The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) envelope (E) protein harbors a conserved BH3-like motif

Vincent Navratil [1], Loïc Lionnard [2], Sonia Longhi [3], Christophe Combet [4] and Abdel Aouacheria [2,*]

[1] PRABI, Rhône Alpes Bioinformatics Center, UCBL, Lyon1, Université de Lyon, Lyon, France.

[2] ISEM, Institut des Sciences de l'Evolution de Montpellier, Université de Montpellier, UMR 5554, CNRS, IRD, EPHE, Place Eugène Bataillon, 34095 Montpellier, France

[3] CNRS and Aix-Marseille Univ, Laboratoire Architecture et Fonction des Macromolecules Biologiques (AFMB), UMR 7257, Marseille, France.

[4] Centre de Recherche en Cancérologie de Lyon, UMR Inserm U1052, CNRS 5286, Université Claude Bernard Lyon 1, Centre Léon Bérard, Lyon, France.

* Corresponding author

Abdel Aouacheria, ISEM, Institut des Sciences de l'Evolution de Montpellier, Université de Montpellier, CNRS, EPHE, IRD, Montpellier, France. Phone : +33467143751

abdel.aouacheria@umontpellier.fr

Abstract

Following the Severe Acute Respiratory Syndrome coronavirus (SARS-CoV) in 2002 and Middle East Respiratory Syndrome coronavirus (MERS-CoV) in 2012, the novel coronavirus SARS-CoV-2 emerged at the end of 2019 as a highly pathogenic infectious agent that rapidly spread around the world. SARS-CoV-2 shares high sequence homology with SARS-CoV and causes acute, highly lethal pneumonia coronavirus disease 2019 (COVID-19) with clinical symptoms similar to those reported for SARS-CoV. Like other betacoronaviruses, SARS-CoV-2 encode four major structural proteins: Spike (S), Membrane (M), Nucleocapsid (N) and Envelope (E). SARS-CoV E protein is abundant in infected cells and plays a crucial role in viral particle assembly. Moreover, SARS coronaviruses lacking E are attenuated in vivo, suggesting that CoV E may act as

a critical virulence factor not only in SARS-CoV but also in the case of the new coronavirus SARS-CoV-2. Ectopic expression of SARS-CoV E was previously shown to trigger apoptosis (cell suicide) of T lymphocytes, lymphopenia being a common feature observed in fatal cases following viral infections. Importantly, T-cell apoptosis was shown to involve interaction between the C-terminal region of SARS-CoV E and the Bcl-2 family member Bcl-xL, which acts as a potent anti-apoptotic protein. Here we provide the first observation that the SARS-CoV E and SARS-CoV-2 E proteins share a conserved Bcl-2 Homology 3 (BH3)-like motif in their C-terminal region, a well-studied motif shown to be necessary for SARS-CoV E binding to Bcl-xL. We used available sequence data for SARS-CoV-2 and related coronaviruses, in combination with structural information, to study the structure to biological activity relationships of SARS-CoV-2 E in relation with its BH3-like motif. Our analysis of the SARS-CoV E interactome further revealed that the predicted SARS-CoV-2 network is extensively wired to the Bcl-2 apoptotic switch. Research is therefore needed to establish if SARS-CoV-2 E targets prosurvival Bcl-2 homologs to modulate cell viability, as part of a coronavirus strategy to interfere with apoptosis. The identification of small molecules (or the repurposing of existing drugs) able to disrupt SARS-CoV-2 E BH3-mediated interactions might provide a targeted therapeutic approach for COVID-19 treatment. Recombinant SARS-CoV-2 expressing an E protein with a deleted or mutated BH3-like motif might also be of interest for the design of a live, attenuated vaccine.

Keywords: coronavirus; covid-19; apoptosis; Bcl-2/BH3; protein-protein interactions; structure-function relationships; virus-host interactions

Introduction

A novel virus of the *Coronaviridae* family, designed as Severe Acute Respiratory Syndrome coronavirus 2 (SARS-CoV-2), has been identified as being the causative agent of coronavirus disease 2019 (COVID-19) (*1*). While its exact origins are still debated (*2-4*), it is known that SARS-CoV-2 belongs to the β genus (subgenus: *Sarbecovirus*) of the order Nidovirales, subfamily *Orthocoronavirinae*. Coronaviruses of this subfamily encode at least four major structural proteins in their genomes: Spike (S), Membrane (M), Nucleocapsid (N) and Envelope (E) (*3-6*). The CoV E protein is the smallest of these structural proteins, consisting of only 76-109 amino acids (for a recent

review, see (*7*)). Recent biochemical and structural data revealed that SARS-CoV E is a viroporin endowed with ion channel and membrane-permeabilizing activities (*8*), forming homopentameric transmembrane α-helical bundles in lipid bilayers (*9*). This protein is abundant in infected cells where it plays a crucial role in viral particle assembly (*10, 11*). Compared to *wild-type* controls, recombinant coronaviruses lacking the E gene have significantly lower viral titers and reduced viral maturation and propagation (*12-14*). E-deficient SARS-CoV is also attenuated *in vivo*, exhibiting diminished pathogenic effects (*15, 16*) and inability to grow in the central nervous system (*17*), suggesting that CoV E may act as an important virulence factor in coronaviruses. The C-terminal region of SARS-CoV E is mostly α-helical (*9*) and harbors at its extremity a so-called (PSD95/DLG1/ZO-1) PDZ-binding motif (PBM) (*18*) that was identified as a critical virulence determinant (*19*). PDZ are common globular domains that interact with short linear motifs (SLiMs) usually present at the C-terminus of their target proteins and prevalent within intrinsically disordered regions (*20*). Through its C-terminal PBM, SARS-CoV E was shown to engage into interactions with multiple host proteins such as syntenin (*19*) and PALS1 (*18*).

In a study published in 2005, Yang et al. (*21*) provided evidence that ectopic expression of SARS-CoV E induced T-cell apoptosis, which was inhibited by the prosurvival Bcl-2 family member Bcl-xL (BCL2L1). The authors identified a Bcl-2 Homology 3 (BH3) motif encompassing residues 51 to 58 (LVKPTVYV) in the C-terminal moiety of the SARS-CoV E protein and subsequent mutagenesis experiments revealed that this amino acid stretch was responsible for interaction with Bcl-xL. BH3 motifs form a newly defined class of SLiMs (*22*) that have well-studied roles in the apoptotic network of protein-protein interactions, by binding to globular proteins of the Bcl-2 family (*23, 24*). Elaborating on previous works (*22, 25*), we provide here the first observation that the C-terminal region of SARS-CoV-2 E also possesses this BH3-like sequence. While great attention has been paid to the role of BH3 motifs in cancer (*26-28*), research efforts merit to be undertaken to determine biological, clinical and therapeutic relevance of modulating CoV E BH3-like motifs and their associated protein-protein interaction networks in the context of coronavirus pathogenesis. The prophylactic interest of using SARS-CoV-2 vaccines with an E protein harboring a mutated or deleted BH3-like motif should also be explored.

Results and Discussion

The BH3-like motif identified by Yang et al. (*21*) in the SARS-CoV E sequence (GenBank Acc.# AAP73416) appears to slightly diverge from other BH3 signatures, as we previously noticed (*25*). This result is not surprising given that, like many other well-established SLiMs, BH3 motifs are degenerate, with diverse lengths (7-15 amino acids) and variable sequences at most of their amino acid positions, especially when structurally and evolutionarily distant BH3-containing proteins are considered altogether (*22*) (see the multiple sequence alignment provided in Fig. 1). We used this CoV E sequence (GenBank Acc.# AAP73416) from SARS coronavirus ZJ01 as a reference in our subsequent analysis. The percent sequence identity between this reference SARS-CoV E sequence and SARS-CoV-2 E is about 95% (pairwise similarity being 98,7%). In comparison, the overall sequence similarities between the E proteins of SARS-CoV-2 and human coronavirus Middle East respiratory syndrome (MERS-CoV), CoV-HKU1, CoV-OC43 and CoV-4408 are 69,5%, 68,3%, 67,8 and 65,5%, respectively.

