, Kleine-Weber H, Schroeder S, *et al.*: **SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor.** *Cell.* 2020; 181(2): 271–280. e8.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 15. Tian X, Li C, Huang A, *et al.*: **Potent binding of 2019 novel coronavirus spike protein by a SARS coronavirus-specific human monoclonal antibody.** *Emerg. Microbes Infect.* 2020; 9(1): 382–385.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 16. Lee JH, Choi M, Jung Y, *et al.*: **A novel rapid detection for SARS-CoV-2 spike 1 antigens using human angiotensin converting enzyme 2 (ACE2).** *Biosens. Bioelectron.* 2021; 171: 112715.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 17. Awadasseid A, Wu Y, Tanaka Y, *et al.*: **Current advances in the development of SARS-CoV-2 vaccines.** *Int. J. Biol. Sci.* 2021; 17(1): 8–19.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 18. Brogi S: **Computational Approaches for Drug Discovery.** *Molecules.* 2019; 24(17): 3061.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 19. Di Giacomo S, Mercatelli D, Rakhimov A, *et al.*: **Preliminary report on severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) Spike mutation T478K.** *J. Med. Virol.* 2021; 93(9): 5638–5643.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 20. Pondé RAA: **Physicochemical effect of the N501Y, E484K/Q, K417N/T, L452R and T478K mutations on the SARS-CoV-2 spike protein RBD and its influence on agent fitness and on attributes developed by emerging variants of concern.** *Virology.* 2022; 572: 44–54.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 21. Seghrouchni F, Contini S, Markova R, *et al.*: **Design of immunogenic peptides from Mycobacterium tuberculosis genes expressed during macrophage infection.** *Tuberculosis.* 2009; 89(3): 210–217.
[PubMed Abstract](#) | [Publisher Full Text](#)
 22. Sunita SA, Singh Y, Shukla P: **Computational tools for modern vaccine development.** *Hum. Vaccin. Immunother.* 2019; 16(3): 723–735.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 23. Sanami S, Zandi M, Pourhossein B, *et al.*: **Design of a multi-epitope vaccine against SARS-CoV-2 using immunoinformatics approach.** *Int. J. Biol. Macromol.* 2020; 164: 871–883.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 24. Chen HZ, Tang LL, Yu XL, *et al.*: **Bioinformatics analysis of epitope-based vaccine design against the novel SARS-CoV-2.** *Infect. Dis. Poverty.* 2020; 9(1): 88.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 25. Takagi A, Matsui M: **Identification of HLA-A*02:01-Restricted Candidate Epitopes Derived from the Nonstructural Polyprotein 1a of SARS-CoV-2 That May Be Natural Targets of CD8+ T Cell Recognition In Vivo.** *J. Virol.* 2021; 95(5): e01837–e01820.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 26. Abraham Peele K, Srihansa T, Krupanidhi S, *et al.*: **Design of multi-epitope vaccine candidate against SARS-CoV-2: a in-silico study.** *J. Biomol. Struct. Dyn.* 2021; 39(10): 3793–3801.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 27. He J, Huang F, Zhang J, *et al.*: **Vaccine design based on 16 epitopes of SARS-CoV-2 spike protein.** *J. Med. Virol.* 2021; 93(4): 2115–2131.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 28. Alam A, Khan A, Imam N, *et al.*: **Design of an epitope-based peptide vaccine against the SARS-CoV-2: a vaccine-informatics approach.** *Brief. Bioinform.* 2021; 22(2): 1309–1323.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 29. Chukwudozie OS, Gray CM, Fagbayi TA, *et al.*: **Immuno-informatics design of a multimeric epitope peptide based vaccine targeting SARS-CoV-2 spike glycoprotein.** *PLoS One.* 2021; 16(3): e0248061.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 30. Behmard E, Soleymani B, Najafi A, *et al.*: **Immunoinformatic design of a COVID-19 subunit vaccine using entire structural immunogenic epitopes of SARS-CoV-2.** *Sci. Rep.* 2020; 10: 20864.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 31. Wang E, Chakraborty AK: **Design of immunogens for eliciting antibody responses that may protect against SARS-CoV-2 variants.** *PLoS Comput. Biol.* 2022; 18(9): e1010563.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 32. Kumavath R, Barh D, Andrade BS, *et al.*: **The Spike of SARS-CoV-2: Uniqueness and Applications.** *Front. Immunol.* 2021; 12: 663912.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 33. Jhun H, Park HY, Hisham Y, *et al.*: **SARS-CoV-2 Delta (B.1.617.2) Variant: A Unique T478K Mutation in Receptor Binding Motif (RBM) of Spike Gene.** *Immune Netw.* 2021; 21(5): e32.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 34. Lee CH, Koohy H: **In silico identification of vaccine targets for 2019-nCoV.** *F1000Res.* 2020; 9: 145.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 35. Jandaghi A, Samiei A, Khaghanzadeh N: **Human Leukocyte Antigen as a Predictor of COVID-19 Severity.** *Hormozgan Med. J.* 2022; 26(3): 156–162.
[Publisher Full Text](#)
 36. Basir HRG, Majzoobi MM, Ebrahimi S, *et al.*: **Susceptibility and Severity of COVID-19 Are Both Associated With Lower Overall Viral-Peptide Binding Repertoire of HLA Class I Molecules, Especially in Younger People.** *Front. Immunol.* 2022; 13.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 37. Vigón L, Galán M, Torres M, *et al.*: **Association between HLA-C alleles and COVID-19 severity in a pilot study with a Spanish Mediterranean Caucasian cohort.** *PLoS One.* 2022; 17(8): e0272867.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 38. Hernández-Doño S, Sánchez-González RA, Trujillo-Vizuet MG, *et al.*: **Protective HLA alleles against severe COVID-19: HLA-A*68 as an ancestral protection allele in Tapachula-Chiapas, Mexico.** *Clin. Immunol. Orlando Fla.* 2022; 238: 108990.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 39. Dobrijević Z, Gligorjević N, Šunderić M, *et al.*: **The association of human leukocyte antigen (HLA) alleles with COVID-19 severity: A systematic review and meta-analysis.** *Rev. Med. Virol.* 2022; 33: e2378.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 40. Menezes L, Garcia LR, Casseb SMM: **HLA-I allelic variants related to resistance and susceptibility to severe SARS-CoV-2 infection: Variantes alélicas do HLA I relacionados a resistência ou suscetibilidade a infecção grave pelo SARS-CoV-2.** *Concilium.* 2023; 23(5): 293–307.
[Publisher Full Text](#)
 41. Baruah V, Bose S: **Immunoinformatics-aided identification of T cell and B cell epitopes in the surface glycoprotein of 2019-nCoV.** *J. Med. Virol.* 2020; 92(5): 495–500.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 42. Bhattacharya M, Sharma AR, Patra P, *et al.*: **Development of epitope-based peptide vaccine against novel coronavirus 2019 (SARS-COV-2): Immunoinformatics approach.** *J. Med. Virol.* 2020; 92(6): 618–631.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 43. Joshi A, Joshi BC, Mannan MA, *et al.*: **Epitope based vaccine prediction for SARS-COV-2 by deploying immuno-informatics approach.** *Inform. Med. Unlocked.* 2020; 19: 100338.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 44. de Moura RR, Agreli A, Santos-Silva CA, *et al.*: **Immunoinformatic approach to assess SARS-CoV-2 protein S epitopes recognised by the most frequent MHC-I alleles in the Brazilian population.** *J. Clin. Pathol.* 2021; 74(8): 528–532.
[PubMed Abstract](#) | [Publisher Full Text](#)
 45. Samad A, Yamchi A, Madanchi H, *et al.*: **Designing a multi-epitope vaccine against SARS-CoV-2: an immunoinformatics approach.** *J. Biomol. Struct. Dyn.* 2020; 40: 14–30.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 46. Fakhkhari M, *et al.*: **Design of SARS-CoV-2 protein S peptides recognized by the most frequent HLA alleles in the Moroccan population using an immunoinformatics approach. [Dataset].** *figshare.* 2024.
[Publisher Full Text](#)

