The scientific research progress of novel coronavirus—— SARS-COV-2

LiuQing Yang[1]

[1]China Agricultural University

Abstract. A COVID-19 outbreak suddenly appeared in Wuhan, China, in December 2019, and then spread around the world quickly. So far, there have been a series of studies on SARS-COV-2 which has been confirmed as the cause of the outbreak. On account of the characteristic of spreading in droplet, SARS-COV-2 could be transmitted from person to person, causing the epidemic to become more and more severe all over the world. For SARS-COV-2, the spike S protein is essential for successfully infecting cells. In fact, most developmental strategies of vaccines are based on the structure of S proteins as well as host cell receptors. There are also vaccines based on the role of RNA molecules of SARS-COV-2 in host cells or the immune response of human body against the virus. This paper summarizes some research results of scholars on SARS-COV-2, aiming to provide people with a clear idea to understand SARS-COV-2, and hoping to make some contributions to the fight against the virus.

1 Background

In December 2019, a disease (COVID-19) broke out in Wuhan China suddenly. The novel virus that causes this disease is a coronavirus, a variant of the SARS virus called 2019-nCOV or SARS-COV-2. Since 1960, six coronaviruses have been found to cause diseases in humans. SARS-COV-2 is the seventh one after SARS-COV and MERS-COV [1]. SARS-COV-2 is most likely to come from wild animals such as bats and minks. Civet cats are likely intermediate hosts for the virus. SARS-COV-2 shows some particular pathogenic, epidemiological and clinical features, which has not been completely understood to date. The virus also has wider and higher transmission in the community versus SARS and MERS, as well as milder infection and lower mortality caused by the two others viruses [2]. The virus is highly infectious and can be transmitted directly in droplets produced when an infected person talks or coughs, or it is able to form aerosols and spread as aerosols. In addition, the virus can stay in the air for several hours, therefore it is easily deposit on objects, touching mouth, nose as well as other mucous membranes after being touched by a person's hand, and this process is called contact propagation. Because of the human-to-human nature of the virus, the disease then quickly spread around the world. As of 11 October, 2020, the number of infected people worldwide had reached 37 million, with 1 million deaths, according to the World Health Organization. The case fatality rate (CFR) also varies from 2.3% to 14.8% depending on the demographics of the nation or region, age, severity of the disease, and comorbidities [3]. On the whole, the CFR was reported to be around 11% worldwide among SARS-COV-2-positive cases, and the ratio of serious as well as critical cases is around 2% [3].

Similar to those infected by pathogenic severe acute respiratory syndrome coronavirus (SARS-COV) in 2003 and Middle East respiratory syndrome coronavirus (MERS COV) in 2012, people infected with SARS-COV-2 show a range of symptoms, including dry cough, fever, headache and pneumonia [4]. In addition, the innate and adaptive immune systems of humans infected with the virus respond, causing “cytokine storms” in which large amounts of cytokines and chemokines are released that damage tissue cells. For SARS-COV-2, a corresponding detection method has been developed which is essential to stop the virus from spreading continuously. A common method is fluorescent Polymerase Chain Reaction (PCR), which amplifies specific DNA fragments in vitro by PCR, and then performs gene comparison, so as to achieve the purpose of detection. What’s more, since the outbreak of the virus, many papers have been published on its structural characteristics, pathogenic mechanism and antibody research. Also, researches on vaccines have made great progresses. There are three main aspects to the vaccines. The first is that the vaccine works on a key protein or receptor of the virus to inhibit its entry into host cells. The second is to prevent the transcription and replication of viral nucleic acids in host cells. The third is to protect the tissues from the damage causing by the release of cytokines by the immune system.

The situation of COVID-19 is still severe, and the number of infected people continues to rise all around the world so far. Research into the virus is also still under way which need more exploration. In this paper, I will mainly review the structure, infection mechanism, detection method and vaccines of the virus SARS-COV-2, and

2018319010414@cau.edu.cn

provide the scientific defensive measures to contribute to win the battle against the COVID-19.

