# State-of-the-art tools to identify druggable protein ligand of SARS-CoV-2

Sayed Abdul Azeez<sup>1</sup>, Zahra Ghalib Alhashim<sup>1,2</sup>, Waad Mohammed Al Otaibi<sup>1</sup>, Hind Saleh Alsuwat<sup>1</sup>, Abdallah M. Ibrahim<sup>1,3</sup>, Noor B. Almandil<sup>4</sup>, J. Francis Borgio<sup>1,5</sup>

<sup>1</sup>Department of Genetic Research, Institute for Research and Medical Consultation (IRMC), Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia <sup>2</sup>College of Clinical Pharmacy, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia

<sup>3</sup>Department of Fundamentals of Nursing, College of Nursing, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia

<sup>4</sup>Department of Clinical Pharmacy Research, Institute for Research and Medical Consultation (IRMC), Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia

<sup>5</sup>Department of Epidemic Diseases Research, Institute for Research and Medical Consultation (IRMC), Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia

Submitted: 6 February 2020 Accepted: 25 March 2020

Arch Med Sci 2020; 16 (3): 497–507 DOI: https://doi.org/10.5114/aoms.2020.94046 Copyright © 2020 Termedia & Banach

#### Abstract

**Introduction:** The SARS-CoV-2 (previously 2019-nCoV) outbreak in Wuhan, China and other parts of the world affects people and spreads coronavirus disease 2019 (COVID-19) through human-to-human contact, with a mortality rate of > 2%. There are no approved drugs or vaccines yet available against SARS-CoV-2.

**Material and methods:** State-of-the-art tools based on in-silico methods are a cost-effective initial approach for identifying appropriate ligands against SARS-CoV-2. The present study developed the 3D structure of the envelope and nucleocapsid phosphoprotein of SARS-CoV-2, and molecular docking analysis was done against various ligands.

**Results:** The highest log octanol/water partition coefficient, high number of hydrogen bond donors and acceptors, lowest non-bonded interaction energy between the receptor and the ligand, and high binding affinity were considered for the best ligand for the envelope (mycophenolic acid: log *P* = 3.00;  $\Delta G = -10.2567$  kcal/mol; p*Ki* = 7.713 µM) and nucleocapsid phosphoprotein (1-[(2,4-dichlorophenyl)methyl]pyrazole-3,5-dicarboxylic acid: log *P* = 2.901;  $\Delta G = -12.2112$  kcal/mol; p*Ki* = 7.885 µM) of SARS-CoV-2.

**Conclusions:** The study identifies the most potent compounds against the SARS-CoV-2 envelope and nucleocapsid phosphoprotein through state-of-the-art tools based on an in-silico approach. A combination of these two ligands could be the best option to consider for further detailed studies to develop a drug for treating patients infected with SARS-CoV-2, COVID-19.

**Key words:** COVID-19, SARS-CoV-2, druggable protein, ligand, nucleocapsid phosphoprotein, molecular docking, envelope protein, phylogenetic tree.

#### Introduction

In December 2019, an unknown pneumonia spread amongst a group of people in Wuhan, China, now termed as coronavirus disease 2019

#### Corresponding author:

Dr. J. Francis Borgio Department of Genetic Research Institute for Research and Medical Consultation (IRMC) Imam Abdulrahman Bin Faisal University Dammam, Saudi Arabia Phone: +966 13 3330864 E-mail: fbalexander@iau.edu.sa borgiomicro@gmail.com



(COVID-19). COVID-19 patients were reported with a cluster of acute respiratory illness and higher interleukin 2 (IL-2), IL-7, IL-10, granulocyte colony-stimulating factor (GCSF), interferon gamma-induced protein 10 (IP10), monocyte chemoattractant protein 1 (MCP1), macrophage inflammatory protein  $1\alpha$  (MIP1A), and tumor necrosis factor  $\alpha$  (TNF- $\alpha$ ) in plasma [1, 2]. It was caused by an unknown beta coronavirus, initially called as 2019-nCoV; later the unknown beta coronavirus was named SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2), which formed a clade within the subgenus Sarbecovirus [2, 3]. Apart from the well-known MERS-CoV (Middle East respiratory syndrome coronavirus) and SARS-CoV (severe acute respiratory syndrome coronavirus), the SARS-CoV-2 is the seventh member of the coronavirus family that infects humans [4]. The genome of SARS-CoV-2 has 89% and 82% nucleotide similarity with bat SARS-like-CoVZXC21 and of human SARS-CoV, respectively. The phylogenetic trees of spike, membrane, envelope, orf1a/b, and nucleoprotein from SARS-CoV-2 are clustered closely with those of the bat, civet, and human SARS-CoV. The external subdomain of the spike's receptor of SARS-CoV-2 has 40% amino acid similarity with other SARS-related CoV [5]. The entire orf3b of SARS-CoV-2 encodes a novel short protein. Moreover, new orf8 of SARS-CoV-2 probably encodes a secreted protein with an  $\alpha$ -helix, a  $\beta$ -sheet(s) having six strands [5]. The phylogenetic analysis of the complete viral genome (29,903 nucleotides) revealed that WH-Human-1 coronavirus (WHCV) or SARS-CoV-2 was most closely related (89.1% nucleotide similarity) to a group of SARS-like coronaviruses (genus Betacoronavirus, subgenus Sarbecovirus) that were previously sampled from bats in China and that have a history of genomic recombination [6]. A recent study confirmed that the SARS-CoV-2 uses the ACE2 cell entry receptor, similar to SARS-CoV [7].

Considering the outbreak and the high need for treatment strategies, we have carried out an in-silico approach to identify the best ligand against the SARS-CoV-2 envelope and nucleocapsid phosphoprotein.

# Material and methods

# Sequence retrieval and secondary structure prediction

The amino acid sequence of the Wuhan seafood market pneumonia virus envelope protein (Accession no QHD43418.1), nucleocapsid phosphoprotein (Accession no QHD43423.2), were retrieved from the NCBI database on 28<sup>th</sup> Jan 2020. Wuhan seafood, SARS (severe acute respiratory syndrome), MERS (Middle East respiratory syndrome), and porcine reproductive and respiratory syndrome and other sequences were retrieved from NCBI, and sequence alignment was done by MAFFT software [8] for both envelop and nucleocapsid phosphoprotein, and phylogeny was constructed using MEGA7 [9–11].

# Homology modelling

The sequences of envelope protein and nucleocapsid phosphoprotein were searched against the protein database using BLAST-P [12]. The proteins having PDB Id: 1ssk.1.A for nucleocapsid phosphoprotein [13] and 5x29.1.A for envelope protein [https://swissmodel.expasy.org/repository/uniprot/A3EX99] were selected for use as a template for 3D modelling of the envelope protein and nucleocapsid phosphoproteins of SARS-CoV-2. FASTA sequences were obtained for target and template selection.

# 3D structure prediction and validation

Homology modelling structure prediction was carried out using the Automated SWISS MODEL server [14]. The modelled PDB file was visualised using PyMOL and validated using PROCHECK [15]. 3D models were validated on the basis of Ramachandran plot [16] statistics using the RAMPAGE server as described earlier [17] and ERRAT2 [18]. From the generated models, the one with highest number of residues in the allowed region and minimum number of residues in the disallowed region were considered as a suitable model for envelope protein and nucleocapsid phosphoprotein of SARS-CoV-2 and then used for further analysis. The active site was predicted using the MOE (Molecular Operating Environment) tool site finder [19]. The two predicted models of 3D atomic coordinates of the receptor were used for computations to verify potential sites for ligand binding and docking.

