Immunoinformatic paradigm predicts macrophage and T-cells epitope responses against globally conserved spike fragments of SARS CoV-2 for universal vaccination

Smarajit Maiti [a,b,*], Amrita Banerjee [a,*], Dipannita Santra [a], Mehak Kanwar [a,1]

[a] Department of Biochemistry and Biotechnology, Cell and Molecular Therapeutics Laboratory, Oriental Institute of Science and Technology, Midnapore, India
[b] Agricure Biotech Research Society, Epidemiology and Human Health Division, Midnapore 721 101, India

ABSTRACT

Background: Different quickly-developed vaccines are introduced against COVID-19 with inconclusive results especially against some recent variants. Eventually, somewhere COVID-19 cases decline and in some countries it revived with some new mutant-variants (i.e. D614G, Delta and Omicron).
Objectives: Proposing a universal vaccination strategy by screening globally-conserved SARS-CoV-2 spike-epitopes.
Methods: Presently, several conserved (186-countries) sequences including multiple-variants (ClustalX2) epitopic-regions (SVMTriP and IEDB) and in-silico mutants of SARS-CoV-2 spike-protein-fragments (Cut1-4) were screened for their stability against proteases, antigenicity (VaxiJen V2.0 and for glycosylation effects NetOGlyc-NetNGlyc), MHCI/II reactivity (IEDB-TOOLS) and CD4+ responses by molecular-docking (Haddock2.4/PatchDock). We also examined Molecular-Dynamic-Simulation (myPresto verson-5) of MHC-II 3LQZ with 3-Cuts and T-cell 2-molecules (1KGC/4JRX) with SM3-Cut. The MD-simulation was run with 5000-cycles after 300 k-heating/1-atm pressure adjustment for the system-equilibration. Finally, 1000 fs production was run.
Results: The cut4-mutant (SRLFRKSNLKPFERD) showed the highest combined-score 48.23548 and Immunogenicity-Score of 92.0887. The core-sequence SRLFRKSNL showed the highest Median-Percentile-Rank (7-HLA-allele) of 19. CD4+ immunogenicity also confirms the representation of the CUT4TM2 epitope SRLFRKSNL by MHC Class II. The epitope YNYKYRLFR from CUT4 showed an IC50 of ~30 nM with allele HLA-DRB1*11:01 and HLA-DRB5*01:01 with plenty H-bonding. Cut4 double-mutants strongly interact with the exposed T-cell surface and are facilitated by its receptors. The MD-simulation data suggest that TM2 has a maximum RMSD value of 1.7 Å, DM2 is at 1.55 Å and SM3 is at 1.5 Å. These variations correspond to structural adjustments and involve binding/unbinding chemical interactions. The RMSD plot shows that 1KGC T-cell molecule is at 2.2 Å and the 4JRX is at 1.2 Å, which increases with the simulation time.
Conclusions: Screening of conserved SARS-CoV-2 spike fragments helps to find the most stable antigenic-determinant which with some mutations showed better antigenicity. Further studies are necessary to develop global vaccination strategies against COVID-19.

1. Introduction

At the end of 2020, at least 30 vaccines are in the process of clinical trials. Some of those, like Oxford's AZD1222/AstraZeneca/Oxford's AZD1222/Moderna's mRNA-1273 and Sinovac's CoronaVac vaccines were reported with inconclusive results [1]. Although, protection against SARS-CoV-2 infection in humans is not recognized. While the memory cellular response to the mRNA −1273 vaccine is not clear yet, this vaccine elicited primary CD4 type 1 helper T cell and studies of vaccine induced B cells are ongoing [2]. Some countries have introduced

self-generated vaccines to their populations. Covishield (Oxford-Astra-Zeneca) and Covaxin (produced by Bharat biotech) from India have been introduced to a significant number of populations (ref https://www.bbc.com/news/world-asia-india-55748124). From another study, a booster dose of Pfizer vaccine strongly increased the levels of anti-SARS-CoV-2 neutralizing antibodies [3]. Those have some diverse outcomes, criticism and no definite clinical-trial outcome. In a natural way, the infection and the case fatality rate (CFR) are declining but ample evidence of multiple mutations and highly evolving SARS CoV-2 urges the development of a more effective vaccine effective against global variants. A wide range of studies have been proposed for the anti-tuberculosis vaccine, Bacile Calmette- Guerin (BCG), which has inherent immunostimulatory properties against SARS-Cov-2 [4]. In some countries with the largest populations, induction of herd immunity by mass vaccination has been effective in the prevention of these types of diseases [5]. The autoimmune suppressive effects of this virus prevent or delay the immunotherapy processes. In multiple vaccinations or other immunotherapy cannot affect at ease. A negative outcome was also noticed during drug-induced inhibition of some subsets of B-cell that hinders the initiation of innate and CD8+ cell responses [6]. None of the COVID-19 vaccines is fully effective with safety to get priority at this point of time. Only rigorous clinical trial and post application follow-up can recognize serious health hazards [7].

The report suggests that ChAdOx1 nCoV-19 has an acceptable safety profile against symptomatic COVID-19 [8]. But that statistical outcome is generated from a small sample size. The candidate vaccine mRNA-1273 encodes the stabilized prefusion SARS-CoV-2 spike protein and generated some immune responses [9]. Not only the generation of specific subsets of IgG but the MHC, T and B cell responses have also been important for the optimized immune responses.

Th1-targeted antibody, CD8+ and CD4+ T cell responses and neutralizing IgG antibody responses were noticed with S protein-specific viral neutralization [10]. Though, vaccine mechanisms depend on the native S-protein for inducing potent neutralizing antibody responses alongside T-cell responses. It is clear that differences in the presentation of S protein affect on the nature of the immune system and immune responses [11]. Bachmann et al. (2021) suggest that the topography of SARS-CoV-2 virus, where S protein is present in the surface of the virus and embedded in a fluid membrane, is such that neutralizing epitopes are loosely floating. Another quality of the S protein is the N-linked glycosylation sites, which can mask epitopes. It was also found that glycan masked epitopes are present on the S2 subunit domain but not in the S1 domain, which includes RBD. These findings help to understand the basic biology of the virus and the development of suitable vaccines [12].

This suggested T and B cell mediated acquired immunity is one of the main natural immune therapies against COVID-19. However, very few studies detected MHC class recognition of spike protein epitope. Our earlier study has indicated some spike epitope capable in the generation of MHC responses [13]. Contrast, it is also evident that the antibodies may suppress viral replication through neutralization but might also augment SARS CoV-2 pathogenesis through a process termed antibody-dependent enhancement [14]. Many of the studies elicited vaccine efficacy in polyfunctional S-specific (some time exposed amino acids of S proteins) CD4+ (mainly Th1) and CD8+ T-cell responses, with a T effectors memory phenotype. In some other strategies, DNA/MVA immunizations elicit higher T-cell responses [15].

The current study was aimed to find the role of a small cut fragment of the globally conserved SARS Cov-2 Spike RBD and its in-silico mutants in the generation of higher antigenicity. Further, antigen presentation by HLA and T-cell responses were tested to confirm the antibody-mediated immunogenic responses. Present results are significant to develop a potent and universal vaccination strategy against the COVID-19. Further studies are necessary.

