The Protein Journal (2020) 39:198–216
https://doi.org/10.1007/s10930-020-09901-4

The Proteins of Severe Acute Respiratory Syndrome Coronavirus-2 (SARS CoV-2 or n-COV19), the Cause of COVID-19

Francis K. Yoshimoto[1]

Published online: 23 May 2020

Abstract
The devastating effects of the recent global pandemic (termed COVID-19 for "coronavirus disease 2019") caused by the severe acute respiratory syndrome coronavirus-2 (SARS CoV-2) are paramount with new cases and deaths growing at an exponential rate. In order to provide a better understanding of SARS CoV-2, this article will review the proteins found in the SARS CoV-2 that caused this global pandemic.

Keywords Proteins · Virus · SARS CoV-2

1 Introduction

Severe acute respiratory syndrome coronavirus-2 (SARS CoV-2) is the virus that caused the global pandemic that was first reported [1] on December 31, 2019 [2]. Taxonomically, SARS CoV-2 belongs to the realm *Riboviria*, order *Nidovirales*, suborder *Cornidovirineae*, family *Coronaviridae*, subfamily *Orthocoronavirinae*, genus *Betacoronavirus* (lineage B), [3] subgenus *Sarbecovirus*, and the species *Severe acute respiratory syndrome-related coronavirus*.

The genome of SARS CoV-2 (NCBI Reference Sequence: NC_045512.2) [4] is similar to the genome of the coronavirus that caused the SARS epidemic in 2003 (SARS CoV, NCBI Reference sequence: NC_004718.3) [5, 6]. Much of the understanding of the proteins found in SARS CoV-2 are based on the numerous research studies reported on SARS CoV and other related viruses (e.g. MERS CoV) [7, 8]. However, among the recent coronavirus outbreaks in the new millennium (SARS CoV: 2002–2003, MERS CoV: 2012, SARS CoV-2: 2020), SARS CoV-2 mysteriously had the most devastating global impact. Understanding the proteins present in these viruses enable a more rational approach to designing more effective antiviral drugs [9, 10]. The majority of proteins of SARS CoV have been characterized in detail. The proteins of SARS CoV consist of two large polyproteins: ORF1a and ORF1ab (that proteolytically cleave to form 16 nonstructural proteins), four structural proteins: spike (S), envelope (E), membrane (M), and nucleocapsid (N), and eight accessory proteins: ORF3a, ORF3b (NP_828853.1, not present in SARS CoV-2), ORF6, ORF7a, ORF7b, ORF8a, ORF8b, and ORF9b (NP_828859.1, not present in SARS CoV-2). Although accessory proteins have been viewed as dispensable for viral replication in vitro, some have been shown to play an important role in virus-host interactions in vivo [11]. Similar to SARS CoV, SARS CoV-2 lacks the hemagglutinin esterase gene, which is found in human coronavirus (hCoV) HKU1, a lineage A betacoronavirus [3]. The spike protein, envelope protein, membrane protein, nucleocapsid protein, 3CL protease, papain like protease, RNA polymerase, [10] and helicase protein have been suggested to be viable antiviral drug targets [12]. SARS CoV-2 is an RNA virus and its RNA genome is 30 kb in length. SARS CoV-2 is thought to have originated from its closest relative, BatCov RaTG13 (GenBank: MN996532), [13] which was isolated from horseshoe bats [14].

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10930-020-09901-4) contains supplementary material, which is available to authorized users.

✉ Francis K. Yoshimoto
francis.yoshimoto@utsa.edu

[1] Department of Chemistry, The University of Texas at San Antonio (UTSA), San Antonio, TX 78249-0698, USA

2 Discussion: Proteins of SARS CoV-2

SARS CoV-2 (NC_045512.2) has a total of 11 genes with 11 open reading frames (ORFs) (Table 1): ORF1ab, ORF2 (Spike protein), ORF3a, ORF4 (Envelope protein), ORF5 (Membrane protein), ORF6, ORF7a, ORF7b, ORF8, ORF9 (Nucleocapsid protein), and ORF10.

2.1 Polyprotein Expressed by ORF1ab

The first gene (ORF1ab) expresses a polyprotein. The ORF1ab polyprotein is comprised of 16 nonstructural proteins (NSPs) (Table 2).

2.1.1 NSP1 (Leader Protein)

Nonstructural protein 1 (NSP1) is the first protein of the polyprotein of SARS CoV-2 (Fig. 1—sequence alignment

Table 1 The genes expressed by SARS CoV-2 (NC_045512.2)

Number(#)	Gene	GeneID	Location	Protein	[LOCUS]
1(7,096)	ORF1ab	43,740,578	266–21,555	ORF1ab polyprotein	[BCB15089.1/BCB97900.1]
1(4,405)	ORF1a	43,740,578	266–13,483	ORF1a polyprotein	[YP_009725295.1]
2(1,273)	ORF2 (S)	43,740,568	21,563–25,384	Spike protein (S protein)	[BCA87361.1]
3(275)	ORF3a	43,740,569	25,393–26,220	ORF3a protein	[BCA87362.1]
4(75)	ORF4 (E)	43,740,570	26,245–26,472	Envelope protein (E protein)	[BCA87363.1]
5(222)	ORF5 (M)	43,740,571	26,523–27,191	Membrane protein (M protein)	[BCA87364.1]
6(61)	ORF6	43,740,572	27,202–27,387	ORF6 protein	[BCA87365.1]
7(121)	ORF7a	43,740,573	27,394–27,759	ORF7a protein	[BCA87366.1]
8(43)	ORF7b	43,740,574	27,756–27,887	ORF7b protein	[BCB15096.1]
9(121)	ORF8	43,740,577	27,894–28,259	ORF8 protein	[BCA87367.1]
10(419)	ORF9 (N)	43,740,575	28,274–29,533	Nucleocapsid phosphoprotein (N protein)	[BCA87368.1]
11(38)	ORF10	43,740,576	29,558–29,674	ORF10 protein	[BCA87369.1]

#Represents the number of amino acids in each gene

Table 2 The nonstructural proteins (NSPs) found in the polyprotein of SARS CoV-2

#	Name	Accession	Amino acids	Proposed function
(i)	NSP1	YP_009725297.1	180 amino acids	Induce host mRNA (leader protein) cleavage
(ii)	NSP2	YP_009725298.1	638 amino acids	Binds to PHBs 1, 2
(iii)	NSP3[a]	YP_009725299.1	1945 amino acids	Release NSPs 1, 2, 3 (Papain like proteinase)
(iv)	NSP4	YP_009725300.1	500 amino acids	Membrane rearrangement
(v)	NSP5[a]	YP_009725301.1	306 amino acids	Cleaves at 11 sites of (3C-like proteinase) NSP polyprotein
(vi)	NSP6	YP_009725302.1	290 amino acids	Generates autophagosomes
(vii)	NSP7	YP_009725303.1	83 amino acids	Dimerizes with NSP8
(viii)	NSP8	YP_009725304.1	198 amino acids	Stimulates NSP12
(ix)	NSP9	YP_009725305.1	113 amino acids	Binds to helicase(?)
(x)	NSP10	YP_009725306.1	139 amino acids	Stimulates NSP16(?)
(xi)	NSP11	YP_009725312.1	13 amino acids	Unknown
(xii)	NSP12[a]	YP_009725307.1	932 amino acids	Copies viral RNA (RNA polymerase) methylation (guanine)
(xiii)	NSP13	YP_009725308.1	601 amino acids	Unwinds duplex RNA (Helicase)
(xiv)	NSP14	YP_009725309.1	527 amino acids	5′-cap RNA (3′ to 5′ exonuclease, guanine N7-methyltransferase)
(xv)	NSP15[a]	YP_009725310.1	346 amino acids	Degrade RNA to (endoRNAse/endoribonuclease) evade host defense
(xvi)	NSP16	YP_009725311.1	298 amino acids	5′-cap RNA (2′-O-ribose-methyltransferase—potential antiviral drug target) methylation (adenine)

[a]Indicates possible targets of antiviral compounds

```
NP_828860.2      1   MESLVLGVNEKTHVQLSLPVLQVRDVLVRGFGDSVEEALSEAREHLKNGTCGLVELEKGVLPQLEQPYVFIKRSDALSTN  80
YP_009725297.1   1   MESLVPGFNEKTHVQLSLPVLQVRDVLVRGFGDSVEEVLSEARQHLKDGTCGLVEVEKGVLPQLEQPYVFIKRSDARTAP  80

NP_828860.2      81  HGHKVVELVAEMDGIQYGRSGITLGVLVPHVGETPIAYRNVLLRKNGNKGAGGHSYGIDLKSYDLGDELGTDPIEDYEQN  160
YP_009725297.1   81  HGHVMVELVAELEGIQYGRSGETLGVLVPHVGEIPVAYRKVLLRKNGNKGAGGHSYGADLKSFDLGDELGTDPYEDFQEN  160

NP_828860.2      161 WNTKHGSGALRELTRELNGG  180 NSP1 of SARS CoV
YP_009725297.1   161 WNTKHSSGVTRELMRELNGG  180 NSP1 of SARS CoV-2
```

Fig. 1 Alignment of the primary amino acid sequence of NSP1 of SARS CoV (top, NP_828860.2) and SARS CoV-2 (YP_009725297.1). Sequence identity: 84.4%. Sequence similarity: 93.9%—determined using LALIGN software (and for subsequent alignments, Figs. 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, and 26, see Supporting Information for output data) [16].

of NSP1 for SARS CoV with SARS CoV-2). This protein is also known as the leader protein. This protein is also found in SARS coronavirus and is known to be a potent inhibitor of host gene expression. NSP1 binds to the 40S ribosome of the host cell to inactivate translation and promotes host mRNA degradation selectively, while the viral SARS CoV mRNA remain intact [15]. Figure 1 shows the amino acid sequence alignment for the NSP1 proteins of SARS CoV (from genome: NCBI Reference Sequence: NC_004718.3) and SARS CoV-2.

2.1.2 NSP2

Nonstructural protein 2 (NSP2) is the second protein of the polyprotein of SARS CoV-2 (Fig. 2). This protein is conserved in SARS CoV, the related beta coronavirus to SARS CoV-2. In SARS CoV, NSP2 was found to bind to two host proteins: prohibitin 1 and prohibitin 2 (PHB1 and PHB2) [17]. PHB1 and PHB2 proteins are known to play roles in cell cycle progression, cell migration, cellular differentiation, apoptosis, and mitochondrial biogenesis. The binding of NSP2 to PHB1 and PHB2 proteins suggest that NSP2 plays a role in disrupting the host cell environment.

2.1.3 NSP3 (Papain like Proteinase)

NSP3 is the papain-like proteinase protein (Fig. 3). This protein is nearly 200 kDa in size and is the largest protein (not including the polyproteins ORF1a and ORF1ab) encoded by the coronaviruses. With such a long sequence, it possesses several conserved domains: ssRNA binding, ADPr binding, G-quadruplex binding, ssRNA binding, protease (papain-like protease), and NSP4 binding), and transmembrane domain. Among the 16 nonstructural proteins, NSP3, NSP4, and NSP6 have transmembrane domains [18]. The papain like protease 1 (PL1 protease) of alpha coronavirus (alpha CoV) Transmissible Gastroenteritis Virus (TGEV), which is part of NSP3, was shown to cleave the site between NSP2 and NSP3. Furthermore, this papain like protease domain is responsible for the release of NSP1, NSP2, and NSP3 from the N-terminal region of polyproteins 1a and 1ab from coronaviruses [19]. Considering this important protease activity to release essential proteins for viral activity, the inhibition of NSP3 protease activity is an important target for antiviral activity [20]. Tanshinones, a class of natural products have been found to inhibit NSP3 protease activity.

```
NP_828861.2      1   AVTRYVDNNFCGPDGYPLDCIKDFLARAGKSMCTLSEQLDYIESKRGVYCCRDHEHEIAWFTERSDKSYEHQTPFEIKSA  80
YP_009725298.1   1   AYTRYVDNNFCGPDGYPLECIKDLLARAGKASCTLSEQLDFIDTKRGVYCCREHEHEIAWYTERSEKSYELQTPFEIKLA  80

NP_828861.2      81  KKFDTFKGECPKFVFPLNSKVKVIQPRVEKKKTEGFMGRIRSVYPVASPQECNNMHLSTLMKCNHCDEVSWQTCDFLKAT  160
YP_009725298.1   81  KKFDTFNGECPNFVFPLNSIIKTIQPRVEKKKLDGFMGRIRSVYPVASPNECNQMCLSTLMKCDHCGETSWQTGDFVKAT  160

NP_828861.2      161 CEHCGTENLVIEGPTTCGYLPTNAVVKMPCPACQDPEIGPEHSVADYHNHSNIETRLRKGGRTRCFGGCVFAYVGCYNKR  240
YP_009725298.1   161 CEFCGTENLTKEGATTCGYLPQNAVVKIYCPACHNSEVGPEHSLAEYHNESGLKTILRKGGRTIAFGGCVFSYVGCHNKC  240

NP_828861.2      241 AYWVPRASADIGSGHTGITGDNVETLNEDLLEILSRERVNINIVGDFHLNEEVAIILASFSASTSAFIDTIKSLDYKSFK  320
YP_009725298.1   241 AYWVPRASANIGCNHTGVVGEGSEGLNDNLLEILQKEKVNINIVGDFKLNEEIAIILASFSASTSAFVETVKGLDYKAFK  320

NP_828861.2      321 TIVESCGNYKVTKGKPVKGAWNIGQQRSVLTPLCGFPSQAAGVIRSIFARTLDAANHSIPDLQRAAVTILDGISEQSLRL  400
YP_009725298.1   321 QIVESCGNFKVTKGKAKKGAWNIGEQKSILSPLYAFASEAARVVRSIFSRTLETAQNSVRVLQKAAITILDGISQYSLRL  400

NP_828861.2      401 VDAMVYTSDLLTNSVIIMAYVTGGLVQQTSQWLSNLLGTTVEKLRPIFEWIEAKLSAGVEFLKDAWEILKFLITGVFDIV  480
YP_009725298.1   401 IDAMMFTSDLATNNLVVMAYITGGVVQLTSQWLTNIFGTVYEKLKPVLDWLEEKFKEGVEFLRDGWEIVKFISTCACEIV  480

NP_828861.2      481 KGQIQVASDNIKDCVKCFIDVVNKALEMCIDQVTIAGAKLRSLNLGEVFIAQSKGLYRQCIRGKEQLQLLMPLKAPKEVT  560
YP_009725298.1   481 GGQIVTCAKEIKESVQTFFKLVNKFLALCADSIIIGGAKLKALNLGETFVTHSKGLYRKCVKSREETGLLMPLKAPKEII  560

NP_828861.2      561 FLEGDSHDTVLTSEEVVLKNGELEALETPVDSFTNGAIVGTPVCVNGLMLLEIKDKEQYCALSPGLLATNNVFRLKGG 638 NSP2 SARS CoV
YP_009725298.1   561 FLEGETLPTEVLTEEVVLKTGDLQPLEQPTSEAVEAPLVGTPVCINGLMLLEIKDTEKYCALAPNMMVTNNTFTLKGG 638 NSP2 SARS CoV-2
```

