Review article

Structural features of coronavirus SARS-CoV-2 spike protein: Targets for vaccination

Ariane Sternberg[a], Cord Naujokat[b],*

[a] *Center and Network for Targeted Oncology, Muehlackerweg 8, D-69239 Heidelberg, Germany*
[b] *Institute of Immunology, University of Heidelberg, Im Neuenheimer Feld 305, D-69120 Heidelberg, Germany*

ABSTRACT

Various human pathogenic viruses employ envelope glycoproteins for host cell receptor recognition and binding, membrane fusion and viral entry. The spike (S) glycoprotein of betacoronavirus SARS-CoV-2 is a homotrimeric class I fusion protein that exists in a metastable conformation for cleavage by host cell proteases furin and TMPRSS2, thereby undergoing substantial structural rearrangement for ACE2 host cell receptor binding and subsequent viral entry by membrane fusion. The S protein is densely decorated with N-linked glycans protruding from the trimer surface that affect S protein folding, processing by host cell proteases and the elicitation of humoral immune response. Deep insight into the sophisticated structure of SARS-CoV-2 S protein may provide a blueprint for vaccination strategies, as reviewed herein.

1. Introduction

Coronaviruses, such as pandemic SARS-CoV and SARS-CoV-2, are highly pathogenic for humans and can induce a severe acute respiratory syndrome (SARS). Infection with SARS-CoV-2 which emerged in December 2019 in Wuhan, China, can progress to severe pneumonia, multi organ failure and death (COVID-19 disease) [1–3]. Like other human pathogenic enveloped viruses, coronaviruses use unique envelope protein complexes for host cell receptor recognition and binding, and subsequent viral and host cell membrane fusion, leading to cell entry [4–10]. Host cell entry of coronaviruses is mediated by a transmembrane homotrimeric class I fusion glycoprotein, the spike protein (S protein), which exists in a metastable prefusion conformation and comprises two functional subunits for binding to the host cell receptor angiotensin-converting enzyme 2 (ACE2) (S1 subunit) and for fusion of the viral and host cell membranes (S2 subunit) [11,12]. The S protein of SARS-CoV is cleaved by a host cell protease, the transmembrane protease/serine subfamily member 2 (TMPRSS2), an airway and alveolar cell serine protease preferentially expressed on epithelial cells of the respiratory tract, such as type II pneumocytes [13–15]. TMPRSS2-mediated cleavage and priming of SARS-CoV S protein is required for binding to ACE2, membrane fusion and cell entry that requires a concerted action of a viral and host cell machinery comprising S protein, TMPRSS2 and ACE2 [13–15]. SARS-CoV-2 also employs TMPRSS2 for priming of its S protein (S) and S-driven cell entry *via* ACE2 [16–21].

Further, the S1/S2 boundary of SARS-CoV-2 S harbors multiple arginine residues not found in SARS-CoV and SARS-CoV-related S proteins. This S1/S2 boundary constitutes the cleavage site for the subtilisin-like host cell protease furin, which is ubiquitously expressed in humans [18,19,22].

The distal S1 subunit of S comprises the receptor-binding domains (RBDs) and contributes to stabilization of the prefusion state of the membrane-anchored S2 subunit that contains the fusion machinery [19]. For ACE2 receptor engagement, the RBDs located at the apex of S1 undergo hinge-like conformational movements that transiently expose (open status, "up") or hide (closed status, "down") the subdomains required for receptor binding, whereby the open status allows for receptor engagement, followed by shedding of S1 and refolding of S2 for membrane fusion [18,19]. Although the RBDs of the S1 subunit are more exposed on the viral surface than the S2 fusion machinery and are likely to be subject to selection pressure from immune surveillance, the S2 fusion machinery is densely decorated with heterogeneous N-linked glycans protruding from the S2 surface that may interfere with the elicitation of humoral immune responses and the accessibility to neutralizing antibodies [19]. In addition, the RBDs of S1 also contain N-linked glycans and unexpected O-linked glycans attached to the surface of S1 RBDs that also may interfere with the elicitation of neutralizing antibodies upon immune exposure or vaccination [23,24]. In individuals convalescent from COVID-19, the adaptive immunity to SARS-CoV-2 is largely mediated by CD4+ T cells with a T cell receptor
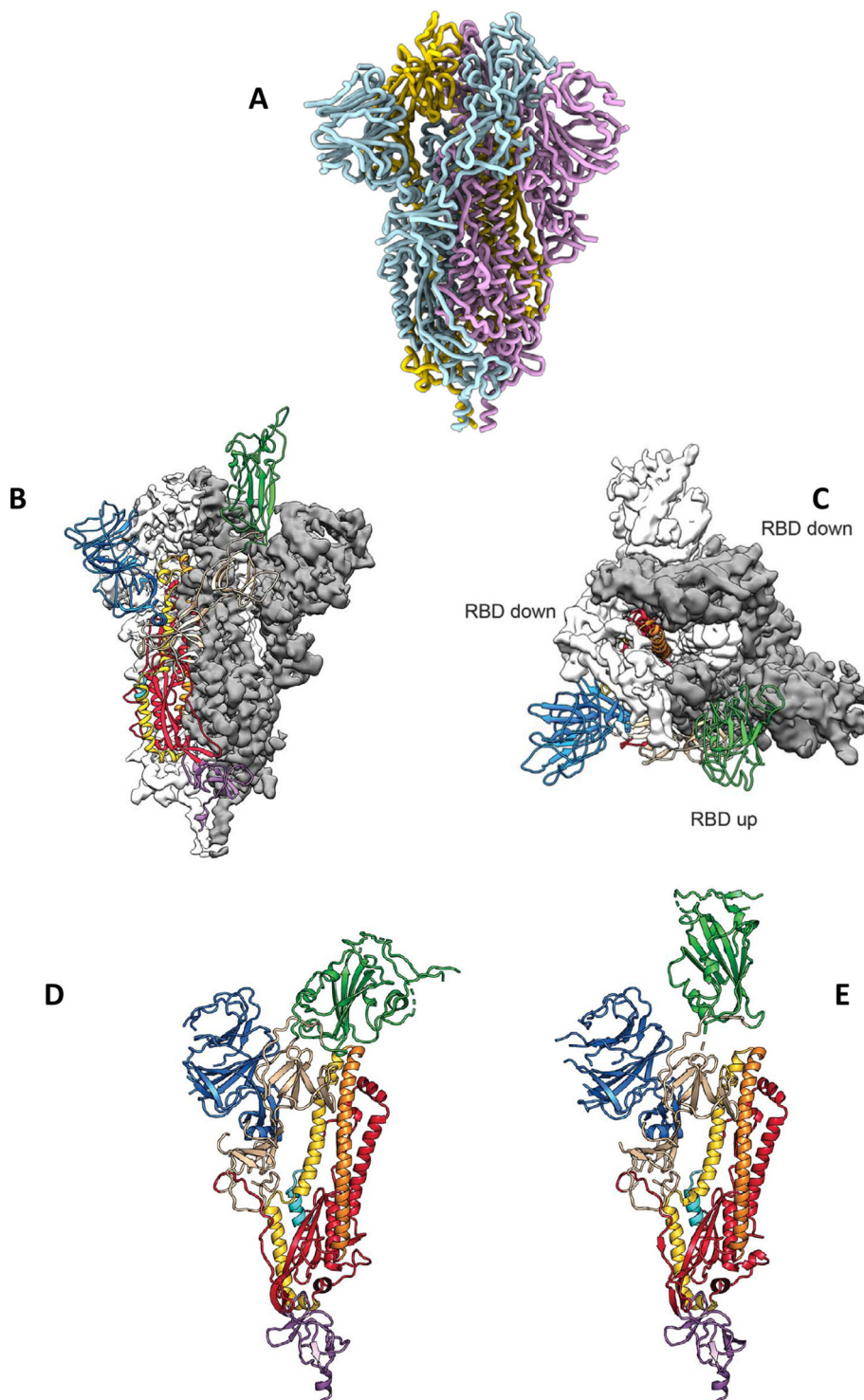
Fig. 1. Structural features of the SARS-CoV-2 spike (S) protein. (A) Ribbon diagram of the homotrimeric S, adopted from [19], (with permission from Elsevier Inc.). (B) Side view of the prefusion structure of S, with a single RBD in open ("up") conformation (green), adopted from [18], (with permission from Science.org). (C) Top view of the prefusion structure of S, with two single RBDs in closed ("down") conformation (white and grey) and one single RBD in open ("up") conformation (green), adopted from [18], (with permission from Science.org). (D) Single monomer of S, with the RBD in closed ("down") conformation (green), adopted from [18], (with permission from Science.org). (E) Single monomer of S, with the RBD in open ("up") conformation (green), adopted from [18], (with permission from Science.org). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