# Preparation of ligand for docking analysis

Chemical compounds were taken from the National Centre for Biotechnology Information (NCBI) Pub-Chem database. All the ligands involved in our report were accumulated from the ones available in the literature [20–23], and others are listed in Table I. The ligands for envelop protein (1175, 2CBU, 2AAC, 1JR1) and nucleocapsid phosphoprotein (4UCE, 4UCC, 4UCD, 4UC8) were downloaded from a protein databank in Structure Data File (SDF) format and later converted to Protein Data Bank (PDB) coordinate files using Marvin space software, and ligands were saved in .mol format with the aim of opening these files in MOE software. Energy minimisation was done using

Protein	Ligand	Number of bonds	HbA	HbD	Log P	∆G [kcal/mol]	р <i>Кі</i> [µМ]
Envelope .	E1	5	5	4	-2.194	-7.1939	5.509
	E2	5	5	2	3.00*	-10.2567	7.713
	E3	6	4	5	-3.899	-7.9052	8.105
	E4	6	4	5	-3993	-6.7359	8.761
Nucleocapsid phosphoprotein - -	N1	4	5	2	1.733	-10.3805	7.067
	N2	2	5	2	2.901*	-12.2112	7.885
	N3	3	5	2	2.248	-9.3889	7.284
	N4	1	2	1	-1.411	-8.6312	5.725

 Table I. The properties of the ligands in the active site of envelope protein and nucleocapsid phosphoprotein of Wuhan coronavirus, 2019-nCoV

\*Significant druggable protein ligand; HbA – hydrogen bond acceptors, HbD – hydrogen bond donors, log P – The log octanol/water partition coefficient, pKi – estimated binding affinity, E1 –  $\beta$ -D fucose, E2 – mycophenolic acid, E3 – castanospermine, E4 – deoxynojirimycinls, N1 – M72: 1-[(4-fluorophenyl)methyl]pyrazole-3,5-dicarboxylic acid, N2 – M76: 1-[(2,4-dichlorophenyl)methyl]pyrazole-3,5-dicarboxylic acid, N2 – M76: 1-[(2,4-dichlorophenyl)methyl]pyrazole-3,5-dicarboxylic acid, N4 – P1: phenylalanine.

MOE tools to first protonate the structure by using default parameters pH 7 and temp 300°C. The selected ligand molecules were passed through a Lipinski filter.

# Molecular docking

For molecular docking the two modelled structures of selected antiviral molecules with envelope protein and nucleocapsid phosphoprotein were 3D protonated, and then docking was performed; we selected ligand (β-D-fucose; mycophenolic acid; castanospermine; deoxynojiri-1-[(4-fluorophenyl)methyl]pyrazole-3,5mycin; dicarboxylic acid; 1-[(2,4-dichlorophenyl)methyl] pyrazole-3,5-dicarboxylic acid; 1-[(2 chlorophenyl) methyl]pyrazole-3,5-dicarboxylic acid, and the PHENYLALANINE atom. Settings were selected in MOE software as rescoring1 at London dG and rescoring2 at GBVI/WSA dG, and the ligand interaction was performed with protein [24]. Four ligands were used for envelope protein, and another four ligands were used for nucleocapsid phosphoprotein. Energy minimisation was done for both ligands and proteins. Envelope protein before energy minimisation E: 5471.98, RMS: 14.93, and after energy minimisation E: 2433.49, RMS: G = 0.0709512, E: 2489.62, RMS G = 0.0700238, E: 2477.92, RMS: G = 0.0713067, and E: 2562.35, RMS: G = 0.124056 with  $\beta$ -D-fucose, mycophenolic acid, castanospermine, and deoxynojirimycinls ligands, respectively. For nucleocapsid phosphoprotein before energy minimisation: E: 2673.4, RMS: G = 17.3825, and after energy minimisation E:475.537, RMS G = 0.0875944, E:428.511, RMS G = 0.0805305, E: 372.844, RMS G = 0.0508421, and E: 390.26, RMS G = 0.0939766 with M72, M76, M81, and P1 ligands, respectively.

# Results

The amino acid sequences of envelope protein and nucleocapsid phosphoprotein were blasted against the PDB-BLAST database to identify an appropriate template for homology modelling. The protein having PDB Id: 1ssk.1.A (seq. identity 92.37, seq. similarity 0.61) and 5x29.1.A (seq. identity 91.38, seq. similarity 0.54) were selected as a template for 3D modelling of the envelope protein and nucleocapsid phosphoprotein. The SWISS MODEL server was used to predict the 3D structure of the envelope protein and nucleocapsid phosphoprotein. Models were built based on target-template alignment using ProMod3 in the SWISS MODEL server. The best models of envelope protein and nucleocapsid phosphoprotein were selected based on the best QMEAN score (0.01) and highest resolution 2.48Å, and were validated using the RAMPAGE sever.

The protein structure's stereochemical stability was calculated with the help of a Ramachandran plot. The Ramachandran plot explained the 3D structure of the envelope protein and nucleocapsid phosphoprotein, showing 84% and 90.4% amino acid residue of predicted structure are in the favoured region for the nucleocapsid phosphoprotein and envelope protein, respectively. Also, amino acid residues in the allowed region were 6.1% (nucleocapsid phosphoprotein) and 13.3% (envelope protein), and the remaining number of residues in the outlier region was 3.6% (nucleocapsid phosphoprotein; Figure 1 B) and 2.2% (envelope protein; Figure 1 B). The overall guality factors for nucleocapsid phosphoprotein and envelope protein of the predicted models at ERRAT2 were 94 and 87, respectively.



**Figure 1. A**, **B** – Ramachandran plot from RAMPAGE of Wuhan coronavirus, SARS-CoV-2 protein. **A** – Envelope protein. **B** – Nucleocapsid phosphoprotein. The phi ( $\phi$ ) values of amino acid residues are on the *x*-axis. The psi ( $\psi$ ) values are on the *y*-axis. **C**, **D** – 3D structure of envelope protein (**C**) and nucleocapsid phosphoprotein (**D**) after homology modelling

# Molecular docking

Envelope protein and nucleocapsid phosphoprotein of SARS-CoV-2 were prepared for molecular docking and were analysed by MOE software initially by 3D protonation, energy minimisation, and prediction of active site for the eight ligands by keeping the parameters at their defaults. Then the ligands (E1 to E4 and N1 to N4) were docked separately with the envelope protein and nucleocapsid phosphoprotein of SARS-CoV-2 (Figures 2, 3) using MOE software. The results from molecular docking suggested that the E2: mycophenolic acid (log P =3.00;  $\Delta G = -10.2567$  kcal/mol; pKi = 7.713  $\mu$ M) was the most potent druggable protein ligand of the SARS-CoV-2 envelope protein (Figures 2 A, B), while N2, 1-[(2,4-dichlorophenyl)methyl]pyrazole-3,5-dicarboxylic acid (Log P = 2.901;  $\Delta G = -12.2112$ 

kcal/mol;  $pKi = 7.885 \mu$ M) was the most potent druggable protein ligand of SARS-CoV-2 nucleo-capsid phosphoprotein protein (Table I, Figure 2).

#### Discussion

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a global pandemic health threat. SARS-CoV-2 was identified as a new strain of the Beta-CoVs genera, and is a member of the zoonotic origin coronavirus group. It causes coronavirus disease-2019 (COVID-19), which is the greatest concern in all the countries involved in the outbreak for health and economy reasons. SARS-CoV-2 is distinct from the severe acute respiratory syndrome virus [2, 3, 25–27]. However, the phylogenetic analysis of the envelope protein and nucleocapsid phosphoprotein revealed





**Figure 2.** Significant druggable protein ligand complex of envelope protein and nucleocapsid phosphoprotein of SARS-CoV-2. E2 ligand has five hydrogen bonds: one with Asn\_64, two with lys\_63, one with val\_49, and one with ILE\_46. N2 has two hydrogen bonds: one with Thr 49, and the other with Tyr112, in addition to one arene-arene interaction with Tyr 109

that these proteins are close to the nucleocapsid phosphoprotein of bat coronavirus and severe acute respiratory syndrome-related coronavirus (Figures 4-6). Hence, the study was designed to predict potent ligands against druggable envelope and nucleocapsid phosphoprotein of SARS-CoV-2. The 3D models of the envelope protein and nucleocapsid phosphoprotein of SARS-CoV-2 were predicted, validated, and used for docking studies. The docking studies help in the prediction of the preferred orientation of a ligand with the binding site on a protein and are used for conformation of various chemical compounds at the target site of the protein. The most potent identified compounds for envelope protein, mycophenolic acid and nucleocapsid phosphoprotein, 1-[(2,4-dichlorophenyl)methyl]pyrazole-3,5-dicarboxylic acid) with highest log octanol/water partition coefficient (Log P), high number of hydrogen bond donors and acceptors, lowest non-bonded interaction energy ( $\Delta G$ ) between the receptor and the ligand, and high binding affinity (pKi), indicate that they are the most potent compounds against the SARS-CoV-2 envelope and nucleocapsid phosphoprotein.

The coronavirus nucleocapsid phosphoprotein is a multifunctional structural protein; during virion assembly it interacts with the viral membrane and forms complexes with genomic RNA. The coronavirus nucleocapsid phosphoprotein plays an important role in coronavirus transcription and assembly as well as the coronavirus lifecycle [28–34]. The most potent identified compound, 1-[(2,4-dichlorophenyl)methyl]pyrazole-3,5-dicarboxylic acid], may inhibit any of its multifarious activities and functions during virion assembly; however, detailed studies are needed on the inhibitory effect of these compounds on the interaction of nucleocapsid phosphoprotein with the viral membrane, and formation of complexes with genomic RNA during SARS-CoV-2 transcription and virion assembly.

