The SARS-CoV-2 Spike Glycoprotein Biosynthesis, Structure, Function, and Antigenicity: Implications for the Design of Spike-Based Vaccine Immunogens

Liangwei Duan [1,2], Qianqian Zheng [1,2], Hongxia Zhang [1,2], Yuna Niu [1,2], Yunwei Lou [1,2] and Hui Wang [1,2]*

[1] Henan Key Laboratory of Immunology and Targeted Drugs, School of Laboratory Medicine, Xinxiang Medical University, Xinxiang, China, [2] Henan Collaborative Innovation Center of Molecular Diagnosis and Laboratory Medicine, Xinxiang Medical University, Xinxiang, China

The ongoing pandemic of coronavirus disease 2019 (COVID-19), caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), poses a grave threat to global public health and imposes a severe burden on the entire human society. Like other coronaviruses, the SARS-CoV-2 genome encodes spike (S) glycoproteins, which protrude from the surface of mature virions. The S glycoprotein plays essential roles in virus attachment, fusion and entry into the host cell. Surface location of the S glycoprotein renders it a direct target for host immune responses, making it the main target of neutralizing antibodies. In the light of its crucial roles in viral infection and adaptive immunity, the S protein is the focus of most vaccine strategies as well as therapeutic interventions. In this review, we highlight and describe the recent progress that has been made in the biosynthesis, structure, function, and antigenicity of the SARS-CoV-2 S glycoprotein, aiming to provide valuable insights into the design and development of the S protein-based vaccines as well as therapeutics.

Keywords: SARS-CoV-2, spike glycoprotein, receptor-binding domain, synthesis, structure, membrane fusion, neutralizing antibodies, immunogen design

INTRODUCTION

The coronavirus disease 2019 (COVID-19) global pandemic represents an unprecedented public health, social and economic challenge (1, 2). The etiological agent of COVID-19 is a new member of the Coronaviridae family that is closely related to severe acute respiratory syndrome coronavirus (SARS-CoV) and was recently referred to as SARS-CoV-2 by the Coronavirus Study Group of the International Committee on Taxonomy of Viruses (3). The virus has spread rapidly and sustainably around the global resulting in over twenty-one million cases and more than 750,000 deaths as of August 15, 2020 (4).

Coronaviruses (CoVs) are enveloped positive-sense RNA viruses (5). Enveloped CoVs entering host cells and initiating infection is achieved through the fusion of viral and cellular membranes

(6, 7). Membrane fusion is mediated by the large type I transmembrane S glycoprotein on the viral envelope and the cognate receptor on the surface of host cells (8–10). The surface-exposed location of the S glycoprotein not only allows it to carry out membrane fusion but also renders it a direct target for host immune responses, making it the major target of neutralizing antibodies (11). Because of its central roles in viral infection and eliciting protective humoral and cell-mediated immune responses in hosts during infection (10), the S protein is the primary target for vaccine design as well as antiviral therapeutics (12).

Here, we provide a comprehensive overview of the wealth of research related to the SARS-CoV-2 S glycoprotein biosynthesis, structure, function, and antigenicity, aiming to provide useful insights into the design and development of the S protein-based vaccines as well as therapeutics to prevent or treat the ongoing global spread of SARS-CoV-2/COVID-19.

SYNTHESIS, PROCESSING AND TRAFFICKING OF THE SARS-COV-2 S GLYCOPROTEIN

The SARS-CoV-2 S glycoprotein is synthesized as a 1273-amino acid polyprotein precursor on the rough endoplasmic reticulum (RER) (**Figure 1**) (13). The unprocessed precursor harbors an endoplasmic reticulum (ER) signal sequence located at the N terminus, which targets the S glycoprotein to the RER membrane and is removed by cellular signal peptidases in the lumen of the ER (14, 15). A single stop-transfer, membrane-spanning sequence located at the C terminus of the S protein prevents it from being fully released into the lumen of the ER and subsequent secretion from the infected cell (16, 17). Co-translationally, N-linked, high-mannose oligosaccharide side chains are added during synthesis (18, 19). Shortly after synthesis, the S glycoprotein monomers trimerize, which might be thought to facilitate the transport from the ER to the Golgi complex. Once in the Golgi complex, most of the high-mannose oligosaccharide side chains are modified to more complex forms (20, 21), and O-linked oligosaccharide side chains are also added (22, 23).

In the trans-Golgi network, the SARS-CoV-2 S glycoprotein is proteolytically cleaved by cellular furin or furin-like proteases at the S1/S2 cleavage site, comprising multiple arginine residues that are not found in the closely related SARS-CoV (24, 25). Cleavage at the S1/S2 site yields a surface subunit S1, which attaches the virus to the host cell surface receptor, and a transmembrane subunit S2, which mediates the fusion of viral and host cell membranes (10). The S1 and S2 subunits remain associated through noncovalent interactions in a metastable prefusion state (11). Furin-like cleavage is essential for the S-protein mediated cell-cell fusion and viral infectivity, and is required for efficient SARS-CoV-2 infection of human lung cells (24) and airway epithelial cells (26).

Following cleavage, an ER retrieval signal (ERRS) consisting of a conserved KxHxx motif (27) located at the extreme C terminus ensures that the mature SARS-CoV-2 S protein

accumulates near the ER-Golgi intermediate compartment (ERGIC) (27, 28), where driven by interactions with another structural protein, the membrane (M) protein, the S protein participates in virus particle assembly and is incorporated into virus envelope (**Figure 1**) (29, 30). Besides, a fraction of mature SARS-CoV-2 S proteins travel through the secretory pathway to the plasma membrane, where they can mediate fusion of infected with uninfected cells to form multinucleated giant cells (syncytia) (24, 31). This may allow direct spreading of the virus between cells and potentially alter the virulence of SARS-CoV-2 (24).

Notably, a deletion of ~20 amino acid containing the ERRS from the cytoplasmic tail of the SARS-CoV-2 S protein has been shown to increase the infectivity of single-cycle vesicular stomatitis virus (VSV)-S pseudotypes (9) and replication-competent recombinant VSVs bearing the S glycoprotein (32, 33), which likely could be translated to single-cycle human immunodeficiency virus (HIV)-S or other retrovirus-S pseudotypes straightforward (33). Presumably, this deletion may enhance the cell surface expression of the SARS-CoV-2 S glycoprotein (32), thereby facilitating the S protein incorporation into pseudovirions and replication-competent virions.

SARS-COV-2 S PROTEIN STRUCTURE AND FUNCTION

As mentioned above, the SARS-CoV-2 S glycoprotein plays pivotal roles in viral infection and pathogenesis. Mature S glycoprotein on the viral surface is a heavily glycosylated trimer, each protomer of which is composed of 1260 amino acids (residues 14-1273) (**Figure 2A**). The surface subunit S1 is composed of 672 amino acids (residues 14–685) and organized into four domains: an N-terminal domain (NTD), a C-terminal domain (CTD, also known as the receptor-binding domain, RBD), and two subdomains (SD1 and SD2) (**Figure 2A**) (34). The transmembrane S2 subunit is composed of 588 amino acids (residues 686-1273) and contains an N-terminal hydrophobic fusion peptide (FP), two heptad repeats (HR1 and HR2), a transmembrane domain (TM), and a cytoplasmic tail (CT), arranged as FP-HR1-HR2-TM-CT (**Figure 2A**) (34).

As a typical class I viral fusion protein (35), the SARS-CoV-2 S glycoprotein shares common structural, topological and mechanistic features with other class I fusion proteins, including HIV envelope (Env) glycoprotein and influenza virus haemagglutinin (HA) (36–38). Like other class I viral fusion proteins, the SARS-CoV-2 S glycoprotein is also a conformational machine that mediates viral entry by rearranging from a metastable unliganded state, through a pre-hairpin intermediate state, to a stable postfusion state (38, 39). Since the first genome sequence of SARS-CoV-2 became publicly available (40), a number of structures have been determined for the SARS-CoV-2 S glycoprotein trimer fragments in both the prefusion and postfusion states (**Figures 2B–D**) (11, 34, 41).