In order to gain insight into the conservation and evolution of CoV E BH3-like sequences, we performed multiple sequence alignments starting with the closest relatives to the reference sequence in terms of phylogenetic proximity (betacoronaviruses) (Fig. 2) and adding more dissimilar sequences from the three other genera of *Orthocoronavirinae* (alpha-, gamma- and deltacoronaviruses). CoV E protein sequences were found in all clades but with various levels of similarity in the C-terminal domain where the BH3-like motif occurs (Fig. 3). Based on the amino acid sequences of the BH3-like region, we analyzed the phylogenetic relationships between the CoV E proteins from SARS-CoV-2 and other coronaviruses derived from various species (Fig. 4). CoV E sequences from gammacoronavirus rooted the phylogenetic tree, providing evidence that their BH3-like region is more distantly related to the reference ZJ01 sequence than the homologous region from the other coronavirus genera (i.e., alpha- and deltacoronaviruses). Deltacoronavirus sequences were found to be more akin to sequences from betacoronaviruses than alphacoronavirus sequences. SARS-CoV-2 E sequence clusters with bat CoV isolate RaTG13 and pangolin CoV isolate Guangdong, within a group (100% bootstrap support) comprising the reference ZJ01 sequence, other bat-SARS-like coronaviruses and isolate 007/2004 derived from civet. This result has been initially viewed as being indicative of coronavirus crossing of the animal-human species barrier (*3, 4, 29*), with bats as reservoir hosts and pangolins as intermediate hosts facilitating SARS-CoV-2 transfer to humans, although sampling bias (i.e., under coverage of

coronavirus diversity in other species) may obscure this interpretation (*2*). Sequences from other human coronaviruses of zoonotic origin (human CoV OC43, 4408, HKU1 and MERS) fell into distinct phylogenetic clades with high probability values.

Motif-based sequence analysis suggested that coronaviruses have evolved a conserved and specific BH3-like motif (data not shown). Instead of conforming to the L^{1}-X^{2}-X^{3}-X^{4}-X^{5}-D^{6} core pattern (with conserved leucine and aspartic acid residues) characteristic of most BH3 sequences (*22, 25*), the BH3 signature inferred from betacoronavirus E sequences is atypical and reads as L^{1}-X^{2}-X^{3}-P^{4}-X^{5}-ϕ^{6}-Ω^{7}-ϕ^{8}-Ω^{9} (with ϕ indicating a preference for aliphatic residues and Ω for aromatic residues) (see logo analysis in Fig. 2, *bottom*). Mutagenesis of key positions 1, 6 and 8 of this BH3-like sequence consensus (corresponding to positions 51, 56 and 58 in the actual sequence) was reported to abrogate the ability of SARS-CoV E to interact with Bcl-xL (*21*). The Pro residue at position 4 of the consensus (position 54 in the actual sequence) is totally conserved and a study indicated that it is critical for targeting the protein to the Golgi (*30*). The influence of the other residues present in the CoV E BH3-like motifs (including the potentially important Asn48 residue located three residues upstream of L51) on protein stability, subcellular targeting, function and interaction is unknown and awaits further characterization.

To predict structure-activity relationship, we constructed a structural model of SARS-CoV-2 E guided by the homologous structure of SARS-CoV E (PDB code: 5X29) (*9*). Note that only four positions are different between full-length SARS-CoV E and SARS-CoV-2 E: two within the BH3-like helix (Thr55 and Val56 in SARS-CoV E, Ser55 and Phe56 in SARS-CoV-2) and two in a site immediately N-terminal to the PBM (where a pair of residues formed by Glu69 and Gly70 in SARS-CoV E is found instead of a single Arg69 residue in SARS-CoV-2 E). As shown in Fig. 5, in both the experimental structure (Fig. 5A) and the derived structural model (Fig. 5B), position of Leu51 is clearly inter-helical whereas the highly conserved Pro54 residue is located at the beginning of the terminal α-helix. Except for Tyr57, the BH3-like region (in red) appears to be only marginally involved in the inter-chain interactions needed to form the protein homo-pentamer (*9*). Rather, its external orientation, and extended (i.e. irregular) structure in the homo-pentamer (*31*) makes the CoV E BH3-like segment highly prone to serve as a small binding interface able to engage in transient interactions with protein partners, as expected for a SLiM (*22*). It is possible that the loop that exposes the Leu51 residue may

adopt a regular secondary structure upon binding to targets to form an extended α-helix, with the conserved Pro54 inducing formation of a π-helix.

Annotation of SARS-CoV-2 E from 3,449 SARS-CoV-2 genomes deposited in GISAID EpiCoV identified at least 10 mutated positions in the translated E sequences from SARS-CoV-2 isolates (Fig. 5C), five of them lying in regions known to interact with host proteins: two within the BH3-like core region and three in the PBM. These substitutions were mapped onto the SARS-CoV-2 E structural model (Fig. 5B). We found that the mutated residues tend to locate on the surface of the protein, in line with previous observations that the hydrophobic core of proteins from RNA viruses shows greater conservation than functional sites on the protein surface (*32, 33*). Interestingly, we also detected a partial deletion spanning eight amino acids (including a Cys residue) in a turn present upstream of the BH3-like region. In addition to shortening the transmembrane domain, which may affect subcellular localization (e.g. by enhancing retention in the Golgi (*34*)), such deletion might interfere with the binding properties of the SARS-CoV-2 E BH3-like motif by introducing local conformational rearrangements.

Last, our analysis of the SARS-CoV-2 interactome using the VirHostNet 2.0 interology webservice (*35*) (Fig. 6) revealed that the predicted SARS-CoV-2 network appears to be wired to the Bcl-2 apoptotic switch by the nonstructural protein NS7a and by protein E. Interestingly, NS7a contacts the five major anti-apoptotic Bcl-2 proteins (Bcl-xL/BCL2L1, BCL2, BCL-w/BCL2L2, MCL-1 and Bfl-1/BCL2A1), suggesting that this protein might act as a universal prosurvival Bcl-2 protein modulator after infection. Because many viruses have evolved strategies to interfere with apoptosis, research is needed to establish if the E protein from SARS-CoV-2 directly or indirectly targets prosurvival Bcl-2 homologs (beyond Bcl-xL) to modulate cell viability. Contradictory results have been published regarding CoV E ability to induce or repress apoptosis. Overexpressed epitope-tagged SARS-CoV E was found to promote apoptotic cell death in the Jurkat cell line (*21*), and 17Cl-1 cells infected with the mouse hepatitis virus (MHV) (murine coronavirus) exhibited a higher level of apoptosis (*36*). In contrast, cells infected by a coronavirus lacking E (SARS-CoVΔE) were more primed for apoptosis than those infected with the *wt* strain (*16*). This latter study also suggested that CoV E may have an anti-apoptotic function by suppressing the Unfolded Protein Response (UPR), likely as a survival mechanism important for virus dissemination. It may well be that CoV E activity could protect cells from cell death at early stages post-infection (by regulating the UPR) and induce apoptosis at a later stage (via Bcl-xL sequestration at the Golgi (*21*) or through its

viroporin activity). Yang et al. (*21*) further demonstrated that the interaction of SARS-CoV E with Bcl-xL is mediated by the BH3-like region of SARS-CoV E and the BH3-binding motif of Bcl-xL. Anti-apoptotic Bcl-2 proteins normally accommodate the BH3 motif from typical BH3-only proteins in a hydrophobic receptor cleft present at their surface (*25-28*). Hence, it remains to be determined whether SARS-CoV-2 E BH3 could bind in the hydrophobic groove of Bcl-xL or at another, distinct interface. It might also be interesting to assess the ability of BH3-mimetics like ABT-737, which tightly binds to the hydrophobic BH3-binding groove of Bcl-xL (*37*), to inhibit binding of SARS-CoV-2 E to Bcl-xL.

In conclusion, the SARS-CoV E protein has been reported to interact with the anti-apoptotic Bcl-2 family protein Bcl-xL through a BH3-like motif located in its C-terminal portion. Considering the importance of Bcl-2 homologs and BH3-containing proteins in physiopathological settings, the relevance of this interaction and its extrapolation to CoV E from the novel coronavirus SARS-CoV-2 call for an urgent assessment. In particular, detailed structural characterization and the identification of small molecule drugs able to disrupt this interaction could provide a targeted therapeutic approach where coronavirus treatments are currently lacking or scarce at best. Recombinant SARS-CoV-2 deleted or mutated in the BH3-like region of E might also prove valuable for the design of a live, attenuated vaccine, with an E protein sufficiently intact to maintain structural integrity necessary for the virus life cycle, but mutated enough to abolish protein-protein interactions involved in viral pathogenicity.

Material and methods

Database mining, sequence alignments and phylogenetic reconstruction

UniProtKB and GenBank nucleotide and protein sequence databases were searched for sequences related to human SARS-CoV E (Supplemental Table 1). Pangolin CoV E nucleic sequence was *de novo* assembled using metaSPADes v3.13.0 (*38*) from PRJNA573298 project raw reads data (*39*) after carefully removing reads mapped on *Manis javanica* genomic sequences (INSDC: JSZB00000000.1) using bwa. Homologous sequences were aligned with the ClustalW alignment program implemented in the graphical multiple sequence alignment editor SeaView (*40*) or at NPS@ (*41*) using BLOSUM alignment matrices followed by manual optimization after visual inspection. Consensus sequence logos were generated with the online software

WebLogo. A phylogenetic tree was calculated on 67 sites, with 1000 bootstrap replicates, using the BIONJ algorithm (*42*) embedded in Seaview v5.0.2.

