

## **Virion structure and mechanism of propagation of coronaviruses including SARS-CoV 2 (COVID -19 ) and some meaningful points for drug or vaccine development**

Swapan Kumar Ghosh

Molecular Mycopathology Lab, Cancer Research Unit, Ramakrishna Mission Vivekananda Centenary College (Autonomous), Rahara, Kolkata- 700118, WB, India.

Corresponding address: gswapan582@gmail.com

### **Abstract**

SARS-CoV-2 or COVID-19, a new seventh human corona virus, has out-broken in Wuhan, China since 31<sup>st</sup> December 2019, and quickly escalated to take the form of pandemic which killed many human beings throughout almost all countries across continents. The rapidity of its transmission from human to human is far greater than all previous human corona viruses which came into existence like SARS-CoV, MERS-CoV, etc. The nucleotide sequence of SARS-CoV-2 (isolates Wuhan-Hu-1) is 29,875 bp in ss-RNA. Symptoms of SARS-CoV-2 infected pneumonia include from asymptomatic to high fever and/or respiratory illnesses. Coronavirus virion (spherical/round /elliptical in shape) consists of three parts- outer membrane or envelope, nucleocapsid and genome (RNA). SARS-CoV-2 was shown to use receptor, angiotensin converting enzyme 2 (ACE2) for attachment to the cells through its surface spike (S) protein (S1), and the virion enters into the host cell through two routes- direct membrane fusion and endocytotic pathway. The RNA of SARS-CoV acts directly as mRNA and here minus(-) 1 programmed ribosomal frameshift (-1PRF) is being operated by slippery sequence and pseudoknot, so it translates 16 nonstructural proteins including RNA dependent RNA replicase. Then genomic RNA replicated continuously on – strand RNA template and subgenomic RNA transcribed discontinuously on –RNA template to sgmRNA. Subgenomic RNAs/sgmRNAs synthesize all structural proteins. This article takes into consideration the details of established theories of viral structure, viral attachment, mode of entry into human cells, different models of replication and transcription of virus genome proposed by eminent scientists over the years, and makes an in depth examination highlighting meaningful points or important target cites of viral propagation or synthesis, which are conserved, for prompt development of potent drugs or vaccine to counter COVID-19 for which human race is anxiously and eagerly waiting.

**KEYWORDS:** SARS-CoV-2(COVID-19), + ssRNA, ribosomal – framshift, Pseudoknot, replication, model, drug, vaccine

## INTRODUCTION

Coronaviruses (CoVs) are positive-single stranded RNA viruses. An electron microscopical image shows that these viruses appear crown-like (Latin term *coronam* means crown) due to emergence of spike from the envelope, and they are important pathogens for human and other vertebrates including birds, bat, mouse, etc and they generally infect respiratory system and other non-pulmotary systems like gastrointestinal, cardiovascular, hepatic, and central nervous systems (1,2). Six human coronaviruses such as HCoV-229E, HCoVOC43 (since the 1960s), HCoV-NL63 (since 2004), HKU1(since 2005), SARS-CoV(since 2002-2003) and MARS-CoV(since 2012) were recorded before the emergence of the seventh one which is currently on a rampant mood to bring the whole world to a standstill. Only mild upper respiratory disease of us which was caused by other human corona viruses (HCoV-OC43, HCoV-229E, HCoV-NL63 and HKU1), was noted, and in few cases some of them become dangerous for elders, infants and young children. In Guangdong Province of China, in late 2002. severe acute respiratory syndrome coronavirus (SARS-CoV) out broke.and it, at the end of 2003, became spread to many countries including North America, South America, and Europe, infecting 8,096 people and taking the death toll to 774 (10% fatality rate) (3,4). After that the people of Middle East, Europe and Africa faced another serious fatal respiratory disease known as Middle East Respiratory Syndrome (MERS) coronavirus (MERS-CoV), with 55% fatality rate (5,6). More than 2428 individuals were infected and 838 people were died by MERS coronavirus(7). Theses two coronaviruses( SARS -CoV and MERS -CoV) generally attack lower respiratory tract and manifest severe respiratory syndrome in human (8,9). In the year of 2013, Jiang and co- workers (10) had alarmed the chance of a re- appearance of SARS-CoV in the form of even more dangerous pandemic, and now, what is unfolding in front of us because of the outbreak of dangerous respiratory disease in Wuhan since 31st December 2019 bears testimony to the prediction of Jiang and co-workers (10). This novel coronavirus 2019 has added another human coronavirus in the portfolio of coronaviruses that ravaged human lives and left indelible traumatic marks on human society.WHO(11) on 2 January, 2020, tentatively designated the new virus as 2019 novel coronavirus (2019 -nCoV) which has been identified molecularly by at least five independent laboratories in China (12,13), as SAR-Cov-2. The whole genome sequence of the 2019 -nCoV (Wuhan-Hu-1 isolate) showed that it includes 29903 nucleotide bp (GenBank: accession No NC\_045512.2 ) in ss-RNA,(14) while other few isolates of this virus are also submitted and published in genbank. The proteins encoded by gRNA of SARS-CoV-2 SARS-CoV and bat SARS-like coronaviruses, are similar in length but a notable difference was found in the longer spike protein (S) of SARS-CoV-2, bat SARS-like coronaviruses and SARS-CoV(15).The phylogenetic tree of human corona viruses along with other non-human corona viruses was constructed with their genomics which is presented in fig.1A (16).The novel CoV, SARS-CoV-2 is grouped into betacoronaviruses on genomic analysis, and is responsible for pneumonia in human, but it manifest milder symptom than SARS and MERS. Some workers argue that SARS-CoV-2 may have came from bats, but there is no definite evidence about the intermediate host that transmitted it to us (17,14). Symptoms of SARS-CoV-2 infected pneumonia include from asymptomatic to high fever and/or respiratory illnesses (throat soar, dry cough, problem in respiration) lymphopenia, and radiologic abnormality including predominant ground glass opacities in lower chest (18).The disorders of taste, smell or both, and dyspnea were

also noted in some patients in Germany (19). Although, SARS-CoV-2 has highest rapidity of transmission from human to human but mortality rate of it is below 5 % (20,21), and it may cause asymptomatic infections. The innate immune responses in human lung tissues due to SARS-CoV-2 was not significantly triggered the expression of any types of IFN at all. In addition, SARS-CoV-2 infection upregulated only 63.16% of inflammatory mediators (pro-inflammatory cytokines and chemokines) while much expression of all types of IFN and 84.62% upregulation of inflammatory mediators were noted in SARS-CoV infection (22). Probably due to this suboptimally induced innate immune response of patient, SARS-CoV-2 replicates quickly and accumulates in large quantity in the respiratory tract early and that is why this virus transmits efficiently from person-to-person via droplets or contact with secretions of patients containing high viral loads (23,24). It was noted that within the 48-hour interval, SARS-CoV-2, produced 3.20 times greater virion than that of SARS-CoV from the infected lung tissues (22). COVID-19 patients, in some cases, showed very mild or no symptoms. This may be the result of having poor level of innate immune response in the system of concerned patients. Infections in such patients, actually, makes it difficult to break the chain of viral transmission and keep the situation under control (22). This virus takes entry into us through the mucosa of the respiratory tract, mouth and the eyes. It is important to note that person-to-person transmission of SARS-CoV2 is happening and this is the reason why it took the form of pandemic quickly in almost all countries through rapid spreading (11,25-28). However, the actual mechanism which confers this viral characteristics of high transmissibility in all geographically different countries and asymptomatic infection is now under intense research. As of 10 June, 2020, a total of 7145539 cases have been reported to be infected and 408025 patients died among them globally(29). The estimated incubation period of SARs-Cov2 in human is claimed to have been recorded from the studies of some scientists (30-32). From their works. it was realized that maximum incubation period is 14 days(30) and therefore, Leung(30) suggested that quarantine period should be of three weeks (21days) instead of 14 days. Now, its highly contagious nature and mortality rate have brought the entire world to its knees and challenged the very existence of human beings. The concerned governments of all countries have deployed war-like emergency strategies, and the leaders in charge are leaving no stones unturned to reduce the speed of transmission of the disease, so that the infectivity curve of COVID-19 could be brought under control. WHO is monitoring the epidemiological, symptomological characteristic data from time to time and circulating these data through its website under situation reports. The situation compelled us to increase our safety measures to the highest level and maintain necessary personal and collective hygiene procedures properly in order to keep this deadly virus away. Moreover, in a bid to curb rapid transmission governments are taking extreme precautionary measures, and some of them are being imposition of lock down for minimum 21 days, social distancing or “physical distancing”, which is to keep safe distance of at least two meters between one person and the other while in public place frequent hand sanitization by alcohol based sanitizer or washing hands by soap or water, and covering mouth and nose by using a face mask while venturing outside of their home (33). Under such grim situation, potential drug and vaccine discovery to combat COVID-19 have become extremely urgent if human beings want to leave this depressing and disturbing period behind and move forward to see the light of a better future. Therefore, the objectives of this article are to have in depth discussion of virion structure, mode of attachment to human cell, entry into the cell, replication and transcription of RNA of corona virus including SARS-CoV-2 (COVID-19), and simultaneously this article wants to highlight some significantly meaningful points which

will further help in the gradual development and final discovery of appropriate drugs and vaccines as soon as possible.

## VIRION STRUCTURE

The virions of coronaviruses are round /elliptical/spherical having diameters of 60- 140 nm (34,35). Like other corona viruses, each virion of SARS-CoV-2 consists of three parts –a) outer membrane or envelope, b) nucleocapsid and c) genome or nucleic acid. The very diagnostic character of envelope is the club-shaped spike (S) projecting from outer side of virion. Nucleocapsid, having helically symmetrical, resides within the envelope of the virion.,and it is not common among +sense RNA viruses (Fig 1B)., The carboxy-terminal region of the M protein develops an extra internal layer, so, the viral membrane is unusually thick (34). Each virion of SARS-CoV2 contains four structural proteins -i) the spike (S), ii) membrane (M), iii) envelope (E) and iv) nucleocapsid (N) proteins. These proteins are encoded within the 3' end of the viral genomic RNA. i)The S protein – It is made up of 1,273 amino acids that construct spikes on the surface of SARS-CoV2(14) (NC\_045512.2) and S of SARS-CoV has 23 N-linked glycosylation sites(36). Its molecular weight (mwt) is about 150 kDa, and it is made up of trimeric glycoprotein. It has two parts- S1 and S2. Former makes the receptor protein domain and later forms the stalk of spike. The trimeric S glycoprotein behaves as a class I fusion protein (37) and attaches to the host receptor (38,39). In SARS-CoV-2, the S2 subunit is highly conserved. ii) The M protein- the topological model of M protein was presented byArmstrong et al (40). It is a smaller protein having three transmembrane domains and mwt ranges between 25-30 kDa (35,40,41). The differential maturation and subcellular localization of SAR-CoV surface proteins S, M and E were reported by Nal et al,(41). and they found that it (M) is divided into a small ectodomain(N-terminal) and a larger endodomain(C terminal) which expands 6–8 nm into the virion. The work of Neuman et al.(35) gave a clear structure and function of the M protein which has two different conformations, as it promotes membrane curvature and on the other hand it binds to the nucleocapsid.The M proteins of SARS-CoV2 contain 222 amino acids14(NC\_045512.2 ), iii) The E protein- SARS-CoV2 has 75 amino acids in N- proteins( NC\_045512.2). Its mwt ranges from 8–12 kDa having in small quantities and it is a transmembrane protein. The E protein of SARS-CoV consists of a C-terminal endodomain (CTEND) and an N-terminal ectodomain(NTED) and shows ion channel activity (42) and iv) The N protein –The highly phosphorylated N proein binds with g RNA maks a flexible, probably helical, nucleocapsid.The SARS-CoV2 shows that it contains 419 aminoacids (NC\_045512.2 ). Chang et al.(43) exhibited a modular organization of SARS-CoVs nucleocapsid protein and consists of two separate domains, an N-terminal domain (NTD) and a C-terminal domain (CTD), both bind RNA, but mechanism of binding are different, and it was reported that N protein also binds with M proteins(44). Another very important role of N protein is to encounter the interferon (IFN) and viral encoded repressor of RNA interference, as a result a microenvironment is created for smooth synthesis of virion(45.46). SARS-CoV-2 bears structural characteristics like other coronaviruses as described above.

The whole genomic analysis of SARS-CoV2 (isolate Wuhan-Hu-1) shows that it includes 29903 nucleotide(nt) bp (accession No NC\_045512.2 ) in non-segmented ss-RNA(14), while other few isolates of this virus have been also submitted and published in genbank and they are 29881 bp (isolate 2019-nCoV WHU01, GenBank: the accession no MN988668), 29838 bp (isolate 2019-nCoV\_HKU-SZ-002a\_2020. the accession no MN938384), etc. The isolate Wuhan-Hu-1 of SARS-CoV2 exhibits that it contains many genes and they are ORF1ab (21,290nt), ORF1a (13,218nt), S gene (3822nt), ORF3a (828nt), E gene (228nt), M gene (669nt), ORF6 (186nt), ORF7a (366nt), ORF7b (132nt), ORF8 (366nt), N gene/ORF9 (1260nt) and ORF10 gene (117nt). At 5' end UTR (untranscribed region) contains 265nt while at 3' end UTR has 229nt (GenBank: accession No NC\_045512.2). This polycistronic genome of SARs-CoV2 can be categorized into two parts: the first 2/3<sup>rd</sup> of the genome (ORF1ab & 1a genes) can translate to 16 non-structural proteins and while the remaining part of the genome(3-terminal) encodes the structural proteins (S, M, E and N) and some accessory proteins that are translated from a nested set of subgenomic mRNAs (sgmRNAs)47(GenBank: accession No NC\_045512.2)(Fig 1C). Previously, same thing is also recorded in other corona viruses(8,48). Notable thing is that SARS-CoV-2 bears 79% identity with SARS-CoV and 50% identity with MERS-CoV while 88% identity to two bat-derived SARS-like bat coronaviruses [bat-SL-CoVZC (45) and bat-SL-CoVZXC (21)] collected and identified in 2018 in Zhoushan, eastern China. SARS-CoV-2 is much divergent from SARS-CoV and it has been recognized as a new human-infecting betacoronavirus (15), so it is known as SAR-CoV-2. In general, a coronavirus genome (CoVs) contains a non-segmented, +ssRNA genome of 30 kb having a 5' 7mG( methylated guanine) cap structure along with a 3' polyadenylated (A) tail, permitting its function as an mRNA for synthesis of the replicase polyproteins polyprotein 1a/1ab (pp1a/pp1ab) in the infected cells. The genome encodes five open reading frameS(ORFs) that are replicated to genomic RNA and transcribed to six + subgenomic mRNA (+sgmRNA). The 5' end of the genomic RNA has the untranslated leader (UTL) sequence with the TRS- L( Transcription Regulation Sequence) in the downstream part. The TRS-L is very similar to sequences (TRS-B) that can be found in front of each open reading frame. It is well known that almost all nsps (nsp1 -16) have their specific function in the replication of Coronairuses but the role of few nsps are unclear.

## **VIRAL PROPAGATION IN HOST CELL**

### **Viral binding to host cell**

It is the first and foremost important step for infection or entry of virus into host cell. Here ligand and receptor interaction happens. The specific receptor present on the cell surface is to interact with specific virus or viral part. The presence of receptors are also tissue or organ specific. The angiotensin-converting enzyme 2 (ACE2) of human acting as receptor binds with spike of SARS-CoV and SARS-CoV-2(18) A list of corona viruses along with receptor is given in table 1.



