REVIEW

UDC 578.1:578.233.7 doi: https://doi.org/10.15407/ubj94.04.005

STRUCTURAL PROTEINS IN THE MECHANISMS OF BETACORONAVIRUS VIRION ASSEMBLY

I. ZALOILO[1]✉, O. ZALOILO[2], Y. RUD[3], L. BUCHATSKYI[3]

[1]National University of Life and Environmental Sciences of Ukraine, Kyiv;
[2]Institute of Fisheries of the National Academy of Agrarian Sciences of Ukraine, Kyiv;
[3]ESC Institute of Biology and Medicine, Taras Shevchenko National University of Kyiv, Ukraine;
✉e-mail: zaloilo76@gmail.com, iridolpb@gmail.com

Received: 08 August 2022; **Revised:** 31 August 2022; **Accepted:** 04 November 2022

The emergence of SARS-CoV-2 caused an urgent need to investigate the molecular mechanisms of its reproduction. However, the detailed step-by-step mechanism of SARS-CoV-2 virion assembly has not been described yet. In the present review the data on the role of structural proteins in the efficient assembly of betacoronavirus particles are analyzed.

K e y w o r d s: SARS-CoV-2, betacoronavirus, structural proteins, virion assembly.

Betacoronavirus is one of four known genera of the subfamily of *Orthocoronavirinae*, family of *Coronaviridae*, order of *Nidovirales*. These are small enveloped RNA viruses, the diameter of which usually does not exceed 100 nm. Members of the genus *Betacoronavirus* contain a positive single-stranded RNA, which consists of an average of 30 kb. Betacoronaviruses are of zoonotic origin and some of them can infect humans. The most dangerous member of the genus *Betacoronavirus* is the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which appeared in 2019 and caused a global pandemic of COVID-19. During several years of active study of this betacoronavirus, the main efforts of scientists were directed mainly to the study of peculiarities of its penetration into the host cell, in particular, the peculiarities of the interaction of the spike protein with a specific human receptor called angiotensin-converting enzyme 2 (ACE2) [1].

Along with the mentioned aspect, a clear understanding of the mechanisms of assembly of mature virions is crucial for finding alternative ways to fight off betacoronaviruses. Experimental works devoted to this topic contain mostly fragmentary, sometimes contradictory data and purely hypothetical conclusions. Therefore, generalization and systematization of existing information on the mechanisms of assembly of betacoronavirus virions is an urgent problem today.

It is clearly that various members of the genus *Betacoronavirus* can be characterized by individual peculiarities of self-reproduction. However, it is worth mentioning that the currently most studied SARS-CoV-2 has an extremely high level of genomic similarity with other betacoronaviruses: it is 70% identical to the Middle East respiratory syndrome virus MERS-CoV [2], and this reaches 83% in the case of SARS- CoV-1 [3]. Comparisons of purely structural proteins show up to 50% and 85% similarity of SARS-CoV-2 with MERS-CoV and SARS-CoV-1, respectively [4-5]. It is interesting that one of the most studied members of the genus *Betacoronavirus*, the murine hepatitis virus (MHV), turned out to be quite similar to SARS-CoV-2 [6]. Therefore, given this high similarity, it is quite possible to extrapolate the basic principles of functioning of one type of betacoronavirus to other members of this genus. The presented work contains information about the effect of structural N, E, M and S proteins on the most studied final phases of viral reproduction, the result of which is assembly of new SARS-CoV-1, SARS-CoV-2 and MHV virions.

I. RNA replication site of betacoronaviruses

A number of studies [5, 7, 8] showed that the intermediate "double-stranded" step in the synthesis of betacoronavirus RNA occurs directly in double-membrane vesicles (DMV). These structures are formed on the nuclear periphery and are modified membranes of the host cell separated from other organelles, although the possibility of their connection with the endoplasmic reticulum is not rejected [9, 10]. Single membrane vesicles and convoluted membranes are also considered alternative sites for the synthesis of viral RNA, however, despite the presence of these organelles in the studied cells, no signs of replication were found in them. Thus, DMV is currently the only established site of synthesis of full-length RNA and fragments of subgenomic RNA of coronaviruses in the host cell [8].

There have been attempts to create a unified model of coronavirus reproduction [10, 11]; however, obtained modern data do not allow establishing a detailed mechanism of DMV formation. The results of SARS-CoV-1 studies carried out in 2013 [12] showed that the leading role in the induction of these organelles is played by non-structural proteins 3, 4 and 6: the non-structural protein 3 is the "switch" of induction, while proteins 4 and 6 function as regulators of this process. However, the vesicles produced under transfection conditions, although being structurally similar to their counterparts in infected cells, were characterized by significantly smaller sizes. According to the authors [12], such a result indicates that DMV formation is affected by a significant number of factors including auxiliary proteins and the presence of RNA. Similar conclusions about the effect of non-specific proteins were also proposed in a research paper [13] based on the results of the study on murine hepatitis virus. It is interesting that the authors indicated the existence of a certain structural variability of double-membrane vesicles induced by murine hepatitis virus and the discrete nature of RNA synthesis occurring in them. The described phenomena may be the result of different durations of coronavirus adaptations to a specific host cell.

The most studied DMVs in which viral RNA synthesis occurs are SARS-CoV-2 vesicles. These organelles are the result of modification of cell membranes: the layers of such structures are separated by 18 nm. The diameter of the inner membrane reaches 340 nm, which on average is similar to the size of the vesicle membranes in SARS-CoV-1 [14]. In [15] it was shown that the auxiliary protein Orf3a has a significant effect on increasing the number of double-membrane organelles for SARS-CoV-2 replication.

The peculiarities of the location for viral RNA synthesis (in a double-membrane vesicle, which, in turn, is surrounded by tortuous structures) can be probably an evolutionarily acquired adaptation aimed at avoiding the recognition of the coronavirus by the host's immune system [8, 10, 12].

II. Viral RNA release from double-membrane vesicles

Newly synthesized RNA is released from the inner DMV membrane and forms a complex with nucleocapsid (N) protein in the cytosol. This mechanism of RNA transport from double-membrane vesicles is practically the same for MHV and SARS-CoV-2 betacoronaviruses. In the paper [7], the authors indicate that DMVs contain specific pores with a diameter of up to 3 nm located between two membranes, which with a high probability ensure the release of mRNA into the cytoplasm for further translation. It is important to note that the number of such structures in DMV formed during MHV infection significantly exceeds that for SARS-CoV-2: the extrapolated average values were 8 and 0.5 pores per one vesicle, respectively. Obviously, such a small number of DMV "portals" in SARS-CoV-2 may indicate the presence of yet unknown alternative ways of RNA transport.

DMV pore structures formed by MHV are known to have a hexagonal symmetrical shape and in their cytoplasmic part consists mainly of non-structural protein 3. The rest of the pore is probably built by non-structural proteins 4 and 6 with transmembrane domains in the structure, which are the probable key to the induction of the same double-membrane vesicles [12]. As it is assumed, it is the non-structural protein 3 that directs the synthesized viral RNA to the pore structure through an unknown mechanism during infection with both MHV and SARS-CoV-2 [11, 13].