repertoire specific for S epitopes, leading to the robust generation of neutralizing IgG, IgM and IgA antibodies against the RBDs and the ectodomain trimer of S1 [25,26]. Further, a recently designed human monoclonal IgG1 neutralizing antibody raised against and binding to a conserved epitope of the RBDs of S prevents infection of host cells [27], finally underscoring that understanding the structural features of S is key for vaccine design and development against SARS-CoV-2 infection.

2. Structural features of the SARS-CoV-2 S protein

Using sophisticated approaches, including high-resolution cryogenic electron microscopy (cryo-EM) at < 4.0 Å, the labs of McLellan and Veesler recently uncovered the structural properties of SARS-CoV-2 S protein (S) [18,19] (Fig. 1A–C). S constitutes a tramsmembrane homotrimeric glycoprotein of ~180 kDa that belongs to the class I of trimeric fusion proteins found in other human pathogenic coronaviruses, including MERS-CoV and SARS-CoV. S is composed of two subunits, the apical V-shaped S1 ectotrimer subunit that harbors one ACE2-recognition motif per monomer (the receptor binding domain, RBD), and the S2 subunit required for fusion of the viral and cellular membranes (Fig. 2B, left) after being processed by the host cell protease furin at a polybasic cleavage site (with a four amino acid residue
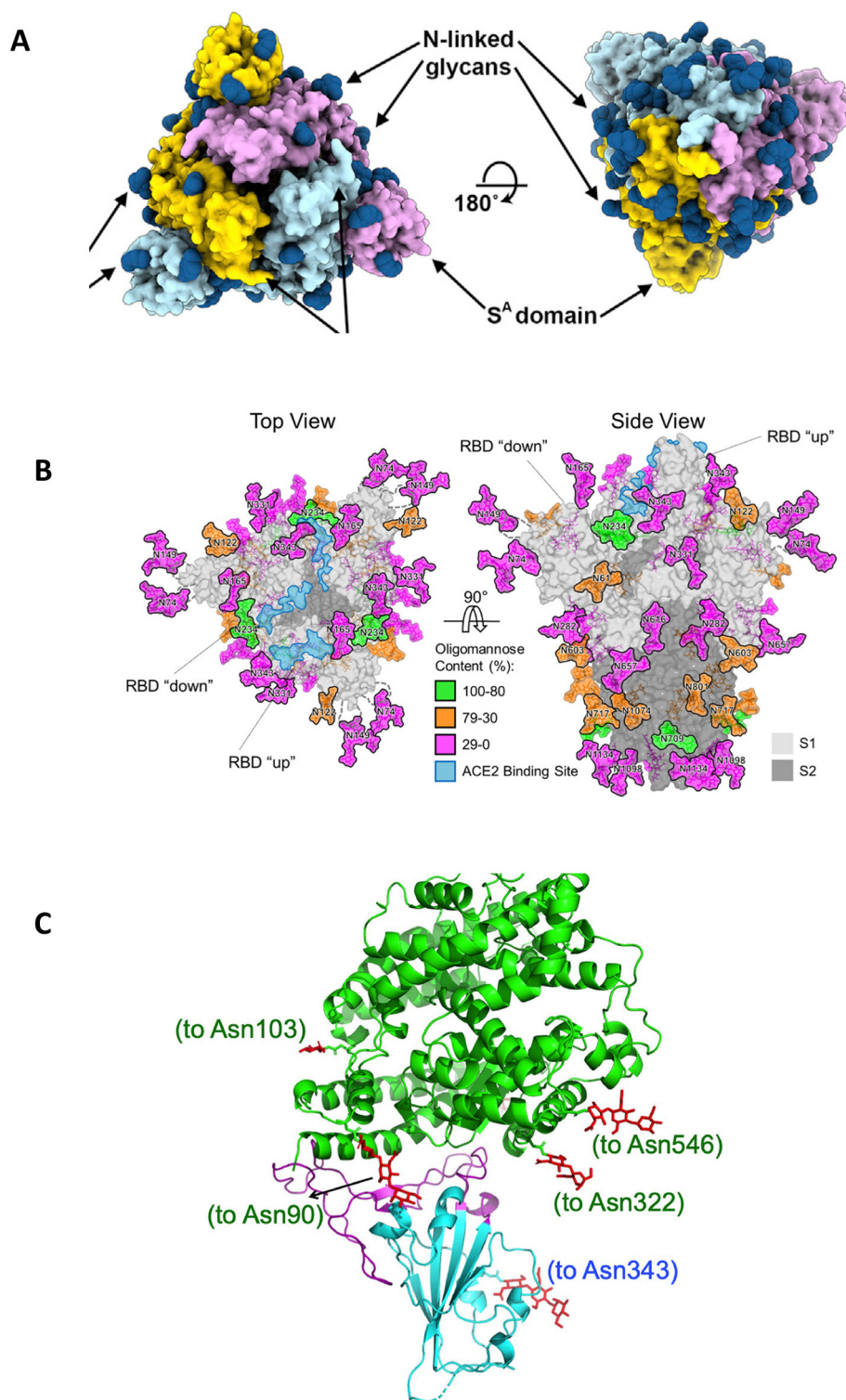
Fig. 2. N-linked glycosylation of S, and the RBD of S binding to ACE2. (A) left: top view of the S1 homotrimer, with N-linked glycans as dark blue spheres; (A) right: bottom view of the S2 homotrimer, with N-linked glycans as dark blue spheres. Adopted from [19], (with permission from Elsevier Inc.). (B) Left: top view of the S1 homotrimer, with N-linked glycans colored according to their oligomannose content (green to pink), with the ACE2 binding site in light blue. (B) Right: side view of the S homotrimer, with N-linked glycans colored according to their oligomannose content (green to pink), with the ACE2 binding site in light blue, and the S1 (light grey) and S2 (dark grey) subunits. Adopted from [36], (with permission from Science.org). (C) The RBD (light blue), with its twisted five-stranded antiparallel β sheet of β1, β2, β3, β4 and β7 strands, binding to the bottom side of the small lobe of ACE2 (green helices). Adopted from [34], (with permission from Nature.com). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