**Table 1: A list of receptors used by different genera of coronavirus**

S.N	Virus genera	Virus	Receptor	Reference
1.	Alpha corona virus	TGEV(Transmissible gastroenteritis coronavirus)	APN(aminopeptidase N)	(39)
2.	Alpha corona virus	CCoV(Canine Corona virus)	APN(aminopeptidase N)	(8)
2.	Alpha corona virus	FIPV (feline infectious peritonitis coronavirus)	APN(aminopeptidase N)	(50)
3	Alpha corona virus	HCoV-NL63)(humancoronavirus NL63 )	ACE2(angiotensin-converting enzyme 2 )	(51)
4.	betacoronaviruses	SARS-CoV(Severe acute respiratory syndrome coronavirus) )	ACE2(angiotensin-converting enzyme 2)	(49)
5.	betacoronaviruses	SARS-CoV2 (Severe acute respiratory syndrome coronavirus 2) or COVID-19	ACE2(angiotensin-converting enzyme 2 )	(18,47)
6.	betacoronaviruses	MHV(murine hepatitis virus),	Ceacam1a (carcinoembryonic antigen-related cell adhesion molecule 1a)	(52)
7.	betacoronaviruses	MERS-CoV (Middle East respiratory syndrome coronavirus )	DPP4(dipeptidyl peptidase 4)	(53)
8.	betacoronaviruses	human coronavirus OC43	9-O-acetylated sialic acid, HLA class 1( human leukocyte antigen), IFN-inducible transmembrane (IFITM) proteins in endocytic vesicles	(8,38, 54)

The RBD(receptor binding domain) of the S1 protein i.e ligand interact with ACE2 for binding. Both SARS-CoV and MERS-CoV deploy the S1 CTD as RBD to recognize the receptor ACE(55). The region in SARS-CoV-2 S protein that is responsible for ACE2 interaction remains under research but recent work of Wang et al.(56) that utilized flow cytometry and immunostaining assays, first revealed that the S1 CTD serves as the main location in SARS-CoV-2 that binds to the ACE2 receptor. The S protein of SARS-CoV-2 exists in the form of a trimer, each monomer contains approximately 1,300 amino acids(57,58). Wrapp et al.(57) and Walls et al(58) have just reported the cryo-Electron Microscopic structure of spike trimer of this virus in two independent studies. The spike proteins of SARS-CoV-2 (SARS2-S; 1,273 residues,

strain Wuhan-Hu-1) and SARS-CoV (SARS-S, 1,255 residues, strain Urbani) are 77.5% similar by sequences of amino acid, are very identical in structure(57,58) and easily bind the human ACE2 protein as a host receptor(59,60) through their S1 B domain. Modern research also stresses on the key function of ACE2 in facilitating the entry of SARS-CoV-2(58,61,62). The S protein is next to become broken by host proteases into a S1 subunit(N-terminal) and a S2 subunit (membrane-bound C-terminal)(63). The research on crystal structures of SARS-CoV RBD complex with ACE2 exhibited that SARS-CoV RBD has a core structure and a receptor-binding motif (RBM) and that the RBM interacts with the outer surface of the claw-like structure of ACE2(64,65). Importantly, scientists noted two virus-binding hot spots on human ACE2(64,65) and also recorded mutations on these two hot spots, and these regions select the host range of SARS-CoV. Moreover, Wan et al (60) recorded selective amino acids at the position of 442, 472, 479, 480, and 487 of RBD of SARS-CoV-2 that enhance viral binding to human ACE2 and they gave the sequences of amino acids of RBD of SARS-CoV2 and compared with others corona viruses (Civet -SARS-CoV2002, SARS -CoV2002, Bat -SARV2013) and showed ACE 2 contacting residues, fully conserved residues, strongly conserved residues and non-conserved residues. The amino acid sequence residue of RBD (SARs-CoV2) was from 319-529 but in case of RBD (SARs-CoV2002) it was 306-515. On the other hand, Lan et al (66) compared amino acid sequence residue RBD 387-516 of SARs-CoV2 with RBD amino acid sequence residue 374-502 of SARs-CoV 2002. Moreover, from their constructed complex of RBD(SARS-CoV2) and human ACE2, hot spot 31 is matchable with Gln493 in SARS-CoV2 RBD. As per the structure of this designed RBD, residue 479 is situated near virus-binding hot spot Lys31 (i.e., hot spot 31) on our ACE2. A salt bridge exists between Lys31 and Glu35 in a hydrophobic condition(60). The gross similarities between SARS-CoV-2 spike and SARS-CoV spike (isolated from human, civet, or bat) are found from 76 to 78% for the whole protein, while it was near 73% to 76% for the RBD, and 50% to 53% for the RBM(60). Furthermore, they reported that RBD has 14 ACE2-contacting residues, but 4 are partially conserved and 9 are fully conserved among SARS-CoV2 and SARS-CoV from human. Similarly, Lan et al.(66) presented gross topology of SARS-CoV-2 spike monomer citing N-terminal domain (NTD), receptor-binding domain (RBD, 333-527 aa), receptor-binding motif (RBM, 438-506 aa), fusion peptide (FP), subdomain 1 (SD1), subdomain 2(SD2), heptad repeat 1 (HR1), heptad repeat2 (HR2), intracellular domain(IC) and transmembrane region (TM), (Fig 2A). Moreover, we came to know from their important work on the SARS-CoV-2 RBD that it has a twisted five-stranded anti-parallel  $\beta$  sheet ( $\beta$ 1-  $\beta$ 4 and  $\beta$ 7) with loops and short connecting helices. The extended insertion (  $\beta$ 5 and  $\beta$ 6 strands,  $\alpha$ 4 and  $\alpha$ 5 helices and loops) is the receptor-binding motif (RBM) containing most of the contacting residues of SARS-CoV-2 for ACE2 binding. The RBD contains nine cysteine amino acids. The N-terminal peptidase domain of ACE2 consists of two lobes which form the site which binds with substrate between them (Fig 2B). Lan et al.(66) reported that the ACE2-interacting amino acids on the SARS-CoV-2 was 17 amino acid positions and their positions are K418, G447, Y450, Y454, L456, F457, A476, F486, N487, Y 489, Q 493, G496. Q498, T500, N501, G502 and Y505(Fig 2C ). Wang et al.(56) have

done an atomic comparison of the two viral ligands (SARS-CoV-2-CTD and SARS-RBD) binding the receptor ACE2. It revealed more interactions in SARS-CoV-2-CTD/ACE2 than in SARS-RBD/ACE2. It was also validated by other workers (57). Walls et al.(58) also reported that the SARS CoV 2-S1B receptor binding domain (residues 338-506) has a core domain and a subdomain (residues 438-498) for receptor binding and looping out from the antiparallel beta sheet core domain structure that directly engages the receptor. An irreversible conformational change of spike proteins is noted to be triggered due to receptor interaction which results membrane fusion.

The ACE2, was recorded as ACE homologue. The protein encoded by the gene is grouped to the angiotensin-converting enzyme family of dipeptidylcarboxypeptidases. Genomic structure comparison suggests that ACE2 and ACE genes have been created by duplication of a common ancestor(67). ACE2 has 805 amino acids and is a type I transmembrane glycoprotein (metalloproteinase). The ACE2 gene is located on the X chromosome. ACE2 consists of two domains: an amino-terminal catalytic domain and a carboxy-terminal domain. The catalytic domain has an functional zone known as the zinc metallopeptidase domain (HEXH motif) and exhibits 41.8% sequence similarity with the amino domain of ACE(49) (Fig. 2B). The affinity of binding of ACE2 with the RBD (receptor binding domain) of SARS-CoV-2 is many times higher than its affinity with the RBD of the SARS-CoV(49). In addition, some workers (68) reported that the M protein also play a vital role during early stages of HCoV-NL63 infection, and that the concerted action of the two proteins (M and S) is a prerequisite for effective infection. The variation of binding capacity at different temperatures demonstrated that at both 4 and 37°C, virions can bind to the surface of host cells as noted for the LPV(B-lymphotropic papovavirus)(69) while in case of HIV-170, the affinity of virion-binding at 37°C is higher than that at 4°C but internalization only occurs at 37°C. Kuba et al.(71) using a flow cytometry assay exhibited that SARS-CoV S protein is internalized by VeroE6 cells together with ACE2 at 37°C.

The correlation between receptor expression in human and infectivity of coronavirus was searched by some scientists. ACE2 expression in human tissues has a positive correlation with possibility of SARS-CoV infection, including lung and intestine (72,73). SARS-CoV has much tendency to infect ciliated epithelial cells expressing ACE2(74). Wang and Cheng (75) noted that SARS-CoV-2 upregulated the expression of ACE2 in lung tissue, as a results viral replication and transmission becomes increase and but a negative correlation between ACE2 expression and SARS-CoV-2 (COVID-19) severity and fatality at a population level was reported by Chen et al.(47). More research is required to determine the actual receptor–ligand interaction. More than one receptor may play here. The proper understanding of this process can quicken the development of effective vaccines and antiviral drug development.

### Entry into host cell

After attachment on host receptor, next important step in viral disease is the invasion of the viral RNA into the host cells. As we know that ACE2 is common receptor for both SARS-CoV and SARS-CoV2, so here pathway of entry of SARS –CoV2 will be discussed along with other corona viruses. Now, it is confirmed that genomic RNA of CoVs invades the host cells via two paths: (i) the direct membrane fusion or non-endocytic pathway- here viruses put their genomes



into the cytoplasm of host cell after their envelopes join or fuse with the plasma membrane at the cell surface and it is pH independent (ii) endocytotic pathway(76). It occurs through endocytic machinery. It is influenced by the acidic endosomal pH, which triggers the fusion of viral and endosomal membranes and ultimately of the viral gRNA enters into the cytoplasm.. SARS-CoV has adapted to enter into cells by direct fusion at the plasma membrane(77)(Fig 3A). In this case, virus fusogenic mechanism was described by Liu et al.(78). According to them, as soon as the S1 subunit of SARS-CoV S protein attaches to ACE2, the S2 subunit rearranges conformation by invading the fusion peptide(FP) into the plasma membrane. The HR2 domain reacts with the HR1 trimer to make 6-HB core, directing the fusion between the cell membrane and envelope of CoV(78). Further work has confirmed that entry of SARS-CoV is pH-dependent(79), and that the endosomal protease cathepsin L(80) is engaged here, advocating that this virus follows endocytosis. So, this pathway has been vigorously worked out.

Some scientists worked out by flow cytometry internalization assay to ascertain that the RBD spike protein binding directs the endocytosis of SARS-CoV by infected cells. So they confirmed that cellular endocytosis is a mechanism for SARS-CoV entry (81,82). Wang et al. (82) propose that the RBD S- protein induces the ACE2 directed cellular endocytosis signal pathway, as a result, gRNA of SARS-CoV invades the host cells. As the binding of SARS-CoV RBD spike protein triggered ACE2 internalization, now question may arise whether N-linked glycosylation on the RBD could affect ACE2 internalization or not. It was noticed that deletion of N-glycans from RBD-Fc can still induce ACE2 internalization(83). It is very interesting to note that after internalization of spike S or virus, down regulation of receptor occurred. For instance, the down-regulation of the receptor (CD46) occurred due to attachment of measles hemagglutinin, to imbalance the complement pathways and immune systems(84). So, after utilizing the receptors for cell entry, some viruses influence down-regulation of the receptor to disturb its normal function, causing severe disease.

The entry of SARS-CoV has been also recorded through Clathrin-mediated as well as clathrin- and caveolae-non-mediated pathway(81,82). The study of Inoue et al (81) indicated that the pathway of SARS-CoV entry into host cells is mediated by clathrin. Initially viral S1 binds to receptor ACE2 on plasma membrane, next a pit is formed by invagination of plasma membrane along with S1- receptor complex and its membrane is moulded by clathrin triskelions. These pits form early endosomes which are acidic (pH 6.5 to 6.0) and early endosomes becomes late endosomes and more acidic pH 5.5)(Fig 3A). The acidic condition of endosomes facilitates viral infection(85,86). The endosomal acid proteases (cathepsin L) mediates the cleavage of S into S1 and S2 to enhance its fusion activity. After the fusion peptide (FP) inserts into the endosomal membrane, a six-helix bundle (6-HB) core is built by interacting among the heptad repeat 1 and 2 (HR1 and HR2) domains and the S2 protein. This core structure facilitates the the cellular plasma membrane and viral envelope to come close to each other for fusion (Fig 3B). As a result the ACE2/ virus complex becomes translocated to endosomes, and the virus is uncoated by the help of endosomal acid protease, cathepsin L(80,77). Finally, Wang et al.(82) suggest that SARS-CoV penetrate cells through receptor-mediated, clathrin-independent, caveolae-independent endocytosis, likely through a process which is associated with lipid rafts. In different cell types, same virus invades through different endocytotic pathways. For example, the entry of SARS-CoV is mediated by the clathrin-dependent endocytosis pathway into HepG2 cells(81) while it relies on a novel clathrin- and caveolae-independent endocytotic pathway for

entry into VeroE6 cells(82). According to Burkard et al,(87) coronaviruses take entry into host cell through Endo/Lysosomal pathway in a proteolytic processing of fusion proteins by lysosomal protease. They proposed two models like early (e.g. HCoV NL 63 ) and late coronavirus fusion(e.g. HCoV). At present, questions may arise whether CoV induces autophagy, and whether the autophagy machinery or ATG proteins are engaged for the infection and replication of CoVs. Viral attack was not arrested by the knockdown of ATG5 gene (88), so it confirmed us that the autophagy was not involved in the viral infection or replication.

### **Replication of genome, subgenomic RNA or m RNA formation, nsps and structural protein and RTC formation**

The whole genome sequences, replication, transcription and translation to proteins or enzymes of SAR –CoV2 have been recorded since December 2019 (accession No NC\_045512.2) (22, 24, 65, 56, 47) but actual mechanism of these processes of this virus have not been reported. As it belongs to same group of SARS-CoV, we assume that this virus follows same mechanism of replication, transcription and translation of SARS-CoV or other corona viruses, so, mechanism of synthesis of coronaviruses including SARs-CoV has been reviewed here. The replication or synthesis entirely happens in the cytosol of host cell. In the first step of viral synthesis, translation of pp1a and pp1ab for production 16 non-structural proteins or enzymes that are necessary for viral further biosynthesis. As corona viruses including SARS-CoV 2 belongs to + SS RNA, their genome can be directly utilized as mRNA and they use RNA directed RNA transcription in their replication.

### **Translation of pp1a and pp1ab and nsps formation**

After entry and uncoating, RNA genome attaches with host ribosome and exploits host RNA, amino acid pool and other factors, and initiates translation of the corona virus genome at the replicase-(RdRp) ORF1a start codon and translation of ORF1a gives polyprotein 1a (pp1a), and – 1RFS(1 ribosomal frameshifting) directs the translation of ORF1b to produce pp1ab, ultimately through co- and posttranslational processing or cleaving by virus-encoded proteinases like papain-like (PLpro) that reside in nsp3 and chymotrypsin-like main proteinase (Mpro) in nsp5 of pp1a, into 16 nonstructural proteins(nsps); which are utilized later for replication of genomic RNA and translation of six structural proteins. ORF1a and 1b are linked by a -1ribosomal frameshift site (89). During synthesis, nsp1 to nsp11 are encoded in ORF1a, and nsp12 to nsp16 are encoded in ORF1b(90)( Fig 4A). Nsps 4–16 are conserved in CoVs (91). The functions of 16 nsps are listed in table 2.

