Novel Coronavirus (COVID-19) outbreak: A review on most targetable proteins (S and N) of SARS-CoV-2

Abstract

In this review specific aspect of CoV-host interactions have been discussed. Newly identified COVID-19 virus belongs to β-Coronavirus , where genome sequencing of Receptor Binding Domain (RBD) of SARS-CoV-2 shows a close relation with bats viruses. Therefore, similar to MERS and SARS it is also having an animal origin (like bats or civet cats). In contrast to SARS it is highly contagious and rate of transmission is high in person to person. In current situation, there is an urgent need of effective vaccine to cure against life threatens infection. This review focuses to understand the most likely targetable proteins i.e. Spike (S) and Nucleocapsid (N). Due to emergence of Coronavirus outbreak globally, present research based on drug-genes signatures available in the literature. Corona viridae is a family of viruses mainly caused upper respiratory infections from common colds to more severe illnesses such as Severe Acute Respiratory Syndrome (SARS). The S-glycoprotein found on the surface of Coronavirus and one of the important enzyme that gain attention due to involvement in viral attachment and entry into host cells. Similarly Nucleocapsid (N) protein is highly immunogenic phosphor protein; it is filamentous and single-stranded viral RNA. It plays role in genome replication and involve in cell signaling pathways. This review is to be addressed the main challenge i.e. pathogenesis of SARS-CoV-2 by understanding both S protein and N protein.

Keywords: SARS-CoV-2 ▪ Spike (S) protein ▪ Nucleocapsid (N) protein ▪ Receptor Binding Domain (RBD) ▪ Coronavirus ▪ Severe Acute Respiratory Syndrome (SARS)

Introduction

Coronaviruses (Latin: corona=crown), named for the crown-like spikes (S glycoprotein) on their surface [1]. It belongs to Coronaviridae family, single strand enveloped viruses and included in the Nidovirales order [2]. It is sub divided into four groups-alpha, beta, gamma, and delta on the basis of genomic structure. Coronavirus not only infect mammals also has broad host range including avian. They may cause mainly upper respiratory tract infection and also cause gastrointestinal, hepatic disease and central nervous system [3]. All corona viruses like SARS-CoV and MERS-CoV have had animal origin generally either bats or rodents and caused diseases to humans [4]. So, genome sequencing of new Coronavirus showed 96.2% sequence identity with Bat CoV RaTG13 [5]. Till December 2019, only six different Coronavirus es were known. Four of these (HCoV-NL63, HCoV-229E, HCoV-OC43 and HKU1) usually caused mild common cold-type symptoms in immune-competent people and the other two (SARS-CoV and MERS-CoV) have caused pandemics in the past two decades. Genome of 2019-nCov was sequenced, it shared 79.5% of the genetic sequence of the SARS-CoV that caused the 2002-2003 pandemic [6] and the International Committee on Taxonomy of Viruses renamed the 2019-nCov as SARS-CoV-2 [7]. Generally Coronaviruses are positive-stranded RNA viruses with a 27-kb to 31-kb genome [8]. Similarly, SARS-CoV and MERS-CoV have positive-sense RNA genomes of 27.9 kb and 30.1 kb, respectively [9]. About two-thirds of the Coronavirus genome (~20,000 bases) encodes the viral replicase that involve in viral RNA synthesis. The replicase gene is comprised of two large open reading frames, designated ORF1a and ORF1b [10], it also encodes 3' structural proteins includes large glycoprotein Spike (S), Envelope (E), a small glycoprotein Membrane (M) mainly embedded in the membrane and phosphorylated Nucleocapsid (N). These structural proteins have genes that evoke pathogenicity in host cells and involve in viral process and replication [11-14].