2. Materials and methods

2.1. Structure retrieval, analysis and prediction

The structures of different CUT4 and mutated peptide sequences were taken from our previous analysis. The binding sites of these CUTs with ACE2 Receptor Binding Domain (RBD) were analyzed using PyMol [16]. For an immunoinformatics study, T cell receptor, MHC I and MHC II structures were downloaded from RCSB PDB [17]. Different mutated CUTs were analyzed for T cell receptor interaction analysis using PDB ID 1kgc, 2ak4, 3kxf and 4jrx. For different mutated CUTs interaction with MHC class I, PDB ID 1i4f was used. Different epitopes from mutated CUTs were also analyzed for MHC class I binding using the PDB structures of 6iex, 3x13, 4lnr, 6j1w and 6vb3. Different mutated CUTs interaction with MHC class II was analyzed with PDB ID 4p5m and 3c5j. Different mutated CUTs epitopes were analyzed with 1jk8, 3pdo, 3lqz, 1i4f, 5ni9 and 1BX2 MHC class II structures.

2.2. Docking studies

Protein-Protein docking studies were performed using Haddock 2.4 [18] and PatchDock [19] servers to check the binding affinity with T cell receptor, MHC class I and MHC class II molecules. The results were analyzed using Pymol [16] and the CUTs which had the best binding affinity were chosen. The best mutated structures from the docking result were accepted. The interaction study of the accepted unmutated cut and the predicted epitopes with the best interacting T cell receptor, MHC I and MHC II according to IEDB tools was done using HADDOCK 2.4 [18].

2.3. Epitope study

Epitope prediction using SVMTRIP [20] was done and the probable epitopes were taken for the finalized structure. All the probable epitopes were analyzed using all the IEDB Tools [21]. The T cell Epitope prediction and analysis were performed using IEDB tools [13]. Epitope analysis using SVMTRIP [20] and analyzing results using all IEDB tools were performed. Antigenicity was checked using VaxiJen V2.0 [22] keeping the threshold value as 0.4. A Glycosylation study using NetO-Glyc and NetNGlyc [23] was performed.

The half-life of the selected peptide in blood was predicted using PlifePred server (http://crdd.osdd.net/raghava/plifepred/).

From 2020 to 2022, SARS-CoV 2 typically mutates time by time. So we selected only the VOCs (variant of concerns) sequence. The spike glycoprotein sequence of wild type SARS-CoV 2 was retrieved from the NCBI [National Center for Biotechnology Information] biological database and the other variant spike protein sequences were retrieved from RCSB PDB [protein data bank].

Prediction of epitope positions of different variants was done by using SVMTRIP server. The epitope location, sequence and probable antigenic property score were shown in table form [S Table 5]. For suitable vaccine preparation, we performed sequence alignment of different epitopes which were retrieved from different variants using Clustal X2 software. All the epitopes have conserved sequence, no mutation found [Fig. 6].

We compared all the other proteins of SARS-CoV-2 like nucleocapsid [YP_009724397.2], envelope [YP_009724392.1] and membrane protein [YP_009724393.1] with the spike glycoprotein using SVMTRIP server. The maximum epitope location found in spike glycoprotein as compared with others [S Table 6]. So, spike glycoprotein was a better selection for epitope study. All the sequences of different proteins were retrieved from the NCBI [National Center for Biotechnology Information] biological database.

2.4. Molecular dynamics simulation

The molecular dynamics (MD) method has been widely applied to dissect the microscopic behavior of various biophysical phenomena. Molecular dynamics (MD) is a promising computational approach to investigating dynamical behavior of molecular systems at the atomic level. For molecular dynamics simulation studies we used the new simulator, myPresto/omegagene [24] which uses the Amber99SB-ILDN force field [25] for the protein, and it is freely available at the following URLs: http://presto.protein.osaka-u.ac.jp/myPresto/. This software is a member of a molecular simulation software suite termed myPresto, which was developed and maintained over the past several decades. The myPresto/omegagene software provides functionalities for efficient conformational sampling with explicitly solvated all-atom systems that use our original generalized ensemble methods, such as the virtual-system coupled multicanonical MD method (V-McMD) [26], and the virtual-system coupled adaptive umbrella sampling method (V-AUS) [27]. They are collectively called the virtual-system coupled sampler. The 5000 cycles of conjugated gradient algorithm and 5000 cycles of steepest descent algorithm were used and after 300 k heating, 1 atm pressure was used for system equilibration. Finally, 1000 fs production was run.

For structural stability analysis for docking and simulations, provides an RMSD (Root Mean Square Deviation) or root-mean-square deviation is a standard measure of structural distance between coordinates. The root-mean-square deviation of atomic positions, or simply root-mean-square deviation (RMSD), is the measure of the average distance between the atoms (e.g. backbone atoms of a protein) of superimposed proteins. We examined MD simulation of MHC class II 3LQZ with 3 different CUTs and T-cell 2 molecules (1KGC and 4JRX) with SM3 cut.

2.5. Global mutation study on SARS CoV-2 spike protein

Initially, 186 gene sequences of nCOV2 spike glycoproteins from different countries were retrieved from GISAID database (https://www.gisaid.org/). Then all the sequences were subjected to multiple sequence alignment for the identification of unique sequences using ClustalX2. From all the sequences, 103 sequences were selected with unique sequences and converted to protein sequences using SMS Sequence Manipulation Suite (bioinformatics.org/sms2/translate.html) online server. Among them, 24 unique protein sequences were selected for further analysis. Among them, two sequences were selected that have sequential differences in the Receptor Binding Domain (RBD) in comparison to CUT4. Finally, sequences of CUT4, Model 1 and Model 15 were selected for epitope analysis. Model 1 represented the RBD sequences of Estonia, Latvia, Hong Kong, Costarica, Iran, Mexico, Mongolia, Japan, Italy, Egypt, Ireland, Denmark, Germany, France, India, DRC, Serbia, Pakistan, England and Wuhan (wild type). Whereas, the sequences of Model 15 represented Finland. As per the Epitope analysis and antigenicity of epitope, the structure of selected sequences (Model 1 and Model15) was individually aligned with CUT4 and the respective epitope sequences and locations were analyzed using Pymol [16].

3. Results and discussions

3.1. Different peptide CUT analysis

According to our previous study, a peptide with an amino acid length of 84 (cut 4) has been selected for mutation analysis (data submitted elsewhere). Among 11 different mutations (4 single mutations, 5 double mutations and 2 triple mutations), three were selected for their antigenicity and other immune-reactivity testing. The CUT4 peptide was shown to competitively inhibit the nCOV2-ACE2 attachment in the RB domain (data submitted elsewhere). These peptide CUT4 mutations were further studied for T-Cell reorganization, MHC class I and II attachment, protease cutting-sites, glycosylation effects stability for stable antigenic responses and that was compared with the corresponding mutants.

Sequence	allele	pdb id	score	rank
FERDISTEIY	HLA-B*40:01	6IEX	0.84839	0.09
KYRLFRKSNL	HLA-B*08:01	3X13	0.375792	0.26
NGVEGFNCYF	HLA-B*35:01	4LNR	0.370024	0.35
TRNIDATSTG	HLA-A*30:01	6J1W	0.008784	9
FQPTNGVGYQ	HLA-B*15:01	6VB3	0.712881	0.12
FQPSNGVGYQ	HLA-B*15:01	6VB3	0.760086	0.09