Structural analysis

The SARS-CoV-2 E protein has been modeled using M4T (*43*) and the known 3D structure of SARS-CoV E as template (PDB code: 5X29). The software PyMOL (https://www.pymol.org) was used for preparing the illustrations.

Interactome analysis

The predicted protein-protein interaction network of SARS-CoV-2 was constructed by using the VirHosNet 2.0 interology web service (*35*) on the COVID-19/SARS-CoV-2 special release (March 2020). We used the SARS-CoV-2 proteins annotated from Uniprot (https://covid-19.uniprot.org/) as a seed and inferred interology using gold standard protein-protein interaction annotated in VirHostNet. Interologs were defined based on known SARS-CoV protein-protein interactions by using 95% percent identity and 95 % protein coverage blastp thresholds. The putative SARS-CoV2 interactome comprised 114 proteins and 218 protein-protein interactions.

Acknowledgments

This work was performed using the computing facilities of the LBBE/PRABI-AMSB. We gratefully acknowledge the authors, originating and submitting laboratories of the sequences from GISAID's EpiCov Database (see Acknowledgment Table).

Funding

This research was supported by CNRS and Montpellier University.

Conflict of interest

None declared.

Legend to figures

Figure 1. Alignment of BH3 regions from Bcl-2 family members, classical BH3-only proteins and other BH3-bearing sequences. Accession identifiers, species

abbreviations, sequence names and PubMed Identifiers are given (the figure was adapted from (*22*)). Numbers refer to starting and ending positions of each BH3 region. The conserved Leu and Asp residues are highlighted in red. The additional BH3 sequences reported in human BID-BH3-B, mouse NOXA and human ATR were omitted for the sake of clarity. ANTI: anti-apoptotic; PRO: pro-apoptotic; Act: activator BH3-only protein; Sen: sensitizer BH3-only protein. STRUCT: experimentally determined structured protein; STRUCT (pred): protein predicted to be structured; IDP: protein experimentally shown as being disordered; IDP (pred): protein predicted to be intrinsically disordered; MIX: protein experimentally shown to contain both disordered and structured regions. MIX (pred): protein predicted to contain both disordered and structured regions. Residue conservation is indicated in shades of colors (darker represents higher degree of conservation). Fixed amino acid positions can hardly be identified in BH3 and BH3-like sequences, and even the typical Leu and Asp residues are not universally conserved among all BH3-bearing proteins. Thus, structural determinations of complexes between BH3 peptides (or BH3-containing proteins) and Bcl-2 homologous proteins and experimental (biochemical and/or cell-based) evidence for direct interactions between them should be considered as additional criteria to validate BH3 motifs, in addition to sequence-based inference (*25*). Note that the BH3 motif is an unusual SLiM type as it is found both in structured proteins, including Bcl-2 homologs, and in intrinsically disordered proteins (IDPs), such as classical BH3-only proteins (*22*). Species abbreviations are as follows: Hs, *Homo sapiens;* Mm, *Mus musculus;* Dm, *Drosophila melanogaster;* Hv, *Hydra vulgaris;* Lb, *Lubomirskia baikalensis;* Aq, *Amphimedon queenslandica;* Am, *Ambystoma mexicanum;* Ce, *Caenorhabditis elegans;* Pl *Photorhabdus luminescens;* NDV, *Newcastle disease virus;* HHV8*, Human herpesvirus 8 (Kaposi's sarcoma-associated herpesvirus);* HBV, *Hepatitis B Virus;* HCV, *Hepatitis C Virus;* Sc, *Saccharomyces cerevisiae;* Sp, *Schizosaccharomyces pombe;* SARS: *human SARS-CoV;* SARS2: *human SARS-CoV-2.*

<u>Figure 2.</u> Topology of CoV E proteins from betacoronaviruses and sequence alignment showing the BH3-like motif. The CoV E protein consists of three domains: (i) the amino (N)-terminal domain; (ii) the transmembrane domain (TMD); (iii) and the carboxy (C)-terminal domain, which comprises a BH3-like motif and a PDZ-binding motif (PBM) at its extremity. The CoV E protein sequences of selected betacoronaviruses were aligned. The region encompassing the BH3-like region is shown. The conserved Leu and Pro residues are highlighted in red. Red circles indicate coronaviruses infecting humans. Bottom: logo map of the aligned amino acid segments.

Figure 3. Sequence alignment of the BH3-like region of CoV E proteins from other coronaviruses. The CoV E protein sequences of selected alpha-, gamma- and deltacoronavirus were aligned. The region encompassing the BH3-like region (as identifiable in the human SARS-CoV ZJ01 reference sequence) is shown. The conserved Leu and Pro residues are highlighted in red. Red circles: coronaviruses infecting humans. Bottom: logo map of the aligned amino acid segments.

Figure 4. Phylogenetic tree of CoV E proteins from SARS-CoV-2 and other representative coronaviruses. Red circles indicate viruses infecting human hosts. SARS-CoV-2 sequences are in white. The tree was calculated using the BioNJ method on the C-terminal domain of CoV E proteins encompassing the BH3-like motif (or homologous site). Numbers indicate bootstrap percentages after 1000 replications; values below 50% are not reported. Branch lengths are proportional to distances between sequences.

Figure 5. Structural analysis of SARS-CoV and SARS-CoV-2 E proteins.

(A) Experimentally determined NMR structure of the SARS-CoV E protein pentamer (PDB code: 5X29). (B) Homology-driven model of the SARS-CoV-2 E protein based on 5X29. Residues from the BH3-like core segment (LVKPTVYVY) are shown in red. The positions of the conserved Leu and Pro residues are indicated (note that the numbering refers to the actual positions within the protein sequence and not to the positions *1* and *4* within the BH3-like sequence consensus described in the text). Deleted amino acids from the EPI_ISL_41375225865_26065 translated sequence are colored in white. Other detected point mutations are indicated. Note that the 3D model lacks ten residues at the C-terminus, as they were also absent in 5X29. Thus, mutations detected in the PBM were simply put at the top. (C) Sequence alignment between SARS-CoV-2 E reference sequence and translated short reads from coronavirus genomes. Except for the reference human SARS-CoV-2 E sequence (GenBank accession number YP_0097243921), other sequences with EPI-ISL-* accession numbers (Supplemental Table 2) were retrieved from GISAID (https://www.gisaid.org/) (*44*) and annotated using tblastn. Non-consensus residues (different from the reference SARS-CoV-2 E sequence) are highlighted in red. The deletion of eight residues detected in one putative SARS-CoV-2 E sequence is framed. Asterisks in the consensus line indicate positions that are fully conserved. Colons indicate positions that have strongly similar residues and dots are assigned to weakly similar positions.

Figure 6. SARS-CoV-2:human predicted interactome.

SARS-CoV-2/human interologs are anchored on the known SARS-CoV/human protein-protein interaction network. Each node corresponds to a pair of homologous proteins: a SARS-CoV homolog (red node) and a human host protein (blue node). Node size is proportional to the number of cellular partners in the infected cell. Each edge in the interolog network represents an interaction conserved in both species (interolog). The predicted protein-protein interaction between SARS-CoV-2 E (VEMP_CVHSA) and BCL2L1 (BCL2L1_HUMAN) is highlighted in yellow. Nodes that represent anti-apoptotic Bcl-2 family members are colored in green.