The coronavirus envelope protein plays a crucial role for the lifecycle of the virus. The small integral membrane protein, the coronavirus envelope protein, is important for the development of the disease in the host through viral assembly, to exit the host cell by viral budding, viral propagation, envelope formation by taking portions of the host cell membranes, and the release of infectious virus from the host cell [33-35]. Hence, the SARS-CoV-2 envelope protein was considered for the docking study to identify the most potent compound; the study revealed that mycophenolic acid may an appropriate druggable protein ligand of SARS-CoV-2 to inhibit the development of a COVID-19 by blocking the viral assembly. Complete wet lab analysis is needed to elucidate the impact of the mycophenolic acid on the virus' exit



Figure 3. 2D and 3D protein-ligand interaction of envelope protein and nucleocapsid phosphoprotein of SARS-CoV-2

State-of-the-art tools to identify druggable protein ligand of SARS-CoV-2



**Figure 4.** Phylogenetic analysis of the nucleocapsid phosphoprotein of SARS-CoV-2 by Maximum Likelihood method. "The evolutionary history was inferred by using the Maximum Likelihood method based on the JTT matrix-based model [11]. The bootstrap consensus tree inferred from 500 replicates [10] is taken to represent the evolutionary history of the taxa analysed [10]. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. Initial tree(s) for the heuristic search were obtained automatically by applying neighbour-joining and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. The analysis involved 78 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 43 positions in the final dataset. Evolutionary analyses were conducted in MEGA7 [9]." Nucleocapsid phosphoprotein sequence used for constructing the phylogenetic tree: MSDNGPQNQRNAPRITFGGPSDSTGSNQNGERSGARSKQRRPQGLPNNTASWFTALTQHGKEDLK-FPRGQGVPINTNSSPDDQIGYYRRATRRIRGGDGKMKDLSPRWYFYYLGTGPEAGLPYGANKDGIIWVATEGALNTPKDHIG-TRNPANNAAIVLQLPQGTTLPKGFYAEGSRGGSQASSRSSRSRNSTPGSSRGTSPARMAGNGGDAALALLLDRLNQLE-SKMSGKGQQQGQTVTKKSAAEASKKPRQKRTATKAYNVTQAFGRRGPEQTQGNFGDQELIRQGTDYKHWPQIAQFAPSAS-AFFGMSRIGMEVTPSGTWLTYTGAIKLDDKDPNFKDQVILLNKHIDAYKTFPPTEPKKDKKKKADETQALPQRQKKQQTVTLL-PAADLDDFSKQLQQSMSSADSTQA

