

# Immune Epitopes of SARS-CoV-2 Spike Protein and Considerations for Universal Vaccine Development

Nicholas Magazine,<sup>\*,1</sup> Tianyi Zhang,<sup>\*,1</sup> Anang D. Bungwon,<sup>\*</sup> Michael C. McGee,<sup>\*</sup> Yingying Wu,<sup>†</sup> Gianluca Veggiani,<sup>\*,‡</sup> and Weishan Huang<sup>\*,§</sup>

<sup>\*</sup>Department of Pathobiological Sciences, School of Veterinary Medicine, Louisiana State University, Baton Rouge, LA; <sup>†</sup>Department of Mathematics, University of Houston, Houston, TX; <sup>‡</sup>Division of Biotechnology and Molecular Medicine, Louisiana State University, Baton Rouge, LA; and <sup>§</sup>Department of Microbiology and Immunology, College of Veterinary Medicine, Cornell University, Ithaca, NY

## ABSTRACT

Despite the success of global vaccination programs in slowing the spread of COVID-19, these efforts have been hindered by the emergence of new SARS-CoV-2 strains capable of evading prior immunity. The mutation and evolution of SARS-CoV-2 have created a demand for persistent efforts in vaccine development. SARS-CoV-2 Spike protein has been the primary target for COVID-19 vaccine development, but it is also the hotspot of mutations directly involved in host susceptibility and virus immune evasion. Our ability to predict emerging mutants and select conserved epitopes is critical for the development of a broadly neutralizing therapy or a universal vaccine. In this article, we review the general paradigm of immune responses to COVID-19 vaccines, highlighting the immunological epitopes of Spike protein that are likely associated with eliciting protective immunity resulting from vaccination in humans. Specifically, we analyze the structural and evolutionary characteristics of the SARS-CoV-2 Spike protein related to immune activation and function via the TLRs, B cells, and T cells. We aim to provide a comprehensive analysis of immune epitopes of Spike protein, thereby contributing to the development of new strategies for broad neutralization or universal vaccination. *ImmunoHorizons*, 2024, 8: 214–226.

## INTRODUCTION

SARS-CoV-2, the etiological pathogen of COVID-19, was responsible for over 760 million cases of illness and 6.9 million deaths as of July 2023 (1). As a single-stranded positive-sense betacoronavirus, SARS-CoV-2 exploits the host machinery for replication (2). Spike protein is a key envelope glycoprotein for SARS-CoV-2 infection. The SARS-CoV-2 Spike protein assembles into trimers (Fig. 1A) and can be broadly split into two subunits, S1 and S2,

divided by a furin cleavage site (Fig. 1B). The S1 subunit is primarily responsible for initial angiotensin-converting enzyme 2 (ACE2) interaction whereas the S2 subunit enables membrane fusion (3). The S1 subunit contains a receptor-binding domain (RBD), which directly binds to ACE2, as well as an N-terminal domain (NTD), which is considered primarily structural in Spike function (4). In contrast, the S2 domain contains two heptad repeat regions (HR1 and HR2) and a fusion peptide (FP) (5). The FP acts immediately after the RBD-mediated Spike–ACE2 interaction by inserting itself

Received for publication January 31, 2024. Accepted for publication February 1, 2024.

**Address correspondence and reprint requests to:** Dr. Weishan Huang, Louisiana State University, 1909 Skip Bertman Drive, Baton Rouge, LA 70803. E-mail address: huang1@lsu.edu

ORCID: 0009-0000-0830-0655 (A.D.B.); 0000-0003-4279-9802 (Y.W.); 0000-0002-4064-0112 (G.V.); 0000-0002-1330-1131 (W.H.).

<sup>1</sup>N.M. and T.Z. contributed equally to this work.

This work was supported in part by grants from the National Institutes of Health (P20 GM130555, R21 AI166898, and R01 AI151139); a Big Idea Research Grant from the Provost's Fund for Innovation in Research at Louisiana State University; and a Research Publication Grant in Engineering, Medicine, and Science from the American Association of University Women.

N.M. and T.Z. analyzed and interpreted data, made the figures, and wrote the manuscript; G.V., Y.W., M.C.M., and A.D.B. reviewed and edited the manuscript; W.H. conceived the research, secured funding, made the figures, and wrote the manuscript. All authors approved the final version of the manuscript.

**Abbreviations used in this article:** ACE2, angiotensin-converting enzyme 2; FP, fusion peptide; HR, heptad repeat region; N, nucleocapsid; NTD, N-terminal domain; RBD, receptor-binding domain; VOC, variant of concern.

This article is distributed under the terms of the [CC BY-NC 4.0 Unported license](https://creativecommons.org/licenses/by-nc/4.0/).

Copyright © 2024 The Authors

into the host membrane, causing a drastic conformational change that brings the HR1 and HR2 domains into proximity with the cell membrane, leading to membrane fusion and viral internalization (6).

Spike protein recognizes and binds to targeted host receptors, predominantly the ACE2 receptor (3, 7), initiating viral entry into host cells via endocytosis. After entry, the viral RNA escapes the host endosome and exploits the translational machinery of the cell for viral protein production (8), followed by assembly and release of the virions from the host cells (9, 10) (Fig. 1B). Because it relies on the translational mechanisms of host cells, Spike protein is subject to post-translational modifications, which may alter Spike-mediated host–virus interactions and host immune responses. Particularly, glycosylation, the addition of glycan molecules to a protein or other molecules, has been demonstrated to affect the interaction of a variety of biomolecules with Spike protein (11).

Since the identification of its role in the initiation of SARS-CoV-2 infection, Spike protein has been extensively investigated as a primary target for vaccine design and therapeutics (5, 12–14). The Spike-based vaccine-elicited host immune cell responses can be broadly divided into three arms: innate immune response, in part as a result of TLR recognition; T cell response; and B cell response (15–20). In brief, innate receptor ligands such as TLR ligands in the vaccine can activate innate immune cells, particularly APCs, which further prime the activation of CD4<sup>+</sup> and CD8<sup>+</sup> T cells. Among the activated CD4<sup>+</sup> T cells, T follicular helper cells contribute to B cell activation, plasma cell maturation, and Ab production, whereas Th1 cells help potentiate APC function, CD8<sup>+</sup> T cell activation, and cytotoxicity. The vaccine-elicited Abs provide protection through the neutralization of viruses, whereas the CTLs eliminate infected host cells to promote clearance of infections (Fig. 2). The immune subset specialization is heavily influenced by the vaccine adjuvants, whereas the specificity of the vaccine-elicited Abs and CTLs is defined largely by the antigenic immune epitopes of the vaccine (21, 22).

Host–virus interactions mediated by Spike protein appear to be the primary driver of SARS-CoV-2 evolution (23, 24). In the 3-y period following the initial outbreak of COVID-19, most circulating SARS-CoV-2 strains have accumulated upward of 40 mutations in their Spike proteins. Given that most vaccinations administered to the global population specifically target Spike proteins, this may in part explain the sharp decrease in vaccine efficacy over a relatively short period of time (25, 26). With cases of breakthrough SARS-CoV-2 infections in vaccinated populations increasing recently (27), it has become apparent that the development of universal vaccines that are resilient against emerging SARS-CoV-2 variants will be required moving forward.

A universal vaccine is largely defined as one that is safe, long-lasting, and effective against multiple variants (28). A detailed understanding of the underlying factors of host immune responses to SARS-CoV-2 is critical for the development of a universal COVID-19 vaccine. In this study, we explore

the characteristics of immunogenic epitopes on Spike protein, aiming to identify regions of the SARS-CoV-2 Spike protein that are of particular interest in the context of universal vaccine development. Specifically, we explore the interactions between Spike immunogenic motifs and immune receptors on APCs, T cells, and B cells and discuss the efficacy and sustainability of specific T and B cell immune epitopes for universal vaccine development.

## MATERIALS AND METHODS

### Genetic entropy analysis

The genetic variability within individual amino acids of SARS-CoV-2 was measured using the Shannon entropy of codons occurring at the indicated positions. The Shannon entropy value was denoted as “genetic entropy” and was retrieved using the NextStrain (nextstrain.org) COVID-19 analysis tool. Data used for this analysis were obtained from the aligned sequence of SARS-CoV-2 genomes deposited in the GISAID (gisaid.org) as of July 2023. For peptide sequences, the average genetic entropy value of codons throughout the peptide was calculated and reported as mean genetic entropy. The genetic entropy of all deposited sequences, and the sequences of the BA.5 Omicron lineage, were reported separately.

### Solvent accessibility analysis

“Solvent accessibility” refers to the solvent-accessible surface area of indicated individual amino acids. This value was calculated using the PyMOL by the “Get\_area” function, which calculates the surface area in square Angstroms of the selection, given with default setting (dot density = 2, solvent radius = 1.4 Å, dot\_solvent = true). This function was performed for each resolved residue within the closed-state SARS-CoV-2 Spike protein (Protein Data Bank identifier 6VXX) (29). For peptides, average solvent accessibility (total solvent accessibility of all resolved residues within the peptide divided by the total number of resolved residues) was denoted as “mean solvent accessibility.”

### Experimentally validated epitope determination and frequency of response

Data from the Immune Epitope Database (IEDB.org) were used for the determination of experimentally validated T and B cell epitopes (30). T cell epitopes were queried using SARS-CoV-2 Spike sequence as the epitope source and searching for epitopes that had positive outcomes in T cell assays. These results were further filtered by having human origin, occurring within SARS-CoV-2 and Spike protein, and having linear epitopes. The database was then queried for patient samples that had been tested for T cell epitope propensity. The frequency of response for each epitope was calculated by dividing the total number of positive tests by the overall test number. Epitopes demonstrating a frequency of response exceeding 80% and having 10 or more deposited tested individuals were reported. Experimentally validated neutralizing B cell epitopes were found by querying the

database for SARS-CoV-2 anti-Spike Abs of human origin that have also been tested for SARS-CoV-2–neutralizing capabilities. The linear epitopes of samples meeting these criteria and showing successful neutralization in over 50% of tests were reported. Calculation of the propensity of occurrence for all binding Abs was performed by querying IEDB for all human-origin Abs demonstrating SARS-CoV-2 Spike binding capacity. Data were retrieved in July 2023.

### ***IgG/IgA overlap calculation***

Data from the IEDB were used to calculate the relative propensity for Ig binding to various Spike regions. A database export containing linear epitopes of human-derived Abs to SARS-CoV-2 Spike between 4–20 amino acids was processed through IgPred Server (iiitd.edu) (31), and the resultant IgG/IgA propensity score was reported. The value for each residue is reported as the mean score for fragments overlapping the target area. Data were retrieved in July 2023.