Fig. 2 The primary amino acid sequence alignment of NSP2 for SARS CoV (NP_828861.2) and SARS CoV-2 (YP_009725298.1). These proteins have 68.3% sequence identity (90.0% similar)

Fig. 3 The primary amino acid sequence alignment of NSP3 for SARS CoV (NP_828862.2) and SARS CoV-2 (YP_009725299.1). Sequence identity: 76.0%, sequence similarity: 91.8%

2.1.4 NSP4 (Contains Transmembrane Domain 2)

NSP4 interacts with NSP3 and possibly host proteins to confer a role related to membrane rearrangement in SARS CoV. Moreover, the interaction between NSP4 and NSP3 is essential for viral replication [18]. The sequence alignment for NSP4 proteins for SARS CoV and SARS CoV-2 is shown in Fig. 4.

2.1.5 NSP5 (3C-like proteinase)

The NSP5 protein based on the Middle East Respiratory Syndrome (MERS) coronavirus has been characterized. NSP5 cleaves at 11 distinct sites to yield mature and intermediate nonstructural proteins (NSPs) [21]. The amino acid sequence alignment for NSP5 of SARS CoV and SARS CoV-2 is shown in Fig. 5.

2.1.6 NSP6 (Putative Transmembrane Domain)

The NSP6 protein of the avian coronavirus (infectious bronchitis virus, IBV) was shown to generate autophagosomes from the endoplasmic reticulum (ER) (Fig. 6b shows sequence alignment with SARS CoV-2 NSP6). Autophagosomes facilitate assembly of replicase proteins. Furthermore, NSP6 limited autophagosome/lysosome expansion, which in turn prevents autophagosomes from delivering viral components for degradation in lysosomes [22]. With SARS CoV, NSP6 was shown to induce membrane vesicles [23]. The amino acid sequence alignment for NSP6 of SARS CoV and SARS CoV-2 is shown in Fig. 6.

2.1.7 NSP7

NSP7 is required to form a complex with NSP8 (next section) and NSP12 to yield the RNA polymerase activity of NSP8 [24]. The primary amino acid sequence alignment

```
NP_904322.1      1   KIVSTCFKLMLKATLLCVLAALVCYIVMPVHTLSIHDGYTNEIIGYKAIQDGVTRDIISTDDCFANKHAGFDAWFSQRGG  80
YP_009725300.1   1   KIVNNWLKQLIKVTLVFLFVAAIFYLITPVHVMSKHTDFSSEIIGYKAIDGGVTRDIASTDTCFANKHADFDTWFSQRGG  80

NP_904322.1      81  SYKNDKSCPVVAAIITREIGFIVPGLPGTVLRAINGDFLHFLPRVFSAVGNICYTPSKLIEYSDFATSACVLAAECTIFK  160
YP_009725300.1   81  SYTNDKACPLIAAVITREVGFVVPGLPGTILRTTNGDFLHFLPRVFSAVGNICYTPSKLIEYTDFATSACVLAAECTIFK  160

NP_904322.1      161 DAMGKPVPYCYDTNLLEGSISYSELRPDTRYVLMDGSIIQFPNTYLEGSVRVVTTFDAEYCRHGTCERSEVGICLSTSGR  240
YP_009725300.1   161 DASGKPVPYCYDTNVLEGSVAYESLRPDTRYVLMDGSIIQFPNTYLEGSVRVVTTFDSEYCRHGTCERSEAGVCVSTSGR  240

NP_904322.1      241 WVLNNEHYRALSGVFCGVDAMNLIANIFTPLVQPVGALDVSASVVAGGIIAILVTCAAYYFMKFRRVFGEYNHVVAANAL  320
YP_009725300.1   241 WVLNNDYYRSLPGVFCGVDAVNLLTNMFTPLIQPIGALDISASIVAGGIVAIVVTCLAYYFMRFRRAFGEYSHVVAFNTL  320

NP_904322.1      321 LFLMSFTILCLVPAYSFLPGVYSVFYLYLTFYFTNDVSFLAHLQWFAMFSPIVPFWITAIYVFCISLKHCHWFFNNYLRK  400
YP_009725300.1   321 LFLMSFTVLCLTPVYSFLPGVYSVIYLYLTFYLTNDVSFLAHIQWMVMFTPLVPFWITIAYIICISTKHFYWFFSNYLKR  400

NP_904322.1      401 RVMFNGVTFSTFEEAALCTFLLNKEMYLKLRSETLLPLTQYNRYLALYNKYKYFSGALDTTSYREAACCHLAKALNDFSN  480
YP_009725300.1   401 RVVFNGVSFSTFEEAALCTFLLNKEMYLKLRSDVLLPLTQYNRYLALYNKYKYFSGAMDTTSYREAACCHLAKALNDFSN  480

NP_904322.1      481 SGADVLYQPPQTSITSAVLQ  500 NSP4 SARS CoV
YP_009725300.1   481 SGSDVLYQPPQTSITSAVLQ  500 NSP4 SARS CoV-2
```

Fig. 4 The primary amino acid sequence alignment of NSP4 for SARS CoV (NP_904322.1) and SARS CoV-2 (YP_009725300.1). Sequence identity: 80.0%, sequence similarity: 95.0%

```
NP_828863.1      1   SGFRKMAFPSGKVEGCMVQVTCGTTTLNGLWLDDTVYCPRHVICTAEDMLNPNYEDLLIRKSNHSFLVQAGNVQLRVIGH  80
YP_009725301.1   1   SGFRKMAFPSGKVEGCMVQVTCGTTTLNGLWLDDVVYCPRHVICTSEDMLNPNYEDLLIRKSNHNFLVQAGNVQLRVIGH  80

NP_828863.1      81  SMQNCLLRLKVDTSNPKTPKYKFVRIQPGQTFSVLACYNGSPSGVYQCAMRPNHTIKGSFLNGSCGSVGFNIDYDCVSFC  160
YP_009725301.1   81  SMQNCVLKLKVDTANPKTPKYKFVRIQPGQTFSVLACYNGSPSGVYQCAMRPNFTIKGSFLNGSCGSVGFNIDYDCVSFC  160

NP_828863.1      161 YMHHMELPTGVHAGTDLEGKFYGPFVDRQTAQAAGTDTTITLNVLAWLYAAVINGDRWFLNRFTTTLNDFNLVAMKYNYE  240
YP_009725301.1   161 YMHHMELPTGVHAGTDLEGNFYGPFVDRQTAQAAGTDTTITVNVLAWLYAAVINGDRWFLNRFTTTLNDFNLVAMKYNYE  240

NP_828863.1      241 PLTQDHVDILGPLSAQTGIAVLDMCAALKELLQNGMNGRTILGSTILEDEFTPFDVVRQCSGVTFQ  306 NSP5 SARS CoV
YP_009725301.1   241 PLTQDHVDILGPLSAQTGIAVLDMCASLKELLQNGMNGRTILGSALLEDEFTPFDVVRQCSGVTFQ  306 NSP5 SARS CoV-2
```

Fig. 5 The primary amino acid sequence of NSP5 for SARS CoV (NP_828863.1) and SARS CoV-2 (YP_009725301.1). Sequence identity: 96.1%, sequence similarity: 99.7%

```
NP_828864.1      1   GKFKKIVKGTHHWMLLTFLTSLLILVQSTQWSLFFFVYENAFLPFTLGIMAIAACAMLLVKHKHAFLCLFLLPSLATVAY  80
YP_009725302.1   1   SAVKRTIKGTHHWLLLTILTSLLVLVQSTQWSLFFFLYENAFLPFAMGIIAMSAFAMMFVKHKHAFLCLFLLPSLATVAY  80

NP_828864.1      81  FNMVYMPASWVMRIMTWLELADTSLSGYRLKDCVMYASALVLLILMTARTVYDDAARRVWTLMNVITLVYKVYYGNALDQ  160
YP_009725302.1   81  FNMVYMPASWVMRIMTWLDMVDTSLSGFKLKDCVMYASAVVLLILMTARTVYDDGARRVWTLMNVLTLVYKVYYGNALDQ  160

NP_828864.1      161 AISMWALVISVTSNYSGVVTTIMFLARAIVFVCVEYYPLLFITGNTLQCIMLVYCFLGYCCCCYFGLFCLLNRYFRLTLG  240
YP_009725302.1   161 AISMWALIISVTSNYSGVVTTVMFLARGIVFMCVEYCPIFFITGNTLQCIMLVYCFLGYFCTCYFGLFCLLNRYFRLTLG  240

NP_828864.1      241 VYDYLVSTQEFRYMNSQGLLPPKSSIDAFKLNIKLLGIGGKPCIKVATVQ  290 NSP6 SARS CoV
YP_009725302.1   241 VYDYLVSTQEFRYMNSQGLLPPKNSIDAFKLNIKLLGVGGKPCIKVATVQ  290 NSP6 SARS CoV-2
```

Fig. 6 Amino acid sequence alignment between the NSP6 proteins of SARS CoV (top: NP_828864.1) and SARS-CoV-2 (bottom: YP_009725302.1). Sequence identity: 88.2%, sequence similarity: 98.3%

for the NSP8 proteins for SARS CoV and SARS CoV-2 is shown in Fig. 7. Only one amino acid residue is different (arginine vs. lysine) but the charge is conserved at this location.

2.1.8 NSP8

NSP8 is a peptide cofactor that makes a heterodimer with NSP7 (the other peptide cofactor), and this NSP7-NSP8 heterodimer complexes with NSP12. In addition to the NSP7-NSP8 heterodimer, an NSP8 monomer unit also complexes with NSP12, which ultimately forms the RNA polymerase complex. The cryo-EM structure of this complex has been solved [25]. The amino acid sequence alignment for NSP8 of SARS CoV and SARS CoV-2 is shown in Fig. 8.

2.1.9 NSP9

NSP9 from the porcine reproductive and respiratory syndrome virus (PRRSV) has been found to interact with the DEAD-box RNA helicase 5 (DDX5) cellular protein [26]. This interaction between NSP9 and DDX5 has been shown to be important for viral replication—when the DDX5 gene was silenced in MARC-145 cells, the virus titers were lower by tenfold. Figure 9 shows the amino acid sequence alignment between the two NSP9 proteins from SARS CoV and SARS CoV-2.

2.1.10 NSP10

NSP10 has been shown to interact with NSP14 in SARS coronavirus, and this interaction stimulates activity of NSP14. NSP 14 is known to function as an S-adenosylmethionine (SAM)-dependent (guanine-N7) methyl transferase (N7-MTase) [27]. Furthermore, NSP10 has also been shown to stimulate the activity of NSP16, which is a 2′-O-methyltransferase [28]. Figure 10 shows the amino acid sequence alignment between the two NSP10 proteins from SARS CoV and SARS CoV-2.

2.1.11 NSP11

The function of NSP11 seems to be unknown. NSP11 is made of thirteen amino acids and the first nine amino acids (sadaqsfln) are identical to the first nine in NSP12. Figure 11 shows the amino acid sequence alignment between the two NSP12 proteins from SARS CoV and SARS CoV-2.