Fig. 1. Interaction of different selected CUT4 mutant (SM3, DM2 and TM2) with MHC class I (1i4f). Interaction showed SM3 and DM2 at the nearby location of 1i4f Epitope presenting site but no interactions were observed (a, b). TM2 formed H-bond with 1i4f (c). Different Epitope from these selected mutants (d) showed that Epitope4 to interact with MHC class I (6j1w). It formed H-bond between TYR7 of Epitope4 and SER445 of MHC I.
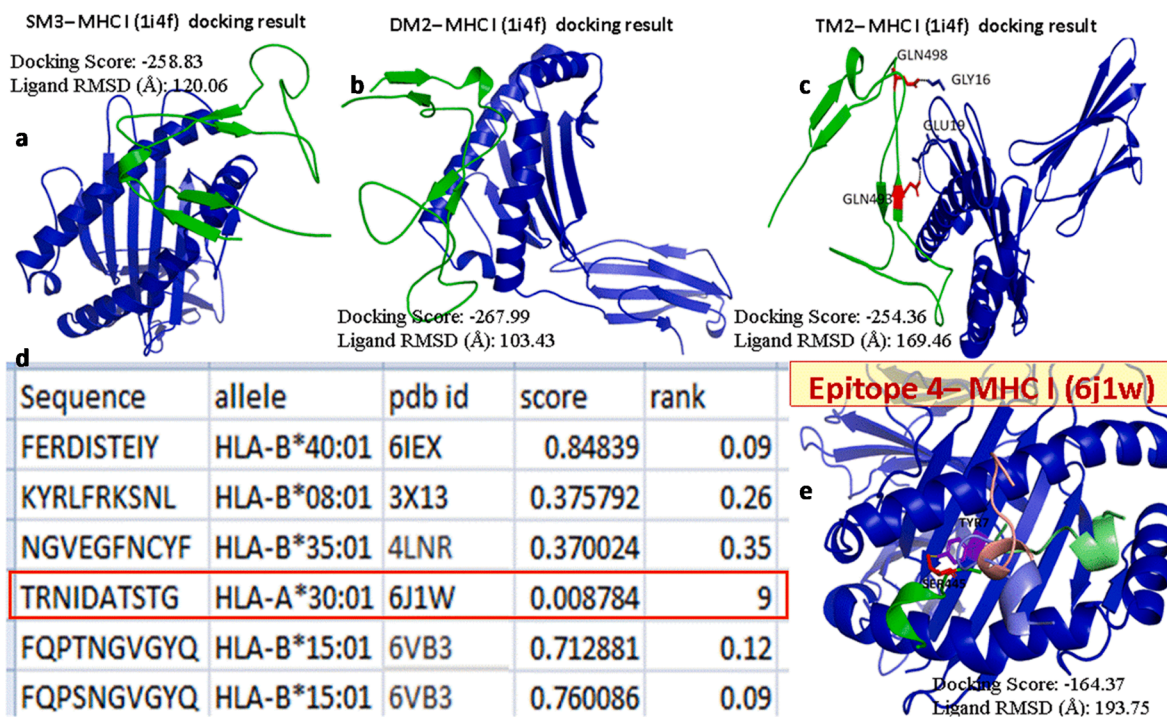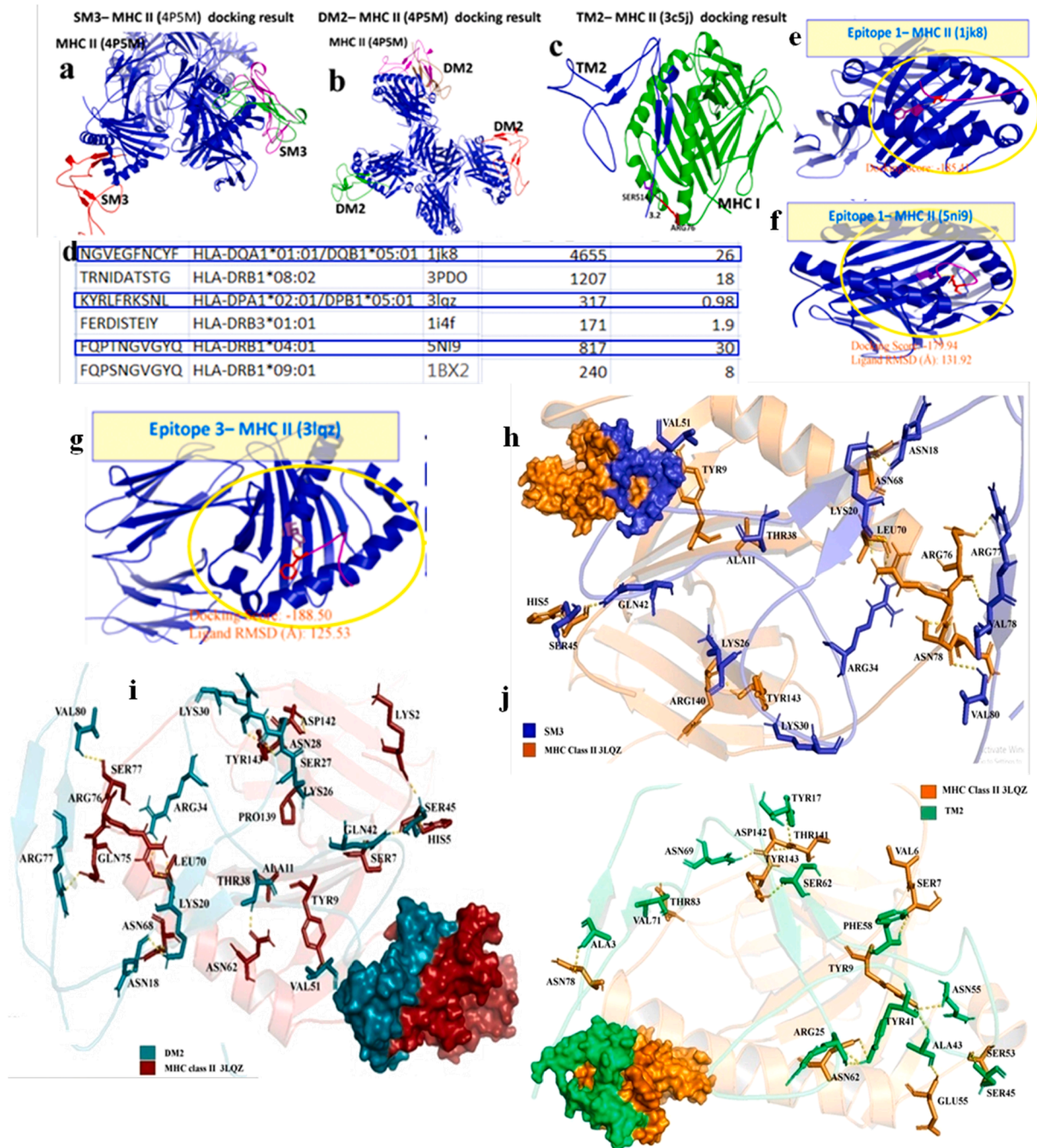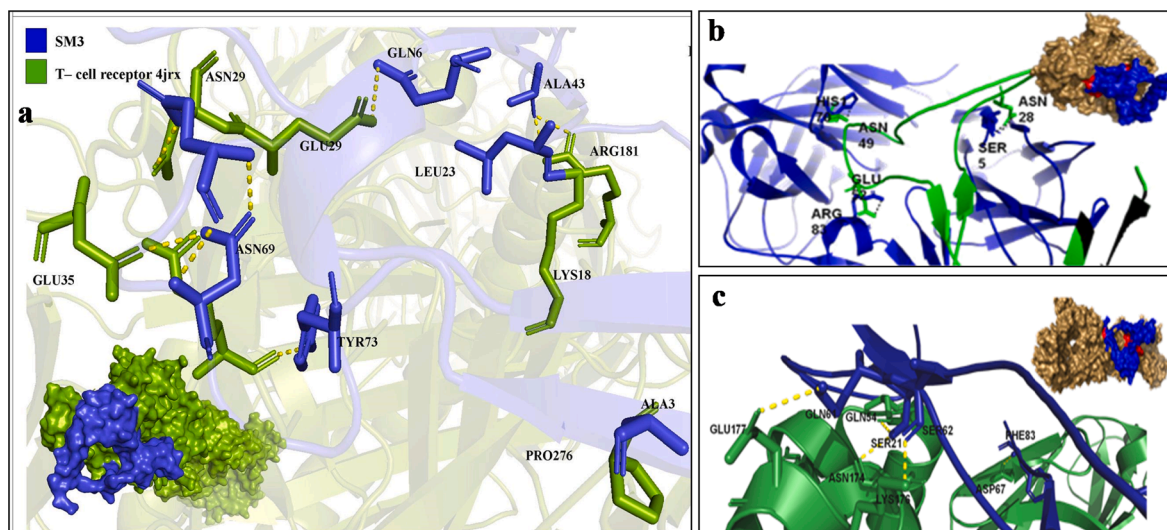
Fig. 2. Interaction of different selected CUT4 mutant (SM3, DM2 and TM2) with MHC class II (4P5M and 3C5J). Interaction showed SM3 and DM2 at the nearby location of 4P5M Epitope presenting site but no interactions were observed (a, b). TM2 formed H-bond with 3C5J (c). Different Epitope from these selected mutants (d) showed that Epitope1, 3 and 5 to interact with MHC class II (1jk8, 3lqz and 5ni9) (e, f, g). Interaction of different selected CUT4 mutant (SM3, DM2 and TM2) with MHC class II (3LQZ) (h, i,j).