Open Peer Review

Current Peer Review Status: 

Version 1

Reviewer Report 23 September 2024

<https://doi.org/10.5256/f1000research.164583.r309414>

© 2024 Vieira G. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Gustavo Fioravanti Vieira 

Bioinformatic Core, Immunogenetics Laboratory, Genetics Department, Biosciences Institute, Post-Graduation Program in Health and Human Development, Federal University of Rio Grande do Sul, Universidade La Salle, Porto Alegre, Canoas, Brazil

While it is very important to design vaccines and immunotherapeutic approaches considering the genetics of the immunized population, some aspects of the work need to be improved or reconsidered before indexing:

- The introduction is too basic and lacks a stronger foundation with references supporting the influence of HLA on the variation in severity, particularly in relation to COVID.
- What was the criterion for the selected alleles? "Most frequent alleles" is too vague.
- NetMHCpan is already at version 4.1 for both class I and II MHC.
- For class I, it was essential to use nonamers instead of decamers, as nonamers are the standard fragment length for most alleles. Alternatively, the appropriate size should have been used for each different allele.
- Vaxijen is suitable for an initial analysis, but it does not include MHC-related information in its prediction, making it insufficient for this case. Ideally, you should use sequences already described as immunogenic in IEDB.
- I didn't understand why the toxicity of the peptides was assessed. Is the idea to administer them synthetically?