**Table 2 List of non-structural proteins (Naps) and their functions with references**

Non-structural proteins(Naps)	Function	Reference
Nsp1	Decaying of Host mRNA, Interfering Translation, arresting cell cycle, inhibition of INF signaling	(93,94)

Nsp2 , Nsp11	Yet to known	-
Nsp3	PL1 pro and PL2pro (papain-like proteases), in the assembly of virus and has poly (ADP-ribose) binding abilities, and deubiquitylating activity in its protease domain	(92,95)
Nsp4 and Nsp6	DMV formation, Potential transmembrane scaffold protein	(96, 97)
Nsp5	3CLpro( chymotrypsin-like Protease) or Mpro (Main protease )	(92)
Nsp7	ssRNA binding	(98,99)
Nsp 8	Primate	(99)
Nsp9	Partial function of replicase enzyme	(100)
Nsp10	Partial function of of replicase enzyme	(100)
Nsp12	RdRpolymerase	(101)
Nsp13	Nucleoside triphosphatase, RNA 5 triphosphatase, helicase activity	(102,103)
Nsp14	3-5 exoribonuclease, RNA cap formation ( guanine-N7 )-, methyltransferase , role in RNA proof reading	(104,105,91)
Nsp15	Viral endoribonuclease, NendoU	(106,103)
Nsp16	RNA cap formation (2 O-transferase)	(107)

It has been recently reported that, a spike mutation of CoV, which probably happened in late November 2019, directed jumping to humans. In particular, Angeletti et al (108) compared the SAR-CoV-2 gene sequence with that of SARS-CoV. A genetic analysis of SAR-CoV-2 and SARS-CoV, revealed that the transmembrane helical segments in the ORF1ab encoded 2 (nsp2) and nsp3 and exhibited that at the position 723 a serine replaced a glycine residue, while in the position 1010 proline replaced isoleucine(108). The viral mutations is responsible for potential the relapse of this viral disease.

### Programmed ribosomal frameshifting

An alternative way of translation of merged proteins encoded by two overlapping ORFs is Programmed Ribosomal frameshifting (PRFS). In this case, the frameshift appears at low frequency and ribosome's slipping by one base in either the 3'(+1) or 5'(-1) directions during translation is key matter. Specifically, -1 ribosomal frameshifting appears more frequently in RNA viruses (Fig 4B, C). Eukaryotic ribosomal frameshift signals usually composed of two elements: a heptanucleotide (a 7-bp) slippery sequence (XX XY YYN) where X indicates any nucleotide, Y indicates A or U, and Z is A, U, or C (the frame of the initiator AUG is denoted by the spacing), and an H-type RNA pseudoknot (PK) located downstream (**109,110**). Each Pseudoknot of SARV-CoV contains 79 nucleotides of a secondary RNA substructure that contains two stacked stems 1(S1, 18nt) and stem2 ( S2, 30 nt) connected by a ssloop(single stranded) (L1, 3nt ) and a big stem-loop (SL1, 28 nt) (111) (Fig 4D). Later Plant et al.(112) also mentioned the third stem (S3) in SARS-CoV and MHV, and third pseudonot structures were

analysis by mutation. A stem-loop SL1 with a proximal ds-segment (double-stranded) segment which exists in SARS-CoV, is absent in other characterized coronavirus *1a/1b* pseudoknots(11,113). The  $\pm$  ribosomal frameshifting is generally occurred due to presence of slippery sequence and pseudoknot but, there are alternative ways for formation of “out-of-frame proteins”. Alternative splicing can cause frameshifting (113).

However, in programmed -1 ribosomal frame shifting, the ribosome is bound to move one nucleotide backwards into an overlapping reading frame and to translate an completely new amino acids sequence. It is very common SARS-CoVs(114). In SARS-CoVs, the stimulatory structure is to be an mRNA pseudoknot as described in the infectious bronchitis virus(IBV) where a ‘slippery’ sequence of the type UUUAAC and a H-type pseudoknot are able to induce translation of the zone 1b of the polyprotein 1a/1b(115). The SARS-CoV genome has been found to be responsible for encoding an mRNA segment that directs -1 ribosomal frameshifting (36). Smith et al.(116) and Tu et al.(117) recorded that ribosomal movement was stalled at a pseudoknot necessary for frameshifting (Fig 4B,C). Many models are present to analyze -1 ribosomal frameshifting operation by pseudoknots. It is found that RNA pseudoknot structures force to stall ribosomes over a slippery sequence, at which the ribosome-bound tRNAs realign in the -1 frame, but that pausing alone is not sufficient to operate frameshifting (111,118). To discuss the in-depth mechanism of it at the gene level, Namy et al (89) introduced a method of purifying “rabbit reticulocyte lysate (RRL) ribosomes”. Cryo-EM (Cryo-electron microscopy) study of purified mammalian 80S ribosomes from rabbit reticulocytes halted at a coronavirus pseudoknot revealed an “intermediate of the frameshifting process”. The translating 80S ribosome, halted at the IBV pseudoknot (Fig 4C), contains a P-site tRNA and elongation factor 2 (eEF2). Frank, and Agrawal (119) noted that the 70S complex exhibited “a ratchet-like rearrangement of the ribosome” which is connected to the trapping of eEF2 in *E. coli*, and in 80S complex, “a ratchet like rearrangement of the ribosome” connected to the trapped eEF2 in yeast by sordain antibiotic treatment (120). So, they (119,120) proposed a model where “the ratchet-like rearrangement of the ribosome” is a part of a mechanism for moving the tRNAs during the translocation. Similarly, experimental results observed by Namy et al, (89) where no antibiotic was applied, also showed a ratchet like rearrangement of the ribosome in eukaryote. Plant et al.(121) have proposed a “golden mean” model in which viruses utilize both “programmed ribosomal frameshifting” and “translational attenuation” to maintain the relative ratios of their encoded proteins.

### RTC formation

The SARS-CoV RTC (Replication-transcriptional complex) is the assemblage of many of the nsps including RNA-dependent RNA polymerase (RdRp) and unidentified host factor. RTCs are attached in double-membrane vesicles (DMVs) suitable for RNA synthesis to create an environment, and ultimately replication of RNA and transcription of the sub-genomic RNAs (sgRNAs) sequences are held. These DMVs are virus induced structure from intracellular membranes. The viral “RNA-dependent RNA polymerase (RdRp)” is the key enzyme of this ‘replication/transcription complex’ (RTC). RNA replication is seemed to happen on DMVs (122). One common feature among plus-strand RNA viruses is that RTCs is linked with cell membrane (123). The mechanism by which the DMVs are developed are yet to known. In

the work of Bechill et al.(124) it was found that uPR(unfolded protein response)) may mediate DMV formation as it is promoted during coronavirus infections. Snijder et al. (122) and Snijder et al.(125) recorded that the DMVs of SARS-CoV are most likely derived from the endoplasmic reticulum (ER). The results of Knoops et al. (126) showed that DMVs are likely to originate from the part of a reticulovesicular network of modified ER membranes. Later on, these networks become a large single-membrane vesicles. gRNA ( genomic RNA) which are synthesised in infected cell, are packed into virions on membranes which are situated between the endoplasmic reticulum (ER) and the Golgi apparatus ( “ER–Golgi intermediate compartment (ERGIC)” (127). Proteins ( nsp3, nsp5 and nsp8) engaged in the replication of virus are situated in adjacent reticular structures but not inside of DMVs. RNA, which is either “replicative intermediates” or “dead end” double-stranded RNA, was found in DMVs. The non-ionic detergents disrupt all RTC and make SARS-CoV RNA to susceptible for breaking by nuclease, and this reminds us the intact membrane structure of RTC for RNA synthesis (128). It is found that pool of replicated genomic RNAs of SARS-CoV are connected with the RTC, while sg RNAs after synthesis come out from the RTC structures. Moreover, it has been recorded that RTC activity relies on a host factor which is present in cytosol, that indicates that crosstalk between RTC and cytosol, takes place via channels that help transport through membranes (128). The single DMV can move freely but when many are “captured” by the DMV/CM(convoluted membrane) assemblies, they are unable to move. It was also reported that disruption of microtubule-dependent transport of DMVs, did not interfere on RNA replication (129). In coronaviruses the RTC formation might be one of strategies for involving of cellular and viral proteins or enzymes for successful replication of g RNA and other parts of virus and also for creating a safe place from the attack of host defense (129).

### **RNA replication and transcription of the sub-genomic RNAs (sgRNAs) sequences.**

During genome replication and sgRNA synthesis, two replicative intermediate (RIs) such as RI -1 and RI- 2 are operated. From RI-1, – sense full genome strand was formed on the template on + sense ss RNA by viral encoded the RNA dependent RNA polymerase (RdRp). In RI-2, - sense RNA genome strand is used as template to generate + sense genome of virus by viral encoded RdRp ( Fig 5A) . The TIs (Transcription intermediates) and TFs(transcription forms) are main transcription structures that are responsible for synthesis of subgenomic mRNA (130). Here the exiting recent model of coronavirus RNA synthesis reveals that minus strand RNAs come from copying the + RNA continuously to form genomic templates and discontinuously to create subgenomic templates (131).

### **Formation of subgenomic RNA template(-) and subgenomic +stand mRNA**

**There are many models for** formation of subgenomic RNA template(-) and subgenomic +stand mRNA. Taking consideration of transcription of MHV, three models were proposed over the years. They are as follows-1. Loop out model: It was the first model in which the RNA polymerase starts the replication of the leader RNA from the 3' end of the minus (-)-stranded



RNA template. As soon as replication of leader RNA completed, the polymerase enzyme "jumps" over to the different initiation sites for several mRNA species, so, it is probable to create "loop out" of the RNA template. This model is not supported by maximum workers, 2. Post-transcriptional processing model (Fig 5B). This model indicates that synthesis of the leader RNA and mRNAs are independent; after the finishing of synthesis, the leader RNA is then bound to the body sequences of the mRNAs by a mechanism which is yet to know. UV transcription inactivation studies were performed and the results indicated that the formation of mass of subgenomic RNAs were not possible post-transcriptionally by cis-splicing of a genome-length precursor molecule (132) and 3. "leader primed transcription model" (Fig 5C): This model exhibits that the leader RNA is replicated and "falls off" from the template. Then a viral RNA dependent RNA polymerase (RdRp) binds with this free leader RNA and initiates mRNA synthesis at several initiation points. This model indicates that the leader RNA serves as a primer for RNA synthesis. Lai et al. (133) gave this model and it is known as "leader primed transcription model" (134). Out of these three models, this one was accepted by many authors by few modification.

At the 5' end of the genome and all subgenomic mRNAs, there is an identical "leader" that has sequence which is 65 to 98 nt in many CoVs (135). As this leader sequence is located to the 5' end of the genomic RNA, viral RNA synthesis has a mechanism which directs the leader RNA to join the "body leader" of each sg mRNAs that are situated at the 3' end of the genomic RNA. Moreover, at the end of the leader and before the body of each ORF of the genomic or subgenomic mRNAs there is one TRS ("transcription regulating sequence") or "intergenic sequence," ("IGS" or IS) (135-137). TRS which is situated near the leader sequence at the 5' end, then it known as TRS-L and TRS which is preceding each viral gene or ORF is called TRS-B. If we go in details regarding TRS, we found that each TRS has a conserved core sequence (CS) bearing 6–7 nt and variable 5' and 3' flanking sequences (the 5' TRS and 3' TRS, respectively) (138). The CS is similar for the genome leader (CS-L) and all mRNA coding sequences (CS-B). The study of secondary structure of the TRS-L region from TGEV (transmissible gastroenteritis virus) and BCoV (bovine coronavirus) exhibited that the CS-L is extended like a hairpin-like structure which is suitable for replication and transcription (139). The transcription mechanism is monitored by TRSs (140). The viral RdRp complex recognizes TRSs. At these locations, the polymerase complex either reads through to the next TRS or detaches from the template strand, then binds with the leader TRS, situated in the 5' UTR, and as a result, the synthesis of a set of sg-strand RNA bearing an antileader RNA sequence, become completed. The TRS of nidovirus is an AU-rich element (eg 5' UCUAAAC3' in MHV) that is situated at (i) the 3' end of the common leader, (ii) the 5' end of each mRNA 'body' segment, and (iii) fusion between leader and body in the sg mRNA, so here like MHV, it assumed that nucleotide base pairing between minus- and plus sense copies of this regulatory sequence might guide the co-transcriptional fusion of sg mRNA leader and body by discontinuous transcription manner (133,134). Although different models are present to analyze the fusion of the common 5' leader sequence to the different 3' body segments in arterivirus and coronavirus sg mRNAs,

but almost co-transcriptional fusion of leader and body is being reported in all models (135,141,142). The RNA-dependent RNA-polymerase has been seemed to halt after a body TRS(TRS-B) of a particular gene is replicated during (-) strand synthesis, therefore, switching to the TRS-L occurs and thus a common L sequence to each sg mRNA is added.

The gRNA and 6 sgmRNAs are formed during replication of MHV-A59. It was observed that 1stsgmRNA- 6th sgmRNA are 9.6, 7.4, 3.4, 3.0, 2.4 and 1.7 kb respectively. All together form a 3' co-terminal nested set. The sgmRNAs, in size, ranges from 1/3rd (1stsgmRNA) to about 1/20th (6<sup>th</sup> sgmRNA) of the genomic RNA. It was also recorded that + strands (genomes and subgenomic mRNA) are formed in huge amounts but out of them about 1% minus strands of both genome- and subgenomic length act as the templates for genome and subgenomic mRNA synthesis(143,144). It was also reported that the newly synthesized viral genomic RNAs may act as template for the formation of sg-length minus strands. Experimental studies conducted by Stanley and Dorothea Sawicki and coworkers (130), where they used MHV as a model, revealed that both –strand gRNA and sg-length minus strands are formed during very early stage of infection. Each sg mRNA is formed from a corresponding transcription intermediate (TI) which bears the sg-length minus-strand template. Many sg mRNAs are synthesized from these complexes in constant amounts but in non-equimolar. It appears that the ratio of the synthesis of genomic RNA to sg mRNAs is constant in the whole replication stage (130,131).

Next question has been raised whether the discontinuous process of RNA replication occurs during minus- or plus strand synthesis. To solve the matter or question two prominent, models -i) 'leader-primed transcription' with discontinuous plus strand RNA synthesis (Fig 5D) and ii) 'discontinuous extension of minus-strand RNA synthesis'(Fig 6) and they are opposite to each other. In both cases the TRS elements acts as major role. A base-pairing interaction between the sense copy of the TRS in the genomic leader (leader TRS) and the antisense copy of the TRSs at the 5' end of each of the sg mRNA body segments (anti-body TRS) is a common phenomenon in these two models. The 'leader-primed transcription' model with discontinuous + strand RNA synthesis (133,134) exhibited that transcription is started from the 3' end of the anti-genome(-strand RNA) to create a leader primer of which the 3'-terminal leader TRS would base pair to the various anti-body TRSs in the anti-genome. Later on, the leader transcript would be elongated to develop the sg mRNA. Thus, this model gives us that the discontinuous process proceeds during plus-strand synthesis and that the body TRS complements in the – strand g RNA or anti-genome invariably play as promoters for transcription. There was no detection of sg-length -strands in coronavirus-infected cells validated the leader-primed transcription model with discontinuous + strand RNA synthesis (133). Moreover, the subsequent discovery of such molecules, for both coronaviruses and arteriviruses (145,146), generated the idea of alternative models. Brian et al. first time first reported a nested set of sg -strand RNAs in TGEV infected cells, (145) and they gave us the idea that the sg-length -strands were intermediates in the production of sg mRNAs. Subsequently, Stanley and Sawicki (147) gave us discontinuous extension of - strand RNA synthesis as an alternative way (Fig. 6). They confirmed that minus-

strand sg RNA synthesis was operated discontinuously, with attenuation of nascent strand RNA synthesis occurring in the different body TRS regions of the genomic template, and while plus strands were synthesized continuously. A small portion of the polymerase complex detaches from the +strand template when it comes to TRS to develop an -sg RNA. So, the synthesis of a longer -sg RNA, i.e., the product of the upstream TRS, is attenuated by the presence of a downstream TRS because only a little percentage of the polymerase complexes reach the upstream TRS. The nascent sg-length minus strand, bearing an anti-body TRS at its 3' end, become moved to the leader TRS (TRS-L) in the genomic template and the CS-L could base-pairing process proceeds between CS-L and CS-B (cCS-B) as the CS is similar to the genome leader (CS-L) and all mRNA coding sequences (CS-B), and it allows for leader-body joining (Fig 6) and the polymerase (RdRp) would continue and duplicate the leader sequence to put an antileader (complementary to the 5' plus-strand leader sequence) on the 3' end of the nascent, subgenomic minus strand (Fig 6). Here the anti-TRS plays as a primer to finish transcription and duplicates the leader RNA (141). Ultimately the subgenomic minus strands are produced and they perform as templates for the transcription of subgenomic mRNAs (Fig 6). The genomic and subgenomic RNA synthesis take place in TRC/DVC. Later the RTCs employing in plus-strand synthesis became aged and released their minus-strand templates, which are then decayed, and fresh RTCs are produced in infection (131). Both genetical and biochemical researches validated this model (130,131,148).