The overall architecture of the prefusion SARS-CoV-2 S ectodomain stabilized by two consecutive proline mutations in
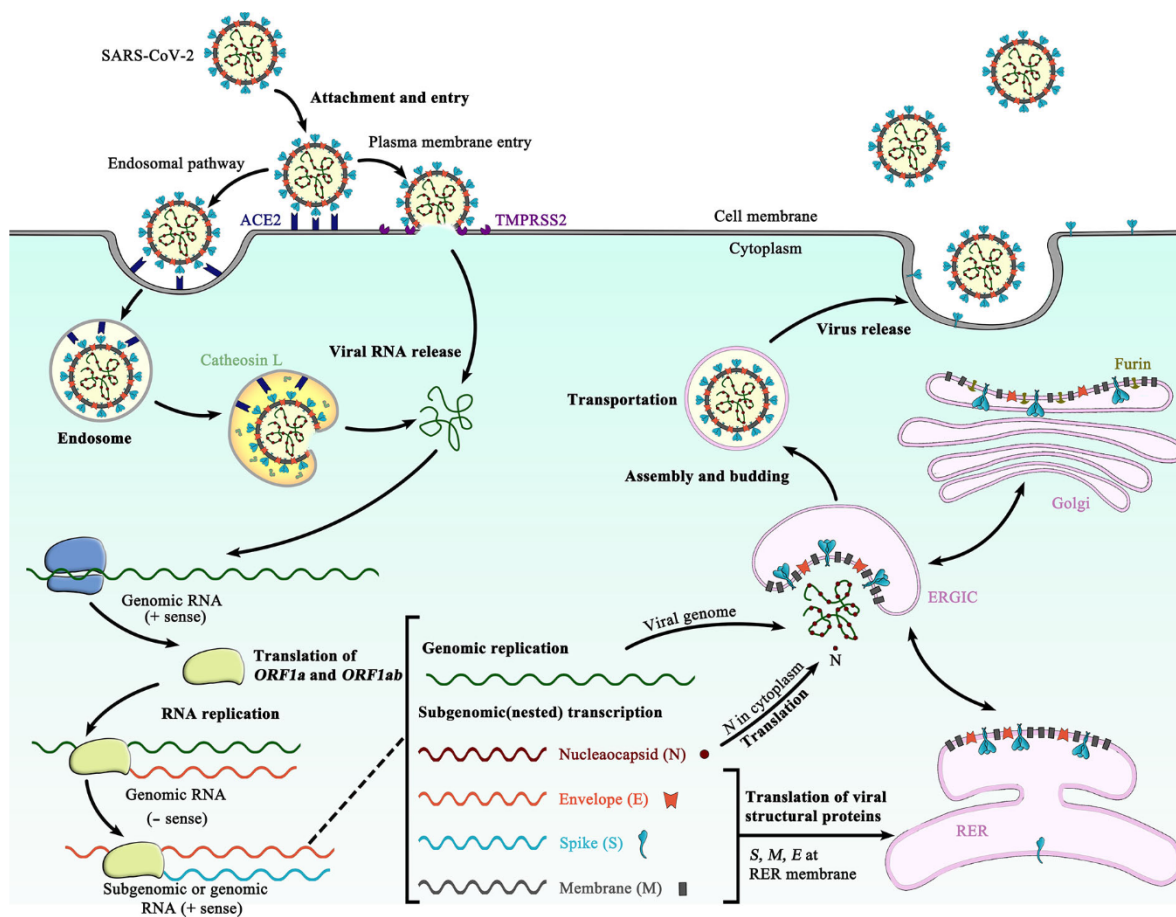
FIGURE 1 | Schematic representation of the life cycle of SARS-CoV-2. The life cycle of SARS-CoV-2 begins with membrane fusion occurring at the plasma membrane or within acidified endosomes after endocytosis, which is mediated by conformational changes in the S glycoprotein triggered by angiotensin-converting enzyme 2 (ACE2) binding. Following viral entry, SARS-CoV-2 releases its genomic RNA into the host cell cytoplasm. Genome RNA is first translated into viral replicase polyproteins (pp1a and 1ab), which are further cleaved by viral proteases into a total of 16 nonstructural proteins. A replication-transcription complex (RTC) is formed based on many of these nonstructural proteins. In the process of genome replication and transcription mediated by RTC, the negative-sense (− sense) genomic RNA is synthesized and used as a template to produce positive-sense (+ sense) genomic RNA and subgenomic RNAs. The nucleocapsid (N) structural protein and viral RNA are replicated, transcribed, and synthesized in the cytoplasm, whereas other viral structural proteins, including the S protein, membrane (M) protein and envelope (E) protein, are transcribed and then translated in the rough endoplasmic reticulum (RER) and transported to the Golgi complex. In the RER and Golgi complex, the SARS-CoV-2 glycoprotein is subjected to co-translational and post-translational processing, including signal peptide removal, trimerization, extensive glycosylation and subunit cleavage. The N protein is subsequently associated with the positive sense genomic RNA to become a nucleoprotein complex (nucleocapsid), which together with S, M, and E proteins as well as other viral proteins, is further assembled and followed by budding into the lumen of the ER-Golgi intermediate compartment (ERGIC) to form mature virions. Finally, the mature virions are released from the host cell, waiting for a new life cycle to start. This figure is adapted from the template in BioRender (https://biorender.com/).

two conformations determined by single particle cryo-electron microscopy (cryo-EM) is a ~160 Å long trimer with a triangular cross-section, with the S1 subunit adopting a "V" shape contributing to the overall triangular appearance and the S2 subunit forming the stalk (**Figures 2B, C**) (11, 34). The structural difference between these two conformations only lies in the position of one of the three S1 RBDs (**Figures 2B, C**) (11). When all three RBDs are in the "down" position, the resulting S ectodomain trimer assumes a closed conformation, in which the receptor-binding surface of the S1 RBD is buried at the interface between protomers and cannot be accessible by its receptor (**Figure 2B**) (11). The S ectodomain trimer with one single RBD in the "up" position assumes a partially open conformation and represents the functional state, as the receptor-binding surface of the "up" RBD can be fully exposed (**Figure 2C**) (11, 34). The structural information provides a blueprint for structure-based design of vaccine immunogens and entry inhibitors of SARS-CoV-2.

In the closed SARS-CoV-2 S ectodomain trimer, inter-protomer interactions occur through the S1 CTD packed against the other two S1 CTDs and one NTD from an adjacent protomer because of domain swapping and through S2, primarily between helical interactions formed by the upstream and central helices from each subunit around the trimer axis (**Figure 2B**) (11). The S1 subunits rest above the S2 trimer,
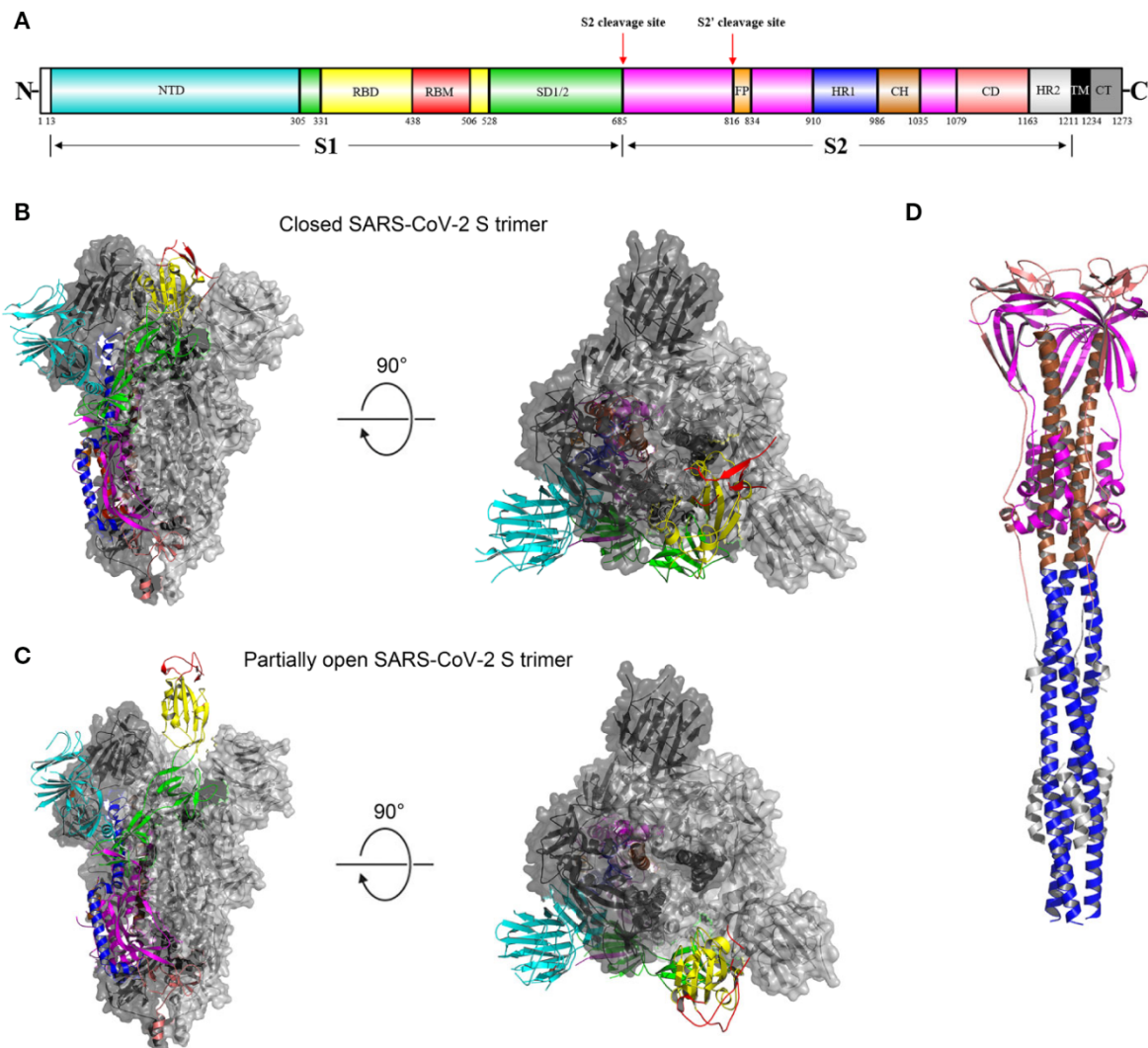
FIGURE 2 | Overall structures of the SARS-CoV-2 S glycoprotein trimer in different conformations. **(A)** Schematic representation of the domain arrangement of the SARS-CoV-2 S protein precursor. SS, signal peptide; NTD: N-terminal domain; RBD: receptor-binding domain; RBM: receptor-binding motif; SD1/2: subdomain 1 and 2; FP, fusion peptide; HR1, heptad repeat 1; CH, central helix; CD, connector domain; HR2, heptad repeat 2; TM, transmembrane domain; CT, cytoplasmic tail. Arrows denote protease cleavage sites. **(B)** Side and top views of the prefusion structure of the SARS-CoV-2 S ectodomain trimer with all three RBDs in the down conformation (PDB ID: 6VXX). One protomer is shown in ribbon representation colored corresponding to the schematic in **(A)**, a second protomer in light gray surface representation, and the third protomer in dark gray surface representation. **(C)** is identical to **(B)** except that a single RBD assumes the up conformation and is shown in ribbon representation (PDB ID: 6VYB). **(D)** Overall structure of the SARS-CoV-2 S2 trimer in the postfusion conformation is shown in ribbon representation colored corresponding to the schematic in **(A)** (PDB ID: 6XRA). The glycans were omitted for clarity.

stabilizing the later in the prefusion conformation (**Figure 2B**) (11). When the S ectodomain trimer adopts a partially open conformation, the RBD in the "up" position will abolish the contacts with the S2 subunit of an adjacent protomer, destabilizing the partially open conformation (**Figure 2C**) (11, 34). This will be beneficial to the dissociation of the S1 subunit and facilitate conformational rearrangements that the S2 trimer undergoes to mediate viral entry.

