Replication-transcription complex of coronaviruses: functions of individual viral non-structural subunits, properties and architecture of their complexes

E.L. Mishchenko ✉, V.A. Ivanisenko

Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
✉ elmish@bionet.nsc.ru

Abstract. Coronaviruses (CoVs) belong to the subfamily Orthocoronavirinae of the family Coronaviridae. CoVs are enveloped (+) RNA viruses with unusually long genomes. Severe acute respiratory syndrome CoV (SARS-CoV), Middle East respiratory syndrome CoV (MERS-CoV), and the novel coronavirus (2019-nCoV, SARS-CoV-2) have been identified as causing global pandemics. Clinically tested vaccines are widely used to control rapidly spreading, acute, and often severe infections; however, effective drugs are still not available. The genomes of SARS-CoV-2 and SARS-CoV are approximately 80 % identical, while the genomes of SARS-CoV-2 and MERS-CoV are approximately 50 % identical. This indicates that there may be common mechanisms of coronavirus pathogenesis and, therefore, potential therapeutic targets for each virus may be the same. The enzymes and effector proteins that make up the replication-transcription complex (RTC) of coronaviruses are encoded by a large replicase gene. These enzymes and effector proteins represent promising targets for potential therapeutic drugs. The enzyme targets include papain- and 3C-like cysteine proteinases that process two large viral polyproteins, RNA-dependent RNA polymerase, RNA helicase, viral genome-modifying enzymes, and enzymes with 3′–5′ exoribonuclease or uridylate-specific endonuclease activity. Currently, there are many studies investigating the complex molecular mechanisms involved in the assembly and function of the RTC. This review will encompass current, modern studies on the properties and complexes of individual non-structural subunits of the RTC, the structures of individual coronavirus RTC subunits, domain organization and functions of subunits, protein-protein interactions, properties and architectures of subunit complexes, the effect of mutations, and the identification of mutations affecting the viability of the virus in cell culture.

Key words: non-structural proteins CoVs; subunits of replicase CoVs; replication-transcription complex of CoVs; architecture of non-structural protein complexes CoVs.

Репликационно-транскрипционный комплекс коронавирусов: функции индивидуальных вирусных неструктурных субъединиц, свойства и архитектура их комплексов

Е.Л. Мищенко ✉, В.А. Иванисенко

Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
✉ elmish@bionet.nsc.ru

Аннотация. Коронавирусы относятся к семейству Coronaviridae подсемейства Orthocoronavirinae и представляют собой оболочечные (+) РНК-вирусы с необычно длинным геномом. В настоящее время часто идентифицируются, вызывая продолжающуюся пандемию во всем мире, коронавирусы Severe Acute respiratory syndrome CoV (SARS-CoV), Middle East respiratory syndrome CoV (MERS-CoV) и новый коронавирус (2019-nCoV, SARS-CoV-2). Для сдерживания быстро распространяющейся, острой и часто тяжело протекающей инфекции широко применяют прошедшие клинические испытания вакцины, однако эффективных лекарств по-прежнему нет. Геномы SARS-CoV-2 и SARS-CoV идентичны на ~80 %, а SARS-CoV-2 и MERS-CoV – на ~50 %. Это свидетельствует об общих механизмах патогенеза коронавирусов и одних и тех же потенциальных терапевтических мишенях. Ферменты и эффекторные белки, входящие в состав репликационно-транскрипционного комплекса (РТК) коронавирусов, кодируются весьма крупным геном репликазы и представляют собой перспективные мишени действия потенциальных эффективных лекарств. Эти мишени включают папаин- и 3С-подобные цистеиновые протеиназы, осуществляющие процессинг двух больших вирусных полипротеинов, РНК-зависимую РНК-полимеразу, РНК-хеликазу, ферменты, модифицирующие вирусный геном, ферменты, обладающие 3′–5′-экзорибонуклеазной и уридилат-специфичной эндонуклеазной активностью, а также важные эффекторные белки. В настоящее время изучение сложных молекулярных механизмов сборки и функционирования

РТК находится на пике изучения. Обзор посвящен актуальным и современным исследованиям свойств индивидуальных неструктурных субъединиц РТК и их комплексов и включает изучение структур индивидуальных субъединиц РТК коронавирусов, доменной организации субъединиц и их функций, белок-белковых взаимодействий, свойств и архитектуры комплексов субъединиц, влияния мутаций, а также выявления мутаций, влияющих на жизнеспособность вируса в клеточной культуре.
Ключевые слова: неструктурные белки коронавирусов (CoVs); субъединицы репликазы CoVs; репликационно-транскрипционный комплекс CoVs; архитектура комплексов неструктурных белков CoVs.

Introduction

The 2019 coronavirus infection has spread globally, often causing severe respiratory, intestinal, and systemic illnesses. Coronaviruses (CoVs) belong to the Orthocoronavirinae subfamily of the Coronaviridae family. The subfamily is further divided into α-, β-, ϒ-, and δ-coronaviruses. Severe acute respiratory syndrome CoV (SARS-CoV), Middle East respiratory syndrome CoV (MERS-CoV), novel coronavirus (2019-nCoV, SARS-CoV-2), mouse hepatitis virus (MHV), and bovine coronavirus (BCoV) are all β-coronaviruses (Malik et al., 2020). Coronaviruses are enveloped viruses with an unusually long, single-stranded (+) RNA genome (26–32 kb). The SARS-CoV-2 genome is similar to the SARS-CoV genome (sequence identity ~80 %), while the SARS-CoV-2 and MERS-CoV genomes are less similar (sequence identity ~50 %) (Lu et al., 2020). The structure and function of proteins are preserved at levels as low as 30 % of amino acid sequence identity (Rost, 1999). This indicates that there may be common mechanisms of pathogenesis among the CoVs and, therefore, the viruses may have the same potential therapeutic targets. The 5′ proximal region of each CoV genome includes a cap, a 5′ untranslated region (UTR) and a long replicase gene encoding 16 non-structural proteins (comprising two-thirds of the genome). The 3′ regions encode structural proteins, including S (spike), E (surface), M (membrane), and N (nucleocapsid), auxiliary proteins (the number of these varies among CoVs), a 3′ UTR, and a poly(A) tract.