III. The role of nucleocapsid (N) protein in betacoronavirus RNA packaging

As is known, the structural N protein of coronaviruses is associated with genomic viral RNA and another structural M protein, which probably controls the processes of RNA packing and genome encapsulation [16]. The structure of N protein is provided in Fig. 1. In SARS-CoV-2, it is a cytosolic protein with a weight of 46 kDa consisting of 419

amino acids, which is approximately 85% identical to the general structure of the N-protein of SARS-CoV-1 [17, 5].

The N protein can be divided into five domains: an N-terminal, an intrinsically disordered region that probably contacts with the RNA packing signal, an RNA-binding domain that accompanies RNA, which was released from vesicles into the cytosol, and a disordered linker region, which is rich in arginine and serine, a dimerization domain and a disordered carboxylated C-terminal region, which is involved in the interaction of N and M proteins [16, 18]. The paper [19] suggests that all 5 mentioned domains are suitable for binding to RNA and forming a nucleocapsid. Taking into account the parallel contribution of the dimerization domain to the formation of specific stoichiometric domains, it is possible to predict the process of RNA-independent oligomerization of a higher order [20].

In the paper [16], the N-terminal domain is considered to be the location that ensures the formation of a complex of N protein with SARS-CoV-2 RNA, since it is characterized by a high affinity for binding to the viral RNA. Previous studies of the SARS-CoV-1 nucleocapsid protein also concluded that the N-region is necessary for RNA packaging [21].

The studies of the assembly processes of virions of murine hepatitis virus show that N protein is localized near the exit from the pore, i.e. the RNA transported through the channel from the double-membrane vesicle can, with a high probability, bind to N protein [22]. Given a high level of structural similarity, the formation of the "RNA-N protein" complex was also considered for SARS-CoV-2 at the first stages of its study, according to the same scheme. However, quite soon a question arose for this betacoronavirus about the possibility of forming such a complex not only with full-length RNA, which contains 30 kb, but also with subgenomic variants, which, according to [13], compose the majority of newly synthesized RNAs.

The formation of the RNA-N protein complex in SARS-CoV-1 and MHV betacoronaviruses is preceded by special packing of the newly synthesized ribonucleic acid molecule. The study [17] established the presence of a signal region of MHV genomic RNA, which "switches on" the process of its packaging: it is a substructure of 69 nucleotides, which is part of a 190-nucleotide segment located approximately 20 kb away below the 5′-terminus of the genome. It is obvious that a similar packing step is inherent in the similar complex formation of SARS-CoV-2. The study [23] proposed the following model of packing initiation for this betacoronavirus: N protein undergoes "liquid-liquid" phase separation with the viral genome. At the same time, the nucleocapsid protein condenses with specific RNA sequences in the first 100 nucleotides of the 5′-terminus. The formed protein condensates exclude unpackaged RNA sequences from the composition. Thus, the authors hypothetically propose the aforementioned process of "liquid-liquid" phase separation (LLPS) as a signal for SARS-CoV-2 RNA packaging.

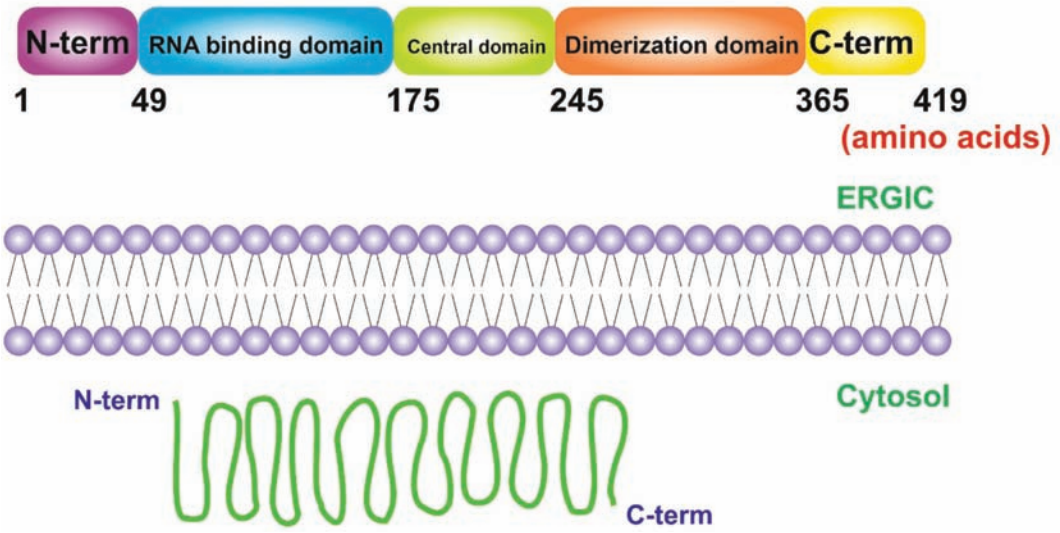

Fig. 1. Structural and topological scheme of N protein

The necessity of the presence of nucleocapsid protein for the complete assembly of virions is not a definitively established fact. For example, in MHV, one cellular mRNA can be incorporated into a viral particle by interacting with the structural M protein, i.e., the presence of N protein is not necessary for virion formation for this betacoronavirus (unlike SARS-CoV-1) [24]. However, the probability of "RNA-N protein" complex formation for SARS-CoV-2 is considered very low, although this mechanism also needs experimental verification with unequivocal results.

The questions regarding the processes accompanying nucleocapsid penetration into the coronavirus cavity still remain largely open. The first electron microscopic studies of coronaviruses [25] showed the presence of single-chain ribonucleoproteins (RNPs) with a diameter of up to 15 nm in the "RNA-N protein" complex. Much later, electron cryomicroscopy [26] allowed detecting a lattice-like organization of clusters of viral RNPs in SARS-CoV-1, and even the distances between individual molecules (4-5 nm) were found. However, similar studies on MHV did not show the presence of such RNP lattices [27]. According to recent works [28, 29], the structure of the "RNA-N protein" complex for SARS-CoV-2 is proposed to be considered as a set of so-called sub-complexes of about 12 N proteins (their exact number is not established), which consist of about 800 nucleotides forming a RNP. The row of such structures is located along a sufficiently long RNA and resembles a necklace on a thread. In contrast to the number of N proteins, the number of RNPs was established quite clearly: 30-35 per virion [28, 30]. The further process consists in the self-assembly of N protein with the formation of a tetramer, and then, the oligomerization of this protein in the presence of RNA [31].

Many works [16, 18, 19] affirm that several domains are involved in the processes of N protein condensation from viral RNA, however, only the central (linker) region has a special affinity. The same domain with the regulatory participation of the N-terminus is responsible for its phase separation from the coronavirus RNA from a liquid to a gel-like condensate [32], although, the condensate is not thickened with excessive phosphorylation of the N-domain (observed during cell infection) [16].

The later step is characterized by an interaction of the N protein complex with a RNA with the structural membrane M protein [18]. In SARS-CoV-1, such contact occurs directly between N- and M-proteins, probably through their carboxylated terminal ends [21, 33]. The specifics of this process for SARS-CoV-2 have not yet been described.