Prefusion structures of human coronavirus HKU1 (HCoV-HKU1) and mouse hepatitis virus S protein ectodomains without two consecutive proline mutations reveal only fully closed

conformation (37, 42), similar to that observed for a full-length, wild-type prefusion form of the SARS-CoV-2 S glycoprotein (41). Notably, it is well established that trimeric prefusion HIV-1 Env primarily resides in a closed configuration that is conformationally masked to evade antibody-mediated neutralization (43, 44) and can spontaneously sample a transient, functional configuration (45). It can thus be speculated that native CoV S glycoproteins on mature and infectious virions share a similar conformational masking feature (46), concealing the receptor-binding surface (for those utilizing CTDs as RBDs) (**Figure 2C**), which is further discussed below.

Several lines of research have established that angiotensin-converting enzyme 2 (ACE2) is an entry receptor for SARS-CoV-2 (47–49). Detailed interactions between the SARS-CoV-2 RBD and its receptor ACE have been revealed by several structures of ACE2 in complex with RBD (50–53). Structurally, RBD consists of two subdomains: a core and an external subdomain (51, 52). An extended loop (residues 438-506), which lies on one edge of the core subdomain, presents a gently concave surface to cradle the N-terminal helix (α1) of ACE2. Analysis of the interface between the SARS-CoV-2 RBD and ACE2 reveals that a total of 17 residues in RBD are in contact with 20 amino acids in ACE2, forming a network of hydrophilic interactions that are suggested to predominate the virus-receptor engagement (51). Outside this extended loop, residue Lys417 located in helix α3 of the core subdomain, was shown to form ionic interactions with Asp30 of ACE2. As the extended loop contains almost all the amino acids of the SARS-CoV-2 RBD that contact ACE2, it is referred to as the receptor-binding motif (RBM) (51).

It has been proposed that inhibiting the interaction between RBD and ACE2 might be useful in treating SARS-CoV-2 infection. Recombinant soluble ACE2 (54) and ACE2-Fc (55, 56) have been shown to have potential applications in the prevention and treatment of SARS-CoV-2 infection *in vitro*. As the interaction between the RBD and ACE2 is extensive, small molecules probably cannot be used as entry inhibitors to effectively block the virus entry by targeting the interaction interface. However, peptides would be able to engage most of the residues belonging to RBM (57). A pioneering study demonstrated that a 23-amino acid peptide (residues 21-43), derived from the N-terminal helix (α1) of ACE2, specifically associates with the SARS-CoV-2 RBD with low nanomolar affinity and disables receptor interactions (57), representing a promising strategy for preventing the virus from invading human cells. In another study, a 65-amino acid peptide (residues 19-83), derived from the N-terminal back-to-back helices (α1 and 2) and composed of most of the residues of ACE2 that mediate interactions with the S protein, shows a similar but probably more potent inhibitory effect (58).

The formation of a trimer-of-hairpins structure (also known as six-helix bundle) comprising HR1 and HR2 in the postfusion conformation is a unifying feature of class I viral fusion proteins (37). The crystal structure of a protein construct in which SARS-CoV-2 HR1 and HR2 were connected by a six-residue hydrophilic flexible linker was determined to be a canonical six-helix bundle structure with a rod-like shape ~115 Å in length and ~25 Å in diameter (59). Three HR1 helices form a parallel central coiled-coil with three HR2 helices packing in an oblique, antiparallel manner against deep hydrophobic grooves on the surface of the central coiled-coil (59). Notably, when a full-length S protein construct bearing the native furin-like cleavage site was transiently expressed by Expi293F cells, the purified S proteins contained the dissociated S2 trimer in the postfusion conformation (41). The cryo-EM structure of this trimeric postfusion S2 shows that the central helix (CH) extended regular helices from the central coiled-coil, oriented toward target cells (**Figure 2D**) (41), which forms the longest central triple helical coiled-coil (~180Å) among all known class I transmembrane subunit structures.

The SARS-CoV-2 S trimer in the pre-hairpin intermediate state is very unstable and is just transiently present *in vivo* after triggering by ACE2 engagement, stymieing structural characterization of the S protein in this state (60). However, although this fusion-intermediate phase is very short, it is enough for inhibitory peptides to associate with the pre-hairpin intermediate and block the six-helix bundle formation (39). Furthermore, it has already been shown that the HR1 regions in various human CoVs are highly conserved (61), and therefore could serve as an attractive target for the design and development of potent and broad-spectrum inhibitors of pan-CoVs, including SARS-CoV-2. A highly potent pan-coronavirus fusion inhibitor, EK1C4, has been reported to have good prophylactic and therapeutic potential against SARS-CoV-2 infection (59).

GLYCAN SHIELD OF THE SARS-COV-2 S GLYCOPROTEIN

As mentioned earlier, the SARS-CoV-2 S proteins are heavily decorated by heterogeneous N-linked glycans projecting from the S trimer surface. The SARS-CoV-2 S sequence encodes up to 22 N-linked glycan sequons per protomer, which likely plays an important role in protein folding (19) and host immune evasion as a glycan shield (62). Of the 22 potential N-linked glycosylation sites on the S protein, 14 were identified to be predominantly occupied by processed, complex-type glycans (63). The remaining eight sites were found to be dominated by oligomannose-type glycans, which are divergent from those founded on host glycoproteins (63). Although glycosylation sites (N165, N234, N343) proximal to the receptor-binding sites on the SARS-CoV-2 S protein can be observed, ACE2 bound to the glycosylated and deglycosylated S ectodomains with nearly identical affinity (1.7 nM vs 1.5 nM) determined by a biolayer interferometry binding assay (64). This observation suggests that the high binding affinity between the SARS-CoV-2 S protein and ACE2 does not depend on the S protein glycosylation.

When the site-specific N-linked glycans are mapped onto the prefusion structure of the SARS-CoV-2 S ectodomain (63), the resulting model exhibited substantially higher levels of glycan-free surface than that revealed by structures of fully glycosylated, trimeric HIV-1 Env ectodomains (65, 66). This suggests that the SARS-CoV-2 S protein is covered by a less dense and less effective glycan shield compared to viral glycoproteins from HIV-1 (36, 66) and Lassa virus (67), which may be beneficial for the induction of humoral immunity and could be good news for a SARS-CoV-2 vaccine (68).

Notably, it has been shown that multiple major viral surface antigens have neutralizing epitopes that are partly or even exclusively composed of carbohydrate moieties (69, 70), exemplified by the HIV-1 Env spike, which could be recognized by a large number of carbohydrate-binding

antibodies, including 2G12, PG9, PG16, CH04, PGT121, PGT128, PGT135, and PGT145 (70, 71). In the case of SARS-CoV-2, more recently a potent neutralizing antibody against both SARS-CoV and SARS-CoV-2, S309, has been shown to recognize a highly conserved glycan-containing RBD epitope (72). These observations suggest that carbohydrate moieties could be immunogenic and highlight the need for immunogens to display the glycans important for the recognition of neutralizing antibodies (73); in support of this, specific N-linked glycans on Hemagglutinin has been shown to be essential for the elicitation of broadly neutralizing antibodies against Influenza (74). Accordingly, there has been mounting interest in exploring the potential of immunogenic glycan moieties as vaccine candidates against multiple viruses, including SARS-CoV-2 (75, 76).

SARS-COV-2 S GLYCOPROTEIN-MEDIATED MEMBRANE FUSION

Membrane fusion and viral entry of SARS-CoV-2 is initiated by binding of RBD in the viral S glycoprotein transiently sampling the functional conformation to ACE2 on the surface of target cells (**Figure 1**) (10). After receptor engagement at the plasma membrane or ensuing virus endocytosis by the host cell (8), a second cleavage (S2′ cleavage site) is generated, which is mediated by a cellular serine protease TMPRSS2 (48) or endosomal cysteine proteases cathepsins B and L (10) (**Figure 1**). Protease cleavage at S2′ site frees the fusion peptide from the new S2 N-terminal region, further destabilizes the SARS-CoV-2 S glycoprotein and may initiate S2-mediated membrane fusion cascade. Following the second cleavage, the fusion peptide at the N terminus of the S2 trimer is inserted into the host membrane (8), forming the pre-hairpin intermediate state (39). Since the pre-hairpin intermediate state is extremely unstable, the S2 fusion protein is refolded quickly and irreversibly into the stable postfusion state (39, 77). These large conformational rearrangements pull the viral and host cell membrane into close proximity, leading ultimately to the membrane fusion (8, 39).

INSIGHTS INTO THE DESIGN AND DEVELOPMENT OF S PROTEIN-BASED VACCINES

Since SARS-CoV-2 was identified as the causative agent of COVID-19, and its first genome sequence was released immediately and freely by a Chinese research group (40), SARS-CoV-2 vaccine candidates based on various vaccine platforms, such as inactivated or live attenuated vaccines, DNA and mRNA vaccines, viral vector-based vaccines, and recombinant protein-based vaccines, have been developed (12, 78). Most of these vaccine strategies are based on the full-length S glycoprotein, the major viral surface antigen (12). When a

vaccine strategy requires that the SARS-CoV-2 S protein be recombinantly expressed in the human body, the ERRS should be omitted to enhance the cell surface expression level of the resulting protein.