### ***B cell epitope antigenicity prediction***

B cell epitope antigenicity prediction was performed on the basis of methods used by Sikora et al. (32). Accessibility analysis over the SARS-CoV-2 Spike protein (6VXX) was performed by rigid-body docking of the Fab fragment of Ab CR3022 (33). This analysis was performed on both the native Spike structure and the same Spike structure excluding glycans from the model. The BepiPred 2.0 webserver was used for the prediction of linear B cell epitopes (34). The outputs of these analyses were then multiplied across their respective residues. The maximal output value was then set as 1, and all other values were normalized to the maximum across the entire data set to produce epitope scores for all residues in the predicted epitope regions.

### ***Data availability***

The findings of this study are based on metadata deposited within the GISAID (gisaid.org) and IEDB (iedb.org) databases as of July 2023. The closed state SARS-CoV-2 Spike protein structure (Protein Data Bank identifier 6VXX) was used to present the three-dimensional structure of Spike protein in this article. The PyMOL Molecular Graphics System (version 2.0; Schrödinger, LLC, New York, NY) was used for structural analysis and rendering.

## **RESULTS**

### ***Potential interactions of SARS-CoV-2 Spike protein with toll-like receptors***

The role of Spike protein in the interaction between SARS-CoV-2 and host receptor ACE2 is well characterized (Fig. 1). However, very little is known about Spike interaction with host innate immune receptors. TLRs regulate the activation and localization of cells of the innate immune system, ultimately governing the initial host response to both infection and vaccination. Following SARS-CoV-2 infection or the administration of therapeutics that directly

or indirectly introduce SARS-CoV-2 Spike protein into a host system, an innate immune response is initiated, in part via TLR activity (18, 35) (Fig. 2). Although this reaction is integral to the host immune response, multiple studies have shown excessive TLR activation to be a driver of uncontrolled inflammation and ultimately tissue injury upon host exposure to SARS-CoV-2 (36, 37). In the context of SARS-CoV-2 therapies, cell-surface TLRs that are capable of recognizing glycolipids, proteins, and peptides (such as those in SARS-CoV-2 Spike) are of particular importance. The involvement of TLRs, including TLR1, TLR4, and TLR6, in COVID-19 pathogenesis, has been demonstrated (36–39). TLR2 primarily recognizes lipoproteins, and its role in SARS-CoV-2 Spike interaction has been suggested to be minor and partially dependent on the formation of heterodimers with both TLR1 and TLR6 (40). Patients with severe COVID-19 have been demonstrated to have high serum levels of cytokines downstream of the MyD88 signal cascade (37, 41). Additionally, vaccines with supplemental adjuvants targeting TLR1 and TLR4 have been demonstrated to provide a higher level of protection than controls (42–44).

Structure-based predictions have suggested that multiple Spike residues likely interact with TLR1, TLR4, and TLR6 (Fig. 3A–3C) (45–51 and K. Dick, K. K. Biggar, and J. R. Green, manuscript posted on bioRxiv, DOI: 10.1101/2020.03.29.014381). The sites predicted to be involved in direct interaction with TLRs are predominantly located in the Spike S1 subunit flanking the RBD, including the NTD and regions to the C-terminus of RBD on the S1. Note that there is no preference toward hydrophobic or hydrophilic interaction (Fig. 3D). Most residues predicted to be interacting with TLR1, TLR4, or TLR6 can be found within the Spike S1 subunit, with a particular density within regions 30–117, 204–305, and 536–657. Notably, the Shannon entropy, which represents the degree of genetic variation between all SARS-CoV-2 variants of all residues within this region, is very low, with no single residue exceeding 0.05, even when exclusively considering the Omicron BA.5 lineage (Fig. 3D). Low genetic entropy indicates a low rate of genetic variability within this region, suggesting that these potential TLR-interacting residues are relatively conserved.

The described regions are expected to interact with TLRs and help induce an enhanced initial response of APCs that is conducive to higher vaccine efficacy. However, i.m. injection of the soluble Spike protein does not induce overt systemic or local inflammation (52), suggesting that these predicted TLR-binding regions on Spike protein may not skew TLR activation. However, it is possible that these regions, in alternative forms, may elicit TLR signaling activation, contributing to an enhanced initial response. Even if they are validated as immunogenic, in addition to their potential benefits, these TLR-interacting motifs should also be considered for potential vaccine-associated adverse effects. The inclusion of TLR agonists as adjuvants has been a popular strategy in vaccine development, with numerous demonstrated successes (53, 54). It is, however, critical to understand the agonistic properties of endogenous TLR

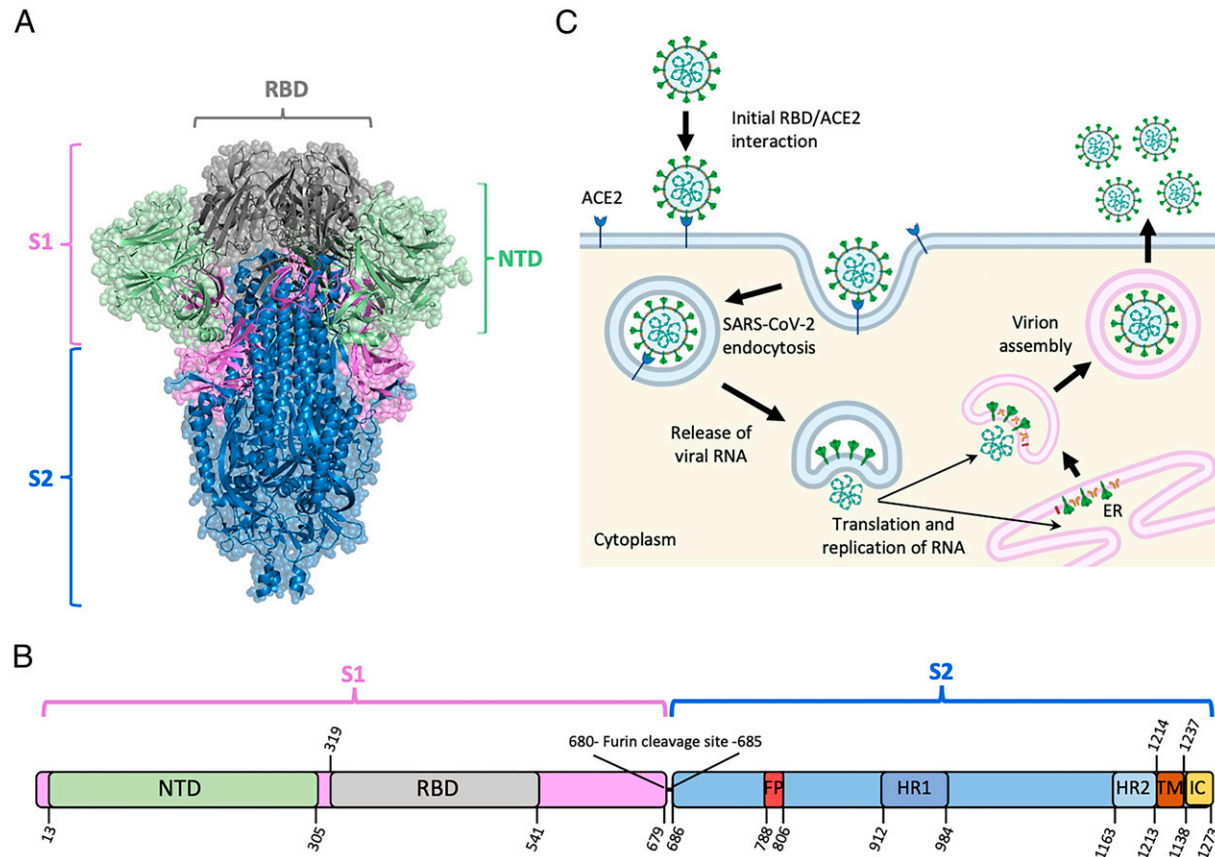


**FIGURE 1. Schematics of Spike structure and SARS-CoV-2 infection.**

(A) Structure of the SARS-CoV-2 Spike protein (Protein Data Bank identifier 6VXX). (B) Diagram of SARS-CoV-2 Spike (S) gene. (C) The primary mechanism of SARS-CoV-2 cell entry and propagation. IC, intracellular domain; TM, transmembrane region.

ligands in vaccine Ags to guide the selection of additional adjuvants for emerging diseases. Although the inclusion of TLR agonists is likely beneficial to vaccine efficacy for most populations, they pose a potential danger of provoking vaccine-associated acute inflammation mediated by an excessive innate immune response.

### T cell epitopes of SARS-CoV-2 Spike protein

Being a primary driver of the adaptive immune response, T cells modulate humoral and cellular immunity to SARS-CoV-2 infection and play a critical role in the maintenance of long-term SARS-CoV-2 immunity (17, 55, 56). Studies have demonstrated the persistence of T cell-mediated immune responses against SARS-CoV-2 variants of concern (VOCs) (57–59), including the Omicron BA.5 strain (60). This cross-reactive T cell immunity may be the primary driver of resistance to these newer strains conferred by early COVID-19 vaccines and/or prior infections (SARS-CoV-2 or other related viruses carrying shared T cell epitopes), suggesting a critical role of T cell immune epitopes for universal vaccine development.

As T cells primarily act to limit viral replication by mediating APC activation, humoral immune response, and destruction of infected cells (Fig. 2), they likely do not prevent SARS-CoV-2 infection. Although the prevention of disease is the ideal outcome

of vaccination, the prevention of severe clinical manifestations and the further spread of disease is also a desirable result (61). This assertion is bolstered by the greatly decreased incidence of hospitalization for vaccinated patients upon clinical presentation of COVID-19 following Omicron infection, despite having received vaccination for non-Omicron SARS-CoV-2 lineages (62–64).