3.2. Class I and class II immunogenicity analysis

According to the results, CUT4, CUT4SM3 and CUT4DM2 showed similar results in Class I immunogenicity (S Table 1). Though the length of the peptide remains the same, score values differ due to single, double and triple mutations. The highest score of 0.27813 was found for CUT4 unmutated; CUT4SM3 also showed a positive value of 0.07923. Whereas, score −0.07827 was found for both CUT4DM2 and CUT4TM2 fragments.

For Class II analysis, single and double mutations did not affect the CD4 immunogenicity property. Whereas, CUT4TM2 showed better results than others (S Table 2). The peptide sequence SRLFRKSNLKPFERD showed the highest combined score of 48.23548, and immunogenicity Score of 92.0887. The peptide core sequence SRLFRKSNL of CUT4TM2 showed the highest Median Percentile Rank (among 7-alleles) of 19. For all of these 7 interactive alleles; HLA-DRB1:03:01, HLA-DRB1:07:01, HLA-DRB1:15:01, HLA-DRB3:01:01, HLA-DRB3:02:02, HLA-DRB4:01:01 and HLA-DRB5:01:01, CUT4TM2 showed the highest scores of interactions with 6 except HLA-DRB5:01:01. The Class II molecules present epitopes on CD4 T cells. So, an increase in CD4 immunogenicity confirms the representation of CUT4TM2 epitope SRLFRKSNL by MHC Class II.

Fig. 3. Interaction of different T cell receptors with SM3 mutant cut. SM3 interacted with 1kgc directly and through H-bond interaction.

3.3. Interaction analysis of CUT4 and other mutants with MHC class I and II

The different segments were analyzed for stabilized interaction with different alleles of MHC class II. All the data were selected on the basis of percentile rank below 10% and IC50 value below 50 nM. Here stabilization matrix alignment method or SMM-align method has been used for IC50 value calculation (S Table 3). Only the lower IC50 values were represented in S Table 3, the rest of the data is not shown. The lower IC50 value indicated a higher chance of forming stabilized structures with MHC II alleles. Here, an epitope YNYKYRLFR from CUT4 showed an IC50 values of 30 nM and 28 nM with alleles HLA-DRB1*11:01 and HLA-DRB5*01:01 respectively. Similar results were found with CUT4DM2. For the same epitope, CUT4SM3 showed an IC50 value of 28 nM using only allele HLA-DRB5*01:01. But CUT4TM2 showed a higher value at the range of 44–47 nM with the same allele HLA-DRB5*01:01. The unique epitope ATSTGNYNY in CUT4SM3 showed a lower IC50 value of 30 nM with HLA-DRB5*01:01. Whereas neural network based IC50 analysis, alignment represented that CUT4 and CUT4DM2 had the lowest and similar IC50 values with epitope sequences YRLFRKSNL, YNYKYRLFR and YRLFRKSNL. The Sturniolo method is one of the best methods to predict the best fit between epitope and MHC Class II alleles. The Sturniolo score also represented CUT4 and CUT4DM2 with the similar values where CUT4SM3 showed lower values. According to the current study, all the IC50 values were less than 50 nM. So, from the above analysis it could be predicted that, just like CUT4 unmutated, different epitopes were equally capable of forming a stabilized structure with different alleles of MHC Class II. For MHC Class I analysis, no significant changes were observed (S Table 4). So, the Class II type of interaction and representation was observed in our epitope efficiency analysis.

Molecular docking analysis, CUT4SM3 and CUT4DM2 were found to interact with MHC Class I allele 1i4f with the docking scores of −258.83 and −267.99, and ligand RMSD of 120.06 and 103.43 respectively (Fig. 1). At those positions, CUTs were found to block the MHC Class I peptide presenting site, but no effective bond formation was observed. Whereas, CUT4TM2 was found to interact with 1i4f with the amino acid involvement in H-bond were, GLN498 of CUT4TM2 with GLY16 1i4f and GLN 493 of CUT4TM2 with GLU19 of 1i4f. The docking score and ligand RMSD of that interaction were −254.36 and 169.46 respectively. Whenever the CUTs were analyzed for epitope, 6 fragments were found to interact with different alleles of MHC Class I. Based on the IEDB results, all the fragments were docked with their respective alleles as represented in Fig. 1. This effective H-bonding was observed for epitope

segment 4 with a sequence of TRNIDATSTG. It interacted with HLA-A*30:01 (PDB ID: 6J1W) with a score of 0.008784 and ranked 9. SER445 of CUT4TM2 formed an H-bond with the TYR7 of 6J1W. The docking score was −164.37 and Ligand RMSD was 193.75. The rest of the epitope segments were found exactly at the epitope presenting domain of the respective MHC Class I alleles but no H-bonding was observed.

Similarly, molecular docking with MHC class II was also performed. The CUT4SM3 and CUT4DM2 were found in the epitope presenting domain of MHC class II allele 4P5M without any H-bonding (Fig. 2). Whereas, CUT4TM2 forms an H-bond through SER14 to ARG76 of MHCII (PDB ID: 3c5j). But more affinity was observed for different epitopes as presented in Fig. 1. The THR470 of Epitope 1 formed an H-bond with TYR9 (PDB ID: 1kj8) within 2.9 Å distance, docking score −185.41. PHE490 of Epitope 3 formed an H-bond with TYR9 (PDB ID: 3lqz) within 2.0 Å distance, a docking score −188.50 and a ligand RMSD value of 125.53 and finally, GLN506 of Epitope 5 formed H-bond with GLN9 (PDB ID: 5NI9) within 2.7 Å distance, docking score −179.94 and ligand RMSD value of 131.92.

3.4. T Cell receptor binding analysis of CUT4 and other mutated CUTs

All the CUTs, SM3, DM2 and TM2 showed remarkable affinity for T cell receptors. Several strong bonds as well as H-bonds were formed as presented in Fig. 3. SM3 showed interactions with two different T cell receptors like 1kgc and 4jrx. The H-binding pattern and surface attachment pattern are presented in Fig. 3a & b. Most of the surface area has been exposed to DM2 in the peptide presenting site of receptor 1kgc. TM2 showed the highest representation by 4 different T cell receptors, like 1kgc, 4jrx, 2ak4 and 3kxf. The representation of different CUTs is facilitated by T cell receptors.