According to some workers, the RNA synthesis of coronaviruses must have to involve by cis-acting RNA elements, that are situated in structured 5' and 3' untranslated Regions (UTRs), and into the associated coding sequences (149-151). These RNA elements have stem-loops (SL1, SL2, SL3, SL4 and SL5) structures which are conserved but degree of conservation among different coronaviruses are variable, and all have some role in replication of RNA, discontinuous synthesis of sgRNA and sgRNA (149,151,152). Very primary research work using defective interfering RNAs from alpha-, beta-, and gamma-coronaviruses limited the 3' cis-acting RNA elements needed for synthesis of RNA of CoVs coronavirus to the 3' UTR plus the poly(A) tail (150) but later it was found in other gene like at down stream N gene stop codon where BSL (Bulged Stem Loop) and H-type PK (pseudoknot) are involved. So, a model for the initiation of replication of negative-strand RNA of CoVs was postulated (140,153) (Fig 7).

It may be assumed that the production of subgenome length minus-strand (sgl-s) templates with the same 5' and 3' ends takes place from the genome-length minus strands by splicing but this possibility was not supported by experimental results (131). Finally, many workers have supported discontinuous extension of minus-strand RNA synthesis' model and analyzed in detail in their reviews (141,148). But there are some limitations of this model to explain some facts or questions like what does constitute the signal that attenuates or arrest minus-strand synthesis at each of the body TRS motifs? How does "the relocation of the 3' end of the nascent minus strand to the 5' end of the genomic template" happen (131)? Sola et al. (140) proposed a three step model of transcription of corona virus on the basis of experimental data on corona

virus transcription as shown by some workers (154-156). The steps are: (i) Complex formation (ii) Base pairing scanning and (iii) Template switch (for details ref 140).

The regulation of transcription process have suggested that it is regulated by many factors through monitoring the template switch frequency during discontinuous transcription (131,149). These factors are like complementarily between the leader TRS and the body TRS-B(154), TRS secondary structure to the 3' end, RNA-RNA or protein-RNA interactions eg N-protein RNA chaperone (149,155), cell factor RNA helicase DDX1.(158), the two polymerases nsp8 and nsp12(159), etc. It is very interesting to note that Nucleocapsid phosphorylation and RNA helicase DDX1 recruitment enable coronavirus transition from discontinuous to continuous transcription (158).

After synthesis of nested sgRNAs (10 sgRNAs in SARS-CoV-2), each of them undergoes translation by host ribosome, amino acids and other factors and express its gene as protein. As a results, four known structural proteins(sps) like S, E, M and N and other proteins in coronaviruses but in case of SARS-CoV2, in addition to sps, others are ORFs proteins (3a, 6, 7a,7b, 8 and 10) in SARS-CoV2 (NC\_045512.2, NCBI, Baltimore) (Fig 6). Throughout the discussion, we found that all RNAs and protein synthesis take place in solely cytoplasm, there is no involvement of host nucleus. But researches of some workers reported that some viral proteins (e.g N, nsp 3, 6, 9b etc) have been located in nucleus. N protein enters into nucleus through nuclear pore complex and it later induces the cell cycle blockage and impediment of cytokinesis. (160,161). Protein 6 inhibits nuclear import of factors like STAT1(162) and nullifies IFN signaling pathways(163). Although viral proteins are observed in nucleus of diseased cells, its actual mechanism of entry into nucleus is unknown.

### **Assembly and release of virion**

SARS-CoV structural proteins then moved to inside of the ERGolgi intermediate compartment (ERGIC).The envelope glycoproteins (E)after synthesis are attached on the membrane of the ER(endoplasmic reticulum) or Golgi, and the nucleocapsid is crated by N- proteins. Newly synthesized genomic RNA is then packed in nucleocapsid. The phosphorylation of N protein helps to induce structural change enhancing the affinity for viral versus nonviral RNA. N protein attaches the viral genome in a “beads-on-a-string type conformation”. Specific interaction between leader RNA and nucleocapsid protein of CoVs has been reported (164) and the positive relationship between the nucleocapsid protein and the genomic packaging signal in MHV was also recorded (165). The C-terminal RNA binding domain is involved in this packaging signal (166). Then N-proteins are incorporated during budding to develop complete virion (127). Hurst et al (167) observed an interaction between the N protein and nsp3 which is associated with replicase-transcriptase complex (RTC) and showed that N protein also binds with M proteins(8). M proteins are found to play an important function during viral assembly by interacting with other structural proteins of SARS-COV. At last, the vesicles containing the virions join with the cell membrane and spread the virus through exocytosis from the diseased

cells. E-proteins of SARS-CoV also serves a vital function on assembly and release of the virus and pathogenesis (42). Finally after assembly of viral parts and maturation of virion in ERGIC (Endoplasmic reticulum Golgi body intermediate compartment or Golgi body), buds with virions come out from Golgi body and form vesicles. These vesicles with virions approach to plasma membrane and by exocytosis virions are released from infected host cell (Fig 8).

### Meaningful points to note for development of antiviral drugs and vaccines against SARS-CoV-2

Till now, no approved effective antiviral drug or vaccine has been developed against SARS-CoV2 (COVID-19). Scientists of many private companies and Government institutional laboratories of some countries putting their hard labour around the clock to develop effective remedies. After thorough search and rigorous reading of numerous research papers related to genomic structures, ligand receptor interaction, mode of entry in human cell, replication and transcription, RTC formation, assembly of viral parts, maturation and release of coronona virus, particularly human coronona viruses, such as SARS –CoV, MERS –CoV, SARs –CoV-2, we have attempted to highlight some meaningful points that may be utilized for the production of drugs or vaccines and developed drugs must target the conserved zone of virus–

i) A lot of compounds like polypeptides, peptides, antibiotics etc are able to interact with receptor ACE2 and inactivate the latter (168,169). We may use some of them against SARS-CoV2 because they block the S-protein-binding site, or change the conformation of ACE2 which is not suitable for binding or fusion. So, now we can remember the words of Li et al. (49)- “if SARS returns as a threat to human health, these studies may contribute to its control”(49). The chloroquine or hydroxychloroquine have been considered by *in vitro* trial in Vero E6 cell line as effective drug against Covid-19 as it disturbs the glycosylation of a virus cell surface receptor, ACE2 on Vero E6 cell line (170). So, hydroxy chloroquine sulfate (HCQS) has been approved as one of the important drugs for the treatment of severe SARSCoV-2 infections (171).

ii) The virus utilizes TMPRSS2, host serine protease to prime S protein as a result it facilitates the fusion of viral and cellular membranes and leads to the entry of virus into the cell. Here, scientists may develop or search already existing any serine protease inhibitor for inhibiting the viral entry (172,173). We may mention here serine protease inhibitors like “camostat mesylate” and K11777, that inhibit TMPRSS2 and partially arrest SARS-CoV infection of lung epithelial cells. (174,175). Hoffmann et al. (62) recently reported that TMPRSS2 is hindered by a protease inhibitor which has been clinically tested. Some peptide inhibitors have been formulated which are active to destabilize HR regions of S2 and they have exhibited their effectiveness in both *in vitro* and *in vivo* trial (176,177).

iii) The increase of pH of endosome may be one target for drug development. Chloroquine when applied against Chikungunya virus *in vitro* assay, was successful to check the viral infection as this drug makes an alkaline environment in the cell (178). A recent work revealed that both chloroquine and the antiviral drug remdesivir are effective against SARS-CoV-2 *in vitro* and



advocated that these drugs might be applied in human patients against this virus(171). Adedeji et al.(179) screened some of 14,000 compounds of the Maybridge Hit Finder small-molecule library. These compounds are at par with Lipinski's rule of five(180). Out of these compounds, they found three compounds-SSAA09E2, SSAA09E1 and SSAA09E30 arrest SARS-CoV entry by the following principles: (i) checking of early SARS-S-ACE2 interactions (ii) inactivation of cathepsin-L, and (iii) hindering fusion respectively.

iv) Studies have shown that M<sub>pro</sub> of different coronaviruses are highly conserved in terms of both sequences and 3D structures. The main protease of SARS-CoV-2 and SARS-CoV are almost similar (96.1%) (59). These features, together with its functional importance, have considered M<sub>pro</sub> an important object for the design of anti-coronaviral drugs(172). Very recently, Li et al.(59) selected the structure of SARS-CoV-2 main protease as a homologous target for drug molecule screening on basis of bioinformatics analysis, and suggested that out of 8,000, four drugs such as Prulifloxacin (fluoroquinolone antibiotic), Tegobuvir, Bictegravir and Nelfinavir (anti-HIV drugs) exhibited maximum binding conformations with the main protease of virus(181). Similarly, Li et al.(182) shorted out available drugs which may be potential inactivator for SARS-CoV-2 M<sub>pro</sub> protease and these drugs have higher mutation tolerance than widely used drugs lopinavir. In this respect, we may mention the review work of Ghosh et al (183) which highlighted the repurposing of drugs against SARS-CoV2, SARS-CoV and MERS-CoV.

v) The functional domains exist in the replicase polyproteins are conserved in all CoVs, so, they must be good targets for anti viral drugs or vaccine(90). According to Amici et al.(184). Indothethacin showed antiviral activity against SARS-CoV blocking viral synthesis in early stage in Vero 6E cell line.

vi) Coronavirus RNA is synthesized in a RTC or DVC. Drugs may be developed for changing its micro-environment or degradation of this DMV(185).

vii) The -1 ribosomal frameshifting is very essential for SARS-CoV to synthesize the replication-transcription complex and the 1a/1b ribosomal frameshift signal is conserved (111). So it can be attractive approach that this mRNA structure is an important target for drug development. ANXA2 has been suggested as an antiviral regulator which specifically binds to the frameshift signal as frameshift signal binding compounds(186) and other compounds could be applied targeting this conserved frameshift signal.

viii) During RNA replication of SARS-CoV, nucleic acid unwinding by the viral helicase is a critical point, so this point may be targeted for inhibition of RNA replication. Adedeji et al.(187) applied some inhibitors to this point and become successful to block RNA replication.

ix) Here discontinuous RNA transcription may be a good target for drug design in such a way that a compound may be applied to inhibit this RNA transcription. Similarly, nucleoside inhibitors may be applied to arrest SARS-CoV-2 replication specifically without damaging the cell. Moreover, TRSs have immense role in regulating RNA synthesis and they have conserved sequences (CS), which may be targeted for anti-SARS-CoV2 drug development.

x) Chloroquine drug has a power for inhibiting proteolytic processing of the M protein and affects construction of virus and budding. Besides, this drug changing pH of cell can damage the viral protein (188) and interferes the recognition of viral antigen by dendritic cells, which operates by a Toll-like receptor-dependent pathway that needs changes of pH of endosomes to low acidic (189). Chloroquine may be good drug against COVID-19 (171).

xi) Viral proteins such as nsp12(RdRp), N protein, nsp14 etc have conserved zones, as these three proteins have vital role in RNA synthesis, drug designer may target their conserved sequences. As for example nsp14 of corona viruses is highly susceptible to ribavirin and 5-fluorouracil agents (116).

xii) Blocking of entry of viral proteins (N, nsp3, nsp 6, etc) into nucleus may be a good strategy to arrest the viral synthesis and to restore our immune system which became antagonized by viral proteins. Ivermectin, an FDA-approved anti-parasitic and broad spectrum anti-viral drug, has been found as an inhibitor of SARS-CoV-2 isolate Australia/VIC01/2020) in cell line (Vero/hSLAM). It reduced nearly 5000-fold viral RNA synthesis at 48 h. It binds with the carrier (IMP $\alpha$ / $\beta$ 1) of viral protein which carries the protein inside the nucleus. So, binding to the viral protein becomes inhibited and preventing it from entering the nucleus. As a result our antiviral responses become normal or more efficient antiviral response. If patients are treated by this drug early in infection, it reduces the viral load, checks severity of COVID-19 and arrests "person-person transmission" (190).

xiii) The newly replicated genomic RNA and structural proteins of SARS-CoV are assembled into virions in ERGIC or endoplasmic reticular or Golgi complex membrane. Virions are shed off from infected cells by exocytosis, so these two works (virion assembly and shredding) may be arrested by any inhibitor.

xiv) Proper and faithful vaccine development is most likely effective to control of SARS-CoV-2, learning from previous antiviral passive immunization (like RBD based vaccine) and active immunization (live attenuated vaccine). The convalescent sera after collecting from patients may be a target for passive immunization for treatment of SARS-CoV2. The antibodies against S1 from convalescent sera generated in SARS patients might have a chance to block SARS-CoV2 entry (62). Some research works exhibited that SARS-CoV S induced polyclonal antibodies in animal model like mice and patients effectively sequestered SARS-CoV-2 S-dependent entry into cells (58, 62). Here Wang et al. (191) reported a human monoclonal antibody (47D11) that neutralized SARS-CoV-2. It firmly attaches a conserved epitope on the spike receptor binding domain. The RBD in the S1 subunit of the SARS-CoV spike (S) protein has CND (critical neutralizing domain) that is capable of triggering effective neutralizing antibody response and cross-protection against different isolates of SARS-CoV2, so scientists target it for vaccine development. Zhu et al. (192), have compiled in their review paper all RBD based SARS vaccines that were the most effective and safest. Recently Fast and Chen (193) have applied bio-informatics tools to detect B-cell and T-cell epitopes in SARS-CoV-2 based on viral protein antigen presentation and antibody binding properties and reported that the spike protein of SARS-CoV2 bears a lot of both T-cell and B-cell epitopes. Neutralizing

antibodies can clear virus or protect an uninfected host that is exposed to the virus. We may recall here the words of Li et al (49) “ if SARS returns as a threat to human health, these studies may contribute to its control”. Hence, this antibody application may prevent and/or treat COVID-19 (191). A recombinant SARS-CoV-2 spike protein vaccine combined with other top epitopes could be a meaningful step for development of vaccine against SARS-CoV-2. The ray of hope is coming from some laboratories for drug and vaccine production, examples may be cited here as the mRNA base vaccine (mRNA -1273) under a phase 1 clinical trial against SARS-Co-2 since 25th February, 2020 with 1st dose already applied on human on 16th March at Kaiser Permanente Washington Health Research Institute (KPWHRI) in Seattle under guidance of NIH(national Institution for Health, USA(194). Other numerous vaccines developed are being undertrial in different institutions in different countries like Ad5-nCoV in China, ChAdOx1 nCoV-19 in Oxford University, INO-4800 in Philadelphia and Kansas City, BNT162 a1, b1, b2, and c2 in Germany, bacTRL-Spike in Canada, Covid-19 S-Trimer & a nanoparticle vaccine SARS-CoV-2 rS in Australia, and BCG trial in Brasil, Germany, France, Denmark, Hungary and South Africa (195), etc. Previous antiviral vaccine like live attenuated vaccines was applied to arrest dangerous diseases caused by avian and porcine CoVs. By gaining experiences of previous vaccines, a live attenuated vaccine could be invented against COVID-19 or SARS-CoV-2, because we can multiply this virus in high titers in Vero 6E or other cell lines.

## CONCLUSION

In conclusion, now the SARS-CoV-2 or COVID-19 pandemic appears to be out of control in some countries. Although the development of drugs and vaccine are very urgent, we have to remember it needs time and patience to discover the appropriate drug. In such a moment new quick tests to identify SARS-CoV- 2 patients at the earliest stages of disease are also necessary as these tests will lead quarantine and isolation procedures to arrest the transmission of this disease. We are hopeful that the elaborate discussion of virion structure , molecular mechanism of propagation and clues for drug or vaccine development embedded in this article will help us for quick proper and effective drug or vaccine discovery. The whole world is anxiously waiting to defeat the COVID-19 pandemic and win the game of survival by discovering drugs and vaccines against SARS-CoV-2.