Theoretically, the native HIV-1 Env trimer present on the surface of intact virions is thought to be a most ideal immunogen (60), as most of the neutralizing antibodies thus far described could recognize and bind to the prefusion form of trimeric HIV-1 Env, although it is with great difficulty that such neutralizing antibodies against this glycan-covered, sequence-variable native form are induced (36). For SARS-CoV-2, different lines of research have shown that convalescent sera from SARS-CoV and SARS-CoV-2 patients showed no or limited cross-neutralization activity against these two viruses by pseudotyped and authentic viral infection assays, despite significant cross-reactivity in binding to the S glycoproteins of both viruses (9, 79–81). Similar results were also observed in infected or immunized animals (48, 79, 81). Together with the finding that although the SARS-CoV-2 S protein shares a high degree of amino acid sequence identity with that of SARS-CoV (~76% overall), the RBM is less conserved (~47% identity) than any other functional region or domain (82), it can thus been surmised that the RBM has the most immunodominant neutralizing epitope(s) of the whole S protein, capable of readily eliciting strong neutralizing antibody responses. However, the native trimeric SARS-CoV-2 S protein could conceal each of its immunodominant RBMs by adopting the closed conformation (41, 83). Therefore, SARS-CoV-2 evades immune surveillance also through conformational masking, which is well-documented for HIV-1 (43, 44); while at the same time, the S protein could transiently sample the functional state to engage ACE2, consistent with the notion that the fusion glycoprotein of highly pathogenic viruses have evolved to perform its functions while evading host neutralizing antibody responses.

Another concern for vaccine candidates based on the full-length S glycoprotein of SARS-CoV-2 is raised by the observation that the S1 subunit could spontaneously dissociate from the S glycoprotein probably as a trimer that still assumes the RBD closed conformation, leaving only the postfusion S2 trimer (41). The resulting S1 and S2 subunits might expose immunodominant, nonneutralizing epitopes that are utilized by SARS-CoV-2 to serve as decoys to distract the host immune system, inducing a large proportion of ineffective antibody responses, as documented for HIV-1 (60) and respiratory syncytial virus (RSV) (84).

It should be noted that although vaccine candidates based on the full-length S protein of the closely related SARS-CoV could elicit neutralizing antibody responses against infection of SARS-CoV, they may also induce harmful immune responses, including liver damage of the vaccinated animals, infection of human immune cells by SARS-CoV, and antibody-dependent enhancement of SARS-CoV infection (85–89). Therefore, although the S proteins of both SARS-CoV and SARS-CoV-2 are thought to be promising vaccine immunogens for generating protective immunity, optimizing antigen design is critical to ensure an optimal immune response through exposing more

neutralizing epitopes and displaying fewer potentially weakly or non-neutralizing epitopes (90). Vaccines containing or expressing the full-length S protein or its soluble ectodomain form should thus be engineered to sample a RBD(s) "up" conformation while the rest is still kept in the prefusion state (91, 92).

Apart from recombinant, soluble, stabilized ectodomains that are engineered to expose the immunodominant RBD by adapting the RBD(s) "up" conformation, RBD proteins of SARS-CoV and SARS-CoV-2 have also been widely used as recombinant protein-based vaccines (85, 93–95). The RBD of SARS-CoV is highly immunogenic (96, 97) and is targeted by most of the neutralizing monoclonal antibodies that have been characterized (98). Based on the observation that a 193-amino acid fragment (residues 318-510) was previously identified to be the minimal RBD region of SARS-CoV (99), a corresponding 194-amino acid fragment (residues 331-524) can be readily selected as the minimal RBD region of SARS-CoV-2 and has already been characterized (100). This minimal form of RBDs of both viruses could serve as a vaccine candidate (100).

However, a conserved cysteine residue is located immediately upstream of the minimal RBD fragments of both viruses and always forms a disulfide bond in nearly all published structures containing this residue (101, 102); this is also the case for Middle East respiratory syndrome coronavirus (MERS-CoV) (103, 104) and HCoV-HKU1 (37), consistent with the observation that all RBDs of these viruses share a conserved structural core. The disulfide bond contributes to stabilization of the RBD structure and likely modulates the protein immunogenicity. This notion is consistent with the observation that mice immunized with a longer form of the SARS-CoV RBD (residues 318-536) produced a higher titer of neutralizing antibodies compared with mice immunized with the minimal RBD region (residues 318-510) (105). Therefore, when each of the minimal RBD fragments of SARS-CoV and SARS-CoV-2 is used as vaccine candidates, the critical cysteine residue should not be ignored and thus should be included (106).

Besides the RBD, which has been shown to a major target for human neutralizing antibody responses (107), the NTD was recently identified to be a new vulnerable site of the SARS-CoV-2 S protein for antibody neutralizing and therefore could also serve as a recombinant protein-based vaccine (108–110). As expected, NTD-specific neutralizing antibodies could target the S protein in both closed and open conformations (108). In addition, the apparent accessibility of the fusion peptide and HR1 region in published structures of the SARS-CoV-2 S ectodomain trimer as well as their high sequence conservation among CoVs suggests that they would be good immunogen candidates for epitope-focused vaccine design aimed at raising broadly CoV neutralizing antibodies (46). The epitope-focused vaccine design has proven to be successful in generating neutralizing antibodies against RSV fusion glycoprotein (111). However, neutralizing antibodies targeted against these two regions still need to be isolated in infected individuals to support this notion.

Unlike wild-type full-length S protein of SARS-CoV-2, the above monomeric fragments do not induce any infection-

enhancing antibodies or harmful immune or inflammatory responses (106, 112), all of which could be potentially avoided through structure-based immunogen design to improve immunogenicity (113, 114). However, wide-type full-length or soluble ectodomain form of the SARS-CoV-2 S protein could trigger stronger cellular immune responses (115), which have been demonstrated to play an important role in controlling diseases caused by CoVs (116, 117), including SARS-CoV-2 (118), and are probably also an important determinant of effective vaccines against SARS-CoV-2 (115, 119). Additionally, when more than one RBD of the S protein trimer is engineered to be locked in the "up" conformation (120, 121), the antigenicity and immunogenicity of the resulting RBDs would be significantly enhanced compared with monomeric RBD form (97, 122). Moreover, improved protection is likely to be achieved when vaccinated with full-length or soluble ectodomain form of the SARS-CoV-2 S protein in that both forms can elicit neutralizing antibodies directed against non-RBD sites, as observed for MERS-CoV (123).

Genetic variation has been used by many viruses that have RNA genomes (124), including HIV and influenza, as a mechanism to avoid antibody-mediated immunity, and is partially responsible for the great difficulty in developing effective and durable vaccines against these viruses (36). As an RNA virus, however, SARS-CoV-2 has a very low mutation rate overall (125) likely because CoVs have a genetic proofreading mechanism (126). All reported variations occurred in the SARS-CoV-2 S glycoprotein have a prevalence of no more than 1% (127), with an exception of D614G, which has become the most prevalent genotype in the global COVID-19 pandemic (127). Fortunately, although the D614G mutation of the SARS-CoV-2 S protein has been shown to enhance viral infectivity (128–130), until now there is no evidence that infection with SARS-CoV-2 carrying the G614 mutant will be associated with disease severity (127, 131). Furthermore, assays using both monoclonal and polyclonal antibodies generated from individuals naturally infected with D614- or G614-carrying viruses demonstrated that the D614G mutation retains or even increases viral susceptibility to neutralization (127, 130, 132, 133). This suggests that the D614G mutant maintains or favors an open, functional conformational state (134).

Although at an extremely low frequency, natural variations, including L452R A475V, V483A, and F490L that render the S glycoprotein resistant to certain neutralizing antibodies targeting the RBD, emerged under no selection pressure exerted by approved vaccines or neutralizing antibodies or entry inhibitors (127, 132). However, it has been shown that SARS-CoV-2 escape mutants could be easily selected and quickly amplified under the selection pressure of single antibody treatment (135). These observations suggest that a combination of at least two neutralizing antibodies that recognize and bind to distinct and non-overlapping epitopes on the SARS-CoV-2 S glycoprotein (e.g., RBD and NTD, as well as HR and glycan) is required to restrict the possible occurrence of viral escape mutants and potential subsequent loss of single antibody-mediated neutralization (135–138). When these

observations are taken into consideration for vaccine design and development, an ideal SARS-CoV-2 immunogen should contain as many exposed neutralizing epitopes as possible, although the RBD also possesses extra epitope(s) besides the epitope in the RBM region (72, 139–141).

CONCLUDING REMARKS AND PROSPECTS

SARS-CoV-2 is a highly contagious pathogen that continues to spread quickly around the globe, causing COVID-19 to be one of the worst pandemics in recorded history. A safe and efficacious vaccine represents one of the best ways to reduce or eliminate the COVID-19 pandemic (142). Unfortunately, no vaccines for any of the known human CoVs have been licensed (143, 144), although several potential SARS-CoV and MERS-CoV vaccines have advanced into human clinical trials for years (117, 145), suggesting the development of effective vaccines against human CoVs has always been challenging. However, it has been shown that both SARS-CoV and SARS-CoV-2 could readily induce neutralizing antibodies following natural infection or immunization (146–149). Moreover, a growing number of neutralizing monoclonal antibodies targeting the SARS-CoV-2 S glycoprotein with high potency have been isolated from plenty of convalescent donors (33) as well as humanized mice (136, 141), some of which have been shown to afford protection against SARS-CoV-2 challenge in animal models. It thus seems that vaccine candidates designed to elicit such neutralizing antibodies are feasible. It is widely accepted that the S protein of SARS-CoV-2 is a most promising immunogen for producing protective immunity (150). However, it is likely that the S protein has evolved to perform its functions while evading host neutralizing antibody responses and thus should be engineered to ensure an optimal immune response (151, 152). The immunogen design strategies described in this review based on the wealth of the SARS-CoV-2 S glycoprotein research related to its biosynthesis, structure, function, antigenicity as well as immunogenicity will likely contribute to the ultimate success of safe and efficacious vaccines against SARS-CoV-2/COVID-19.