The replication-transcription complex (RTC) of CoVs is a complex consisting of viral and, probably, cellular proteins. The RTC produces the (+) RNA genome and a set of subgenomic CoVs RNA in infected cells. The CoV replicase gene has two overlapping open reading frames, ORF1a and ORF1b, which encode the viral components of the RTC. Expression of the gene leads to the formation of the pp1a polyprotein, which is encoded by ORF1a. A ribosomal frameshift of –1 before the ORF1a translation termination codon and ORF1b are required for the formation of the pp1ab polyprotein, which is a continuation of pp1a. Polyproteins pp1a and pp1ab are processed by two viral cysteine proteinases, papain-like proteinase PL^{pro} (PLP) and 3C-like proteinase 3CL^{pro} (M^{pro}), which results in the release of intermediate precursors and 16 mature highly conserved non-structural proteins (nsps) capable of associating with each other and being subunits of the RTC. The pp1a polyprotein includes nsp1–nsp11, while pp1ab includes all pp1a nsps, as well as nsp12–nsp16 (Nagvi et al., 2020).

Structural and functional properties of conserved non-structural RTC subunits and their complexes

The molecular mechanisms of the assembly and function of the RTC has not been studied. However, the structural and functional properties of conserved non-structural RTC subunits and their complexes have been extensively researched and are extremely important for the identification of key drug targets against CoVs:

nsp1 interacts with the 40S ribosome subunit and inhibits translation initiation of host proteins, including interferon response factors. Interaction of nsp1 with ribosomes also leads to the degradation of host RNA. Thus, nsp1 suppresses cellular defence antiviral mechanisms (Kamitani et al., 2006; Narayanan et al., 2008);

nsp2 is not part of the RTC in cell culture. However, the absence of nsp2 in cells infected with the MHVΔnsp2 or SARS-CoVΔnsp2 deletion mutants reduces the production of the virus and viral RNA (Graham et al., 2005);

nsp3 and **nsp5** are proteinases that process the pp1a and pp1ab polyproteins, resulting in the release of individual RTC components. The PLP nsp3 domain(s) process the N-proximal regions of pp1a and pp1ab. The MHV nsp3 has two domains, PL1^{pro} (PL1P) and PL2^{pro} (PL2P). The PL1P domain cleaves the nsp1/nsp2 and nsp2/nsp3 sites, while the PL2P domain cleaves the nsp3/nsp4 site (Hughes et al., 1995; Kanjanahaluethai et al., 2000). The SARS-CoV nsp3 has a single PL2P domain that cleaves all three nsp sites (Thiel et al., 2003). The SARS-CoV PLP is an intracellular immune response antagonist. PLP blocks the activation of transcription factors IRF3 and NF-κB, which induces the expression of IFN(I) and antiviral genes. It does this by indirectly inhibiting IKKi and TBK1 kinases that activate IRF3 and stabilizing IκBα, an inhibitor of NF-κB (Frieman et al., 2009). PLP also hydrolyzes elements of ubiquitin and the product of interferon-stimulating gene 15 of the ubiquitin-like protein, thereby blocking the cellular mechanism of post-translational ubiquitination and, in turn, enhancing viral replication (Daczkowski et al., 2017). However, nsp3 stabilizes the host E3 ubiquitin ligase RCHY1 through the interaction of its SUD and PLP domains with RCHY1. This activates the RCHY1-mediated degradation of p53, a cellular inhibitor of SARS-CoV replication (Ma-Lauer et al., 2016). SARS-CoV nsp3 interacts with nsp5, nsp6, nsp12, nsp13, nsp14, and nsp16 in the yeast two-hybrid (Y2H) system and is thought to serve as a scaffold for RTC assembly (Imbert et al., 2008);

nsp5 is a 3CL^{pro} (M^{pro}). M^{pro} plays a key role in the processing of pp1a and pp1ab polyproteins, cleaving the central and C-proximal regions of pp1ab at 11 highly conserved sites, which releases mature nsp4–nsp16 proteins (Ziebuhr et al., 2000; Thiel et al., 2003; Goyal B., Goyal D., 2020). M^{pro} is only active as a dimer. Self-elimination of MERS-CoV M^{pro} at the nsp4/nsp5 and nsp5/nsp6 sites occurs as a result of the ligand-induced formation of an "immature dimer" during the convergence of M^{pro} III domains within two polyproteins (Tomar et al., 2015). Structural analysis of the SARS-CoV-2 M^{pro} complexes with known antiviral inhibitors, Boceprevir (peptidomimetic NS3/4A protease of hepatitis C virus) and

Е.Л. Мищенко
В.А. Иванисенко

Свойства вирусных неструктурных субъединиц
репликационно-транскрипционного комплекса коронавирусов

2022
26•2

GC376 (inhibitor of CoV replication), revealed atomic-level interactions between M^{pro} and these inhibitors. Such studies are important for the optimization and design of effective drugs against CoVs (Fu et al., 2020);

nsp6 interacts with nsp2, nsp8, and nsp9 in the Y2H system (Brunn et al., 2007). Six of the predicted hydrophobic domains of MHV nsp6 and SARS-CoV nsp6 are transmembrane domains (Oostra et al., 2008). MHV nsp6 and SARS-CoV nsp6 are localized in the membranes of the endoplasmic reticulum (ER) and induce the formation of autophagosomes from ER membranes and activate autophagy (Cottam et al., 2011). Co-transfection of nsp3, nsp4, and nsp6 induces a change in the internal membranes of the host cell through the formation of double-membrane vesicles (DMVs), similar to the DMVs induced by SARS-CoV (Angelini et al., 2013);