A plus-strand (+) RNA virus have genetic diversity and applies to their RNA synthesis

Sana Gul*

Department of Chemistry, Federal Urdu University of Arts, Sciences and Technology, Pakistan

**Author for correspondence: sana.gul@fuuast.edu.pk*

machinery. Only one enzyme is conserved i.e. RNA-dependent RNA polymerase (RdRp), whereas in other domains replicative and accessory protein may vary. In most of the targets, both Spike (S) and Nucleocapsid (N) protein gains attention due to fusion of virus on the host and involve in replication, respectively [15]. The S protein comprises into two components, S1 (amino-terminal include amino acids 270 to 510) contains the Receptor Binding Domain (RBD); while S2 (carboxyl-terminal) contains the fusion peptide and due to conformational changes it allows membrane fusion to host cells, allowing entry of virus [16-21]. Novel Coronavirus cause severe illness as Angiotensin-Converting Enzyme 2 (ACE2) is found in the lower respiratory tract of humans, resulting into fever, destruction of alveolies and characterized atypical pneumonia [22]. ACE2, is contributing as cell receptor for SARS-CoV-2 during entry of virus into host cells and regulates both the cross-species and human-to-human transmission [23].

Despite limited information on this new virus, previous findings support that SARS uses Angiotensin-Converting Enzyme 2 (ACE2) to gain entry in to cells [18,24], in similar fashion nCoV uses dipeptidyl peptidase 4 (DPP4 or CD26) as a functional receptor [25]. This finding may be important as the requirement for ACE2 and may be responsible for the pathogenicity of SARS-CoV. Still pathogenesis of nCoV results into pandemic and the reason is still unknown. Once SARS-CoV-2 enters into host cells due to interactions of SARS-S RBD with the cell surface receptor ACE2 [26,27]. These interactions are further followed by endocytosis at low pH, resulting in the cleavage of SARS-S by a host protease called cathepsin L, thus exposing the S2 domain for membrane fusion, the two subunits arrange and fold into a metastable pre-fusion conformation [28-33]. SARS-S also regulates cell stress responses and apoptosis [34].

Thus, it is important to understand the mechanisms of entry of Coronavirus and further followed by fusion to the membrane to eliminate pathogenesis and can prevent Coronavirus infection at earliest [35]. It is critical to understand S protein fusion because viruses exhibit tropism for specific cells *in vivo* [36]. It is important to evaluate the role of host proteins in relevant primary cell types when experimentally feasible. Evidence suggests that tissue expression of the ACE2 receptor corresponds to the localization of virus during infection in infected individuals [37-40]. Also, the efficiency of infection in humans also correlates with the ability of the ACE2 to support viral replication [37,41-43].

Further studies demonstrated that when the S protein binds to the ACE2 receptor, followed by type 2 Transmembrane Protease TMPRSS2 leading to cleavage of ACE2 and activation of the spike protein [44,45]. Similarly influenza virus use same mechanism to facilitate viral entry into the target cell. It has been suggested that cells in which ACE2 and TMPRSS2 are simultaneously present are most susceptible to entry by SARS-CoV [46]. Early indications are that SARS-CoV-2 virus also requires ACE2 and TMPRSS2 to enter cells [6]. Viral entry triggers the host's immune response, and the inflammatory cascade is initiated by Antigen-Presenting Cells (APC). The APC performing two functions: (1) presenting the foreign antigen to CD4+-T-helper (Th1) cells, and (2) releasing interleukin-12 to further stimulate the Th1 cell. The Th1 cells stimulate CD8+-T-killer (Tk) cells that will target any cells containing the foreign antigen. In addition, activated Th1 cells stimulate B-cells to produce antigen-specific antibodies [47]. Immuno-compromised patient fails to combat infection via immune response and their lung samples showed alveolar damage resulting into fatality [17].

Another defense mechanism is Autophagy. It is a cellular stress response that functions to recycle proteins and organelles [48,49]. It is studied in Sindbis virus and herpes simplex virus-1 and showed an important defense mechanism against infection with those viruses [50,51]. It might be possible recovered patient may have autophagy defense mechanism.