AUTHOR CONTRIBUTIONS

FUNDING

ACKNOWLEDGMENTS

REFERENCES

1. Keni R, Alexander A, Nayak PG, Mudgal J, Nandakumar K. COVID-19: Emergence, Spread, Possible Treatments, and Global Burden. *Front Public Health* (2020) 8:216. doi: 10.3389/fpubh.2020.00216
2. Liang Y, Wang M-L, Chien C-S, Yarmishyn AA, Yang Y-P, Lai W-Y, et al. Highlight of Immune Pathogenic Response and Hematopathologic Effect in SARS-CoV, MERS-CoV, and SARS-Cov-2 Infection. *Front Immunol* (2020) 11:1022. doi: 10.3389/fimmu.2020.01022
3. Coronaviridae Study Group of the International Committee on Taxonomy of Viruses. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat Microbiol* (2020) 5:536–44. doi: 10.1038/s41564-020-0695-z
4. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis* (2020) 20:533–4. doi: 10.1016/S1473-3099(20)30120-1
5. Cui J, Li F, Shi ZL. Origin and evolution of pathogenic coronaviruses. *Nat Rev Microbiol* (2019) 17:181–92. doi: 10.1038/s41579-018-0118-9
6. Harrison SC. Viral membrane fusion. *Virology* (2015) 479-480:498–507. doi: 10.1016/j.virol.2015.03.043
7. Shang J, Wan Y, Luo C, Ye G, Geng Q, Auerbach A, et al. Cell entry mechanisms of SARS-CoV-2. *Proc Natl Acad Sci U S A* (2020) 117:11727–34. doi: 10.1073/pnas.2003138117
8. Tang T, Bidon M, Jaimes JA, Whittaker GR, Daniel S. Coronavirus membrane fusion mechanism offers a potential target for antiviral development. *Antiviral Res* (2020) 178:104792. doi: 10.1016/j.antiviral.2020.104792
9. Ou X, Liu Y, Lei X, Li P, Mi D, Ren L, et al. Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nat Commun* (2020) 11:1620. doi: 10.1038/s41467-020-15562-9
10. Belouzard S, Millet JK, Licitra BN, Whittaker GR. Mechanisms of coronavirus cell entry mediated by the viral spike protein. *Viruses* (2012) 4:1011–33. doi: 10.3390/v4061011
11. Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* (2020) 181:281–92.e6. doi: 10.1016/j.cell.2020.02.058
12. Amanat F, Krammer F. SARS-CoV-2 Vaccines: Status Report. *Immunity* (2020) 52:583–9. doi: 10.1016/j.immuni.2020.03.007
13. Liu Z, Xiao X, Wei X, Li J, Yang J, Tan H, et al. Composition and divergence of coronavirus spike proteins and host ACE2 receptors predict potential intermediate hosts of SARS-CoV-2. *J Med Virol* (2020) 92:595–601. doi: 10.1002/jmv.25726
14. Breitling J, Aebi M. N-linked protein glycosylation in the endoplasmic reticulum. *Cold Spring Harb Perspect Biol* (2013) 5:a013359. doi: 10.1101/cshperspect.a013359
15. Braakman I, Hebert DN. Protein folding in the endoplasmic reticulum. *Cold Spring Harb Perspect Biol* (2013) 5:a013201. doi: 10.1101/cshperspect.a013201
16. Broer R, Boson B, Spaan W, Cosset FL, Corver J. Important role for the transmembrane domain of severe acute respiratory syndrome coronavirus spike protein during entry. *J Virol* (2006) 80:1302–10. doi: 10.1128/JVI.80.3.1302-1310.2006
17. Chung CY, Majewska NI, Wang Q, Paul JT, Betenbaugh MJ. SnapShot: N-Glycosylation Processing Pathways across Kingdoms. *Cell* (2017) 171:258–e1. doi: 10.1016/j.cell.2017.09.014

18. Aebi M. N-linked protein glycosylation in the ER. *Biochim Biophys Acta* (2013) 1833:2430–7. doi: 10.1016/j.bbamcr.2013.04.001

19. Helenius A, Aebi M. Intracellular functions of N-linked glycans. *Science* (2001) 291:2364–9. doi: 10.1126/science.291.5512.2364

20. Schwarz F, Aebi M. Mechanisms and principles of N-linked protein glycosylation. *Curr Opin Struct Biol* (2011) 21:576–82. doi: 10.1016/j.sbi.2011.08.005

21. Zhang X, Wang Y. Glycosylation Quality Control by the Golgi Structure. *J Mol Biol* (2016) 428:3183–93. doi: 10.1016/j.jmb.2016.02.030

22. Andersen KG, Rambaut A, Lipkin WI, Holmes EC, Garry RF. The proximal origin of SARS-CoV-2. *Nat Med* (2020) 26:450–2. doi: 10.1038/s41591-020-0820-9

23. Joshi HJ, Narimatsu Y, Schjoldager KT, Tytgat HLP, Aebi M, Clausen H, et al. SnapShot: O-Glycosylation Pathways across Kingdoms. *Cell* (2018) 172:632–e2. doi: 10.1016/j.cell.2018.01.016

24. Hoffmann M, Kleine-Weber H, Pohlmann S. A Multibasic Cleavage Site in the Spike Protein of SARS-CoV-2 Is Essential for Infection of Human Lung Cells. *Mol Cell* (2020) 78:779–84 e5. doi: 10.1016/j.molcel.2020.04.022

25. Coutard B, Valle C, de Lamballerie X, Canard B, Seidah NG, Decroly E. The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade. *Antiviral Res* (2020) 176:104742. doi: 10.1016/j.antiviral.2020.104742

26. Bestle D, Heindl MR, Limburg H, Pilgram O, Moulton H, Stein DA, et al. TMPRSS2 and furin are both essential for proteolytic activation and spread of SARS-CoV-2 in human airway epithelial cells and provide promising drug targets. *Life Sci Alliance* (2020) 3:e202000786. doi: 10.26508/lsa.202000786

27. Lontok E, Corse E, Machamer CE. Intracellular targeting signals contribute to localization of coronavirus spike proteins near the virus assembly site. *J Virol* (2004) 78:5913–22. doi: 10.1128/JVI.78.11.5913-5922.2004

28. Stertz S, Reichelt M, Spiegel M, Kuri T, Martinez-Sobrido L, Garcia-Sastre A, et al. The intracellular sites of early replication and budding of SARS-coronavirus. *Virology* (2007) 361:304–15. doi: 10.1016/j.virol.2006.11.027

29. McBride CE, Li J, Machamer CE. The cytoplasmic tail of the severe acute respiratory syndrome coronavirus spike protein contains a novel endoplasmic reticulum retrieval signal that binds COPI and promotes interaction with membrane protein. *J Virol* (2007) 81:2418–28. doi: 10.1128/JVI.02146-06

30. Ujike M, Huang C, Shirato K, Makino S, Taguchi F. The contribution of the cytoplasmic retrieval signal of severe acute respiratory syndrome coronavirus to intracellular accumulation of S proteins and incorporation of S protein into virus-like particles. *J Gen Virol* (2016) 97:1853–64. doi: 10.1099/jgv.0.000494

31. Malik YA. Properties of Coronavirus and SARS-CoV-2. *Malays J Pathol* (2020) 42:3–11.

32. Dieterle ME, Haslwanter D, Bortz III RH, Wirchnianski AS, Lasso G, Vergnolle O, et al. A replication-competent vesicular stomatitis virus for studies of SARS-CoV-2 spike-mediated cell entry and its inhibition. *Cell Host Microbe* (2020) 28:486–96.e6. doi: 10.1016/j.chom.2020.06.020

33. Schmidt F, Weisblum Y, Muecksch F, Hoffmann H-H, Michailidis E, Lorenzi JC, et al. Measuring SARS-CoV-2 neutralizing antibody activity using pseudotyped and chimeric viruses. *J Exp Med* (2020) 217:e20201181. doi: 10.1084/jem.20201181

34. Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh CL, Abiona O, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* (2020) 367:1260–3. doi: 10.1126/science.abb2507

35. Bosch BJ, van der Zee R, de Haan CA, Rottier PJ. The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. *J Virol* (2003) 77:8801–11. doi: 10.1128/jvi.77.16.8801-8811.2003

36. Pancera M, Zhou T, Druz A, Georgiev IS, Soto C, Gorman J, et al. Structure and immune recognition of trimeric pre-fusion HIV-1 Env. *Nature* (2014) 514:455–61. doi: 10.1038/nature13808

37. Kirchdoerfer RN, Cottrell CA, Wang N, Pallesen J, Yassine HM, Turner HL, et al. Pre-fusion structure of a human coronavirus spike protein. *Nature* (2016) 531:118–21. doi: 10.1038/nature17200

38. Rey FA, Lok SM. Common Features of Enveloped Viruses and Implications for Immunogen Design for Next-Generation Vaccines. *Cell* (2018) 172:1319–34. doi: 10.1016/j.cell.2018.02.054

39. Harrison SC. Viral membrane fusion. *Nat Struct Mol Biol* (2008) 15:690–8. doi: 10.1038/nsmb.1456

40. Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, et al. A new coronavirus associated with human respiratory disease in China. *Nature* (2020) 579:265–9. doi: 10.1038/s41586-020-2008-3