IV. Functions of M protein in the processes of virion assembly

The transmembrane M glycoprotein as well as N protein is a structural protein of betacoronaviruses, which, while interacting with itself and with other viral structural proteins, performs leading functions in the mechanisms of virion assembly. The M protein is formed by 222 amino acids with a total weight of 25 kDa. The spatial structure of the molecule looks like a triple helix [18, 34], has three transmembrane domains with a long C-terminus (located in the cytoplasm) and a short N-terminus (located outside the virion) [35] (Fig. 2).

Results of cryo-electron microscopy allowed describing two conformations of the membrane structural protein for MHV and SARS-CoV-1 viruses: compact (in the form of a conventional ellipse) and so-called "long" (2 nm longer than the ellipsoid. At the same time, the authors [36] consider that the compact form is responsible for the general regulation of membrane curvature, while the function of the elongated one is mainly associated with the rigidity and fine regulation of membrane curvature. Since the "long" conformation has an endodomain end, it is obviously involved in the formation of the M-N protein complex. Similar conformations of M protein are quite likely for SARS-CoV-2, given its 90% similarity to SARS-CoV-1 [37].

The so-called "M-M" complex (the product of M protein binding to itself) is a necessary condition for the processes of virion assembly. The mechanism of such interaction is best studied for mouse hepatitis virus [38-40]. Use of numerous mutated variants of M protein showed [41] that these proteins are able to selectively exclude "foreign" proteins from the membrane in the process of a complex mechanism of interaction. As a result, a special structure with a clear pattern formed by 1-4 M-proteins is formed on the viral envelope. The main function of the C-terminus of an M-protein molecule is to ensure the lateral M-M interaction [41].

The M-M complex in SARS-CoV-1 is formed as a result of the interaction of the cytosolic domain of M-protein with its dimers [36]. The studies [42] show that, unlike MHV, the more important role in the formation of virions in SARS-CoV-1 is not

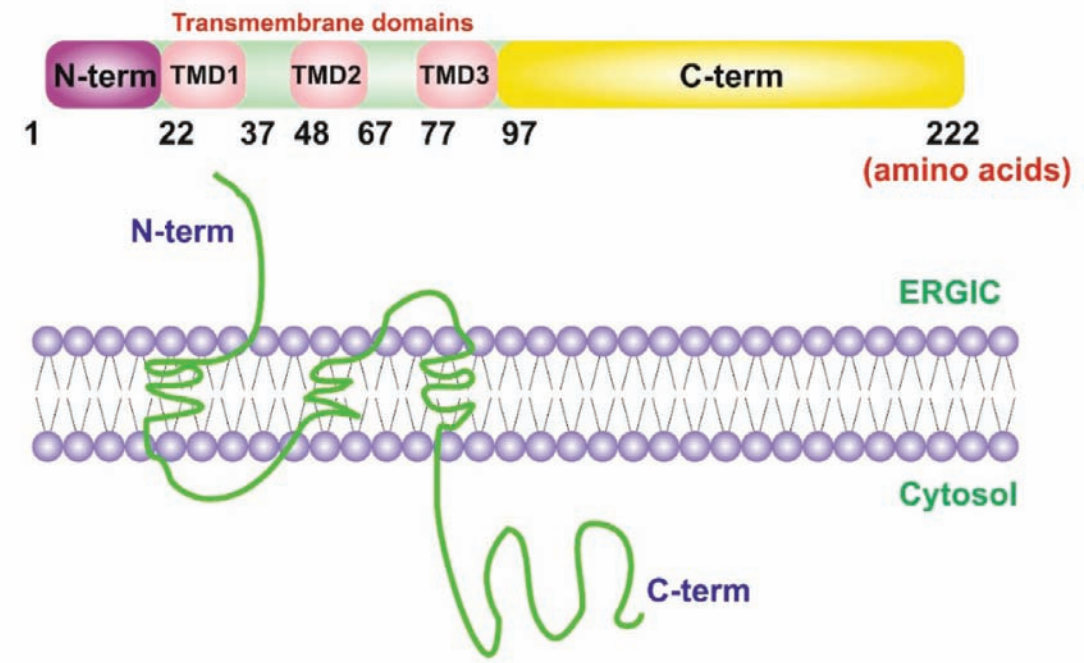

Fig. 2. Structural and topological scheme of M protein

played by the transmembrane domains but rather by the carboxylated C-terminus of the molecule: the authors alternately deleted different domains of M-protein and observed the release of vesicles in the presence of only a C-segment. Earlier studies [33, 36] showed that the interaction of the ends of this protein in SARS-CoV-1 occurred by electrostatic way, since their carboxyl groups were hydrophilic and had a charge. As a result of this mechanism, the virion also includes a nucleocapsid, i.e., a complex of N protein with RNA. A similar functional advantage of the C-terminus over the transmembrane segments of the protein molecule was observed in the assembly of SARS-CoV-2 virions [43]. The purpose of M protein binding to the nucleocapsid complex for this coronavirus is considered in [44]: the authors assume that this complex is a certain stabilizer of the C-terminus, which, in turn, ensures an increase in the number of formed virions.

It is worth noting that the localization of M protein at each step of this process is also important for the assembly and separation of virions. The assembly of coronavirus virions is known to occur in a gap between the Golgi apparatus and the endoplasmic reticulum. A number of works [39, 45] showed the presence of M protein in the Golgi apparatus, the accumulation of these molecules and their modifications. This protein was also found in vesicles located along the secretory pathway. For SARS-CoV-2 and MHV, the ability of M-protein clustering followed by its interaction with its own modified (glycosylated) form has been proven. The authors see the reasons for this recycling of M protein into the Golgi apparatus for binding to other structural proteins in the still unknown functions of the N-terminus domain of the molecule.

V. Involvement of E-protein in the assembly of betacoronavirus virions

The envelope (E) protein of coronaviruses is a relatively small protein of the transmembrane type with a weight of 8-12 kDa, which contains 75 amino acids. An E-protein molecule has the following structure: a transmembrane (hydrophobic) domain formed of 25 amino acids, a luminal N-terminus of 7-12 amino acids, and a hydrophilic carboxyl C-terminus [18, 46]. Spatially, E protein is oriented with its long C-terminus to the cytoplasm, and its short N-terminus to the intermediate compartment of the edoplasmic reticulum and the Golgi apparatus (ERGIC) (Fig. 3).

The study [47] of the E-protein of SARS-CoV-1 using immunofluorescence microscopy showed that its terminuses in the secondary structure of the molecule form a so-called beta-hairpin - a location, the secondary structure of which looks like two threads combined with amino acids oriented in antiparallel directions. The authors suggest that this structure is

in some way responsible for the spatial orientation of E-protein relative to the Golgi apparatus. The location of this structural protein of SARS-CoV-2 and MHV in the Golgi apparatus was observed in the medial and cis-cisterns [48], which does not contradict the proposed concept. The work [49] showed the possibility of separate expression of E-protein with a labeled carboxyl terminus for SARS-CoV-2. The authors [50] affirm that such autonomy may endanger the natural arrangement of the entire molecule, because it is known that the envelope (E) protein interacts both with other structural proteins of betacoronaviruses and with host proteins [46, 51]. The question of the structural and functional features of post-translational modifications of E protein, in particular, and the effect of such transformations on virion assembly remains open [51].