2 The structure of the SARS-COV-2 and the host cell receptor

According to phylogenetic tree, the coronavirus are divided into α、β、γ and δ four genera. Both SARS-COV and SARS-COV-2 belong to the genus β and have a high degree of homology in structure. SARS-COV-2 is one of positive sense RNA viruses and has four proteins, including M (membrane protein)、E (envelope protein)、N (nucleocapsid protein) and S (spike protein) (Fig.1). The M protein is a type of glycoprotein that spans the envelope of a virus and is involved in its assembly. The E protein is involved in the formation of the viral envelope and can be bound to the M protein to form a complex taking shape a small viral assembly mechanism which plays an important role in the morphogenesis as well as assembly of viruses. The N protein recognizes RNA molecules from the viral genome and is bound to it to form a nucleocapsid. The spike S protein plays a crucial role in completing the infection process by being bound to the host cell's receptor. According to previous studies, spike S protein is a homologous trimeric glycoprotein containing S1 subunit and S2 subunit. A total of three S proteins are connected by non-covalent bonds to form projections that protrude from the surface of the virion. The S1 subunit has two important parts, the N-terminal domain (NTD) and the receptor-binding domain (RBD) (Fig.2.A). The structure of RBD is small (~21kDa), showing asymmetric conformation after 3D reconstruction and classification [5]. Besides, there are two conformations in RBD, namely the up conformation and the low conformation (Fig.2.B). In this sense, the up conformation means that the S protein is in a state where it can be bound to the receptor, while the lower conformation means the opposite. Despite the high homology of SARS-COV and SARS-CoV-2, there are still some differences between them, and one of the biggest (though still relatively small) differences lies in the position of RBDs in their respective downward conformation [5]. In this conformation, SARS-COV tends to be closer to NTD while SARS-COV-2 seems to be closer to the central lumen of the S1 trimer. The SARS-COV-2 RBD has a twisted four-stranded antiparallel β sheet (β1, β2, β3 and β6) with short connecting helices and loops forming as the core [4]. There are nine cysteine residues in RBD, and six of them can form three disulfide bonds to stabilize the structure of RBD as well as connect the distal regions. In the S2 subunit, there are protease cutting sites, fusion peptides, central screw, connector fields and other structures (Fig.2.A). Compared with the asymmetric structure of S1 subunit, S2 subunit presents a symmetrical trimer structure, which is involved in virus and host cell membrane fusion.

For the receptor, one of the most common receptors for SARS-COV-2 to invade host cells is angiotensin converting enzyme 2 (ACE2). ACE2 is a very important titanium enzyme in the renin-angiotensin system which can maintain the balance of the system. Besides, ACE2 can also regulate blood pressure, maintain fluid balance, and participate in some vital activities such as cell proliferation. The catalytic regions of ACE2 and ACE have 42% amino acid sequence homology, and both can express in endothelial cells, whereas ACE2 has higher organ specificity. ACE2 mainly exists in mucosal cells, therefore it is commonly found in the cells of the oral cavity, nasal cavity, eyes, stomach and other organs. On account of the distribution of the ACE2, once the virus infects these cells, it will continue to spread in the form of respiratory sol, droplets and so on. In fact, ACE2 is difficult to obtain stably in vitro because it is a membrane protein. However, it has been found that ACE2 can be bound to an amino acid transporter B0AT1 in the intestinal tract to form complexes (ACE2-B0AT1) to stabilize ACE2, which contributes to analyze the structure of ACE2. With this being the case, researchers can find that ACE2 is a homologous dimer with 805 amino acids which is composed of N-terminal peptidase domain (PD) and C-terminal collectrin-like domain (CLD) (Fig.3). Such a dimerized ACE2 can contain two S protein trimers and each of them passes through one ACE2 monomer [6]. There is a metal titanium enzyme structure region in outer domain of ACE2 which has catalytic activity. The CLD has a transmembrane (TM) helix at the end and an intracellular segment with about 40 residues. The N-terminal as well as catalytic sites of ACE2 are extracellular oriented to metabolize circulating peptides, and the N-terminal doesn't have catalytic activity. The PD region of ACE2 can be cleaved to Ang I to produce Ang-(1-9) and further processed by other enzymes [7]. Moreover, the structural information of ACE2 is actually only exhibited in the PD region [6]. There is a cervical region between the PD and TM helices, and the ferrin like folding region is in this cervical region [6]. Such a structure is conducive to the rotation of the PD region, which can achieve a change in conformation while the rest of the complex remains unchanged.