41. Cai Y, Zhang J, Xiao T, Peng H, Sterling SM, Walsh RM, et al. Distinct conformational states of SARS-CoV-2 spike protein. *bioRxiv* (2020). doi: 10.1101/2020.05.16.099317

42. Walls AC, Tortorici MA, Bosch BJ, Frenz B, Rottier PJM, DiMaio F, et al. Cryo-electron microscopy structure of a coronavirus spike glycoprotein trimer. *Nature* (2016) 531:114–7. doi: 10.1038/nature16988

43. Kwong PD, Doyle ML, Casper DJ, Cicala C, Leavitt SA, Majeed S, et al. HIV-1 evades antibody-mediated neutralization through conformational masking of receptor-binding sites. *Nature* (2002) 420:678–82. doi: 10.1038/nature01188

44. Lewis GK, Finzi A, DeVico AL, Pazgier M. Conformational Masking and Receptor-Dependent Unmasking of Highly Conserved Env Epitopes Recognized by Non-Neutralizing Antibodies That Mediate Potent ADCC against HIV-1. *Viruses* (2015) 7:5115–32. doi: 10.3390/v7092856

45. Munro JB, Gorman J, Ma X, Zhou Z, Arthos J, Burton DR, et al. Conformational dynamics of single HIV-1 envelope trimers on the surface of native virions. *Science* (2014) 346:759–63. doi: 10.1126/science.1254426

46. Pallesen J, Wang N, Corbett KS, Wrapp D, Kirchdoerfer RN, Turner HL, et al. Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen. *Proc Natl Acad Sci U S A* (2017) 114:E7348–E57. doi: 10.1073/pnas.1707304114

47. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* (2020) 579:270–3. doi: 10.1038/s41586-020-2012-7

48. Hoffmann M, Kleine-Weber H, Schroeder S, Kruger N, Herrler T, Erichsen S, et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* (2020) 181:271–80.e8. doi: 10.1016/j.cell.2020.02.052

49. Letko M, Marzi A, Munster V. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat Microbiol* (2020) 5:562–9. doi: 10.1038/s41564-020-0688-y

50. Shang J, Ye G, Shi K, Wan Y, Luo C, Aihara H, et al. Structural basis of receptor recognition by SARS-CoV-2. *Nature* (2020) 581:221–4. doi: 10.1038/s41586-020-2179-y

51. Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S, et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* (2020) 581:215–20. doi: 10.1038/s41586-020-2180-5

52. Wang Q, Zhang Y, Wu L, Niu S, Song C, Zhang Z, et al. Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell* (2020) 181:894–904 e9. doi: 10.1016/j.cell.2020.03.045

53. Yan R, Zhang Y, Li Y, Xia L, Guo Y, Zhou Q. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* (2020) 367:1444–8. doi: 10.1126/science.abb2762

54. Monteil V, Kwon H, Prado P, Hagelkruys A, Wimmer RA, Stahl M, et al. Inhibition of SARS-CoV-2 Infections in Engineered Human Tissues Using Clinical-Grade Soluble Human ACE2. *Cell* (2020) 181:905–13 e7. doi: 10.1016/j.cell.2020.04.004

55. Lei C, Qian K, Li T, Zhang S, Fu W, Ding M, et al. Neutralization of SARS-CoV-2 spike pseudotyped virus by recombinant ACE2-Ig. *Nat Commun* (2020) 11:2070. doi: 10.1038/s41467-020-16048-4

56. Iwanaga N, Cooper L, Rong L, Beddingfield B, Crabtree J, Tripp RA, et al. Novel ACE2-IgG1 fusions with increased activity against SARS-CoV-2. *bioRxiv* (2020). doi: 10.1101/2020.06.15.152157

57. Zhang G, Pomplun S, Loftis AR, Loas A, Pentelute B. The first-in-class peptide binder to the SARS-CoV-2 spike protein. *bioRxiv* (2020). doi: 10.1101/2020.03.19.999318

58. Renzi F, Ghersi D. ACE2 fragment as a decoy for novel SARS-Cov-2 virus. *bioRxiv* (2020). doi: 10.1101/2020.04.06.028647

59. Xia S, Liu M, Wang C, Xu W, Lan Q, Feng S, et al. Inhibition of SARS-CoV-2 (previously 2019-nCoV) infection by a highly potent pan-coronavirus fusion inhibitor targeting its spike protein that harbors a high capacity to mediate membrane fusion. *Cell Res* (2020) 30:343–55. doi: 10.1038/s41422-020-0305-x

60. Duan L, Du J, Liu X. Insights into vaccine development for acquired immune deficiency syndrome from crystal structures of human immunodeficiency virus-1 gp41 and equine infectious anemia virus gp45. *Protein Sci* (2015) 24:1549–59. doi: 10.1002/pro.2750

61. Xia S, Yan L, Xu W, Agrawal AS, Algaissi A, Tseng CK, et al. A pan-coronavirus fusion inhibitor targeting the HR1 domain of human coronavirus spike. *Sci Adv* (2019) 5:eaav4580. doi: 10.1126/sciadv.aav4580

62. Walls AC, Tortorici MA, Frenz B, Snijder J, Li W, Rey FA, et al. Glycan shield and epitope masking of a coronavirus spike protein observed by cryo-electron microscopy. *Nat Struct Mol Biol* (2016) 23:899–905. doi: 10.1038/nsmb.3293

63. Watanabe Y, Allen JD, Wrapp D, McLellan JS, Crispin M. Site-specific glycan analysis of the SARS-CoV-2 spike. *Science* (2020) 369:330–3. doi: 10.1126/science.abb9983

64. Sun Z, Ren K, Zhang X, Chen J, Jiang Z, Jiang J, et al. Mass spectrometry analysis of newly emerging coronavirus HCoV-19 spike S protein and human ACE2 reveals camouflaging glycans and unique post-translational modifications. *Engineering (Beijing)* (2020). doi: 10.1016/j.eng.2020.07.014

65. Watanabe Y, Berndsen ZT, Raghwani J, Seabright GE, Allen JD, Pybus OG, et al. Vulnerabilities in coronavirus glycan shields despite extensive glycosylation. *Nat Commun* (2020) 11:2688. doi: 10.1038/s41467-020-16567-0

66. Stewart-Jones GB, Soto C, Lemmin T, Chuang GY, Druz A, Kong R, et al. Trimeric HIV-1-Env Structures Define Glycan Shields from Clades A, B, and G. *Cell* (2016) 165:813–26. doi: 10.1016/j.cell.2016.04.010

67. Watanabe Y, Raghwani J, Allen JD, Seabright GE, Li S, Moser F, et al. Structure of the Lassa virus glycan shield provides a model for immunological resistance. *Proc Natl Acad Sci U S A* (2018) 115:7320–5. doi: 10.1073/pnas.1803990115

68. Zhang Y, Zhao W, Mao Y, Wang S, Zhong Y, Su T, et al. Site-specific N-glycosylation Characterization of Recombinant SARS-CoV-2 Spike Proteins using High-Resolution Mass Spectrometry. *bioRxiv* (2020). doi: 10.1101/2020.03.28.013276

69. Horiya S, MacPherson IS, Krauss IJ. Recent strategies targeting HIV glycans in vaccine design. *Nat Chem Biol* (2014) 10:990–9. doi: 10.1038/nchembio.1685

70. Seabright GE, Doores KJ, Burton DR, Crispin M. Protein and Glycan Mimicry in HIV Vaccine Design. *J Mol Biol* (2019) 431:2223–47. doi: 10.1016/j.jmb.2019.04.016

71. Crispin M, Ward AB, Wilson IA. Structure and Immune Recognition of the HIV Glycan Shield. *Annu Rev Biophys* (2018) 47:499–523. doi: 10.1146/annurev-biophys-060414-034156

72. Pinto D, Park YJ, Beltramello M, Walls AC, Tortorici MA, Bianchi S, et al. Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody. *Nature* (2020) 583:290–5. doi: 10.1038/s41586-020-2349-y

73. Dubrovskaya V, Tran K, Ozorowski G, Guenaga J, Wilson R, Bale S, et al. Vaccination with Glycan-Modified HIV NFL Envelope Trimer-Liposomes Elicits Broadly Neutralizing Antibodies to Multiple Sites of Vulnerability. *Immunity* (2019) 51:915–29 e7. doi: 10.1016/j.immuni.2019.10.008

74. Huang Y, Owino SO, Crevar CJ, Carter DM, Ross TM. N-Linked Glycans and K147 Residue on Hemagglutinin Synergize To Elicit Broadly Reactive H1N1 Influenza Virus Antibodies. *J Virol* (2020) 94:e01432–19. doi: 10.1128/JVI.01432-19

75. Crispin M, Doores KJ. Targeting host-derived glycans on enveloped viruses for antibody-based vaccine design. *Curr Opin Virol* (2015) 11:63–9. doi: 10.1016/j.coviro.2015.02.002

76. Wang D. Coronaviruses' sugar shields as vaccine candidates. *Curr Trends Immunol* (2020) 21:17–23.

77. Yuan Y, Cao D, Zhang Y, Ma J, Qi J, Wang Q, et al. Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains. *Nat Commun* (2017) 8:15092. doi: 10.1038/ncomms15092

78. Thanh Le T, Andreadakis Z, Kumar A, Gomez Roman R, Tollefsen S, Saville M, et al. The COVID-19 vaccine development landscape. *Nat Rev Drug Discov* (2020) 19:305–6. doi: 10.1038/d41573-020-00073-5

79. Yang R, Lan J, Huang B, A R, Lu M, Wang W, et al. Lack of antibody-mediated cross-protection between SARS-CoV-2 and SARS-CoV infections. *EBioMedicine* (2020) 58:102890. doi: 10.1016/j.ebiom.2020.102890