References

1. Lunardi LW, Bragatte MAS, Vieira GF: The influence of HLA/HIV genetics on the occurrence of elite controllers and a need for therapeutics geotargeting view. *Braz J Infect Dis*. 2021; **25** (5): 101619 [PubMed Abstract](#) | [Publisher Full Text](#)
2. Pretti MAM, Galvani RG, Vieira GF, Bonomo A, et al.: Class I HLA Allele Predicted Restricted Antigenic Coverages for Spike and Nucleocapsid Proteins Are Associated With Deaths Related to COVID-19. *Front Immunol*. 2020; **11**: 565730 [PubMed Abstract](#) | [Publisher Full Text](#)

Is the work clearly and accurately presented and does it cite the current literature?

Partly

Is the study design appropriate and is the work technically sound?

No

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Not applicable

Are all the source data underlying the results available to ensure full reproducibility?

Partly

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Immunoinformatics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.

Author Response 12 Feb 2025

Khalid Sadki

Here is our response to all comments.

- 1: Some references will be added according to the first comment concerning the introduction section.
- 2: We chose to set an allele frequency threshold of 3% based on HLA polymorphism research conducted in Moroccan communities; hence, any allele with a frequency higher than 3% is regarded as more prevalent. The method section will be updated with this information.
- 3: In our study, we used NetMHCpan v4.0 for MHC class I predictions and NetMHCIpan v3.2 for class II. After comparing these with version 4.1, we found that the differences in peptide binding affinities were minimal and statistically insignificant for the targeted mutations. The same peptides were consistently identified across both versions, with no impact on their classification or our conclusions. In the revised version, we will use NetMHCpan v4.1 to update the analyses rather than NetMHCIpan v3.2.
- 4: Studies have demonstrated that 10-mer peptides form stable complexes with a variety of HLA alleles, including HLA-A02:01, which is common in many populations, even though 9-mers are generally accepted as the standard [Ahn et al., 2024, doi: 10.1016/j.crstbi.2024.100148; Nerli & Sgourakis, 2020, doi: 10.1016/j.crstbi.2024.100148].

10.3389/fmedt.2020.553478]. Additionally, in the context of SARS-CoV-2, 10-mer peptides have been identified as immunogenic epitopes, eliciting robust CD8+ T cell responses [Ahn et al., 2024, doi: [10.1016/j.crstbi.2024.100148](https://doi.org/10.1016/j.crstbi.2024.100148); Abd El-Baky et al., 2023, doi: [10.3390/vaccines11030548](https://doi.org/10.3390/vaccines11030548); Kiyotani et al., 2020, doi: [10.1038/s10038-020-0771-5](https://doi.org/10.1038/s10038-020-0771-5); Arshad et al., 2024, doi: [10.1016/j.heliyon.2024.e24186](https://doi.org/10.1016/j.heliyon.2024.e24186); de Moura RR et al., 2021, doi: [10.1136/jclinpath-2020-206946](https://doi.org/10.1136/jclinpath-2020-206946)]. Furthermore, the study by Motozono et al. (2015) demonstrated that MHC class I molecules can undergo structural adaptations to accommodate 10-mer peptides, maintaining stable peptide-MHC interactions, which supports their relevance in T-cell-mediated immune responses [Motozono et al., 2015, doi: [10.1074/jbc.M114.622522](https://doi.org/10.1074/jbc.M114.622522)]. By incorporating 10-mers in our study, we aimed to capture a broader repertoire of potential T-cell epitopes, accounting for the structural and functional diversity of HLA alleles. This approach ensures comprehensive epitope coverage, enhancing the relevance of our findings for diverse populations. Nevertheless, we will also consider 9-mer peptides in future analyses to further optimize epitope selection.

- 5: We agree that Vaxijen is primarily used for initial antigenicity screening and does not incorporate MHC-related information in its predictions. However, studies begin with Vaxijen to assess antigenicity, particularly when dealing with a large number of generated peptides. This helps minimize the inclusion of peptides that might not be antigenic, allowing for a more efficient selection process before conducting more specific immunogenicity and MHC binding predictions. In the revised version, we will also incorporate IEDB to further validate our findings and confirm the immunogenic potential of the selected peptides.
- 6: While the current focus of our study is on identifying immunogenic peptides, evaluating their toxicity using bioinformatics methods is a preliminary step to ensure the peptides are safe for potential future applications, such as synthetic administration in peptide-based vaccines. Toxicity analysis is commonly used in this type of study to support the safety assessment of candidate peptides. This computational analysis helps identify any harmful properties at an early stage, reducing the risk of toxicity in further experimental or therapeutic developments. We will clarify this in the manuscript to explain the rationale for including this assessment.

Competing Interests: I have no actual or potential conflict of interest in relation to this work.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research