insertion, RRAR, at positions 681–684) that harbors multiple arginine residues and is located at the boundary between the S1 and S2 subunit [18,19,22]. Such polybasic cleavage sites are present in S proteins of human low pathogenic coronaviruses OC43 and HKU1, and in the S protein of the human high pathogenic coronavirus MERS.CoV [22], but are not present in SARS-CoV and SARS-CoV-related group 2b betacoronaviruses found in humans, civets, raccoon dog, pangolin and bats that possess a monobasic S1/S2 cleavage site processed upon entry of host cells [14,19,22,28–32]. The polybasic cleavage site of S may contribute to the high virulence of SARS-CoV-2, because furin and furin-like

proteases required for proteolytic activation of S are ubiquitously expressed in humans, providing expanded tissue tropism of SARS-CoV-2 [18,19,22]. All 9 N-linked glycans protruding from the surface of one S2 monomer (Fig. 2A, right, Fig. 2B, right) are conserved among SARS-CoV and SARS-CoV-2, and the N-linked glycosylation sequons in S2 are mostly conserved across glycoproteins of SARS-CoV-related viruses [19], suggesting that these structures of S2 interfere with the elicitation of neutralizing antibodies and promote immune evasion [24].

The S1 subunit of S is a 160-Å-long ectodomain trimer with a triangular cross-section [19] (Fig. 2A, left, Fig. 2B, left). At the apex of

each S1 monomer, one RBD for ACE2 engagement is located. The RBDs undergo hinge-like conformational movements that transiently expose (open or "up" status) (Fig. 1B, C, E) or hide (closed or "down" status) (Fig. 1C, D) the determinants of receptor binding; the open status is required for ACE2 engagement [18,19].

The structural features of the RBD required for binding to ACE2 were recently determined using high-resolution X-ray crystallography [33,34] (Fig. 2C). Therefore, the RBD (residues Arg319-Phe541) and the N-terminal peptidase domain of ACE2 (residues Ser19-Asp615) were expressed in insect cells and subsequently purified. The structure of the complex was determined by molecular replacements using the RBD of SARS-CoV S and domains of ACE2 as search models, and was refined to a resolution of 2.45 Å [33]. The final complex contained residues Thr333-Gly526 of the RBD of SARS-CoV-2 S and residues Ser19-Asp615 of the ACE2 N-terminal peptidase domain [33]. This strategy revealed that the RBD of SARS-CoV-2 S contains a twisted five-stranded antiparallel β sheet of β1, β2, β3, β4 and β7 strands with connecting helices and loops that build the core of the RBD [33]. In the core, between the β4 and β7 strands, there is an extended insertion which contains short β5 and β6 strands and α4 and α5 helices and loops. This extended insertion constitutes the receptor-binding motif (RBM) of the RBD that contains the contacting residues that bind to ACE2 [33]. The extended RBM contacts the bottom side of a small lobe of ACE2, with a concave outer surface in the RBM that accommodates the N-terminal helix of ACE2 [33,34] (Fig. 2C, blue part). Compared to the RBM of SARS-CoV S, the RBM of SARS-CoV-2 S forms a larger binding interface and more contacts with ACE2 as well as higher binding affinity to ACE2 (equilibrium dissociation constant, K_D, 4.7 nM *versus* 31 nM, respectively) [33–35], pointing out the higher infectivity and virulence of SARS-CoV-2 compared to SARS-CoV.

The S gene encodes 22 N-linked glycan sequons per monomer that affect S protein folding, processing by host cell proteases, elicitation of humoral immune response, and immune evasion [36]. Monomers of the S1 subunit are slightly more decorated with N-linked glycans (13) (Fig. 2A, left; Fig. B, left) than S2 monomers (9) (Fig. 2A, right; Fig. 2B, right), finally existing 66 N-linked glycans at N-linked glycosylation sites in one S homotrimer [19,36] (Fig. 2B, right). Compared to spike proteins of other human pathogenic coronaviruses, including HCoV-NL63 [37], MERS-CoV and SARS-CoV [38], and viral envelope glyco-proteins, such as HIV-1 envelope glycoprotein [39] and Lassa virus GPC [40], S is less densely glycosylated [36] that may limit immune evasion and promote the elicitation of humoral immunity.

3. Humoral immunity against S

Extensive coronavirus S protein glycan shielding that obstructs the protein surface contributes to epitope masking and immune evasion by hiding specific epitopes from antibody neutralization [36–38]. Since S is less densely decorated with N-linked glycans compared to S proteins of other human pathogenic coronaviruses [36–38], S is likely to be highly immunogenic and a major target of neutralizing antibodies. Despite the high degree of structural homology between the S protein RBD of SARS-CoV and SARS-CoV-2, certain existing monoclonal antibodies raised against the SARS-CoV RBD (S230, m396, and 80R) failed to bind the RBD of SARS-CoV-2 S [18,33], suggesting that antibody cross-reactivity may be limited between the RBD of SARS-CoV and SARS-CoV-2. In contrast, sera obtained from patients convalescent from SARS-CoV infection, and rabbit sera raised against the S1 subunit of SARS-CoV S protein inhibited S-driven entry into simian Vero target cells [16]. Sera from mice immunized with a stabilized SARS-CoV S protein also significantly inhibited cell entry of SARS-CoV-2 into target cells [19], indicating that cross-neutralizing antibodies targeting conserved epitope of S proteins can be elicited upon vaccination. This is in line with the recent finding that CR3022, a human antibody isolated from a convalescent SARS-CoV patient and targeting the RBD of SARS-CoV S protein, binds to a highly conserved epitope distal from the ACE2

receptor binding site that enables cross-reactive binding between SARS-CoV-2 and SARS-CoV S proteins. Structural modeling approaches further demonstrate that the binding epitope can only be accessed by CR3022 when at least two RBDs of S protein are in the open conformation and slightly rotated [41]. Similar and extended results were obtained using a monoclonal antibody (S309) isolated from a convalescent SARS-CoV individual that potently neutralizes SARS-CoV-2 by engaging the RBD of S [42], indicating cross-neutralization by antibodies obtained from convalescent SARS-CoV individuals on SARS-CoV-2 by engagement of conserved S protein epitopes.