80. Anderson DE, Tan CW, Chia WN, Young BE, Linster M, Low JH, et al. Lack of cross-neutralization by SARS patient sera towards SARS-CoV-2. *Emerg Microbes Infect* (2020) 9:900–2. doi: 10.1080/22221751.2020.1761267

81. Lv H, Wu NC, Tsang OT, Yuan M, Perera R, Leung WS, et al. Cross-reactive Antibody Response between SARS-CoV-2 and SARS-CoV Infections. *Cell Rep* (2020) 31:107725. doi: 10.1016/j.celrep.2020.107725

82. Jaimes JA, Andre NM, Chappie JS, Millet JK, Whittaker GR. Phylogenetic Analysis and Structural Modeling of SARS-CoV-2 Spike Protein Reveals an Evolutionary Distinct and Proteolytically Sensitive Activation Loop. *J Mol Biol* (2020) 432:3309–25. doi: 10.1016/j.jmb.2020.04.009

83. Turonova B, Sikora M, Schurmann C, Hagen WJH, Welsch S, Blanc FEC, et al. In situ structural analysis of SARS-CoV-2 spike reveals flexibility mediated by three hinges. *Science* (2020). doi: 10.1126/science.abd5223

84. McLellan JS, Chen M, Joyce MG, Sastry M, Stewart-Jones GB, Yang Y, et al. Structure-based design of a fusion glycoprotein vaccine for respiratory syncytial virus. *Science* (2013) 342:592–8. doi: 10.1126/science.1243283

85. Du L, He Y, Zhou Y, Liu S, Zheng BJ, Jiang S. The spike protein of SARS-CoV–a target for vaccine and therapeutic development. *Nat Rev Microbiol* (2009) 7:226–36. doi: 10.1038/nrmicro2090

86. Jiang S, He Y, Liu S. SARS vaccine development. *Emerg Infect Dis* (2005) 11:1016–20. doi: 10.3201/1107.050219

87. Ma C, Su S, Wang J, Wei L, Du L, Jiang S. From SARS-CoV to SARS-CoV-2: safety and broad-spectrum are important for coronavirus vaccine development. *Microbes Infect* (2020) 2:245–53. doi: 10.1016/j.micinf.2020.05.004

88. Karthik K, Senthilkumar TMA, Udhayavel S, Raj GD. Role of antibody-dependent enhancement (ADE) in the virulence of SARS-CoV-2 and its mitigation strategies for the development of vaccines and immunotherapies to counter COVID-19. *Hum Vaccin Immunother* (2020) 1–6. doi: 10.1080/21645515.2020.1796425

89. Arvin AM, Fink K, Schmid MA, Cathcart A, Spreafico R, Havenar-Daughton C, et al. A perspective on potential antibody-dependent enhancement of SARS-CoV-2. *Nature* (2020) 584:353–63. doi: 10.1038/s41586-020-2538-8

90. Lurie N, Saville M, Hatchett R, Halton J. Developing Covid-19 Vaccines at Pandemic Speed. *N Engl J Med* (2020) 382:1969–73. doi: 10.1056/NEJMp2005630

91. Mandolesi M, Sheward DJ, Hanke L, Ma J, Pushparaj P, Vidakovics LP, et al. SARS-CoV-2 protein subunit vaccination elicits potent neutralizing antibody responses. *BioRxiv* (2020). doi: 10.1101/2020.07.31.228486

92. Corbett KS, Edwards DK, Leist SR, Abiona OM, Boyoglu-Barnum S, Gillespie RA, et al. SARS-CoV-2 mRNA vaccine design enabled by prototype pathogen preparedness. *Nature* (2020). doi: 10.1038/s41586-020-2622-0

93. Chen WH, Strych U, Hotez PJ, Bottazzi ME. The SARS-CoV-2 Vaccine Pipeline: an Overview. *Curr Trop Med Rep* (2020) 1–4. doi: 10.1007/s40475-020-00201-6

94. Jiang S, Bottazzi ME, Du L, Lustigman S, Tseng CT, Curti E, et al. Roadmap to developing a recombinant coronavirus S protein receptor-binding domain vaccine for severe acute respiratory syndrome. *Expert Rev Vaccines* (2012) 11:1405–13. doi: 10.1586/erv.12.126

95. Wang N, Shang J, Jiang S, Du L. Subunit Vaccines Against Emerging Pathogenic Human Coronaviruses. *Front Microbiol* (2020) 11:298. doi: 10.3389/fmicb.2020.00298

96. Premkumar L, Segovia-Chumbez B, Jadi R, Martinez DR, Raut R, Markmann A, et al. The receptor binding domain of the viral spike protein is an immunodominant and highly specific target of antibodies in SARS-CoV-2 patients. *Sci Immunol* (2020) 5:eabc8413. doi: 10.1126/sciimmunol.abc8413

97. Dai L, Zheng T, Xu K, Han Y, Xu L, Huang E, et al. A Universal Design of Betacoronavirus Vaccines against COVID-19, MERS, and SARS. *Cell* (2020) 182:722–33 e11. doi: 10.1016/j.cell.2020.06.035

98. Jiang S, Hillyer C, Du L. Neutralizing Antibodies against SARS-CoV-2 and Other Human Coronaviruses. *Trends Immunol* (2020) 41:355–9. doi: 10.1016/j.it.2020.03.007

99. Wong SK, Li W, Moore MJ, Choe H, Farzan M. A 193-amino acid fragment of the SARS coronavirus S protein efficiently binds angiotensin-converting enzyme 2. *J Biol Chem* (2004) 279:3197–201. doi: 10.1074/jbc.C300520200

100. Tai W, He L, Zhang X, Pu J, Voronin D, Jiang S, et al. Characterization of the receptor-binding domain (RBD) of 2019 novel coronavirus: implication for development of RBD protein as a viral attachment inhibitor and vaccine. *Cell Mol Immunol* (2020) 17:613–20. doi: 10.1038/s41423-020-0400-4

101. Prabakaran P, Gan J, Feng Y, Zhu Z, Choudhry V, Xiao X, et al. Structure of severe acute respiratory syndrome coronavirus receptor-binding domain complexed with neutralizing antibody. *J Biol Chem* (2006) 281:15829–36. doi: 10.1074/jbc.M600697200

102. Tortorici MA, Veesler D. Structural insights into coronavirus entry. *Adv Virus Res* (2019) 105:93–116. doi: 10.1016/bs.aivir.2019.08.002

103. Lu G, Hu Y, Wang Q, Qi J, Gao F, Li Y, et al. Molecular basis of binding between novel human coronavirus MERS-CoV and its receptor CD26. *Nature* (2013) 500:227–31. doi: 10.1038/nature12328

104. Walls AC, Xiong X, Park YJ, Tortorici MA, Snijder J, Quispe J, et al. Unexpected Receptor Functional Mimicry Elucidates Activation of Coronavirus Fusion. *Cell* (2019) 176:1026–39 e15. doi: 10.1016/j.cell.2018.12.028

105. Chen WH, Tao X, Agrawal A, Algaissi A, Peng BH, Pollet J, et al. Yeast-Expressed SARS-CoV Recombinant Receptor-Binding Domain (RBD219-N1) Formulated with Alum Induces Protective Immunity and Reduces Immune Enhancement. *bioRxiv* (2020). doi: 10.1101/2020.05.15.098079

106. Yang J, Wang W, Chen Z, Lu S, Yang F, Bi Z, et al. A vaccine targeting the RBD of the S protein of SARS-CoV-2 induces protective immunity. *Nature* (2020). doi: 10.1038/s41586-020-2599-8

107. Robbiani DF, Gaebler C, Muecksch F, Lorenzi JCC, Wang Z, Cho A, et al. Convergent antibody responses to SARS-CoV-2 in convalescent individuals. *Nature* (2020) 584:437–42. doi: 10.1038/s41586-020-2456-9

108. Chi X, Yan R, Zhang J, Zhang G, Zhang Y, Hao M, et al. A potent neutralizing human antibody reveals the N-terminal domain of the Spike protein of SARS-CoV-2 as a site of vulnerability. *Science* (2020) 369:650–5.

109. Brouwer P, Caniels T, van Straten K, Snitselaar J, Aldon Y, Bangaru S, et al. Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability. *Science* (2020) 369:643–50. doi: 10.1126/science.abc5902

110. Liu L, Wang P, Nair MS, Yu J, Rapp M, Wang Q, et al. Potent neutralizing antibodies directed to multiple epitopes on SARS-CoV-2 spike. *Nature* (2020) 584:450–6. doi: 10.1038/s41586-020-2571-7

111. Correia BE, Bates JT, Loomis RJ, Baneyx G, Carrico C, Jardine JG, et al. Proof of principle for epitope-focused vaccine design. *Nature* (2014) 507:201–6. doi: 10.1038/nature12966

112. Quinlan BD, Mou H, Zhang L, Guo Y, He W, Ojha A, et al. The SARS-CoV-2 receptor-binding domain elicits a potent neutralizing response without antibody-dependent enhancement. *bioRxiv* (2020). doi: 10.1101/2020.04.10.036418

113. Kwong PD. What Are the Most Powerful Immunogen Design Vaccine Strategies? A Structural Biologist's Perspective. *Cold Spring Harb Perspect Biol* (2017) 9: a029470. doi: 10.1101/cshperspect.a029470

114. Anasir MI, Poh CL. Structural Vaccinology for Viral Vaccine Design. *Front Microbiol* (2019) 10:738. doi: 10.3389/fmicb.2019.00738

115. Grifoni A, Weiskopf D, Ramirez SI, Mateus J, Dan JM, Moderbacher CR, et al. Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals. *Cell* (2020) 181:1489–501.e15. doi: 10.1016/j.cell.2020.05.015