2.1.12 NSP12 (RNA Dependent RNA Polymerase)

NSP12 is the RNA-dependent RNA polymerase that copies viral RNA. As mentioned, NSP12 makes a complex with an NSP7-NSP8 heterodimer and an NSP8 monomer to confer

```
NP_828865.1      1   SKMSDVKCTSVVLLSVLQQLRVESSSKLWAQCVQLHNDILLAKDTTEAFEKMVSLLSVLLSMQGAVDINRLCEEMLDNRA  80
YP_009725303.1   1   SKMSDVKCTSVVLLSVLQQLRVESSSKLWAQCVQLHNDILLAKDTTEAFEKMVSLLSVLLSMQGAVDINKLCEEMLDNRA  80

NP_828865.1      81  TLQ  83 NSP7 SARS CoV
YP_009725303.1   81  TLQ  83 NSP7 SARS CoV-2
```

Fig. 7 The primary amino acid sequence alignment of NSP7 SARS CoV (NP_828865.1) and SARS CoV-2 (YP_009725303.1). Sequence identity: 98.8%, sequence similarity: 100%

```
NP_828866.1      1   AIASEFSSLPSYAAYATAQEAYEQAVANGDSEVVLKKLKKSLNVAKSEFDRDAAMQRKLEKMADQAMTQMYKQARSEDKR  80
YP_009725304.1   1   AIASEFSSLPSYAAFATAQEAYEQAVANGDSEVVLKKLKKSLNVAKSEFDRDAAMQRKLEKMADQAMTQMYKQARSEDKR  80

NP_828866.1      81  AKVTSAMQTMLFTMLRKLDNDALNNIINNARDGCVPLNIIPLTTAAKLMVVVPDYGTYKNTCDGNTFTYASALWEIQQVV  160
YP_009725304.1   81  AKVTSAMQTMLFTMLRKLDNDALNNIINNARDGCVPLNIIPLTTAAKLMVVIPDYNTYKNTCDGTTFTYASALWEIQQVV  160

NP_828866.1      161 DADSKIVQLSEINMDNSPNLAWPLIVTALRANSAVKLQ  198 NSP8 SARS CoV
YP_009725304.1   161 DADSKIVQLSEISMDNSPNLAWPLIVTALRANSAVKLQ  198 NSP8 SARS CoV-2
```

Fig. 8 The primary amino acid sequence alignment of NSP8 for SARS CoV (NP_828866.1) and SARS CoV-2 (YP_009725304.1). Sequence identity: 97.5%, sequence similarity: 100.0%

```
NP_828867.1      1   NNELSPVALRQMSCAAGTTQTACTDDNALAYYNNSKGGRFVLALLSDHQDLKWARFPKSDGTGTIYTELEPPCRFVTDTP  80
YP_009725305.1   1   NNELSPVALRQMSCAAGTTQTACTDDNALAYYNTTKGGRFVLALLSDLQDLKWARFPKSDGTGTIYTELEPPCRFVTDTP  80

NP_828867.1      81  KGPKVKYLYFIKGLNNLNRGMVLGSLAATVRLQ  113 NSP9 SARS CoV
YP_009725305.1   81  KGPKVKYLYFIKGLNNLNRGMVLGSLAATVRLQ  113 NSP9 SARS CoV-2
```

Fig. 9 The primary amino acid sequence alignment of NSP9 for SARS CoV (NP_828868.1) and SARS CoV-2 (YP_009725305.1). Sequence identity: 97.3%, sequence similarity: 99.1%

processivity of NSP12. NSP12 exhibits poor processivity in RNA synthesis—that is the presence of NSP7 and NSP8 lowers the dissociation rate of NSP12 from RNA [29]. The amino acid sequence alignment between the two NSP12 proteins from SARS CoV and SARS CoV-2 is shown in Fig. 12.

2.1.13 NSP13 (Helicase)

SARS CoV was used to characterize the helicase enzyme, NSP13, which unwinds duplex RNA [30]. The crystal structure of NSP13 of SARS CoV has been reported [31]. Furthermore, it has been shown that binding of NSP12 with NSP13 can enhance the helicase activity of NSP13. In addition to its helicase activity, NSP13 of SARS CoV is also known to possess 5′-triphosphatase activity, which is responsible for introducing the 5′-terminal cap of the viral mRNA [32]. Both eukaryotic and most viral mRNA have a 5′-terminal cap structure: m7G(5)ppp(5)N-. This 5′-terminal cap is the site of recognition for translation and plays a role in splicing, nuclear export, translation, and stability of mRNA. This process of incorporating the 5′-terminal cap will be discussed in the next section: (xiv) NSP14. The sequence alignment for NSP13 of SARS CoV and SARS CoV-2 is shown in Fig. 13. Interestingly, only

```
NP_828868.1     1   AGNATEVPANSTVLSFCAFAVDPAKAYKDYLASGGQPITNCVKMLCTHTGTGQAITVTPEANMDQESFGGASCCLYCRCH  80
YP_009725306.1  1   AGNATEVPANSTVLSFCAFAVDAAKAYKDYLASGGQPITNCVKMLCTHTGTGQAITVTPEANMDQESFGGASCCLYCRCH  80

NP_828868.1     81  IDHPNPKGFCDLKGKYVQIPTTCANDPVGFTLRNTVCTVCGMWKGYGCSCDQLREPLMQ  139 NSP10 SARS CoV
YP_009725306.1  81  IDHPNPKGFCDLKGKYVQIPTTCANDPVGFTLKNTVCTVCGMWKGYGCSCDQLREPMLQ  139 NSP10 SARS CoV-2
```

Fig. 10 The primary amino acid sequence alignment of NSP10 for SARS CoV (NP_828868.1) and SARS CoV-2 (YP_009725306.1). Sequence identity: 97.1%, sequence similarity: 99.3%

```
NP_904321.1     1   SADASTFLNGFAV  13 NSP11 SARS CoV
YP_009725312.1  1   SADAQSFLNGFAV  13 NSP11 SARS CoV-2
```

Fig. 11 The primary amino acid sequence alignment of NSP11 for SARS CoV (NP_904321.1) and SARS CoV-2 (YP_009725312.1). Sequence identity: 84.6%, sequence similarity: 100.0%

```
NP_828869.1     1   SADASTFLNRVCGVSAARLTPCGTGTSTDVVYRAFDIYNEKVAGFAKFLKTNCCRFQEKDEEGNLLDSYFVVKRHTMSNY  80
YP_009725307.1  1   SADAQSFLNRVCGVSAARLTPCGTGTSTDVVYRAFDIYNDKVAGFAKFLKTNCCRFQEKDEDDNLIDSYFVVKRHTFSNY  80

NP_828869.1     81  QHEETIYNLVKDCPAVAVHDFFKFRVDGDMVPHISRQRLTKYTMADLVYALRHFDEGNCDTLKEILVTYNCCDDDYFNKK  160
YP_009725307.1  81  QHEETIYNLLKDCPAVAKHDFFKFRIDGDMVPHISRQRLTKYTMADLVYALRHFDEGNCDTLKEILVTYNCCDDDYFNKK  160

NP_828869.1     161 DWYDFVENPDILRVYANLGERVRQSLLKTVQFCDAMRDAGIVGVLTLDNQDLNGNWYDFGDFVQVAPGCGVPIVDSYYSL  240
YP_009725307.1  161 DWYDFVENPDILRVYANLGERVRQALLKTVQFCDAMRNAGIVGVLTLDNQDLNGNWYDFGDFIQTTPGSGVPVVDSYYSL  240

NP_828869.1     241 LMPILTLTRALAAESHMDADLAKPLIKWDLLKYDFTEERLCLFDRYFKYWDQTYHPNCINCLDDRCILHCANFNVLFSTV  320
YP_009725307.1  241 LMPILTLTRALTAESHVDTDLTKPYIKWDLLKYDFTEERLKLFDRYFKYWDQTYHPNCVNCLDDRCILHCANFNVLFSTV  320

NP_828869.1     321 FPPTSFGPLVRKIFVDGVPFVVSTGYHFRELGVVHNQDVNLHSSRLSFKELLVYAADPAMHAASGNLLLDKRTTCFSVAA  400
YP_009725307.1  321 FPPTSFGPLVRKIFVDGVPFVVSTGYHFRELGVVHNQDVNLHSSRLSFKELLVYAADPAMHAASGNLLLDKRTTCFSVAA  400

NP_828869.1     401 LTNNVAFQTVKPGNFNKDFYDFAVSKGFFKEGSSVELKHFFFAQDGNAAISDYDYYRYNLPTMCDIRQLLFVVEVVDKYF  480
YP_009725307.1  401 LTNNVAFQTVKPGNFNKDFYDFAVSKGFFKEGSSVELKHFFFAQDGNAAISDYDYYRYNLPTMCDIRQLLFVVEVVDKYF  480

NP_828869.1     481 DCYDGGCINANQVIVNNLDKSAGFPFNKWGKARLYYDSMSYEDQDALFAYTKRNVIPTITQMNLKYAISAKNRARTVAGV  560
YP_009725307.1  481 DCYDGGCINANQVIVNNLDKSAGFPFNKWGKARLYYDSMSYEDQDALFAYTKRNVIPTITQMNLKYAISAKNRARTVAGV  560

NP_828869.1     561 SICSTMTNRQFHQKLLKSIAATRGATVVIGTSKFYGGWHNMLKTVYSDVETPHLMGWDYPKCDRAMPNMLRIMASLVLAR  640
YP_009725307.1  561 SICSTMTNRQFHQKLLKSIAATRGATVVIGTSKFYGGWHNMLKTVYSDVENPHLMGWDYPKCDRAMPNMLRIMASLVLAR  640

NP_828869.1     641 KHNTCCNLSHRFYRLANECAQVLSEMVMCGGSLYVKPGGTSSGDATTAYANSVFNICQAVTANVNALLSTDGNKIADKYV  720
YP_009725307.1  641 KHTTCCSLSHRFYRLANECAQVLSEMVMCGGSLYVKPGGTSSGDATTAYANSVFNICQAVTANVNALLSTDGNKIADKYV  720

NP_828869.1     721 RNLQHRLYECLYRNRDVDHEFVDEFYAYLRKHFSMMILSDDAVVCYNSNYAAQGLVASIKNFKAVLYYQNNVFMSEAKCW  800
YP_009725307.1  721 RNLQHRLYECLYRNRDVDTDFVNEFYAYLRKHFSMMILSDDAVVCFNSTYASQGLVASIKNFKSVLYYQNNVFMSEAKCW  800

NP_828869.1     801 TETDLTKGPHEFCSQHTMLVKQGDDYVYLPYPDPSRILGAGCFVDDIVKTDGTLMIERFVSLAIDAYPLTKHPNQEYADV  880
YP_009725307.1  801 TETDLTKGPHEFCSQHTMLVKQGDDYVYLPYPDPSRILGAGCFVDDIVKTDGTLMIERFVSLAIDAYPLTKHPNQEYADV  880

NP_828869.1     881 FHLYLQYIRKLHDELTGHMLDMYSVMLTNDNTSRYWEPEFYEAMYTPHTVLQ  932 NSP12 SARS CoV
YP_009725307.1  881 FHLYLQYIRKLHDELTGHMLDMYSVMLTNDNTSRYWEPEFYEAMYTPHTVLQ  932 NSP12 SARS CoV-2
```

Fig. 12 The primary amino acid sequence alignment of NSP12 for SARS CoV (NP_828869.1) and SARS CoV-2 (YP_009725307.1). Sequence identity: 96.4%, sequence similarity: 99.4%

one amino acid residue is different out of the 601 amino acids in these two proteins (isoleucine vs. valine).

2.1.14 NSP14 (3′ to 5′ Endonuclease, N7-Methyltransferase)

NSP14 from coronavirus is known to have 3′-5′ exoribonuclease activity and N7-methyltransferase activity [33]. The guanine-N7-methyltransferase activity is part of the process for introducing the 5′-cap of the virus, which involves multiple steps: [1] the gamma-phosphate of the 5′end of nascent mRNA is removed by the RNA triphosphatase (NSP13), [32], [2] a GMP moiety derived from a covalent enzyme-GMP intermediate is transferred to the resulting mRNA with a diphosphate end, [3] the GpppA cap is methylated with S-adenosyl-methionine, which is catalyzed by the guanine-N7-methyltransferase (NSP14) to yield the cap-0 structure, [34] and [4] 2′-O-methylation by NSP16 of adenine gives the cap-1 structure [35]. It is currently unknown which enzyme incorporates the GMP group involved in the second step, and it is possible that the virus uses the host guanylyltransferase enzyme [36]. Figure 14 shows the amino acid sequence alignment between the NSP14 proteins of SARS CoV and SARS CoV-2.