Consequently, monoclonal antibodies obtained from convalescent SARS-CoV-2 (COVID-19) individuals display neutralization activities against SARS-CoV-2 by targeting highly immunogenic epitopes of S, such as the RBD. In individuals convalescent from COVID-19, adaptive immune responses to SARS-CoV-2 are mediated by CD4 + T cells with a T cell receptor repertoire specific for S epitopes, leading to the robust generation of neutralizing IgG, IgM and IgA antibodies against the RBD and the ectodomain trimer of S [25,26]. By high-throughput single-cell RNA and VDJ sequencing of antigen-enriched B cells from 60 convalescent patients, various potent neutralizing antibodies were identified with the most potent one, BD-368-2, exhibiting strong neutralizing activity against SARS-CoV-2 [26]. Other monoclonal antibodies, B38 and H4, isolated from an individual convalescent from COVID-19 display neutralizing activity against SARS-CoV-2 by binding to the RBD-ACE2 interface [43], and RBD-specific monoclonal antibodies derived from single B cells of SARS-CoV-2 infected individuals exhibit potent neutralization activity that correlates with their competitive capacity with ACE2 for RBD binding [44]. Surprisingly, these monoclonal antibodies failed to bind the RBD of SARS-CoV and MERS-CoV S proteins [44], pointing out their specificity to the RBD of S.

A recently developed hybridoma-derived humanized monoclonal IgG1 neutralizing antibody (47D11) binds to the receptor-binding subdomain (residues 438–498) of S that loops out from the antiparallel β sheet core domain structure of the RBM of S that directly engages the binding domain of ACE2 [27]. Infection of simian VeroE6 cells with SARS-CoV-2 was effectively inhibited and neutralized at an IC_{50} value of 0.57 μg/ml [27], demonstrating high neutralizing activity of 47D11. Immunization of llama camelids, which are able to produce heavy-chain-only antibodies with a single variable domain (VHH) instead of two variable domains (VH and VL) that make up the equivalent antigen-binding fragment (Fab) of conventional immunoglobulin G (IgG) antibodies, with prefusion-stabilized SARS-CoV-1 S protein and MERS-CoV S protein in an alternating mode, resulted in the obtainment of cross-neutralizing VHHs targeting the RBD of S [45]. After engineering the VHH antibody into a bivalent Fc-fusion, the antibody construct was able to neutralize SARS-CoV-2 S pseudoviruses [45], revealing that S proteins of various coronaviruses including SARS-CoV-2 are highly immunogenic and can elicit effective humoral immune responses across mammalian species.

Recent work describes detection and isolation of potent neutralizing monoclonal antibodies from humans convalescent from SARS-CoV-2 infection and COVID-19 disease that may be engineered and used for passive immunization and therapeutic intervention [44,46–49]. For instance, monoclonal antibodies derived from single B cells of SARS-CoV-2 infected individuals showed potent anti-SARS-CoV-2 neutralization activity that correlated with their competitive capacity with ACE2 for RDB binding [44]. SARS-CoV-2-neutralizing monoclonal antibodies isolated from infected patients hospitalized with severe COVID-19 disease displayed strong binding to the RBD and the N-terminal domain (NTD) of S, indicating that both of these S epitopes at the apex of S are highly immunogenic [46]. Another recent study demonstrated exclusive NTD specificity (epitope 4A8 of NTD) of neutralizing monoclonal antibodies isolated from convalescent COVID-19 patients [47], whereas predominant molecular targets of neutralizing monoclonal antibodies isolated from convalescent COVID-19 patients seem to be epitopes of the RBD of S [48,49] that correlates with the adaptive

CD4 + T cell-mediated immune response to the RBD of S [25].

4. Vaccination strategies using S

Outbreak and pandemic of the betacoronaviruses SARS-CoV (2002/2003 in China) and MERS-CoV (2012 in Saudi Arabia) has led to the design and development of vaccination strategies mainly using recombinant viral S proteins as antigen [50–52].

Due to its high antigenicity and its proven ability to elicit robust humoral immune responses and neutralizing antibodies in individuals convalescent from SARS-CoV-2 infection and COVID-19 disease [25,26,43,44], S appears as an ideal candidate for vaccination against SARS-CoV-2 infection [53–56], and constitutes an improved immunogen when stabilized in its prefusion conformation [18]. In a high-yield production approach, more than 100 structure-guided S variants based upon a previously determined cryo-EM structure of the prefusion S were designed, expressed and produced in Chinese hamster ExpiCHO cells [57]. The best prefusion S variant, termed HexaPro, was extremely stable in the prefusion state, retained the S2 subunit conformation and preserved its high antigenicity due to its stable prefusion conformation [57].

Recently, two synthetic DNA-based vaccine candidates expressing different forms of S were developed and investigated in rhesus macaques [58], and mice and guinea pigs [59]. A series of prototype DNA vaccines expressing six variants of the S: 1) full-length S, 2) deletion of the cytoplasmic tail of S, 3) deletion of the transmembrane domain and cytoplasmic tail reflecting the soluble ectodomain of S, 4) S1 domain with a foldon trimerization tag, 5) RBD of S with a foldon trimerization tag, and 6) a prefusion stabilized soluble ectodomain of S with deletion of the furin cleavage site, two proline mutations and a foldon trimerization tag, were produced [58]. Adult rhesus macaques in groups of 4 animals were immunized with one of the six prototype DNA vaccines, respectively, and each animal received 5 mg DNA vaccine by the intramuscular route without adjuvant at week 0 and week 3; ten animals not vaccinated served as control group [58]. After a boost immunization at week 5, S-specific binding antibodies and neutralizing antibodies (NAb) could be obtained from the animals, with median titers of the NAb comparable in the magnitude to NAb titers in a cohort of 9 convalescent macaques and in a cohort of 27 humans convalescent from SARS-CoV-2 infection [58]. Further, a Th1-biased cellular immune response of S-specific IFN-γ + CD4 + T cells to pooled S peptides was detected in the majority of vaccinated animals at week 5 [58]. Three weeks after the boost immunization, animals of the vaccine group and the control group were challenged with SARS-CoV-2, administered by the intranasal and the intratracheal route. In the broncho-alveolar lavage (BAL) and nasal swabs (NS) of the control group, high levels of viral RNA could be detected as compared to significant lower viral RNA levels in the vaccine group, and 8 of 25 vaccinated animals exhibited no detectable viral RNA in BAL and NS at any timepoint following the challenge [58], demonstrating high protective efficacy of the S-expressing DNA vaccine against intranasal and intratracheal SARS-CoV-2 infection in rhesus macaques.

In a similar study, a synthetic DNA plasmid, termed pGX9501/INO-4800, was designed to encode S that matches with > 99.9% amino acid sequence identity of the recently published S sequences [59]. Intramuscular administration of INO-4800 in Balb/c mice on days 0 and 14 resulted in the elicitation of neutralizing IgG antibodies at day 21 that bind to S protein antigens, including S1 and S2 subunits, and RBD, as well as to the S-ACE2 interface, with limited cross-reactivity to SARS-CoV S protein antigens [59]. Similar results were obtained with INO-4800 in Hartley guinea pigs. Neutralizing antibodies were found in the BAL fluids of the animals at day 28 after vaccination, revealing strong lung tropism [56]. Moreover, a cellular immune response against S epitopes mediated by CD4 + and CD8 + IFN-γ + T cells was detected on day 14 after vaccination [59].