116. Liu WJ, Zhao M, Liu K, Xu K, Wong G, Tan W, et al. T-cell immunity of SARS-CoV: Implications for vaccine development against MERS-CoV. *Antiviral Res* (2017) 137:82–92. doi: 10.1016/j.antiviral.2016.11.006

117. Yong CY, Ong HK, Yeap SK, Ho KL, Tan WS. Recent Advances in the Vaccine Development Against Middle East Respiratory Syndrome-Coronavirus. *Front Microbiol* (2019) 10:1781. doi: 10.3389/fmicb.2019.01781

118. Oja AE, Saris A, Ghandour CA, Kragten NA, Hogema BM, Nossent EJ, et al. Divergent SARS-CoV-2-specific T and B cell responses in severe but not mild COVID-19. *bioRxiv* (2020). doi: 10.1101/2020.06.18.159202

119. Ahmed SF, Quadeer AA, McKay MR. Preliminary Identification of Potential Vaccine Targets for the COVID-19 Coronavirus (SARS-CoV-2) Based on SARS-CoV Immunological Studies. *Viruses* (2020) 12:254. doi: 10.3390/v12030254

120. Henderson R, Edwards RJ, Mansouri K, Janowska K, Stalls V, Gobeil S, et al. Controlling the SARS-CoV-2 Spike Glycoprotein Conformation. *bioRxiv* (2020). doi: 10.1101/2020.05.18.102087

121. Hsieh CL, Goldsmith JA, Schaub JM, DiVenere AM, Kuo HC, Javanmardi K, et al. Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. *Science* (2020) 369:1501–5. doi: 10.1126/science.abd0826

122. Liu W, Chen YH. High epitope density in a single protein molecule significantly enhances antigenicity as well as immunogenicity: a novel strategy for modern vaccine development and a preliminary investigation about B cell discrimination of monomeric proteins. *Eur J Immunol* (2005) 35:505–14. doi: 10.1002/eji.200425749

123. Wang L, Shi W, Joyce MG, Modjarrad K, Zhang Y, Leung K, et al. Evaluation of candidate vaccine approaches for MERS-CoV. *Nat Commun* (2015) 6:7712. doi: 10.1038/ncomms8712

124. Lauring AS, Andino R. Quasispecies theory and the behavior of RNA viruses. *PLoS Pathog* (2010) 6:e1001005. doi: 10.1371/journal.ppat.1001005

125. Fauver JR, Petrone ME, Hodcroft EB, Shioda K, Ehrlich HY, Watts AG, et al. Coast-to-Coast Spread of SARS-CoV-2 during the Early Epidemic in the United States. *Cell* (2020) 181:990–6 e5. doi: 10.1016/j.cell.2020.04.021

126. Sevajol M, Subissi L, Decroly E, Canard B, Imbert I. Insights into RNA synthesis, capping, and proofreading mechanisms of SARS-coronavirus. *Virus Res* (2014) 194:90–9. doi: 10.1016/j.virusres.2014.10.008

127. Korber B, Fischer WM, Gnanakaran S, Yoon H, Theiler J, Abfalterer W, et al. Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell* (2020) 182:812–27. doi: 10.1016/j.cell.2020.06.043

128. Zhang L, Jackson CB, Mou H, Ojha A, Rangarajan ES, Izard T, et al. The D614G mutation in the SARS-CoV-2 spike protein reduces S1 shedding and increases infectivity. *bioRxiv* (2020). doi: 10.1101/2020.06.12.148726

129. Hu J, He CL, Gao Q, Zhang GJ, Cao XX, Long QX, et al. The D614G mutation of SARS-CoV-2 spike protein enhances viral infectivity. *bioRxiv* (2020). doi: 10.1101/2020.06.20.161323

130. Yurkovetskiy L, Pascal KE, Tompkins-Tinch C, Nyalile T, Wang Y, Baum A, et al. SARS-CoV-2 Spike protein variant D614G increases infectivity and retains sensitivity to antibodies that target the receptor binding domain. *bioRxiv* (2020). doi: 10.1101/2020.07.04.187757

131. Grubaugh ND, Hanage WP, Rasmussen AL. Making Sense of Mutation: What D614G Means for the COVID-19 Pandemic Remains Unclear. *Cell* (2020) 182:794–5. doi: 10.1016/j.cell.2020.06.040

132. Li Q, Wu J, Nie J, Zhang L, Hao H, Liu S, et al. The Impact of Mutations in SARS-CoV-2 Spike on Viral Infectivity and Antigenicity. *Cell* (2020) 182:1284–94. doi: 10.1016/j.cell.2020.07.012

133. Weissman D, Alameh M-G, LaBranche CC, Edwards RJ, Sutherland L, Santra S, et al. D614G Spike Mutation Increases SARS CoV-2 Susceptibility to Neutralization. *medRxiv* (2020). doi: 10.1101/2020.07.22.20159905

134. Mansbach RA, Chakraborty S, Nguyen K, Montefiori D, Korber B, Gnanakaran S. The SARS-CoV-2 Spike Variant D614G Favors an Open Conformational State. *bioRxiv* (2020). doi: 10.1101/2020.07.26.219741

135. Baum A, Fulton BO, Wloga E, Copin R, Pascal KE, Russo V, et al. Antibody cocktail to SARS-CoV-2 spike protein prevents rapid mutational escape seen with individual antibodies. *Science* (2020) 369:1014–8. doi: 10.1126/science.abd0831

136. Hansen J, Baum A, Pascal KE, Russo V, Giordano S, Wloga E, et al. Studies in humanized mice and convalescent humans yield a SARS-CoV-2 antibody cocktail. *Science* (2020) 369:1010–4. doi: 10.1126/science.abd0827

137. Jiang S, Hillyer C, Du L. Neutralizing Antibodies against SARS-CoV-2 and Other Human Coronaviruses: (Trends in Immunology 41, 355-359; 2020). *Trends Immunol* (2020) 41:545. doi: 10.1016/j.it.2020.04.008

138. Ho M. Perspectives on the development of neutralizing antibodies against SARS-CoV-2. *Antib Ther* (2020) 3:109–14. doi: 10.1093/abt/tbaa009

139. Yuan M, Wu NC, Zhu X, Lee CD, So RTY, Lv H, et al. A highly conserved cryptic epitope in the receptor binding domains of SARS-CoV-2 and SARS-CoV. *Science* (2020) 368:630–3. doi: 10.1126/science.abb7269

140. Lv Z, Deng YQ, Ye Q, Cao L, Sun CY, Fan C, et al. Structural basis for neutralization of SARS-CoV-2 and SARS-CoV by a potent therapeutic antibody. *Science* (2020) 369:1505–9. doi: 10.1126/science.abc5881

141. Wang C, Li W, Drabek D, Okba NMA, van Haperen R, Osterhaus A, et al. A human monoclonal antibody blocking SARS-CoV-2 infection. *Nat Commun* (2020) 11:2251. doi: 10.1038/s41467-020-16256-y

142. Li H, Zhou Y, Zhang M, Wang H, Zhao Q, Liu J. Updated Approaches against SARS-CoV-2. *Antimicrob Agents Chemother* (2020) 64:e00483–20. doi: 10.1128/AAC.00483-20

143. Saif LJ. Vaccines for COVID-19: perspectives, prospects, and challenges based on candidate SARS, MERS, and animal coronavirus vaccines. *Euro Med J* (2020). doi: 10.33590/emj/200324

144. Xu J, Jia W, Wang P, Zhang S, Shi X, Wang X, et al. Antibodies and vaccines against Middle East respiratory syndrome coronavirus. *Emerg Microbes Infect* (2019) 8:841–56. doi: 10.1080/22221751.2019.1624482

145. Conte C, Sogni F, Affanni P, Veronesi L, Argentiero A, Esposito S. Vaccines against Coronaviruses: The State of the Art. *Vaccines (Basel)* (2020) 8:309. doi: 10.3390/vaccines8020309

146. Juno JA, Tan H-X, Lee WS, Reynaldi A, Kelly HG, Wragg K, et al. Immunogenic profile of SARS-CoV-2 spike in individuals recovered from COVID-19. *medRxiv* (2020). doi: 10.1101/2020.05.17.20104869

147. Smith TRF, Patel A, Ramos S, Elwood D, Zhu X, Yan J, et al. Immunogenicity of a DNA vaccine candidate for COVID-19. *Nat Commun* (2020) 11:2601. doi: 10.1038/s41467-020-16505-0

148. Yu J, Tostanoski LH, Peter L, Mercado NB, McMahan K, Mahrokhian SH, et al. DNA vaccine protection against SARS-CoV-2 in rhesus macaques. *Science* (2020) 369:806–11. doi: 10.1126/science.abc6284

149. Ravichandran S, Coyle EM, Klenow L, Tang J, Grubbs G, Liu S, et al. Antibody signature induced by SARS-CoV-2 spike protein immunogens in rabbits. *Sci Transl Med* (2020). doi: 10.1126/scitranslmed.abc3539

150. Zhang J, Zeng H, Gu J, Li H, Zheng L, Zou Q. Progress and Prospects on Vaccine Development against SARS-CoV-2. *Vaccines (Basel)* (2020) 8:153. doi: 10.3390/vaccines8020153

151. Burton DR. What Are the Most Powerful Immunogen Design Vaccine Strategies? Reverse Vaccinology 2.0 Shows Great Promise. *Cold Spring Harb Perspect Biol* (2017) 9:a030262. doi: 10.1101/cshperspect.a030262

152. Graham BS, Gilman MSA, McLellan JS. Structure-Based Vaccine Antigen Design. *Annu Rev Med* (2019) 70:91–104. doi: 10.1146/annurev-med-121217-094234