Epitope pattern of Receptor Binding Domain (RBD) analysis within various nCOV2 spike proteins found in 186 countries worldwide (Fig. 4). According to CUT4, Model 1 represented Estonia, Latvia, Hong Kong, Costarica, Iran, Mexico, Mongolia, Japan, Italy, Egypt, Ireland, Denmark, Germany, France, India, DRC, Serbia, Pakistan, England and Wuhan (wild type) and Model 15 represented the Finland. Though the sequential diversity was present among CUT4, Model1 and Model15, a sequential similarity was found among 4 epitopes out of 5. The structural alignment also showed similarities among all except epitope 4.

5
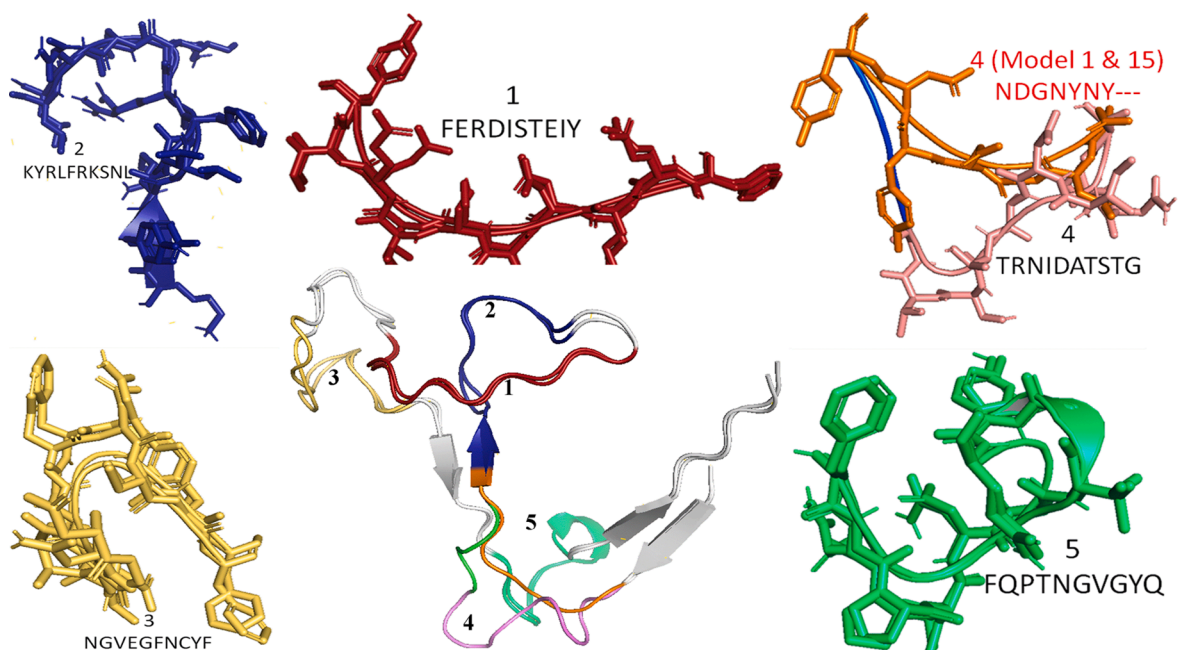
Model 1: representative of Estonia, Latvia, Hong kong, Costarica, Iran, Mexico, Mongolia, Japan, Italy, Egypt, Ireland, Denmark, Germany, France, India, DRC, Serbia, Pakistan , England and Wuhan (wild type)
Model 15: representative of Finland

Rank	Location	Epitope	Score	Antigenicity	Remarks	Model 1 (Representative of most of the countries)	Model 15 (Finland)
1	32 - 41	FERDISTEIY	1.000	**-0.7585**	Probable Non antigen	FERDISTEIY	FERDISTEIY
2	20 - 29	KYRLFRKSNL	0.844	**-0.2129**	Probable Non antigen	-YRLFRKSNL	-YRLFRKSNL
3	49 - 58	NGVEGFNCYF	0.608	**0.6268**	Probable antigen	NGVEGFNCYF	NGVEGFNCYF
4	6 - 15	TRNIDATSTG	0.595	**1.2404**	Probable antigen	NDGNYNY---	NDGNYNY---
5	65 - 74	FQPTNGVGYQ	0.593	**0.4180**	Probable antigen	FQPTNGVGYQ	FQPTNGVGYQ

Fig. 4. Epitope pattern of Receptor Binding Domain (RBD) analysis within various nCOV2 spike proteins found in 186 countries worldwide.

3.5. Dynamics stability analysis of MHC class II and T –cell with cut segments

The conformational dynamics of MHC class II and T-cell with 3 important CUTs were investigated through the root mean square deviation (RMSD) based on carbon atoms. The RMSD value of MHC class II with SM3 cut was estimated for an 1000 fs (fsec) period and is presented in Fig. 5. As can be seen in Fig. 6 from the RMSD plot, we observed the 1KGC T-cell molecule is at 2.2 Å and the 4JRX molecule is at 1.2 Å, which simultaneously increases with time. We used three different RMSD classifications for docking solutions: (a) good solution when RMSD $\leq$ 2.0 Å [28], (b) acceptable solutions when RMSD is between 2.0 and 3.0 Å, and (c) bad solutions when RMSD $\geq$ 3.0 Å. So, from the RMSD plot analysis, we predicted that 4JRX t-cell docking with cut SM3 have the smallest distance and well specific docking binding.

From Fig. 5 graph plot, we observed MHC class II binding with 3 different CUTs (SM3, DM2 and TM2). TM2 has a maximum RMSD value of 1.7 Å, DM2 is at 1.55 Å and SM3 has a minimum RMSD value of 1.5 Å. These variations correspond to structural adjustments to get better binding conformation between molecules and involve binding/unbinding chemical interactions. However, all the three docking structures are good binding stability and determine good stable dynamics of the system.