T cells primarily recognize short, linear peptides ranging from 8 to 15 aa in length presented on HLA molecules (65, 66). There are two major HLA groups, class I and class II, which present Ags predominantly to the TCRs of CD8<sup>+</sup> and CD4<sup>+</sup> T cells, respectively (67, 68). Class II HLA complexes are typically found only on APCs and present peptides from exogenous sources, whereas class I HLA complexes are expressed ubiquitously and recognize peptides from endogenous Ags (68). The linear nature of T cell immune epitopes decreases the complexity of vaccine development targeting TCRs, as variation in protein/peptide structure is of minimal concern. However, TCR recognizes peptide Ags only if they are presented by the HLA. HLA genes are expressed in a polymorphic manner, and various populations may be deficient in a specific HLA group (69, 70). It is speculated that the outcome of TCR-stimulating vaccines is heterogeneous, but a potential cross-reactive long-

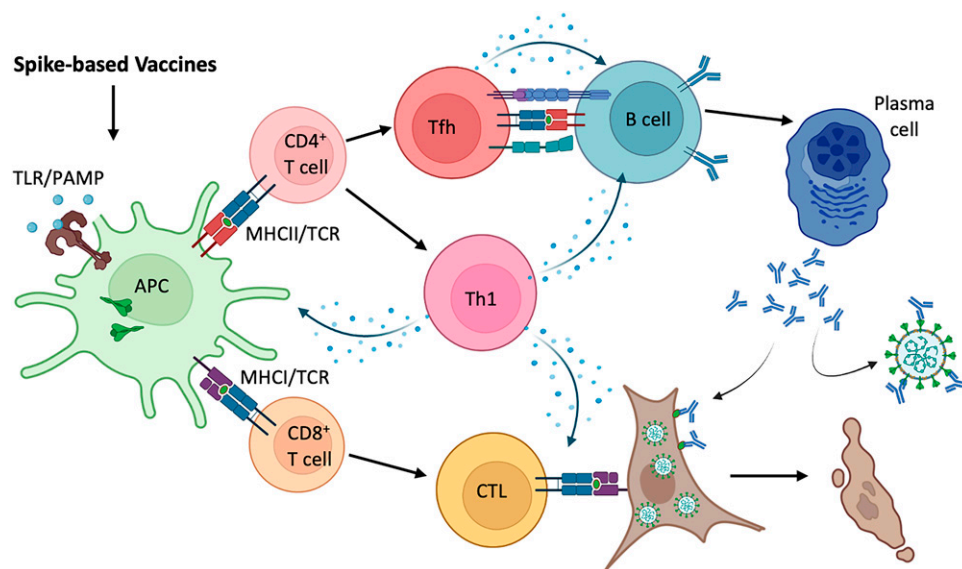


**FIGURE 2. A general overview of host immune responses to Spike-based vaccines.**

Upon host exposure to Spike-based vaccines, APCs sample, process, and present the antigens to  $CD4^+$  and  $CD8^+$  T cells via MHC class II and MHC class I, respectively. Activated  $CD4^+$  T cells can differentiate into follicular helper T (Tfh) and Th1 cells, which further mediate B cell activation and plasma cell maturation. Th1 cells also aid in priming APC activation and maturation to enhance antigen presentation. Activated  $CD8^+$  T cells also receive assistance from Th1 cells to potentiate development into CTLs. While plasma cell-produced Abs can neutralize viruses and bind to infected cells to mediate Ab-associated cell death, CTLs function through direct contact to eliminate infected cells. Therefore, Spike-based vaccine-induced Abs and T cells provide protective immunity against SARS-CoV-2 infections. PAMP, pathogen-associated molecular pattern.

term effect may be achieved in some individuals. A potential strategy to cope with the HLA genetic variance is to identify peptide sequence combinations that can be presented by HLAs expressed by the majority of the global population (with some combinations capable of exceeding 99% global population coverage) (71–74).

Although the binding of a peptide to HLA complex is necessary for the initiation of a T cell response, this is not sufficient. Bioinformatic analysis of T cell epitopes (compiled at the Immune Epitope Database) shows a relatively narrow epitope repertoire for T cells as compared with HLA complexes (30, 75). As of July 2023, a query in IEDB showed a pervasive subset of T cell epitopes (having occurred in 10 or more individuals and taking place in over 80% of individuals tested for these epitopes), as demonstrated by the proportion of their incidence among individuals who have been tested for them (Fig. 4A, 4B). Interestingly, 9 of these 21 peptides show high average levels of genetic entropy (with a Shannon entropy exceeding 0.05), suggesting that T cell interaction could contribute to SARS-CoV-2 evolution (Fig. 4A). These nine epitopes all fall within the NTD and RBD. Structurally, both RBD and NTD are critical in SARS-CoV-2 interactions with host receptors. Mutations in S1 including the NTD and RBD are abundant (23, 24), likely due to their active involvement in host–virus interaction that cause selection pressure imposed by host receptors (notably ACE2) and Abs. Conversely, T cell epitopes in the Spike S2 subunit (residues 751–767, 936–952, 1011–1028, and 1171–1179) exhibit comparatively lower rates of genetic variability, likely because the S2 domain is less engaged in direct interactions between the

virus and host receptor/Abs. Given the relatively low genetic variability, TCR epitopes falling within the S2 region may be target candidates of cross-reactive T cell memory responses for universal vaccine development.

### **B cell epitopes of SARS-CoV-2 Spike protein**

B cells play a critical role in the adaptive immune response arising from SARS-CoV-2 infection (15, 76, 77). B cells have functions including cytokine production, the secretion of neutralizing Abs, and Ag presentation. In the context of COVID-19, the primary known protective role of B cells is the secretion of neutralizing Abs to prevent both initial infection and viral propagation (78). Neutralizing Abs act by binding SARS-CoV-2, primarily Spike protein, and preventing cellular entry of viral particles by sterically blocking their interaction with the host receptors (79, 80). Anti-Spike Abs have been the primary scientific focus of COVID-19 pathology and therapeutic development (81–83). Experimental analysis of patient sera samples has been used to classify Abs that are commonly found in individuals infected with SARS-CoV-2. Bioinformatic analysis of anti-Spike binding Abs revealed over 4000 human Abs with distinct Ag binding sites (30). Although multiple regions of the Spike protein are capable of eliciting Ab responses, only a small subset of these Abs appears to be capable of inhibiting SARS-CoV-2 infection when used in neutralization assays. The identification of neutralizing Ab epitopes, especially broadly neutralizing Ab epitopes, will aid the development of future therapeutics and universal vaccines.

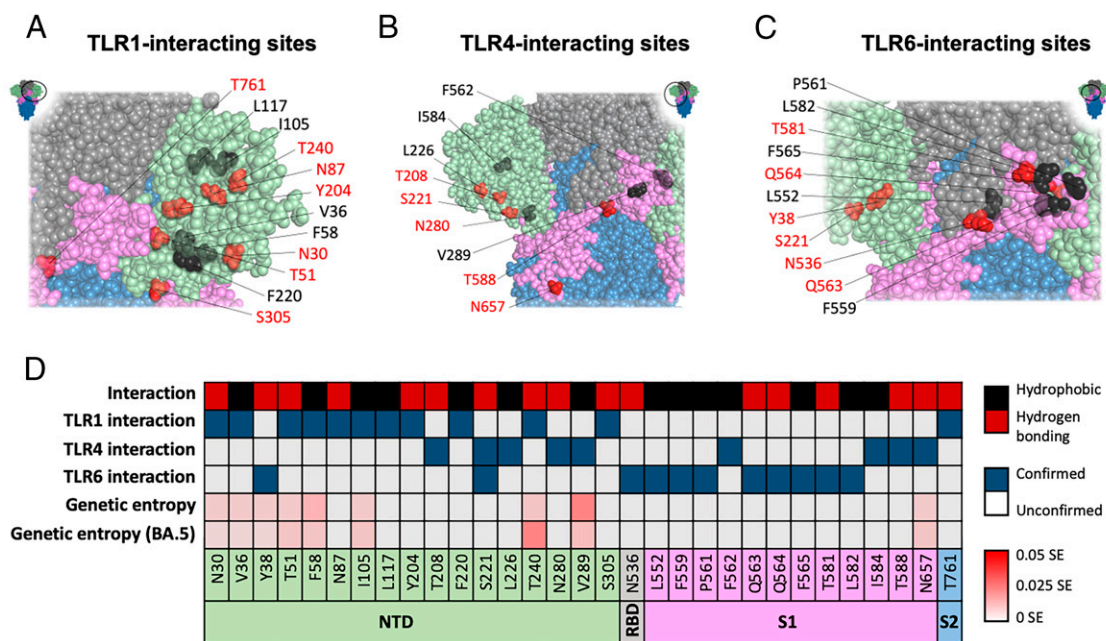


**FIGURE 3. Predicted TLR-interacting sites in SARS-CoV-2 Spike.**

Sites on SARS-CoV-2 Spike protein interacting with TLR1 (A), TLR4 (B), and TLR6 (C). Inset shows full-length Spike protein with a circle indicating the relative position of zoomed-in area. Sites interacting via hydrophobic interaction or hydrogen bonding are displayed in black or red, respectively. (D) List of the TLR-interacting sites, interaction type, TLR interaction, genetic entropy of all reposit sequences, and genetic entropy within Omicron strain BA.5. Blue boxes indicate experimentally confirmed TLR-interacting sites. Sites with Shannon entropy (SE) exceeding 0.05 are considered substantially variable among the described SARS-CoV-2 strains.

To best summarize the SARS-CoV-2 Spike protein epitopes recognized by human Abs, we queried the IEDB database and further cross-referenced our search for Abs whose neutralizing activity was validated via in vitro neutralization assays. This analysis highlighted the distribution and frequency of serum anti-Spike Abs having linear epitopes as well as the position of each recognized epitope (Fig. 5). Notably, a large number of the observed neutralizing Abs targeted regions throughout the RBD of SARS-CoV-2 Spike protein (Fig. 5A). However, this observation may partially result from sampling bias created by higher rates of database deposition for Abs that are found to have neutralizing activity due to their binding to the RBD. Nonetheless, when we compared the relative propensity of regions to which classified serum Abs bind SARS-CoV-2 Spike protein and the locations of Ab binding sites that have been experimentally validated to effectively neutralize SARS-CoV-2 (Fig. 5B), there appear to be only three recurring regions to which patient serum Abs bind and effectively neutralize SARS-CoV-2, the majority of which occur outside of Spike protein's RBD (Fig. 5B). Taking into consideration the genetic entropy and solvent accessibility of these regions, many of the epitopes recognized by patients' Abs have high solvent accessibility and experience high rates of mutation within the RBD (Fig. 5B). This observation closely aligns with the notion that mutations in the RBD of SARS-CoV-2 play a part in immune escape from Ab neutralization and thus viral evolution (23). Interestingly, the neutralizing epitopes of Spike protein that are located

outside the RBD appear to be conserved (low genetic entropy), despite their intermediate to high solvent accessibility. In particular, epitopes in the S2 domain (spanning residues 789–876 and 1140–1166) displayed low genetic entropy and relatively high solvent accessibility and were commonly recognized by patients' Abs, resulting in good neutralization (Fig. 5B, 5C). Therefore, such epitopes would likely be promising targets for universal vaccine development because patient Abs targeting these areas would likely be abundant, highly neutralizing, and less likely to lose effectiveness against newly emerging SARS-CoV-2 lineages.