The viral E-protein is a viroporin: it oligomerizes subsequently forming an ion channel, through which virions are released [46, 52]. The question of the specificity of such a portal to certain cations and the peculiarities of virion release as a result of the formation of an ion flow through the channel remains controversial [53, 54]. Thus, a study of the transmembrane domain of the E-protein of MHV showed that virion release can cause not only the actual flow of ions through the pore, but also can cause mediated membrane cleavage [55]. It is worth noting that, for SARS-CoV-2, the location of the transmembrane domain, which is directly involved in the formation of the ion channel, has not been established yet.

The leading role in the formation of vesicles, inside which the SARS-CoV-1 and SARS-CoV-2 viral protein is contained, definitely belongs to E and M proteins [56]. For SARS-CoV-2, as it was recently shown, the deletion of the carboxyl terminus of E- and M-proteins inhibits their potential for interaction, while the attachment of ubiquitin monomers to the amino groups of the N-terminus of M protein, on the contrary, is a stabilizing factor for similar interprotein interactions [43]. The authors of this work also recognize an indirect effect of the consequences of M-E protein interactions on the release of vesicles. In the work [50], these factors are explained by the possible actions of E-protein in membrane rupture and steps of the virion assembly mechanism with the involvement of membrane structural protein. The versatility of M-E protein interaction in coronaviruses was clearly illustrated in the work [57]: the E proteins of SARS-CoV-1 and even the avian infectious bronchitis virus demonstrated a high ability to interact with the M protein of MHV.

The E protein interacts with nucleocapsid N protein through the carboxylated terminus. The functional role of such contacts has not been established yet, although there is an assumption that the inclusion of ribonucleoproteins is enhanced as a result of the E-N interaction [48]. Despite numerous gaps in the current understanding of the effect of E protein on virion assembly and release, the importance of its role in the mentioned processes is beyond doubt.

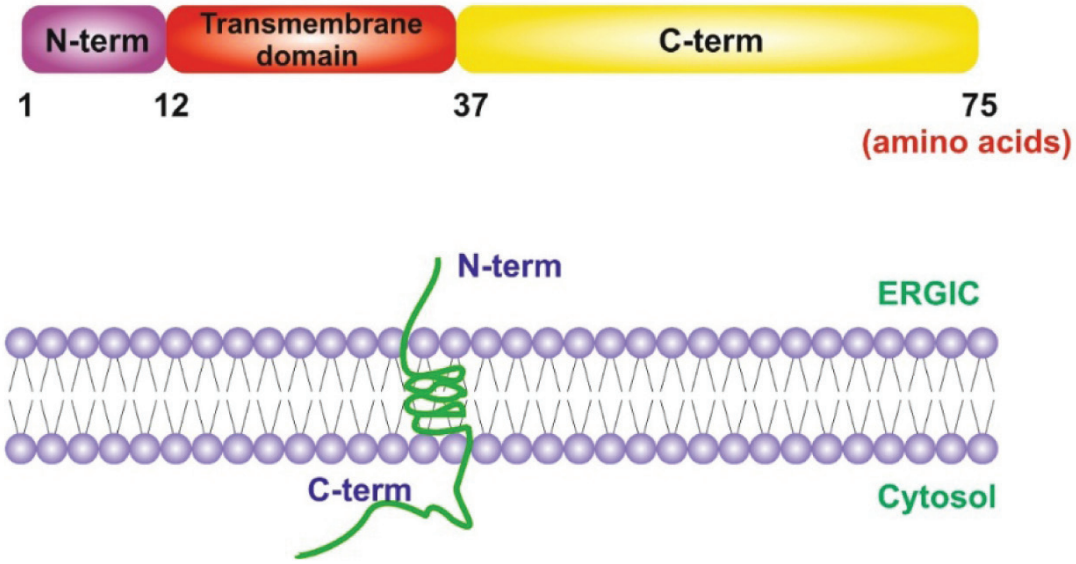

Fig. 3. Structural and topological scheme of E protein

VI. S-protein in the assembly mechanisms of betacoronavirus virions

The spike S-protein of coronaviruses plays a major role in the process of virus penetration into the host cell. The structural and functional features of this structural protein have been studied in depth, including for SARS-CoV-2 and its mutated forms, since it is the S-protein that can be an effective target for targeted vaccines [58, 59]. It is a transmembrane glycoprotein assembled in the form of a trimer directly on the viral membrane. The structure of S-protein is represented by two subunits: S1 and S2 (Fig. 4). S1 contains the N-terminus and the receptor-binding domain (RBD), which directly interacts with the ACE2 receptor of the host cell for further entry of the virion into it. The S2 subunit contains a transmembrane domain and the C-terminus, which plays a major role in membrane fusion [60].

The functional spectrum of S protein is quite broad and is not the subject of this review. Despite its absolute importance for viral reproduction, only isolated facts of the indirect effect of this protein on virion assembly have been described so far. For example, the work [44] mentions that E-protein causes retention of glycoproteins (including the structural S-protein) inside the cell by slowing down the secretory pathway of it. The authors hypothesize that such a forced concentration of the spike protein in the intermediate compartment may indicate a certain importance of this protein for virion assembly, but no direct confirmation of this was obtained. There is also no information on the interaction of E and S proteins. In general, taking into account the results of numerous studies in recent years, it is possible to assume the effect of S-protein on the general kinetics of coronavirus assembly or on the stability of the processes of its budding on the cell membrane, but not on the direct participation of this protein in virion separation or in membrane bending.

Summarizing the review of the effect of structural proteins on virion assembly, it is worth noting that the set of necessary proteins varies depending on the betacoronavirus type. For example, the presence of N, E and M proteins is mandatory for the stable production of virus-like particles of SARS-CoV-1 and SARS-CoV-2, [42, 61]: N and M participate in the formation of vesicles, while E protein affects an increase in the number of virions and ensures their release into the intercellular space (confirmed for SARS-CoV-1) [62]. However, earlier studies of virus-like particles obtained by recombinant expression of the main structural proteins showed a dependence only on M and E proteins during assembly after infection of insect cells [63]. As for the MHV virus, the effect of only M and E proteins on virion assembly was also confirmed [64].

Studies of the minimal set of structural proteins for SARS-CoV-2 assembly are still ongoing. The