3.6. In silico analysis of CUT4 and other mutant stability at physiological level

The stability of CUT4 and its different mutants in physiological level has been studied through proteasomal cleavage site analysis, N terminal and C-terminal glycosylation analysis and half life prediction in blood.
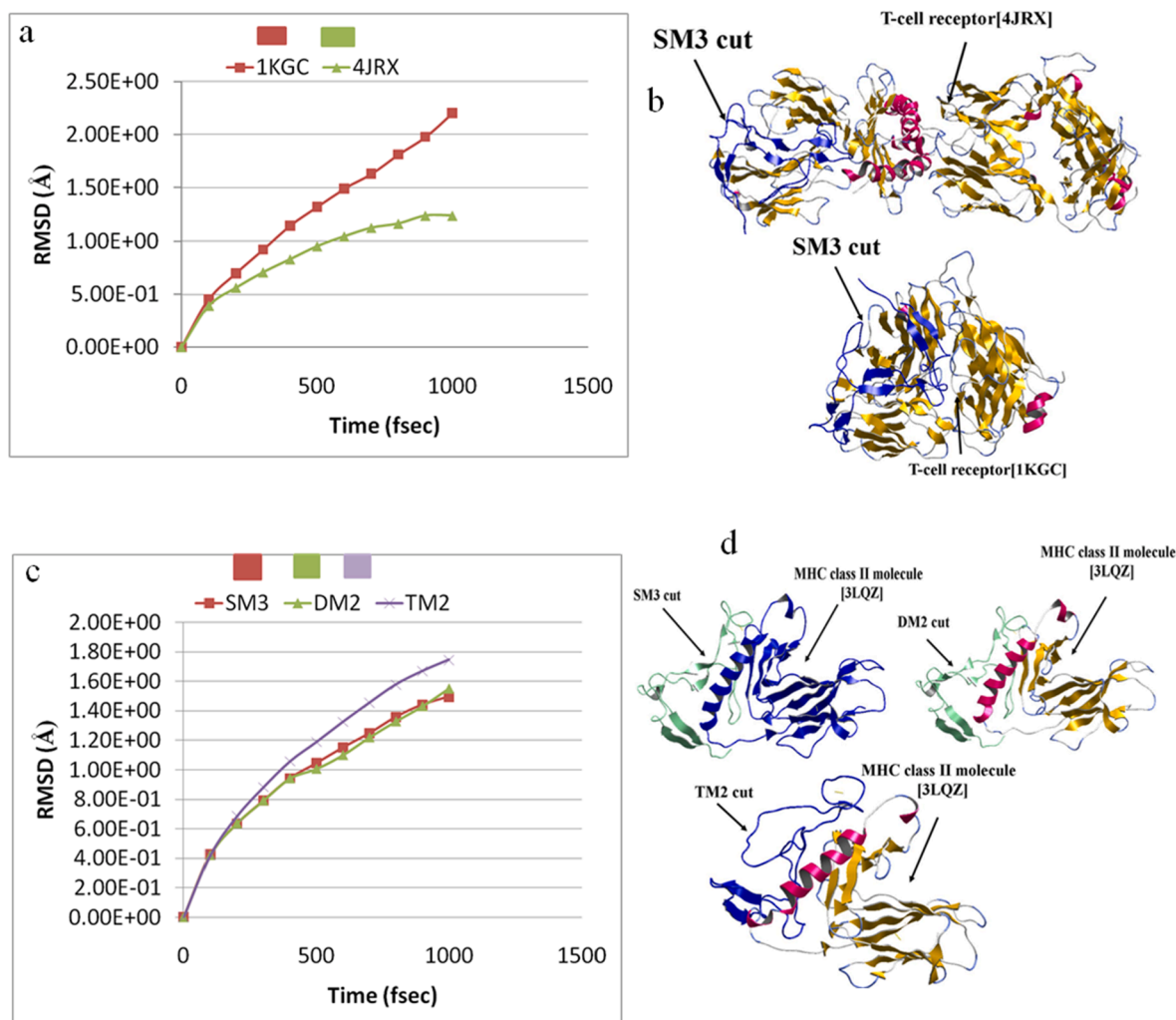
Fig. 5. a) Comparison Cα-RMSD- structural and dynamic stability analysis of T-cell 1KGC (red) with SM3 at the crystal structures (RMSD = 2.2 Å) and 4JRX (green) with SM3 (RMSD = 1.2 Å); b) molecular dynamic posture of T-cell with SM3; c) Comparison Cα-RMSD- structural and dynamic stability analysis of MHC class II SM3 (red), DM2 (green) and TM2 (violet) with 3LQZ; d) molecular dynamic posture of MHC class II with SM3, DM2 and TM2.

There was no proteasomal cleavage sites found in CUT4 before and after mutation according to IEDB analysis. Whereas, according to a neural network-based program, NetCTL, predicted that there were some one Proteasomal processing site but the score of 0.1984 stands insignificant. So, the peptides may remain stable in physiological conditions. All the un-mutated and mutated CUT4 showed no glycosylation at both ends (N and C terminal), except TM2 with C terminal glycosylation. These indicated that these peptides were effectively bound to T cell receptors, and enhanced further immunological reactions as they were not glycosylated.

Finally, the presence of CUT4 epitope was analyzed among different nCOV2 spike glycoproteins from 186 countries worldwide (Fig. 4). Though sequential diversity was present among different nCOV2 spike whole proteins, the Receptor Binding Domain (RBD) showed higher sequential similarity. As a result of which, two sequences and their specific predicted tertiary models were selected as representatives of all 186 countries. Model 1 was representative of all countries except Finland. Model 15 was the representative of Finland. The 5 epitopes present within the CUT4 were also found within Model 1 and Model15. Epitope 1, 2, 3, and 5 were sequentially and structurally the same. The epitope 4 was found to be different. This result was also alike in 103 previously selected protein sequences (S Fig. 1.) According to the antigenicity analysis, epitope 3, 4, and 5 were the potent antigens with antigenicity values of 0.6268, 1.2404, and 0.4639. This epitope-based

analysis could be applicable for worldwide varieties of nCOV2. The effect of mutation was reflected in antigenicity analysis (S Table 5). Comparative studies among CUT4, SM3, DM2 and TM2 represented that, TM2 has the highest antigenicity of 0.6501. Whereas, CUT4, SM3 and DM2 have scores of 0.6478, 0.6373 and 0.6197, respectively. On the other hand, epitope mutation showed significant differences (S Table 5). In this study, both the wild and mutant type epitopes were found to be reactive with the CD4+ cells, which expected cells mediated a higher order of immune responses. SARS CoV-2 develops immunosuppression in autoimmunity, which leads to terminated and/or delayed immunotherapy. Although CD4+ T cells are the key player, but they are ineffective until the naive B-cells upregulation. This is noticed in some severe patients recovering who have a CD20-depleted state and can overcome immunosuppression [6]. A virus mediated couple of vaccination strategies have been tested. The recombinant adenovirus vector has been used to carry SARS-CoV-2) spike glycoprotein and used as a heterologous vaccine that induces strong humoral and cellular immune responses [29]. Vaccine hypersensitivity reactions (VAH) are one of the demerits of this procedure. In some earlier cases of respiratory syndrome (SARS, MERS or respiratory syncytial virus) this adverse effect was noticed [30].

It is speculated that two human adenovirus-vector vaccines, three attenuated vaccines, and one peptide vaccine are in limited use and some more are under clinical trial but having no satisfactory results till
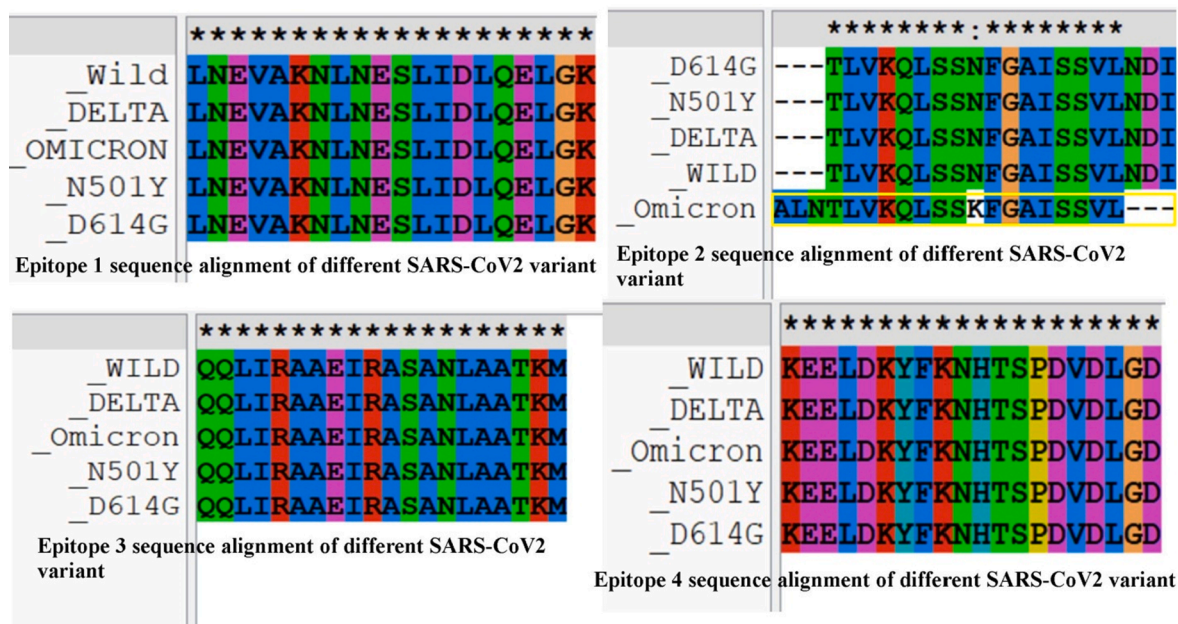
Epitope 1 sequence alignment of different SARS-CoV2 variant