Next, most targetable protein known as nucleocapsid N protein; works as an antigen and binds with viral RNA genome. It involve in both RNA binding and replication by modulating transcription process. It also contains two non-interacting structural domain, one is N-terminal domain and serve for RNA binding, where second is C-terminal domain exits in dimer form in solution and considered as dimerization domain [52]. Initially ribonucleoprotein core is form by binding of N protein with viral RNA genome, on the onset of infection this ribonucleoprotein core enters into host cell and interacts with host proteins [53-57]. Coronavirus

es replicate entirely in the cytoplasm of cells. In newly recognized human Coronavirus , N protein identified as the causative agent of Severe Acute Respiratory Syndrome (SARS) [35] and found most abundantly protein in infected cells. It contains RNA-binding motif and involve in cellular signaling. It is phosphoprotein involve in formation of viral core, packaging and transcription of viral RNA and have high affinity with viral RNA and form ribonucleocapsid structure [58]. In infections of SARS- CoV, N protein has been involved in the induction of Mitogen-Activated Kinases (MAPKs), especially p38MAPK [59,60]. Additionally both, AP-1 signal pathway is activated in response of stress, inflammation or viral infection and weak induction of Akt signaling pathways were also found [22,61-63]. AP-1 involve in cellular transduction and regulate cellular process, due to stimuli it activates innate or adaptive immunity. Due to deregulation of AP-1 pathway it results into numerous lymphomas.

In Coronavirus infection, viral mRNA parasitize host machinery and affects host transcription and translation processes, due to similar structure of viral mRNAs to eukaryotic hosts, they capture host machinery to translate the viral mRNA [22]. In vrius life cycle, Nucleocapsid protein not only responsible for mRAN generation, M and E protein together intimately participate in genome condensation and packaging [64-66].

Conclusion

In order to understand virus pathogenesis, there are three strategies to employ either to target Receptor-Binding Domain (RBD) of S protein, second to target ACE2 receptor by antibodies to block S protein binding and prevent virus entry, and last to stop viral replication via targeting AP-1 pathway to stop both transcription and translation process of genomic material.

References

Rabi FA, Al Zoubi MS, Kasasbeh GA, et al. SARS-CoV-2 and Coronavirus Disease 2019: what we know so far. *Pathogens*. 9, 231 (2020).

Gorbalenya AE, Enjuanes L, Ziebuhr J, et al. Nidovi- rales: Evolving the largest RNA virus genome. *Virus Res*. 117, 17-37 (2006).

Perlman SR, Lane TE, Buchmeier MJ. Lippincott Williams and Wilkins Philadelphia. 1, 331-348 (1999).

Fan Y, Zhao K, Shi ZL, et al. Bat Coronaviruses in China. *Viruses*. 11, 210 (2019).

Zhou P, Yang XL, Wang XG, et al. A pneumonia outbreak associated with a new Coronavirus of probable bat origin. *Nature*. 579, 270-273 (2020).

Zhou P, Yang XL, Wang XG, et al. Discovery of a novel Coronavirus associated with the recent pneumonia outbreak in humans and its potential bat origin. *Microbiology*. (2020). [Available at: https://www.biorxiv.org/content/10.1101/2020.01.22.914952v2]

Gorbalenya AE, Baker SC, Baric RS, et al. Severe acute respiratory syndrome-related Coronavirus: The species and its viruses-A statement of the Coronavirus study group. *Microbiology*. 10, 937-862 (2020).

Snijder EJ, Van Der Meer Y, Zevenhoven-Dobbe J, et al. Ultrastructure and origin of membrane vesicles associated with the severe acute respiratory syndrome Coronavirus replication complex. *J Virol.* 80, 5927-5940 (2006).

De Wit E, van Doremalen N, Falzarano D, et al. SARS and MERS: Recent insights into emerging Coronavirus es. *Nat Rev Microbiol*. 14, 523-534 (2016).

Ziebuhr J. In Coronavirus replication and reverse genetics. *Springer*. 57-94 (2005).