### Glycosylation in SARS-CoV-2 Spike immune epitopes

Glycosylation is the most common post-translational modification occurring within most organisms. During the process of glycosylation, carbohydrate molecules are covalently attached to the side chains of a number of amino acids (84, 85). Most commonly, this occurs at either oxygen (in Ser, Thr, and Tyr) or nitrogen (in Arg and Asn), resulting in O-glycosylation and N-glycosylation, respectively. In eukaryotes, O-glycosylation occurs primarily in the Golgi apparatus, whereas N-glycosylation occurs primarily in the endoplasmic reticulum (86, 87). The reasons for which a protein may undergo glycosylation are multifaceted. The added carbohydrate motifs can function as adhesion molecules via interaction with cellular lectins, which may ultimately drive the transportation and subcellular location of the modified proteins (88, 89). Glycans may also obscure the presentation of surface molecules, which may alter the antigenicity of

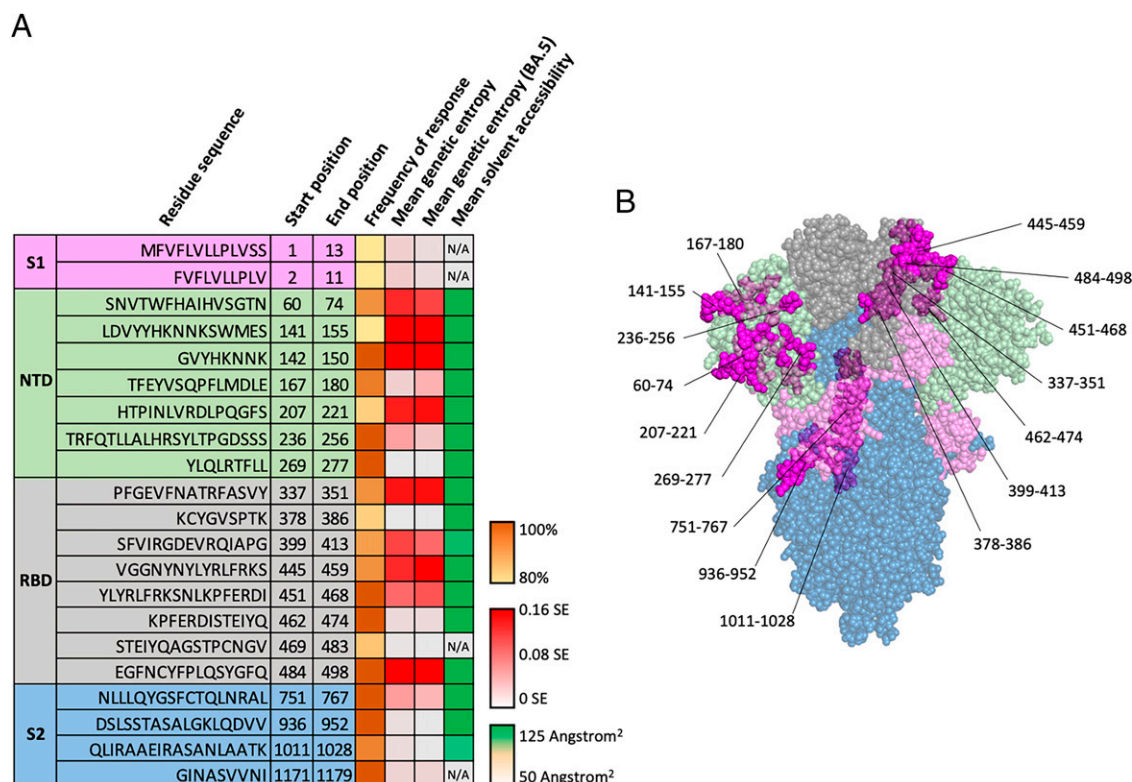


**FIGURE 4. TCR immune epitopes in SARS-CoV-2 Spike.**

(A) List of TCR-interacting epitopes, starting/ending positions, frequency of response as seen in experimental assays (T cell activation or peptide/tetramer binding), mean residue genetic entropy of all reposit sequences, mean residue genetic entropy within Omicron strain BA.5, and mean residue solvent accessibility. The frequency of response represents the proportion of positive assays as compared with the total number of assays for a given epitope (higher being more common) with data obtained from IEDB.org. Genetic entropy was calculated on the basis of the Shannon entropy of GISAID-reposited SARS-CoV-2 sequences and is reported as the mean calculated residue entropy through the given epitope. Solvent accessibility is reported as mean square angstroms of solvent accessibility for residues through each epitope. (B) Sites on the SARS-CoV-2 Spike reported to have TCR interaction. Surface regions are displayed in magenta; some residues are not on the surface and are occluded by the colored domains, becoming purple or violet; some residues cannot be shown because of their inaccessibility on the structure resolved.

proteins and peptides presented by various pathogens (90, 91). Finally, glycosylation can substantially alter the structural presentation of a protein by influencing its local conformation as well as its thermostability (92, 93). During the production of SARS-CoV-2 proteins in host cells (Fig. 1C), the host molecular machinery also heavily glycosylates viral proteins, including Spike protein. In total, the SARS-CoV-2 Spike protein has been predicted to contain as many as 27 glycosites (predominantly N-glycosylation), the majority of which have been confirmed by mass spectroscopy (Fig. 6A, 6B) (11, 94–96).

Because of their size, glycans substantially alter the surface accessibility of the residues to which they are attached. This steric hindrance has major consequences for epitope accessibility and ultimately the manner in which a protein is engaged by the human immune system, especially for Ab binding (97, 98). A comparison of glycosites found on the SARS-CoV-2 Spike protein within immunologically relevant sites reveals that glycosylation occurs at several sites recognized by the immune system.

For example, the N801 glycosite falls within a B cell epitope (along with predicted glycosites S673 and N1158) (Fig. 6A). N657 might interact with TLR4, although this interaction has only been demonstrated computationally (Fig. 3D). Note that the majority of the glycosites (23 of 27) in Spike protein are highly conserved (see genetic entropy, Fig. 6A). This conservation may be a result of these glycosites having structural and/or functional roles within the SARS-CoV-2 Spike protein.

Interestingly, N801 has been consistently confirmed as glycosylated (94) and demonstrated as a potent epitope for Ab recognition, including IgG and IgA (30) (Fig. 6A). Notably, this glycosite is located within the FP (Fig. 6A) and is also proximal to a region that frequently has been targeted by neutralizing Abs in the S2 domain (Fig. 5A, 5B). An important observation is that the immune epitope containing N801 is well conserved (low genetic entropy) in all VOCs, which is likely associated with its location within the FP region of the S2 domain. The FP is required for virus–cell membrane fusion during viral entry

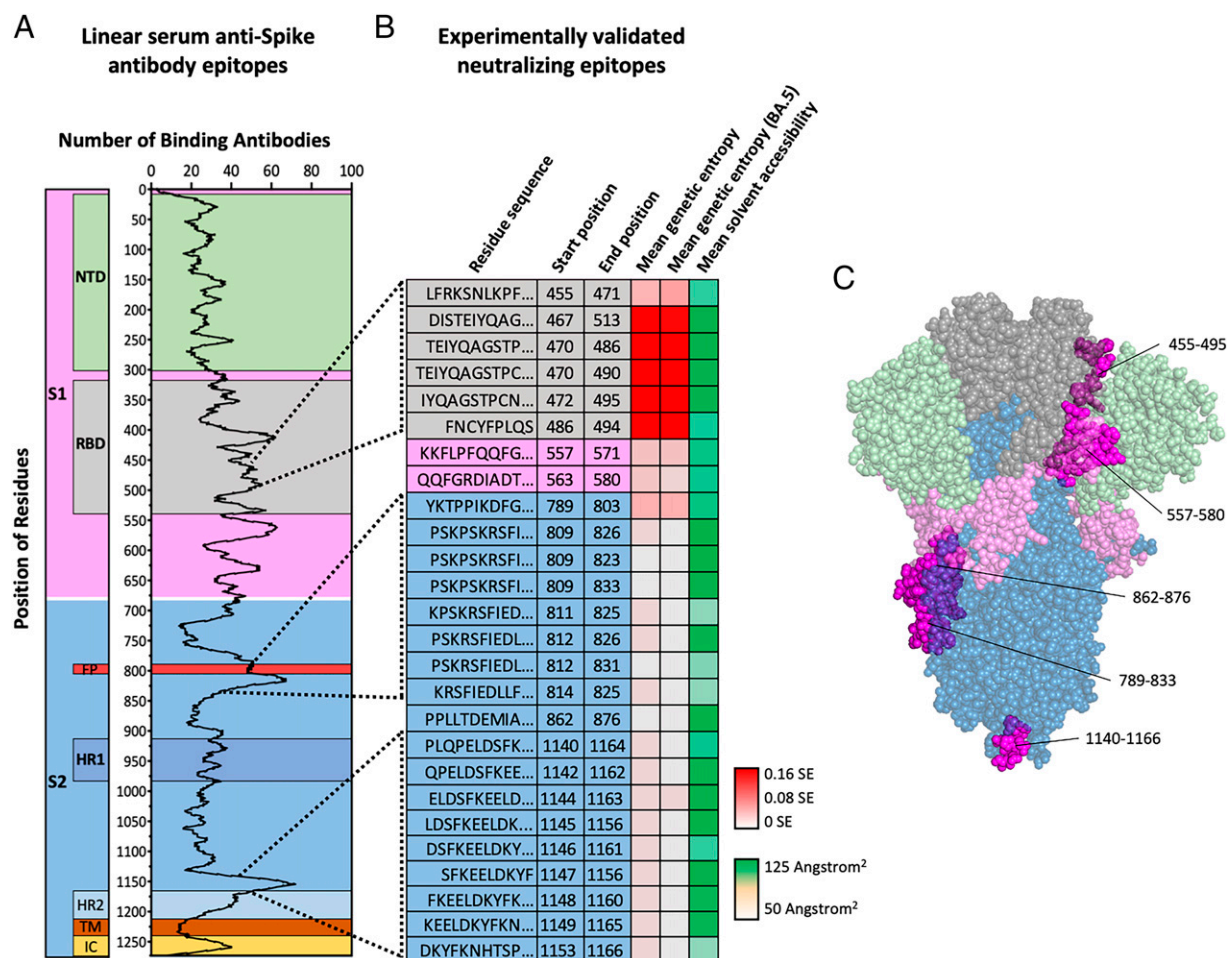


**FIGURE 5. B cell receptor/Ab immune epitopes in SARS-CoV-2 Spike.**

Comparison of linear epitopes (A) observed as binding targets and (B) experimentally validated as neutralizing Ab epitopes by Abs found in database-reposited SARS-CoV-2 patient sera. The first nine amino acid sequences of the epitopes, starting/ending positions, mean genetic entropy of all trains, mean genetic entropy of BA.5, and mean residue solvent accessibility are shown. Genetic entropy was calculated on the basis of the Shannon entropy (SE) of GISAID-reposited SARS-CoV-2 sequences. Values are reported as the mean calculated residue entropy through the given epitope. Solvent accessibility is reported as mean square angstroms of solvent accessibility for residues through each epitope. (C) Sites of linear epitopes of neutralizing Abs on the SARS-CoV-2 Spike protein. Surface regions are displayed in magenta; some residues are not on the surface and are occluded by the colored domains, becoming purple or violet; some residues cannot be shown because of their inaccessibility on the structure resolved.

and functionally conserved in coronaviruses (99), making it a promising candidate epitope for broad neutralization and universal vaccine design.