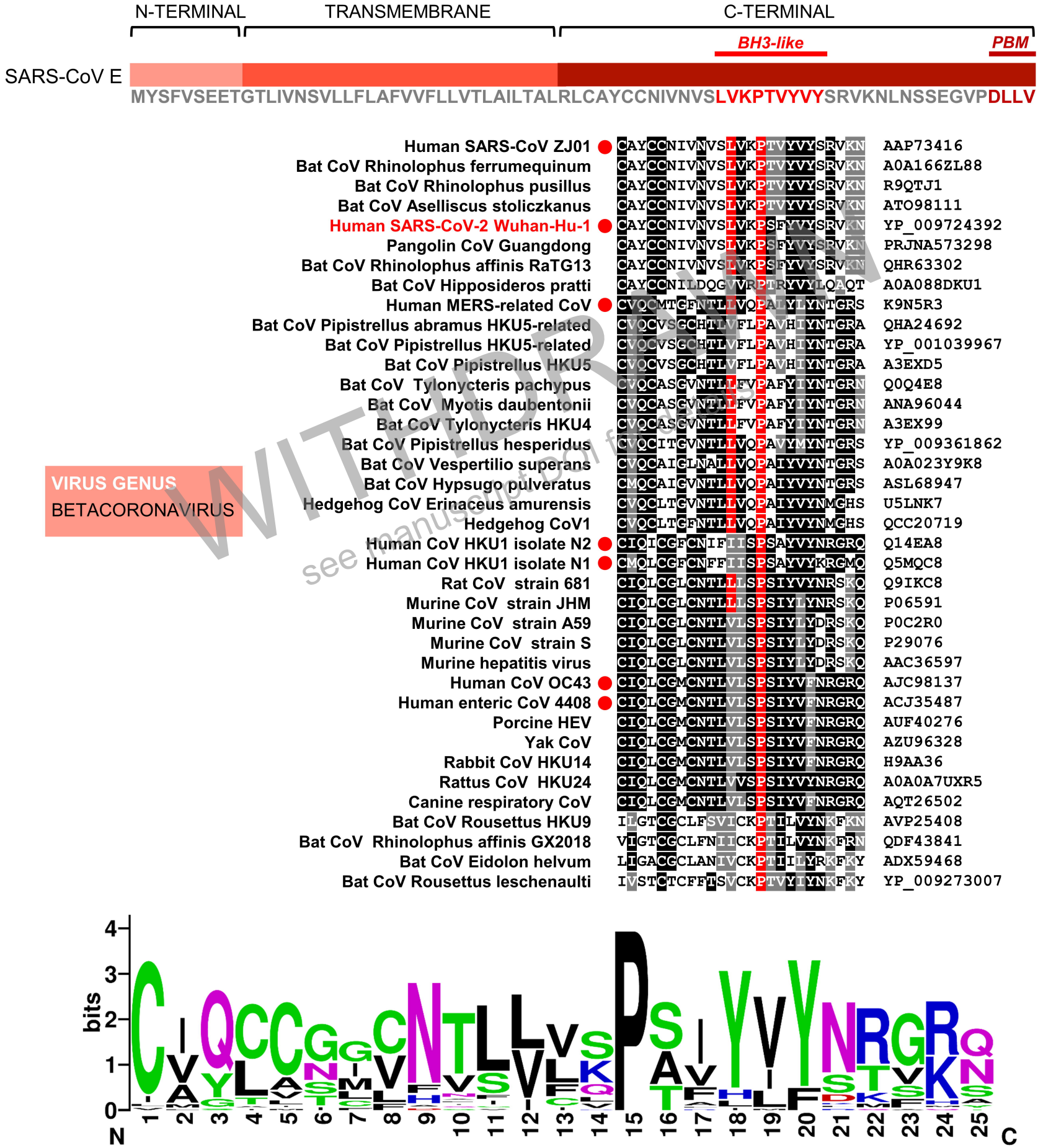
References

1. N. Zhu *et al.*, A Novel Coronavirus from Patients with Pneumonia in China, 2019. *The New England journal of medicine* **382**, 727-733 (2020).
2. K. G. Andersen, A. Rambaut, W. I. Lipkin, E. C. Holmes, R. F. Garry, The proximal origin of SARS-CoV-2. *Nature medicine*, (2020).
3. P. Zhou *et al.*, A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270-273 (2020).
4. F. Wu *et al.*, A new coronavirus associated with human respiratory disease in China. *Nature* **579**, 265-269 (2020).
5. C. A. de Haan, P. J. Rottier, Molecular interactions in the assembly of coronaviruses. *Adv Virus Res* **64**, 165-230 (2005).
6. P. S. Masters, The molecular biology of coronaviruses. *Adv Virus Res* **66**, 193-292 (2006).
7. D. Schoeman, B. C. Fielding, Coronavirus envelope protein: current knowledge. *Virology journal* **16**, 69 (2019).
8. L. Wilson, C. McKinlay, P. Gage, G. Ewart, SARS coronavirus E protein forms cation-selective ion channels. *Virology* **330**, 322-331 (2004).
9. W. Surya, Y. Li, J. Torres, Structural model of the SARS coronavirus E channel in LMPG micelles. *Biochim Biophys Acta Biomembr* **1860**, 1309-1317 (2018).
10. D. X. Liu, Q. Yuan, Y. Liao, Coronavirus envelope protein: a small membrane protein with multiple functions. *Cellular and molecular life sciences : CMLS* **64**, 2043-2048 (2007).
11. H. Vennema *et al.*, Nucleocapsid-independent assembly of coronavirus-like particles by co-expression of viral envelope protein genes. *The EMBO journal* **15**, 2020-2028 (1996).
12. J. Ortego, J. E. Ceriani, C. Patino, J. Plana, L. Enjuanes, Absence of E protein arrests transmissible gastroenteritis coronavirus maturation in the secretory pathway. *Virology* **368**, 296-308 (2007).
13. K. M. Curtis, B. Yount, R. S. Baric, Heterologous gene expression from transmissible gastroenteritis virus replicon particles. *Journal of virology* **76**, 1422-1434 (2002).
14. J. Ortego, D. Escors, H. Laude, L. Enjuanes, Generation of a replication-competent, propagation-deficient virus vector based on the transmissible gastroenteritis coronavirus genome. *Journal of virology* **76**, 11518-11529 (2002).
15. M. L. DeDiego *et al.*, A severe acute respiratory syndrome coronavirus that lacks the E gene is attenuated in vitro and in vivo. *Journal of virology* **81**, 1701-1713 (2007).

16. M. L. DeDiego *et al.*, Severe acute respiratory syndrome coronavirus envelope protein regulates cell stress response and apoptosis. *PLoS pathogens* **7**, e1002315 (2011).
17. M. L. Dediego *et al.*, Pathogenicity of severe acute respiratory coronavirus deletion mutants in hACE-2 transgenic mice. *Virology* **376**, 379-389 (2008).
18. K. T. Teoh *et al.*, The SARS coronavirus E protein interacts with PALS1 and alters tight junction formation and epithelial morphogenesis. *Molecular biology of the cell* **21**, 3838-3852 (2010).
19. J. M. Jimenez-Guardeno *et al.*, The PDZ-binding motif of severe acute respiratory syndrome coronavirus envelope protein is a determinant of viral pathogenesis. *PLoS pathogens* **10**, e1004320 (2014).
20. K. Luck, S. Charbonnier, G. Trave, The emerging contribution of sequence context to the specificity of protein interactions mediated by PDZ domains. *FEBS letters* **586**, 2648-2661 (2012).
21. Y. Yang *et al.*, Bcl-xL inhibits T-cell apoptosis induced by expression of SARS coronavirus E protein in the absence of growth factors. *The Biochemical journal* **392**, 135-143 (2005).
22. A. Aouacheria, C. Combet, P. Tompa, J. M. Hardwick, Redefining the BH3 Death Domain as a 'Short Linear Motif'. *Trends in biochemical sciences* **40**, 736-748 (2015).
23. M. Kvansakul, M. G. Hinds, The structural biology of BH3-only proteins. *Methods in enzymology* **544**, 49-74 (2014).
24. J. A. Glab, G. W. Mbogo, H. Puthalakath, BH3-Only Proteins in Health and Disease. *Int Rev Cell Mol Biol* **328**, 163-196 (2017).
25. A. Aouacheria, V. Rech de Laval, C. Combet, J. M. Hardwick, Evolution of Bcl-2 homology motifs: homology versus homoplasy. *Trends in cell biology* **23**, 103-111 (2013).
26. S. Cory, A. W. Roberts, P. M. Colman, J. M. Adams, Targeting BCL-2-like Proteins to Kill Cancer Cells. *Trends Cancer* **2**, 443-460 (2016).
27. J. D. Huska, H. M. Lamb, J. M. Hardwick, Overview of BCL-2 Family Proteins and Therapeutic Potentials. *Methods Mol Biol* **1877**, 1-21 (2019).
28. J. Montero, A. Letai, Why do BCL-2 inhibitors work and where should we use them in the clinic? *Cell death and differentiation* **25**, 56-64 (2018).
29. T. T. Lam *et al.*, Identifying SARS-CoV-2 related coronaviruses in Malayan pangolins. *Nature*, (2020).
30. J. R. Cohen, L. D. Lin, C. E. Machamer, Identification of a Golgi complex-targeting signal in the cytoplasmic tail of the severe acute respiratory syndrome coronavirus envelope protein. *Journal of virology* **85**, 5794-5803 (2011).
31. Y. Li, W. Surya, S. Claudine, J. Torres, Structure of a conserved Golgi complex-targeting signal in coronavirus envelope proteins. *The Journal of biological chemistry* **289**, 12535-12549 (2014).
32. S. Warren, X. F. Wan, G. Conant, D. Korkin, Extreme evolutionary conservation of functionally important regions in H1N1 influenza proteome. *PloS one* **8**, e81027 (2013).
33. S. Cheng, C. L. Brooks, 3rd, Viral capsid proteins are segregated in structural fold space. *PLoS computational biology* **9**, e1002905 (2013).
34. S. Munro, An investigation of the role of transmembrane domains in Golgi protein retention. *The EMBO journal* **14**, 4695-4704 (1995).
35. T. Guirimand, S. Delmotte, V. Navratil, VirHostNet 2.0: surfing on the web of virus/host molecular interactions data. *Nucleic acids research* **43**, D583-587 (2015).
36. S. An, C. J. Chen, X. Yu, J. L. Leibowitz, S. Makino, Induction of apoptosis in murine coronavirus-infected cultured cells and demonstration of E protein as an apoptosis inducer. *Journal of virology* **73**, 7853-7859 (1999).
37. T. Oltersdorf *et al.*, An inhibitor of Bcl-2 family proteins induces regression of solid tumours. *Nature* **435**, 677-681 (2005).