## Contribution of author

As SKG, is sole author, all works like data/literature collection, writing Ms, fig drawn, editing etc are done by myself.

**ACKNOWLEDGEMENT** Author is grateful to Principal, RKMVC College, Rahara. Author declares no competing financial and nonfinancial interest.

**Funding Source** The funding source is Department of Science & Technology and Biotechnology(DSTBT), GoW B, Sanction No.840(sac) /ST/P/S&T/1G-11/2015 dated 11.01.2016. The funder has no role in decision to submit it for publication.

## REFERENCES

- 1 Ge, X.Y. et al. Isolation and characterization of a bat SARS -like coronavirus that uses the ACE2 receptor. *Nature* **503**(7477):535 -538(2013).
- 2 Chen, Y, Guo, D. Molecular mechanisms of coronavirus RNA capping and methylation. *Viol. Sin.*31(1): 3 -11(2016).
- 3 Butler, D. SARS veterans tackle coronavirus. *Nature* **490**:20(2012).
- 4 Peiris, J.S. et al. Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet* **361**:1319-1325(2003).
- 5 Zaki, A.M. et al. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N Engl. J. Med.* **367**:1814-1820(2012).
- 6 WHO. Novel coronavirus infection - update (Middle East respiratory syndrome- coronavirus). Available online: [http://www.who.int/csr/don/2013\\_05\\_29\\_ncov/en/index.html](http://www.who.int/csr/don/2013_05_29_ncov/en/index.html). (accessed May 30, 2013).
- 7 Rahman, A, Sarkar, A. Risk factors for fatal middle east respiratory syndrome coronavirus infections in Saudi Arabia: analysis of the WHO Line List, 2013– 2018. *Am. J. Public Health* **109**(9):1288–93(2019)
- 8 Fehr, A.R., Perlman S. Coronaviruses: an overview of their replication and pathogenesis. *Methods Mol. Biol.* **1282**: 1 -23(2015).
- 9 Su, S. et al. Epidemiology, genetic recombination, and pathogenesis of coronaviruses. *Trends Microbiol.***24**(6):490 -502(2016).
- 10 Jiang, S., Lu L., Du, L. Development of SARS vaccines and therapeutics is still needed (editorial). *Future Virol.* **8**: 1-2(2013).
- 11 WHO Announces COVID-19 Outbreak a pandemic. <http://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/news/news/2020/3/who-announces-covid-19-outbreak-a-pandemic> (accessed on 21 March 2020).
- 12 <http://virological.org/>
- 13 <https://www.gisaid.org/>
- 14 Wu, F, S. Zhao, B. Yu, et al., A new coronavirus associated with human respiratory disease in China. *Nature* 2020, <https://doi.org/10.1038/s41586-020-2008-3>.
- 15 Lu, R, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding *Lancet* 2020; **395**(10224): 565-574.
- 16 ChenY, Liu Q, Guo D . Emerging coronaviruses: Genome structure, replication, and pathogenesis. *J Med Virol.* 2020; **92**:418–423.
- 17 Zhou, P, Yang X-L, Wang, X-G, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 2020 <https://doi.org/10.1038/s41586-020-2012-7>, February .
- 18 Yuan, M, Yin W, Tao Z, Tan W, Hu Y (2020) Association of radiologic findings with mortality of patients infected with 2019 novel coronavirus in Wuhan, China. *PLoS One* 2020; **15**(3): e0230548. <https://doi.org/10.1371/journal.pone.0230548>.
- 19 Wölfel, R. et al. SARS-CoV-2 Replication at Different Body Sites. *Nature* 2020; **581**: 465–469.
- 20 WHO <http://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/news/news/2020>, 4<sup>th</sup> April ,2020
- 21 Guan WJ, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. *N Engl J Med* 2020; **382**(18). 1708-1720.

- 22 Chu H., et al, Comparative replication and immune activation profiles of SARS-CoV-2 and SARS-CoV in human lungs: an *ex vivo* study with implications for the pathogenesis of COVID-19. *Clin Infect D* 2020 DOI: 10.1093/cid/ciaa410, April.
- 23 Chan JF, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet* 2020; **395**: 514–23.
- 24 To KK, et al. Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. *Lancet Infect Dis* 2020; **20**: 565–74.
- 25 Chan JF, et al. Simulation of the clinical and pathologic manifestations of Coronavirus Disease 2019 (COVID-19) in golden Syrian hamst model: implications for disease pathogenesis and transmissibility. *Clin Infect D* 2020. doi: 10.1093/cid/ciaa325
- 26 Huang C, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020; **395**:497–506.
- 27 CDCP Centers for disease control and prevention <https://www.cdc.gov/coronavirus/2019-ncov/> January 28, 2020.
- 28 Phan LT, et al. Importation and human-to-human transmission of a novel coronavirus in Vietnam. *N Engl J Med*. DOI: 10.1056/NEJMc2001272
- 29 WHO, <https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200610-covid-19-sitrep-142> , 10 June,2020
- 30 Leung C 2020 Estimating the distribution of the incubation period of 2019 novel coronavirus (COVID-19)infection between travelers to Hubei, China and non-travelers med Rxiv preprint doi: <https://doi.org/10.1101/2020.02.13.20022822>
- 31 Li Q, et al (2020) Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. *N Engl J Med*. doi:10.1056/NEJMoa2001316
- 32 Backer JA, Klinkenberg D, Wallinga J. Incubation period of 2019 novel coronavirus (COVID-19) infections among travellers from Wuhan, China. *Eurosurveillance*. 6 Feb 2020. doi:10.2807/1560-7917.ES.2020.25.5.2000062.
- 33 CDC-Centers for disease control and prevention 4/4/2020 <https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/social-distancing.html>),
- 34 Barcena M, et al. Cryo-electron tomography of mouse hepatitis virus: insights into the structure of the coronavirus. *Proc Natl Acad Sci ,USA*, 2009 **106**:582–587
- 35 Neuman BW, et al . A structural analysis of M protein in coronavirus assembly and morphology. *J Struct Biol* 2011; **174**:11–22.
- 36 Rota PA, et al. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science* 2003; **300**:1394-1399.
- 37 Bosch BJ, et al. The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. *J Virol* 2003; **77**:8801–8811
- 38 Collins AR, et al (1982) Monoclonal antibodies to murine hepatitis virus-4 (strain JHM) define the viral glycoprotein responsible for attachment and cell–cell fusion. *Virology* 1982; **119**: 358–371.
- 39 Delmas B, Laude H . Assembly of coronavirus spike protein into trimers and its role in epitope expression. *J Virol* 1990 ; **64**:5367–5375
- 40 Armstrong J, Niemann H, Smeekens S et al . Sequence and topology of a model intracellular membrane protein, E1 glycoprotein, from a coronavirus. *Nature* 1984; **308**:751–752
- 41 Nal B, et al. Differential maturation and subcellular localization of severe acute respiratory syndrome coronavirus surface proteins S, M and E. *J Gen Virol* 2005; **86**:1423–1434.



- 42 Nieto-Torres JL, et al. Severe acute respiratory syndrome coronavirus envelope protein ion channel activity promotes virus fitness and pathogenesis. *PLoS Pathog* 2014; 10:e1004077. doi:10.1371/journal.ppat.1004077
- 43 Chang C, et al. Modular organization of SARS coronavirus nucleocapsid protein. *J Biomed Sci.* 2006; **13**(1): 59 -72.
- 44 Sturman LS, Holmes KV, Behnke J. Isolation of coronavirus envelope glycoproteins and interaction with the viral nucleocapsid. *J Virol* 1980; **33**:449–462.
- 45 Cui L, et al. The nucleocapsid protein of coronaviruses acts as a viral suppressor of RNA silencing in mammalian cells. *J Virol* 2015; **89**(17):9029 -9043.
- 46 Cui J, Li F, Shi ZL. Origin and evolution of pathogenic Coronaviruses. *Nat Rev Microbiol.* 2019; **17**(3):181–92.
- 47 .Chen J, Jiang Q, Xia X, Liu K, Yu Z, Wanyu Tao , Gong W ., Han J-D J. Individual variation of the SARS-CoV2 receptor ACE2 gene expression and regulation. 2020 (www.preprints.org) 12 March .Preprint.
- 48 Hussain S, et al. Identification of novel subgenomic RNAs and noncanonical transcription initiation signals of severe acute respiratory syndrome coronavirus. *J Virol* 2005; **79**(9):5288 -5295.
- 49 Li W, et al., Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus, *Nature* 2003; **426**: 450e454, <https://doi.org/10.1038/nature02145>.
- 50 Tresnan DB, Levis R, Holmes KV. 1996. Feline aminopeptidase N serves as a receptor for feline, canine, porcine, and human coronaviruses in sero group I. *J Virol* 1996; **70**: 8669–8674.
- 51 Hofmann H, Pyrc K, van der Hoek L, Geier M, Berkhout B, Pöhlmann S. Human coronavirus NL63 employs the severe acute respiratory syndrome coronavirus receptor for cellular entry. *Proc Natl Acad Sci* 2005; **102**:7988–7993.
- 52 Saeki K, Ohtsuka N, Taguchi F. Identification of spike protein residues of murine coronavirus responsible for receptor-binding activity by use of soluble receptor-resistant mutants. *J Virol* 1997; **71**: 9024 –9031.
- 53 Mou H, Raj VS, van Kuppeveld FJ, Rottier PJ, Haagmans BL, Bosch BJ.. The receptor binding domain of the new Middle East respiratory syndrome coronavirus maps to a 231-residue region in the spike protein that efficiently elicits neutralizing antibodies. *J Virol* 2013; **87**: 9379 –9383.
- 54 Zhao Z, et al. Coronavirus replication does not require the autophagy gene ATG5. *Autophagy* **3**: 581-585(2007).
- 55 Lu, G. et al. Molecular basis of binding between novel human coronavirus MERS-CoV and its receptor CD26. *Nature* **500**: 227–231(2013).
- 56 Wang, Q, et al. Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2, *Cell* 2020, <https://doi.org/10.1016/j.cell.2020.03.045>
- 57 Wrapp, D. et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*; **367**: 1260–1263(2020).
- 58 Walls, A. C. et al. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* 2020; S0092-8674(20)30262-2. <https://doi.org/10.1016/j.cell.2020.02.058>
- 59 Li, X., Genga M, Penga Y, Menga L , Lua, S. Molecular immune pathogenesis and diagnosis of COVID-19. *J Pharmaceutical Analysis* **10**: 102-108(2020).
- 60 Wan, Y., Shang, J., Graham, R., Baric, R.S., Li, F. Receptor recognition by the novel coronavirus from Wuhan: an analysis based on decade-long structural studies of SARS coronavirus. *J Virol* 2020; **94**:e00127-20. <https://doi.org/10.1128/JVI.00127-20>
- 61 Letko M, Marzi A, Munster V. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat. Microbiol* 2020; <https://doi.org/10.1038/s41564-020-0688-y>
- 62 Hoffmann M et al. ACE2 is the SARS-CoV-2 Receptor Required for Cell Entry *Cell* 2020 Mar5.
- 63 Lu G, Wang Q, Gao GF Bat-to-human: spike features determining ‘host jump’ of coronaviruses SARS-CoV, MERS-CoV, and beyond. *Trends Microbiol* 2015; **23**: 468–478.

- 64 Li F. Structural analysis of major species barriers between humans and palm civets for severe acute respiratory syndrome coronavirus infections. *J Virol* 2008; **82**:6984 – 6991.
- 65 Wu KL, Peng GQ, Wilken M, Geraghty RJ, Li F. Mechanisms of host receptor adaptation by severe acute respiratory syndrome coronavirus. *J Biol Chem* 2012; **287**:8904 – 8911.
- 66 Lan J, Ge J, Yu J, et al.. Structure of the SARS-CoV-2 spike receptor binding domain bound to the ACE2 receptor *Nature* 2020; <https://doi.org/10.1038/s41586-020-2180-5>.
- 67 Donoghue M, et al. A Novel Angiotensin-Converting Enzyme-Related Carboxypeptidase (ACE2) Converts Angiotensin I to Angiotensin 1-9 *Circ Res* 2000; **87**(5), E1–9
- 68 Naskalska A, Dabrowska A, Szczepanski A, Milewska A, Jasik KP, Pyrc K. Membrane protein of human coronavirus NL63 is responsible for interaction with the adhesion receptor. *J Virol* 2019;**93**:e00355-19. <https://doi.org/10.1128/JVI.00355-19>.
- 69 Haun G, Keppler OT, Bock CT, Herrmann M, Zentgraf H, And Pawlita M The Cell Surface Receptor Is a Major Determinant Restricting the Host Range of the B-Lymphotropic Papovavirus *J Virol* 1993; **67** (12):.7482-7492.
- 70 Moore JP, McKeating JA, Norton WA, Sattentau QJ. Direct measurement of soluble CD4 binding to human immunodeficiency virus type 1 virions: gp120 dissociation and its implications for virus-cell binding and fusion reactions and their neutralization by soluble CD4. *J Virol* 1991; **65**: 1133–1140.
- 71 Kuba K, et al A crucial role of angiotensin converting enzyme 2 (ACE2) in SARS coronavirus–induced lung injury *Nature Medicine* 2005; **11**(8):.875- 879.
- 72 Hamming I., Timens W, Bulthuis ML, Lely AT, Navis G J, van Goor H. . Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. *J Pathol* 2004; **203**:631–637.
- 73 Harmer D, Gilbert M, Borman R, Clark KL. Quantitative mRNA expression profiling of ACE 2, a novel homologue of angiotensin converting enzyme. *FEBS Lett* 2002; **532**:107–110.
- 74 Jia HP, Look DC, Shi L, et al. ACE2 Receptor Expression and Severe Acute Respiratory SyndromeCoronavirus Infection Depend on Differentiation of Human Airway Epithelia. *J Virol* 2005; **79**: 14614–14621.
- 75 Wang PH, Cheng Y. Increasing host cellular receptor—angiotensin-converting enzyme 2 (ACE2) expression by coronavirus may facilitate 2019-nCoV infection. *BioRxiv* 2020 Feb 27. doi:10.1101/2020.02.24.963348. Preprint.
- 76 Zumla A, Chan JF, Azhar EI, Hui DS, Yuen KY. Coronaviruses – drug discovery and therapeutic options. *Nat Rev Drug Discov* 2016; **15**: 327-47.
- 77 Simmons G, Reeves JD, Rennekamp AJ, Amberg SM, Piefer AJ, Bates P. Characterization of severe acute respiratory syndrome associated coronavirus (SARS-CoV) spike glycoprotein-mediated viral entry. *Proc Natl Acad Sci USA*; 2004; **101**:4240-4245.
- 78 Liu S, et al. Interaction between heptad repeat 1 and 2 regions in spike protein of SARS-associated coronavirus: implications for virus fusogenic mechanism and identification of fusion inhibitors. *Lancet* 2004; **363**: 938-47.
- 79 Yang ZY, et al. pH-dependent entry of severe acute respiratory syndrome coronavirus is mediated by the spike glycoprotein and enhanced by dendritic cell transfer through DC-SIGN. *J Virol* 2004; **78**:5642-5650.
- 80 Huang I C,. et al. SARS coronavirus, but not human coronavirus NL63, utilizes cathepsin L to infect ACE2-expressing cells. *J Biol Chem* 2006; **281**:3198–3203.
- 81 Inoue Y, et al. Clathrin dependent entry of severe acute respiratory syndrome coronavirus into target cells expressing ACE2 with the cytoplasmic tail deleted. *J Virol* 2007; **81**:8722–8729.
- 82 Wang H, et al. SARS coronavirus entry into host cells through a novel clathrin- and caveolae-independent endocytic pathway *Cell Research* 2008; **18**:290-301.
- 83 Wang LF, Shi Z, Zhang S, Field H, Daszak P, Eaton B. Review of bats and SARS. *Emerg Infect Dis.* 2006; **12**(12):1834 -1840.