In addition to residue N801, several other locations should be noted. Glycosylation sites N61, N74, N149, N165, N343, and N1074 overlap with regions known to engage in TCR interaction, whereas S673 and N1158 also fall within regions known to elicit neutralization upon Ab binding (Fig. 6A). Particularly, N1158 falls within a region frequently validated as neutralizing Ab epitope (see Fig. 5B). It should be noted that not all glycosites are confirmed by mass spectrometry, with O-linked glycosites being particularly erratic in their identification by mass spectrometry (95, 96, 100–102).

### Consideration of Spike immune epitopes in universal vaccine development

Taken together, our analyses suggest that regions proximal to N149, N165, N343, N801, and N1158 are immunogenic and relatively conserved. These regions process immune epitopes that can interact with TCRs or Igs and have low genetic entropy, making them promising Ags for universal vaccine development. Moreover, these regions benefit from low genetic entropy, likely resulting from the structural and/or functional roles of these glycosites.

Such a relative conservation would likely reduce the probability of escape from vaccine-induced immunity by future SARS-CoV-2 variants. However, it is possible that this very same

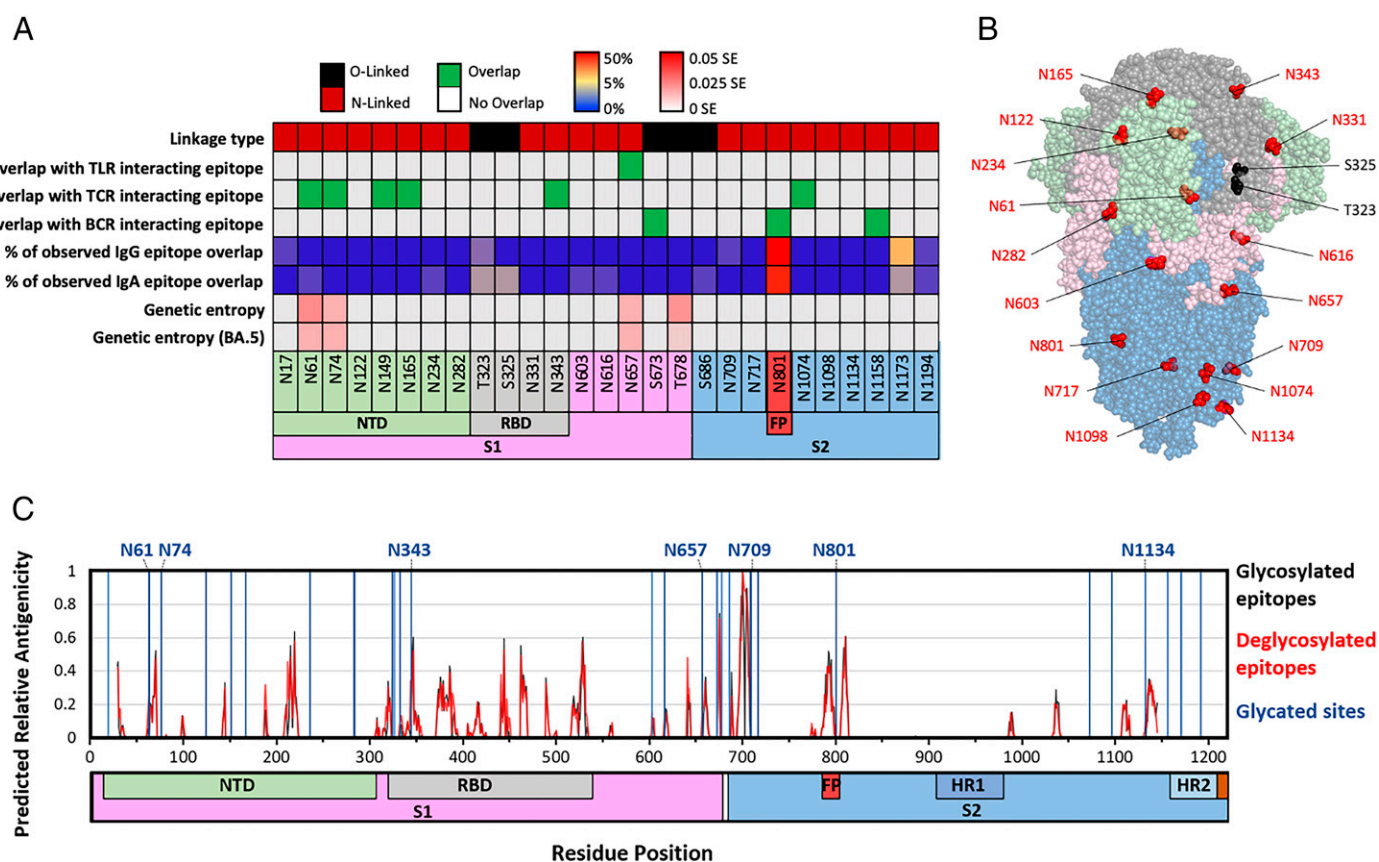


**FIGURE 6. Glycosites and their relationships with immune epitopes in SARS-CoV-2 Spike.**

(A) List of glycosites, linkage type, overlaps with immune epitopes interacting with TLRs, TCRs, or BCRs; proportion of IgG epitope overlap from patient serum; proportion of IgA overlap from patient serum; and genetic entropy of all repositioned sequences or only those of the Omicron BA.5 subset. IgA/IgG overlap is reported as the percentage of observed patient serum Abs that overlap with a given site. (B) Glycosylated sites of the SARS-CoV-2 Spike protein. Some residues cannot be shown because of their absence from the resolved structure. O-linked and N-linked sites are displayed in black and red, respectively. (C) Plots of predicted relative antigenicity over SARS-CoV-2 Spike protein for both glycosylated and deglycosylated Spike. Glycosites are denoted with vertical blue lines.

glycosylation may also play significant roles in host–virus interactions and the immune responses mounted against these sites. Therefore, it is important to further investigate the effects of glycosylation in peptide-based vaccines targeting these regions, because the presence or absence of a glycan molecule may substantially alter T cell recognition and/or Ab binding. Comparative studies of peptide-based vaccines that include glycosylated versus deglycosylated Ags will provide valuable insights into the impact of glycosylation on vaccine efficacy and immune response.

When it comes to selecting vaccine Ags, it is essential to consider both the immunogenicity of the “naked” Ags as well as their post-translational modifications, because they significantly influence the function and interaction of proteins with the immune system. An example of such steric alterations can be observed in B cell epitopes of the SARS-CoV-2 Spike protein. Linear epitope prediction of the SARS-CoV-2 Spike sequence reveals a wide array of potential binding targets. Nonetheless, it is crucial to assess these targets in the context of their surface accessibility, which contributes to their potential for interaction

with immune receptors, especially Abs. The relative capability of a peptide sequence to bind Abs can be approximated by multiplying its relative antigenicity by its relative surface accessibility (32). To further compare the immunogenicity of the glycosylated and nonglycosylated Spike regions, we computed the relative antigenicity of each spike region using methods adapted from Sikora et al. (32). Such computational analyses revealed minimal differences in the predicted immunogenicity of the nonglycosylated versus glycosylated epitopes for eliciting B cell responses (Fig. 6C). However, recent studies comparing deglycosylated and fully glycosylated Spike proteins in immunizations of hamsters and mice found that the deglycosylated Spike protein led to enhanced protective immunity (103). Furthermore, removal of glycosites in the RBD or especially in the S2 domain resulted in the exposure of more conserved epitopes in both regions, eliciting broadly protective immunity (104). These effects may be contributed in part by the better accessibility and/or altered stability of the peptide epitopes in the “naked” Spike protein.

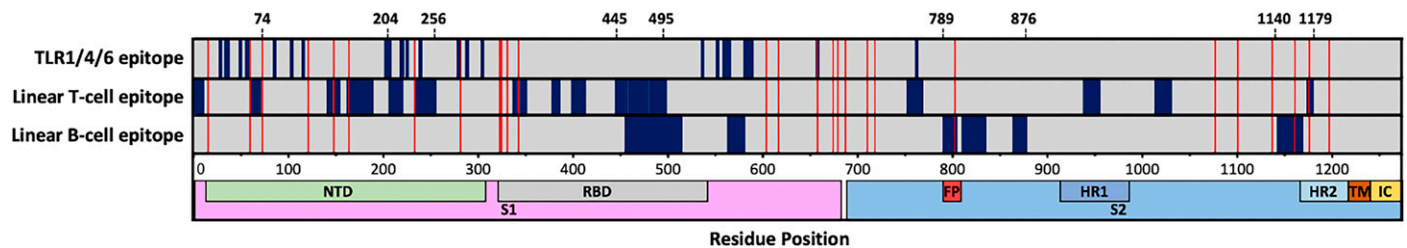


**FIGURE 7. Linear immune epitope map of SARS-CoV-2 Spike protein.**

The predicted TLR-binding regions, T cell epitopes, and B cell linear neutralizing epitopes as summarized in this work are displayed. Glycosites are depicted with vertical red lines.

To further reveal the relative locations of the predicted glycosites, as compared to immunological epitopes summarized in this work, we assembled a linear immune epitope map of SARS-CoV-2 Spike protein (Fig. 7). This map highlights the proximity of the glycosites to the predicted TLR binding sites, TCR epitopes, and linear epitopes for neutralizing Abs.

## DISCUSSION

A robust understanding of SARS-CoV-2 Spike immune epitopes is critical for the development of universal vaccines and broadly effective therapies that target the Spike protein. Our analysis demonstrates that although an immunological response is observed against nearly every portion of SARS-CoV-2 Spike, only a fraction of these sites elicit a response that is effective in preventing SARS-CoV-2 infection. This observation demonstrates that vaccination using proteins or peptides that target the entirety of Spike protein is likely a less optimal solution than the use of pools of immunogenic epitopes. Such a strategy not only may produce an enhanced immune response to current SARS-CoV-2 variants but also will likely provide a stronger and/or more sustainable immunity to emerging SARS-CoV-2 lineages. Our bibliographic and computational analyses indicate that the bulk of a physiologically relevant immune response can likely be elicited by the administration of a subset of Spike protein peptides, such as 1–74, 204–256, 445–495, 789–876, and 1140–1179 (Fig. 7). Regions 1–74 and 204–256 benefit from having interactions with both TLRs and T cells. Regions 445–495 and 1140–1179 contain linear epitopes for neutralizing Abs and TCR interactions. Both 789–876 and 1140–1179 benefit from high solvent accessibility and very low relative genetic entropy (Fig. 5B). The low genetic entropy of these regions suggests that these regions are relatively well conserved among SARS-CoV-2 variants and unlikely to escape immunorecognition even in the case of novel VOCs. The presence of two glycosylation sites within 1140–1179 might make them less susceptible to mutations. Similarly, region 789–876, housing the glycosite N801 and FP domain, is likely to remain well conserved.