Furthermore, a vaccine candidate consisting in chimpanzee-derived adenoviral vector (ChAdOx1), expressing full-length S (GenBank accession number YP_009724390.1) and termed ChAdOx1 nCoV-19, has been shown to induce a robust humoral and cellular immune response in rhesus macaques after a single vaccination [60]. S-specific neutralizing antibodies and T-cell responses against full-length S could be detected 14 days post vaccination. A significantly reduced viral load in broncho-alveolar lavage fluid and respiratory tract tissue of vaccinated animals challenged with SARS-CoV-2 compared with control animals, and no pneumonia was observed in vaccinated rhesus macaques [60]. ChAdOx1 nCoV-19 is currently under investigation in a phase II/III clinical trial in the UK [60].

In a dose-escalation, open label, non-randomized, first-in-human trial, 108 healthy humans with mean age of 36 years received single low dose (n = 36), middle dose (n = 36) and high dose (n = 36) of a recombinant replication-defective adenovirus type-5 (Ad5) vectored vaccine expressing S, with full-length S gene based on Wuhan-Hu-1 strain (GenBank accession number YP_009724390) [61]. This resulted in the occurrence of frequent (in more than 80% of the participants) adverse reactions within the first 7 days after vaccination, including fever, fatigue, headache and muscle pain, but elicited robust cellular and humoral immune responses, with neutralizing antibodies binding to the RBD of S (from day 14, peaking at 28 after vaccination), and CD4 + and CD8 + T cells specific for S epitopes (from day 14 after vaccination) [61]. These inaugural results suggest that vaccination of humans using S as antigen will be successful. Consequently, a phase II clinical trial with Ad5 at low or middle dose has started in China [62], and Canada has approved a phase I/II clinical trial in humans with Ad5 [63]. However, pre-existing immunity against the Ad5 vector could reduce immunogenicity, potentially limiting efficacy in populations in which adenovirus type-5 is endemic, with a reported seroprevalence of 30–80% [64,65].

Currently, other vaccine candidates based on S or its RBD are in rapid development (Table 1), and different antigen delivery platforms, including recombinant protein vaccines, replicating or non-replicating viral-vector-based vaccines, and DNA or mRNA vaccines are under investigation [53–56,66–69]. An overview of the current status of the pipeline of COVID-19 vaccine candidates is provided in Ref. [55] and Refs. [66–69]. For example, phase I/II clinical first-in-humans trials using m-RNA vaccines encoding full-length S, the RBD of S [70–72] or a recombinant trimeric S subunit for vaccination [73] have been started recently (Table 1), pointing out a pivotal role of S and its RBD for vaccination against SARS-CoV-2.

5. Conclusion

The outbreak of coronavirus SARS-CoV-2 in Wuhan, China, in December 2019, the cause of COVID-19 disease, represents a pandemic threat to global health and has major consequences on global economy if SARS-CoV-2 spread and virulence is not contained, or effective treatments are not developed [74]. The pandemic has spread to more than 188 countries with more than 10,000,000 confirmed cases, more than 500,000 confirmed deaths and more than 5,000,000 total recoveries worldwide as of June 30th 2020 [75]. Since no specific drug against SARS-CoV-2 infection or COVID-19 disease is available or approved to date [76,77], it is mandatory to rapidly develop and provide successful vaccines against SARS-CoV-2 that should be available soon for large populations. Usually, it takes 10–15 years of vaccine development by the classical way, using inactivated or live attenuated vaccines after the generation of long-term safety and efficacy data [78–80]. Clinical development of vaccines begins with phase I trials to evaluate the safety of vaccine candidates, followed by phase II trials to establish doses and formulations to prove efficacy, and finally followed by phase III trials to prove and demonstrate safety and efficacy in larger human cohorts [53]. In an extraordinary situation like the COVID-19 pandemic, this procedure should be compressed, and an accelerated regulatory approval pathway might be developed [53].

Table 1
Examples of vaccines targeting S that are in clinical development.

Vaccine	Type	Target/Antigen	Clinical Status	Reference
ChAdOx1 nCoV-19	Chimpanzee-derived adonoviral vectored vaccine expressing full-length S	Full-length S, (GenBank accession number YP_009724390.1)	Phase II/III clinical trial. UK, University of Oxford	[60]
Ad5	Recombinant replication-defective adenovirus type-5 (Ad5) vectored vaccine expressing full-length S	Full-length S, based on Wuhan-Hu-1 strain (GenBank accession number YP_009724390)	Phase I, dose-escalation, open label, non-randomized, first-in-human clinical trial. China	[61]
Ad5	Recombinant replication-defective adenovirus type-5 (Ad5) vectored vaccine expressing full-length S	Full-length S, based on Wuhan-Hu-1 strain (GenBank accession number YP_009724390)	Phase II, low and middle dose. China	[62]
Ad5	Recombinant replication-defective adenovirus type-5 (Ad5) vectored vaccine expressing full-length S	Full-length S, based on Wuhan-Hu-1 strain (GenBank accession number YP_009724390)	Phase I/II, low and middle dose. Canada	[63]
mRNA-1273	Lipid nanoparticle (LNP)-encapsulated mRNA-based vaccine encoding full-length, prefusion stabilized S	Full-length, prefusion stabilized S	Phase II/III clinical trial. USA	[70,71]
Different mRNAs: BNT 162a1, BNT 162b1, BNT 162b2, BNT 162c2	Lipid nanoparticle (LNP)-encapsulated mRNA-based vaccines encoding either full-length S or RBD of S	Full length S, or RBD of S	Phase I/II, randomized, placebo-controlled, observer-blind, dose-finding clinical trial. Low, middle and high dose. USA, Germany	[72]
SCB-2019	Recombinant trimeric S subunit, with different adjuvants	Trimeric S subunit	Phase I, randomized, double blind, placebo controlled, first-in-human clinical trial. USA	[73]

The spike protein of SARS-CoV-2 (S) is a homotrimeric class I fusion protein protruding from the viral surface that is required for host cell receptor (ACE2) recognition and binding, and fusion of the viral and cellular membrane, leading to viral cell entry. Since S is highly exposed on the viral surface, it is likely to be subject to immune surveillance by T cells and professional antigen-presenting cells, leading to the elicitation of neutralizing antibodies against specific epitopes and domains of S. This is finally not surprising, because the human cellular and humoral immune system is capable of generating a response very specific to the structure of a foreign invading virus or protein, and the general adaptive immune response consisting of antigen capture and presentation by professional antigen-presenting cells, followed by T cell receptor-mediated antigen recognition and B cell-driven production of antibodies specific for the given antigen, is fortunately also operative in SARS-CoV-2 infection, obviously by choosing S epitopes as major antigens [25,81,82].