2.1.15 NSP15 (endoRNAse)

NSP15 of SARS coronavirus has been biochemically characterized as an endoribonuclease that cleaves RNA at uridylates at the 3′-position to form a 2′-3′ cyclic phosphodiester product [37]. The NSP15 protein specifically targets and degrades the viral polyuridine sequences to prevent the host immune sensing system from detecting the virus [38]. The crystal structure of NSP15 has been reported for SARS CoV [39]

```
NP_828870.1     1   AVGACVLCNSQTSLRCGACIRRPFLCCKCCYDHVISTSHKLVLSVNPYVCNAPGCDVTDVTQLYLGGMSYYCKSHKPPIS  80
YP_009725308.1  1   AVGACVLCNSQTSLRCGACIRRPFLCCKCCYDHVISTSHKLVLSVNPYVCNAPGCDVTDVTQLYLGGMSYYCKSHKPPIS  80

NP_828870.1     81  FPLCANGQVFGLYKNTCVGSDNVTDFNAIATCDWTNAGDYILANTCTERLKLFAAETLKATEETFKLSYGIATVREVLSD  160
YP_009725308.1  81  FPLCANGQVFGLYKNTCVGSDNVTDFNAIATCDWTNAGDYILANTCTERLKLFAAETLKATEETFKLSYGIATVREVLSD  160

NP_828870.1     161 RELHLSWEVGKPRPPLNRNYVFTGYRVTKNSKVQIGEYTFEKGDYGDAVVYRGTTTYKLNVGDYFVLTSHTVMPLSAPTL  240
YP_009725308.1  161 RELHLSWEVGKPRPPLNRNYVFTGYRVTKNSKVQIGEYTFEKGDYGDAVVYRGTTTYKLNVGDYFVLTSHTVMPLSAPTL  240

NP_828870.1     241 VPQEHYVRITGLYPTLNISDEFSSNVANYQKVGMQKYSTLQGPPGTGKSHFAIGLALYYPSARIVYTACSHAAVDALCEK  320
YP_009725308.1  241 VPQEHYVRITGLYPTLNISDEFSSNVANYQKVGMQKYSTLQGPPGTGKSHFAIGLALYYPSARIVYTACSHAAVDALCEK  320

NP_828870.1     321 ALKYLPIDKCSRIIPARARVECFDKFKVNSTLEQYVFCTVNALPETTADIVVFDEISMATNYDLSVVNARLRAKHYVYIG  400
YP_009725308.1  321 ALKYLPIDKCSRIIPARARVECFDKFKVNSTLEQYVFCTVNALPETTADIVVFDEISMATNYDLSVVNARLRAKHYVYIG  400

NP_828870.1     401 DPAQLPAPRTLLTKGTLEPEYFNSVCRLMKTIGPDMFLGTCRRCPAEIVDTVSALVYDNKLKAHKDKSAQCFKMFYKGVI  480
YP_009725308.1  401 DPAQLPAPRTLLTKGTLEPEYFNSVCRLMKTIGPDMFLGTCRRCPAEIVDTVSALVYDNKLKAHKDKSAQCFKMFYKGVI  480

NP_828870.1     481 THDVSSAINRPQIGVVREFLTRNPAWRKAVFISPYNSQNAVASKILGLPTQTVDSSQGSEYDYVIFTQTTETAHSCNVNR  560
YP_009725308.1  481 THDVSSAINRPQIGVVREFLTRNPAWRKAVFISPYNSQNAVASKILGLPTQTVDSSQGSEYDYVIFTQTTETAHSCNVNR  560

NP_828870.1     561 FNVAITRAKIGILCIMSDRDLYDKLQFTSLEIPRRNVATLQ  601 NSP13 SARS CoV
YP_009725308.1  561 FNVAITRAKVGILCIMSDRDLYDKLQFTSLEIPRRNVATLQ  601 NSP13 SARS CoV-2
```

Fig. 13 The primary amino acid sequence of NSP13 SARS CoV (NP_828870.1) and SARS CoV-2 (YP_009725308.1). Sequence identity: 99.8%, sequence similarity: 100.0%

```
NP_828871.1     1   AENVTGLFKDCSKIITGLHPTQAPTHLSVDIKFKTEGLCVDIPGIPKDMTYRRLISMMGFKMNYQVNGYPNMFITREEAI  80
YP_009725309.1  1   AENVTGLFKDCSKVITGLHPTQAPTHLSVDTKFKTEGLCVDIPGIPKDMTYRRLISMMGFKMNYQVNGYPNMFITREEAI  80

NP_828871.1     81  RHVRAWIGFDVEGCHATRDAVGTNLPLQLGFSTGVNLVAVPTGYVDTENNTEFTRVNAKPPPGDQFKHLIPLMYKGLPWN  160
YP_009725309.1  81  RHVRAWIGFDVEGCHATREAVGTNLPLQLGFSTGVNLVAVPTGYVDTPNNTDFSRVSAKPPPGDQFKHLIPLMYKGLPWN  160

NP_828871.1     161 VVRIKIVQMLSDTLKGLSDRVVFVLWAHGFELTSMKYFVKIGPERTCCLCDKRATCFSTSSDTYACWNHSVGFDYVYNPF  240
YP_009725309.1  161 VVRIKIVQMLSDTLKNLSDRVVFVLWAHGFELTSMKYFVKIGPERTCCLCDRRATCFSTASDTYACWHHSIGFDYVYNPF  240

NP_828871.1     241 MIDVQQWGFTGNLQSNHDQHCQVHGNAHVASCDAIMTRCLAVHECFVKRVDWSVEYPIIGDELRVNSACRKVQHMVVKSA  320
YP_009725309.1  241 MIDVQQWGFTGNLQSNHDLYCQVHGNAHVASCDAIMTRCLAVHECFVKRVDWTIEYPIIGDELKINAACRKVQHMVVKAA  320

NP_828871.1     321 LLADKFPVLHDIGNPKAIKCVPQAEVEWKFYDAQPCSDKAYKIEELFYSYATHHDKFTDGVCLFWNCNVDRYPANAIVCR  400
YP_009725309.1  321 LLADKFPVLHDIGNPKAIKCVPQADVEWKFYDAQPCSDKAYKIEELFYSYATHSDKFTDGVCLFWNCNVDRYPANSIVCR  400

NP_828871.1     401 FDTRVLSNLNLPGCDGGSLYVNKHAFHTPAFDKSAFTNLKQLPFFYYSDSPCESHGKQVVSDIDYVPLKSATCITRCNLG  480
YP_009725309.1  401 FDTRVLSNLNLPGCDGGSLYVNKHAFHTPAFDKSAFVNLKQLPFFYYSDSPCESHGKQVVSDIDYVPLKSATCITRCNLG  480

NP_828871.1     481 GAVCRHHANEYRQYLDAYNMMISAGFSLWIYKQFDTYNLWNTFTRLQ  527 NSP14 SARS CoV
YP_009725309.1  481 GAVCRHHANEYRLYLDAYNMMISAGFSLWVYKQFDTYNLWNTFTRLQ  527 NSP14 SARS CoV-2
```

Fig. 14 The primary amino acid sequence alignment of NSP14 of SARS CoV (NP_828871.1) and SARS CoV-2(YP_009725309.1). Sequence identity: 95.1%, sequence similarity: 99.1%

and SARS CoV-2 [40]. NSP15 uses manganese as a cofactor to promote endoribonuclease activity [41]. It has been suggested that NSP15 degrades viral dsRNA to prevent host recognition [42]. The amino acid sequence alignment of NSP15 from SARS CoV and SARS CoV-2 is shown in Fig. 15.

2.1.16 NSP16 (2′-O-Ribose-Methyltransferase)

NSP16 for coronavirus has been biochemically [43] (feline coronavirus, FCoV) and structurally [44] (complex of NSP10-NSP16 for SARS CoV) characterized. The viral RNA has a 5′-cap, which protects it from mRNA degradation by 5′-exoribonucleases, promotes mRNA translation, and prevents the viral RNA from being recognized by innate immunity mechanisms [44]. The RNA cap is an N7-methylated guanine nucleotide connected through a 5′-5′ triphosphate bridge to the first transcribed nucleotide (adenine). NSP16 methylates the 2′-hydroxy group of adenine using S-adenosylmethionine as the methyl source. Figure 16 shows the amino acid sequence alignment between the two NSP16 proteins from SARS CoV and SARS CoV-2.

2.2 Spike Protein (Surface Glycoprotein)

The spike protein (Fig. 17—sequence alignment between SARS CoV and SARS CoV-2) is a glycoprotein, which mediates attachment of the virus to the host cell. The structure of the spike (S) protein has been determined. This protein recognizes the human angiotensin-converting enzyme 2 (ACE2) protein on the host cell surface [45–47]. SARS CoV spike mouse polyclonal antibodies potently inhibited SARS CoV-2 spike protein mediated entry into cells [47]. Interestingly, a furin cleavage site (highlighted in Fig. 17: QTQTNSPRRARSVASQSIIA) was located in the S protein of SARS CoV-2, which was lacking in the S protein of SARS CoV. This difference in site could possibly explain the difference in pathogenicity of these two viruses [47].

2.3 ORF3a Protein

The ORF3a protein from SARS CoV is an ion channel protein related to NLRP3 inflammasome activation. ORF3a interacts with TRAF3, which in turn activates ASC ubiquitination, and as a result, leads to activation of caspase 1 and IL-1β maturation [48]. The amino acid sequence alignment between the two ORF3a proteins from SARS CoV and SARS CoV-2 is shown in Fig. 18.

2.4 Envelope Protein

The envelope protein is a small integral membrane protein in coronaviruses, which can oligomerize and create an ion channel [49]. The four structural proteins of coronaviruses are: S protein, M protein, E protein, and N protein [50].

```
NP_828872.1      1   SLENVAYNVVNKGHFDGHAGEAPVSIINNAVYTKVDGIDVEIFENKTTLPVNVAFELWAKRNIKPVPEIKILNNLGVDIA  80
YP_009725310.1   1   SLENVAFNVVNKGHFDGQQGEVPVSIINNTVYTKVDGVDVELFENKTTLPVNVAFELWAKRNIKPVPEVKILNNLGVDIA  80

NP_828872.1     81   ANTVIWDYKREAPAHVSTIGVCTMTDIAKKPTESACSSLTVLFDGRVEGQVDLFRNARNGVLITEGSVKGLTPSKGPAQA  160
YP_009725310.1  81   ANTVIWDYKRDAPAHISTIGVCSMTDIAKKPTETICAPLTVFFDGRVDGQVDLFRNARNGVLITEGSVKGLQPSVGPKQA  160

NP_828872.1     161  SVNGVTLIGESVKTQFNYFKKVDGIIQQLPETYFTQSRDLEDFKPRSQMETDFLELAMDEFIQRYKLEGYAFEHIVYGDF  240
YP_009725310.1  161  SLNGVTLIGEAVKTQFNYYKKVDGVVQQLPETYFTQSRNLQEFKPRSQMEIDFLELAMDEFIERYKLEGYAFEHIVYGDF  240

NP_828872.1     241  SHGQLGGLHLMIGLAKRSQDSPLKLEDFIPMDSTVKNYFITDAQTGSSKCVCSVIDLLLDDFVEIIKSQDLSVISKVVKV  320
YP_009725310.1  241  SHSQLGGLHLLIGLAKRFKESPFELEDFIPMDSTVKNYFITDAQTGSSKCVCSVIDLLLDDFVEIIKSQDLSVVSKVVKV  320

NP_828872.1     321  TIDYAEISFMLWCKDGHVETFYPKLQ  346 NSP15 SARS CoV
YP_009725310.1  321  TIDYTEISFMLWCKDGHVETFYPKLQ  346 NSP15 SARS CoV-2
```

Fig. 15 The primary amino acid sequence alignment of NSP15 of SARS CoV (NP_828872.1) and SARS CoV-2 (YP_009725310.1). Sequence identity: 88.7%, sequence similarity: 97.7%

```
NP_828873.2      1   ASQAWQPGVAMPNLYKMQRMLLEKCDLQNYGENAVIPKGIMMNVAKYTQLCQYLNTLTLAVPYNMRVIHFGAGSDKGVAP  80
YP_009725311.1   1   SSQAWQPGVAMPNLYKMQRMLLEKCDLQNYGDSATLPKGIMMNVAKYTQLCQYLNTLTLAVPYNMRVIHFGAGSDKGVAP  80

NP_828873.2     81   GTAVLRQWLPTGTLLVDSDLNDFVSDADSTLIGDCATVHTANKWDLIISDMYDPRTKHVTKENDSKEGFFTYLCGFIKQK  160
YP_009725311.1  81   GTAVLRQWLPTGTLLVDSDLNDFVSDADSTLIGDCATVHTANKWDLIISDMYDPKTKNVTKENDSKEGFFTYICGFIQQK  160

NP_828873.2     161  LALGGSIAVKITEHSWNADLYKLMGHFSWWTAFVTNVNASSSEAFLIGANYLGKPKEQIDGYTMHANYIFWRNTNPIQLS  240
YP_009725311.1  161  LALGGSVAIKITEHSWNADLYKLMGHFAWWTAFVTNVNASSSEAFLIGCNYLGKPREQIDGYVMHANYIFWRNTNPIQLS  240

NP_828873.2     241  SYSLFDMSKFPLKLRGTAVMSLKENQINDMIYSLLEKGRLIIRENNRVVVSSDILVNN  298 NSP16 SARS CoV
YP_009725311.1  241  SYSLFDMSKFPLKLRGTAVMSLKEGQINDMILSLLSKGRLIIRENNRVVISSDVLVNN  298 NSP16 SARS CoV-2
```

Fig. 16 The primary amino acid sequence alignment of NSP16 of SARS CoV (NP_828873.2) and SARS CoV-2 (YP_009725311.1). Sequence identity: 93.3%, sequence similarity: 99.0%

Fig. 17 The primary amino acid sequence alignment of the spike proteins from SARS CoV (NP_828851.1) and SARS CoV-2 (BCA87361.1). Sequence identity: 76.0%, sequence similarity: 91.5%

The E protein has been shown to play multiple roles in the viral replication cycle: [1] viral assembly, [51] [2] virion release, [52] and [3] viral pathogenesis [53]. Interestingly, in the sequence alignment of the E proteins from SARS CoV and SARS CoV-2 (Fig. 19), there is a glutamate residue (E69) with a negative charge in SARS CoV that corresponds to a positively charged arginine in SARS CoV-2 (R69).