SARS-CoV-2 will continue to evolve, and new vaccines to address emerging variants will be needed. This demand will call for the creation of novel vaccines that are effective against

a wide range of SARS-CoV-2 variants as opposed to a specific strain. COVID-19 vaccine-elicited Ab responses are critical for protective immunity, but they wane rapidly within 1 y after vaccination (105). Ag-specific T cells do not block the initial infection, but induction of T cells specific against conserved domains of SARS-CoV-2 Spike protein demonstrated potential in protecting against SARS-CoV-2 variants that escaped Ab neutralization (106). We speculate that administration of peptide pools encompassing regions within the SARS-CoV-2 Spike protein that are conserved and immunogenic (via interactions with TLRs, TCRs, and Abs) will likely produce broad protection from SARS-CoV-2 variants. An important consideration in the design of universal vaccines targeting immunogenic Spike regions would be the potential overlap of these Spike regions with autoantigens. New onset of autoimmune diseases has been reported after COVID-19 infection or vaccination (107, 108), although the underlying mechanisms are unclear. There is an urgent need to further identify and characterize the immune epitopes that overlap with human self-Ags, including self-encoded nucleic acids and proteins, post-translational modifications, and epitopes in the commensal microbiome. Autoantigen-overlapping epitopes should be excluded from future vaccine designs to avoid vaccine-induced adverse effects. Beyond Spike protein, other conserved SARS-CoV-2 proteins such as the nucleocapsid (N) protein may be included as vaccine Ags. Although N protein-specific Abs may not be neutralizing Abs, the Abs and T cells specific against the conserved N protein may help provide broader protection through Ab-dependent cellular cytotoxicity and T cell immunity. Vaccines that contain an enriched pool of conserved and immunogenic epitopes, while avoiding autoantigen overlaps, will likely not only improve clinical outcomes but also provide enhanced protection, decrease healthcare costs, improve patient satisfaction, and limit vaccine-associated autoimmune effects.

## DISCLOSURES

W.H. received research support from MegaRobo Technologies Co., which was not used in this study. The other authors have no financial conflicts of interest.

## ACKNOWLEDGMENTS

We thank members of the Huang and Veggiani labs for helpful discussions.

## REFERENCES

- Dong, E., H. Du, and L. Gardner. 2020. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.* 20: 533–534.
- Alexandersen, S., A. Chamings, and T. R. Bhatta. 2020. SARS-CoV-2 genomic and subgenomic RNAs in diagnostic samples are not an indicator of active replication. *Nat. Commun.* 11: 6059.
- Lan, J., J. Ge, J. Yu, S. Shan, H. Zhou, S. Fan, Q. Zhang, X. Shi, Q. Wang, L. Zhang, and X. Wang. 2020. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* 581: 215–220.
- Zhang, J., T. Xiao, Y. Cai, and B. Chen. 2021. Structure of SARS-CoV-2 spike protein. *Curr. Opin. Virol.* 50: 173–182.
- Wang, M.-Y., R. Zhao, L.-J. Gao, X.-F. Gao, D.-P. Wang, and J.-M. Cao. 2020. SARS-CoV-2: structure, biology, and structure-based therapeutics development. *Front. Cell. Infect. Microbiol.* 10: 587269.
- Jackson, C. B., M. Farzan, B. Chen, and H. Choe. 2022. Mechanisms of SARS-CoV-2 entry into cells. *Nat. Rev. Mol. Cell Biol.* 23: 3–20.
- Yan, R., Y. Zhang, Y. Li, L. Xia, Y. Guo, and Q. Zhou. 2020. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* 367: 1444–1448.
- Harrison, A. G., T. Lin, and P. Wang. 2020. Mechanisms of SARS-CoV-2 transmission and pathogenesis. *Trends Immunol.* 41: 1100–1115.
- V'kovski, P., A. Kratzel, S. Steiner, H. Stalder, and V. Thiel. 2021. Coronavirus biology and replication: implications for SARS-CoV-2. *Nat. Rev. Microbiol.* 19: 155–170.
- Kapoor, K., T. Chen, and E. Tajkhorshid. 2022. Posttranslational modifications optimize the ability of SARS-CoV-2 spike for effective interaction with host cell receptors. *Proc. Natl. Acad. Sci. U. S. A.* 119: e2119761119.
- Zhao, X., H. Chen, and H. Wang. 2021. Glycans of SARS-CoV-2 spike protein in virus infection and antibody production. *Front. Mol. Biosci.* 8: 629873.
- Bangaru, S., G. Ozorowski, H. L. Turner, A. Antanasijevic, D. Huang, X. Wang, J. L. Torres, J. K. Diedrich, J. H. Tian, A. D. Portnoff, et al. 2020. Structural analysis of full-length SARS-CoV-2 spike protein from an advanced vaccine candidate. *Science* 370: 1089–1094.
- Salvatori, G., L. Luberto, M. Maffei, L. Aurisicchio, G. Roscilli, F. Palombo, and E. Marra. 2020. SARS-CoV-2 spike protein: an optimal immunological target for vaccines. *J. Transl. Med.* 18: 222.
- Sternberg, A., and C. Naujokat. 2020. Structural features of coronavirus SARS-CoV-2 spike protein: targets for vaccination. *Life Sci.* 257: 118056.
- Chen, S., F. Guan, F. Candotti, K. Benlagha, N. O. S. Camara, A. A. Herada, L. K. James, J. Lei, H. Miller, M. Kubo, et al. 2022. The role of B cells in COVID-19 infection and vaccination. *Front. Immunol.* 13: 988536.
- Gilbert, S. C. 2012. T-cell-inducing vaccines – what's the future. *Immunology* 135: 19–26.
- Moss, P. 2022. The T cell immune response against SARS-CoV-2. *Nat. Immunol.* 23: 186–193.
- Kayesh, M. E. H., M. Kohara, and K. Tsukiyama-Kohara. 2021. An overview of recent insights into the response of TLR to SARS-CoV-2 infection and the potential of TLR agonists as SARS-CoV-2 vaccine adjuvants. *Viruses* 13: 2302.
- Tejaro, J. R., and D. L. Farber. 2021. COVID-19 vaccines: modes of immune activation and future challenges. *Nat. Rev. Immunol.* 21: 195–197.
- Mistry, P., F. Barmania, J. Mellet, K. Peta, A. Strydom, I. M. Viljoen, W. James, S. Gordon, and M. S. Pepper. 2021. SARS-CoV-2 variants, vaccines, and host immunity. *Front. Immunol.* 12: 809244.
- McKee, A. S., and P. Marrack. 2017. Old and new adjuvants. *Curr. Opin. Immunol.* 47: 44–51.
- Bendelac, A., and R. Medzhitov. 2002. Adjuvants of immunity: harnessing innate immunity to promote adaptive immunity. *J. Exp. Med.* 195: F19–F23.
- Harvey, W. T., A. M. Carabelli, B. Jackson, R. K. Gupta, E. C. Thomson, E. M. Harrison, C. Ludden, R. Reeve, A. Rambaut, S. J. Peacock, D. L. Robertson; COVID-19 Genomics UK Consortium. 2021. SARS-CoV-2 variants, spike mutations and immune escape. *Nat. Rev. Microbiol.* 19: 409–424.
- Magazine, N., T. Zhang, Y. Wu, M. C. McGee, G. Veggiani, and W. Huang. 2022. Mutations and evolution of the SARS-CoV-2 spike protein. *Viruses* 14: 640.
- Lin, D. Y., Y. Gu, B. Wheeler, H. Young, S. Holloway, S. K. Sunny, Z. Moore, and D. Zeng. 2022. Effectiveness of Covid-19 vaccines over a 9-month period in North Carolina. *N. Engl. J. Med.* 386: 933–941.
- Feikin, D. R., M. M. Higdon, L. J. Abu-Raddad, N. Andrews, R. Araos, Y. Goldberg, M. J. Groome, A. Huppert, K. L. O'Brien, P. G. Smith, et al. 2022. Duration of effectiveness of vaccines against SARS-CoV-2 infection and COVID-19 disease: results of a systematic review and meta-regression. *Lancet* 399: 924–944.
- Tan, S. T., A. T. Kwan, I. Rodríguez-Barraquer, B. J. Singer, H. J. Park, J. A. Lewnard, D. Sears, and N. C. Lo. 2023. Infectiousness of SARS-CoV-2 breakthrough infections and reinfections during the Omicron wave. *Nat. Med.* 29: 358–365.
- Erbelding, E. J., D. J. Post, E. J. Stemmy, P. C. Roberts, A. D. Augustine, S. Ferguson, C. I. Paules, B. S. Graham, and A. S. Fauci. 2018. A universal influenza vaccine: the strategic plan for the National Institute of Allergy and Infectious Diseases. *J. Infect. Dis.* 218: 347–354.
- Walls, A. C., Y. J. Park, M. A. Tortorici, A. Wall, A. T. McGuire, and D. Veasler. 2020. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. [Published erratum appears in 2020 *Cell* 183: 1735] *Cell* 181: 281–292.e6.
- Vita, R., S. Mahajan, J. A. Overton, S. K. Dhanda, S. Martini, J. R. Cantrell, D. K. Wheeler, A. Sette, and B. Peters. 2018. The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Res.* 47: D339–D343.
- Gupta, S., H. R. Ansari, A. Gautam, G. P. S., and Raghava, C. Open Source Drug Discovery Consortium. 2013. Identification of B-cell epitopes in an antigen for inducing specific class of antibodies. *Biol. Direct* 8: 27.
- Sikora, M., S. von Bulow, F. E. C. Blanc, M. Gecht, R. Covino, and G. Hummer. 2021. Computational epitope map of SARS-CoV-2 spike protein. *PLoS Comput. Biol.* 17: e1008790.
- Yuan, M., N. C. Wu, X. Zhu, C. D. Lee, R. T. Y. So, H. Lv, C. K. P. Mok, and I. A. Wilson. 2020. A highly conserved cryptic epitope in the receptor binding domains of SARS-CoV-2 and SARS-CoV. *Science* 368: 630–633.
- Jespersen, M. C., B. Peters, M. Nielsen, and P. Marcatili. 2017. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res.* 45: W24–W29.
- Dai, J., Y. Wang, H. Wang, Z. Gao, Y. Wang, M. Fang, S. Shi, P. Zhang, H. Wang, Y. Su, and M. Yang. 2022. Toll-like receptor signaling in severe acute respiratory syndrome coronavirus 2-induced innate immune responses and the potential application value of Toll-like receptor immunomodulators in patients with coronavirus disease 2019. *Front. Microbiol.* 13: 948770.
- Aboudounya, M. M., and R. J. Heads. 2021. COVID-19 and Toll-like receptor 4 (TLR4): SARS-CoV-2 may bind and activate TLR4 to increase ACE2 expression, facilitating entry and causing hyperinflammation. *Mediators Inflamm.* 2021: 8874339.
- Mabrey, F. L., E. D. Morrell, and M. M. Wurfel. 2021. TLRs in COVID-19: how they drive immunopathology and the rationale for modulation. *Innate Immun.* 27: 503–513.
- Khanmohammadi, S., and N. Rezaei. 2021. Role of Toll-like receptors in the pathogenesis of COVID-19. *J. Med. Virol.* 93: 2735–2739.