S is less densely decorated with N-linked glycans compared to S proteins of other human pathogenic coronaviruses, which display extensive S protein glycan shielding that obstructs the protein surface and thereby contributes to epitope masking and immune evasion. Therefore, S appears highly immunogenic and as a target for vaccination. In fact, a considerable number of recent studies discussed herein have demonstrated high immunogenicity of S and its RBD, leading to adaptive T cell-mediated immune responses and, finally, to the elicitation of neutralizing antibodies in humans and various mammals that can prevent SARS-CoV-2 infection and rechallenge.

In conclusion, deep understanding of the structural features of S will facilitate the design and development of successful vaccines against coronavirus SARS-CoV-2 for large populations.

Declaration of competing interest

None.

References

[1] J. Liu, X. Zheng, Q. Tong, W. Li, B. Wang, K. Sutter, M. Trilling, M. Lu, U. Dittmer, D. Yang, Overlapping and dicrete aspects of the pathology and pathogenesis of the emerging human pathogenic coronaviruses SARS-CoV, MERS-CoV, and 2019-nCoV, J. Med. Virol. 92 (2020) 491–494.

[2] C. Jiang, X. Yao, Y. Zhao, J. Wu, P. Huang, C. Pan, S. Liu, C. Pan, Comparative review of respiratory diseases caused by coronaviruses and influenza a viruses during epidemic season, Microbes. Infect. May 13 (2020), https://doi.org/10.1016/j.micinf.2020.05.005.

[3] Y. Yang, F. Peng, R. Wang, K. Guan, T. Jiang, G. Xu, J. Sun, C. Chang, The deadly coronaviruses: the 2003 SARS pandemic and the 2020 novel coronavirus epidemic in China, J. Autoimmun. 109 (2020) 102434.

[4] U.F. Greber, Virus and host mechanics support membrane penetration and cell entry, J. Virol. 90 (2016) 3802–3805.

[5] Y. Tong, D. Lavillette, Q. Li, J. Zhong, Role of hepatitis C virus envelope glycoprotein E1 in virus entry and assembly, Front. Immunol. 9 (2018) 1411.

[6] J. Arii, Y. Kawaguchi, The role of HSV glycoproteins in mediating cell entry, Adv. Exp. Med. Biol. 1045 (2018) 3–21.

[7] E. Malito, S. Chandramouli, A. Carfi, From recognition to execution – the HCMV pentamer from receptor binding to fusion triggering, Curr. Opin. Virol. 31 (2018) 43–51.

[8] K.D. Azam, B. Lee, Differential features of fusion activation within the *Paramyxoviridae*, Viruses 12 (2020) 161.

[9] J.K. Millet, G.R. Whittaker, Physiological and molecular triggers for SARS-CoV membrane fusion and entry into host cells, Virology 517 (2018) 3–8.

[10] D.J. Schoeman, B.C. Fielding, Coronavirus envelope protein: current knowledge, Virol. J. 16 (2019) 69.

[11] A.C. Walls, M.A. Tortorici, B.J. Bosch, B. Frenz, P.J.M. Rottier, F. DiMaio, F.A. Rey, D. Veesler, Cryo-electron microscopy structure of a coronavirus spike glycoprotein trimer, Nature 531 (2016) 114–117.

[12] R.N. Kirchdoerfer, C.A. Cottrell, N. Wang, J. Pallesen, H.M. Yassine, H.L. Turner, K.S. Corbett, B.S. Graham, J.S. McLellan, A.B. Ward, Pre-fusion structure of a human coronavirus spike protein, Nature 531 (2016) 118–121.

[13] S. Matsuyama, N. Nagata, K. Shirato, M. Kawase, M. Takeda, F. Tagushi, Efficient activation of the severe acute respiratory syndrome coronavirus spike protein by the transmembrane protease TMPRSS2, J. Virol. 84 (2010) 12685-12664.

[14] A. Shulla, T. Heald-Sargent, G. Subramanya, J. Zhao, S. Perlman, T. Gallagher, A transmembrane serine protease is linked to the severe acute respiratory syndrome coronavirus receptor and activates virus entry, J. Virol. 85 (2011) 873–882.

[15] I. Glowacka, S. Bertram, M.A. Müller, P. Allen, E. Soilleux, S. Pfefferle, et al., Evidence that TMPRSS2 activates the severe acute respiratory syndrome coronavirus spike protein for membrane fusion and reduces viral control by the humoral immune response, J. Virol. 85 (2011) 4122–4134.

[16] M. Hoffmann, H. Kleine-Weber, S. Schroeder, N. Krüger, T. Herrler, S. Erichsen, et al., SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor, Cell 181 (2020) 271–280.

[17] P. Zhou, X.L. Yang, X.G. Wang, B. Hu, L. Zhang, W. Zhang, et al., A pneumonia outbreak associated with a new coronavirus of probable bat origin, Nature 579 (2020) 270–273.

[18] D. Wrapp, N. Wang, K.S. Corbett, J.A. Goldsmith, C.L. Hsieh, O. Abiona, et al., Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation, Science 367 (2020) 1260–1263.

[19] A.C. Walls, Y.J. Park, M.A. Tortorici, A. Wall, A.T. McGuire, D. Veesler, Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein, Cell 181 (2020) 281–292.

[20] C.G.K. Ziegler, S.J. Allon, S.K. Nyquist, I.M. Mbano, V.N. Miao, C.N. Tzouanas, et al., SARS-CoV-2 receptor ACE2 is an interferon-stimulated gene in human airway epithelial cells and is detected in specific cell subsets across tissues, Cell Apr 27 (2020), https://doi.org/10.1016/j.cell.2020.04.035.

[21] Q. Wang, Y. Zhang, L. Wu, S. Niu, C. Song, Z. Zhang, et al., Structural and functional

basis of SARS-CoV-2 entry by using human ACE2, Cell 181 (2020) 894–904.

[22] M. Hoffmann, H. Kleine-Weber, S. Pöhlmann, A multibasic cleavage site in the spike protein of SARS-CoV-2 is essential for infection of human lung cells, Mol. Cell 78 (2020) 1–6.

[23] A. Shajahan, N.T. Supekar, A.S. Gleinich, P. Azadi, Deducing the N- and O- glyco-sylation profile of the spike protein of novel coronavirus SARS-CoV-2, Glycobiology (2020), https://doi.org/10.1093/glycob/cwaa042 May 4.

[24] D.J. Vigerust, V.L. Shepherd, Virus glycosylation: role in virulence and immune interactions, Trends Microbiol. 15 (2007) 211–218.

[25] A. Grifoni, D. Weiskopf, S.I. Ramirez, J. Mateus, J.M. Dan, C. Rydyznski Moderbacher, et al., Targets of T cell responses to SARS-CoV-2 coronavirus in hu-mans with COVID-19 disease and unexposed individuals, Cell 181 (2020) 1489–1501.

[26] Y. Cao, B. Su, X. Guo, W. Sun, Y. Deng, L. Bao, et al., Potent neutralizing antibodies against SARS-CoV-2 identified by high-throughput single-cell sequencing of con-valescent patients' B cells, Cell May 13 (2020), https://doi.org/10.1016/j.cell.2020. 05.025.