- 84 Oldstone MB, et al.. Measles virus infection in a transgenic model: virus-induced immunosuppression and central nervous system disease. *Cell*. 1999;**98**:629–640.
- 85 Marsh M, Helenius A. Virus entry: open sesame. *Cell* 2006;**124**:729–740 , 29.
- 86 Pelkmans L, Helenius A. Insider information: what viruses tell us about endocytosis. *Curr Opin Cell Biol* 2003; **15**:414–422.
- 87 Burkard C, et al. (2014) Coronavirus Cell Entry Occurs through the Endo-/Lysosomal Pathway in a Proteolysis-Dependent Manner. *PLoS Pathog* 2014;**10**(11): e1004502. doi:10.1371/journal.ppat.1004502.
- 88 Cottam EM, et al. Coronavirus nsp6 proteins generate autophagosomes from the endoplasmic reticulum via an omegasome intermediate. *Autophagy*. 2011;**7**: 1335–47.
- 89 Namy O, Moran SJ, Stuart DL, Gilbert RJ, Brierley L.. A mechanical explanation of RNA pseudoknot function in programmed ribosomal frameshifting. *Nature* 2006; **441**:244–247.
- 90 Ziebuhr J, Snijder EJ, Gorbalenya AE. Virus-encoded proteinases and proteolytic processing in the Nidovirales. *J Gen Virol* 2000; **81**:853–879.
- 91 Denison MR, Graham RL, Donaldson EF, Eckerle LD, Baric RS. Coronaviruses . An RNA proofreading machine regulates replication fidelity and diversity. *RNA Biol* 2011; **8**:270–79.
- 92 Snijder EJ, et al. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. *J Mol Biol* 2003, **331**:991–1004.
- 93 Huang C et al. Alphacoronavirus transmissible gastroenteritis virus nsp1 protein suppresses protein translation in mammalian cells and in cell-free HeLa cell extracts but not in rabbit reticulocyte lysate. *J Virol* 2011; **85**:638–643.
- 94 Kamitani W, et al (2009) A two-pronged strategy to suppress host protein synthesis by SARS coronavirus Nsp1 protein. *Nat Struct Mol Biol* 2009;**16**:1134–1140.
- 95 Ratia K, et al. Severe acute respiratory syndrome coronavirus papain-like protease: structure of a viral deubiquitinating enzyme. *Proc. Natl Acad. Sci. USA* 2006;**103**, 5717–5722.
- 96 Clementz M A, Kanjanahaluethai A, O'Brien T E, Baker S C. Mutation in murine coronavirus replication protein nsp4 alters assembly of double membrane vesicles. *Virology* 2006; **375**, 118–129.
- 97 Gadlage MJ, et al Murine hepatitis virus nonstructural protein 4 regulates virus-induced membrane modifications and replication complex function. *J Virol* 2010; **284**: 280–290.
- 98 Egloff M P, et al. The severe acute respiratory syndrome-coronavirus replicative protein nsp9 is a single-stranded RNA-binding subunit unique in the RNA virus world. *Proc. Natl Acad. Sci. USA* 2004;**101** ,3792–3796.
- 99 Zhai, Y. et al. Insights into SARS-CoV transcription and replication from the structure of the nsp7-nsp8 hexadecamer. *Nature Struct. Mol. Biol.* **12**, 980–986(2005).
- 100 Deming, D. J., Graham, R. L., Denison, M. R. & Baric, R. S. Processing of open reading frame 1a replicase proteins nsp7 to nsp10 in murine hepatitis virus strain A59 replication. *J. Viro* 2007; **81**: 10280–10291.
- 101 Xu X, Liu Y, Weiss S, Arnold E, Sarafianos SG, Ding J. Molecular model of SARS coronavirus polymerase: implications for biochemical functions and drug design. *Nucleic Acids Res* 2003;**31**:7117–30.
- 102 Ivanov KA, et al. Major genetic marker of nidoviruses encodes a replicative endoribonuclease. *Proc Natl Acad Sci, U S A* 2004; **101**:12694–12699
- 103 Ivanov KA, et al. Multiple enzymatic activities associated with severe acute respiratory syndrome coronavirus helicase. *J Virol* 2004; **78**:5619–5632.,
- 104 Eckerle L D, Lu X, Sperry S M, Choi L, Denison M R. High fidelity of murine hepatitis virus replication is decreased in nsp14 exoribonuclease mutants. *J. Virol.* 2007; **81** , 12135–12144.
- 105 Chen Y, et al. Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase. *Proc. Natl Acad. Sci. USA* 2009; **106**(9):3484–9348
- 106 Bhardwaj K, et al (2006) RNA recognition and cleavage by the SARS coronavirus endoribonuclease. *J Mol Biol* 2006; **361**:243–256.

- 107 Decroly E, *et al.* Coronavirus nonstructural protein 16 is a cap-0 binding enzyme possessing (nucleoside-2'-O)-methyltransferase activity. *J. Virol.* 2008; **82**: 8071–8084.
- 108 Angeletti S, Benvenuto D, Bianchi M, Giovanetti M, Pascarella S, Ciccozzi M. COVID-19: The role of the nsp2 and nsp3 in its pathogenesis. *J Med Virol* 2020; 10.1002/jmv.25719. doi: 10.1002/jmv.25719.
- 109 Reil H, Kollmus H, Weidle Uh, Hauser Ha. Heptanucleotide Sequence Mediates Ribosomal Frameshifting in Mammalian Cells. *J Virol* 1993; **67**( 9): 5579-5584.
- 110 Dinman J, Icho DT, Wickner RB. 1991. A 1 ribosomal frameshift in a double-stranded RNA virus of yeast forms a gag-pol fusion protein. *Proc. Natl. Acad. Sci. USA* 1991; **88**:174–178.
- 111 Ramos FD, Carrasco M, Doyle T , Brierley I. Programmed–1 ribosomal frameshifting in the SARS coronavirus. *Biochem Soc Transact* ;2004; **32**, part 6.
- 112 Plant EP, Sims AC, Baric RS, Dinman JD, Taylor DR. Altering SARS coronavirus frameshift efficiency affects genomic and subgenomic RNA production. *Viruses*. 2013; **5**:279–94.
- 113 Ketteler R. On programmed ribosomal frameshifting: the alternative proteomes *Front Gen* 2012; **3**(1): :Article 242 <https://doi.org/10.3389/fgene.2012.00242>.
- 114 Brierley I and Ramos FJD Programmed ribosomal frameshifting in HIV-1 and the SARS–CoV. *Virus Res*. 2006; **119**(1): 29–42.
- 115 Brierley I., Digard P, Inglis SC. *Cell* (Cambridge, Mass.) 1889; **57**, 537–5478.
- 116 Smith EC, Blanc H, Vignuzzi M, Denison MR. Coronaviruses lacking exoribonuclease activity are susceptible to lethal mutagenesis: evidence for proofreading and potential therapeutics. *PLOS Pathog* 2013; **9**: e1003565.
- 117 Tu, C., Tzeng, T.-H. & Bruenn, J. A. Ribosomal movement impeded at a pseudoknot required for frameshifting. *Proc. Natl Acad. Sci. USA* 1992; **89**: 8636–8640.
- 118 Lopinski J D, Dinman J D , Bruenn, J A Kinetics of ribosomal pausing during programmed - 1 ribosomal frameshifting. *Mol Cell Biol* 2000; **20**: 1095–1103.
- 119 Frank J, Agrawal R K A. ratchet-like inter-subunit reorganization of the ribosome during translocation. *Nature* 2000; **406**: 318–322.
- 120 Spahn C M T, Gomez-Lorenzo MG, A Grassucci RA *et al.* Domain movements of elongation factor eEF2 and the eukaryotic 80S ribosome facilitate tRNA translocation. *EMBO J* 2004; **23**: 1008–1019.
- 121 Plant EP, Dinman JD. The role of programmed-1 ribosomal frameshifting in coronavirus propagation. *Front Biosci* 2008; **13**:4873–4881.
- 122 Snijder EJ, van der Meer Y, Zevenhoven-Dobbe J, Onderwater JJ, van der Meulen J, Koerten HK, Mommaas AM. Ultrastructure and origin of membrane vesicles associated with the severe acute respiratory syndrome coronavirus replication complex. *J Virol* 2006; **80**:5927–5940.
- 123 Miller S, Krijnse-Locker J. Modification of intracellular membrane structures for virus replication. *Nat Rev Microbiol* 2008; **6**:363–374.
- 124 Bechill J, Chen Z, Brewer JW, Baker SC. Coronavirus infection modulates the unfolded protein response and mediates sustained translational repression. *J Virol* 2008; **82**: 4492–4501.
- 125 Snijder E J, van Tol H, Roos N, Pedersen KW. Non-structural proteins 2 and 3 interact to modify host cell membranes during the formation of the arterivirus replication complex. *J Gen Virol*. 2001; **82**, 985–994.
- 126 Knoops K, Kikkert M, Worm SH, *et al.* SARS-coronavirus replication is supported by a reticulovesicular network of modified endoplasmic reticulum. *PLoS Biol* 2008; **6**:e226.
- 127 de Haan C A, Rottier P J. Molecular interactions in the assembly of coronaviruses. *Adv. Virus Res*. **64**,165–230.
- 128 van Hemert MJ, van den Worm SHE, Knoops K, *et al.* SARS-coronavirus replication /transcription complexes are membrane-enclosed and need a host factor for activity in vitro. *PLoS Pathog* 2008; **4**: e1000054. doi:10.1371/journal.ppat.1000054
- 129 Hagemeijer M, *et al.* Dynamics of Coronavirus Replication-Transcription Complexes† *J Virol* 2010; **84**(4): 2134–2149.



- 130 Sawicki D, Wang T, Sawicki S. The RNA structures engaged in replication and transcription of the A59 strain of mouse hepatitis virus. *J Gen Virol* 2001; **82**:385–396.
- 131 Sawicki SG, Sawicki DL, Siddell SG. A contemporary view of coronavirus transcription. *J Virol*. 2007; **81**(1):20–29.
- 132 Yokomori K, Banner L R, Lai M M. Coronavirus mRNA transcription: UV light transcriptional mapping studies suggest an early requirement for a genomic-length template. *J Virol* 1992; **66**, 4671–4678.
- 133 Lai MMC, Patton CD, Stohlman SA. Replication of mouse hepatitis virus: negative strand RNA and replicative form RNA are of genome length. *J Virol* 1982; **44**:487–492.
- 134 Baric R S, Stohlman S A, Lai M M C. Characterization of replicative intermediate RNA of mouse hepatitis virus: presence of leader RNA sequences on nascent chains. *J Virol* 1983; **48**: 633–640.
- 135 van der Most RG, Spaan WJM. Coronavirus replication, transcription, and RNA recombination In: Siddell, SG., editor. *The Coronaviridae*. New York: Plenum; 1995: 11–31.
- 136 Snijder E J, Meulenberg J J M. The molecular biology of arteriviruses. *J Gen Virol* 1998; **79**: 961–979.
- 137 van Marle G, Dobbe JC, Gultyaev AP, Luytjes W, Spaan WJM, Snijder EJ. Arterivirus discontinuous mRNA transcription is guided by base pairing between sense and antisense transcription-regulating sequences. *Proc Natl Acad Sci* 1999; **96**:12056–61.
- 138 Alonso S, Izeta A, Sola I, Enjuanes L. Transcription regulatory sequences and mRNA expression levels in the coronavirus transmissible gastroenteritis virus. *J Virol* 2002; **76**:1293–308.
- 139 Dufour D, Mateos-Gomez PA, Enjuanes L, Gallego J, Sola I. Structure and functional relevance of a transcription-regulating sequence involved in coronavirus discontinuous RNA synthesis. *J Virol* 2011; **85**:4963–73.
- 140 Sola I, Almazán F, Zúñiga S, Enjuanes L. Continuous and Discontinuous RNA Synthesis in Coronaviruses. *Annu Rev Virol* 2015; **2**(1): 265–288.
- 141 Sawicki SG, Sawicki DL. 2005. Coronavirus transcription: a perspective. *Curr Top Microbiol Immunol* 2005; **287**:31–55.
- 142 Lai MMC, Holmes KV 2001. Coronaviridae: the viruses and their replication, in: Knipe, D.M. and Howley, P.M. (Eds.), *Fields virology* 2001 Lippincott Williams & Wilkins, Philadelphia, pp. 1163–1179.
- 143 Sawicki SG, Sawicki DL. Coronavirus minus-strand RNA synthesis and effect of cycloheximide on coronavirus RNA synthesis. *J Virol* 1986; **57**:328–334.
- 144 Sawicki SG, Sawicki DL. Coronavirus transcription: subgenomic mouse hepatitis virus replicative intermediates function in RNA synthesis. *J Virol* 1990; **64**:1050–1056.
- 145 Sethna PB, Brian DA. Coronavirus genomic and subgenomic minus-strand RNAs copartition in membrane-protected replication complexes. *J Virol* 1997; **71**:7744–7749.
- 146 den Boon J A, Kleijnen M F, Spaan W J M, Snijder E J. Equine arteritis virus subgenomic mRNA synthesis: analysis of leader-body junctions and replicative-form RNAs. *J Virol* 1996; **70**: 4291–4298.
- 147 Sawicki S G, Sawicki DL. Coronaviruses use discontinuous extension for synthesis of subgenome-length negative strands. *Adv Exp Med Biol* 1995; **380**: 499–506.
- 148 Pasternak A O, Spaan WJM, Snijder EJ. Nidovirus transcription: how to make sense? *J Gen Virol* 2006; **87**:1403–1421.
- 149 Sola I, Mateos-Gomez PA, Almazan F, Zúñiga S, Enjuanes L. RNA-RNA and RNA-protein interactions in coronavirus replication and transcription. *RNA Biol* 2011; **8**:237–48.
- 150 Masters PS. Genomic cis-acting elements in coronavirus RNA replication. In: Thiel, V., editor. *Coronaviruses: Molecular and Cellular Biology*. Norfolk, UK: Caister Academic; 2007. p. 65–80.
- 151 Madhugiri R, Fricke M, Marz M, Ziebuhr J. RNA structure analysis of alphacoronavirus terminal genome regions. *Virus Res* 2014; **194**:76–89.