nsp7 and **nsp8** interact with each other. SARS-CoV nsp7 and nsp8 co-crystallize to form the nsp7/8 hexadecameric supercomplex. The assembly of the supercomplex involves the formation of two different nsp7/8 heterodimers, D1 and D2, which differ in nsp8 conformation. D1 and D2 each dimerize to form the heterotetramers T1 and T2. The interaction of two T1 with two T2, in the order T1–T2–T1′–T2′ and with ring closure through the T1–T2′ interaction, leads to the construction of the full supercomplex. The supercomplex has a unique architecture: 16 molecules (8 nsp7 molecules and 8 nsp8 molecules) interact tightly with each other, forming a hollow cylindrical structure in which two nsp8 conformations coexist. The positive charge of the inner channel of the cylinder and its diameter (30 Å) indicates the ability of the nsp7/8 supercomplex to surround and interact with double-stranded RNA (dsRNA) (Zhai et al., 2005). The SARS-CoV nsp7/8 hexadecameric supercomplex can be formed in solution at an equimolar nsp7: nsp8 ratio (Zhai et al., 2005; Velthuis et al., 2012).

In solution, hexadecameric SARS-CoV nsp7/8 associates with dsRNA (Kd ~1.2 μM). The association of nsp7/8 with dsRNA is mediated by nsp8 and enhanced by nsp7 (Velthuis et al., 2012). On its own, SARS-CoV nsp8 possesses primer-independent RNA-dependent RNA polymerase (RdRp) activity and initiates, with low fidelity, the *de novo* synthesis of short (less than 6 nucleotides) complementary oligomers (primers) on single-stranded RNA (ssRNA) templates (Imbert et al., 2006). SARS-CoV nsp8 and nsp7/8 complex also exhibit primer-dependent RdRp activity (Velthuis et al., 2012). The FCoV (feline coronavirus) nsp7/8 complex is a 2 : 1 heterotrimer formed by the association of two nsp7 molecules and one nsp8 molecule. This complex does not form a hollow structure, either in crystalline form or in solution. FCoV nsp7/8 has primer-independent RdRp activity (Xiao et al., 2012);

nsp9 is an ssRNA-binding protein (Egloff et al., 2004; Miknis et al., 2009). Both monomeric and dimeric forms of PDCoV (porcine δ coronavirus) nsp9 and PEDV (porcine epidemic diarrhoea virus related to α coronaviruses) nsp9 (Zeng et al., 2018), as well as the dimeric form of SARS-CoV nsp9 (Miknis et al., 2009), have been found in solution during *in vitro* experiments. Studies of the crystal structures of SARS-CoV nsp9 (Egloff et al., 2004; Miknis et al., 2009), PDCoV nsp9, and PEDV nsp9 (Zeng et al., 2018) have revealed dimeric forms of nsp9. The monomer SARS-CoV nsp9 is characterized by a different structure compared to other proteins involved in the replicative complexes of RNA viruses, the features of which are similar to the structures of oligosaccharide/oligonucleotide-binding proteins (Egloff et al., 2004). Mutations affecting the dimerization of SARS-CoV nsp9 weaken its interaction with ssRNA, which is lethal for SARS-CoV replication in cell culture (Miknis et al., 2009);

nsp10 interacts with dsRNA, dsDNA, and ssRNA with micromolar affinity (Joseph et al., 2006). Crystal structure studies of SARS-CoV nsp10 showed that the monomer structure includes two zinc fingers, a new discovery among zinc finger protein structures. Motifs of zinc-binding nsp10 sequences have been identified. Twelve identical monomers form a unique spherical dodecameric architecture, which is hypothesized to be the functional form of nsp10 (Joseph et al., 2006; Su et al., 2006). Through two-hybrid analysis in mammalian cells, interactions of SARS-CoV nsp10 with nsp14 and nsp16 have been revealed (Pan et al., 2008);

nsp11 is a short peptide resulting from the cleavage of the pp1a polyprotein by the 3CL^{pro}/M^{pro} proteinase at the nsp10/nsp11 site. Nsp11 is encoded by the region of genomic RNA where the translational reading frame shift occurs (ORF1a to ORF1b). This shift results in the formation of nsp12–nsp16 proteins from the pp1ab polyprotein. SARS-CoV-2 nsp11 contains 13 amino acid residues and has a disordered conformation, the dynamics of which have been studied in the presence of lipid-membrane mimetics. In the presence of SDS micelles, the disordered conformation of nsp11 is transformed into an α-helix (Gadhave et al., 2021);

nsp14 is bifunctional. The N-terminal domain of nsp14 has 3′–5′ exoribonuclease activity (ExoN), and its C-terminal domain has (guanine-N7) methyltransferase activity (N7-MTase). ExoN corrects the low fidelity of synthesis of the complementary RNA strand by viral RdRp nsp12 and catalyzes the removal of 3′-terminal erroneous nucleotides in dsRNA. N7-MTase catalyzes the methylation of the viral RNA cap at the N7 guanine position in the presence of S-adenosylmethionine (methyl group donor). N7-MTase has an S-adenosylmethionine binding motif which recognizes the cap of viral GpppRNA and methylates guanine N7 GpppRNA to form ^{7Me}GpppRNA (cap-0). Cap-0 plays an important role in blocking the degradation of viral RNA by 5′–3′ exoribonucleases, translation initiation, and immune system control escape (Chen et al., 2009; Tahir, 2021). Mutants of the catalytic motif MERS-CoV ExoN and SARS-CoV-2 ExoN are not viable in cell culture (Ogando et al., 2020). SARS-CoV nsp10 associates with the ExoN domain of SARS-CoV nsp14, increasing the ExoN activity of nsp14 by more than 35 times without affecting its N7-MTase activity (Bouvet et al., 2012). Structural studies of the SARS-CoV nsp10/14 complex showed that one nsp10 molecule associates with the ExoN domain of nsp14, stabilizing and enhancing the activity of the ExoN. The architecture of the nsp10/14 complex has been studied and has been found to include two regions of contact between the nsp10 molecule and the ExoN domain of nsp14 (Ma et al., 2015);