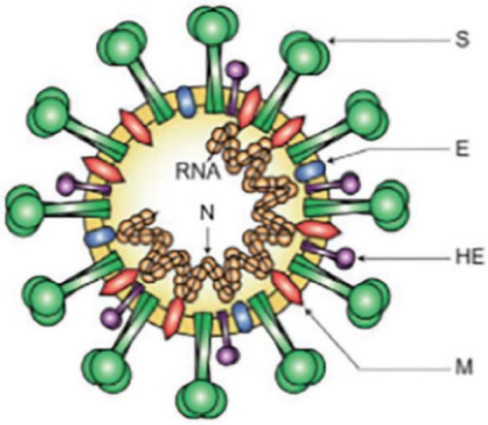

Fig.1 Structure of the coronavirus [8].

The locations of S、M、N、E proteins are indicated in the figure. The S protein protrudes from the capsule in a trismeric shape with a protruding tip. The M protein straddles the viral envelope. The N protein binds to the genomic RNA in the cytoplasm to form a helically coiled nucleocapsid. The E protein is a small envelope protein.

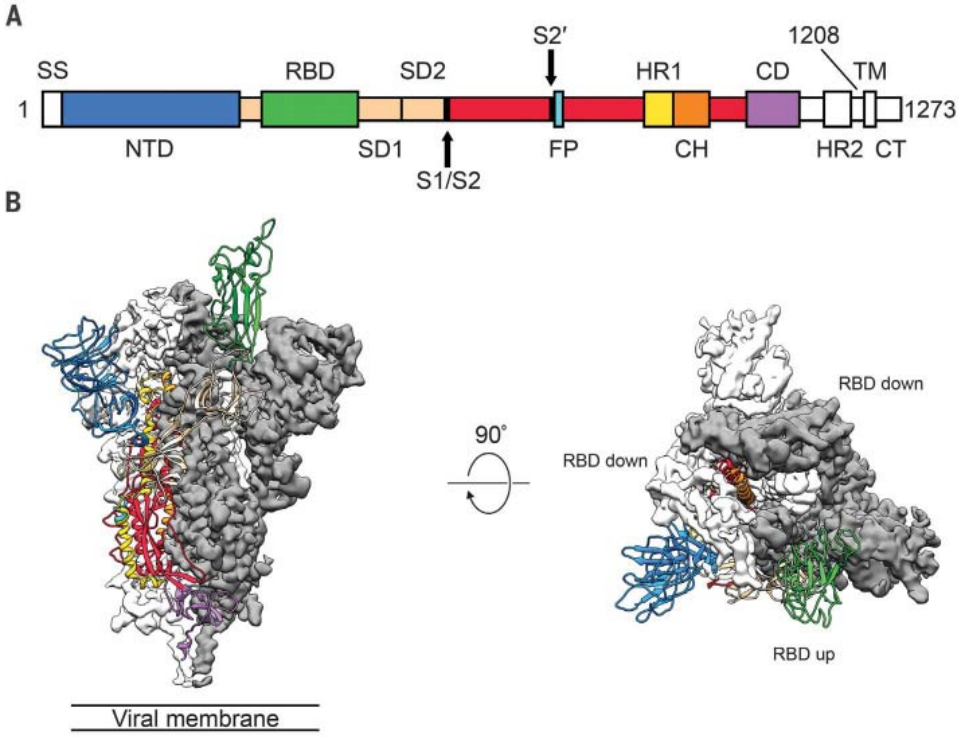

Fig. 2 Structure of 2019-nCoV S in the prefusion conformation [5].

(A) NTD, N-terminal domain. RBD, receptor-binding domain. RBM, receptor-binding motif. SD1, subdomain 1. SD2, subdomain 2. FP, fusion peptide. HR1, heptad repeat 1. HR2, heptad repeat 2. TM, transmembrane region. IC, intracellular domain [5].

(B) This figure shows the side and top views of SARS-COV-2 S protein. It contains one down conformation of RBD and two up conformations of RBD.