39. Khan, S., M. S. Shafiei, C. Longoria, J. W. Schoggins, R. C. Savani, and H. Zaki. 2021. SARS-CoV-2 spike protein induces inflammation via TLR2-dependent activation of the NF- $\kappa$ B pathway. *Elife* 10: e68563.
40. Zheng, M., R. Karki, E. P. Williams, D. Yang, E. Fitzpatrick, P. Vogel, C. B. Jonsson, and T. D. Kanneganti. 2021. TLR2 senses the SARS-CoV-2 envelope protein to produce inflammatory cytokines. *Nat. Immunol.* 22: 829–838.
41. Bhatraju, P. K., B. J. Ghassemieh, M. Nichols, R. Kim, K. R. Jerome, A. K. Nalla, A. L. Greninger, S. Pipavath, M. M. Wurfel, L. Evans, et al. 2020. Covid-19 in critically ill patients in the Seattle region - case series. *N. Engl. J. Med.* 382: 2012–2022.
42. Liu, Z. M., M. H. Yang, K. Yu, Z. X. Lian, and S. L. Deng. 2022. Toll-like receptor (TLRs) agonists and antagonists for COVID-19 treatments. *Front. Pharmacol.* 13: 989664.
43. Atalis, A., M. C. Keenum, B. Pandey, A. Beach, P. Pradhan, C. Vantucci, L. O'Farrell, R. Noel, R. Jain, J. Hosten, et al. 2022. Nanoparticle-delivered TLR4 and RIG-I agonists enhance immune response to SARS-CoV-2 subunit vaccine. *J. Control. Release* 347: 476–488.
44. Cheng, K., M. Gao, J. I. Godfroy, P. N. Brown, N. Kastelowitz, and H. Yin. 2015. Specific activation of the TLR1-TLR2 heterodimer by small-molecule agonists. *Sci. Adv.* 1: e1400139.
45. Choudhury, A., and S. Mukherjee. 2020. In silico studies on the comparative characterization of the interactions of SARS-CoV-2 spike glycoprotein with ACE-2 receptor homologs and human TLRs. *J. Med. Virol.* 92: 2105–2113.
46. Ohashi, K., V. Burkart, S. Flohe, and H. Kolb. 2000. Cutting edge: heat shock protein 60 is a putative endogenous ligand of the Toll-like receptor-4 complex. *J. Immunol.* 164: 558–561.
47. Krishnan, M., J. Choi, A. Jang, S. Choi, J. Yeon, M. Jang, Y. Lee, K. Son, S. Y. Shin, M. S. Jeong, and Y. Kim. 2022. Molecular mechanism underlying the TLR4 antagonistic and antiseptic activities of papiliocin, an insect innate immune response molecule. *Proc. Natl. Acad. Sci. U. S. A.* 119: e2115669119.
48. Dyavar, S. R., R. Singh, R. Emani, G. P. Pawar, V. D. Chaudhari, A. T. Podany, S. N. Avedissian, C. V. Fletcher, and D. B. Salunke. 2021. Role of Toll-like receptor 7/8 pathways in regulation of interferon response and inflammatory mediators during SARS-CoV2 infection and potential therapeutic options. *Biomed. Pharmacother.* 141: 111794.
49. Zhou, H., J. Zhao, and S. Perlman. 2010. Autocrine interferon priming in macrophages but not dendritic cells results in enhanced cytokine and chemokine production after coronavirus infection. *mBio* 1: e00219-10.
50. Durai, P., H. J. Shin, A. Achek, H. K. Kwon, R. G. Govindaraj, S. Panneerselvam, D. Yesudhas, J. Choi, K. T. No, and S. Choi. 2017. Toll-like receptor 2 antagonists identified through virtual screening and experimental validation. *FEBS J.* 284: 2264–2283.
51. Bhattacharya, M., A. R. Sharma, B. Mallick, G. Sharma, S. S. Lee, and C. Chakraborty. 2020. Immunoinformatics approach to understand molecular interaction between multi-epitopic regions of SARS-CoV-2 spike-protein with TLR4/MD-2 complex. *Infect. Genet. Evol.* 85: 104587.
52. Goepfert, P. A., B. Fu, A. L. Chabanon, M. I. Bonaparte, M. G. Davis, B. J. Essink, I. Frank, O. Haney, H. Janoszyk, M. C. Keefer, et al. 2021. Safety and immunogenicity of SARS-CoV-2 recombinant protein vaccine formulations in healthy adults: interim results of a randomised, placebo-controlled, phase 1-2, dose-ranging study. *Lancet Infect. Dis.* 21: 1257–1270.
53. Lahiri, A., P. Das, and D. Chakravorty. 2008. Engagement of TLR signaling as adjuvant: towards smarter vaccine and beyond. *Vaccine* 26: 6777–6783.
54. Steinhausen, F., T. Kinjo, C. Bode, and D. M. Klinman. 2011. TLR-based immune adjuvants. *Vaccine* 29: 3341–3355.
55. Le Bert, N., A. T. Tan, K. Kunasegaran, C. Y. L. Tham, M. Hafezi, A. Chia, M. H. Y. Chng, M. Lin, N. Tan, M. Linster, et al. 2020. SARS-CoV-2-specific T cell immunity in cases of COVID-19 and SARS, and uninfected controls. *Nature* 584: 457–462.
56. Mateus, J., A. Grifoni, A. Tarke, J. Sidney, S. I. Ramirez, J. M. Dan, Z. C. Burger, S. A. Rawlings, D. M. Smith, E. Phillips, et al. 2020. Selective and cross-reactive SARS-CoV-2 T cell epitopes in unexposed humans. *Science* 370: 89–94.
57. Tarke, A., C. H. Coelho, Z. Zhang, J. M. Dan, E. D. Yu, N. Methot, N. I. Bloom, B. Goodwin, E. Phillips, S. Mallal, et al. 2022. SARS-CoV-2 vaccination induces immunological T cell memory able to cross-recognize variants from Alpha to Omicron. *Cell* 185: 847–859.e11.e811.
58. Gao, Y., C. Cai, A. Grifoni, T. R. Müller, J. Niessl, A. Olofsson, M. Humbert, L. Hansson, A. Österborg, P. Bergman, et al. 2022. Ancestral SARS-CoV-2-specific T cells cross-recognize the Omicron variant. *Nat. Med.* 28: 472–476.
59. Keeton, R., M. B. Tincho, A. Ngomti, R. Baguma, N. Benede, A. Suzuki, K. Khan, S. Cele, M. Bernstein, F. Karim, et al. 2022. T cell responses to SARS-CoV-2 spike cross-recognize Omicron. *Nature* 603: 488–492.
60. Park, K., S. J. Choi, and E. C. Shin. 2022. Omicron subvariants, including BA.4 and BA.5, substantially preserve T cell epitopes of ancestral SARS-CoV-2. *Immune Netw.* 22: e29.
61. Rennert, L., Z. Ma, C. S. McMahan, and D. Dean. 2022. Effectiveness and protection duration of Covid-19 vaccines and previous infection against any SARS-CoV-2 infection in young adults. *Nat. Commun.* 13: 3946.
62. Christensen, P. A., R. J. Olsen, S. W. Long, R. Snehal, J. J. Davis, M. Ojeda Saavedra, K. Reppond, M. N. Shyer, J. Cambric, R. Gadd, et al. 2022. Signals of significantly increased vaccine breakthrough, decreased hospitalization rates, and less severe disease in patients with coronavirus disease 2019 caused by the Omicron variant of severe acute respiratory syndrome coronavirus 2 in Houston, Texas. *Am. J. Pathol.* 192: 642–652.
63. Chenchula, S., P. Karunakaran, S. Sharma, and M. Chavan. 2022. Current evidence on efficacy of COVID-19 booster dose vaccination against the Omicron variant: a systematic review. *J. Med. Virol.* 94: 2969–2976.
64. Meggiolaro, A., M. Sane Schepisi, S. Farina, C. Castagna, A. Mammone, A. Siddu, P. Stefanelli, S. Boccia, and G. Rezza. 2022. Effectiveness of vaccination against SARS-CoV-2 Omicron variant infection, symptomatic disease, and hospitalization: a systematic review and meta-analysis. *Expert Rev. Vaccines*. 21: 1831–1841.
65. Hemmer, B., T. Kondo, B. Gran, C. Pinilla, I. Cortese, J. Pascal, A. Tzou, H. F. McFarland, R. Houghten, and R. Martin. 2000. Minimal peptide length requirements for CD4<sup>+</sup> T cell clones—implications for molecular mimicry and T cell survival. *Int. Immunol.* 12: 375–383.
66. Stryhn, A., M. Kongsgaard, M. Rasmussen, M. N. Harndahl, T. Osterbye, M. R. Bassi, S. Thybo, M. Gabriel, M. B. Hansen, M. Nielsen, et al. 2020. A systematic, unbiased mapping of CD8<sup>+</sup> and CD4<sup>+</sup> T cell epitopes in yellow fever vaccines. *Front. Immunol.* 11: 1836.
67. Rock, K. L., E. Reits, and J. Neefjes. 2016. Present yourself! By MHC class I and MHC class II molecules. *Trends Immunol.* 37: 724–737.
68. Neefjes, J., M. L. M. Jongmsa, P. Paul, and O. Bakke. 2011. Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nat. Rev. Immunol.* 11: 823–836.
69. Piazza, A., P. Menozzi, and L. L. Cavalli-Sforza. 1980. The HLA-A, B gene frequencies in the world: migration or selection? *Hum. Immunol.* 1: 297–304.
70. Bodmer, J. 1996. World distribution of HLA alleles and implications for disease. *Ciba Found. Symp.* 197: 233–253. ; discussion 253–238.
71. Shi, J., J. Zhang, S. Li, J. Sun, Y. Teng, M. Wu, J. Li, Y. Li, N. Hu, H. Wang, and Y. Hu. 2015. Epitope-based vaccine target screening against highly pathogenic MERS-CoV: an in silico approach applied to emerging infectious diseases. *PLoS One* 10: e0144475.
72. Liu, G., B. Carter, T. Bricken, S. Jain, M. Viard, M. Carrington, and D. K. Gifford. 2020. Computationally optimized SARS-CoV-2 MHC