- 152 Liu P, Li L, Keane SC, Yang D, Leibowitz JL, Giedroc DP. Mouse hepatitis virus stem-loop 2 adopts a uYNMG(U)a-like tetraloop structure that is highly functionally tolerant of base substitutions. *J Virol* 2009; **83**:12084–93.
- 153 Zust R, Miller TB, Goebel SJ, Thiel V, Masters PS. Genetic interactions between an essential 3'cis-acting RNA pseudoknot, replicase gene products, and the extreme 3' end of the mousecoronavirus genome. *J Virol* 2008; **82**:1214–28.
- 154 Sola I, Moreno JL, Zuñiga S, Alonso S, Enjuanes L. Role of nucleotides immediately flanking the transcription-regulating sequence core in coronavirus subgenomic mRNA synthesis. *J Virol* 2005; **79**:2506–16
- 155 Enjuanes L, Almazan F, Sola I, Zuñiga S. Biochemical aspects of coronavirus replication and virushost interaction. *Annu Rev Microbiol* 2006; **60**:211–30.
- 156 Moreno JL, Zuñiga S, Enjuanes L, Sola I. Identification of a coronavirus transcription enhancer. *J Virol* 2008; **82**:3882–93.
- 157 Mateos-Gomez PA, Morales L, Zuñiga S, Enjuanes L, Sola I. Long-distance RNA-RNA interactions in the coronavirus genome form high-order structures promoting discontinuous RNA synthesis during transcription. *J Virol* 2013; **87**:177–86.
- 158 Wu CH, Chen PJ, Yeh SH. Nucleocapsid phosphorylation and RNA helicase DDX1 recruitment enables coronavirus transition from discontinuous to continuous transcription. *Cell Host Microbe* 2014; **16**:462–72.
- 159 Plant EP, Sims AC, Baric RS, Dinman JD, Taylor DR. Altering SARS coronavirus frameshift efficiency affects genomic and subgenomic RNA production. *Viruses* 2013; **5**:279–94.
- 160 Surjit M, Liu B, Chow VT, Lal SK. The nucleocapsid protein of severe acute respiratory syndromecoronavirus inhibits the activity of cyclin-cyclin-dependent kinase complex and blocks Sphase progression in mammalian cells. *J Biol Chem* 2006; **281**:10669–81.
- 161 Cawood R, Harrison SM, Dove BK, Reed ML, Hiscox JA. Cell cycle dependent nucleolar localization of the coronavirus nucleocapsid protein. *Cell Cycle* 2007; **6**:863–67.
- 162 Hussain S, Gallagher T. SARS-coronavirus protein 6 conformations required to impede protein import into the nucleus. *Virus Res* 2010; **153**:299–304.
- 163 Frieman M, Yount B, Heise M, Kopecky-Bromberg SA, Palese P, Baric RS. Severe acute respiratory syndrome coronavirus ORF6 antagonizes STAT1 function by sequestering nuclear import factors on the rough endoplasmic reticulum/Golgi membrane. *J Virol* 2007; **81**:9812–24.
- 164 Stohlman SA, Baric RS, Nelson GN et al. Specific interaction between coronavirus leader RNA and nucleocapsid protein. *J Virol* 1988; **62**:4288–4295.
- 165 Molenkamp R, Spaan WJ Identification of a specific interaction between the coronavirus mouse hepatitis virus A59 nucleocapsid protein and packaging signal. *Virology* 1997; **239**:78–86.
- 166 Kuo L, Masters PS Functional analysis of the murine coronavirus genomic RNA packaging signal. *J Virol* 2013; **87**: 5182–5192.
- 167 Hurst KR, Koetzner CA, Masters PS Characterization of a critical interaction between the coronavirus nucleocapsid protein and nonstructural protein 3 of the viral replicase-transcriptase complex. *J Virol* 2013; **87**:9159–9172.
- 168 Dales N A, et al. Substrate-based design of the first class of angiotensin-converting enzyme-related carboxypeptidase (ACE2) inhibitors. *J Am Chem Soc* 2002; **124**, 11852–11853.
- 169 Widjaja I. et al. Towards a solution to MERS: protective human monoclonal antibodies targeting different domains and functions of the MERS-coronavirus spike glycoprotein. *Emerg Microbes Infect* 2019; **8**, 516-530.
- 170 Vincent MJ, et al. Chloroquine is a potent inhibitor of SARS coronavirus infection and spread. *Virol J* 2005; **2**:69.
- 171 Gao J, Tian Z, Yang X. Breakthrough: chloroquine phosphate has shown apparent efficacy in treatment of COVID-19 associated pneumonia in clinical studies. *Biosci Trends* Feb 2020 doi:10.5582/bst.2020.01047 [Epub ahead of print].

- 172 Xia B, Kang X . Activation and maturation of SARS-CoV main protease. *Protein Cell* 2011; **2**:282-290.
- 173 Matsuyama S, Taguchi F. 2002. Receptor-induced conformational changes of murine coronavirus spike protein. *J Virol* 2002;**76**:11819–11826.
- 174 Kawase M, Shirato K, van der Hoek L, Taguchi F, Matsuyama S. . Simultaneous Treatment of Human Bronchial Epithelial Cells with Serine and Cysteine Protease Inhibitors Prevents Severe Acute Respiratory Syndrome Coronavirus Entry. *J Virol* 2012; **86**:6537–6545.
- 175 Zhou Y, Vedantham P, Lu K, Agudelo J, et al . Protease inhibitors targeting coronavirus and filovirus entry. *Antiviral Res* 2015 ; **116**:76–84.
- 176 Gao J, et al. Structure of the fusion core and inhibition of fusion by a heptad repeat peptide derived from the S protein of Middle East respiratory syndrome coronavirus. *J Virol* 2013; **87**, 13134–13140.
- 177 Channappanavar, R. et al. Protective effect of intranasal regimens containing peptidic middle east respiratory syndrome coronavirus fusion inhibitor against MERS-CoV infection. *J. Infect Dis* 2015; **212**, 1894–1903.
- 178 Khan M, Santhosh SR, Tiwari M, Lakshmana Rao PV, Parida M. Assessment of in vitro prophylactic and therapeutic efficacy of chloroquine against Chikungunya virus in Vero cells. *J Med Virol* 2010; **82**:817–24.
- 179 Adediji AO, Severson W, Jonsson C , Singh K, Weiss SR, Sarafianosa SG . Novel Inhibitors of Severe Acute Respiratory Syndrome Coronavirus Entry That Act by Three Distinct Mechanisms. *J Virol* 2013; **87** ( 14) : 8017–8028
- 180 Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* 2001; **46**:3–26.
- 181 Davis DA, et al. Activity of human immunodeficiency virus type 1 protease inhibitors against the initial autocleavage in Gag-Pol polypeptide processing. *Antimicrob Agents Chemother* 2012; **56**: 3620-3628.
- 182 Li Y., Zhang J, Wang N, Li H, Shi Y, Guo G , Liu K, Zeng H, Zou Q. Therapeutic Drugs Targeting 2019-nCoV Main Protease by High-Throughput Screening. *BioRxiv* doi: <https://doi.org/10.1101/2020.01.28.922922> preprint.
- 183 Ghosh AK, Brindisi M, Shahabi D, E Chapman ME, Mesecar AD. Drug Development and Medicinal Chemistry Efforts Toward SARS-Coronavirus and Covid-19 Therapeutics. *Chem Med Chem* 2020; **15**(11):907-932.
- 184 Amici C, et al Indomethacin has a potent antiviral activity against SARS coronavirus *Antiviral therapy* 2006; **11**:1021-1030.
- 185 Gosert R, Kanjanahaluethai A, Egger D, Bienz K, Baker SC. RNA replication of mouse hepatitis virus takes place at double membrane vesicles. *J Virol* 2002; **76**:3697–3708.
- 186 Dinman J D . Mechanisms and implications of programmed translational frameshifting. *Wiley Interdiscip Rev RNA*. **3**, 661–673(2012).
- 187 Adediji AO, Singh K, Calcaterra NE, Dediego ML, Enjuanes L, Weiss S, Sarafianos SG. Severe acute respiratory virus syndrome coronavirus replication inhibitor that interferes with the nucleic acid unwinding of the viral helicase. *Antimicrob Agents Chemother* **56**: 4718 – 4728(2012).
- 188 Randolph VB, Winkler G, Stollar V. Acidotropic amines inhibit proteolytic processing of flavivirus prM protein. *Virology*; **174**:450–8(1990).
- 189 Diebold SS, Kaisho T, Hemmi H, Akira S, Reis e Sousa C. Innate antiviral responses by means of TLR7-mediated recognition of single-stranded RNA. *Science* 2004;**303**:1529–31 .
- 190 Caly L, Druce JD, Catton MG, Jans DA, Wagstaff KM. The FDA approved Drug Ivermectin inhibits the replication of SARS-CoV-2 in vitro, *Antiviral Research*, 2020; <https://doi.org/10.1016/j.antiviral.2020.104787>.
- 191 Wang, C. et al. A human monoclonal antibody blocking SARS-CoV-2 infection. *Bio Rxiv* 2020; doi: <https://doi.org/10.1101/2020.03.11.987958> preprint.

192 Zhu, X., Liu, Q., Du, L., Lu, L., Jiang, S. Receptor-binding domain as a target for developing SARS vaccines *J Thorac Dis* 5(S2): S142-S148(2013).

193 Fast, E., Chen. B. Potential T-cell and B-cell Epitopes of 2019-nCoV. *BioRxiv* 2020; <https://doi.org/10.1101/2020.02.19.955484> Preprint.

194 <https://www.nih.gov/news-events/news-releases/nih-clinical-trial-investigational-vaccine-covid-19-begins>.

195 Bastian, H. First in Human: Covid-19 Vaccines & Tales of Phase 1 Clinical Trials Past, *PLOS Blog* 16<sup>th</sup> March (2020).

Figures with legend

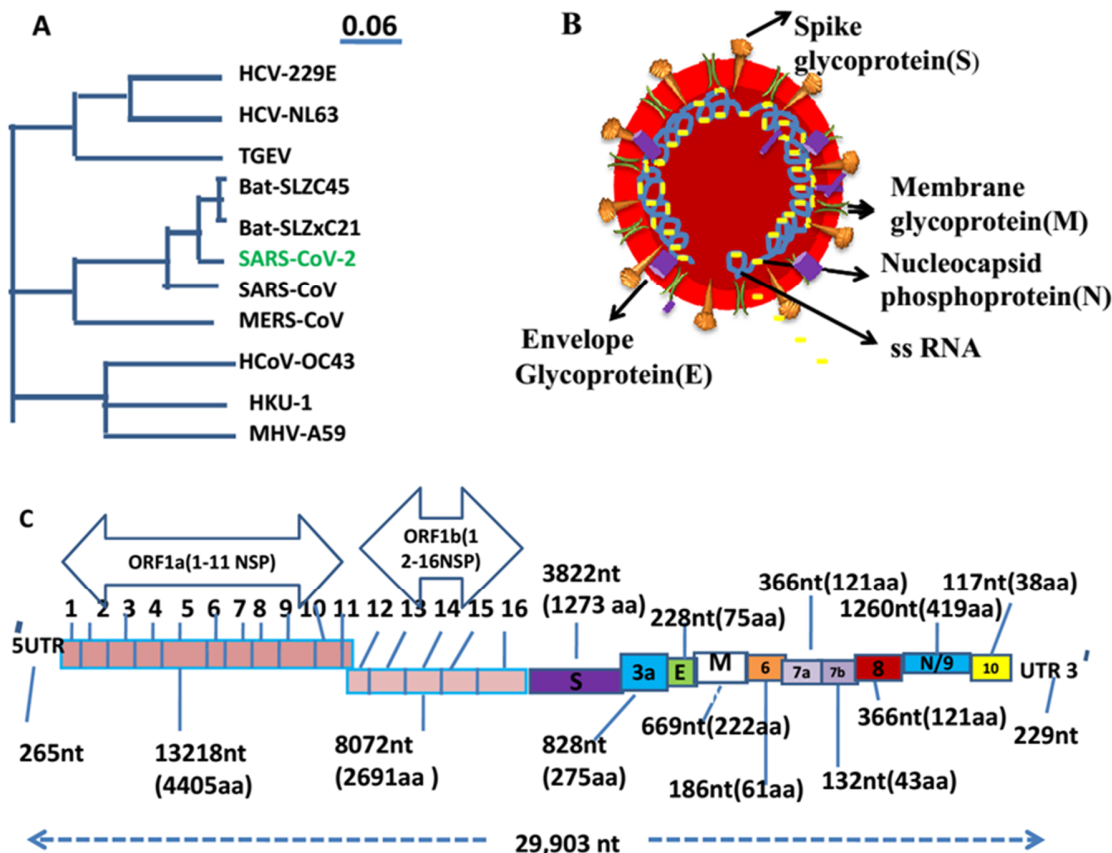


Fig. 1 SARS-CoV-2(COVID19) with genome and structural proteins A. Phylogenetic tree showing relationship among seven human corona viruses and other few coronaviruses(Fig drawn on the information from reference.16 ) B.A complete virion showing different parts ( Fig drawn on the basis of Transmission electron micrograph & Scanning electron micrograph of SARS-CoV-2 virus particles displayed at <https://www.flickr.com/photos/niid/49530315718> at the NIAID. Integrated Research Facility (IRF) in Fort Detrick, Maryland and information from NC\_045512.2, NCBI, Baltimore, C. An overall genome +ssRNA showing different non-structural proteins(nsp1-16) and structural proteins (S-Spike, E-Envelope, M-membrane, N-nucleocapsid n and ORF genes with nucleotides(nt) and amino acid number(aa) of product proteins( Fig drawn on the base of

information of reference 14 and GenBank: NC\_045512.2, NCBI, Baltimore).

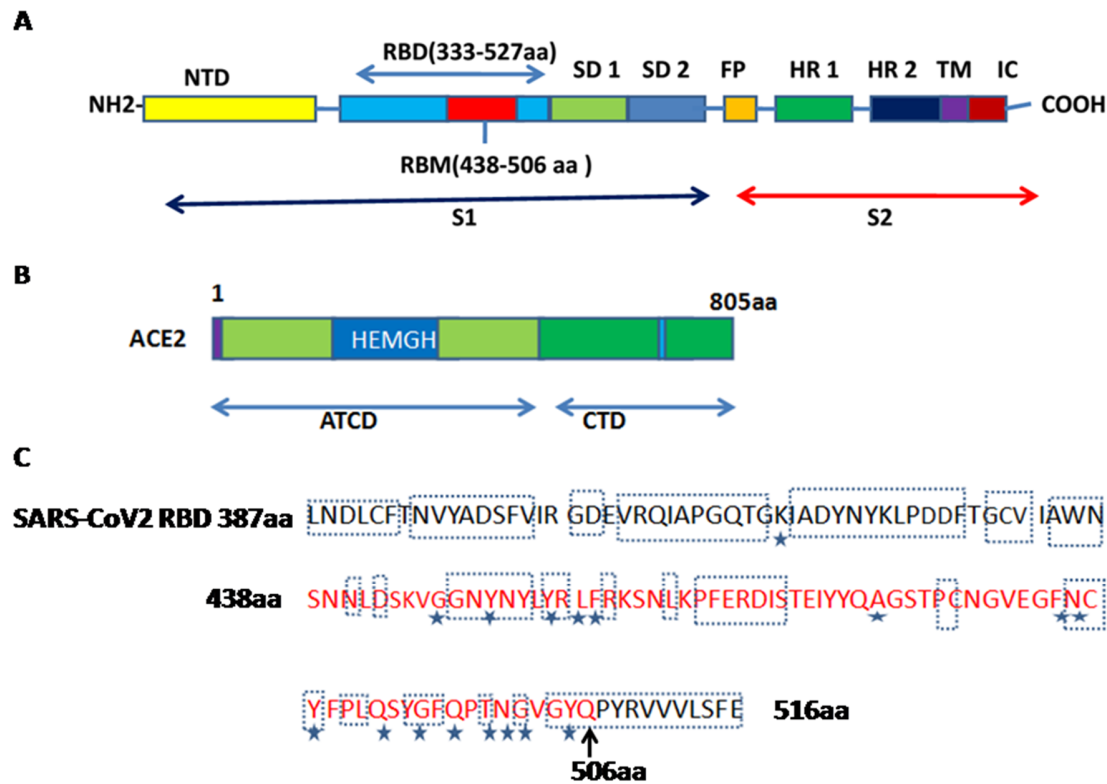


Fig. 2 Ligand (S-protein) of virus and human receptor ACE2(angiotensin converting enzyme 2).A.Topological structure of S-protein (Monomer) showing NTD(N-terminal domain). RBD, RBM(receptor binding motif) and S2 with SD1 (subdomain 1), SD2(subdomain 2), FD (fusion peptide), HR1(Heptad Repeat 1), HR2(Heptad repeat 2), TM(trans membrane region) and IC (intracellular domain) (Fig drawn on the basis of reference 66.) B. virion receptor, ACE2 showing its parts like HEMGH, CTD( carboxy-terminal domain, NTD ( N-terminal catalytic domain a) C. Amino acid residues of RBD of SARS-CoV-2 with showing contacting portions with ACE2(stars indicate the contacting positions) and similarity with RBD of SARS-CoV.(Dotted box showing similarity of amino acids between SAR-CoV-2 RBD and SAR-CoV RBD. Red zone shows RBM(438-506)aa. (Fig drawn on the base of information of reference 66).