[27] C. Wang, W. Li, D. Drabek, N.M.A. Okba, R. van Haperen, A.D.M.E. Osterhaus, et al., A human monoclonal antibody blocking SARS-CoV-2 infection, Nat. Commun. 11 (2020) 2251.

[28] B.J. Bosch, W. Bartelink, P.J. Rottier, Cathepsin L functionally cleaves the severe acute respiratory syndrome coronavirus class I fusion protein upstream of rather than adjacent to the fusion peptide, J. Virol. 82 (2008) 8887–8890.

[29] S. Belouzard, V.C. Chu, G.R. Whittaker, Activation of the SARS coronavirus spike protein via sequential proteolytic cleavage at two distinct sites, Proc. Natl. Acad. Sci. U. S. A. 106 (2009) 5871–5876.

[30] T.T. Lam, M.H. Shum, H.C. Zhu, Y.G. Tong, X.B. Ni, Y.S. Liao, et al., Identifying SARS-CoV-2 related coronaviruses in Malayan pangolins, Nature Mar 26 (2020), https://doi.org/10.1038/s41586-020-2169-0.

[31] T. Zhang, Q. Wu, Z. Zhang, Probable pangolin origin of SARSCoV-2 associated with the COVID-19 outbreak, Curr. Biol. 30 (2020) 1346–1351.

[32] X. Li, J. Zai, Q. Zhao, Q. Nie, Y. Li, B.T. Foley, et al., Evolutionary history, potential intermediate animal host, and cross-species analyses of SARS-CoV-2, J. Med. Virol. Feb 27 (2020), https://doi.org/10.1002/jmv.25731.

[33] J. Lan, J. Ge, J. Yu, S. Shan, H. Zhou, S. Fan, et al., Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor, Nature 581 (2020) 215–220.

[34] J. Shang, G. Ye, K. Shi, Y. Wan, C. Luo, H. Aihara, et al., Structural basis of receptor recognition by SARS-CoV-2, Nature 581 (2020) 221–224.

[35] J. Shang, Y. Wan, C. Luo, G. Ye, Q. Geng, A. Auerbach, et al., Cell entry mechanisms of SARS-CoV-2, Proc. Natl. Acad. Sci. U. S. A. (2020), https://doi.org/10.1073/pnas.2003138117 May 6.

[36] Y. Watanabe, J.D. Allen, D. Wrapp, J.S. McLellan, M. Crispin, Site-specific glycan analysis of the SARS-CoV-2 spike, Science (2020), https://doi.org/10.1126/science.abb9983 May 4.

[37] A.C. Walls, M.A. Tortorici, B. Frenz, J. Snijder, W. Li, F.A. Rey, et al., Glycan shield and epitope masking of a coronavirus spike protein observed by cryo-electron mi-croscopy, Nat. Struct. Mol. Biol. 23 (2016) 899–905.

[38] A.C. Walls, X. Xiong, Y.J. Park, M.A. Tortorici, J. Snijder, J. Quispe, et al., Unexpected receptor functional mimicry elucidates activation of coronavirus fu-sion, Cell 176 (2019) 1026–1039.

[39] A.J. Behrens, S. Vasiljevic, L.K. Pritchard, D.J. Harvey, R.S. Andev, S.A. Krumm, et al., Composition and antigenic effects of individual glycan sites of a trimeric HIV-1 envelope glycoprotein, Cell Rep. 14 (2016) 2695–2706.

[40] Y. Watanabe, J. Raghwani, J.D. Allen, G.E. Seabright, S. Li, F. Moser, et al., Structure of the Lassa virus glycan shield provides a model for immunological re-sistance, Proc. Natl. Acad. Sci. U. S. A. 115 (2018) 7320–7325.

[41] M. Yuan, N.C. Wu, X. Zhu, C.C.D. Lee, R.T.Y. So, H. Lv, et al., A highly conserved cryptic epitope in the receptor binding domains of SARS-CoV-2 and SARS-CoV, Science 368 (2020) 630–633.

[42] D. Pinto, Y.J. Park, M. Beltramello, A.C. Walls, M.A. Tortorici, S. Bianchi, et al., Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody, Nature (2020), https://doi.org/10.1038/s41586-020-2349-y May 18.

[43] Y. Wu, F. Wang, C. Shen, W. Peng, D. Li, C. Zhao, et al., A noncompeting pair of human neutralizing antibodies block COVID-19 virus binding to its receptor ACE2, Science (2020), https://doi.org/10.1126/science.abc2241 May 13.

[44] B. Ju, Q. Zhang, J. Ge, R. Wang, J. Sun, X. Ge, et al., Human neutralizing antibodies elicited by SARS-CoV-2 infection, Nature (2020), https://doi.org/10.1038/s41586-020-2380-z May 26.

[45] D. Wrapp, D. De Vlieger, K.S. Corbett, G.M. Torres, N. Wand, W. Van Breedam, et al., Structural basis for potent neutralization of betacoronaviruses by single-do-main camelid antibodies, Cell (2020), https://doi.org/10.1016/j.cell.2020.04.031 May 5.

[46] L. Liu, P. Wang, M.S. Nair, J. Yu, Y. Huang, M.A. Rapp, et al., Potent Neutralizing Monoclonal Antibodies Directed to Multiple Epitopes on the SARS-CoV-2 Spike, bioRxiv Jun 18 (2020), https://doi.org/10.1101/2020.06.17.153486 06.17. 153486.

[47] X. Chi, R. Yan, J. Zhang, G. Zhang, Y. Zhang, M. Hao, et al., A neutralizing human antibody binds to the N-terminal domain of the spike protein of SARS-CoV-2, Science (2020) eabc6952, https://doi.org/10.1126/science.abc6952 Jun 22.

[48] P.J.M. Brouwer, T.G. Caniels, K. van der Straten, J.L. Snitselaar, Y. Aldon, S. Bangaru, et al., Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability, Science (2020) eabc5902, https://doi.org/10. 1126/science.abc5902 Jun 15.

[49] R. Shi, C. Chan, X. Duan, Z. Chen, P. Liu, J. Song, et al., A human neutralizing antibody targets the receptor binding site of SARS-CoV-2, Nature (2020), https://

doi.org/10.1038/s41586-020-2381-y May 26.

[50] J.E. Martin, M.K. Louder, L.A. Holman, I.J. Gordon, M.E. Enama, B.D. Larkin, et al., A SARS DNA vaccine induces neutralizing antibody and cellular immune responses in healthy adults in a phase I clinical trial, Vaccine 26 (2008) 6338–6343.

[51] C.Y. Wong, H.K. Ong, S.K. Yeap, K.L. Ho, W.S. Tan, Recent advances in the vaccine development against middle east respiratory syndrome-coronavirus, Front. Microbiol. 10 (2019) 1781.