2.5 Membrane Protein

The SARS coronavirus membrane (M) protein is an integral membrane protein that plays an important role in viral assembly [54]. In addition, the SARS coronavirus M protein has been shown to induce apoptosis [55]. The M protein interacts with the nucleocapsid (N) protein to encapsidate the RNA genome [56]. Figure 20 shows the

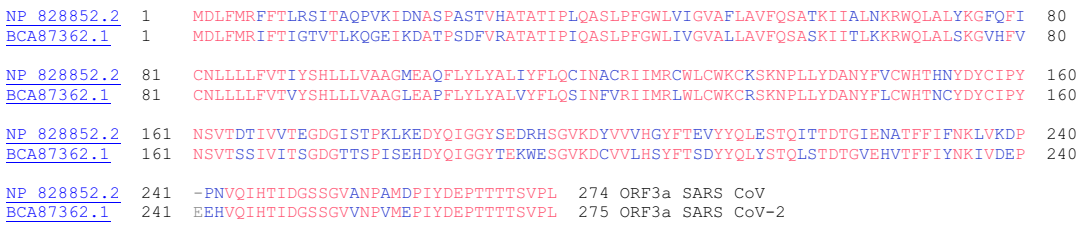

Fig. 18 The primary amino acid sequence alignment of the ORF3a proteins from SARS CoV (NP_828852.2) and SARS CoV-2 (BCA87362.1). Sequence identity: 72.4%, sequence similarity: 90.2%

```
NP_828854.1  1   MYSFVSEETGTLIVNSVLLFLAFVVFLLVTLAILTALRLCAYCCNIVNVSLVKPTVYVYSRVKNLNSSEGVPDLLV  76 SARS CoV
BCA87363.1   1   MYSFVSEETGTLIVNSVLLFLAFVVFLLVTLAILTALRLCAYCCNIVNVSLVKPSFYVYSRVKNLNSSR-VPDLLV  75 SARS CoV-2
```

Fig. 19 The primary amino acid sequence of the E proteins (ORF4) from SARS CoV (NP_828854.1) and SARS CoV-2 (BCA87363.1). Sequence identity: 94.7%, sequence similarity: 97.4%

```
NP_828855.1  1    MAD-NGTITVEELKQLLEQWNLVIGFLFLAWIMLLQFAYSNRNRFLYIIKLVFLWLLWPVTLACFVLAAVYRINWVTGGI  79
BCA87364.1   1    MADSNGTITVEELKKLLEQWNLVIGFLFLTWICLLQFAYANRNRFLYIIKLIFLWLLWPVTLACFVLAAVYRINWITGGI  80

NP_828855.1  80   AIAMACIVGLMWLSYFVASFRLFARTRSMWSFNPETNILLNVPLRGTIVTRPLMESELVIGAVIIRGHLRMAGHSLGRCD  159
BCA87364.1   81   AIAMACLVGLMWLSYFIASFRLFARTRSMWSFNPETNILLNVPLHGTILTRPLLESELVIGAVILRGHLRIAGHHLGRCD  160

NP_828855.1  160  IKDLPKEITVATSRTLSYYKLGASQRVGTDSGFAAYNRYRIGNYKLNTDHAGSNDNIALLVQ  221 M protein SARS CoV
BCA87364.1   161  IKDLPKEITVATSRTLSYYKLGASQRVAGDSGFAAYSRYRIGNYKLNTDHSSSSDNIALLVQ  222 M protein SARS CoV-2
```

Fig. 20 The primary amino acid sequence of the M proteins (ORF5) from SARS CoV (NP_828855.1) and SARS CoV-2 (BCA87364.1). Sequence identity: 90.5%, sequence similarity: 98.2%

```
NP_828856.1  1   MFHLVDFQVTIAEILIIMRTFRIAIWNLDVIISSIVRQLFKPLTKKNYSELDDEEPMELDYP  63 ORF6 SARS CoV
BCA87365.1   1   MFHLVDFQVTIAEILLIIMRTFKVSIWNLDYIINLIIKNLSKSLTENKYSQLDEEQPMEID--  61 ORF6 SARS CoV-2
```

Fig. 21 The primary amino acid sequence alignment of the ORF6 proteins from SARS CoV (NP_828856.1) and SARS CoV-2 (BCA87365.1). Sequence identity: 68.9%, sequence similarity: 93.4%

```
NP_828857.1  1    MKIILFLTLIVFTSCELYHYQECVRGTTVLLKEPCPSGTYEGNSPFHPLADNKFALTCTSTHFAFACADGTRHTYQLRAR  80
BCA87366.1   1    MKIILFLALITLATCELYHYQECVRGTTVLLKEPCSSGTYEGNSPFHPLADNKFALTCFSTQFAFACPDGVKHVYQLRAR  80

NP_828857.1  81   SVSPKLFIRQEEVQQELYSPLFLIVAALVFLILCFTIKRKTE  122 ORF7a SARS CoV
BCA87366.1   81   SVSPKLFIRQEEV-QELYSPIFLIVAAIVFITLCFTLKRKTE  121 ORF7a SARS CoV-2
```

Fig. 22 The primary amino acid sequence of the ORF7a protein from SARS CoV (NP_828857.1) and SARS CoV-2 (BCA87366.1). Sequence identity: 85.2%, sequence similarity: 95.9%

amino acid sequence alignment of the two ORF5 proteins from SARS CoV and SARS CoV-2.

2.6 ORF6 Protein

The ORF6 protein from SARS coronavirus is an accessory protein that plays an important role in viral pathogenesis [57, 58]. Using a yeast two-hybrid system, ORF6 was shown to interact with NSP8, the nonstructural protein related to promoting RNA polymerase activity [57]. Figure 21 shows the amino acid sequence alignment of the two ORF6 proteins from SARS CoV and SARS CoV-2.

2.7 ORF7a Protein

ORF7a from SARS coronavirus is an accessory protein that is a type I transmembrane protein and its crystal structure has been determined [59]. Figure 22 shows the amino acid sequence alignment between the two ORF7a proteins of SARS CoV and SARS CoV-2.

2.8 ORF7b Protein

The ORF7b accessory protein from SARS coronavirus is localized in the Golgi compartment [60]. Figure 23 shows the sequence alignment between the two ORF7b proteins of SARS CoV and SARS CoV-2.

2.9 ORF8 Protein

SARS CoV-2 has a single ORF8 protein while SARS CoV has two ORF8 proteins: ORF8a and ORF8b [61]. In SARS CoV, the ORF8b protein binds to the IRF association domain (IAD) region of interferon regulatory factor 3 (IRF3), which in turn inactivates interferon signaling [62]. Interestingly, L84S and S62L missense mutations have been reported in various SARS CoV-2 sequences [5]. Figure 24 shows the alignment between the ORF8 protein of SARS CoV-2 with the ORF8a and ORF8b proteins of SARS CoV.

2.10 Nucleocapsid Protein

The nucleocapsid (N) protein of coronaviruses is a structural protein that binds directly to viral RNA and providing stability [63]. Furthermore, the N protein of SARS CoV-2 (Fig. 24) has been found to antagonize antiviral RNAi [64]. In another study, the nucleocapsid protein of SARS CoV was found to inhibit the activity of cyclin-cyclin-dependent kinase (cyclin-CDK) complex. Inactivation of the cyclin-CDK complex results in hypophosphorylation of the retinoblastoma protein and in turn inhibits S phase (genome replication) progression in the cell cycle [65]. Figure 25 shows the amino acid sequence alignment between the two N proteins of SARS CoV and SARS CoV-2.

2.11 ORF10 Protein

ORF10 protein from SARS CoV-2 is comprised of 38-amino acids and its function is unknown. Interestingly, SARS CoV possesses an ORF9b protein (NP_828859.1), which is not present in SARS CoV-2. Figure 26 shows the sequence alignment between ORF10 of SARS CoV-2 with ORF9b of SARS CoV. SARS CoV-2 does not have an ORF10 protein. A summary of the sequence identities and similarities of the discussed proteins from SARS CoV and SARS CoV-2 is shown in Table 4.

```
NP_849175.1   1    MNELTLIDFYLCFLAFLLFLVLIMLIIFWFSLEIQDLEEPCTKV   44 ORF7b SARS CoV
BCB15096.1    1    MIELSLIDFYLCFLAFLLFLVLIMLIIFWFSLELQDHNETCHA-   43 ORF7b SARS CoV-2
```

Fig. 23 The primary amino acid sequence of the ORF7b proteins from SARS CoV (NP_849175.1) and SARS CoV-2 (BCB15096.1). Sequence identity: 85.4%, sequence similarity: 97.2%

```
NP_849176.1   1    --MKLLIVLTCISLCSCI---CTVvQRCASNKPHVLEDPCKVQH---------    ----------------------  39
NP_849177.1   1    mcLKILVRYNTRGNTYSTAWLCAL---------------GKVLPFHRWHTMVQ    TCTPnVTINCQDPAGGALIARCW  61
QJA17759.1    1    --MKFLVFLGIITTVAAFHQECSL-QSCTQHQPYVVDDPCPIHFYSKWYIRVG[31]SCLP-FTINCQEPKLGSLVVRCS  103

NP_849176.1        ----------------------          ORF8a SARS CoV
NP_849177.1   62   YLHEGHQtaaFRDVLVVLNKRtn  84      ORF8b SARS CoV
QJA17759.1    104  FYEDFLE---YHDVRVVLDFI--  121     ORF8 SARS CoV-2
```

Fig. 24 Sequence alignment of ORF8a (NP_849176.1) and ORF8b (NP_849177.1) proteins from SARS CoV (top and middle) with the ORF8 protein (QJA17759.1) from SARS CoV-2 (bottom). Sequence identity and sequence similarity between ORF8a (SARS CoV) and ORF8 (SARS CoV-2): 31.7% and 70.7% in 41 amino acid overlap. Sequence identity and sequence similarity between ORF8b (SARS CoV) and ORF8 (SARS CoV-2): 40.5% and 66.7% in 42 amino acid overlap

```
NP_828858.1   1    MSDNGPQSNQRSAPRITFGGPTDSTDNNQNGGRNGARPKQRRPQGLPNNTASWFTALTQHGKEELRFPRGQGVPINTNSG  80
BCA87368.1    1    MSDNGPQ-NQRNAPRITFGGPSDSTGSNQNGERSGARSKQRRPQGLPNNTASWFTALTQHGKEDLKFPRGQGVPINTNSS  79

NP_828858.1   81   PDDQIGYYRRATRRVRGGDGKMKELSPRWYFYYLGTGPEASLPYGANKEGIVWVATEGALNTPKDHIGTRNPNNNAATVL  160
BCA87368.1    80   PDDQIGYYRRATRRIRGGDGKMKDLSPRWYFYYLGTGPEAGLPYGANKDGIIWVATEGALNTPKDHIGTRNPANNAAIVL  159

NP_828858.1   161  QLPQGTTLPKGFYAEGSRGGSQASSRSSSRSRGNSRNSTPGSSRGNSPARMASGGGETALALLLLDRLNQLESKVSGKGQ  240
BCA87368.1    160  QLPQGTTLPKGFYAEGSRGGSQASSRSSSRSRNSSRNSTPGSSRGTSPARMAGNGGDAALALLLLDRLNQLESKMSGKGQ  239

NP_828858.1   241  QQQGQTVTKKSAAEASKKPRQKRTATKQYNVTQAFGRRGPEQTQGNFGDQDLIRQGTDYKHWPQIAQFAPSASAFFGMSR  320
BCA87368.1    240  QQQGQTVTKKSAAEASKKPRQKRTATKAYNVTQAFGRRGPEQTQGNFGDQELIRQGTDYKHWPQIAQFAPSASAFFGMSR  319

NP_828858.1   321  IGMEVTPSGTWLTYHGAIKLDDKDPQFKDNVILLNKHIDAYKTFPPTEPKKDKKKKTDEAQPLPQRQKKQPTVTLLPAAD  400
BCA87368.1    320  IGMEVTPSGTWLTYTGAIKLDDKDPNFKDQVILLNKHIDAYKTFPPTEPKKDKKKKADETQALPQRQKKQQTVTLLPAAD  399

NP_828858.1   401  MDDFSRQLQNSMSGASADSTQA  422 N protein of SARS CoV
BCA87368.1    400  LDDFSKQLQQSMS--SADSTQA  419 N protein of SARS CoV-2
```

Fig. 25 The primary amino acid sequence of the N protein from SARS CoV (ORF9a, NP_828858.1) and SARS CoV-2 (ORF9, BCA87368.1). Sequence identity: 90.5%, sequence similarity: 97.2%

```
NP_828859.1   1   MDPNQTNVVPPALHLVDPQIQLTITRMEDAMGQGQNSADPKVYPIILRLGSQLSLSMARRN-LDSLEARAFQSTPIVVQM  79
BCA87369.1    1   -----------------------------------MGYINVFAFPFTIYSLLLCRMNSRNYIAQVDVVNFNLT------  38

NP_828859.1   80  TKLATTEELPDEFVVVTAK  98 ORF9b of SARS CoV
BCA87369.1        ------------------  ORF10 of SARS CoV-2
```

Fig. 26 The primary amino acid sequence alignment of the ORF9b protein from SARS CoV and the ORF10 protein from SARS CoV-2 (Accession number: BCA87369.1). Sequence identity: 28.6%, sequence similarity: 52.4%

Table 4 Sequence identity and similarities between SARS CoV-2 proteins and SARS CoV proteins determined through LALIGN (17) (see Supporting Information)