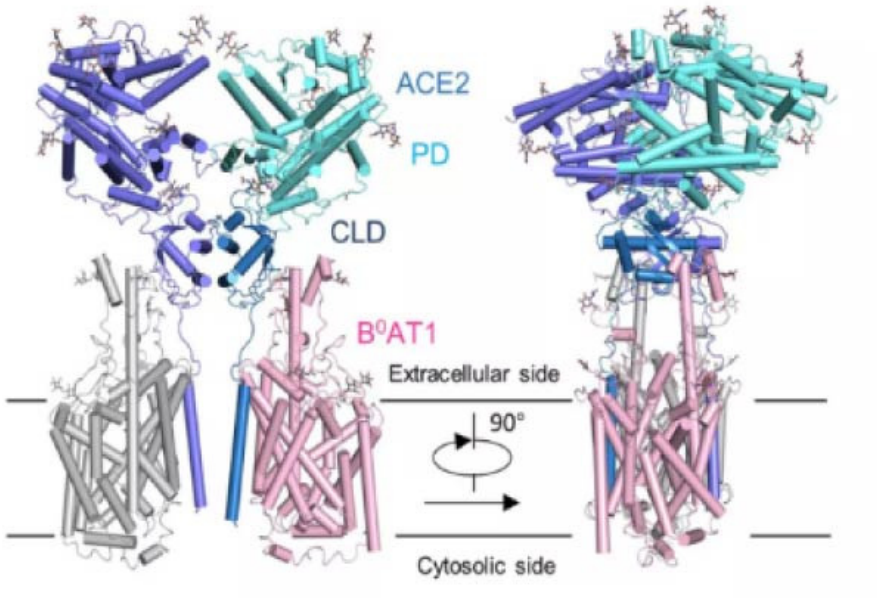

Fig. 3 Cartoon representation of the atomic model of the ACE2-B0 AT1 complex [6].

The sticks are the glycosylation part. Both PD and CLD are part of the structure of ACE2. The PD appears cyan while the CLD is drawn in blue. The structure of the composite is shown from two angles 90 degrees apart from the front and the side.

According to some studies, in addition to ACE2, CD147 on host cells is also likely to be bound to viral spike S protein and causes viral invasion. CD147, also widely known as Basigin or EMMPRIN, is a highly glycosylated transmembrane glycoprotein in the immunoglobulin superfamily, with a molecular weight of about 50kDA-60kDA[9]. Approximately 115 nucleotides are noncoding regions prior to the initiation of the N-terminal, and 39 amino acids at the C-terminal are intracellular domains. The coding region encodes 269 amino acids, 21 amino acids as signal peptides, while the extracellular domain in the middle is composed of 185 amino acids, and 206-229 has 24 amino acids as transmembrane regions [10]. Studies have investigated the possible role of CD147 in the latter invading host cells, and experiments have shown that blocking CD147 on host cells has an inhibitory effect on SARS-COV-2, suggesting that CD147 is able to promote the introduction of the virus [9].

3 The infection and pathogenesis of SARS-COV-2

As previously described, SARS infects cells by binding S protein to the host's ACE2 receptor. Such a process requires the S protein to undergo structural rearrangement to fuse to the host cell membrane. In order to bind to the host cell receptor, the S protein S1 subunit performs articulated conformation movement, leading to conformational changes in the RBD region of S1, thereby exposing the key elements of receptor binding [5]. The receptor can then be bound to the exposed RBDs, causing RBD to form an unstable upward spiral conformation, and eventually making S1 subunit fall off and S2 subunit refold (Fig.4). Then the cleavage site on the S2 subunit is exposed and lysed by the host protease to fuse the virus with the host cell membrane. When the envelope of SARS-COV-2 fuses with the host cell membrane, the virus will release the genetic material -- single-stranded RNA -- in the host cell. These RNA molecules use the ribosomes of the host cell to translate RNA replicase in the cell, which in turn replicate more viral RNA. At the same time, the replicated viral RNA molecules use the ribosomes of the host cell to synthesize the proteins which are needed by the virus. The synthesized viral proteins bind to the viral RNA to form a large number of new viruses that can be secreted outside the cell and continue to infect other cells. In this sense, the viruses continue to multiply and spread in this way after they infect other cells. In addition, ACE2 protein is generally distributed in epithelial tissues, such as mucosa and respiratory organs, due to its organ-specific distribution. Therefore, it is easy for SARS-COV-2 to enter the lungs and respiratory tract through ACE2 protein receptors and start to replicate as well as spread. As a result, patients infected with SARS-COV-2 are prone to lung lesions and develop symptoms such as pneumonia as well as dyspnea. ACE2 is also presented in the cells of the stomach, kidneys and other organs, which means that viruses can infect gastrointestinal cells through ACE2. With this being the case, many patients also develop symptoms of diarrhoea and SARS-COV-2 virus can be detected in their faeces.

Based on clinical observations, most people infected with SARS-COV-2 have no symptoms or mild symptoms of fever, cough and fatigue[11]. Some infected people develop severe symptoms, such as pneumonia, while a small number eventually develop acute respiratory distress syndrome (ARDS) [11]. The symptoms of pneumonia and ARDS are closely related to the structure and function of lung cells. In fact, there are a lot of hydrogen bonds in the alveoli cells, and these hydrogen bonds are broken by surfactant secreted by type II cells.