Pyrc K, Jebbink MF, Berkhout B, et al. Genome structure and transcriptional regulation of human Coronavirus NL63. *Virol J*. 1, 7 (2004).

Casais R, Davies M, Cavanagh D, et al. Gene 5 of the avian Coronavirus infectious bronchitis virus is not essential for replication. *J Virol*. 79, 8065-8078 (2005).

de Haan CAM, Masters PS, Shen X, et al. The group-specific murine Coronavirus genes are not essential, but their deletion, by reverse genetics, is attenuating in the natural host. *Virology*. 296, 177-189 (2002).

Haijema BJ, Volders H, Rottier PJM. Live, attenuated Coronavirus vaccines through the directed deletion of group-specific genes provide protection against feline infectious peritonitis. *J Virol*. 78, 3863-3871 (2004).

Gallagher TM, Buchmeier MJ. Coronavirus spike proteins in viral entry and pathogenesis. *Virology*. 279, 371-374 (2001).

Hwa KY, Lin WM, Hou YI, et al. Peptide mimicrying between SARS Coronavirus spikes protein and human proteins react with SARS patient serum. *J Biomed Biotechnol*. 2008, 326464 (2008).

He R, Adonov A, Traykova-Adonova M, et al. Potent and selective inhibition of SARS Coronavirus replication by aurintricarboxylic acid. *Biochem Biophys Res Commun*. 320, 1199-1203 (2004).

Li W, Moore MJ, Vasilieva N, et al. Angiotensin-converting enzyme 2 is a functional receptor for the SARS Coronavirus . *Nature*. 426, 450-454 (2003).

Wong SK, Li W, Moore MJ, et al. A 193-amino acid fragment of the SARS Coronavirus S protein efficiently binds angiotensin-converting enzyme 2. *J Biol Chem*. 279, 3197-3201 (2004).

Tsai JC, Zelus BD, Holmes KV, et al. The N-terminal domain of the murine Coronavirus spike glycoprotein determines the CEACAM1 receptor specificity of the virus strain. *J Virol*. 77, 841-850 (2003).

Zelus BD, Schickli JH, Blau DM, et al. Conformational changes in the spike glycoprotein of murine Coronavirus are induced at 37ºC either by soluble murine CEACAM1 receptors or by pH 8. *J Virol*. 77, 830-840 (2003).

He R, Leeson A, Andonov A, et al. Activation of AP-1 signal transduction pathway by SARS Coronavirus nucleocapsid protein. *Biochem Biophys Res Commun*. 311, 870-876 (2003).

Wan Y, Shang J, Graham R, et al. Receptor recognition by the novel Coronavirus from Wuhan: an analysis based on decade-long structural studies of SARS Coronavirus . *J Virol*. 94 (2020).

Muller MA, Raj VS, Muth D, et al. Human Coronavirus EMC does not require the SARS-Coronavirus receptor and maintains broad replicative capability in mammalian cell lines. *M Biol*. 3, e00515-612 (2012).

Raj VS, Mou H, Smits SL, et al. Dipeptidyl peptidase 4 is a functional receptor for the emerging human Coronavirus -EMC. *Nature*. 495, 251-254 (2013).

Li F, Berardi M, Li W, et al. Conformational states of the severe acute respiratory syndrome Coronavirus spike protein ectodomain. *J Virol*. 80, 6794-6800 (2006).

Lu L, Liu Q, Zhu Y, et al. Structure-based discovery of Middle East respiratory syndrome Coronavirus fusion inhibitor. *Nature Commun*. 5, 1-12 (2014).

Matsuyama S, Taguchi F. Two-step conformational changes in a Coronavirus envelope glycoprotein mediated by receptor binding and proteolysis. *J Virol*. 83, 11133-11141 (2009).

Simmons G, Gosalia DN, Rennekamp AJ, et al. Inhibitors of cathepsin L prevents severe acute respiratory syndrome Coronavirus entry. *Proc Natl Acad Sci*. 102, 11876-11881 (2005).

