

Preprocessing of the candidate antiviral drugs against COVID-19 in models of SARS cov2 targets

Hussein Abdulkadhim A. Prof. of Pharmacology and Therapeutics. 1

Najah R Hadi Prof. of Pharmacology and Therapeutics2

Mohammad Abdulhusein. BPh. Pharmacist.3

Suhad Traji Zamil. BPh, Msc Pharmacy.4

Suaad Traji Zamil. BPh, Msc Pharmacy.5

- 1.University of Kufa/ College of Medicine. Husseina.ahussein@uokufa.edu.iq
2. University of Kufa/ College of Medicine. drnajahhadi@yahoo.com
- 3.University of Kufa/College of Pharmacy Mohammad_abd2007@hotmail.com
- 4.University of Kufa/College of Pharmacy. Suhadt.alakeli@uokufa.edu.iq
5. University of Kufa/College of Pharmacy. Suaadt.alakeli@uokufa.edu.iq

Abstract

Although many viral infections are self-limiting, other are real health challenges like COVID-19 since many viruses possess just few druggable targets to be treated with small drug molecules.

Coronavirus genome encodes for upto 17 main proteins. Orf1ab encodes for polyprotein. COVID-19 structural proteins are the spike S, membrane M, envelope E and the nucleocapsid N protein while other are non-structural proteins designated as NSP1-13 for non-structural proteins. Among NSP the most important coronavirus targets for developing antiviral drugs are the papain-like protease, PDB ID: 6m03 and RNA polymerase NSP12, PDB ID: 6nur.

NCBI, NIH Genbank, Uniprot, PDB, DrugBank, ChemSpider databases and bioinformatics editor softwares like ICM Molsoft pro and SwissDock were used in addition to the in vitro lab model of viral protease were integrated to retrieve and analyze coronavirus targets and to select the candidate ligands in an attempt to evaluate the inhibitory efficacy of different experimental and approved drugs which were further optimized and searched for the highly similar approved drug. This step aims to adopt drug repurposing to speed the development of antiviral drugs and recommend rational in vivo and clinical studies.

After COVID-19 targets had been analyzed the drugs that shared > 70% similarity to the binding sites of those targets were reversin, pentagastrin, remdesivir, norfloxacin and nitazoxanide against COVID-19 papain-like protease whereas benzylglutathione, lopinavir and hydroxymethylglutathione against RNA polymerase.

The anti-resistance reversin showed the highest inhibitory efficacy against COVID-19 papain-like protease as indicated by the ligand-protease binding energy with Molsoft pro analysis. The calculated inhibitory binding was -137.30 kJ/mol $z > 1.9$ as compared with the tetrazapentadecanoate -129.57 kJ/mol $z = 4.0$, whereas remdesivir, pentagastrin, nitazoxanide and norfloxacin had a moderate antiprotease activity (> -100 kJ/mol). Norfloxacin shows results showed a slight consistency between in vitro and in silico models.

Although benzylglutathione is an experimental compound, however it had the highest RNA polymerase inhibiting efficacy with -129 kJ/mol binding energy which is even higher than lopinavir and Favonavir.

From the overall results, reversin, oligopeptides, quinolones and antiviral drugs may widen the treatment options for COVID-19 if further evaluated in clinical studies.

Keywords: *COVID-19, NSP, RDRP, papain-like protease, replicase and antiviral drugs*

Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS cov2 or COVID-19) is a highly spreading viral infection caused by a novel type of Coronaviridae family called SARS nCov.(1) Similar to other viral infections, Covid-19 has unique virulence characteristics in that viruses possess just a few druggable target proteins (2).like hemagglutinin, neuraminidase, polymerase, proteases, envelope and membrane proteins in addition to few polyproteins associated with their genome (3) Viruses are semi-dormant units that are interactive with drugs only in their cycles of proliferation. (4) These factors diminish the ability to treat viral infections. However in concern to the COVID-19, the viral proteome is composed of 17 main proteins: four are structural proteins, envelope E , membrane glycoprotein M, surface glycoprotein spikes S, and nucleocapsid protein N for + RNA binding whereas hemagglutinin and other non-structural proteins NSP1-13 are encoded with different orf genes (PDB, genbank, NCBI ID: 6lu7, 6lvn, 6lxt, 5r7y,5r7z,5r80,5r81,5r82,5r83,5r84) (5)(6)(7). The following set of COVID-19 proteome contents is a reasonable pharmacological target for developing new or repurposed antiviral drugs. The orf1ab encoded polyprotein: COVID-19 replicase is a polyprotein that has multiple activities like -RNA transcription, RNA template, mRNAs and virion RNA. This polyprotein has also multiple proteinases activity to cleave this polyprotein into functional products. The non-structural protein 1 NSP1 (host translation inhibitor) blocks 40S of human ribosome so that it arrests host transcription and mediate mRNA lysis and spare COVID-19 mRNAs. Non-structural protein 2: through cell PHB and PHB2, NSP2 maintains cells mitochondria viable. Papain-like proteinase PLP: lyses the N-terminus of the replicase polyprotein and activates Lys-48 and Lys-63 linked polyubiquitin chains in cells. PLP mediates viral membrane assembly. In addition, PLP inhibits INF production and NFkB. Non-structural protein 4 NSP4: mediates viral membrane assembly. Proteinase 3CL-PRO: activates lysis C-terminus of replicase polyprotein. Non-structural protein 6 NSP6: Mediates endocytosis and prevents lysosomal fusion. Non-structural protein 7 NSP7: with NSP8 it acts as a primase. Non-structural protein 8 NSP8: synthesizes longer products than oligonucleotide primers. Non-structural protein 9 NSP9: by ssRNA-binding protein it mediates viral replication. Non-structural protein 10 NSP10: mediates COVID-19 transcription by stimulating both nsp14 (exoribonuclease) and nsp16 (O-methyltransferase) for methylating viral mRNAs cap. RNA-directed RNA polymerase NSP12 RDRP: for replication and transcription of the viral RNA genome. Viral Helicase: a Mg-dependent and Zn-binding