#### S.A. Azeez, Z.G. Alhashim, W.M. Al Otaibi, H.S. Alsuwat, A.M. Ibrahim, N.B. Almandil, J.F. Borgio

<b>1</b> 1.2019-000V	MYSFYSEELELIVHSYLLFLAFYYFLLYELAILEALKLGAYGCHIYHYSLYKPSFYYYSKYKHLHSS-KY-POLL)
2. APO40681.1:1-76_E_protein_Severe_acute_respiratory_syndrome-related_coronavirus	MYSFYSEERS LIVHSVLLFLAFYYFLLV LAILEALBLCAYCONIVNVSLVKPSFYIYSAVKALBSSOI - POLLY
3. AGZ48809.1 1-76_envelope_protein_Bat_SARS-like_coronavirus_RsSHC014	11.1.2.1.1.2.1.2.1.1.1.1.1.1.1.1.1.1.1.
<ol> <li>4. QDF43816.1:1-76_small_envelope_protein_Coronavirus_BtRI-BetaCoV/9C2018</li> </ol>	WAREAREELETIAN AND A A A A A A A A A A A A A A A A A
5. NP_828854.1.1-76_protein_E_Severe_acuts_respiratory_syndrome-related_coronavirus	MYSFVSEETGTLIVHSVLLFLAFVVFLLVTLAILTALRLCAYCCHIVHVSLVXFTVYVYSRVKHLHSSESV-POLLV
6 ADE34767 1:1-76 small memorane protein Bat SARS coronavirus HKU3-7	MYBFYBEETERLIVMSVILFLAFVVFLLVTLAILTALRLCAVCCHIVMVSLVKPTVVVSRVKMLMSSERV-PDLLA
7. ATO98160.1.1-76 small envelope protein Bat SARS-like coronavirua	MYSFVBEETGTLIVNSVLLFLAFVVFLLVTLAILTALRUCAVCCHIVNVSLVKFTVVVVSRVKNLNSAEGV-POLLV
8 AAP51230.1.1-76 envelope protein E SARS coronavirus GD01	NYSFYSEETCTLIVNSVLLFLAFNVFLLVTLAILTALRLCAVCCN(VNVSLVKPTVVVSRVKNLNSEEQV-POLL)
9. AlA62312 1:1-76 small envelope protein BIRs-BetaCoV/HuB2013	NY NEVERSE AND
10 10050008 115.76 anualoga contain E 0100 concessionia B100.10	NUMBER OF THE OWNER
11 JATP3502 11 76 contract while CIDP concertion Evalua	
TT. ACCY 2023. 1. HYD_ENVEROPE_DIVENT_CARRO_COLORIBUTINES_EDIVIT	
12 YP_003656566.11+76_envelope_protein_83t_coronavrus_8048-31/8GR02008	MINTY VERETE LIVEN VELLEVELVELVELVELVELVELVELVELVELVELVELVE
13. Q9Q473.1:1-76_RecName:_Fuil_Envelope_smail_memorane_protein_Bat_CoV_279/2005	HYSFYSELETETLIVESVLLFFAFYYFLLVELAILEALSLCAYCCHIVEYSLVKFLYYYSEVKHLESEESY-POLLY
14. ADC35514.1.1-76_envelope_protein_SARS_coronavirus_HKU-39849	MYBFYSEETSELIVMSVLFFLAFVYFLLVELAILEALRUCAYCCMIVMVSLVKPEVYYYSRVKNLMSSEDV-POLLV
15. ADE34768.1:1-76_smail_membrane_protein_Bat_SARS_coronavirus_HKU3-8	市大型とへ営業を実施していた施設へしたというとなんとして入量しかして置いておいたなどのの第1人類へ気に入発を置い入入入業が入業を必要に加る手能の、このについた。
16. AKZ19089.1:1-76_envelope_protein_Bat_SARS-like_coronavirus_YNLF_34C	WAREA 新聞記事業ではAMERATOLETVEAALETAATVILLEVENDOVACOM(AMARATAAAMAAAMAAMAAMAAMAAMAAMAAAMAAAMAAA
17. AHX37560.1.1-76_envelops_protein_Rhinolophus_affinis_coronaulrus	WAREAREEREELCIAMEACTELTEREALACETTERECUTEDECERCEELAMEAREEREELAMEACTELEEREEREEREEREEREEREEREEREEREEREEREEREE
18. ACZ72198.1.1-76_envelope_protein_SARS_coronavirus_ExoN1	MYSFYBEET MELLYMSVLLFLAFVYFLLVELAILSAHRLCAYCCHIVMYSLVKFTVYYSRVKHLMSSENV-POLLI
19 AAR23255.1.1-76 envelope protein SARS coronavirus Sino3-11	MYSFFSEETER LIVHEVLLFLAFVVFLLVILAILIAALRUCAVCCNIVNVSLVKFIVVVVSRVKNLNISEGV-POLLV
20 ABD76313.1:1-76 envelope crotein Bat SARS CoV Rf1(2004	NY SEVERET AT LIVESVLETVATVYFLLVELAIL TALELCAYCCHIVEVALVKPTVYVSRVKNLDSSDCV. POLLA
21 AUA62260 111.76 email emaines crotein BID: BetsCov/U 2012	HVECKEETENII VEEVII CVACVVCI VII ATI MATEL CAVCCHIVEVEI VEETVVVEEVENI NEACHV.
22 JOC24167 1-176 Advalage protein Det cardening DeChasen/2011	
22. AGGT4T01.1.170_ENHOUPS_DOGIN_DAL_CONTINUES_OPERATION_011	
23. AGE/2243.1.1-YD_ENVERODE_DICHEN_SAPIS_COTONEVIDE_NAYD_EXCH1	MIRCORE AND LIVERY LOCAL TO LARD CONSTRUCT OF A CARTER CARE AND CARE
24. AVAZ302.1.1-76_\$mail_enverope_protein_BHR-BetaCov/SA2013	MTSPVSEETULIVESVLEPVAPVVPLLVELAILEALECATGCHIVEVSLVAPIVTVSKVAMMSSOCV-PULL
25. ABM92862.1:1-76_smail_envelop4_protein_SAR3_coronawrus_TJ01	MYEFYBEETER LIVHEYLLFLAFYYFLLVILAILEALHLCAYCCHIVHYBLYKPEYYYHBHYKNLHBBEBY POLLY
26. ACZ72257.1.1-76_envelope_protein_SARS_coronavirus_ExoN1	WYSFYSEETSELIYNSYLLFLAFYYFLLYELAILEALCACCCNIYNYSLYKPEYYYYSRYKALMSSEGY-POLLY
27. ATO98184.1 1-76_small_envelope_protein_Bat_SARS-like_coronavirus	WAREAREAREACTIANSAFELEVEAAAAEFEALAAEFEALAAEFEALAAEFEAAAEFEAAAEFEAAAEFEAAAFEAAEFEAAAFEAAAEFEAAAFEAAAEFEAAAFEAAA
28. AAS44718.1:1-75_small_envelope_E_protein_SARS_coronavirus_TW-GD1	内支援とな器器局単規型に1・構築ストのというとしてもんたらのためになった置くの構成のなどの構成器を含く、調査した。
29. AMA96030.1:1-75_envelope_protein_Bat_coronautrus	NIC型とA.温着単単態面で1.A.種類A.F.F.E.A.V.E.F.F.A.M.F.F.F.A.M.F.C.F.A.C.C.M.1.A.種A.M.F.A.A.M.K.A.A.M.K.M.F.M.S.A.K.A.M.F.M.S.A.M.F.M.F.M.F.M.F.M.F.M.F.M.F.M.F.M.F.M
30. AAR23247.1.1-75_envelope_protein_SARS_coronavirus_Sino1-11	MYRFVREET MILL HRSVLLFUAFVVFULVEUAILEALRUCAYCCYIVMVSUVKPEVYVXERVKHLMSSEEDV.PDLU
31, AAS48454, 1:1-76_small_envelope_protein_SARS_coronavirus_EJ01	UNBEVEET TUTVESTOFFLAFTVELLATLAILTALELCAVCCHIVHVSLVKETVVVSEVKHUNSEETV.
32 ACZ72168.1:1-75 envelope protein SARS poronavirus ExoN1	NYSFYSEETGTLIVNSVLLFLAFVXFLLVTLAILTALRLCAVCCNIVNVSLVKPTVVVVSRVKNLNSSEAV-POLL-
22 5/29 1/20-91 Chain 1 Emplore small membrane ordein SARS-related coroexing	CORRECT OF STREET, STR
24 Stills 41.40 Chain & Employee small memories and the SIDC related concertain	
Se znakejk PSO Change Enterope Enter menorane protein SARSHeater Contravios	
35. AD117913.1.1-77_Dulawe_envelope_protein_cana_cat_coronavirus	AT BEYOND THE FUT HOLY CLIVANA CLICK CONTYNING VALUE AND
36. YP_009072442.1:1-71_envelope_protein_Bat_Hp-betacoronawrus/zrveyang2013	MYSEVEDERE VIVEAVELUVEEVALAILEGUELGAIGONILBERVVEEREVVLEAUEFINKLEP
37. Q0Q4EB 1.1-75_Envelope_smail_membrane_protein_Bat_coronawrus_(BtCoVI133(2005)	ALPEVHEDIBTIIVMEFILEVVCAITLLVCLAVLEAIBLCVGCABBYHTLLEVPAFYIVMTBRNAYFKEDE . MRPP.
38. YP_001039958.1:1-75_smail_membrane_protein_Tylonycleris_bat_coronavirus_HKU4	NFSEAHEDIBEIIABELIERAACVIEFAACVIEFAACVIEVIERVIECABCVBEABEABEREVAACVIEBEBEBEAVACVIEFEBE
39 ANA66044 1 1-64_envelope_protein_Bat_coronavirus	■「ちんれ夏間」回転11人間をとうと重んんなとし載さんのです」「基本1名から人間でや客間人間面にたたんなをというが登録回答題。
40. VP_009513016.1:1-76_E_Betacoronavisus_Ennaceus/VMC/DEU/2012	NTSもA1回動で開催とうAMEととした美AヤビンA1CFA2EMサイトで基々C3FCCA型CC1単価AMEにてA2EサイルA1EM目前に図AAAAA3220000000000000000000000000000000
41. ASL68947.1:1-75_envelope_protein_Hypsupp_bat_coronavirus_HKU25	MLPFV@E@IB@FIVMFFIFEVACAITLLVCMAFLEATRLCM@CAI@VM#LLV@FAIVYMEBR@AVVKF@E-BKPP-
42 QCC20719 1:1-76_E_Hedgehog_coronavirus_1_Hedgehog_coronavirus_1	MLPFVHOOLOSFIVMFFIFTVACAVILLVCMAILTACRLCVOCLTGFMTLLVOPAIVVMMDHSVVVKYON SKPEL
43 ASL69958 1:1-59 envelope protein Hypsupp bat coronavirus HKU25	NUPPYNEDIGEFLYNFFIFEYAGAI ELLYGNAFLEATSLCHOGAIGYNELLYDPALYYY
44. AHY61342 1 1-75 small envelope protein BIVs-BetaCoW9C2013	NURFY DEGISSFIVEFFIFTYACALILLYCHAFLEATELCYDCALGUMALLYDPALYYYNTERSYYYKFDE-BKPP.
45 BB (35013 1:1-75 small envelope protein Bat coronavitus	NEW FYORE FURTHER FOR THE VALUE FOR THE COVERS VIEW AT LY THE FOR THE F
45. BBU35013.1:1-75_small_envelope_protein_Bat_coronavitus	MULTINEED HER FLYER FLYER VOOL TUUR VAN AF SEARCUS WEST HELE ALLY EN ALLY EN ALLY EN AND VESS AND VESS AND VESS
45. BBJ3013.11-75_small_envelope_protein_Bal_coronavitus 46. YP_009361862.11-75_small_envelope_protein_Bal_coronavitus 47. 00127697.400.000	ULPEVOCOLOGE VNFFIFTVACAVILLVCUAFLIATELCVOCITOVNILLVCPAVVUVNITARSVYVKFOE SKPP
45. BEJ35013.11-75_scmall_envelope_protein_Ball_contrautus 46. PF_003516562.11-75_scmall_envelope_protein_Ball_contrautus 47. QGA70697.11-60_envelope_protein_Efinaceus_hedgehog_contrautus_HRU31	MERTY BOOL BETIVEFFIFT VACAVILLVCHAFLTARELCVOOR MELLVOPAVVUUNTERBOVVVEDE BEFP