Kong L, Torrents de la PeÃa A, Deller MC, et al. Complete epitopes for vaccine design derived from a crystal structure of the broadly neutralizing antibodies PGT128 and 8ANC195 in complex with an HIV-1 Env trimer. *Acta Crystallogr*. 71, 2099-2108 (2015).

Lee PS, Ohshima N, Stanfield RL, et al. Receptor mimicry by antibody F045â€“092 facilitates universal binding to the H3 subtype of influenza virus. *Nature Commun*. 5, 3614 (2014).

McLellan JS, Chen M, Joyce MG, et al. Structure-based design of a fusion glycoprotein vaccine for respiratory syncytial virus. *Science*. 342, 592-598 (2013).

Sanders RW, Derking R, Cupo A, et al. A next-generation cleaved, soluble HIV-1 Env trimer, BG505 SOSIP. 664 gp140, expresses multiple epitopes for broadly neutralizing but not non-neutralizing antibodies. *Plos Pathog*. 9, e1003618 (2013).

DeDiego ML, Nieto-Torres JL, JimÃnez-GuardeÃo JM, et al. Severe acute respiratory syndrome Coronavirus envelope protein regulates cell stress response and apoptosis. *Plos Pathog*. 7, e1002315 (2011).

Prentice E, Jerome WG, Yoshimori T, et al. Coronavirus replication complex formation utilizes components of cellular autophagy. *J Bio Chem*. 279, 10136-10141 (2004).

Zhao Z, Thackray LB, Miller BC, et al. Coronavirus replication does not require the autophagy gene ATG5. *Autophagy*. 3, 581-585 (2007).

Harmer D, Gilbert M, Borman R, et al. Quantitative mRNA expression profiling of ACE 2, a novel homologue of angiotensin converting enzyme. *FEBS Lett*. 532, 107-110 (2002).

Chan PKS, To KF, Lo AWI, et al. Persistent infection of SARS Coronavirus in colonic cells *in vitro*. *J Med Virol*. 74, 1-7 (2004).

Ding Y, He L, Zhang Q, et al. Organ distribution of Severe Acute Respiratory Syndrome (SARS) associated Coronavirus (SARSâ€-CoV) in SARS patients: Implications for pathogenesis and virus transmission pathways. *J Pathol*. 203, 622-630 (2004).

Hamming I, Timens W, Bulthuis MLC, et al. Tissue distribution of ACE2 protein, the functional receptor for SARS Coronavirus . A first step in understanding SARS pathogenesis. *J Pathol*. 203, 631-637 (2004).

Li W, Greenough TC, Moore MJ, et al. Efficient replication of severe acute respiratory syndrome Coronavirus in mouse cells is limited by murine angiotensin-converting enzyme 2. *J Virol*. 78, 11429-11433 (2004).

Subbarao K, McAuliffe J, Vogel L, et al. Prior infection and passive transfer of neutralizing antibody prevent replication of severe acute respiratory syndrome Coronavirus in the respiratory tract of mice. *J Virol*. 78, 3572-3577 (2004).

Wentworth DE, Gillim-Ross L, Espina N, et al. Mice susceptible to SARS Coronavirus . *Emerg Infect Dis*. 10, 1293 (2004).

Glowacka I, Bertram S, Maller MA, 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, 4122-4134 (2011).

Heurich A, Hofmann-Winkler H, Gierer S, et al. TMPRSS2 and ADAM17 cleave ACE2 differentially and only proteolysis by TMPRSS2 augments entry driven by the severe acute respiratory syndrome Coronavirus spike protein. *J Virol*. 88, 1293-1307 (2014).

Shulla A, Heald-Sargent T, Subramanya G, et al. A transmembrane serine protease is linked to the severe acute respiratory syndrome Coronavirus receptor and activates virus entry. *J Virol*. 85, 873-882 (2011).