38. S. Nurk, D. Meleshko, A. Korobeynikov, P. A. Pevzner, metaSPAdes: a new versatile metagenomic assembler. *Genome research* **27**, 824-834 (2017).
39. P. Liu, W. Chen, J. P. Chen, Viral Metagenomics Revealed Sendai Virus and Coronavirus Infection of Malayan Pangolins (Manis javanica). *Viruses* **11**, (2019).
40. M. Gouy, S. Guindon, O. Gascuel, SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular biology and evolution* **27**, 221-224 (2010).
41. C. Combet, C. Blanchet, C. Geourjon, G. Deleage, NPS@: network protein sequence analysis. *Trends in biochemical sciences* **25**, 147-150 (2000).
42. O. Gascuel, BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. *Molecular biology and evolution* **14**, 685-695 (1997).
43. D. Rykunov, E. Steinberger, C. J. Madrid-Aliste, A. Fiser, Improved scoring function for comparative modeling using the M4T method. *J Struct Funct Genomics* **10**, 95-99 (2009).
44. Y. Shu, J. McCauley, GISAID: Global initiative on sharing all influenza data - from vision to reality. *Euro Surveill* **22**, (2017).

									PMID:	
BCL-2 homologs	P10415	Hs	BCL2	86	LSPVPPVVHLTL-RQAG--DDFSRRYRRDFAEM	115	ANTI	6093263	STRUCT
	Q07817	Hs	BCL2L1	79	EVIPMAAVKQAL-REAG--DEFELRYRRAFSDL	108	ANTI	8358789	STRUCT
	Q92843	Hs	BCL2L2	35	EGPAADPLHQAM-RAAG--DEFETRFRRTFSDL	64	ANTI	8761287	STRUCT
	Q07820	Hs	MCL1	202	SGATSRKALETL-RRVG--DGVQRNHETAFQGM	231	ANTI	7682708	STRUCT
	Q9HD36	Hs	BCL2L10	28	EPAPSTPEAAVL-RSAA--ARLRQIHRSFFSAY	57	ANTI	11593390	STRUCT
	Q16548	Hs	BCL2A1	46	PGAGPSQEENVA-EEVCRYVNKNLKQL	75	ANTI	8345191	STRUCT
	P41958	Ce	CED9	105	LPCGVQPEHEMM-RVMG--TIFEKKHAENFETF	134	ANTI	7907274	STRUCT
	Q8T8Y5	Dm	BUFFY	128	SMGIVRDVFPAV-QVLG--DELERMHPRIYNGV	157	ANTI	10801447	STRUCT (pred)
	Q9V9C8	Dm	DEBCL	135	SSHVVYEVFPAL-NSMG--EELERMHPRVYTNI	164	PRO	10684252	STRUCT (pred)
	Q9UMX3	Hs	BOK	59	VPGRLAEVCAVL-LRLG--DELEMIRPSVYRNV	88	PRO	9535847	STRUCT (pred)
	Q07812	Hs	BAX	52	QDASTKKLSECL-KRIG--DELDSNMELQRMIA	81	PRO	8358790	STRUCT
	Q16611	Hs	BAK1	67	PSSTMGQVGRQL-AIIG--DDINRRYDSEFQTM	96	PRO	7715730	STRUCT
	Q9HB09	Hs	BCL2L12	217	KEAILRRLVALL-EEEA--EVINQKLASDPALR	235	PRO	11401436	STRUCT (pred)
	Q9BXK5	Hs	BCL2L13	93	PANPESSMEDCL-AHLG--EKVSQELKEPLHKA	122	PRO	11262395	STRUCT (pred)
	Q9BZR8	Hs	BCL2L14	205	EEQILAKIVELL-KYSG--DQLERKLKKDKALM	234	PRO	11054413	STRUCT (pred)
	Q5TBC7	Hs	BCL2L15	46	CSFDVAIIAGRL-RMLG--DQFNGELEASAKNV	75	PRO	12700646	STRUCT (pred)
	P55957	Hs	BID (A)	79	QEDIIRNIARHL-AQVG--DSMDRSIPPGLVNG	108	PRO (Act)	8918887	STRUCT
	A7LM80	Hv	BCL2-like	39	LKPPVIEIAKTL-QRVG--DELESANTDFFKNM	68	ND	24065732	STRUCT (pred)
	I1GDG3	Aq	BCL2-like	89	GESNAKAVGRRL-AELG--DEIDSMYSGHFEDI	118	ND	20686567	STRUCT (pred)
	Q1RPT5	Lb	BCL2-like	64	ASSMASEVGRRL-AEFG--DQVDGQFYQEMEEI	93	ND	16569194	STRUCT (pred)
	B7SB93	Am	BCL2-like	35	VNESPSAAACHL-RRVA--DELIEENRQLFDSM	64	ND	Prediction	STRUCT (pred)
Classical BH3-only	O43521	Hs	BIM	141	DMRPEIWIAQEL-RRIG--DEFNAYYARRVFLN	170	PRO (Act)	9430630	IDP
	Q9BXH1	Hs	PUMA	130	EEQWAREIGAQL-RRMA--DDLNAQYERRRQEE	159	PRO (Act)	11572983	IDP
	Q92934	Hs	BAD	103	NLWAAQRYGREL-RRMS--DEFVDSFKKGLPRP	132	PRO (Sen)	7834748	IDP
	Q13323	Hs	BIK	50	CMEGSDALALRL-ACIG--DEMDVSLRAPRLAQ	79	PRO (Sen)	7478623	STRUCT (pred)
	Q96LC9	Hs	BMF	126	QHQAEVQIARKL-QCIA--DQFHRLHVQQHQQN	155	PRO (Sen)	11546872	IDP
	O00198	Hs	HRK	26	RSSAAQLTAARL-KALG--DELHQRTMWRRRAR	55	PRO (Sen)	9130713	IDP
	Q13794	Hs	NOXA	18	PAELEVECATQL-RRFG--DKLNFRQKLLNLIS	47	PRO (Act)	10807576	IDP
	O61667	Ce	EGL1	62	ISSIGYEIGSKL-AAMC--DDFDAQMMSYSAHA	91	PRO	9604928	IDP
Other BH3-containing proteins	P21980	Hs	TG2	193	CLILLDVNPKFL-KNAG-RDCSRRSSPVYVGRV	223	PRO	15485857	STRUCT
	O94817	Hs	ATG12	48	AGDTKKKIDILL-KAVG--DTPIMKTKKWAVER	77	PRO	22152474	STRUCT
	Q99638	Hs	RAD9A	9	NVKVLGKAVHSL-SRIG--DELYLEPLEDGLSL	38	PRO	10620799	STRUCT
	P26306	Sp	SpRAD9	9	NLRDLARIFTNL-SRID--DAVNWEINKNQIEI	38	PRO	10996309	MIX (pred)
	Q9TY06	Hs	CED13	48	SNTVEYNIGRKL-TVMC--DEFDSELMSYKEEK	77	PRO	15605074	STRUCT (pred)
	O14791	Hs	APOL1	147	LKSELEDNIRRL-RALA--DGVQKVHKGTTIAN	176	PRO	18505729	STRUCT (pred)
	Q9BQE5	Hs	APOL2	88	LKRELEDHIRKL-RALA--EEVEQVHRGTTIAN	117	NEUTRAL	24901046	STRUCT (pred)
	Q9BWW8	Hs	APOL6	49	LKEDLKGNIDKL-RALA--DDIDKTHKKFTKAN	69	PRO	15671246	STRUCT (pred)
	Q14457	Hs	BECN1	105	DGGTMENLSRRL-KVTG--DLFDIMSGQTDVDH	134	PRO	17337444	STRUCT (pred)
	Q8VE73	Hs	CUL7	1555	RNLLNCLVVRIL-KAHG--DEGLHVDRLVYLVL	1584	PRO	10652310	STRUCT (pred)
	Q15303	Hs	ERBB4	974	FSRMARDPQRYL-VIQG--DDRMKLPSPNDSKF	1003	PRO	16778220	STRUCT (pred)
	P04626	Hs	ERBB2	1109	LQRYSEDPTVPL-PSET--DGYVAPLTCSPQPE	1138	PRO	18420586	STRUCT (pred)
	Q9Y287	Hs	ITM2B	80	ALQPDDVYYCGI-KYIK--DDVILNEPSADAPA	108	PRO	12082633	STRUCT (pred)
	Q96BY2	Hs	MOAP1	109	FLAGEGMTVGELSRALGH-ENGSLDPEQGMIP-	139	PRO	11060313	STRUCT (pred)
	Q8IZY5	Hs	BRCC2	1	-------MVTLL-PIEG--QEIHFFEILESECV	23	PRO	15069058	STRUCT (pred)
	Q9Y5Z4	Hs	SOUL	151	NQEQLLTLASIL-REDG--KVFDEKVYYTAGYN	180	PRO	19901022	STRUCT (pred)
	Q12981	Hs	BNIP1	103	KIAIDNLEKAEL-LQGG--DLLRQRKTTKESLA	132	PRO	12766483	STRUCT (pred)
	Q12982	Hs	BNIP2	77	SGEIDLDGLDTP-SENS--NEFEWEDDLPKPKT	106	PRO	12766483	STRUCT (pred)
	Q12983	Hs	BNIP3	99	DIERRKEVESIL-KKNS--DWIWDWSSRPENIP	128	PRO	9575197	MIX (pred)
	O60238	Hs	BNIP3L	123	EVVEGEKEVEAL-KKSA--DWVSDWSSRPENIP	152	PRO	12766483	MIX (pred)
	Q09969	Ce	CEBNIP3	118	EQVKYKLVREML-PPGKNTDWIWDWSSRPENTP	149	PRO	11114722	MIX (pred)
	Q7L3V2	Hs	BOP	107	FDGSPWLLDRFL-AQLG--DYMSFHFEHYQDNI	136	PRO	23055042	MIX (pred)
	P10909	Hs	CLU	312	DCSTNNPSQAKL-RREL--DESLQVAERLTRKY	341	PRO	21527247	MIX (pred)
	Q7Z6Z7	Hs	HUWE	1969	PGVMTQEVGQLL-QDMG--DDVYQQYRSLTRQS	1998	PRO	15989957	MIX (pred)
	Q9NZZ3	Hs	CHMP5	156	PELDEDDLEAEL-DALG--DELLADEDSSYLDE	185	PRO	12594175	MIX (pred)
	Q9NRA0	Hs	SPHK2	243	TVSGDGLLHEVL-NGLL--DRPDWEEAVKMPVG	272	PRO	12835323	MIX (pred)
	Q9NQS1	Hs	AVEN	134	ESQRGTDFSVLL-SSAG--DSFSQFRFAEEKEW	162	ANTI	22754595	IDP (pred)
	Q9C0C7	Hs	AMBRA1	1204	SSPQPSTSRGLL-PEAG--QLAERGLSPRTASW	1233	PRO	28215535	MIX (pred)
	Q13535	Hs	ATR	466	DMNQKSILWSAL-KQKA--ESLQISLEYSGLKN	484	ANTI	26387736	STRUCT
	Q14643	Hs	ITPR1	2583	FMVIIIVLNLIF-GVII--DTFADLRSEKQKKE	2637	PRO	26976600	STRUCT (pred)
	P13693	Hs	TPT1	9	SHDEMFSDIYKI-REIA--DGLCLEVEGKMVSR	38	ANTI	26813996	STRUCT
	P57764	Hs	GMDMD	139	LRQPEHKVLQQL-RSRG--DNVYVVTEVLQTQK	168	PRO	31839993	STRUCT
	Q8NB16	Hs	MLKL	158	IEASLRRLEINM-KEIK--ETLRQYLPPKCMQE	187	PRO	31839993	STRUCT
	Q8K459	Mm	PXT1	1	---------MQL-RHIG--DSVNHRVIQEHLAQ	21	PRO	21460186	STRUCT (pred)
	Q0GET0	Mm	BLM-S	84	MQLDFQQFLMKL-EKLT--DIRPIPDKEFVETY	113	PRO	25263558	STRUCT (pred)
	C8ZGL9	Sc	YNL305CP	279	IVNLFLSILRIL-ANSN--DDN----------	297	PRO	21673659	STRUCT (pred)
	Q8KT65	PI	MCF	904	QAVKDLELKAGL-TSVG--DGFEPRQGSADIHQ	933	PRO	12136122	STRUCT (pred)
	Q9QBF1	HBV-B	P (HBSP)	14	LDDEAGPLEEEL-PRLA--DEGLNRRVAEDLNL	43	PRO	17049490	STRUCT (pred)
	Q05499	HBV-F	HBX	109	EEYIKDCVFKDW-EELG--EEIRLKVFVLGGCR	138	PRO	16274670	STRUCT (pred)
	P88946	HHV8	VIRF1	163	RMLAALRRTRGL-QEIG--KGISQDGHHFLVFR	192	ANTI	22685405	STRUCT
	Q9DLD4	NDV	F0	81	LDAYNRTLTTLL-TPLG--DSIRRIQESVTTSG	110	PRO	21810274	STRUCT
	Q9WBL3	NDV	M	16	PSSSLLAFPIVL-QDTG--DGKKQITPQYRIQR	45	PRO	21810274	STRUCT
	Q6X1B6	NDV	L	2111	EVTILGLRVKDL-NKVG--DVIGLVLRGMVSLE	2140	PRO	21810274	STRUCT (pred)
	Q00269	HCV	CORE	108	GPTDPRRRSRNL-GKVI--DTLTCGFADLMGYI	137	PRO	19605477	MIX (pred)
	AAP73416	SARS-CoV E		40	CAYCCNIVNVSL-VKPT--VYVYSRVKNLNSSE	69	PRO	16048439	MIX
	YP_009724392	SARS-CoV-2 E		40	CAYCCNIVNVSL-VKPS--FYVYSRVKNLNSSR	69	ND	ND	MIX (pred)