This process eventually expands the area of the alveolar wall membrane, which makes gas exchange easier. However, once SARS-COV-2 invades lung cells, it will affect the surfactant secreted by the type II cells, making it difficult for the lungs to exchange gases. As the pressure in the lungs drops, fluid from other parts of the body will move into the lungs and lead to inflammation eventually. The ARDS mentioned above is actually a kind of acute hypoxic respiratory failure with extremely high morbidity and mortality. As mentioned above, once a large amount of fluid enters the lungs, it is easy to cause lung infection. And in the case of severe lung infection, ARDS will be caused. Besides, increased permeability of the pulmonary endothelium to the alveoli leads to fluid flowing out of the alveoli, which will cause noncardiogenic pulmonary edema and decreased oxygenation of the arteries, resulting in ARDS too [3].

In order to resist infection of virus, our body will produce corresponding immune response, including innate immune response and adaptive immune response. However, an overactive immune response can actually do more damage to the body. In fact, under normal circumstances, when foreign pathogens invade, the immune system will activate the immune stress mechanism, and the immune cells will clear these harmful substances. At the same time, pro-inflammatory cytokines are released, allowing more immune cells to participate in the clearance process and suppress the infection. However, once the overactive immune system is activated, it will produce excessive cytokines, increasing inflammation and forming a cytokine. In cytokine storm, the cytokine interleukin-6 levels will be significantly increased, and serum IL-6 driver protein expression will be increased, such as C-reactive protein [12]. As a result, the immune system will spin out of control, rendering pathogens indistinguishable from healthy cells, and lead to multiple organ damage [12].

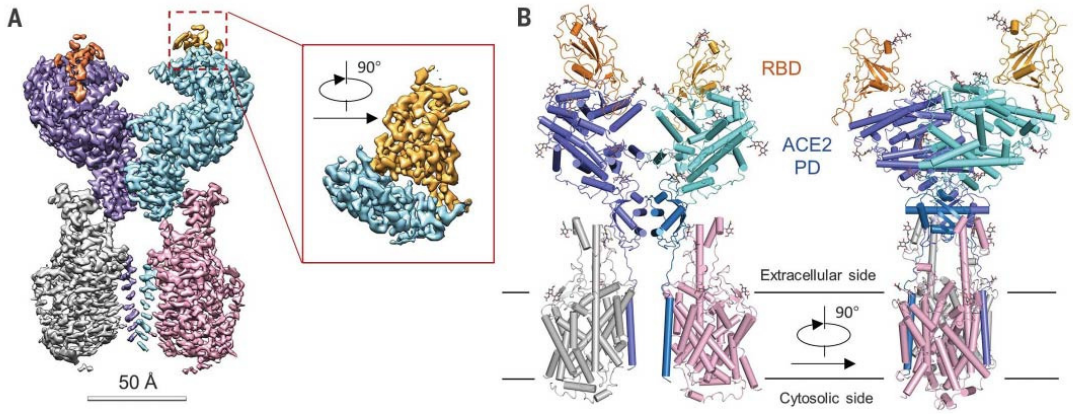

Fig.4 Overall structure of the RBD-ACE2-B0AT1 complex [6].

(A) The structure of the composite shown by Cryo-EM. The vignette on the right enlarge part of the RBD structure. Cyan represents PD in ACE2 while blue represents CLD in ACE2. The RBD protomers are shown in red and gold.

(B) Compared with Fig.3, this figure appears the structure of RBD. The structure of the complex is shown from two angles. The ACE2 and RBD regions are located laterally. The amino acid transporter B0AT1 is on the cell membrane.

4 Some vaccines and antibodies against SARS-COV-2

As of September 2020, no vaccine has been developed for widespread use. The reality is that more than 200 vaccines against SARS-COV-2 are being developed worldwide, and more than 20 vaccines are being evaluated in clinical trials. In fact, several of the vaccines most likely to be used are fall into the two technical routes of adenovirus vaccine and inactivated vaccine. Adenovirus vaccines bind the gene of S protein of SARS-COV-2 to modify replica-deficient adenovirus, after which the binding is introduced into the body to create an immune memory to prevent the infection by the virus. An inactivated vaccine is able to physically or chemically inactivate a virus and trigger an immune response in humans. A recent study has reported an early candidate for an inactivated vaccine called PiCoVacc. The researchers have obtained the SARS-COV-2 strain from a number of patients in different areas and inactivated the strain. Some animals were given different doses of vaccine and then infected with SARS-COV-2 [13]. It's inspiring that the infection was suppressed in the animals that received the small amount of vaccine, while the animals that received the large amount of vaccine could not detect the virus after a few days [13]. And such vaccine has been already in the next phase of clinical trials.