Epitope 2 sequence alignment of different SARS-CoV2 variant

Epitope 3 sequence alignment of different SARS-CoV2 variant

Epitope 4 sequence alignment of different SARS-CoV2 variant

Fig. 6. Different VOCs of SARS-CoV 2 epitope alignment is shown. No mutation found in any epitope positions.

date [31]. Reactivity against the global population a vaccine is required which should comprise B-cell, CTL, HTL epitopes. The epitopes associated with maximum HLA alleles are the requirement of time to ensure viral neutralization. One report claims this outcome is currently under trial [33.]. The CRISPR-mediated genome editing approach has also been used for anti-viral strategies but has yet to take time to be decisive [33]. In our work, we investigated the effects of mutational CUTs spike proteins on MHC Class II and T-cell receptors through molecular dynamic simulation 1000 fs period. Results show that the changes on MD simulation parameters in both MHC class II and T-cell indicate fluctuations and changes in the conformation. RMSD plots of the three complexes of T-cell receptor with SM3, DM2 and TM2 indicated that these systems reached equilibrium quickly and stabilized around 2 Å in all simulations. Through MD analysis, we showed that all the structures remained stable.

In the current study, the globally conserved spike fraction of SARS CoV-2 has been extensively studied for greater immunogenicity and further, in silico, several of its mutants were generated with more immunogenic epitopic fragments which can develop MHC multi HLA-responses and T-cell CD4+ signals. It is hypothesized that multi-epitopic combinations or cocktails can be used for global vaccination strategies. Further studies are necessary in this regard. The present molecular dynamic simulation studies have justifies our primary findings. The global conserved nature of the epitopes has been verified by screening Wuhan-type, D614, Delta and Omicron variants that strengthens our finding. The further attempt in an immuno-pharmacology/vaccines development pharma lab with proper ethical/regulatory settings may help in product development. Unfortunately, it is unavailable in educational-research lab setup. So, further studies are necessary.

Authors contributions

SM – study design, analysis and final manuscript approval, AB – Experiment and data analysis, MK and DS – Experiment and procurement of data, figure preparations.

All Authors read and approved the manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.intimp.2022.108847.

References

[1] O. Sharma, A.A. Sultan, H. Ding, C.R. Triggle, A Review of the Progress and Challenges of Developing a Vaccine for COVID-19, Front. Immunol. 14 (11) (2020) 585354, https://doi.org/10.3389/fimmu.2020.585354. PMID: 33163000; PMCID: PMC7591699.

[2] A.T. Widge, N.G. Rouphael, L.A. Jackson, E.J. Anderson, P.C. Roberts, M. Makhene, J.D. Chappell, M.R. Denison, L.J. Stevens, A.J. Pruijssers, A.B. McDermott, B. Flach, B.C. Lin, N.A. Doria-Rose, S. O'Dell, S.D. Schmidt, K.M. Neuzil, H. Bennett, B. Leav, M. Makowski, J. Albert, K. Cross, V.-V. Edara, K. Floyd, M.S. Suthar, W. Buchanan, C.J. Luke, J.E. Ledgerwood, J.R. Mascola, B.S. Graham, J.H. Beigel, Durability of responses after SARS-CoV-2 mRNA-1273 vaccination, N. Engl. J. Med. 384 (1) (2021) 80–82.

[3] D. Planas, N. Saunders, P. Maes, F. Guivel-Benhassine, C. Planchais, J. Buchrieser, W.H. Bolland, F. Porrot, I. Staropoli, F. Lemoine, H. Péré, D. Veyer, J. Puech, J. Rodary, G. Baele, S. Dellicour, J. Raymenants, S. Gorissen, C. Geenen, B. Vanmechelen, T. Wawina-Bokalanga, J. Martí-Carreras, L. Cuypers, A. Sève, L. Hocqueloux, T. Prazuck, F.A. Rey, E. Simon-Loriere, T. Bruel, H. Mouquet, E. André, O. Schwartz, Considerable escape of SARS-CoV-2 Omicron to antibody neutralization, Nature 602 (7898) (2022) 671–675, https://doi.org/10.1038/s41586-021-04389-z. Epub 2021 Dec 23 PMID: 35016199.

[4] C. Counoupas, M.D. Johansen, A.O. Stella, D.H. Nguyen, A.L. Ferguson, A. Aggarwal, N.D. Bhattacharyya, A. Grey, O. Hutchings, K. Patel, R. Siddiquee, E. L. Stewart, C.G. Feng, N.G. Hansbro, U. Palendira, M.C. Steain, B.M. Saunders, J.K. K. Low, J.P. Mackay, A.D. Kelleher, W.J. Britton, S.G. Turville, P.M. Hansbro, J. A. Triccas, A single dose, BCG-adjuvanted COVID-19 vaccine provides sterilising immunity against SARS-CoV-2 infection, NPJ Vaccines 6 (1) (2021), https://doi.org/10.1038/s41541-021-00406-4.

[5] L.S.F. Frederiksen, Y. Zhang, C. Foged, A. Thakur, The Long Road Toward COVID-19 Herd Immunity: Vaccine Platform Technologies and Mass Immunization

Strategies, Front. Immunol. 21 (11) (2020) 1817, https://doi.org/10.3389/fimmu.2020.01817. PMID: 32793245; PMCID: PMC7385234.

[6] D. Baker, C.A.K. Roberts, G. Pryce, A.S. Kang, M. Marta, S. Reyes, K. Schmierer, G. Giovannoni, S. Amor, COVID-19 vaccine-readiness for anti-CD20-depleting therapy in autoimmune diseases, Clin. Exp. Immunol. 202 (2) (2020) 149–161.

[7] B.F. Haynes, L. Corey, P. Fernandes, P.B. Gilbert, P.J. Hotez, S. Rao, M.R. Santos, H. Schuitemaker, M. Watson, A. Arvin, Prospects for a safe COVID-19 vaccine, Sci. Transl. Med. 12 (568) (2020), https://doi.org/10.1126/scitranslmed.abe0948.

[8] M. Voysey, S.A.C. Clemens, S.A. Madhi, et al., Oxford COVID Vaccine Trial Group. Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK, Lancet. 397 (10269) (2021) 99–111, https://doi.org/10.1016/S0140-6736(20)32661-1. Epub 2020 Dec Erratum in: Lancet. 2021;397(10269): 9PMID: 33306989; PMCID: PMC7723445.

[9] L.A. Jackson, E.J. Anderson, N.G. Rouphael, et al., mRNA-1273 Study Group. An mRNA Vaccine against SARS-CoV-2 - Preliminary Report, N. Engl. J. Med. 383 (20) (2020) 1920–1931, https://doi.org/10.1056/NEJMoa2022483. Epub 2020 Jul 14. PMID: 32663912; PMCID: PMC7377258.