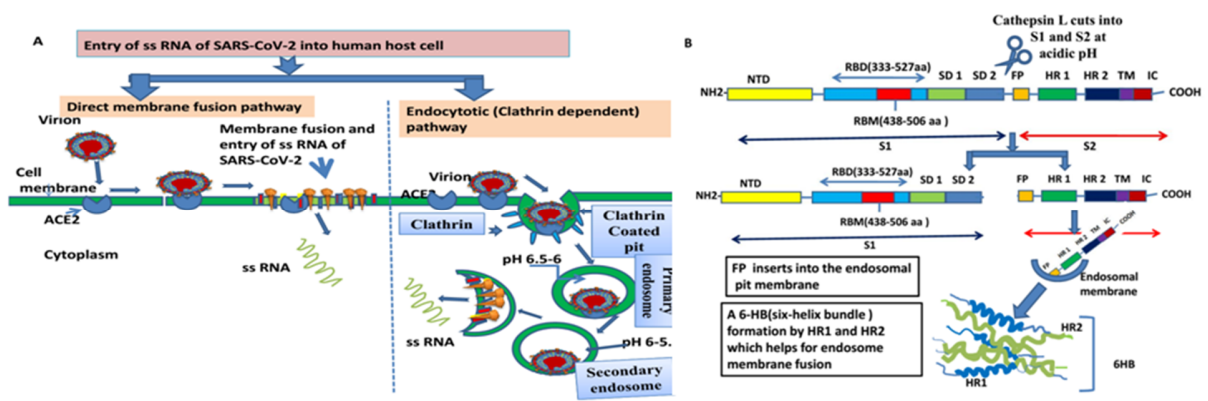


Fig. 3 Entry of human SARS-CoV-2 into host cell (assuming that SARS-CoV2 follows the mechanism of entry like SARS-CoV). A Direct membrane fusion pathway and endocytic (Cathrin mediated ) pathway (Fig drawn on basis of the information of references 76,77, 81, 86, 87) B. Cleavage of S protein into S1 and S2 by Cathepsin L, FP activation and 6-HB formation for viral envelope and membrane fusion in endosome.( Fig drawn on the basis of SARS-COV entry into human cell as by references 80,77).

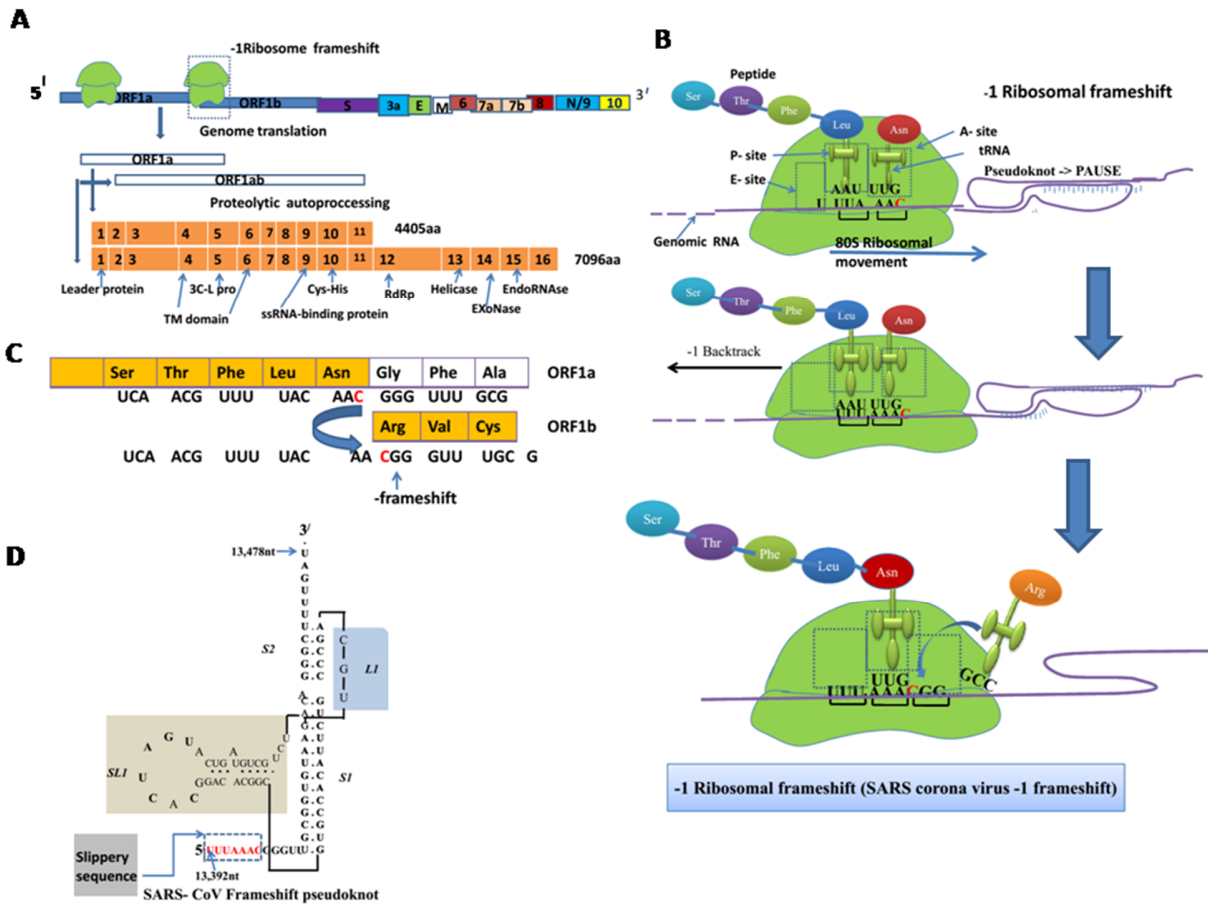




Fig. 4 Translation of RNA genome for production of sixteen nsps(nonstructure proteins) showing -1 Ribosomal Frameshift (-1RF) A. Mechanism of translation for nsp1-16 after proteolytic autoprocessing from ORF(openreading frame)1a and ORF1b of SARS-CoV2(Fig. drawn based on NC\_045512.2) B. Presentation of -1 ribosomal frameshifting by pseudoknot in corona virus(Fig. drawn based on reference 113) C. Codon sequence or base pairing during -1 frameshifting. D. SARS-CoV psedoknot with different parts and slippery sequence (Fig. drawn based on reference 111).

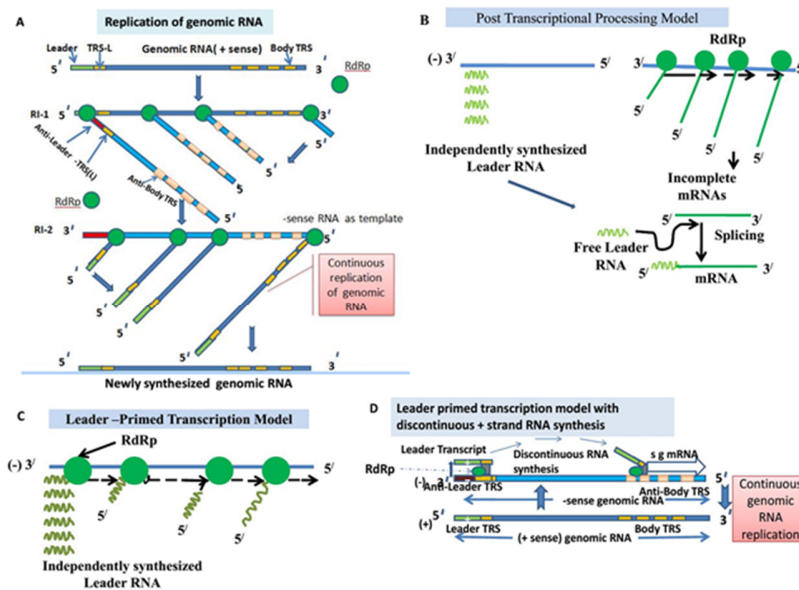


Fig. 5 Genomic RNA replication and different models for RNA synthesis of corona virus A. Continuous Replication of Genomic RNA on -sense RNA template showing RI-1 (Replicative Intermediate) and RI-2, B Post transcriptional processing model, C. Leader -Prime Transcription model (drawn on the basis of reference 134) D. Leader -Prime Transcription model with discontinuous + strand RNA synthesis (Fig drawn based on the information from references 120, 133, 134, 148).

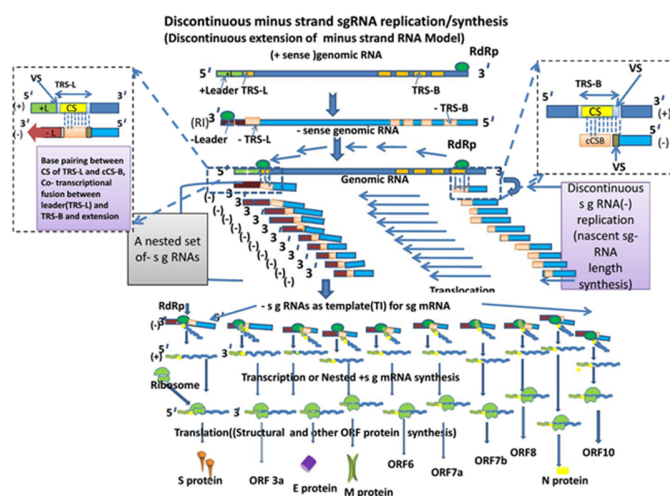


Fig. 6 Discontinuous minus strand sgRNA replication /synthesis (Discontinuous extension of minus strand RNA model, nested sgRNA synthesis and translation of all sgRNA producing all structural proteins of SARS-CoV2. (Fig drawn on the basis of references 122, 148 and NC\_045512.2,NCBI, Baltimore) (Note : in each +sense /antisense(-) full genome only four body TRSs /antibody TRSs have been shown for figure clarity; actually before each ORF gene /structural gene one TRS exists. Each sgRNA possesses one 5' end leader sequence, TRS, and gene for specific protein, 3' end shown in Fig sgRNA).(TRS-L,Transcription Regulating Sequence–Leader; CS, Core sequence;TRS-B (Body); cCSB, core conserved sequence; VS, Variable sequence; RI, Replication intermediate).

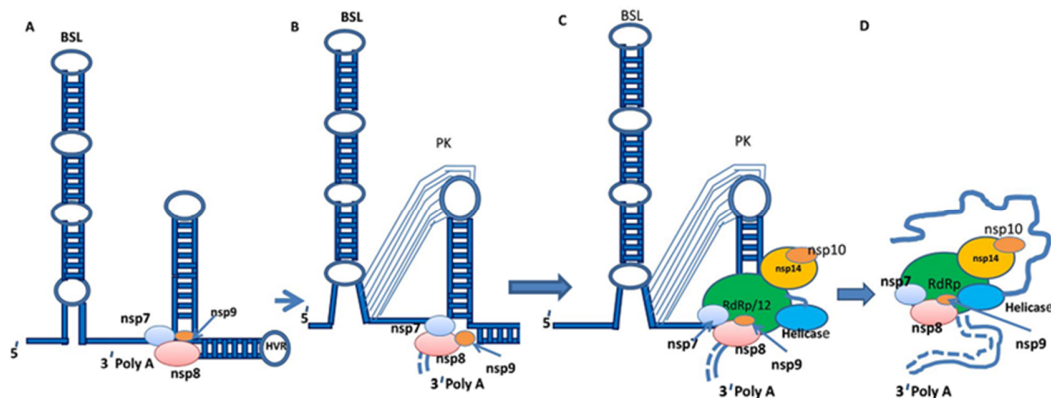


Fig 7 The model for the initiation of coronavirus negative-strand RNA synthesis (proposed by Sola et al(140), Zust et al.(153). A. nsp8, nsp7, and nsp9 form a complex to the genomic RNA 3' end, then nsp8 primase activity initiates RNA synthesis de novo. B. conformational change in the 3'-end RNA structure and the formation of the PK folding C. At the 3' end of the genome the RdRp and associated factors became recognized and promoting the initiation of negative-strand RNA synthesis, nsp 14 acts as proofreader D. Freely synthesis of negative-strand RNA synthesis(Fig drawn on the basis of reference140) (BSL, Bulged Stem Loop; PK, Pseudoknot).

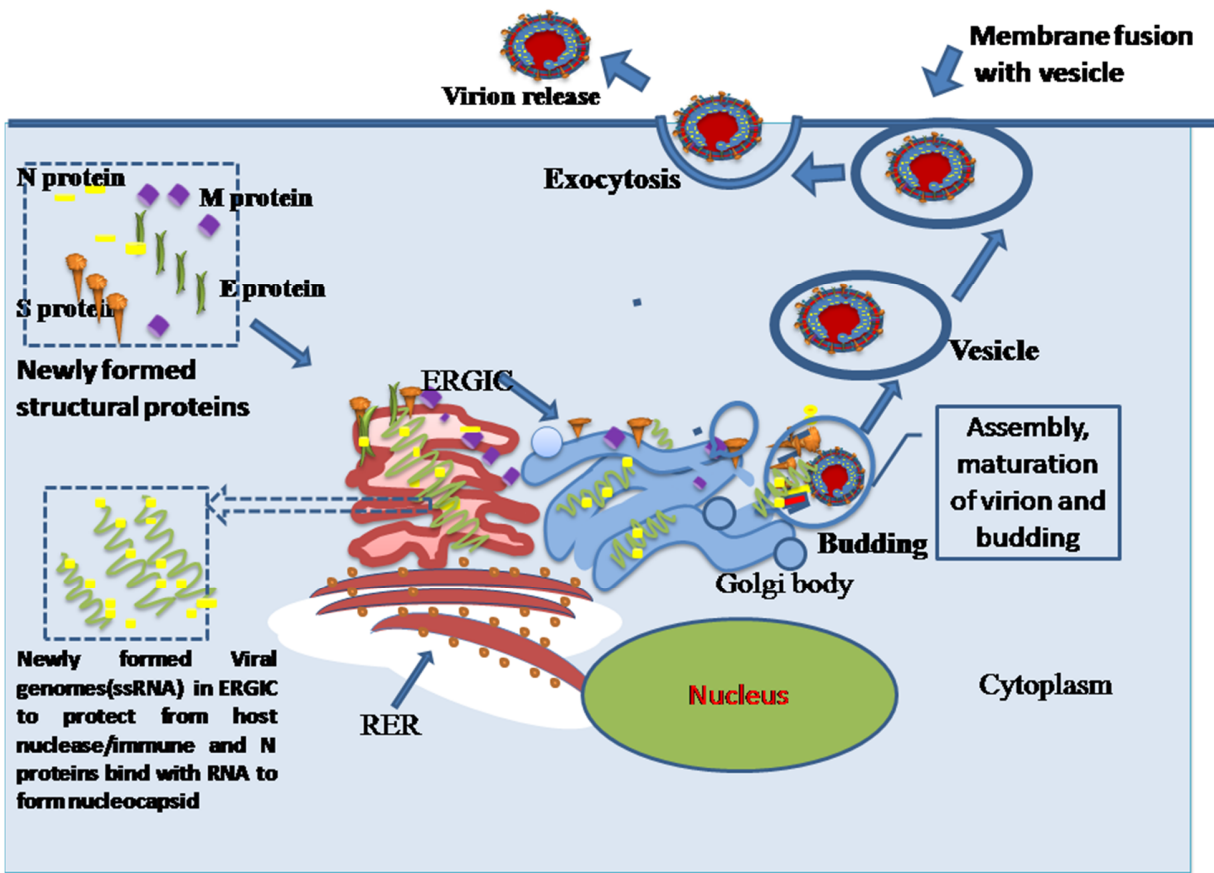


Fig. 8 Assembly of viral parts, maturation of virion in ERGIC(Endoplasmic reticulum Golgi body intermediate compartment or Golgibody), budding, vesicle formation and exocytosis for virion release.