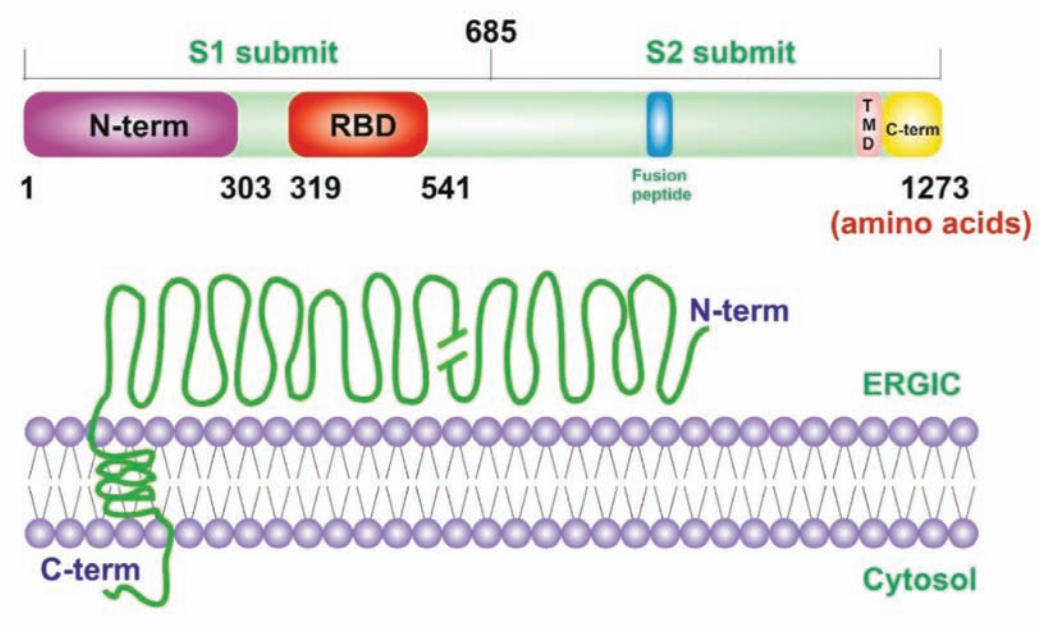

Fig. 4. Scheme of the structure and spatial arrangement of S protein

currently described are fragmentary and sometimes show open contradictions. For example, the work [61] showed that the condition for the formation of vesicles is the presence of a pair of M and E proteins. At the same time, the results obtained in the study [56] indicate that, in addition to E and M proteins, S protein is also required for the production of fully functional virions.

VII. Participation of betacoronavirus structural proteins in ERGIC membrane bending

The presence of main structural proteins in the sites of betacoronavirus virion assembly suggests their effect on the change in membrane curvature in the process of virion formation. The E-protein is located in the ERGIC compartment, where virus-like particles are assembled. Studies of this MHV protein showed its ability to form tube-like structures, which may well cause bending of the intermediate compartment membrane [47]. The dynamic model of the E-protein of SARS-CoV-2 also predicts the participation of its amphipathic helices of the carboxylated C-terminus in the change of membrane curvature [65].

The possible participation of M protein in ERGIC membrane bending is more indirect: the authors [39] believe that interacting with itself, this protein makes a purely energetic contribution to the process of membrane bending. Thus, it is possible to consider the presence of E and M proteins as one of the numerous factors triggering the change in membrane curvature.

It is known that during the assembly of SARS-CoV-1 virions, when M and N proteins interact, a switch is made between the ellipsoid and elongated conformers of M protein. The authors [36] show that only the elongated conformation forms a convex envelope rigid enough to significantly affect the membrane curvature. However, this question remains open, since the fact of bending as a result of binding of the RNA-N protein complex to the terminal domain of M protein was not proven, and RNP accumulation can occur under the already curved areas of the ERGIC membrane caused by E protein or the interaction of E and M proteins.

The presence of the spike protein is not a necessary condition for virion release into the intercellular space, so the role of this protein in membrane bending is unlikely. This concept was confirmed experimentally for SARS-CoV-2 [66]: cryo-electron tomography in situ did not allow recording ERGIC bending in sites where S-proteins accumulated.

Conclusion. The results of the studies presented in this review show that different types of betacoronaviruses are characterized by individual features of virion assembly mechanisms. However, taking into account the significant percentage of structural and functional similarity of the most studied betacoronaviruses such as SARS-CoV-1, SARS-CoV-2 and MHV, it is possible, with a sufficient degree of probability, to extrapolate the proven properties of one type of virus to others to obtain a generalized scheme of virion assembly at least for the three aforementioned betacoronaviruses.

The synthesis of viral genomic RNA occurs inside double-membrane vesicles, while its release occurs through membrane pores. RNA release with a high probability precedes the creation of its complex with nucleocapsid N-proteins. The resulting structure resembles a necklace on a thread, which contains 30-35 viral ribonucleoproteins. The existence of this structure was proven by electron microscopic methods [14], but the ratio of the number of N-proteins to individual RNPs remains unknown. After the “RNA-N protein” complexes are formed in cytosol, liquid phase transitions occur, which are probably one of the regulatory mechanisms of RNP assembly on membranes of the intermediate ERGIC compartment.

According to [44], it is E protein that has the ability to create a local concentration of main structural coronavirus proteins in ERGIC membranes. It is believed that release of these glycosylated viral proteins from the Golgi apparatus to ERGIC occurs synchronously, since their inclusion into the composition of the forming virion can also take place after very short intervals, or even simultaneously. However, the stimulating factors as well as the transport mechanism of structural proteins to the intermediate compartment remain unknown. Another promising direction of studies is establishing the locations of structural proteins that participate in their clustering in ERGIC.

The next step of virion assembly is numerous interactions of the nucleocapsid (RNA-N protein) with the membrane structural M-protein. The formed complex is part of new virions, which already contain free E, M, and S proteins [32]. Another unique property of the membrane protein is the ability of its molecules to interact with each other. It is assumed that precisely such interactions cause membrane bending, which contributes to the separa-