[10] C. Hörner, C. Schürmann, A. Auste, A. Ebenig, S. Muraleedharan, K.H. Dinnon, T. Scholz, M. Herrmann, B.S. Schnierle, R.S. Baric, M.D. Mühlebach, A highly immunogenic and effective measles virus-based Th1-biased COVID-19 vaccine, Proc. Natl. Acad. Sci. U.S.A. 117 (51) (2020) 32657–32666.

[11] F.X. Heinz, K. Stiasny, Distinguishing features of current COVID-19 vaccines: knowns and unknowns of antigen presentation and modes of action, NPJ Vaccines 6 (2021) 104.

[12] M.F. Bachmann, M.O. Mohsen, L. Zha, M. Vogel, D.E. Speiser, SARS-CoV-2 structural features may explain limited neutralizing-antibody responses, NPJ Vaccines 6 (2021) 2.

[13] A. Banerjee, D. Santra, S. Maiti, Energetics and IC50 based epitope screening in SARS CoV-2 (COVID 19) spike protein by immunoinformatic analysis implicating for a suitable vaccine development, J. Transl. Med. 18 (1) (2020) 281, https://doi.org/10.1186/s12967-020-02435-4. PMID: 32650788; PMCID: PMC7351549.

[14] L. Lu, H. Zhang, M. Zhan, J. Jiang, H. Yin, D.J. Dauphars, S.-Y. Li, Y. Li, Y.-W. He, Antibody response and therapy in COVID-19 patients: what can be learned for vaccine development? Sci. China Life Sci. 63 (12) (2020) 1833–1849.

[15] J. García-Arriaza, U. Garaigorta, P. Pérez, A. Lázaro-Frías, C. Zamora, P. Gastaminza, C. Del Fresno, J.M. Casasnovas, C.Ó.S. Sorzano, D. Sancho, M. Esteban, COVID-19 vaccine candidates based on modified vaccinia virus Ankara expressing the SARS-CoV-2 spike induce robust T- and B-cell immune responses and full efficacy in mice, J. Virol. (2021), https://doi.org/10.1128/JVI.02260-20. Epub ahead of print. PMID: 33414159.

[16] G. Janson, C. Zhang, M.G. Prado, A. Paiardini, PyMod 2.0: improvements in protein sequence-structure analysis and homology modeling within PyMOL, Bioinformatics 33 (3) (2017) 444–446, https://doi.org/10.1093/bioinformatics/btw638. PMID: 28158668.

[17] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I. N. Shindyalov, P.E. Bourne, The Protein Data Bank, Nucleic Acids Res. 28 (1) (2000) 235–242, https://doi.org/10.1093/nar/28.1.235. PMID: 10592235; PMCID: PMC102472.

[18] G.C.P. van Zundert, J.P.G.L.M. Rodrigues, M. Trellet, C. Schmitz, P.L. Kastritis, E. Karaca, A.S.J. Melquiond, M. van Dijk, S.J. de Vries, A.M.J.J. Bonvin, The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes, J. Mol. Biol. 428 (4) (2016) 720–725.

[19] D. Schneidman-Duhovny, Y. Inbar, R. Nussinov, H.J. Wolfson, PatchDock and SymmDock: servers for rigid and symmetric docking, Nucleic Acids Res. 33 (Web Server) (2005) W363–W367, https://doi.org/10.1093/nar/gki481.

[20] B. Yao, L. Zhang, S. Liang, C. Zhang, A. Porollo, SVMTriP: A Method to Predict Antigenic Epitopes Using Support Vector Machine to Integrate Tri-Peptide Similarity and Propensity, PLoS ONE 7 (9) (2012) e45152, https://doi.org/10.1371/journal.pone.0045152.

[21] W. Fleri, S. Paul, S.K. Dhanda, S. Mahajan, X. Xu, B. Peters, A. Sette, The Immune Epitope Database and Analysis Resource in Epitope Discovery and Synthetic Vaccine Design, Front. Immunol. 14 (8) (2017) 278, https://doi.org/10.3389/fimmu.2017.00278. PMID: 28352270; PMCID: PMC5348633.

[22] I.A. Doytchinova, D.R. Flower, VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines, BMC Bioinf. 5 (8) (2007) 4, https://doi.org/10.1186/1471-2105-8-4. PMID: 17207271; PMCID: PMC1780059.

[23] R. Azevedo, A.M.N. Silva, C.A. Reis, L.L. Santos, J.A. Ferreira, In silico approaches for unveiling novel glycobiomarkers in cancer, J. Proteomics. 16 (171) (2018) 95–106, https://doi.org/10.1016/j.jprot.2017.08.004. Epub 2017 Aug 4 PMID: 28782717.

[24] K. Kasahara, B. Ma, K. Goto, B. Dasgupta, J. Higo, I. Fukuda, T. Mashimo, Y. Akiyama, H. Nakamura, myPresto/omegagene: a GPU-accelerated molecular dynamics simulator tailored for enhanced conformational sampling methods with a non-Ewald electrostatic scheme, Biophys. Physicobiol. 13 (0) (2016) 209–216, https://doi.org/10.2142/biophysico.13.0_209.

[25] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J.L. Klepeis, R.O. Dror, D. E. Shaw, Improved side-chain torsion potentials for the Amber ff99SB protein force field, Proteins 78 (8) (2010) 1950–1958.

[26] J. Higo, K. Umezawa, H. Nakamura, A virtual-system coupled multicanonical molecular dynamics simulation: Principles and applications to free-energy landscape of protein–protein interaction with an all-atom model in explicit solvent, J. Chem. Phys. 138 (18) (2013) 184106, https://doi.org/10.1063/1.4803468.

[27] J. Higo, B. Dasgupta, T. Mashimo, K. Kasahara, Y. Fukunishi, H. Nakamura, Virtual-system-coupled adaptive umbrella sampling to compute free-energy landscape for flexible molecular docking, J. Comput. Chem. 36 (20) (2015) 1489–1501, https://doi.org/10.1002/jcc.23948.

[28] H. Gohlke, M. Hendlich, G. Klebe, Knowledge-based scoring function to predict protein-ligand interactions, J. Mol. Biol. 295 (2) (2000) 337–356.

[29] D.Y. Logunov, I.V. Dolzhikova, O.V. Zubkova, et al., Safety and immunogenicity of an rAd26 and rAd5 vector-based heterologous prime-boost COVID-19 vaccine in two formulations: two open, non-randomised phase 1/2 studies from Russia, Lancet 396 (10255) (2020) 887–897, https://doi.org/10.1016/S0140-6736(20)31866-3. Epub 2020 Sep 4. Erratum in: Lancet. 2021 Jan 9;397(10269):98. PMID: 32896291; PMCID: PMC7471804.

[30] S.B. Halstead, L. Katzelnick, COVID-19 Vaccines: Should We Fear ADE? J. Infect. Dis. 222 (12) (2020) 1946–1950, https://doi.org/10.1093/infdis/jiaa518. PMID: 32785649; PMCID: PMC7454712.

[31] J. Zhao, S. Zhao, J. Ou, J. Zhang, W. Lan, W. Guan, X. Wu, Y. Yan, W. Zhao, J. Wu, J. Chodosh, Q. Zhang, COVID-19: Coronavirus Vaccine Development Updates, Front. Immunol. 23 (11) (2020), 602256, https://doi.org/10.3389/fimmu.2020.602256. PMID: 33424848; PMCID: PMC7785583.

[33] M.A. Faiq, B-cell engineering: A promising approach towards vaccine development for COVID-19, Med. Hypotheses 144 (2020) 109948, https://doi.org/10.1016/j.mehy.2020.109948.