nsp16 has (nucleoside-2′O) methyltransferase activity (2′O-MTase). 2′O-MTase recognizes the cap-0 of viral RNA and catalyzes the transfer of the methyl group from S-adenosylmethionine to the 2′OH group of the first nucleotide's ribose after N7-methylated guanine, resulting in the

formation of $^{7Me}GpppN_{2'OMe}$-RNA (conversion of cap-0 into cap-1). SARS-CoV nsp10 associates with nsp16 and stimulates the 2′O-MTase activity of nsp16 (Bouvet et al., 2010). Mutagenesis mapping of the surface amino acid residues of SARS-CoV nsp10 involved in nsp10–nsp14 interaction and structural analysis of the nsp10/16 complex revealed overlapping surfaces of nsp10 interacting with nsp14 and nsp16. Nsp10 can serve as a platform that recruits nsp14 or nsp16 to the RTC, stimulating the ExoN activity of nsp14 or the 2′O-MTase activity of nsp16. Therefore, nsp10 is an important regulator of the RTC. Mutations have been identified that disrupt the nsp10–nsp14 and nsp10–nsp16 interactions, some of which lead to a nonviable viral phenotype in cell culture (Bouvet et al., 2014);

nsp12, an RdRp, catalyzes the synthesis of complementary RNA strands on (+) and (–) viral RNA templates and is a key CoV RTC enzyme. CoVs nsp12 initiates *de novo* synthesis from the 3′ end of the viral genome of the full-length (–) RNA strand, as well as for subgenomic (–) RNA transcripts, which have differing 3′ end lengths. In turn, the full-length (–) RNA strands and subgenomic (–) RNA strands serve as templates for the synthesis of the new RNA genome and subgenomic (+) RNA transcripts. Subgenomic (+) RNA is important for the expression of structural and accessory proteins encoded by genes in the 3′ proximal region of the viral genome, which is inaccessible to ribosomes that translate the viral genome (Pasternak et al., 2006). Full-length recombinant SARS-CoV nsp12 associates with short (20–30 nucleotides long) dsRNA and ssRNA (Kd 0.13 and 0.1 μM, respectively), and initiates primer-dependent RNA synthesis on both homo- and heteropolymeric RNA templates of the same length (Velthuis et al., 2010). However, the recombinant SARS-CoV nsp12 does not associate with a primer-template that mimics the 3′-terminal 40 nucleotides of the SARS-CoV UTR and does not exhibit RdRp activity on this primer-template (Subissi et al., 2014).

Cryo-electron microscopy determined the structure of the monomer SARS-CoV-2 nsp12, which includes the nidovirus-specific N-terminal domain of the RdRp-associated nucleotidyl transferase (NiRAN), the interface domain, and the C-terminal RdRp domain. The C-terminal RdRp domain has a conserved right-hand-like architecture, which includes three subdomains: fingers, palm, and thumb. The active centre is formed by conservative motifs of amino acid residues localized in the palm subdomain. The motifs of the channel of entry for nucleotide triphosphates and the primer-template and the exit of the resulting RNA strand converge in the central cavity, where these motifs carry out matrix-dependent RNA synthesis have been determined (Gao et al., 2020).

The association of the nsp7/8 complex with nsp12 leads to the formation of the nsp7/8/12 complex. This complex possesses high RNA-binding capacity, polymerase activity, and processivity. It is also capable of initiating *de novo* RNA synthesis on the 3′ UTR template of the SARS-CoV genome, resulting in the elongation of the RNA product by over 300 nucleotides. For nsp7/8/12 complex-mediated initiation of processive RNA synthesis, three amino acid residues from nsp7 (K7, H36, N37) and one amino acid residue from nsp8 (K58) are required for the interaction of nsp7/8/12 with the RNA template, while four amino acid residues from nsp8 (D99, P116, P183, R190) interact with nsp12 (Subissi et al.,

2014). Moreover, nsp7/8/12 is able to associate with nsp14 to form the nsp7/8/12/14 multicomplex. This ensemble of non-structural proteins possesses high RNA polymerase activity and is involved in 5′ RNA capping; however, it does not have ExoN activity (Subissi et al., 2014). The structure of the SARS-CoV-2 and SARS-CoV nsp7/8/12 complexes was determined by cryo-electron microscopy. These complexes, including the nsp12 monomer, nsp8 monomer, and nsp7/nsp8 heterodimer, have similar architecture: the nsp8-1 subunit interacts with the RdRp finger subdomain, while the nsp7 and nsp8-2 subunits interact with the RdRp thumb subdomain (Gao et al., 2020);

nsp13 possesses helicase and nucleoside-triphosphatase (NTPase) activities; nsp13 interacts with nsp7, nsp8, and nsp12 in the Y2H system (Brunn et al., 2007; Pan et al., 2008). The crystal structure of MERS-CoV nsp13 was determined. The structure includes an N-terminal zinc-binding domain (ZBD) rich in Cys/His residues that coordinate three zinc ions, as well as C-terminal helicase RecA1 and RecA2 domains which contain parallel β-chains (Hao et al., 2017). The nsp13 helicase separates dsRNA and dsDNA strands with overhanging (5–20 nucleotides long) 5′ ends; nsp13 interacts with the single-stranded 5′ end of the partial nucleic acid duplex of dsRNA and dsDNA and unwinds it in the 5′–3′ direction, using the hydrolysis energy of the 5′-deoxy- and ribonucleotide triphosphates (Ivanov et al., 2004; Adedeji et al., 2012, 2016). The RNA 5′-triphosphatase activity of nsp13 catalyzes the cleavage of the ϒ phosphate from the 5′-terminal nucleotide of RNA and is thought to be involved in the capping of viral RNA (Ivanov et al., 2004). Studies of SARS-CoV nsp13 have shown that the unwinding of DNA duplexes occurs in discrete intervals of ~9.3 base pairs (bp) at a rate of 30 intervals per second. Therefore, the unwinding speed of DNA duplexes is approximately 280 bp/s.