VIRUS GENUS
ALPHACORONAVIRUS

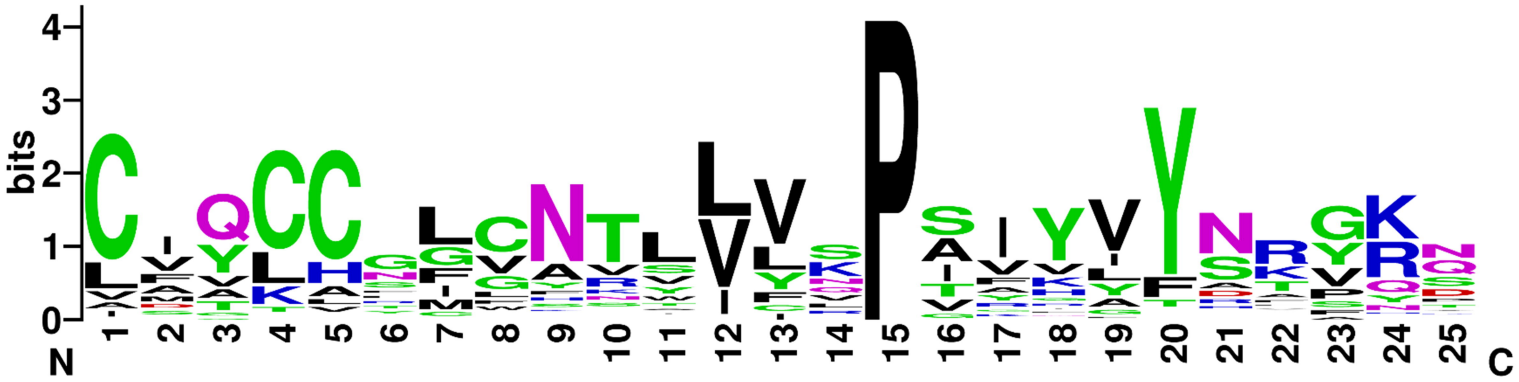
Human CoV NL63	CFTCHYFFSRTLYQPVYKIFLAYQD	QED88021
Human CoV 229E	CFTCHMFCNRTVYGPIKNVYHIYQS	QEO75987
Bat CoV Tylonycteris robustula HKU33	CMVCHHLMSGAVYRPVYRVYELYKD	QCX35162
Bat CoV Nyctalus velutinus	CFACHQLMTSAVYKPVHNAYAIYRD	AIA62267
Bat CoV Myotis ricketti	CIACHQFANRTVYSPIYSAYQWYKD	AIA62248
Bat CoV Scotophilus kuhlii GX2018A	CFTCHRLCNNVVYKPVGKVYGVYKS	QDF43791
Bat CoV Scotophilus kuhlii 512r	CFTCHRLCNSVVYRPVGKFYRVYKS	QHA24698
Bat CoV Miniopterus fuliginosus	LFSCHRLCSNTVYRPVYGAYRAYQD	AIA62214
Porcine epidemic diarrhea virus	CFTCHRLCNSAVYTPIGRLYRVYKS	APH81308
Apodemus chevrieri AcCoV-JC34	VGVCCSLTNKVIVLPVKGVYHLYQD	A0A1X9JJV7
Lucheng Rn rat CoV	VGVCCSLTNKVIVLPVKGVYHLYQD	A0A1L7HIW4
Canine CoV strain Insavc-1	CMVCCNLGRTVLIVPARHAYDAYKN	P36696
Feline CoV strain FIPV WSU-79/1146	CMVCCNLGKTIIVLPARHAYDAYKT	Q52PA5
Mink CoV strain WD1133	CMVCCRLSNVVIIAPARQAYNAYKD	D9J205