Cui J, Li F, Shi ZL. Origin and evolution of pathogenic Coronavirus es. *Nat Rev Microbiol*. 17, 181-192 (2019).

Klionsky DJ, Emr SD. Autophagy as a regulated pathway of cellular degradation. *Science*. 290, 1717-1721 (2000).

Reggiori F, Klionsky DJ. Autophagy in the eukaryotic cell. *Eukaryot*. 1, 11-21 (2002).

Talloczy Z, Jiang W, Virgin HW, et al. Regulation of starvation and virus induced autophagy by the eIF2Î ± kinase signaling pathway. *Proc Natl Acad Sci*. 99, 190-195 (2002.

Liang XH, Kleeman LK, Jiang HH, et al. Protection against fatal Sindbis virus encephalitis by beclin, a novel Bcl-2-interacting protein. *J Virol*. 72, 8586-8596 (1998).

Lai MMC. SARS virus: The beginning of the unraveling of a new Coronavirus . *J Biomed Sci*. 10, 664-675 (2003).

Gorbalenya AE, Snijder EJ, Spaan WJM. Severe acute respiratory syndrome Coronavirus phylogeny: Toward consensus. *J Virol*. 78, 7863-7866 (2004).

Rota PA, Oberste MS, Monroe SS, et al. Characterization of a novel Coronavirus associated with severe acute respiratory syndrome. *Science*. 300, 1394-1399 (2003).

Marra MA, Jones SJM, Astell CR, et al. The genome sequence of the SARS-associated Coronavirus . *Science*. 300, 1399-1404 (2003).

Snijder EJ, Bredenbeek PJ, Dobbe JC, et al. Unique and conserved features of genome and proteome of SARS-Coronavirus , an early split-off from the Coronavirus group 2 lineages. *J Mol Biol*. 331, 991-1004 (2003).

Ng ML, Tan SH, See EE, et al. Early events of SARS Coronavirus infection *in vitro* cells. *J Med Virol*. 71, 323-331 (2003).

You J, Dove BK, Enjuanes L, et al. Subcellular localization of the severe acute respiratory syndrome Coronavirus nucleocapsid protein. *J Gen Viro*. 86, 3303-3310 (2005).

Banerjee S, Narayanan K, Mizutani T, et al. Murine Coronavirus replication-induced p38 mitogen-activated protein kinase activation promotes interleukin-6 production and virus replication in cultured cells. *J Virol*. 76, 5937-5948 (2002).

Mizutani T, Fukushi S, Saijo M, et al. Phosphorylation of p38 MAPK and its downstream targets in SARS Coronavirus -infected cells. *Biochem Biophys Res Commun*. 319, 1228-1234 (2004).

Mizutani T, Fukushi S, Murakami M, et al. Tyrosine dephosphorylation of STAT3 in SARS Coronavirus ae infected Vero E6 cells. *FEBS Lett*. 577, 187-192 (2004).

Mizutani T, Fukushi S, Saijo M, et al. Importance of Akt signaling pathway for apoptosis in SARS-CoV infected Vero E6 cells. *Virology*. 327, 169-174 (2004).

Surjit M, Liu B, Jameel S, et al. The SARS Coronavirus nucleocapsid protein induces actin reorganization and apoptosis in COS-1 cells in the absence of growth factors. *Biochem J*. 383, 13-18 (2004).

Baric RS, Nelson GW, Fleming JO, et al. Interactions between Coronavirus nucleocapsid protein and viral RNAs: Implications for viral transcription. *J Virol*. 62, 4280-4287 (1988).

Kuo L, Masters PS. Genetic evidence for a structural interaction between the carboxy termini of the membrane and nucleocapsid proteins of mouse hepatitis virus. *J Virol*. 76, 4987-4999 (2002).

Narayanan K, Kim KH, Makino S. Characterization of N protein self-association in Coronavirus ribonucleoprotein complexes. *Virus Res*. 98, 131-140 (2003).