In addition to the two advanced vaccines mentioned above, vaccines and antibodies based on other principles are also being developed. As mentioned earlier, when SARS-COV-2 invades host cells, the S protein has to undergo conformational changes first, which means the S1 subunit will change the conformation, while the S2 subunit will be split by cell protease. With this being the case, if the action of cell protease is blocked, it is possible to prevent the invasion of SARS-COV-2. It has been

reported that the cell serine protease TMPRSS2 is essential for SARS-COV-2 to invade host cells because the TMPRSS2 can initiate the division of the S protein [14]. The antibody designed based on this process can produce a neutralizing antibody response against SARS-COV-2 as well as reduce the infection of SARS-COV-2, which may be of great significance for disease control [14]. What's more, the fusion of the virus envelope and the cell membrane is also crucial for the process of virus invasion. For SARS-COV-2, the S2 subunit of S protein has two important domains, including heptad repeat 1 (HR1) and heptad repeat 2 (HR2). These two domains interact to form a six-helical bundle (6-HB) fusion nucleus that allows the virus to fuse with the cell membrane and thus promotes infection. According to this process, it is possible for virus fusion inhibitors to be reasonably designed. In fact, there are eight different residues in the HR1 region. The presence of these residues can enhance the infection effect of the virus, as they are able to promote the interaction between HR1 and HR2 as well as stabilize the structure of 6-HB. However, peptide EK1 can block the formation of virus 6-HB and prevent the infection of SARS-COV-2 [15]. By coupling cholesterol molecule with EK1 peptide, it was found that EK1C4 of one lipid peptide showed strong inhibitory activity on the membrane fusion of SARS-COV-2 and the host cells [15]. Moreover, EK1C4 is highly effective against some live HCOV infection both in vitro and in vivo, which means the EK1C4 might be further developed into a therapeutic and prophylactic drug against coronavirus fusion [15].

Besides the mode of binding to ACE2 receptors, the S protein of SARS-COV-2 can also be bind to CD147 receptors on host cells to achieve invasion. Studies have shown that if CD147 is blocked, the replication of SARS-COV-2 will be inhibited [9]. Based on this principle, Meplazumab, an anti-CD147 antibody, was developed to effectively inhibit the virus from invading host cells [9]. Besides, two specific human monoclonal antibodies, CA1 and CB6, have been isolated from SARS-COV-2 patients in convalescence. Both of these antibodies have strong specific neutralizing activity against SARS-COV-2 in vitro. Further studies have shown that the CB6 can recognize the overlapping epitopes of ACE2 binding sites in the RBD region of S-protein of SARS-COV-2. After the identification, the CB6 can also inhibit virus infection through the interaction of steric hindrance and virus receptor interference [16]. According to previous research, for pulmonary diseases, anti-ACE2 antibodies may actually activate the renin-angiotensin system and lead to the occurrence of severe lung injury or some other negative effects [9]. In contrast, CD147 is highly expressed in some tumors and inflammatory tissues, which means anti-CD147 antibodies have less effect on normal cells [17].

5 Conclusion

The sudden spread of SARS-COV-2 caught people off guard. After unremitting research by scientists, people finally have understood that SARS-COV-2 is actually a novel coronavirus with highly contagious and can spread from person to person. For detection methods, the researchers used PCR to amplify a specific piece of DNA in the organism and then performed a genetic comparison to detect SRAS-COV-2 infection. People infected with SARS-COV-2 suffer from cough, fever, headache, general fatigue or even pneumonia as well as acute respiratory distress [4]. To resist the invasion of the virus, the immune system of the human body will release excessive cytokines, which actually aggravates the damage of human's organs.