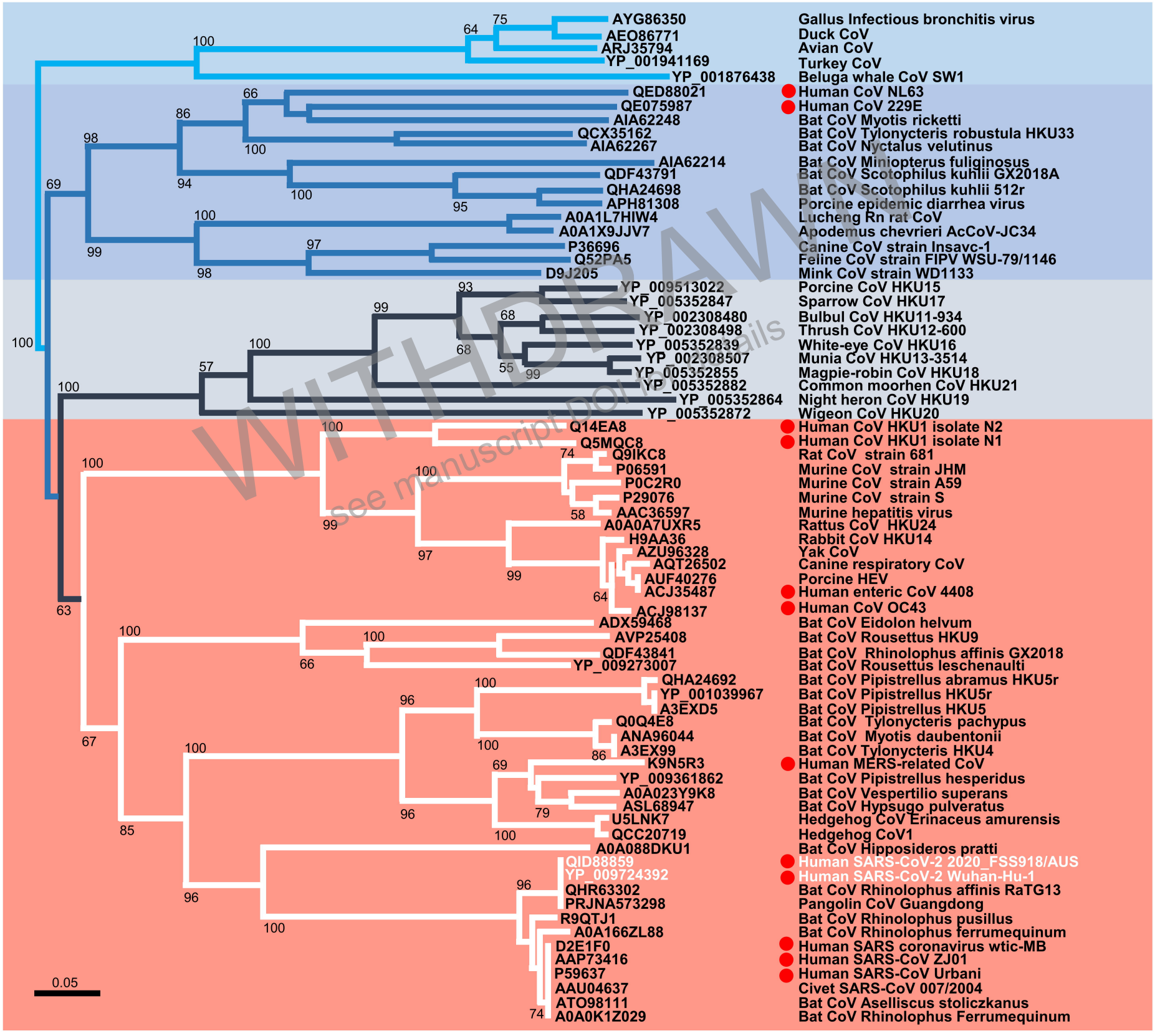
VIRUS GENUS
GAMMACORONAVIRUS

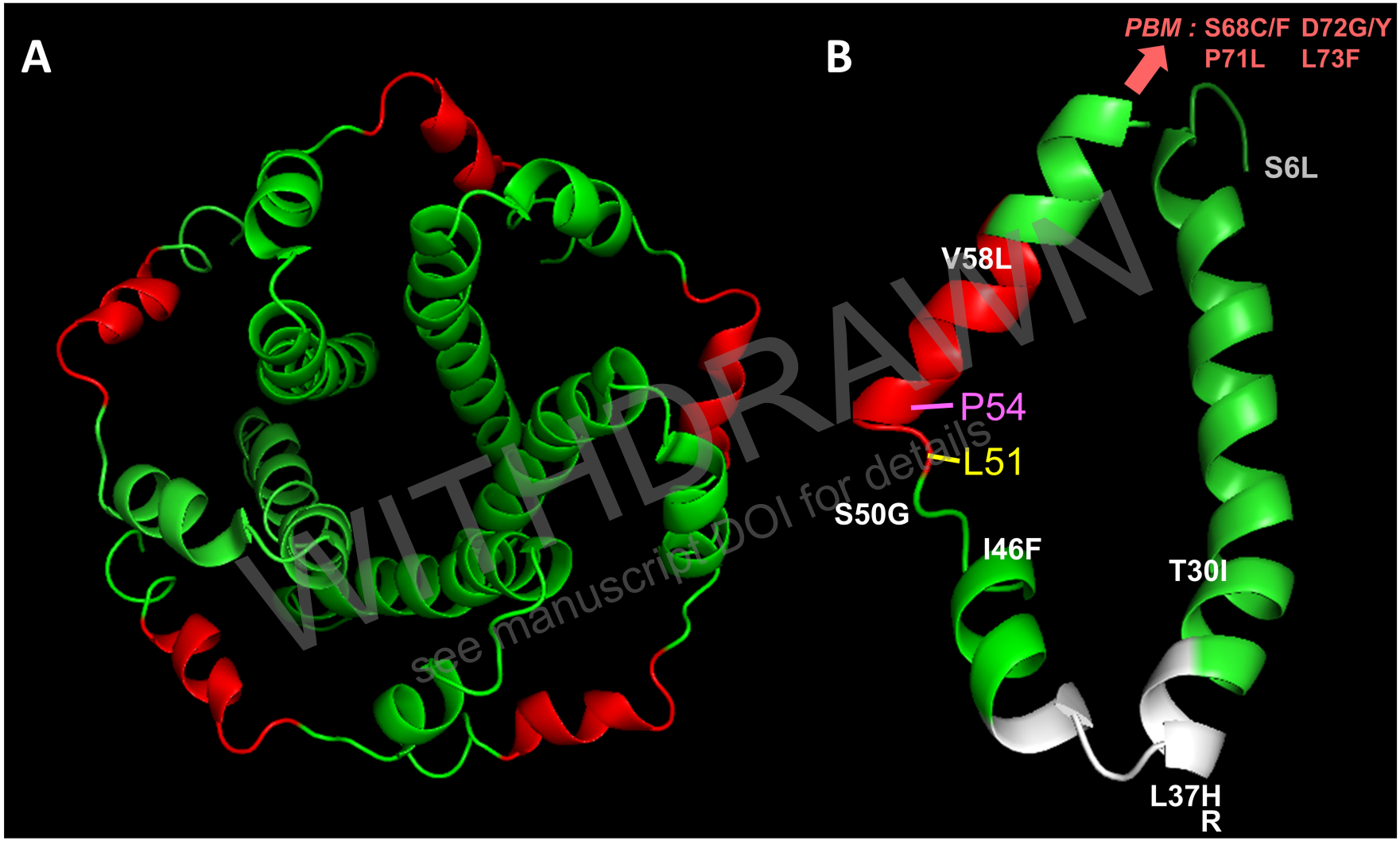
Avian CoV	ADACCLFWYTWLVIPGVKGTAFVYK	ARJ35794
Turkey CoV	ADACCLFWYTWIVVPGAKGAAFVYN	YP_001941169
Gallus Infectious bronchitis virus	ADACCLFWYTWVVVPGAKGTAFVYK	AYG86350
Duck CoV	VDACCLFWYTWVVVPGAKGTAFVYK	AEO86771
Beluga whale CoV SW1	IDVCLLFCKGWAVDPSVRLVSYAYG	YP_001876438