[52] J. Pallesen, N. Wang, K.S. Corbett, D. Wrapp, R.N. Kirchdoerfer, H.L. Turner, et al., Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen, Proc. Natl. Acad. Sci. U. S. A. 114 (2017) E7348–E7357.

[53] F. Amanat, F. Krammer, SARS-CoV-2 vaccines: status report, Immunity 52 (2020) 583–589.

[54] F. Wang, R.M. Kream, G.B. Stefano, An evidence based perspective on mRNA-SARS-CoV-2 vaccine development, Med. Sci. Monit. 26 (2020) e924700.

[55] R. Mukherjee, Global efforts on vaccines for COVID-19: since, sooner or later, we all will catch the coronavirus, J. Biosci. 45 (2020) 68.

[56] J. Robson, Computers and viral diseases. Preliminary bioinformatics studies on the design of a synthetic vaccine and a preventative peptidomimetic antagonist against the SARS-CoV-2 (2019-nCoV, COVID-19) coronavirus, Comput. Biol. Med. 119 (2020) 103670.

[57] C.L. Hsieh, J.A. Goldsmith, J.M. Schaub, A.M. DiVenere, H.C. Kuo, K. Javanmardi, et al., Structure-based Design of Prefusion-Stabilized SARS-CoV-2 Spikes, bioRxiv 2020 May 30 (2020), https://doi.org/10.1101/2020.05.30.125484 05.30.125484.

[58] J. Yu, L.H. Tostanoski, L. Peter, N.B. Mercado, K. McMahan, S.H. Mahrokhian, et al., DNA vaccine protection against SARS-CoV-2 in rhesus macaques, Science (2020), https://doi.org/10.1126/science.abc6284 May 20.

[59] T.R.F. Smith, A. Patel, S. Ramos, D. Elwood, X. Zhu, J. Yan, et al., Immunogenicity of a DNA vaccine candidate for COVID-19, Nat. Commun. 11 (2020) 2601.

[60] N. van Doremalen, T. Lambe, A. Spencer, S. Belij-Rammerstorfer, J.N. Purushotham, J.R. Port, et al., ChAdOx1 nCoV-19 Vaccination Prevents SARS-CoV-2 Pneumonia in Rhesus Macaques, bioRxiv. 2020 May 13 (2020), https://doi. org/10.1101/2020.05.13.093195 2020.05.13.093195.

[61] F.C. Zhu, Y.H. Li, X.H. Guan, L.H. Hou, W.J. Wang, J.X. Li, et al., Safety, tolerability, and immunogenicity of a recombinant adenovirus type-5 vectored COVID-19 vac-cine: a dose-escalation, open-label, non-randomised, first-in-human trial, Lancet (2020), https://doi.org/10.1016/S0140-6736(20)31208-3 May 22.

[62] NCT04341389. Clinical Trials.gov, (2020) (Apr 10).

[63] NCT04398147. ClinicalTrials.gov, (2020) (May 21).

[64] D.H. Barouch, S.V. Kik, G.J. Weverling, R. Dilan, S.L. King, L.F. Maxfield, et al., International seroepidemiology of adenovirus serotypes 5, 26, 35, and 48 in pe-diatric and adult poulations, Vaccine 29 (2011) 5203–5209.

[65] F.C. Zhu, A.H. Wurie, L.H. Hou, Q. Liang, Y.H. Li, J.B.W. Russell, et al., Safety and immunogenicity of a recombinant adenovirus type-5 vector-based Ebola vaccine in healthy adults in Sierra Leone: a single-centre, randomised, double-blind, placebo-controlled, phase 2 trial, Lancet 389 (2017) 621–628.

[66] N. Lurie, M. Saville, R. Hatchett, J. Halton, Developing COVID-19 vaccines at pandemic speed, N. Engl. J. Med. 382 (2020) 1969–1973.

[67] T.T. Le, Z. Andreadakis, A. Kumar, R.G. Román, S. Tollefsen, M. Saville, et al., The COVID-19 vaccine development landscape, Nat. Rev. Drug Discov. 19 (2020) 305–306.

[68] B.S. Graham, Rapid COVID-19 vaccine development, Science 368 (2020) 945–946.

[69] W.H. Chen, U. Strych, P.J. Hotez, M.E. Bottazzi, The SARS-CoV-2 vaccine pipeline: an overview, Curr. Trop. Med. Rep. (2020), https://doi.org/10.1007/s40475-020-00201-6.

[70] NCT04283461. ClinicalTrials.gov, (2020) (Feb 25).

[71] K.S. Corbett, D. Edwards, S.R. Leist, O.M. Abiona, S. Boyoglu-Barnum, R.A. Gillespie, et al., SARS-CoV-2 mRNA Vaccine Development Enabled by Prototype Pathogen Preparedness, bioRxiv Jun 11 (2020), https://doi.org/10. 1101/2020.06.11.145920 06.11.145920.

[72] NCT04368728. ClinicalTrials.gov, (2020) (Apr 30).

[73] NCT04405908. ClinicalTrials.gov, (2020) (May 28).

[74] R. Lu, X. Zhao, J. Li, P. Niu, B. Wang, H. Wu, et al., Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding, Lancet 395 (2020) 565–574.

[75] https://coronavirus.jhu.edu/map.html, Johns Hopkins University, USA, (2020).

[76] D.L. McKee, A. Sternberg, U. Stange, S. Laufer, C. Naujokat, Candidate drugs against SARS-CoV-2 and COVID-19, Pharmacol. Res. 157 (2020) 104859.

[77] A. Sternberg, D.L. McKee, C. Naujokat, Novel drugs targeting the SARS-CoV-2/ COVID-19 machinery, Curr.Top. Med. Chem. May 16 (2020), https://doi.org/10. 2174/1568026620999200517043137.

[78] A.B. Papaneri, R.F. Johnson, J. Wada, L. Bollinger, P.B. Jahrling, J.H. Kuhn, Middle east respiratory syndrome: obstacles and prospects for vaccine development, Expert Rev. Vaccines 14 (2015) 949–962.

[79] J.A. Kaiser, A.D.T. Barrett, Twenty years of progress toward West Nile virus vaccine development, Viruses 11 (2019) 823.

[80] R.L. Graham, E.F. Donaldson, R.S. Baric, A decade after SARS: strategies for con-trolling emerging coronaviruses, Nat. Rev. Microbiol. 11 (2013) 836–848.

[81] D. Weiskopf, K.S. Schmitz, M.P. Raadsen, A. Grifoni, N.M.A. Okba, H. Endeman, et al., Phenotype and kinetics of SARS-CoV-2-specific T cells in COVID-19 patients with acute respiratory distress syndrome, Sci. Immunol 5 (48) (2020) eabd2071, https://doi.org/10.1126/sciimmunol.abd2071 Jun 26.

[82] L. Ni, F. Ye, M.L. Cheng, Y. Feng, Y.Q. Deng, H. Zhao, et al., Detection of SARS-CoV-2-specific humoral and cellular immunity in COVID-19 convalescent individuals, Immunity 52 (e3) (2020) 971–977, https://doi.org/10.1016/j.immuni.2020.04.023 Jun 16.