In a sense, understanding the molecular structure and infection mechanism of SARS-COV-2 is essential in the fight against it. According to researches, SARS-COV-2 is one of positive sense RNA viruses and has four proteins. In fact, the spike S protein with S1 and S2 subunits is the key protein for successful infection (Fig.1). For receptors, ACE2 which is primarily present in mucosal cells, is the most common. Besides, CD147 is also likely to be an important receptor for SARS-COV-2 to invade host cells. After successful infection, SARS-COV-2 releases its own RNA molecules in the host cell. These RNA molecules will use the host ribosomes in the cell to synthesize RNA replicases as well as virus shells, and reproduce a large number of new RNA molecules. In the next place, large numbers of new viral molecules will be synthesized and released outside the cell to infect other cells. Also, vaccines and antibodies against SARS-COV-2 are also being studied, and the vaccines are developed from different perspectives, such as inactivated vaccines, adenovirus vaccines, protease inhibitors, monoclonal antibodies and so on. Although no vaccine has been successfully developed so far, many vaccines have entered phase III clinical trials, and it is believed that they will be successfully developed in the near future.

There is no denying that personal protection is critical to prevent the infection of SARS-COV-2. According to the current researches, SARS-COV-2 most likely came from bats, minks and other wild animals. It is well known that wild animals also carry a variety of unknown bacteria and viruses which means avoiding eating game is important to protect people from strange symptoms. What's more, from the perspective of controlling the source of infection, it is necessary to report those who have symptoms of infection in time to avoid the continuous spread of the virus. In terms of cutting off the transmission route, it is significant to wear a mask to stop the transmission due to the characteristics of droplet transmission of SARS-COV-2. In addition, frequent hand washing, appropriate alcohol disinfection can also play a good protective role.

References

1. Zhu, N., et al., *A Novel Coronavirus from Patients with Pneumonia in China, 2019*. N Engl J Med, 2020. **382**(8): p. 727-733.
2. Vincent J. Munster Ph. D. , M.K., D. V. M. , Neeltje van Doremalen, Ph.D., Debby van Riel, Ph. D., and Emmie de Wit ,Ph.D., *A Novel coronavirus Emerging in China——Key Questions for empact assessment.*
3. Can, A. and H. Coskun, *The rationale of using mesenchymal stem cells in patients with COVID-19-*

related acute respiratory distress syndrome: What to expect. Stem Cells Transl Med, 2020.

4. Lan, J., et al., *Crystal structure of the 2019-nCoV spike receptor-binding domain bound with the ACE2 receptor.* 2020.
5. Wrapp, D., et al., <*Cryo-EM Structure of the 2019-nCoV Spike in the Prefusion Conformation.pdf*>. Science, 2020. **367**(6483).
6. Yan, R., et al., <*Structural basis for the recognition of the SARS-CoV-2 by full-length human ACE2.pdf*>. Science, 2020. **367**(6485).
7. Hamming, I., et al., *The emerging role of ACE2 in physiology and disease.* J Pathol, 2007. **212**(1): p. 1-11.
8. Tok, T.T. and G. Tatar, *Structures and Functions of Coronavirus Proteins: Molecular Modeling of Viral Nucleoprotein* SCIRES Literature. **2**(1).
9. Wang, K., W. Chen, and Y.-S. Zhou, <*SARS-CoV-2 invades host cells via a novel route CD147-spike protein.pdf*>.
10. Curtin, K.D., I.A. Meinertzhagen, and R.J. Wyman, *Basigin (EMMPRIN/CD147) interacts with integrin to affect cellular architecture.* J Cell Sci, 2005. **118**(Pt 12): p. 2649-60.
11. Catanzaro, M., et al., *Immune response in COVID-19: addressing a pharmacological challenge by targeting pathways triggered by SARS-CoV-2.* Signal Transduction and Targeted Therapy, 2020. **5**(1).
12. Moore, J.B. and C.H. June, <*Cytokine release syndrome in severe COVID-19.pdf*>. Science, 2020. **368**(6490).
13. Gao, Q., et al., <*Development of an inactivated vaccine candidate for SARS-CoV-2.pdf*>. 2020. **369**(6499).
14. Hoffmann, M., et al., *SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor.* Cell, 2020. **181**(2): p. 271-280 e8.
15. Xia, 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**(4): p. 343-355.
16. Shi, R., et al., *A human neutralizing antibody targets the receptor-binding site of SARS-CoV-2.* Nature, 2020. **584**(7819): p. 120-124.
17. Kosugi, T., et al., *CD147 (EMMPRIN/Basigin) in kidney diseases: from an inflammation and immune system viewpoint.* Nephrology Dialysis Transplantation, 2015. **30**(7): p. 1097-1103.