The helicase activity of nsp13 increases approximately 2-fold when nsp13 interacts with nsp12, suggesting the interaction of these proteins is involved in the function of the RTC (Adedeji et al., 2012). When SARS-CoV-2 nsp13 interacts with the RTC, specifically with nsp7/2nsp8/nsp12: RNA, the stable complexes 2nsp13–RTC (67 %) and nsp13–RTC (20 %), as well as the dimer (2nsp13–RTC)$_2$ (13 %) are formed. Using cryo-electron microscopy, the architecture of the dominant 2nsp13–RTC complex has been determined. This involves the interaction of the ZBD on the first nsp13 molecule with the N-terminus of nsp8b and the nsp12 thumb subdomain, as well as the interaction of the ZBD of the second nsp13 molecule with the N-terminus of nsp8a. The catalytic RecA1 helicase domain of the first nsp13 molecule is fixed against nsp7 and the nsp8b head (Chen et al., 2020);

nsp15 is a nidovirus uridylate-specific endoribonuclease (NendoU). It cleaves RNA at the 3′ uridylate in unpaired, single-stranded, and looped regions (Bhardwaj et al., 2006; Zang et al., 2018). The crystal structures of SARS-CoV-2, SARS-CoV, and MERS-CoV nsp15 are homologous, functionally active hexamers formed by the dimerization of trimers. Each of the hexamer protomers includes three domains: the N-terminal, middle, and C-terminal catalytic NendoU domains. During trimer assembly, the N-terminal domain of one protomer is packed into a gap between the central and C-terminal domains of the neighbouring protomer. During

Е.Л. Мищенко
В.А. Иванисенко

Свойства вирусных неструктурных субъединиц
репликационно-транскрипционного комплекса коронавирусов

2022
26 • 2

assembly of the hexamer, the N-terminal domains of the protomers of the two trimers are packed back-to-back and are in the centre of the hexamer structure. The C-terminal domains containing active centres are located outward at the vertices of the cloverleaf. This architecture provides nsp15 with six functionally active centres (Zang et al., 2018; Kim et al., 2020).

The uridylate content in RNA, Mn^{+2} and, to a lesser extent, Mg^{+2} increase the affinity of nsp15 for RNA. The hexameric structure of SARS-CoV and MERS-CoV nsp15 is critical for its substrate and catalytic activity. The study of a series of mono-, tri- and hexameric protein mutants revealed weak interactions with RNA (rU_{16-20}) and low catalytic activity in monomers and trimers compared to hexamers and wild-type nsp15 (Bhardwaj et al., 2006; Zang et al., 2018). Nsp8 and the nsp7/8 complex interact with MERS-CoV nsp15, enhancing the binding ability of the nsp15 hexamer for RNA and its catalytic activity (Zang et al., 2018). The NendoU activity of nsp15 is an antagonist of the IFN-induced cellular antiviral response and stimulates the initiation of viral RNA translation (Deng et al., 2018, 2019).

A large number of non-structural subunits and their respective complexes within the RTC is a defining feature of CoVs. An important feature is that some subunits have domains with different enzymatic activities. For example, nsp14 has both ExoN and N7-MTase activity, while nsp13 has helicase and nucleoside triphosphatase activity. Non-structural proteins have a set of activities universal for (+) RNA viruses: proteinases (nsp3, nsp5), RNA-dependent RNA polymerases (nsp12), RNA helicases (nsp13). There are also unique domains involved in mRNA capping, cap modification (nsp14, nsp16). Some proteins possess 3′–5′ exoribonuclease activity (nsp14), which regulate the reliability of RNA genome replication, while some possess uridylate-specific endonuclease activity (nsp15). Many proteins serve as cofactors for important enzymes (nsp7, nsp 8, nsp10) and affect cellular processes. In particular, some proteins suppress the antiviral cellular response (nsp1, nsp3, nsp6, nsp15). The nsp9 and nsp10 structures are unique among the protein structures of the replicative complexes of RNA viruses. Many CoVs subunits and their complexes have a complex architecture. For example, nsp7 and nsp8 form a functional and unique hexadecameric supercomplex, the nsp10 architecture includes 12 identical subunits, and the nsp15 architecture is a functionally active hexamer. The complex architecture of the RTC is defined by the 2 nsp13/2 nsp8/1 nsp7/1 nsp12 model.

Conclusion
The lack of effective drugs against the novel coronavirus infection is a current global challenge. The RTC of CoVs replicates (+) RNA and determines the production of the virus in infected cells; however, the molecular mechanisms of this remain unexplored. Currently, great efforts are being undertaken aimed at creating a structural or functional network of interactions within the CoV proteome, as well as its interactions with the host cell. This would identify a large-scale panel of therapeutic targets. The determination of the structures of individual non-structural RTC subunits and their complexes and the identification of key interacting amino acid residues and types of bonds between them will enable the design of selective and effective inhibitors. Investigations employing

biochemical methods and mutational analyses can identify factors that affect the efficiency of viral genomic RNA production and the virus in an infected cell, the multiple effects of viral proteins on the host cell, and potential key drug targets.

References
Adedeji A.O., Lazarus H. Biochemical characterization of Middle East respiratory syndrome coronavirus helicase. *mSphere*. 2016;1(5): e00235-16. DOI 10.1128/mSphere.00235-16.

Adedeji A.O., Marchand B., Velthuis A.J.W., Snijder E.J., Weiss S., Eoff R.L., Singh K., Sarafianos S.G. Mechanism of nucleic acid unwinding by SARS-CoV helicase. *PLoS One*. 2012;7(5):e36521. DOI 10.1371/journal.pone.0036521.

Angelini M.M., Akhlaghpour M., Neuman B.W., Buchmeier M.J. Severe acute respiratory syndrome coronavirus nonstructural proteins 3, 4, and 6 induce double-membrane vesicles. *mBio*. 2013; 4(4):e00524-13. DOI 10.1128/mBio.00524-13.