VIRUS GENUS
DELTACORONAVIRUS

Porcine CoV HKU15	LFYKCYLGAAYLVRPIIVYYSKPNP	YP_009513022
Munia CoV HKU13-3514	LSYKCFLGAKYLVNPIIVYYSKPQT	YP_002308507
Bulbul CoV HKU11-934	LSYKCFLGARYLVNPIIVYYSKPNP	YP_002308480
Thrush CoV HKU12-600	LSYKCFLGARYLVNPIIVYYSKPNP	YP_002308498
White-eye CoV HKU16	LSYKCFLGARYLVNPIIVYYSKPKT	YP_005352839
Sparrow CoV HKU17	LTYKCYLGARYLVNPIIVYYSKPNP	YP_005352847
Magpie-robin CoV HKU18	LSYKCFLGAKYLVNPIIVYYSKPQT	YP_005352855
Common moorhen CoV HKU21	LVYKCYKGATVLLNPFIIFSSKVDP	YP_005352882
Night heron CoV HKU19	FFTRTYTCILHTVKPIYVYLKPAPV	YP_005352864




γ
α
δ
β
AYG86350
AEO86771
ARJ35794
YP_001941169
YP_001876438
QED88021
QE075987
AIA62248
QCX35162
AIA62267
AIA62214
QDF43791
QHA24698
APH81308
A0A1L7HIW4
A0A1X9JJV7
P36696
Q52PA5
D9J205
YP_009513022
YP_005352847
YP_002308480
YP_002308498
YP_005352839
YP_002308507
YP_005352855
YP_005352882
YP_005352864
YP_005352872
Q14EA8
Q5MQC8
Q9IKC8
P06591
P0C2R0
P29076
AAC36597
A0A0A7UXR5
H9AA36
AZU96328
AQT26502
AUF40276
ACJ35487
ACJ98137
ADX59468
AVP25408
QDF43841
YP_009273007
QHA24692
YP_001039967
A3EXD5
Q0Q4E8
ANA96044
A3EX99
K9N5R3
YP_009361862
A0A023Y9K8
ASL68947
U5LNK7
QCC20719
A0A088DKU1
QID88859
YP_009724392
QHR63302
PRJNA573298
R9QTJ1
A0A166ZL88
D2E1F0
AAP73416
P59637
AAU04637
ATO98111
A0A0K1Z029
Gallus Infectious bronchitis virus
Duck CoV
Avian CoV
Turkey CoV
Beluga whale CoV SW1
Human CoV NL63
Human CoV 229E
Bat CoV Myotis ricketti
Bat CoV Tylonycteris robustula HKU33
Bat CoV Nyctalus velutinus
Bat CoV Miniopterus fuliginosus
Bat CoV Scotophilus kuhlii GX2018A
Bat CoV Scotophilus kuhlii 512r
Porcine epidemic diarrhea virus
Lucheng Rn rat CoV
Apodemus chevrieri AcCoV-JC34
Canine CoV strain Insavc-1
Feline CoV strain FIPV WSU-79/1146
Mink CoV strain WD1133
Porcine CoV HKU15
Sparrow CoV HKU17
Bulbul CoV HKU11-934
Thrush CoV HKU12-600
White-eye CoV HKU16
Munia CoV HKU13-3514
Magpie-robin CoV HKU18
Common moorhen CoV HKU21
Night heron CoV HKU19
Wigeon CoV HKU20
Human CoV HKU1 isolate N2
Human CoV HKU1 isolate N1
Rat CoV strain 681
Murine CoV strain JHM
Murine CoV strain A59
Murine CoV strain S
Murine hepatitis virus
Rattus CoV HKU24
Rabbit CoV HKU14
Yak CoV
Canine respiratory CoV
Porcine HEV
Human enteric CoV 4408
Human CoV OC43
Bat CoV Eidolon helvum
Bat CoV Rousettus HKU9
Bat CoV Rhinolophus affinis GX2018
Bat CoV Rousettus leschenaulti
Bat CoV Pipistrellus abramus HKU5r
Bat CoV Pipistrellus HKU5r
Bat CoV Pipistrellus HKU5
Bat CoV Tylonycteris pachypus
Bat CoV Myotis daubentonii
Bat CoV Tylonycteris HKU4
Human MERS-related CoV
Bat CoV Pipistrellus hesperidus
Bat CoV Vespertilio superans
Bat CoV Hypsugo pulveratus
Hedgehog CoV Erinaceus amurensis
Hedgehog CoV1
Bat CoV Hipposideros pratti
Human SARS-CoV-2 2020_FSS918/AUS
Human SARS-CoV-2 Wuhan-Hu-1
Bat CoV Rhinolophus affinis RaTG13
Pangolin CoV Guangdong
Bat CoV Rhinolophus pusillus
Bat CoV Rhinolophus ferrumequinum
Human SARS coronavirus wtic-MB
Human SARS-CoV ZJ01
Human SARS-CoV Urbani
Civet SARS-CoV 007/2004
Bat CoV Aselliscus stoliczkanus
Bat CoV Rhinolophus Ferrumequinum
0.05
WITHDRAWN
see manuscript pdf for details



C

BH3-like PBM

```
hSARS-CoV-2 Wuhan-Hu-1          YSFVSEETGTLIVNSVLLFLAFVVFLLVTLAILTALRLCAYCCNIVNVSLVKPSFYVYSRVKNLNSSRVPDLLV
EPI_ISL_41301526223_26447       YSFVLEETGTLIVNSVLLFLAFVVFLLVTLAILTALRLCAYCCNIVNVSLVKPSFYVYSRVKNLNSSRVPDLLV
EPI_ISL_41635426213_26437       YSFVSEETGTLIVNSVLLFLAFVVFLLVILAILTALRLCAYCCNIVNVSLVKPSFYVYSRVKNLNSSRVPDLLV
EPI_ISL_41192926245_26469       YSFVSEETGTLIVNSVLLFLAFVVFLLVTLAILTAHRLCAYCCNIVNVSLVKPSFYVYSRVKNLNSSRVPDLLV
EPI_ISL_41887425796_26020       YSFVSEETGTLIVNSVLLFLAFVVFLLVTLAILTARRLCAYCCNIVNVSLVKPSFYVYSRVKNLNSSRVPDLLV
EPI_ISL_41637426195_26419       YSFVSEETGTLIVNSVLLFLAFVVFLLVTLAILTALRLCAYCCNFVNVSLVKPSFYVYSRVKNLNSSRVPDLLV
EPI_ISL_41830326245_26469       YSFVSEETGTLIVNSVLLFLAFVVFLLVTLAILTALRLCAYCCNIVNVGLVKPSFYVYSRVKNLNSSRVPDLLV
EPI_ISL_41728526246_26470       YSFVSEETGTLIVNSVLLFLAFVVFLLVTLAILTALRLCAYCCNIVNVSLVKPSFYLYSRVKNLNSSRVPDLLV
EPI_ISL_41388926245_26469       YSFVSEETGTLIVNSVLLFLAFVVFLLVTLAILTALRLCAYCCNIVNVSLVKPSFYVYSRVKNLNSCRVPDLLV
EPI_ISL_41641226235_26459       YSFVSEETGTLIVNSVLLFLAFVVFLLVTLAILTALRLCAYCCNIVNVSLVKPSFYVYSRVKNLNSFRVPDLLV
EPI_ISL_41796926205_26429       YSFVSEETGTLIVNSVLLFLAFVVFLLVTLAILTALRLCAYCCNIVNVSLVKPSFYVYSRVKNLNSSRVLDLLV
EPI_ISL_41649626245_26469       YSFVSEETGTLIVNSVLLFLAFVVFLLVTLAILTALRLCAYCCNIVNVSLVKPSFYVYSRVKNLNSSRVPGLLV
EPI_ISL_41468626233_26457       YSFVSEETGTLIVNSVLLFLAFVVFLLVTLAILTALRLCAYCCNIVNVSLVKPSFYVYSRVKNLNSSRVPYLLV
EPI_ISL_41757126245_26469       YSFVSEETGTLIVNSVLLFLAFVVFLLVTLAILTALRLCAYCCNIVNVSLVKPSFYVYSRVKNLNSSRVPDFLV
EPI_ISL_41375225865_26065       YSFVSEETGTLIVNSVLLFLAFVVFLLVTLAI--------YCCNIVNVSLVKPSFYVYSRVKNLNSSRVPDLLV
                                **** ********************** ***         ****:***.*******:********* **  :**
consensus
```