46. BBU3013.11-75_smbl_privelos_protein_BBU_contensitus 46. YP_D03015402.11-75_smbl_privelos_protein_BBU_contensitus 47. OQX70007.11-60_envelose_protein_Enizocus_Inedgenog_contensitus_HKU31	ULPFYEODLOFTYWFFFFFWACAVILLYUCAATLWARLCYGCAABFWELLYDFAVWWWEBGRWYWKFBE - MKPP- NLPFYHOGYWBFTYWFFFFFWACAVILYUCAATLWARLCYGCAABFWELLYDFAVWWWWEBGRWYWKFBE - MKPP-
46. BBU39013.11-72_annal_envaloge_protein_BBL_corenatives 46. YP_00351962.11-73_annal_envaloge_protein_BEL_corenatives 47. QGA70097.11-60_enveloge_protein_Enracevs_hedgehog_corenatives_HKU31 17.2058/n2/	ULE PENDED LE METEL VIE FELFEN VACAVILLU VOLAFIL AN RLEV OCH MENNELLU VOFAVYU VIE MENNEN VVV KER BERNELLU VOFAVU VIE MENNEN VVVV KER BERNELLU VOFAVU VIE MENNEN VVVV KER BERNELLU VOFAVU VIE MENNEN VVVV KER BERNELLU VIE MENNEN VVVV KER BERNELLU VIE MENNEN VVVV KER BERNELLU VIE MENNEN VVVVV KER BERNELLU VIE MENNEN VVVVV KER BERNELLU VIE MENNEN VVVVVVV KER BERNELLU VIE MENN
46. BBUJ3013.11-72_annal_invelope_protein_BBU_contensitius 46. VF_D0055162.11-72_annal_invelope_protein_BEU_contensitius 47. QGA70697.11-80_envelope_protein_Eninaceus_hedgehog_contaxinus_HKUD1   B 12016-00V 27.90097281972-1-419_audisecasusi stituationemin_Watan_seatod martet creamesa van	
46. BBU3013.11-72_annul_envides_protein_BSL_coreanius     46. YP_00051602.11-75_annul_envides_protein_BSL_coreanius     47. QGA70097.11-60_enveloce_protein_Enracess_hedgehog_coreanius_HKU31     1208+nCeV     120894724972-1-415_andercapscf_phosphorether_Titleam_testor_markt_preumosa_vess     120924972-1-415_andercapscf_phosphorether_Titleam_testor_markt_preumosa_vess     120924972-1-415_andercapscf_phosphorether_Titleam_testor_markt_phosphoretherapsc	ULE FONDEL MET I VEFFIFEVACAVILLUVULAFLEAARLUVUCA I UVEFALVULUVUFALMEN VVVVETEN VVV
46 BBU301311-72_mml_mwloce_protein_BL_corenations 47 QGA70007.11-80_mwloce_protein_BL_corenations 47 QGA70007.11-80_mwloce_protein_Enacess_hedgetog_corenations_HKU31 47 QGA70007.11-80_mwloce_protein_Enacess_hedgetog_corenations_HKU31 47 QGA70007.11-80_mwloce_protein_United_station_monet_corenations 47 QGA70007.11-80_mwloce_protein_United_station_monet_corenations 47 QGA70007.11-80_mwloce_protein_United_station_monet_corenations 47 QGA70007.11-80_mwloce_protein_United_station_monet_corenations 47 QGA70007.11-80_mwloce_protein_United_station_monet_corenations 47 QGA7007.11-80_mwloce_protein_United_station_monet_corenations 48 QGA7007.11-80_mwloce_protein_United_station_monet_corenation 48 QGA7007.11-80_mwloce_protein_Station_11-80_mwloce_prote	
46. BBU3013.11-72_anal_envalue_protein_BL_corenative 47. QGA70097.11-60_enveloe_protein_BL_corenative 47. QGA70097.11-60_enveloe_protein_Enracese_hedgehog_corenative_HKU31  12058-05W 12059.218772.1-41%_undercosed_enveloe_inter_UNAm_seabod_mansf_presentees_ves 12069.2005.11-41%_undercosed_protein_UNAm_seabod_mansf_presentees_ves 12069.2005.11-41%_undercosed_protein_BL_corenatives 48. PPD-0007248772.1-41%_undercosed_protein_BL_corenatives 49. PDD-0007248772.1-41%_undercosed_protein_Envalues 49. PDD-0007248772.1-41%_undercosed_protein_UNAm_seabod_mansf_presentees_ves 40. PDD-0007248772.1-41%_undercosed_protein_BL_corenatives 40. PDD-00072487724747247472474747474747474747474774747474	ULE FONDEL MET IVN FFIFTU VACAVILLUVGAFLENANCUVGA UVALUVGA UVALUVA UVALUVA UVALUVAUVA UVALUVAUVAUVAUVAUVAUVAUVAUVAUVAUVAUVAUVAUVA
H6 BDJ3013.11-72mail_envides_protein_BIL_coreaatus     H6 VF_D030192.11-73mail_envides_protein_BIL_coreaatus     H7 D047/2007.11-80envides_protein_Enacess_hedgehog_coreaatus_HKU31     T	
K BBJ3013.11-74     Suddecastd_provider_BBL_contravitus     K DF2005192.11-75_unal_envides_Environ_BBL_contravitus     K DF2005192.11-75_unal_envides_Environ_BBL_contravitus     K DF2005192.11-74     DecastDecastC_provider_BBL_contravitus     K DF2005218372.1-419_undecastd_pt_contravitus     K DF2005218372.1-419_undecastd_pt_contravitus     K DF2005218372.1-419_undecastd_pt_contravitus     K DF2005218372.1-419_undecastd_pt_contravitus     K DF2005218372.1-419_undecastd_pt_contravitus     K DF2005218.1-419_undecastd_pt_contravitus	
He BU301111-74_minil_enviole_profile_BL_contention     How The BU301111-74_minil_enviole_profile_BL_contention     How The BU301111-74_minil_enviole_profile_BL_contention     How The BU301111-74_minil_enviole_profile_BL_contention     How The BU30111-74_minil_enviole_profile_BL_contention     How The BU30111-74_minil_enviole_profile_BL_contention     Average 11-141_minil_enviole_profile_BL_SRESHIP     Average 11-141_minil_enviole_profile_BL_SRESHing_contention     Average 11-141_minil_enviole_profile_mBL_SRESHing_contention     Average 11-141_minil_enviole_profile_BL_SRESHing_contention     Average 11-141_minil_enviole_profile_ML_SRESHing_contention     Average 11-141_minil_enviole_profile_profile_ML_SRESHing_contention     Average 21-141_minil_enviole_profile_profile_ML_SRESHing_contention     Average 21-141_minil_enviole_profile_RL_SRESHing_contention     Average 21-141_mini	
H4         BBUJ3013.11-72_mml_mmloc_protein_BU_corenations           H4         FP_D005192.11-72_mml_mmloc_protein_BU_corenations           H7         CGA70097.11-80_mmloce_protein_Enizores_hedgehog_corenations_HKU31           H8         1.2018-nCw           T_2018-nCw         2.002228372.21-410_mudecegent_protein_Winter_Minter_set0.00_marks_proteins_enizor           L0         CO2028.11-410_mudecegent_protein_Winter_Minter_set0.00_marks_proteins_winter           L0         CO2028.11-410_mudecegent_protein_Winter_Minter_set0.00_marks_proteins_winter           L0         CO2028.11-410_mudecegent_protein_Winter_Minter_set0.00_marks_proteins_winter           L0         CO2028.11-410_mudecegent_protein_BU_corenations           L0         Co2028.11-410_mudecegent_protein_BU_corenations           L0         Co2028.11-410_mudecegent_protein_BU_corenations           L0         Co2028.11-410_mudecegent_protein_BU_corenations           L0         Co2028.11-410_mudecegent_protein_BU_set0.000000000           L0         Co1010.11-410_mudecegent_protein_BU_set0.00000000000000000000000000000000000	
H4:         BBUJ3013:11-72, unall_invalues_protein_BBU_contensitius           H4:         FP_D005102:11-75, unall_envolues_protein_BBU_contensitius           H7:         CGR/20097:11-80, envolues_protein_BBU_contensitius           H7:         CGR/20097:11-80, envolues_protein_BBU_contensitius           1:         2018-rdb           2:         2018-rdb           2: <td< td=""><td></td></td<>	
K6 B0J3013.11-72, anall_envides_protein_B0L_corenations     K6 PE_3005192.11-75, anall_envides_protein_B0L_corenations     K7 P026351922.11-75, anall_envides_protein_B0L_corenations     K7 P026724972.11-40_wordecoses(_notebother_NUmar_seator_corenation_HKU31     12016-C0V     120169-C0V     120169-C0	
H4         BBU301311-F2_mml_minutescopester_BBL_contrantics           H4         FP_D0051021-F2_mml_member_protein_BBL_contrantics           H7         D60700711-60_mmeloce_protein_Enacess_hedgehog_contrain/us_HKU31           12018-r0pV         12018-r0pV           12018-r0pV         12018-r0pV <td></td>	
K6 B030513.11-72, anall_envides_protein_B1_contentus     K7 P0205192.11-73, anall_envides_protein_B1_contentus     K7 P0205192.11-73, anall_envides_protein_B1_contentus     K7 P0205192.11-74     Devides P12051     12-2014-02	