- class I and II vaccine formulations predicted to target human haplotype distributions. *Cell Syst.* 11: 131–144.e6.
73. Poran, A., D. Harjanto, M. Malloy, C. M. Arieta, D. A. Rothenberg, D. Lenkala, M. M. van Buuren, T. A. Addona, M. S. Rooney, L. Srinivasan, R. B. Gaynor. 2020. Sequence-based prediction of SARS-CoV-2 vaccine targets using a mass spectrometry-based bioinformatics predictor identifies immunogenic T cell epitopes. *Genome Med.* 12: 70.
  74. Smith, C. C., K. S. Olsen, K. M. Gentry, M. Sambade, W. Beck, J. Garness, S. Entwistle, C. Willis, S. Vensko, A. Woods, et al. 2021. Landscape and selection of vaccine epitopes in SARS-CoV-2. *Genome Med.* 13: 101.
  75. Grifoni, A., J. Sidney, Y. Zhang, R. H. Scheuermann, B. Peters, and A. Sette. 2020. A sequence homology and bioinformatic approach can predict candidate targets for immune responses to SARS-CoV-2. *Cell Host Microbe.* 27: 671–680.e2.
  76. Sokal, A., P. Chappert, G. Barba-Spaeth, A. Roeser, S. Fourati, I. Azzaoui, A. Vandenbergh, I. Fernandez, A. Meola, M. Bouvier-Alias, et al. 2021. Maturation and persistence of the anti-SARS-CoV-2 memory B cell response. *Cell* 184: 1201–1213.e14.
  77. Röltgen, K., and S. D. Boyd. 2021. Antibody and B cell responses to SARS-CoV-2 infection and vaccination. *Cell Host Microbe.* 29: 1063–1075.
  78. Andreano, E., I. Paciello, G. Piccini, N. Manganaro, P. Pileri, I. Hyseni, M. Leonardi, E. Pantano, V. Abbiento, L. Benincasa, et al. 2021. Hybrid immunity improves B cells and antibodies against SARS-CoV-2 variants. *Nature* 600: 530–535.
  79. Wang, C., W. Li, D. Drabek, N. M. A. Okba, R. van Haperen, A. D. M. E. Osterhaus, F. J. M. van Kuppeveld, B. L. Haagmans, F. Grosveld, and B.-J. Bosch. 2020. A human monoclonal antibody blocking SARS-CoV-2 infection. *Nat. Commun.* 11: 2251.
  80. Barnes, C. O., C. A. Jette, M. E. Abernathy, K.-M. A. Dam, S. R. Esswein, H. B. Gristick, A. G. Malyutin, N. G. Sharaf, K. E. Huey-Tubman, Y. E. Lee, et al. 2020. SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature* 588: 682–687.
  81. Xiaojie, S., L. Yu, Y. Lei, Y. Guang, and Q. Min. 2021. Neutralizing antibodies targeting SARS-CoV-2 spike protein. *Stem Cell Res.* 50: 102125.
  82. Kim, C., D.-K. Ryu, J. Lee, Y.-I. Kim, J.-M. Seo, Y.-G. Kim, J.-H. Jeong, M. Kim, J.-I. Kim, P. Kim, et al. 2021. A therapeutic neutralizing antibody targeting receptor binding domain of SARS-CoV-2 spike protein. *Nat. Commun.* 12: 288.
  83. Liu, L., P. Wang, M. S. Nair, J. Yu, M. Rapp, Q. Wang, Y. Luo, J. F. W. Chan, V. Sahi, A. Figueroa, et al. 2020. Potent neutralizing antibodies against multiple epitopes on SARS-CoV-2 spike. *Nature* 584: 450–456.
  84. Reily, C., T. J. Stewart, M. B. Renfrow, and J. Novak. 2019. Glycosylation in health and disease. *Nat. Rev. Nephrol.* 15: 346–366.
  85. An, H. J., J. W. Froehlich, and C. B. Lebrilla. 2009. Determination of glycosylation sites and site-specific heterogeneity in glycoproteins. *Curr. Opin. Chem. Biol.* 13: 421–426.
  86. Stanley, P. 2011. Golgi glycosylation. *Cold Spring Harb. Perspect. Biol.* 3: a005199.
  87. Hart, G. W., R. S. Haltiwanger, G. D. Holt, and W. G. Kelly. 1989. Glycosylation in the nucleus and cytoplasm. *Annu. Rev. Biochem.* 58: 841–874.
  88. Marth, J. D., and P. K. Grewal. 2008. Mammalian glycosylation in immunity. *Nat. Rev. Immunol.* 8: 874–887.
  89. Zhang, L., S. Luo, and B. Zhang. 2016. The use of lectin microarray for assessing glycosylation of therapeutic proteins. *MAbs.* 8: 524–535.
  90. Chawla, H., S. E. Jossi, S. E. Faustini, F. Samsudin, J. D. Allen, Y. Watanabe, M. L. Newby, E. Marcial-Juárez, R. E. Lamerton, J. S. McLellan, et al. 2022. Glycosylation and serological reactivity of an expression-enhanced SARS-CoV-2 viral spike mimetic. *J. Mol. Biol.* 434: 167332.
  91. Kumar, S., V. K. Maurya, A. K. Prasad, M. L. B. Bhatt, and S. K. Saxena. 2020. Structural, glycosylation and antigenic variation between 2019 novel coronavirus (2019-nCoV) and SARS coronavirus (SARS-CoV). *Virusdissease* 31: 13–21.
  92. Shental-Bechor, D., and Y. Levy. 2008. Effect of glycosylation on protein folding: a close look at thermodynamic stabilization. *Proc. Natl. Acad. Sci. U. S. A.* 105: 8256–8261.
  93. Bonzom, C., S. Hüttner, E. Mirgorodskaya, S.-L. Chong, S. Uthoff, A. Steinbüchel, R. M. D. Verhaert, and L. Olsson. 2019. Glycosylation influences activity, stability and immobilization of the feruloyl esterase 1a from *Myceliophthora thermophila*. *AMB Express* 9: 126.
  94. Watanabe, Y., J. D. Allen, D. Wrapp, J. S. McLellan, and M. Crispin. 2020. Site-specific glycan analysis of the SARS-CoV-2 spike. *Science* 369: 330–333.
  95. Gong, Y., S. Qin, L. Dai, and Z. Tian. 2021. The glycosylation in SARS-CoV-2 and its receptor ACE2. *Signal Transduct. Target. Ther.* 6: 396.
  96. Sanda, M., L. Morrison, and R. Goldman. 2021. N- and O-Glycosylation of the SARS-CoV-2 spike protein. *Anal. Chem.* 93: 2003–2009.
  97. Reis, C. A., R. Tauber, and V. Blanchard. 2021. Glycosylation is a key in SARS-CoV-2 infection. *J. Mol. Med. (Berl.)* 99: 1023–1031.
  98. Wintjens, R., A. M. Bifani, and P. Bifani. 2020. Impact of glycan cloud on the B-cell epitope prediction of SARS-CoV-2 spike protein. *NPJ Vaccines* 5: 81.
  99. Dacon, C., C. Tucker, L. Peng, C. D. Lee, T. H. Lin, M. Yuan, Y. Cong, L. Wang, L. Purser, J. K. Williams, et al. 2022. Broadly neutralizing antibodies target the coronavirus fusion peptide. *Science* 377: 728–735.
  100. Brun, J., S. Vasiljevic, B. Gangadharan, M. Hensen, A. V. Chandran, M. L. Hill, J. L. Kiappes, R. A. Dwek, D. S. Alonzi, W. B. Struwe, and N. Zitzmann. 2021. Assessing antigen structural integrity through glycosylation analysis of the SARS-CoV-2 viral spike. *ACS Cent. Sci.* 7: 586–593.
  101. Zhang, Y., W. Zhao, Y. Mao, Y. Chen, S. Zheng, W. Cao, J. Zhu, L. Hu, M. Gong, J. Cheng, and H. Yang. 2021. O-glycosylation landscapes of SARS-CoV-2 spike proteins. *Front. Chem.* 9: 689521.
  102. Zhang, Y., W. Zhao, Y. Mao, Y. Chen, S. Wang, Y. Zhong, T. Su, M. Gong, D. Du, X. Lu, et al. 2021. Site-specific N-glycosylation characterization of recombinant SARS-CoV-2 spike proteins. *Mol. Cell. Proteomics* 20: 100058.
  103. Huang, H.-Y., H.-Y. Liao, X. Chen, S.-W. Wang, C.-W. Cheng, M. Shahed-Al-Mahmud, Y.-M. Liu, A. Mohapatra, T.-H. Chen, J. M. Lo, et al. 2022. Vaccination with SARS-CoV-2 spike protein lacking glycan shields elicits enhanced protective responses in animal models. *Sci. Transl. Med.* 14: eabm0899.
  104. Wu, C.-Y., C.-W. Cheng, C.-C. Kung, K.-S. Liao, J.-T. Jan, C. Ma, and C.-H. Wong. 2022. Glycosite-deleted mRNA of SARS-CoV-2 spike protein as a broad-spectrum vaccine. *Proc. Natl. Acad. Sci. U. S. A.* 119: e2119995119.
  105. Levin, E. G., Y. Lustig, C. Cohen, R. Fluss, V. Indenbaum, S. Amit, R. Doolman, K. Asraf, E. Mendelson, A. Ziv, et al. 2021. Waning immune humoral response to BNT162b2 Covid-19 vaccine over 6 months. *N. Engl. J. Med.* 385: e84.
  106. Kingstad-Bakke, B., W. Lee, S. S. Chandrasekar, D. J. Gasper, C. Salas-Quinchucua, T. Cleven, J. A. Sullivan, A. Talaat, J. E. Osorio, and M. Suresh. 2022. Vaccine-induced systemic and mucosal T cell immunity to SARS-CoV-2 viral variants. *Proc Natl Acad Sci U S A* 119: e2118312119.
  107. Chen, Y., Z. Xu, P. Wang, X. M. Li, Z. W. Shuai, D. Q. Ye, and H. F. Pan. 2022. New-onset autoimmune phenomena post-COVID-19 vaccination. *Immunology* 165: 386–401.
  108. Gracia-Ramos, A. E., E. Martin-Nares, and G. Hernández-Molina. 2021. New onset of autoimmune diseases following COVID-19 diagnosis. *Cells* 10: 3592.