Entry	Protein	Amino acid overlap	Sequence identity	Sequence similarity
1	NSP1	180	84.4%	93.4%
2	NSP2	638	68.3%	90.0%
3	NSP3	1,952	76.0%	91.8%
4	NSP4	500	80.0%	95.0%
5	NSP5	306	96.1%	99.7%
6	NSP6	287	88.2%	98.3%
7	NSP7	83	98.8%	100.0%
8	NSP8	198	97.5%	100.0%
9	NSP9	113	97.3%	99.1%
10	NSP10	139	97.1%	99.3%
11	NSP11	13	84.6%	100.0%
12	NSP12	932	96.4%	99.4%
13	NSP13	601	99.8%	100.0%
14	NSP14	527	95.1%	99.1%
15	NSP15	346	88.7%	97.7%
16	NSP16	298	93.3%	99.0%
17	S protein	1,277	76.0%	91.5%
18	ORF3a	1,381	72.4%	90.2%
19	E Protein	76	94.7%	97.4%
20	M Protein	222	90.5%	98.2%
21	ORF6	61	68.9%	93.4%
22	ORF7a	122	85.2%	95.9%
23	ORF7b	41	85.4%	92.7%
24a	(ORF8 vs 8a)[a]	41	31.7%	70.7%
24b	(ORF8 vs 8b)[a]	42	40.5%	66.7%
25	N Protein	422	90.5%	97.2%
26	(ORF10 vs 9b)[a]	21	28.6%	52.4%

[a](SARS CoV-2 protein vs SARS CoV protein). Other reports have also reported amino acid sequence identities using different algorithms (3,67)

3 Overlapping Genes: ORF9b and Two Proteins with Variation Among SARS CoV-2 Sequences: ORF3b and ORF9c

Overlapping genes in coronavirus have been previously observed [67]. For example, in SARS CoV, the start and end positions in the nucleotide sequence of the N-protein are 28,120 and 29,388 respectively while the ORF9b gene of SARS CoV starts and ends at positions: 28,130 and 28,426 (within the gene sequence of the N-protein) [68]. Similarly, there is a putative ORF9b protein in SARS CoV-2 located within the gene encoding the N-protein, which does not yet have an accession number [4].

In the gene alignment of 2,784 SARS CoV-2 sequences, two variations were recognized in the SARS CoV-2 genome [66]. It was recognized that a premature stop codon at position 14 of ORF3b in SARS CoV-2 in 17.6% of isolates (position E14). Furthermore, there were two mutations that gave rise to premature stop codons in ORF9c (at position Q41 in 0.7% of sequences and at position Q44 in 1.4% of the sequences). The observations of

these stop codons suggested that these genes for ORF3b and ORF9c may not be bonafide gene sequences in SARS CoV-2. With the putative SARS CoV-2 ORF3b protein, only 12 out of 57 overlapping amino acid residues were identical (21% sequence identity) to the ORF3b protein of SARS CoV [3]. In the above sections, ORF3b and ORF9c for SARS CoV-2 were not included in the above analysis. Another protein lacking an accession number is ORF14 [69].

4 Nontranslated (or Untranslated) Regions of SARS CoV-2 Genome

Considering the locations of each gene presented in Table 1, there are regions of the genome that are not translated into proteins, which is related to the non-canonical translational strategy employed by this virus [70]. The nucleotide sequences between the genes are the intergenic regions [71]. For instance, there is a conserved transctiption regulatory sequence (TRS) – a conserved hexanucleotide sequence: (5′-ACGAAC-3′) [71] that could be found in between some of the open reading frames (Table 5, Entries 2, 3, 4, 5, 7, and 9). This particular sequence has previously been identified as the leader-body fusion sites [71]. Furthermore, this sequence is a conserved motif that can be found in subgroup 2b, 2c, and 2d viruses [72]. Another transcriptional regulatory sequence was CUAAAC (e.g. Table 5, Entry 1) [73, 74].

5 Exploration of Treatment Options for COVID-19

An intense effort has been put forth to discover potential treatment options for COVID-19, the disease caused by SARS CoV-2 [75–77]. For instance, the FDA approved drug, ivermectin, is known to inhibit nuclear transport, and has been shown to inhibit the replication of SARS CoV-2 [78]. Other drugs have been repurposed and tested against COVID-19 [79, 80]. Remdesivir is a potential antiviral drug originally developed to treat ebola [81] and has been used to treat COVID-19 [82] by inhibiting viral RNA polymerase activity. Hydroxychloroquine [83] and chloroquine [84] have been used to potentially treat COVID-19. However, the use of these drugs has been known to result in cardiotoxicity [85, 86]. In fact, in a recent observational study, it was determined that hydroxychloroquine administration was not associated with a greatly lowered risk of death from COVID-19 [87].

A recent study identified 332 human proteins that interact with SARS CoV-2 proteins [66]. In this report, the predicted SARS CoV-2 proteins (NSPs 1–16 and ORFs) were expressed with 2xstreptavidin affinity tags. These tagged SARS CoV-2 proteins were expressed in human embryonic kidney (HEK)293T/17 cells and isolated the viral protein-(human protein) interactions using affinity purification-mass spectrometry. A total 332 protein–protein interactions (PPIs between SARS CoV-2 proteins and human proteins) were identified. Of these PPIs, 66

Table 5 Nontranslated RNA sequence of SARS CoV-2 (NCBI Reference Sequence: NC_045512.2)

Entry	Location (position)	Sequence
1	Beginning-ORF1ab (1–265)	1 auuaaagguu uauaccuucc cagguaacaa accaaccaac uuucgaucuc uuguagaucu 61 guucucuaaa cgaacuuuaa aaucugugug gcugucacuc ggcugcaugc uuagugcacu 121 cacgcaguau aauuaauaac uaauuacugu cguugacagg acacgaguaa cucgucuauc 181 uucugcaggc ugcuuacggu uucguccgug uugcagccga ucaucagcac aucuaggguu 241 cguccgggug ugaccgaaag guaag
2	ORF1ab-ORF2 (21,556–21,562)	1 acgaaca
3	ORF2-ORF3a (25,385–25,392)	1 acgaacuu
4	ORF3a-ORF4 (26,221–26,244)	1 gcacaagcug augaguacga acu
5	ORF4-ORF5 (26,473–26,522)	1 acgaacuaaa uauuauauua guuuuucugu uuggaacuuu aauuuuagcc
6	ORF5-ORF6 (27,192–27,201)	1 gugacaacag
7	ORF6-ORF7a (27,388–27,393)	1 acgaac
8	ORF7b-ORF8 (27,888–27,893)	1 acgaac
9	ORF8-ORF9 (28,260–28,273)	1 acgaacaaac uaaa
10	ORF9-ORF10 (29,534–29,557)	1 acucaugcag accacacaag gcag
11	ORF10-end (29,675–29,903)	1 caaucuuuaa ucagugugua acauuaggga ggacuugaaa gagccaccac auuuucaccg 61 aggccacgcg gaguacgauc gaguguacag ugaacaaugc uagggagagc ugccuauaug 121 gaagagcccu aauguguaaa auuaauuuua guagugcuau ccccauguga uuuuaauagc 181 uucuuaggag aaugacaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaa

Table 6 Drugs that potentially target (modulate) proteins that interact with SARS CoV-2 proteins as described in reference [66]

Entry	Viral Protein-(Human Gene)	Compound Name(s)
1	E protein-(BRD2/4)	JQ1,[a] RVX-208[b]
2	N protein-(CSNK2A2)	Silmitasertib (cancer),[c] TMCB[a]
3	NSP5-(HDAC2)	Apicidin,[a] Valproic acid (CNS disease, cancer)[c]
4	NSP6-(ATP6AP1)	Bafilomycin A1[a]
5	NSP6-(SIGMAR1)	E-52862,[b] PD-144418,[a] RS-PPCC,[a] PB28,[a] Haloperidol (CNS disease)[c]
6	NSP6-(SLC6A15)	Loratadine (antihistamine)[3]
7	ORF9C-(TMEM97)	PB28,[a] haloperidol (CNS disease)[c]
8	M protein-(ATP6V1A)	Bafilomycin A1[a]
9	NSP7-(COMT)	Entacapone (Parkinson's disease)[c]
10	NSP7-(PTGES2)	Indomethacin (inflammation/pain)[c]
11	NSP7-(NDUFs)	Metformin (diabetes)[c]
12	ORF9C-(NDUFs)	Metformin[c]
13	NSP12-(RIPK1)	Ponatinib (cancer)[c]
14	NSP13-(PRKACA)	H-89[a]
15	NSP14-(IMPDH2)	Merimepodib[b]
16	NSP14-(GLA)	Migalastat (Fabry disease)[c]
17	NSP14-(IMPDH2)	Mycophenolic acid (organ rejection),[3] ribavirin (virus)[c]
18	ORF8-(DNMT1)	Azacitidine[c]
19	ORF8-(LOX)	CCT 365623[a]
20	ORF9b-(MARK2/3)	Midostaurin,[3] Ruxolitinib[c]
21	ORF9b-(DCTPP1)	ZINC1775962367,[a] ZINC4326719,[a] ZINC4511851[a]
22	ORF9b/NSP13-(MARK3/TBK1)	ZINC95559591[a]
23	ORF9C-(F2RL1)	AC-55541,[a] AZ8838[a]
24	ORF9C-(ABCC1)	Daunorubicin[c]
25	ORF9C-(F2RL1)	GB110[a]
26	ORF9C-(ABCC1)	S-Verapamil (hypertension)[c]
27	ORF9C-(F2RL1)	AZ3451[a]
28	M-Protein-(SLC1A3)	UCPH-101[a]
29	E protein-(BRD2/4)	ABBV-744,[b] dBET6,[a] MZ1,[a] CPI-0610[b]
30	N protein-(LARP1)	Sapanisertib,[b] Rapamycin (organ rejection)[c]
31	NSP2-(FKBP15)	Rapamycin[c]
32	ORF8-(FKBP7/10)	Rapamycin[c]
33	NSP2-(EIF4E2/H)	Zotatifin[b]
34	ORF10-(VCP)	CB5083[b]
35	NSP6-(SIGMAR1)	Chloroquine (malaria)[c]
36	NSP9-NEK9	Dabrafenaib (cancer)[c]
37	NSP13-CEP250	WDB002[b]
38	NSP14-IMPDH2	Sanglifehrin A[a]
39	ORF8-(FKBP7)	FK-506 (organ rejection)[c]
40	ORF8-(FKBP10)	FK-506[c]
41	ORF10-(CUL2)	Pevonedistat[b]
42	ORF10-(VCP)	DBeQ, ML240[a]
43	ORF8-(PLOD1/2)	Minoxidil (hair loss)[c]
44	NSP4/9/ORF6-(NUPs RAE1)	Selinexor (cancer)[c]

Entries 1–28 were determined from chemoinformatics. Entries 29–44 were determined from specialist knowledge

[a]Pre-clinical

[b]Clinical trial

[c]FDA-approved drug. In parentheses after the drug is what the FDA-approved drug is used to treat in the clinic

of them are targetable by compounds. Table 6 shows a set of compounds that target the identified PPIs based on chemoinformatics (entries 1–28) or expertise knowledge (entries 29–44). From the subset of potential antiviral compounds that were tested, two classes of compounds were found to be effective against viral pathogenesis: [1] protein translation inhibitors (i.e. zotatifin, ternatin-4, and PS3061), and [2] Sigma1 and Sigma2 receptor ligands (i.e. approved drugs: clemastine, cloperastine, and progesterone and PB28, which was ~ 20 times more potent than hydroxychloroquine with an IC_{90} of 280 nM in the viral titer assay is undergoing pre-clinical trials for anti-cancer [88] activity).

Moreover, in another collaborative study, a library of 12,000 FDA-approved or clinical-stage drugs were tested against SARS CoV-2 infection in Vero-E6 (African green monkey kidney) cells [89]. Some effective compounds identified in the screen were: PIKfyve kinase inhibitor Apilimod, cysteine protease inhibitors (MDL-28170, Z LVG CHN2, VBY-825, and ONO 5334), and MLN-3897 (a CCR1 antagonist).

Traditional Chinese Medicine (TCM) has also been employed in China to treat COVID-19 [90]. However, due to potential toxic components present in TCM remedies, [91, 92] the use of this strategy should be handled with caution [93]. Ironically, it has been suggested that TCM could have potentially been the cause of COVID-19 [94].

In addition to small molecules, vaccines are also currently being developed against SARS CoV-2, [95] and convalescent plasma transfusions have been used to treat COVID-19 [96]. Nevertheless, more research is needed to develop effective treatments against SARS CoV-2 especially in the context of future outbreaks [97, 98].

6 Conclusion

Although there is some variation in sequence in the proteins, many of the proteins found in SARS CoV-2 (NC_045512.2) are also found in SARS CoV (AY515512.1 or NC_004718.3) with 77.1% of the protein sequences shared in their proteomes [99]. Thus, previous research on related coronavirus proteins enable a better understanding of how we can approach to understand the current coronavirus (SARS CoV-2) that caused the current global pandemic (COVID-19). The general structures of most of the proteins from SARS CoV-2 can be visualized from homology models [100]. Advances in the knowledge of the structures and functions of the proteins in SARS CoV-2 will enable researchers to design better antiviral drugs that target this virus.

Acknowledgements Francis K. Yoshimoto, PhD holds a Voelcker Fund Young Investigator Award from the MAX AND MINNIE TOMERLIN VOELCKER FUND.