Bhardwaj K., Sun J., Holzenburg A., Guarino L.A., Kao C.C. RNA recognition and cleavage by the SARS coronavirus endoribonuclease. *J. Mol. Biol*. 2006;361(2):243-256. DOI 10.1016/j.jmb.2006. 06.021.

Bouvet M., Debarnot C., Imbert I., Selisko B., Snijder E.J., Canard B., Decroly E. *In vitro* reconstitution of SARS-coronavirus mRNA cap methylation. *PLoS Pathog*. 2010;6(4):e1000863. DOI 10.1371/journal.ppat.1000863.

Bouvet M., Imbert I., Subissi L., Gluais L., Canard B., Decroly E. RNA 3′-end mismatch excision by the severe acute respiratory syndrome coronavirus nonstructural protein nsp10/nsp14 exoribonuclease complex. *Proc. Natl. Acad. Sci. USA*. 2012;109(24):9372-9377. DOI 10.1073/pnas.1201130109.

Bouvet M., Lugari A., Posthuma C.C., Zevenhoven J.C., Bernard S., Betzi S., Imbert I., Canard B., Guillemot J.-C., Lécine P., Pfefferle S., Drosten C., Snijder E.J., Decroly E., Morelli X. Coronavirus Nsp10, a critical co-factor for activation of multiple replicative enzymes. *J. Biol. Chem*. 2014;289(37):25783-25796. DOI 10.1074/jbc.M114.577353.

Brunn A., Teepe C., Simpson J.C., Pepperkok R., Friedel C.C., Zimmer R., Roberts R., Baric R., Haas J. Analysis of intraviral protein-protein interactions of the SARS coronavirus ORFeome. *PLoS One*. 2007;2(5):e459. DOI 10.1371/journal.pone.0000459.

Chen J., Malone B., Llewellyn E., Grasso M., Shelton P.M.M., Olinares P.D.B., Maruthi K., Eng E.T., Vatandaslar H., Chait B.T., Kapoor T.M., Darst S.A., Campbell E.A. Structural basis for helicase-polymerase coupling in the SARS-CoV-2 replication-transcription complex. *Cell*. 2020;182(6):1560-1573.e13. DOI 10.1016/j.cell.2020.07.033.

Chen Y., Cai H., Pan J., Xiang N., Tien P., Ahola T., Guo D. Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase. *Proc. Natl. Acad. Sci. USA*. 2009;106(9):3484-3489. DOI 10.1073/pnas.0808790106.

Cottam E.M., Maier H.J., Manifava M., Vaux L.C., Chandra-Schoenfelder P., Gerner W., Britton P., Ktistakis N.T., Wileman T. Coronavirus nsp6 proteins generate autophagosomes from the endoplasmic reticulum via an omegasome intermediate. *Autophagy*. 2011;7(11): 1335-1347. DOI 10.4161/auto.7.11.16642.

Daczkowski C.M., Dzimianski J.V., Clasman J.R., Goodwin O., Mesecar A.D., Pegan S.D. Structural insights into the interaction of coronavirus papain-like proteases and interferon-stimulated gene product 15 from different species. *J. Mol. Biol*. 2017;429(11):1661-1683. DOI 10.1016/j.jmb.2017.04.011.

Deng X., Baker S.C. An "Old" protein with a new story: Coronavirus endoribonuclease is important for evading host antiviral defenses. *Virology*. 2018;517:157-163. DOI 10.1016/j.virol.2017.12.024.

Deng X., Geelen A., Buckley A.C., O'Brien A., Pillatzki A., Lager K.M., Faaberg K.S., Baker S.C. Coronavirus endoribonuclease

activity in porcine epidemic diarrhea virus suppresses type I and type III interferon responses. *J. Virol.* 2019;93(8):e02000-18. DOI 10.1128/JVI.02000-18.

Egloff M.-P., Ferron F., Campanacci V., Longhi S., Rancurel C., Dutartre H., Snijder E.J., Gorbalenya A.E., Cambillau C., Canard B. 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(11):3792-3796. DOI 10.1073/pnas.0307877101.

Frieman M., Ratia K., Johnston R.E., Mesecar A.D., Baric R.S. Severe acute respiratory syndrome coronavirus papain-like protease ubiquitin-like domain and catalytic domain regulate antagonism of IRF3 and NF-kappaB signaling. *J. Virol.* 2009;83(13):6689-6705. DOI 10.1128/JVI.02220-08.

Fu L., Ye F., Feng Y., Yu F., Wang Q., Wu Y., Zhao C., Sun H., Huang B., Niu P., Song H., Shi Y., Li X., Tan W., Qi J., Gao G.F. Both Boceprevir and GC376 efficaciously inhibit SARS-CoV-2 by targeting its main protease. *Nat. Commun.* 2020;11(1):4417. DOI 10.1038/s41467-020-18233-x.

Gadhave K., Kumar P., Kumar A., Bhardwaj T., Garg N., Giri R. Conformational dynamics of 13 amino acids long NSP11 of SARS-CoV-2 under membrane mimetics and different solvent conditions. *Microb. Pathog*. 2021;158:105041. DOI 10.1016/j.micpath.2021.105041.

Gao Y., Yan L., Huang Y., Liu F., Zhao Y., Cao L., Wang T., Sun Q., Ming Z., Zhang L., Ge J., Zheng L., Zhang Y., Wang H., Zhu Y., Zhu C., Hu T., Hua T., Zhang B., Yang X., Li J., Yang H., Liu Z., Xu W., Guddat L.W., Wang Q., Lou Z., Rao Z. Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Science*. 2020;368(6492):779-782. DOI 10.1126/science.abb7498.

Goyal B., Goyal D. Targeting the dimerization of the main protease of coronaviruses: a potential broad-spectrum therapeutic strategy. *ACS Comb. Sci.* 2020;22(6):297-305. DOI 10.1021/acscombsci.0c00058.