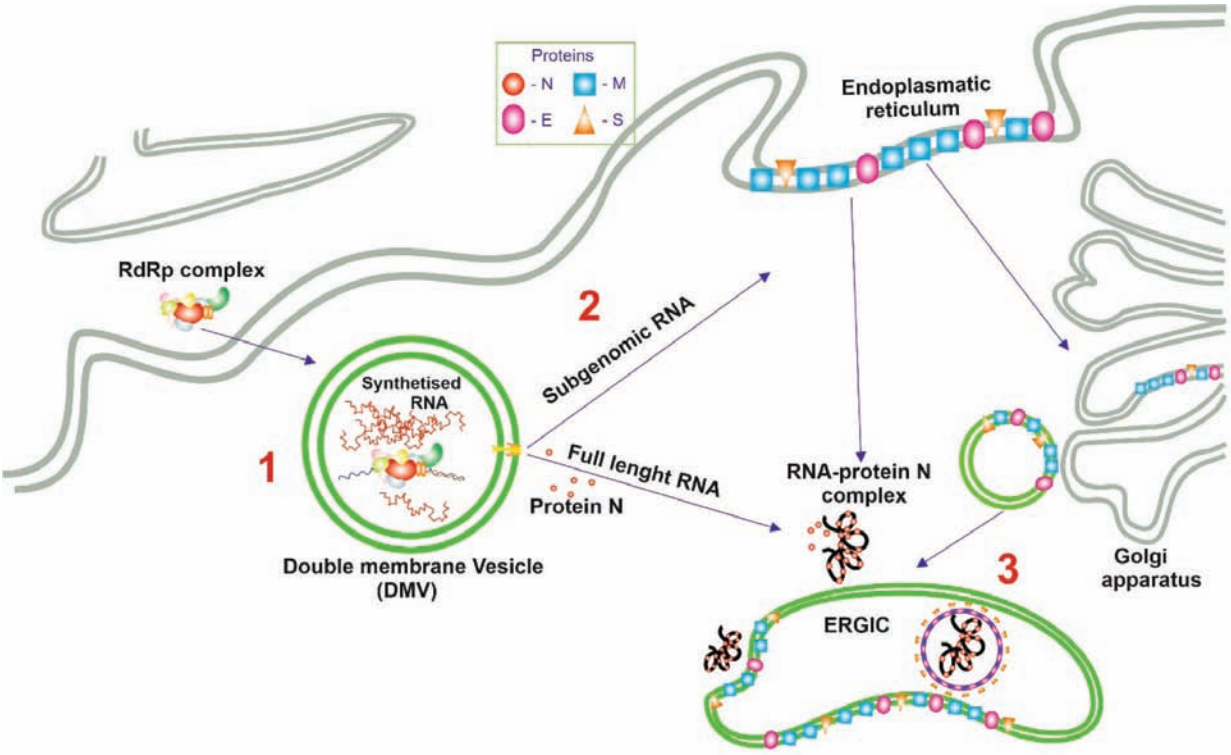

Fig. 5. General diagram of the final steps of betacoronavirus virion assembly: ***1*** *– RNA replication in double-membrane vesicles (DMV).* ***2*** *– Translation of subgenomic RNA to locations with structural and auxiliary proteins.* ***3*** *– Assembly of the "RNA-N protein" complex with other structural proteins (M, E and S) in the intermediate compartment between the endoplasmic reticulum and the Golgi apparatus (ERGIC)*

tion of viral particles inside ERGIC under the parallel effect of membrane and nucleocapsid proteins. This hypothesis [36] does not contradict the current ideas about the betacoronavirus virion assembly, has empirical and indirect experimental confirmation, however, without any direct evidences of the reality of the mentioned mechanism. There is also a question regarding the priority of participation of proteins in changing the membrane curvature of the intermediate compartment between the endoplasmic reticulum and the Golgi apparatus: after all, it is known that virions and vesicles can be formed even in the complete absence of N-protein, so the critical need for E protein to change the ERGIC membrane curvature can be overestimated compared to the functions of M protein. Further processes associated with viral budding obviously occur under the effect of membrane structural E protein, the mechanism of action of which requires further studies and detailing [46]: it is quite likely that currently unknown cellular cofactors can stimulate budding or specific lipid inclusions. The scheme describing the steps of betacoronavirus virion assembly mentioned in the work is shown in Fig. 5.

Thus, despite many years of active research, the final steps of the process of betacoronavirus virion assembly exist in the understanding of the modern scientific community only as a hypothetical general scheme supplemented by fragmentary experimentally proven facts. At the same time, the creation of clear, holistic ideas about these mechanisms is not only interesting from a purely scientific point of view, but will be also definitely useful for practical purposes, in particular, in the process of creating the latest targeted vaccines or other preventive methods against coronavirus infections.

Conflict of interest. Authors have completed the Unified Conflicts of Interest form at http://ukrbiochemjournal.org/wp-content/uploads/2018/12/coi_disclosure.pdf and declare no conflict of interest.

Funding. National project (2021-2023) RW№21BF051-03 funded by the Ministry of Education and Science of Ukraine "Investigation of photophysical and photochemical properties and interaction with cell membranes of RNA-containing viruses in purpose to inactivate them".

СТРУКТУРНІ ПРОТЕЇНИ У МЕХАНІЗМАХ ЗБИРАННЯ ВІРІОНІВ БЕТАКОРОНАВІРУСУ

І. Залоїло[1]✉, О. Залоїло[2], Ю. Рудь[3], Л. Бучацький[3]

[1]Національний університет біоресурсів і природокористування України, Київ;
[2]Інститут рибного господарства НААН України, Київ;
[3]ННЦ «Інститут біології та медицини», Київський національний університет імені Тараса Шевченка, Україна;
✉e-mail: zaloilo76@gmail.com; iridolpb@gmail.com

Поява SARS-CoV-2 спричинила нагальну потребу дослідити молекулярні механізми його розмноження. Однак наразі детального покрокового механізму збирання віріону SARS-CoV-2 не описано. У представленому огляді проаналізовано дані про роль структурних протеїнів у ефективній збірці частинок бетакоронавірусу.

Ключові слова: SARS-CoV-2, бетакоронавірус, структурні протеїни, збирання віріонів.

References

1. Whittaker GR, Daniel S, Millet JK. Coronavirus entry: how we arrived at SARS-CoV-2. *Curr Opin Virol*. 2021; 47: 113-120.
2. Guruprasad L. Human coronavirus spike protein-host receptor recognition. *Prog Biophys Mol Biol*. 2021; 161: 39-53.
3. Nadeem MS, Zamzami MA, Choudhry H, Murtaza BN, Kazmi I, Ahmad H, Shakoori AR. Origin, Potential Therapeutic Targets and Treatment for Coronavirus Disease (COVID-19). *Pathogens*. 2020; 9(4): 307.
4. Voskarides K. SARS-CoV-2: tracing the origin, tracking the evolution. *BMC Med Genomics*. 2022; 15(1): 62.
5. Kaur N, Singh R, Dar Z, Bijarnia RK, Dhingra N, Kaur T. Genetic comparison among various coronavirus strains for the identification of potential vaccine targets of SARS-CoV2. *Infect Genet Evol*. 2021; 89: 104490.
6. Chakravarty D, Das Sarma J. Murine-β-coronavirus-induced neuropathogenesis sheds light on CNS pathobiology of SARS-CoV2. *J Neurovirol*. 2021; 27(2): 197-216.
7. Mendonça L, Howe A, Gilchrist JB, Sun D, Knight ML, Zanetti-Domingues LC, Bateman B, Krebs AS, Chen L, Radecke J, Sheng Y, Li VD, Ni T, Kounatidis I, MoA Koronfel, Szynkiewicz M, Harkiolaki M, Martin-Fernandez ML, James W, Zhang P. SARS-CoV-2 Assembly and Egress Pathway Revealed by Correlative Multi-modal Multi-scale Cryo-imaging. *bioRxiv*. 2020; 2020.11.05.370239.
8. Hartenian E, Nandakumar D, Lari A, Ly M, Tucker JM, Glaunsinger BA. The molecular virology of coronaviruses. *J Biol Chem*. 2020; 295(37): 12910-12934.
9. Knoops K, Kikkert M, van den Worm SHE, Zevenhoven-Dobbe JC, van der Meer Y, Koster AJ, Mommaas AM, Snijder EJ. SARS-coronavirus replication is supported by a reticulovesicular network of modified endoplasmic reticulum. *PLoS Biol*. 2008; 6(9): e226.
10. Snijder EJ, Limpens RWAL, de Wilde AH, de Jong AWM, Zevenhoven-Dobbe JC, Maier HJ, Faas FFGA, Koster AJ, Bárcena M. A unifying structural and functional model of the coronavirus replication organelle: Tracking down RNA synthesis. *PLoS Biol*. 2020; 18(6): e3000715.
11. Wolff G, Melia CE, Snijder EJ, Bárcena M. Double-Membrane Vesicles as Platforms for Viral Replication. *Trends Microbiol*. 2020; 28(12): 1022-1033.
12. Angelini MM, Akhlaghpour M, Neuman BW, Buchmeier MJ. Severe acute respiratory syndrome coronavirus nonstructural proteins 3, 4, and 6 induce double-membrane vesicles. *mBio*. 2013; 4(4): e00524-13.
13. Wolff G, Limpens RWAL, Zevenhoven-Dobbe JC, Laugks U, Zheng S, de Jong AWM, Koning RI, Agard DA, Grünewald K, Koster AJ, Snijder EJ, Bárcena M. A molecular pore spans the double membrane of the coronavirus replication organelle. *Science*. 2020; 369(6509): 1395-1398.
14. Klein S, Cortese M, Winter SL, Wachsmuth-Melm M, Neufeldt CJ, Cerikan B, Stanifer ML, Boulant S, Bartenschlager R, Chlanda P. SARS-CoV-2 structure and replication characterized by in situ cryo-electron tomography. *Nat Commun*. 2020; 11(1): 5885.
15. Qu Y, Wang X, Zhu Y, Wang W, Wang Y, Hu G, Liu C, Li J, Ren S, Xiao MZX, Liu Z,