K6 BBJ3013.11-72mml_envides_protein_Bit_correlations     K7 DBATCORT.11-00_memologe_protein_Bit_correlations     K7 DBATCORT.11-00_memologe_protein_Bit_correlations     K7 DBATCORT.11-00_memologe_protein_United_Bit_correlations     K7 DBATCORT.11-00_memologe_protein_United_Bit_correlations     K7 DBATCORT.11-00_memologe_protein_United_station_correlations     K7 DBATCORT.11-00_memologe.protein_Bit_correlation     K7 DBATCORT.11-01_memologe.protein_Bit_correlation     K7 DBATCORT.11-02_memologe.protein_Bit_correlation     K7 DBATCORT.11-02_memologe.protein_Bit_correlation     K7 DBATCORT.11-02_memologe.protein_Bit_correlation     K7 DBATCORT.11-02_memologe.protein_Bit_correlation_Bit_correlation     K7 DBATCORT.11-02_memologe.protein_Bit_correlation     K7 DBATCORT.11-02_memologe.protein_Bit_correlation     K7 DBATCORT.11-02_memologe.protein_Bit_correlation     K7 DBATCORT.11-02_memologe.protein_DeattCorrelation     K7 DBATCORT.11-02_memologe.protein_Bit_correlation     K7 DBATCORT.11-02_memologe.protein_DeattCorrelation     K7 DBATCORT.11-02_memologe.protein_Bit_correlation     K7 DBATCO	
K6 B03001111-72_mail_cmv/doc_protein_B0_corenatus     K6 VF_00051021-75_mail_cmv/doc_protein_B0_corenatus     K7 026470007.11-60_envelope_protein_Unitin_B0_corenatus     K7 026470007.11-60_envelope_protein_Unitin_seatod_mails_prevenous_veloa     120054400     12005400     12005400     12005400     12005400     12005400     12005400     12005400     12005400     12005400     12005400     12005400     1200540     1200540     1200540     1200540     1200540     1200540     1200540     1200540     1200540     120054     1200540     120054     120054     120054     120054     120054     120054     120054     120054     120054     12005     120054     12005     12005     12005     12005     12005     12005     12005     12005     12005     12005     12005     1200     12005     12005     1200     12005     1200     12005     1200     12005     1200     12005     1200     12005     1200     12005     1200     12005     1200     12005     1200     12005     1200     12005     1200     12005     1200	
K6 BBJ3013.11-72_mail_envides_protein_BBL_coreasius     K7 DBATCORT.11-00_mineloge_protein_BBL_coreasius     K7 DBATCORT.11-00_mineloge_protein_Bit_coreasius     K7 DBATCORT.11-00_mineloge_protein_film.BBL_coreasius     K7 DBATCORT.11-00_mineloge_protein_film.BBL_coreasius     Coreasius_HKU31     I_2016+c0W     I_2	
K6 B030013.11-72, anall_novidos_protein_B0_contravitus     K7 P02003102.11-75, anall_novidos_protein_B0_contravitus     K7 P020721072-11-80_novidosceptorite_NUMan_setDod_matrixt_pretmova_ves     T2008470097.11-80_novidosceptor_NUMan_setDod_matrixt_pretmova_ves     DoddE0000     T2009721072-11-80_novidosceptorite_NUMan_setDod_matrixt_pretmova_ves     DoddE00000     Contravitus_Proteins_B1_contents     APP200011-11-80_novidosceptorite_NUMan_setDod_matrixt_pretmova_ves     APP200011-11-80_novidosceptorite_NUMAN_setDod_matrixtan     APP200011-11-80_novidosceptorites_NUMAN_setDodomatrixtan     APP200011-11-80_novidosceptorites_NUMAN_SetDodomatrixt	
K6 BBJ3013.11-72_mail_envides_protein_BBL_coreasius     K7.0647007.11-80_membors_protein_BBL_coreasius     K7.0647007.11-80_membors_protein_Bits_coreasius     K7.0647007.11-80_membors_protein_Name_sealor_coreasius_HKU31     I_2019-e0W     III-2019-e0W     IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	
Ke Bouldon 1, 11-22, and Linvidos protein, Bill, contrastius     Ke Probability 1, 11-20, and Linvidos, protein, Bill, contrastius     Ke Probability 1, 11-20, and Linvidos, protein, Bill, contrastius     Ke Probability 1, 11-20, and Linvidos, protein, Bill, contrastius     Ke Probability 1, 11-20, and Linvidos, protein, Bill, contrastius     Ke Probability 1, 11-20, and Linvidos, protein, Bill, contrastius     Ke Probability 1, 11-20, and Linvidos, protein, Bill, contrastius     Ke Probability 1, 11-20, and Linvidos, protein, Bill, contrastius     Ke Probability 1, 11-20, and Contrastius, protein, Bill, contrastius     Ke Probability 1, 11-20, and Contrastius, protein, Bill, contrastius     Ke Probability 1, 11-20, and Contrastium 1, 11, 12, 12, 12, 12, 12, 12, 12, 12,	
K4 BBU301311-72_mml_movies.protein_BL_corenation     K4 PE_D0051311-72_mml_movies.protein_BL_corenation     K7 GGA70097.11-60_mevelope_protein_Mina_statuce_proteins_BL_corenations     K7 GGA70097.11-60_mevelope_protein_Mina_statuce_proteins_Mina_     K7 GGA70097.11-60_mevelope_protein_Mina_statuce_proteins_Mina_     K7 GGA70097.11-60_mevelope_protein_Mina_statuce_marks_proteinses_mina     G0403328.11-41_publiceopatis_Mina_statuce_marks_proteinses_mina     G0403328.11-41_publiceopatis_Mina_statuce_marks_proteinses_mina     G0403328.11-41_publiceopatis_Mina_statuce_marks_proteinses_mina     G0403328.11-41_publiceopatis_Mina_statuce_marks_proteinses_mina     G0403328.11-41_publiceopatis_Mina_statuce_marks_proteinses     AvaPable 11-41_publiceopatis_Mina_statuce_marks_proteinses     AvaPable 11-41_publiceopatis_Mina_statuce_marks     AvaPable 11-42_publiceopatis_Mina_statuce_marks_proteinses     Locopatis     Lo	
Ke Bouldons 11-12 and Landscenaria (Johan Sell Contrastics     Ke Provide State (Johan Sell) (Johan Sell) (Johan Sell) (Johan Sell) (Johan Sell)     Ke Johanni (Johan Sell)     Ke Johanni (Johan Sell) (Johan	
Ke BBUJDOTS 11-72mail_enviolog_contem_BBU_contention     Ke Department     K	
He Bouldons 11-12, and Linvidos, protein, Bill, contravitus           He VP, Dobasti Dill, 11-12, and Linvidos, protein, Bill, contravitus           He VP, Dobasti Dill, 11-12, and Linvidos, protein, Bill, contravitus           He VP, Dobasti Dill, 11-12, and Linvidos, protein, Bill, contravitus           He VP, Dobasti Dill, 11-12, and Linvidos, protein, Bill, contravitus           He VP, Dobasti Dill, 11-12, and Linvidos, protein, Bill, contravitus           Linvidos II. 11-12, and Contravitus, protein, Bill, 2014, and 10-12, an	
Ke Bouldons 11-12-2, and Linvidoes protein, Bill, concentration     Ke Decision 21-12-2, and Linvidoes protein, Bill, concentration     Ke Decision 21-12-2, and Linvidoes protein, Bill, concentration     Ke Decision 21-12-2, and Linvidoes protein, Bill, concentration     Londons 11-141, undercaped, protein, Bill, concentration     Ke Decision 21-12-2, and Linvidoes protein, Bill, State 21-20, and 20-20     Ke Decision 21-12-2, and Linvidoes protein, Bill, State 21-20, and 20-20     Ke Decision 21-14-12, undercaped, protein, Bill, State 21-20, and 20-20     Ke Decision 21-14-12, undercaped, protein, Bill, State 21-20, and 20-20     Ke Decision 21-14-22, undercaped, protein, Bill, State 21-20, and 20-20     Ke Decision 21-14-22, undercaped, protein, Bill, State 21-20, and 20-20     Ke Decision 21-14-22, undercaped, protein, Bill, State 21, and 21-20, and 20-20     Ke Decision 21-14-22, undercaped, protein, Bill, State 21, and 21-20, and 20-20     Ke Decision 21-14-22, undercaped, protein, Bill, State 21, and 21-20, and 20-20     Ke Decision 21-14-22, undercaped, protein, Bill, State 21, and 21-20, and 20-20     Ke Decision 21-14-22, undercaped, protein, Bill, State 21, and 21-20, and 20-20     Ke Decision 21-14-22, undercaped, protein, Bill, State 21, and 21-20, and 20-20     Ke Decision 21-14-22, undercaped, protein, Bill, State 21, and 21-20, and 20-20     Ke Decision 21-14-22, undercaped, protein, Bill, State 21, and 21-20, and 20-20     Ke Decision 21-14-22, undercaped, protein, Bill, State 20-20, and 20-20     Ke Decision 21-14-22, undercaped, protein, Bill, and 21-20, and 20-20     Ke Decision 21-20, and 22-20, an	