References

1. Li X, Zai J, Wang X, Li Y (2020) Potential of large "first generation" human-to-human transmission of 2019-nCoV. J Med Virol 92:448–454
2. Gralinski LE, Menachery VD (2020) Return of the Coronavirus: 2019-nCoV. Viruses 12:135
3. Chan JF-W, Kok K-H, Zhu Z, Chu H, To KK-W, Yuan S, Yuen K-Y (2020) Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. Emerg Microb Infect 9:221–236
4. Wang C, Liu Z, Chen Z, Huang X, Xu M, He T, Zhang Z (2020) The establishment of reference sequence for SARS-CoV-2 and variation analysis. J Med Virol 92:667–674
5. Khailany RA, Safdar M, Ozaslan M (2020) Genomic characterisation of a novel SARS-CoV-2. Gene Rep 19:100682
6. Andersen KG, Rambaut A, Lipkin WI, Holmes EC, Garry RF (2020) The proximal origin of SARS-CoV-2. Nat Med 26:450–455
7. Li Y-H, Hu C-Y, Wu N-P, Yao H-P, Li L-J (2019) Molecular characteristics, functions, and related pathogenicity of MERS-CoV proteins. Engineering 5:940–947
8. Song Z, Xu Y, Bao L, Zhang L, Yu P, Qu Y, Zhu H, Zhao W, Han Y, Qin C (2019) From SARS to MERS thrusting coronaviruses into the spotlight. Viruses 11:59
9. Hilgenfeld R (2014) From SARS to MERS: crystallographic studies on coronaviral proteases enable antiviral drug design. FEBS J 281:4085–4096
10. Calligari P, Bobone S, Ricci G, Bocedi A (2020) Molecular investigation of SARS-CoV-2 proteins and their interactions with antiviral drugs. Viruses 12:445
11. Liu DX, Fung TS, Chong KK-L, Shukla A, Hilgenfeld R (2014) Accessory proteins of SARS-CoV and other coronaviruses. Antiviral Res 109:97–109
12. Prajapa M, Sarma P, Shekhar N, Avti P, Sinha S, Kaur H, Kumar S, Bhattacharyya A, Kumar H, Bansal S, Medhi B (2020) Drug targets for corona virus: a systematic review. Indian J Pharmacol 52:56–65
13. Zhou P, Yang X-L, Wang X-G, Hu B, Zhang L, Zhang W, Si H-R, Zhu Y, Li B, Huang C-L, Chen H-D, Chen J, Luo Y, Guo H, Jiang R-D, Liu M-Q, Chen Y, Shen X-R, Wang X, Zheng X-S, Zhao K, Chen Q-J, Deng F, Liu L-L, Yan B, Zhan F-X, Wang Y-Y, Xiao G-F, Shi Z-L (2020) A pneumonia outbreak

associated with a new coronavirus of probable bat origin. Nature 579:270–273
14. Cagliani R, Forni D, Clerici M, Sironi M (2020) Computational inference of selection underlying the evolution of the novel coronavirus, SARS-CoV-2. J Virol. https://doi.org/10.1128/JVI.00411-20
15. Huang C, Lokugamage KG, Rozovics JM, Narayanan K, Semler BL, Makino S (2011) SARS coronavirus nsp1 protein induces template-dependent endonucleolytic cleavage of mRNAs: viral mRNAs are resistant to nsp1-induced RNA cleavage. PLoS Pathog 7:e1002433
16. Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD, Lopez R (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res 47:W636–W641
17. Cornillez-Ty CT, Liao L, Yates JR, Kuhn P, Buchmeier MJ (2009) Severe acute respiratory syndrome coronavirus nonstructural protein 2 interacts with a host protein complex involved in mitochondrial biogenesis and intracellular signaling. J Virol 83:10314–10318
18. Sakai Y, Kawachi K, Terada Y, Omori H, Matsuura Y, Kamitani W (2017) Two-amino acids change in the nsp4 of SARS coronavirus abolishes viral replication. Virology 510:165–174
19. Lei J, Kusov Y, Hilgenfeld R (2018) Nsp3 of coronaviruses: structures and functions of a large multi-domain protein. Antiviral Res 149:58–74
20. Baez-Santos YM, St. John SE, Mesecar AD (2015) The SARS-coronavirus papain-like protease: structure, function, and inhibition by designed antiviral compounds. Antiviral Res 115:21–38
21. Tomar S, Johnston ML, St. John SE, Osswald HL, Nyalapatla PR, Paul LN, Ghosh AK, Denison MR, Mesecar AD (2015) Ligand-induced dimerization of middle east respiratory syndrome (MERS) coronavirus nsp5 protease ($3CL^{pro}$) implications For nsp5 Regulation And The Development Of Antivirals. J. Biol. Chem. 290:19403–19422
22. Cottam EM, Whelband MC, Wileman T (2014) Coronavirus NSP6 restricts autophagosome expansion. Autophagy 10:1426–1441
23. Angelini MM, Akhlaghpour M, Neuman BW, Buchmeier MJ (2013) Severe acute respiratory syndrome coronavirus nonstructural proteins 3, 4, and 6 induce double-membrane vesicles. mBio 13:e00524–e1513
24. te Velthuis AJ, van de Worm SH, Snijder EJ (2012) The SARS-coronavirus nsp7+nsp8 complex is a unique multimeric RNA polymerase capable of both de novo initiation and primer extension. Nucleic Acids Res 40:1737–1747
25. 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 LW, Wang Q, Lou Z, Rao Z (2020) Structure of the RNA-dependent RNA polymerase from COVID-19 virus. Science. https://doi.org/10.1126/science.abb7498
26. Zhao S, Ge X, Wang X, Liu A, Guo X, Zhou L, Yu K, Yang H (2015) The DEAD-box RNA helicase 5 positively regulates the replication of porcine reproductiv e and respiratory syndrome virus by interacting with viral Nsp9 *in vitro*. Virus Res 195:217–224
27. Ma Y, Wu L, Shaw N, Gao Y, Wang J, Sun Y, Lou Z, Yan L, Zhang R, Rao Z (2015) Structural basis and functional analysis of the SARS coronavirus nsp14-nsp10 complex. Proc Natl Acad Sci USA 112:9436–9441
28. Wang Y, Sun Y, Wu A, Xu S, Pan R, Zeng C, Jin X, Ge X, Shi Z, Ahola T, Guo D (2015) Coronavirus nsp10/nsp16 methyltransferase can be targeted by nsp10-derived peptide *in vitro* and *in vivo* to reduce replication and pathogenesis. J Virol 89:8416–8427
29. Subissi L, Posthuma CC, Collet A, Zevenhoven-Dobbe JC, Gorbalenya AE, Decroly E, Snijder EJ, Canard B, Imbert I (2014) One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. Proc Natl Acad Sci USA 111:E3900–E3909
30. Jang K-J, Jeong S, Kang DY, Sp N, Yang YM, Kim D-E (2020) A high ATP concentration enhances the cooperative translocation of the SARS coronavirus helicase nsP13 in the unwinding of duplex RNA. Sci Rep 10:4481
31. Jia Z, Yan L, Ren Z, Wu L, Wang J, Guo J, Zheng L, Ming Z, Zhang L, Lou Z, Rao Z (2019) Delicate structural coordination of the severe acute respiratory syndrome coronavirus Nsp13 upon ATP hydrolysis. Nucleic Acids Res 47:6538–6550
32. Ivanov KA, Thiel V, Dobbe JC, van der Meer Y, Snijder EJ, Ziebuhr J (2004) Multiple enzymatic activities associated with severe acute respiratory syndrome coronavirus helicase. J Virol 78:5619–5632
33. Case JB, Ashbrook AW, Dermody TS, Denison MR (2016) Mutagenesis of *S*-adenosyl-L-methionine-binding residues in coronavirus nsp14 N7-methyltransferase demonstrates differing requirements for genome translation and resistance to innate immunity. J Virol 90:7248–7256
34. Jin X, Chen Y, Sun Y, Zeng C, Wang Y, Tao J, Wu A, Yu X, Zhang Z, Tian J, Guo D (2013) Characterization of the guanine-N7 methyltransferase activity of coronavirus nsp14 on nucleotide GTP. Virus Res 176:45–52
35. Bouvet M, Debarnot C, Imbert I, Selisko B, Snijder EJ, Canard B, Decroly E (2010) In vitro reconstitution of SARS-coronavirus mRNA cap methylation. PLoS Pathog 6:e1000863
36. Nakagawa K, Lokugamage KG, Makino S (2016) Viral and cellular mRNA translation in coronavirus-infected cells. Adv Virus Res 96:165–192
37. Bhardwaj K, Sun J, Holzenburg A, Guarino LA, Kao CC (2006) RNA recognition and cleavage by the SARS coronavirus endoribonuclease. J Mol Biol 361:243–256
38. Hackbart M, Deng X, Baker SC (2020) Coronavirus endoribonuclease targets viral polyuridine sequences to evade activating host sensors. Proc Natl Acad Sci USA 117:8094–8103
39. Bhardwaj K, Palaninathan S, Ortiz Alcantara JM, Li Yi L, Guarino L, Sacchettini JC, Cheng Kao C (2008) Structural and functional analyses of the severe acute respiratory syndrome coronavirus endoribonuclease Nsp15. J Biol Chem 283:3655–3664
40. Kim Y, Jedrzejczak R, Maltseva NI, Wilamowski M, Endres M, Godzik A, Michalska K, Joachimiak A (2020) Crystal structure of Nsp15 endoribonuclease NendoU from SARS-CoV-2. Protein Sci. https://doi.org/10.1002/pro.3873
41. Bhardwaj K, Guarino L, Kao CC (2004) The severe acute respiratory syndrome coronavirus Nsp15 protein is an endoribonuclease that prefers manganese as a cofactor. J Virol 78:12218
42. Deng X, Baker SC (2018) An "Old" protein with a new story: coronavirus endoribonuclease is important for evading host antiviral defenses. Virology 517:157–163
43. Decroly E, Imbert I, Coutard B, Bouvet M, Selisko B, Alvarez K, Gorbalenya AE, Snijder EJ, Canard B (2008) Coronavirus nonstructural protein 16 Is a cap-0 binding enzyme possessing (nucleoside-2'O)-methyltransferase activity. J Virol 82:8071–8084
44. Decroly E, Debarnot C, Ferron F, Bouvet M, Coutard B, Imbert I, Gluais L, Papageorgiou N, Sharff A, Bricogne G, Ortiz-Lombardia M, Lescar J, Canard B (2011) Crystal structure and functional analysis of the sars-coronavirus RNA cap 2'-O-methyltransferase nsp10/nsp16 complex. PLoS Pathog 7:e1002059
45. Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S, Zhang Q, Shi X, Wang Q, Zhang L, Wang X (2020) Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. Nature. https://doi.org/10.1038/s41586-020-2180-5s