Wang C, Fu J, Zhang Y, Li P, Zhang R, Liang Q. ORF3a-mediated incomplete autophagy facilitates severe acute respiratory syndrome Coronavirus-2 replication. *Front Cell Dev Biol*. 2021; 9: 716208.
16. Lu S, Ye Q, Singh D, Cao Y, Diedrich JK, Yates JR 3rd, Villa E, Cleveland DW, Corbett KD. The SARS-CoV-2 nucleocapsid phosphoprotein forms mutually exclusive condensates with RNA and the membrane-associated M protein. *Nat Commun*. 2021; 12(1): 502.
17. Masters PS. Coronavirus genomic RNA packaging. *Virology*. 2019; 537: 198-207.
18. Arya R, Kumari S, Pandey B, Mistry H, Bihani SC, Das A, Prashar V, Gupta GD, Panicker L, Kumar M. Structural insights into SARS-CoV-2 proteins. *J Mol Biol*. 2021; 433(2): 166725.
19. Cubuk J, Alston JJ, Incicco JJ, Singh S, Stuchell-Brereton MD, Ward MD, Zimmerman MI, Vithani N, Griffith D, Wagoner JA, Bowman GR, Hall KB, Soranno A, Holehouse AS. The SARS-CoV-2 nucleocapsid protein is dynamic, disordered, and phase separates with RNA. *Nat Commun*. 2021; 12(1): 1936.
20. Chang CK, Chen CM, Chiang MH, Hsu YL, Huang TH. Transient oligomerization of the SARS-CoV N protein – implication for virus ribonucleoprotein packaging. *PLoS One*. 2013; 8(5): e65045.
21. Hsieh PK, Chang SC, Huang CC, Le TT, Hsiao CW, Kou YH, Chen IY, Chang CK, Huang TH, Chang MF. Assembly of severe acute respiratory syndrome coronavirus RNA packaging signal into virus-like particles is nucleocapsid dependent. *J Virol*. 2005; 79(22): 13848-13855.
22. Ye Q, West AMV, Silletti S, Corbett KD. Architecture and self-assembly of the SARS-CoV-2 nucleocapsid protein. *Protein Sci*. 2020; 29(9): 1890-1901.
23. Iserman C, Roden C, Boerneke M, Sealfo R, McLaughlin G, Jungreis I, Park C, Boppana A, Fritch E, Hou YJ, Theesfeld C, Troyanskaya OG, Baric RS, Sheahan TP, Weeks K, Gladfelter S. Specific viral RNA drives the SARS-CoV-2 nucleocapsid to phase separate. *bioRxiv*. 2020; 2020.06.11.147199.
24. Narayanan K, Chen CJ, Maeda J, Makino S. Nucleocapsid-independent specific viral RNA packaging via viral envelope protein and viral RNA signal. *J Virol*. 2003; 77(5): 2922-2927.
25. Caul EO, Egglestone SI. Coronavirus-like particles present in simian faeces. *Vet Rec*. 1979; 104(8): 168-169.
26. Neuman BW, Adair BD, Yoshioka C, Quispe JD, Orca G, Kuhn P, Milligan RA, Yeager M, Buchmeier MJ. Supramolecular architecture of severe acute respiratory syndrome coronavirus revealed by electron cryomicroscopy. *J Virol*. 2006; 80(16): 7918-7928.
27. Bárcena M, Oostergetel GT, Bartelink W, Faas FG, Verkleij A, Rottier PJ, Koster AJ, Bosch BJ. Cryo-electron tomography of mouse hepatitis virus: Insights into the structure of the coronavirion. *Proc Natl Acad Sci USA*. 2009; 106(2): 582-587.
28. Yao H, Song Y, Chen Y, Wu N, Xu J, Sun C, Zhang J, Weng T, Zhang Z, Wu Z, Cheng L, Shi D, Lu X, Lei J, Crispin M, Shi Y, Li L, Li S. Molecular architecture of the SARS-CoV-2 virus. *Cell*. 2020; 183(3): 730-738.e13.
29. Peng Y, Du N, Lei Y, Dorje S, Qi J, Luo T, Gao GF, Song H. Structures of the SARS-CoV-2 nucleocapsid and their perspectives for drug design. *EMBO J*. 2020; 39(20): e105938.
30. Hardenbrook NJ, Zhang P. A structural view of the SARS-CoV-2 virus and its assembly. *Curr Opin Virol*. 2022; 52: 123-134.
31. Bai Z, Cao Y, Liu W, Li J. The SARS-CoV-2 nucleocapsid protein and its role in viral structure, biological functions, and a potential target for drug or vaccine Mitigation. *Viruses*. 2021; 13(6): 1115.
32. Carlson CR, Asfaha JB, Ghent CM, Howard CJ, Hartooni N, Safari M, Frankel AD, Morgan DO. Phosphoregulation of phase separation by the SARS-CoV-2 N protein suggests a biophysical basis for its dual functions. *Mol Cell*. 2020; 80(6): 1092-1103.e4.
33. Luo H, Wu D, Shen C, Chen K, Shen X, Jiang H. Severe acute respiratory syndrome coronavirus membrane protein interacts with nucleocapsid protein mostly through their carboxyl termini by electrostatic attraction. *Int J Biochem Cell Biol*. 2006; 38(4): 589-599.
34. Satarker S, Nampoothiri M. Structural proteins in severe acute respiratory syndrome coronavirus-2. *Arch Med Res*. 2020; 51(6): 482-491.
35. Hogue BG, Machamer CE. Coronavirus structural proteins and virus assembly. In S. Perlmann, T. Gallagher. E.J. Snijder (Eds.)