Ki BBJ3013.11-72 , mall_envides_protein_BBL_commandus     Ki VP_D035192.11-73 , mall_envides_protein_BBL_commandus     Ki VP_D045192.11-612, mall_envides_protein_BBL_commandus     Ki VP_D045192.11-612, mall_envides_protein_BBL_commandus     ArD D04511.11-812, mall_envides_protein_BBL_SOFTMAN_SOFTTMAN_SOF	
Ke Bouldon 1, 11-72, and Linvidoes protein, Bill, contravitus     Ke December 2, and Linvidoes protein, Bill, and Linvidoes protein, Billonni, Fail, Jubildees protei	
Ke Bouldons 11-12 and Linvidoes protein Bill, concentration     Ke Production Strategy and Linvidoes protein Bill, concentration     Ke Production     Second Strategy and Linvidoes and Linvidoe	
Ke BBJ3013.11-72, Jindle (Inveloce profein, Bill, containing Bill, Co	
Ke Bouldons 11-12 and Linvidoes protein Bill, concentration     Ke Production Statistics and Concentration Statistics     Ke Concentration Statistics     Ke Concentration	
46 B00301311-72_minit_invalues_protein_B0_sets_analysis 47 OG47200711-40_meteope_protein_B0_sets_analysis 47 OG47200711-40_meteope_protein_Minit_B1_contentions 47 OG47200711-40_meteope_protein_Minit_setSoc_marks_preventions_uses 50 P000224872-1-415_undecoped_protein_Winit_setSoc_marks_preventions_uses 50 P000224872-1-415_undecoped_protein_Winit_setSoc_marks_preventions_uses 50 P000224872-1-415_undecoped_protein_Winit_setSoc_marks_preventions_uses 50 P000224872-1-415_undecoped_protein_Winit_SetSoc_marks_preventions_uses 50 P000224872-1-415_undecoped_protein_Winit_SetSoc_marks_preventions_uses 50 P000224872-1-415_undecoped_protein_Winit_SetSoc_marks_preventions_uses 50 P000224872-1-415_undecoped_protein_Winit_SetSoc_marks_preventions_uses 50 P000224872-1-415_undecoped_protein_B1_SetSot_Set_containes 50 P000224872-1-412_undecoped_protein_B1_SetSot_Set_containes 10 P0002784511-42_undecoped_protein_B1_SetSot_Set_containes 11 P000248811-42_undecoped_protein_B1_SetSot_Set_containes 11 P000248811-42_undecoped_protein_B1_SetSot_Set_containes 11 P000248811-42_undecoped_protein_B1_SetSot_Set_containes 11 P000248211-42_undecoped_protein_B1_SetSot_Set_containes 11 P000248211-42_undecoped_protein_B1_SetSot_Set_containes 11 P000248211-42_undecoped_protein_B1_SetSot_Set_containes 11 P000248211-42_undecoped_protein_B1_SetSot_Set_containes 11 P000248211-42_undecoped_protein_B1_SetSot_Set_containes 11 P000248211-42_undecoped_protein_B1_SetSot_Set_containes 11 P000248211-42_undecoped_protein_B1_SetSot_Set_containes 11 P000248211-42_undecoped_protein_B1_SetSot_Set_containes 11 P000248211-42_undecoped_protein_Set_containes 11 P000248211-42_undecoped_protein_SetS_containes 12 P000248211-42_undecoped_protein_SetS_containes 12 P000248211-42_undecoped_protein_SetS_containes 12 P000248211-42_undecoped_protein_SetS_containes 12 P000248211-42_undecoped_protein_SetS_containes 12 P000248211-42_undecoped_protein_SetS_containes 12 P000248211-42_undecoped_protein_SetS_containes 12 P000248211-42_undecoped_protein_SetS_containes 12 P000248211-	
Ke Bouldons 11-12 and Linvidos protein Bil Contrastics     Ke Decision 21-12 and Linvidos protein Bil Contrastics     Ke Decision 21-13 protein protein Bil Contrastics     Ke Decision 21-13 protein protein Bil State Protein Bil Contrastics     Ke Decision 21-13 protein protein Bill Bil State Protein Bill Contrastics     Ke Decision 21-13 protein protein Bill Bill Bill Bill Bill Bill Bill Bi	
Ke Bouldon 1, 11-72, and Linvaloes protein, BitL, constantion     Ke Decision 1, 11-72, and Linvaloes protein, BitL, constantion     Ke Decision 1, 11-72, and Linvaloes protein, BitL, constantion     Ke Decision 1, 11-72, and Linvaloes protein, BitL, constantion     Ke Decision 1, 11-72, and Linvaloes protein, BitL, constantion     Ke Decision 1, 11-72, and Linvaloes protein, BitL, constantion     Ke Decision 1, 11-72, and Linvaloes protein, BitL, constantion     Ke Decision 1, 11-72, and Linvaloes protein, BitL, constantion     Ke Decision 1, 11-72, and Linvaloes protein, BitL, constantion     Ke Decision 1, 11-72, and Linvaloes protein, BitL, constantion     Ke Decision 1, 11-72, and Linvaloes protein, BitL, Steffer, Constantion     Ke Decision 1, 11-72, and Linvaloes protein, BitL, Steffer, Constantion     Ke Decision 1, 11-72, and Linvaloes protein, BitL, Steffer, Constantion     Ke Decision 1, 11-72, and Linvaloes protein, BitL, Steffer, Constantion     Ke Decision 1, 11-72, and Linvaloes protein, BitL, Steffer, Constantion     Ke Decision 1, 11-72, and Linvaloes protein, BitL, Steffer, Constantion     Ke Decision 1, 11-72, and Linvaloes protein, BitL, Steffer, Constantion     Ke Decision 1, 11-72, and Linvaloes protein, BitL, Steffer, Constantion     Ke Decision 1, 11-72, and Linvaloes protein, BitL, Steffer, Constantion     Ke Decision 1, 11-72, and Linvalow protein, BitL, Steffer, Constantion     Ke Decision 1, 11-72, and Linvalow protein, BitL, Steffer, Constantion     Ke Decision 1, 11-72, and Linvalow protein, BitL, Steffer, Constantion     Ke Decision 1, 11-72, and Linvalow protein, BitL, Steffer, Constantion     Ke Decision 1, 11-72, and Linvalow protein, BitL, Steffer, Constantion     Ke Decision 1, 11-72, and Linvalow protein, BitL, Steffer, Constantion     Ke Decision 1, 11-72, and Linvalow protein, BitL, Steffer, Mitch Linvalow protein, BitL, 11-72, and Linvalow protein, BitL, 11-72, an	
48         PED 301311-72         Invalues protein BL contraints           47         PED 301321-72         Invalues protein BL contraints           47         CRATCORT         Invalues protein BL contraints           48         CRATCORT         Invalues protein BL contraints	
Ke Bouldons 11-12 under control (Landon Section (Landon S	
48 B03001111-72_minit_invalues_protein_B1_Streamstr	
<ul> <li>Re Bullion 11: 11: 2018 - 2019</li> <li>Re Bullion 11: 11: 2018 - 2019</li> <li>Re Bullion 11: 2018 - 2018 - 2018</li> <li>Re Bullion 11: 2018 - 2018 - 2018</li> <li>Re Bullion 11: 2018 - 2018</li></ul>	
RE BBJJBJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJ	
<ul> <li>Re Bullion 11: 11: 21: 2018 (2019)</li> <li>Re Bullion 11: 21: 2018 (2019)</li> <li>Re Bullion 11: 2018 (</li></ul>	
46 BBJ3013 11-72	
<ul> <li>Re Bullion 11: 11: 2, and an analysis and an an</li></ul>	
RE BBJJBJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJ	
48 B03001311-12. India. (Invideo: protein, Bit., constants) 47 D08A0007 1 2008-000 2 2008-000 2 2008-000 2 2008-010-000 2 2008-000 2 2008-010-000 2 2008-	
46. VP 2003111-17-2, and _ involve_ protein_Bit_correlations 47. DGATCORF. 11-40, anvelope_protein_Bit_correlations 48. VP 200724272-21-410, anvelope_protein_Bit_correlations 48. VP 200724272-21-410, anvelope_protein_Bit_correlations 50. ONCOLOR 11-410, anvelope_protein_Bit_correlations 50. ONCOLOR 11-410, anvelope_protein_Bit_correlation 50. ONCOLOR 11-420, anvelope_Bit_correlation 50. ONCOLOR 11-420, anvelope_Bit_cor	
RE BBJJBJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJ	