Graham R.L., Sims A.C., Brockway S.M., Baric R.S., Denison M.R. The nsp2 replicase proteins of murine hepatitis virus and severe acute respiratory syndrome coronavirus are dispensable for viral replication. *J. Virol*. 2005;79(21):13399-13411. DOI 10.1128/JVI.79.21.13399-13411.2005.

Hao W., Wojdyla J.A., Zhao R., Han R., Das R., Zlatev I., Manoharan M., Wang M., Cui S. Crystal structure of Middle East respiratory syndrome coronavirus helicase. *PLoS Pathog.* 2017;13(6):e1006474. DOI 10.1371/journal.ppat.1006474.

Hughes S.A., Bonilla P.J., Weiss S.R. Identification of the murine coronavirus p28 cleavage site. *J. Virol*. 1995;69(2):809-813. DOI 10.1128/JVI.69.2.809-813.1995.

Imbert I., Guillemot J.-C., Bourhis J.-M., Bussetta C., Coutard B., Egloff M.-P., Ferron F., Gorbalenya A.E., Canard B. A second, non-canonical RNA-dependent RNA polymerase in SARS coronavirus. *EMBO J.* 2006;25(20):4933-4942. DOI 10.1038/sj.emboj.7601368.

Imbert I., Snijder E.J., Dimitrova M., Guillemot J.-C., Lécine P., Canard B. The SARS-Coronavirus PLnc domain of nsp3 as a replication/transcription scaffolding protein. *Virus Res*. 2008;133(2):136-148. DOI 10.1016/j.virusres.2007.11.017.

Ivanov K.A., Thiel V., Dobbe J.C., Meer Y., Snijder E.J., Ziebuhr J. Multiple enzymatic activities associated with severe acute respiratory syndrome coronavirus helicase. *J. Virol.* 2004;78(11):5619-5632. DOI 10.1128/JVI.78.11.5619-5632.2004.

Joseph J.S., Saikatendu K.S., Subramanian V., Neuman B.W., Brooun A., Griffith M., Moy K., Yadav M.K., Velasquez J., Buchmeier M.J., Stevens R.C., Kuhn P. Crystal structure of nonstructural protein 10 from the severe acute respiratory syndrome coronavirus reveals a novel fold with two zinc-binding motifs. *J. Virol.* 2006;80(16):7894-7901. DOI 10.1128/JVI.00467-06.

Kamitani W., Narayanan K., Huang C., Lokugamage K., Ikegami T., Ito N., Kubo H., Makino S. Severe acute respiratory syndrome coronavirus nsp1 protein suppresses host gene expression by promoting host mRNA degradation. *Proc. Natl. Acad. Sci. USA*. 2006; 103(34):12885-12890. DOI 10.1073/pnas.0603144103.

Kanjanahaluethai A., Baker S.C. Identification of mouse hepatitis virus papain-like proteinase 2 activity. *J. Virol.* 2000;74(17):7911-7921. DOI 10.1128/jvi.74.17.7911-7921.2000.

Kim Y., Jedrzejczak R., Maltseva N.I., Wilamowski M., Endres M., Godzik A., Michalska K., Joachimiak A. Crystal structure of Nsp15 endoribonuclease NendoU from SARS-CoV-2. *Protein Sci.* 2020; 29(7):1596-1605. DOI 10.1002/pro.3873.

Lu R., Zhao X., Li J., Niu P., Yang B., Wu H., Wang W., Song H., Huang B., Zhu N., Bi Y., Ma X., Zhan F., Wang L., Hu T., Zhou H., Hu Z., Zhou W., Zhao L., Chen J., Meng Y., Wang J., Lin Y., Yuan J., Xie Z., Ma J., Liu W.J., Wang D., Xu W., Holmes E.C., Gao G.F., Wu G., Chen W., Shi W., Tan W. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet*. 2020;395(10224):565-574. DOI 10.1016/S0140-6736(20)30251-8.

Ma Y., Wu L., Shaw N., Gao Y., Wang J., Sun Y., Lou Z., Yan L., Zhang R., Rao Z. Structural basis and functional analysis of the SARS coronavirus nsp14–nsp10 complex. *Proc. Natl. Acad. Sci. USA*. 2015;112(30):9436-9441. DOI 10.1073/pnas.1508686112.

Ma-Lauer Y., Carbajo-Lozoya J., Hein M.Y., Müller M.A., Deng W., Lei J., Meyer B., Kusov Y., von Brunn B., Bairad D.R., Hünten S., Drosten C., Hermeking H., Leonhardt H., Mann M., Hilgenfeld R., von Brunn A. p53 down-regulates SARS coronavirus replication and is targeted by the SARS-unique domain and PL^{pro} via E3 ubiquitin ligase RCHY1. *Proc. Natl. Acad. Sci. USA*. 2016;113(35):E5192-E5201. DOI 10.1073/pnas.1603435113.

Malik Y.S., Sircar S., Bhat S., Sharun K., Dhama K., Dadar M., Tiwari R., Chaicumpa W. Emerging novel coronavirus (2019-nCoV) – current scenario, evolutionary perspective based on genome analysis and recent developments. *Vet. Q.* 2020;40(1):68-76. DOI 10.1080/01652176.2020.1727993.

Miknis Z.J., Donaldson E.F., Umland T.C., Rimmer R.A., Baric R.S., Schultz L.W. Severe acute respiratory syndrome coronavirus nsp9 dimerization is essential for efficient viral growth. *J. Virol.* 2009; 83(7):3007-3018. DOI 10.1128/JVI.01505-08.

Naqvi A.A.T., Fatima K., Mohammad T., Fatima U., Singh I.K., Singh A., Atif S.M., Hariprasad G., Hasan G.M., Hassan M.I. Insights into SARS-CoV-2 genome, structure, evolution, pathogenesis and therapies: Structural genomics approach. *Biochim. Biophys. Acta Mol. Basis Dis.* 2020;1866(10):165878. DOI 10.1016/j.bbadis.2020.165878.