46. Shang J, Ye G, Shi K, Wan Y, Luo C, Aihara H, Geng Q, Auerbach A, Li F (2020) Structural basis of receptor recognition by SARS-CoV-2. Nature. https://doi.org/10.1038/s41586-020-2179-y
47. Walls AC, Park Y-J, Tortorici MA, Wall A, McGuire AT, Vessler D (2020) Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. Cell 180:281–292
48. Siu K-L, Yuen K-S, Castano-Rodriguez C, Ye Z-W, Yeung M-L, Fung S-Y, Yuan S, Chan C-P, Yuen K-Y, Enjuanes L, Jin D-Y (2019) Severe acute respiratory syndrome coronavirus ORF3a protein activates the NLRP3 inflammasome by promoting TRAF3-dependent ubiquitination of ASC. FASEB J 33:8865–8877
49. Verdia-Baguena C, Nieto-Torres JL, Alcaraz A, DeDiego ML, Torres J, Aguilella VM, Enjuanes L (2012) Coronavirus E protein forms ion channels with functionally and structurally-involved membrane lipids. Virology 432:485–494
50. Schoeman D, Fielding BC (2019) Coronavirus envelope protein: current knowledge. Virol J 16:69
51. Lim KP, Liu DX (2001) The Missing Link in Coronavirus Assembly retention of the avian coronavirus infectious bronchitis virus envelope protein in the pre-golgi compartments and physical interaction between the envelope and membrane proteins. J Biol Chem 276:17515–17523
52. Ruch TR, Machamer CE (2012) The coronavirus E protein: assembly and beyond. Viruses 4:363–382
53. Weiss SR, Navas-Martin S (2005) Coronavirus pathogenesis and the emerging pathogen severe acute respiratory syndrome coronavirus. Microbiol Mol Biol Rev 69:635–664
54. 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 (2011) A structural analysis of M protein in coronavirus assembly and morphology. J Struct Biol 174:11–22
55. Tsoi H, Li L, Chen ZS, Lau K-F, Tsui SKW, Chan HYE (2014) The SARS-coronavirus membrane protein induces apoptosis via interfering with PDK1-PKB/Akt signalling. Biochem J 464:439–447
56. Siu YL, Teoh KT, Lo J, Chan CM, Kien F, Escriou N, Tsao SW, Nicholls JM, Altmeyer R, Peiris JSM, Bruzzone R, Nal B (2008) The M, E, and N structural proteins of the severe acute respiratory syndrome coronavirus re required for efficient assembly, trafficking, and release of virus-like particles. J Virol 82:11318–11330
57. Kumar P, Gunalan V, Liu B, Chow VTK, Druce J, Birch C, Catton M, Fielding BC, Tan Y-J, Lal SK (2007) The nonstructural protein 8 (nsp8) of the SARS coronavirus interacts with its ORF6 accessory protein. Virology 366:293–303
58. Zhao J, Falcon A, Zhou H, Netlan J, Enjuanes L, Brena PP, Perlman S (2009) Severe acute respiratory syndrome coronavirus protein 6 is required for optimal replication. J Virol 83:2368–2373
59. Nelson CA, Pekosz A, Lee CA, Diamond MS, Fremont DH (2005) Structure and intracellular targeting of the SARS-coronavirus Orf7a accessory protein. Structure 13:75–85
60. Schaecher SR, Mackenzie JM, Pekosz A (2007) The ORF7b protein of severe acute respiratory syndrome coronavirus (SARS-CoV) is expressed in virus-infected cells and incorporated into SARS-CoV particles. J Virol 81:718–731
61. Le TM, Wong HH, Tay FPL, Fang S, Keng C-T, Tan YJ, Liu DX (2007) Expression, post-translational modification and biochemical characterization of proteins encoded by subgenomic mRNA8 of the severe acute respiratory syndrome coronavirus. FEBS J 274:4211–4222
62. Wong HH, Fung TS, Fang S, Huang M, Le MT, Liu DX (2018) Accessory proteins 8b and 8ab of severe acute respiratory syndrome coronavirus suppress the interferon signaling pathway by mediating ubiquitin-dependent rapid degradation of interferon regulatory factor 3. Virology 515:165–175
63. Grunewald ME, Fehr AR, Athmer J, Perlman S (2018) The coronavirus nucleocapsid protein is ADP-ribosylated. Virology 517:62–68
64. Mu J, Xu J, Zhang L, Shu T, Wu D, Huang M, Ren Y, Li X, Geng Q, Xu Y, Qiu Y, Zhou X (2020) SARS-CoV-2-encoded nucleocapsid protein acts as a viral suppressor of RNA interference in cells. Sci China Life Sci 63:10
65. Surjit M, Liu B, Chow VTK, Lal SK (2006) The nucleocapsid protein of severe acute respiratory syndrome-coronavirus inhibits the activity of cyclin-cyclin-dependent kinase complex and blocks S phase progression in mammalian cells. J Biol Chem 281:10669–10681
66. Gordon DE, Jang GM, Bouhaddou M, Xu J, Obernier K, White KM, O'Meara MJ, Rezelj VV, Guo JZ, Swaney DL, Tummino TA, Huettenhain R, Kaake RM, Richards AL, Tutuncuoglu B, Foussard H, Batra J, Haas K, Modak M, Kim M, Haas P, Polacco BJ, Braberg H, Fabius JM, Eckhardt M, Soucheray M, Bennett MJ, Cakir M, McGrego MJ, Li Q, Meyer B, Roesch F, Vallet T, Mac Kain A, Miorin L, Moreno E, Naing ZZC, Zhou Y, Peng S, Shi Y, Zhang Z, Shen W, Kirby IT, Melnyk JE, Chorba JS, Lou K, Dai SA, Barrio-Hernandez I, Memon D, Hernandez-Armenta C, Lyu J, Mathy CJP, Perica T, Pilla KB, Ganesan SJ, Saltzberg DJ, Rakesh R, Liu X, Rosenthal SB, Calviello L, Venkataramanan S, Liboy-Lugo J, Lin Y, Huang X-P, Liu Y, Wankowicz SA, Bohn M, Safari M, Ugur FS, Koh C, Savar NS, Tran QD, Shengjuler D, Fletcher SJ, O'Neal MC, Cai Y, Chang JCJ, Broadhurst DJ, Klippsten S, Sharp PP, Wenzell NA, Kuzuoglu D, Wang H-Y, Trenker R, Young JM, Cavero DA, Hiatt J, Roth TL, Rathore U, Subramanian A, Noack J, Hubert M, Stroud RM, Frankel AD, Rosenberg OS, Verba KA, Agard DA, Ott M, Emerman M, Jura N, von Zastrow M, Verdin E, Ashworth A, Schwartz O, d'Enfert C, Mukherjee S, Jacobson M, Malik HS, Fujimori DG, Ideker T, Craik CS, Floor SN, Fraser JS, Gross JD, Sali A, Roth BL, Ruggero D, Taunton J, Kortemme T, Beltrao P, Vignuzzi M, Garcia-Sastre A, Shokat KM, Shoichet BK, Krogan NJ (2020) A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. Nature. https://doi.org/10.1038/s41586-020-2286-9
67. Rancurel C, Khosravi M, Dunker AK, Romero PR, Karlin D (2009) Overlapping genes produce proteins with unusual sequence properties and offer insight into de novo protein creation. J Virol 83:10719–10736
68. Shukla A, Hilgenfeld R (2015) Acquisition of new protein domains by coronaviruses: analysis of overlapping genes coding for proteins N and 9b in SARS coronavirus. Virus Genes 50:29–38
69. Wu A, Peng Y, Huang B, Ding X, Wang X, Niu P, Meng J, Zhu Z, Zhang Z, Wang J, Sheng J, Quan L, Xia Z, Tan W, Cheng G, Jiang T (2020) Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China. Cell Host Microbe 27:325–328
70. Firth AE, Brierley I (2012) Non-canonical translation in RNA viruses. J Gen Virol 93:1385–1409
71. Hussain S, Pan J, Chen Y, Yang Y, Xu J, Peng Y, Wu Y, Li Z, Zhu Y, Tien P, Guo D (2005) Identification of novel subgenomic RNAs and noncanonical transcription initiation signals of severe acute respiratory syndrome coronavirus. J Virol 79:5288–5295
72. Quan P-L, Firth C, Street C, Henriquez JA, Petrosov A, Tashmukhamedova A, Hutchison SK, Egholm M, Osinubi MOV, Niezgoda M, Ogunkoya AB, Briese T, Rupprecht CE, Lipkin WI (2010) Identification of a severe acute respiratory syndrome coronavirus-like virus in a leaf-nosed bat in Nigeria. MmBio 1:00208–00210

73. Wu F, Zhao S, Yu B, Chen Y-M, Wang W, Song Z-G, Hu Y, Tao Z-W, Tian J-H, Pei Y-Y, Yuan M-L, Zhang Y-L, Dai F-H, Liu Y, Wang Q-M, Zheng J-J, Xu L, Holmes EC, Zhang Y-Z (2020) A new coronavirus associated with human respiratory disease in China. Nature 579:265–269
74. Yang D, Leibowitz JL (2015) The structure and functions of coronavirus genomic 3' and 5' ends. Virus Res 206:120–133
75. Jean S-S, Lee P-I, Hsueh P-R (2020) Treatment options for COVID-19: The reality and challenges. J Microbiol. https://doi.org/10.1016/j.jmii.2020.03.034
76. Costanzo M, De Giglio MAR, Roviello GN (2020) SARS-CoV-2: recent reports on antiviral therapies based on lopinavir/ritonavir, darunavir/umifenovir, hydroxychloroquine, remdesivir, favipiravir and other drugs for the treatment of the new coronavirus. Curr Med Chem 27:32297571
77. Zeng Q-L, Yu Z-J, Gou J-J, Li G-M, Ma S-H, Zhang G-F, Xu J-H, Lin W-B, Cui G-L, Zhang M-M, Li C, Wang Z-S, Zhang Z-H, Liu Z-S (2020) Effect of convalescent plasma therapy on viral shedding and survival in COVID-19 patients. J Infect Dis. https://doi.org/10.1093/infdis/jiaa228
78. Caly L, Druce JD, Catton MG, Jans DA, Wagstaff KM (2020) The FDA-approved drug ivermectin inhibits the replication of SARS-CoV-2 *in vitro*. Antiviral Res 178:104787
79. Esposito S, Noviello S, Pagliano P (2020) Update on treatment of COVID-19: ongoing studies between promising and disappointing results. Le Infezioni Med 2:198–211
80. Kupferschmidt K, Cohen J (2020) Race to find COVID-19 treatments accelerates. Science 367:1412–1413
81. Siegel D, Doerffler E, Clarke MO, Chun K, Zhang L, Neville S, Carra E, Lew W, Ross B, Wang Q, Wolfe L, Jordan R, Soloveva V, Knox J, Perry J, Perron M, Stray KM, Barauskas O, Feng JY, Xu Y, Lee G, Rheingold AL, Ray AS, Bannister R, Strickley R, Swaminathan S, Lee WA, Bavari S, Cihlar T, Lo MK, Warren TK, Mackman RL (2017) Discovery and synthesis of a phosphoramidate prodrug of a pyrrolo[2,1-f][triazin-4-amino] adenine C-nucleoside (GS-5734) for the treatment of ebola and emerging viruses. J Med Chem 60:1648–1661
82. Gordon CJ, Tchesnokov EP, Woolner E, Perry JK, Feng JY, Porter DP, Götte M (2020) Remdesivir is a direct-acting antiviral that inhibits RNA-dependent RNA polymerase from severe acute respiratory syndrome coronavirus 2 with high potency. J Biol Chem. https://doi.org/10.1074/jbc.RA120.013679
83. Meyerowitz EA, Vannier AGL, Friesen MGN, Schoenfeld S, Gelfand JA, Callahan MV, Kim AY, Reeves PM, Poznansky MC (2020) Rethinking the role of hydroxychloroquine in the treatment of COVID-19. FASEB J 34(5):6027–6037
84. Touret F, de Lamballerie X (2020) Of Chloroquine and COVID-19. Antiviral Res 177:104762
85. Wong YK, Yang J, He Y (2020) Caution and clarity required in the use of chloroquine for COVID-19. Lancet Rehumatol 2:255
86. Joyce E, Fabre A, Mahon N (2012) Hydroxychloroquine cardiotoxicity presenting as a rapidly evolving biventricular cardiomyopathy: key diagnostic features and literature review. Eur Heart J 2:77–83
87. Geleris J, Sun Y, Platt J, Zucker J, Baldwin M, Hripcsak G, Labella A, Manson D, Kubin C, Barr RG, Sobieszczyk ME, Schluger NW (2020) Obersvational study of hydroxychloroquine in hospitalized patients with Covid-19. N Engl J Med. https://doi.org/10.1056/NEJMoa2012410
88. Pati ML, Hornick JR, Niso M, Berardi F, Spitzer D, Abate C, Hawkins W (2017) Sigma-2 receptor agonist derivatives of 1-cyclohexyl-4-[3-(5-methoxy-1,2,3,4-tetrahydronaphthalen-1-yl)propyl]piperazine (PB28) induce cell death via mitochondrial superoxide production and caspase activation in pancreatic cancer. BMC Cancer 17:51
89. Riva L, Yuan S, Yin X, Martin-Sancho L, Matsunaga N, Burgstaller-Muehlbacher S, Pache L, De Jesus PP, Hull MV, Chang M, Chan JF-W, Cao J, Poon VK-M, Herbert K, Nguyen T-T, Pu Y, Nguyen C, Rubanov A, Martinez-Sobrido L, Liu W-C, Miorin L, White KM, Johnson JR, Benner C, Sun R, Schultz PG, Su A, Garcia-Sastre A, Chatterjee AK, Yuen K-Y, Chanda SK (2020) A Large-scale drug repositioning survey for SARS-CoV-2 antivirals. bioRxiv. https://doi.org/10.1101/2020.04.16.044016
90. Xu J, Zhang Y (2020) Traditional Chinese Medicine treatment of COVID-19. Complement Ther Clin Pract 39:101165
91. Chen C-H, Dickman KG, Moriya M, Zavadil J, Sidorenko VS, Edwards KL, Gnatenko DV, Wu L, Turesky RJ, Wu X-R, Pu Y-S, Grollman AP (2012) Aristolochic acid-associated urothelial cancer in Taiwan. Proc Natl Acad Sci USA 109:8241–8246
92. Duan L, Guo L, Wang L, Yin Q, Zhang C-M et al (2018) Application of metabolomics in toxicity evaluation of traditional Chinese medicines. Chin Med 13:60
93. Lv W, Piao J-H, Jiang J-G (2012) Typical toxic components in traditional Chinese medicine. Expert Opin Drug Saf 11:985–1002
94. Wassenaar TM, Zou Y (2020) 2019_nCoV/SARS-CoV-2: rapid classification of betacorona viruses and identification of Traditional Chinese Medicine as potential origin of zoonotic coronaviruses. Lett Appl Microbiol 70:342–348
95. Lurie N, Saville M, Hatchett R, Halton J (2020) Developing Covid-19 vaccines at pandemic speed. N Engl J Med. https://doi.org/10.1056/NEJMp2005630
96. Brown BL, McCullough J (2020) Treatment for emerging viruses: convalescent plasma and COVID-19. Transfus Apher Sci. https://doi.org/10.1016/j.transci.2020.102790
97. Xu S, Li Y (2020) Beware of the second wave of COVID-19. Lancet 395:1321–1322
98. Leung K, Wu JT, Liu D, Leung GM (2020) First-wave COVID-19 transmissibility and severity in China outside Hubei after control measures, and second-wave scenario planning: a modelling impact assessment. Lancet 395:1382–1393
99. Ceraolo C, Giorgi FM (2020) Genomic variance of the 2019-nCoV coronavirus. J Med Virol 92:522–528
100. https://swissmodel.expasy.org/repository/species/2697049. Accessed 26 Apr 2020, 6:05 AM