Washington DC, UAS: ASM Press. *Nidoviruse*s. 2008; 179-200.
36. Neuman BW, Kiss G, Kunding AH, Bhella D, Baksh MF, Connelly S, Droese B, Klaus JP, Makino S, Sawicki SG, Siddell SG, Stamou DG, Wilson IA, Kuhn P, Buchmeier MJ. A structural analysis of M protein in coronavirus assembly and morphology. J Struct Biol. 2011;174(1):11-22.
37. Thomas S. The structure of the membrane protein of SARS-CoV-2 resembles the sugar transporter SemiSWEET. *Pathog Immun*. 2020; 5(1): 342-363.
38. Kuo L, Hurst-Hess KR, Koetzner CA, Masters PS. Analyses of coronavirus assembly interactions with interspecies membrane and nucleocapsid protein chimeras. *J Virol*. 2016; 90(9): 4357-4368.
39. de Haan CA, Vennema H, Rottier PJ. Assembly of the coronavirus envelope: homotypic interactions between the M proteins. *J Virol*. 2000; 74(11): 4967-4978.
40. Zhang R, Li Y, Cowley TJ, Steinbrenner AD, Phillips JM, Yount BL, Baric RS, Weiss SR. The nsp1, nsp13, and M proteins contribute to the hepatotropism of murine coronavirus JHM.WU. *J Virol*. 2015; 89(7): 3598-3609.
41. Arndt AL, Larson BJ, Hogue BG. A conserved domain in the coronavirus membrane protein tail is important for virus assembly. *J Virol*. 2010; 84(21): 11418-11428.
42. Huang Y, Yang ZY, Kong WP, Nabel GJ. Generation of synthetic severe acute respiratory syndrome coronavirus pseudoparticles: implications for assembly and vaccine production. *J Virol*. 2004; 78(22): 12557-12565.
43. Yuan Z, Hu B, Xiao H, Tan X, Li Y, Tang K, Zhang Y, Cai K, Ding B. The E3 ubiquitin ligase RNF5 facilitates SARS-CoV-2 membrane protein-mediated virion release. *mBio*. 2022; 13(1): e0316821.
44. Boson B, Legros V, Zhou B, Siret E, Mathieu C, Cosset FL, Lavillette D, Denolly S. The SARS-CoV-2 envelope and membrane proteins modulate maturation and retention of the spike protein, allowing assembly of virus-like particles. *J Biol Chem*. 2021; 296: 100111.
45. Kumar B, Hawkins GM, Kicmal T, Qing E, Timm E, Gallagher T. Assembly and entry of severe acute respiratory syndrome coronavirus 2 (SARS-CoV2): evaluation using virus-like particles. *Cells*. 2021; 10(4): 853.
46. Zaloilo I, Rud Y, Zaloilo O, Buchatskyi L. Coronavirus viroporins: structure and function. *Ukr Biochem J*. 2021; 93(1): 5-17.
47. Cohen JR, Lin LD, Machamer CE. Identification of a Golgi complex-targeting signal in the cytoplasmic tail of the severe acute respiratory syndrome coronavirus envelope protein. *J Virol*. 2011; 85(12): 5794-5803.
48. Pearson GJ, Broncel M, Snijders AP, Carlton JG. Exploitation of the Secretory Pathway by SARSCoV-2 Envelope. *bioRxiv*. 2021.
49. Miserey-Lenkei S, Trajkovic K, D'Ambrosio JM, Patel AJ, Čopič A, Mathur P, Schauer K, Goud B, Albanèse V, Gautier R, Subra M, Kovacs D, Barelli H, Antonny B. A comprehensive library of fluorescent constructs of SARS-CoV-2 proteins and their initial characterisation in different cell types. *Biol Cell*. 2021; 113(7): 311-328.
50. Bracquemond D, Muriaux D. Betacoronavirus assembly: clues and perspectives for elucidating SARS-CoV-2 particle formation and egress. *mBio*. 2021; 12(5): e0237121.
51. Schoeman D, Fielding BC. Coronavirus envelope protein: current knowledge. *Virol J*. 2019; 16(1): 69.
52. Breitinger U, Farag NS, Sticht H, Breitinger HG. Viroporins: structure, function, and their role in the life cycle of SARS-CoV-2. *Int J Biochem Cell Biol*. 2022; 145: 106185.
53. Xu C, Lu P, Gamal El-Din TM, Pei XY, Johnson MC, Uyeda A, Bick MJ, Xu Q, Jiang D, Bai H, Reggiano G, Hsia Y, Brunette TJ, Dou J, Ma D, Lynch EM, Boyken SE, Huang PS, Stewart L, DiMaio F, Kollman JM, Luisi BF, Matsuura T, Catterall WA, Baker D. Computational design of transmembrane pores. *Nature*. 2020; 585(7823): 129-134.
54. Mandala VS, McKay MJ, Shcherbakov AA, Dregni AJ, Kolocouris A, Hong M. Structure and drug binding of the SARS-CoV-2 envelope protein transmembrane domain in lipid bilayers. *Nat Struct Mol Biol*. 2020; 27(12): 1202-1208.
55. Ye Y, Hogue BG. Role of the coronavirus E viroporin protein transmembrane domain in virus assembly. *J Virol*. 2007; 81(7): 3597-3607.
56. Swann H, Sharma A, Preece B, Peterson A, Eldredge C, Belnap DM , Vershinin M, Saffarian S. Minimal system for assembly of SARS-CoV-2 virus like particles. *Sci Rep*. 2020; 10(1): 21877.

57. Tseng YT, Wang SM, Huang KJ, Wang CT. SARS-CoV envelope protein palmitoylation or nucleocapid association is not required for promoting virus-like particle production. *J Biomed Sci*. 2014; 21(1): 34.
58. Saputri DS, Li S, van Eerden FJ, Rozewicki J, Xu Z, Ismanto HS, Davila A, Teraguchi S, Katoh K, Standley DM. Flexible, functional, and familiar: characteristics of SARS-CoV-2 spike protein evolution. *Front Microbiol*. 2020; 11: 2112.
59. Candido KL, Eich CR, de Fariña LO, Kadowaki MK, da Conceição Silva JL, Maller A, Simão RCG. Spike protein of SARS-CoV-2 variants: a brief review and practical implications. *Braz J Microbiol*. 2022; 53(3): 1133-1157.
60. Letko M, Marzi A, Munster V. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat Microbiol*. 2020; 5(4): 562-569.
61. Xu R, Shi M, Li J, Song P, Li N. Construction of SARS-CoV-2 Virus-Like Particles by Mammalian Expression System. *Front Bioeng Biotechnol*. 2020; 8: 862.
62. Siu YL, Teoh KT, Lo J, Chan CM, Kien F, Escriou N, Tsao SW, Nicholls JM, Altmeyer R, Peiris JSM, Bruzzone R, Nal B. The M, E, and N structural proteins of the severe acute respiratory syndrome coronavirus are required for efficient assembly, trafficking, and release of virus-like particles. *J Virol*. 2008; 82(22): 11318-11330.
63. Mortola E, Roy P. Efficient assembly and release of SARS coronavirus-like particles by a heterologous expression system. *FEBS Lett*. 2004; 576(1-2): 174-178.
64. Vennema H, Godeke GJ, Rossen JW, Voorhout WF, Horzinek MC, Opstelten DJ, Rottier PJ. Nucleocapsid-independent assembly of coronavirus-like particles by co-expression of viral envelope protein genes. *EMBO J*. 1996; 15(8): 2020-2028.
65. Kuzmin A, Orekhov P, Astashkin R, Gordeliy V, Gushchin I. Structure and dynamics of the SARS-CoV-2 envelope protein monomer. *Proteins*. 2022; 90(5): 1102-1114.
66. Klein S, Cortese M, Winter SL, Wachsmuth-Melm M, Neufeldt CJ, Cerikan B, Stanifer ML, Boulant S, Bartenschlager R, Chlanda P. SARS-CoV-2 structure and replication characterized by in situ cryo-electron tomography. *Nat Commun*. 2020; 11(1): 5885.