Narayanan K., Huang C., Lokugamage K., Kamitani W., Ikegami T., Tseng C.-T.K., Makino S. Severe acute respiratory syndrome coronavirus nsp1 suppresses host gene expression, including that of type I interferon, in infected cells. *J. Virol.* 2008;82(9):4471-4479. DOI 10.1128/JVI.02472-07.

Ogando N.S., Zevenhoven-Dobbe J.C., Meer Y., Bredenbeek P.J., Posthuma C.C., Snijder E.J. The enzymatic activity of the nsp14 exoribonuclease is critical for replication of MERS-CoV and SARS-CoV-2. *J. Virol.* 2020;94(23):e01246-20. DOI 10.1128/JVI.01246-20.

Oostra M., Hagemeijer M.C., Gent M., Bekker C.P.J., Lintelo E.G., Rottier P.J.M., Haan C.A.M. Topology and membrane anchoring of the coronavirus replication complex: not all hydrophobic domains of nsp3 and nsp6 are membrane spanning. *J. Virol.* 2008;82(24): 12392-12405. DOI 10.1128/JVI.01219-08.

Pan J., Peng X., Gao Y., Li Z., Lu X., Chen Y., Ishaq M., Liu D., Dediego M.L., Enjuanes L., Guo D. Genome-wide analysis of protein-protein interactions and involvement of viral proteins in SARS-CoV replication. *PLoS One*. 2008;3(10):e3299. DOI 10.1371/journal.pone.0003299.

Pasternak A.O., Spaan W.J.M., Snijder E.J. Nidovirus transcription: how to make sense...? *J. Gen. Virol.* 2006;87(6):1403-1421. DOI 10.1099/vir.0.81611-0.

Rost B. Twilight zone of protein sequence alignments. *Protein Eng.* 1999;12(2):85-94. DOI 10.1093/protein/12.2.85.

Е.Л. Мищенко
В.А. Иванисенко

Свойства вирусных неструктурных субъединиц
репликационно-транскрипционного комплекса коронавирусов

2022
26•2

Su D., Lou Z., Sun F., Zhai Y., Yang H., Zhang R., Joachimiak A., Zhang X.C., Bartlam M., Rao1 Z. Dodecamer structure of severe acute respiratory syndrome coronavirus nonstructural protein nsp10. *J. Virol*. 2006;80(16):7902-7908. DOI 10.1128/JVI.00483-06.

Subissi L., Posthuma C.C., Collet A., Zevenhoven-Dobbe J.C., Gorbalenya A.E., Decroly E., Snijder E.J., Canard B., Imbert I. One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. *Proc. Natl. Acad. Sci. USA.* 2014;111(37):E3900-E3909. DOI 10.1073/pnas.1323705111.

Tahir M. Coronavirus genomic nsp14-ExoN, structure, role, mechanism, and potential application as a drug target. *J. Med. Virol.* 2021; 93(7):4258-4264. DOI 10.1002/jmv.27009.

Thiel V., Ivanov K.A., Putics Á., Hertzig T., Schelle B., Bayer S., Weißbrich B., Snijder E.J., Rabenau H., Doerr H.W., Gorbalenya A.E., Ziebuhr J. Mechanisms and enzymes involved in SARS coronavirus genome expression. *J. Gen. Virol*. 2003;84(9):2305-2315. DOI 10.1099/vir.0.19424-0.

Tomar S., Johnston M.L., John S.E.St., Osswald H.L., Nyalapatla P.R., Paul L.N., Ghosh A.K., Denison M.R., Mesecar A.D. Ligand-induced dimerization of Middle East respiratory syndrome (MERS) coronavirus nsp5 protease (3CLpro): implications for nsp5 regulation and the development of antivirals. *J. Biol. Chem*. 2015;290(32): 19403-19422. DOI 10.1074/jbc.M115.651463.

Velthuis A.J.W., Arnold J.J., Cameron C.E., Worm S.H.E., Snijder E.J. The RNA polymerase activity of SARS-coronavirus nsp12 is primer dependent. *Nucleic Acids Res*. 2010;38(1):203-214. DOI 10.1093/nar/gkp904.

Velthuis A.J.W., Worm S.H.E., Snijder E.J. The SARS-coronavirus nsp7+nsp8 complex is a unique multimeric RNA polymerase capable of both *de novo* initiation and primer extension. *Nucleic Acids Res.* 2012;40(4):1737-1747. DOI 10.1093/nar/gkr893.

Xiao Y., Ma Q., Restle T., Shang W., Svergun D.I., Ponnusamy R., Sczakiel G., Hilgenfeld R. Nonstructural proteins 7 and 8 of feline coronavirus form a 2:1 heterotrimer that exhibits primer-independent RNA polymerase activity. *J. Virol.* 2012;86(8):4444-4454. DOI 10.1128/JVI.06635-11.

Zeng Z., Deng F., Shi K., Ye G., Wang G., Fang L., Xiao S., Fu Z., Peng G. Dimerization of coronavirus nsp9 with diverse modes enhances its nucleic acid binding affinity. *J. Virol*. 2018;92(17): e00692-18. DOI 10.1128/JVI.00692-18.

Zhai Y., Sun F., Li X., Pang H., Xu X., Bartlam M., Rao Z. Insights into SARS-CoV transcription and replication from the structure of the nsp7–nsp8 hexadecamer. *Nat. Struct. Mol. Biol.* 2005;12(11):980-986. DOI 10.1038/nsmb999.

Zhang L., Li L., Yan L., Ming Z., Jia Z., Lou Z., Rao Z. Structural and biochemical characterization of endoribonuclease nsp15 encoded by Middle East respiratory syndrome coronavirus. *J. Virol.* 2018; 92(22):e00893-18. DOI 10.1128/JVI.00893-18.

Ziebuhr J., Snijder E.J., Gorbalenya A.E. Virus-encoded proteinases and proteolytic processing in the Nidovirales. *J. Gen. Virol.* 2000; 81(4):853-879. DOI 10.1099/0022-1317-81-4-853.