Figure 5. Representative of the multiple protein sequence alignment of envelope protein (A) and nucleocapsid phosphoprotein (B) of Wuhan novel coronavirus, SARS-CoV-2. Envelope protein and nucleocapsid phosphoprotein sequence used for the sequence alignment are available in Figures 4 and 6, respectively

from the host cell by viral budding, the effect on blocking the envelope formation by taking portions of the host cell membranes, as well as its controlling power on release of infectious virus from the host cell.

There is no defined curative treatment for COVID-19 or any approved vaccines against SARS-CoV-2 infection. The WHO recommendation for the management of MERS-CoV is being in practice: initiation of oxygen therapy to keep the oxygen saturation above 90%, with conservative fluid management in the absence of shock, and an empiric antimicrobial regimen that includes antibiotics and a neuraminidase inhibitor for treatment of influenza. All of those supportive treatments are for the prevention of acute respiratory distress syndrome and for the prevention septic shock [2, 3, 36]. Hence, drug development against SARS-CoV-2 is considered urgent in order to fight COVID-19. The present in-silico approach identifies one potent ligand against the envelope protein and one potent ligand against nucleocapsid phosphoprotein of SARS-CoV-2. A combination of these two ligands might be the best option to

State-of-the-art tools to identify druggable protein ligand of SARS-CoV-2



**Figure 6.** Phylogenetic analysis of the envelope protein of Wuhan coronavirus, SARS-CoV-2 by Maximum Likelihood method. "The evolutionary history was inferred by using the Maximum Likelihood method based on the JTT matrix-based model [11]. The bootstrap consensus tree inferred from 500 replicates [10] is taken to represent the evolutionary history of the taxa analysed [10]. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. Initial tree(s) for the heuristic search were obtained automatically by applying neighbour-joining and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. The analysis involved 78 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 43 positions in the final dataset. Evolutionary analyses were conducted in MEGA7 [9]." Envelope protein sequence used for constructing the phylogenetic tree: MYSFVSEETGTLIVNSVLLFLAFVVFLLVTLAILTALRLCAYCCNIVNVSLVKPSFYVYSRVKNLNSSRVPDLLV

consider for further detailed studies in wet laboratories to develop a drug for treating patients infected with SARS-CoV-2.

#### Acknowledgments

The authors thank the Dean of the Institute for Research and Medical Consultations (IRMC), Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia for her continuous support and encouragement. The authors thank Dr. Balu Kamaraj for his valuable support, and Mr. Ranilo. M. Tumbaga, and Mr. Horace T. Pacifico for their support and assistance.

# **Conflict of interest**

The authors declare no conflict of interest.

#### References

- 1. Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 2020; 395: 497-506.
- 2. WHO, 2020a. Naming the coronavirus disease (COVID-19) and the virus that causes it, Accessed 23 Mar 2020. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-thecoronavirus-disease-(covid-2019)-and-the-virus-thatcauses-it
- Cárdenas-Conejo Y, Liñan-Rico A, García-Rodríguez DA, Centeno-Leija S, Serrano-Posada H. An exclusive 42 amino acid signature in pp1ab protein provides insights into the evolutive history of the 2019 novel human-pathogenic coronavirus (SARS-CoV2). J Med Virol 2020; doi: 10.1002/jmv.25758.
- 4. Zhu N, Zhang D, Wang W, et al. A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med 2020; 382: 727-33.
- 5. Chan JF, Kok KH, Zhu Z, et al. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. Emerging Microbes Infections 2020; 9: 221-36.
- Wu F. Zhao S, Yu B, et al. Complete genome characterisation of a novel coronavirus associated with severe human respiratory disease in Wuhan, China. bioRxiv 2020. doi: https://doi.org/10.1101/2020.01.24.919183.
- Zhou P, Yang X L, Wang X G, et al. Discovery of a novel coronavirus associated with the recent pneumonia outbreak in humans and its potential bat origin. bioRxiv 2020. doi: https://doi.org/10.1101/2020.01.22.914952.
- Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 2013; 30: 772-80.
- 9. Kumar S, Stecher G, Tamura K, MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 2016; 33: 1870-4.
- Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 1985; 39: 783-91.
- Jones DT, Taylor WR, Thornton JM. The rapid generation of mutation data matrices from protein sequences. Bioinformatics 1992; 8: 275-82.
- 12. Altschul SF, Madden TL, Schäffer AA. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997; 25: 3389-402.

- 13. Huang Q, Yu L, Petros AM, et al. Structure of the N-terminal RNA-binding domain of the SARS CoV nucleocapsid protein. Biochemistry 2004; 43: 6059-63.
- 14. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modeling. Electrophoresis 1997; 18: 2714-23.
- Laskowski RA, MacArthur MW, Moss DS, et al. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 1993; 26: 283-91.
- 16. Ramachandran GN, Ramakrishnan C, Sasisekharan V. Stereochemistry of polypeptide chain configurations. J Mol Biol 1963; 7: 95-9.
- 17. Abdulazeez S, Sultana S, Almandi, NB, et al. The rs61742690 (S783N) single nucleotide polymorphism is a suitable target for disrupting BCL11A-mediated foetal-to-adult globin switching. PLoS One 2019; 14: e0212492.
- 18. Colovos C, Yeates TO. Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci 1993; 2: 1511-9.
- 19. Ouizougun-Oubari M, Pereira N, Tarus B, et al. A druggable pocket at the nucleocapsid/phosphoprotein interaction site of human respiratory syncytial virus. J Virol 2015; 89: 11129-43.
- 20. Galloux M, Tarus B, Blazevic I, et al. Characterization of a viral phosphoprotein binding site on the surface of the respiratory syncytial nucleoprotein. J Virol 2012; 86: 8375-87.
- 21. Tran TL, Castagne N, Bhella D, et al. The nine C-terminal amino acids of the respiratory syncytial virus protein P are necessary and sufficient for binding to ribonucleoprotein complexes in which six ribonucleotides are contacted per N protein protomer. J Gen Virol 2007; 88: 196-206.
- 22. Botta L, Rivara M, Zuliani V, et al. Drug repurposing approaches to fight Dengue virus infection and related diseases. Front Biosci Landmark 2018; 23: 997-1019.
- Benvenuto D, Giovannetti M, Ciccozzi A, et al. The 2019 new coronavirus epidemic: evidence for virus evolution. J Med Virol 2020. https://doi.org/10.1002/jmv.25688.
- 24. Chemical Computing Group ULC .Molecular Operating Environment (MOE) 09. 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2017.
- 25. Hui DS, I Azhar E, Madani TA, et al. The continuing 2019nCoV epidemic threat of novel coronaviruses to global health – the latest 2019 novel coronavirus outbreak in Wuhan, China. Int J Infect Dis 2020; 91: 264-6.
- Wang C, Horby PW, Hayden FG, et al. A novel coronavirus outbreak of global health concern. Lancet 2020; 395: 470-3.
- Surjit M, Kumar R, Mishra RN, et al. The severe acute respiratory syndrome coronavirus nucleocapsid protein is phosphorylated and localizes in the cytoplasm by 14-3-3-mediated translocation. J Virol 2005; 79: 11476-86.
- 28. McBride R, Van Zyl M, Fielding BC. The coronavirus nucleocapsid is a multifunctional protein. Viruses 2014; 6: 2991-3018.
- 29. Chang CK, Hou MH, Chang CF, et al. The SARS coronavirus nucleocapsid protein – forms and functions. Antiviral Res 2014; 103: 39-50.
- 30. Cong Y, Ulasli M, Schepers H, Mauthe M. Nucleocapsid protein recruitment to replication-transcription complexes plays a crucial role in coronaviral life cycle. J Virol 2020; 94: e01925-19.
- Sungsuwan S, Jongkaewwattana A, Jaru-Ampornpan P. Nucleocapsid proteins from other swine enteric coronaviruses differentially modulate PEDV replication. Virology 2020; 540: 45-56.

- Boscarino JA, Logan HL, Lacny JJ, et al. Envelope protein palmitoylations are crucial for murine coronavirus assembly. J Virol 2008; 82: 2989-99.
- 33. Stodola JK, Dubois G, Le Coupanec A et al. The OC43 human coronavirus envelope protein is critical for infectious virus production and propagation in neuronal cells and is a determinant of neurovirulence and CNS pathology. Virology 2018; 515: 134-49.
- 34. Schoeman D, Fielding BC. Coronavirus envelope protein: current knowledge. Virol J 2019; 16: 69.
- 35. Westerbeck JW, Machamer CE. The infectious bronchitis coronavirus envelope protein alters Golgi pH to protect the spike protein and promote the release of infectious virus. J Virol 2019; 93: e00015-19.
- 36. World Health OrganizationWHO, 2019. Clinical management of severe acute respiratory infection when Middle East respiratory syndrome coronavirus (MERS-CoV) infection is suspected: interim guidance (No. WHO/MERS/ Clinical/15.1 Revision 1). World Health Organization.