protein that unwinds RNA and DNA. Guanine-N7 methyltransferase: has exoribonuclease activity (on both ssRNA and dsRNA) and a N7-guanine methyltransferase activity. Uridylate-specific endoribonuclease: Mn-dependent, uridyl enzyme. 2'-O-methyltransferase: for viral mRNA cap methylation at 2'-O-ribose site. (8) (9) in addition to the non-protein target against the viral membrane like terpenoids.(10)

Another host factors that were investigated as reasonable targets to treat COVID-19 include angiotensin converting enzymes 2 ACE2 (11)(12)(13) furins and passive and active immunostimulants like INFs and ILs(14)(15).

Many studies analyze viral protein targets virtually for docking sites (16)(17) then the predicted compound with high energy of affinity to target site is optimized for predicting the 3D conformational isomers and analogues.(18) Such lead experimental or drug compounds are then rearranged in a rank according to their target binding energy in kJ/mol with classifying types of ionic, vW, and hydrogen bonds by which they interact with target residues. (19) (20).

Many compounds and approved drugs like antiviral drugs are potential resources to be repurposed against different health challenges like coronavirus infections.

Of these potential drugs are the quinolones, small peptide molecules like reversin and pentagastrin, benzylglutathione and antiviral drugs that inhibit viral replication like lopinavir, favinavir in addition to the protease inhibitors like ritanovir, remdesivir and ribavirin while others are antiparasitic nitro-azole like nitazoxanide

In vitro viral model is a critical technical step for testing antiviral compounds (21) since that direct access of the researchers into the active viruses are forbidden according to the international regulation policy and the strict requisites for level 4 laboratories owing to their health risks (22) so that the artificial viral models including killed virion or its partial components are critical in conducting such researches. (23). In vitro testing is followed to determine pharmacokinetics and target binding dynamics.

Although viruses have just limited number of proteins, however each protein like COVID-19 main protease and NSP12 has in average 3 docking sites for surveying many test compounds. (24) This means that developing a rapid medical treatment against the highly transmittable viral epidemics like COVID-19 is a possible approach.

This study aimed to analyze all COVID-19 proteins and evaluate the highly predictive ligands for further optimization and to study ligand similar to the approved drugs to be re-evaluated and arranged for further clinical studies.

Models, Materials and Methods

This study had been accomplished in serial stages and originally designed to include both computer-based and laboratory assessment of candidate list of drugs against COVID-19.

1. The database retrieval and review study for COVID-19, SARS cov and Human coronavirus Hcov OC43, NL63 whole genome, proteome, envelope and membrane analysis in addition to reviewing relevant host factors like furin and ACE2. NCBI, NIH, Uniprot, GenBank, Proteinpedia, PDB, NeXtProt, Genomix, GeneCard and Viral genome were the used database to specify the target genes, proteins in addition to analyze and optimize the control targets with BLAST analysis to determine synonymous and diverse segments.

The whole genome FASTA format had been downloaded using different bioinformatics software NCBI (appendix IV). Ugene software was used to graph and conduct blast analysis between each corresponding proteins related to the SARS cov, SARS cov2 and Hcov, S, M, E, N, Orf1ab, orf3, orf6, orf7, orf8, orf9, orf10, HA and other NSP proteins that were deposited on softwares were assessed by multiple BLASTwith NCBI to determine the exact synonym and structural variations and to be further confirmed by pdb superimposition with ICM. Viral protein complexes like S-ACE2 and N-RNA were also identified for their interacting sites.

Host relevant proteins furin and ACE2 had been downloaded and assessed for their role in the pathogenesis of steps of viral infection. Those steps were conducted using pdb and ICM molsoft.

2. COVID-19 protein analysis was done using PDB, Uniprot, ICM molsoft pro, NCBI, NIH, SwissDock, BioXLab and Ugene software for determining viral protein 2D, 3D structures, similarity, biological roles , pathological effects, ligands, binding residues, surface map, protein health, strain points and target sites.

3. Ligand prediction. The target proteins were preprocessed for determination of the fullfit ligand with highest binding energy, in addition to determining of the ligand trajectory and types of force field energy at the binding site.

4. Ligand optimization.

Using ICM, molsoft pro the ligand is optimized to best fit binding site after selecting a highly effective pocket among the determined table of protein binding sites and the final structure of the inhibitor is then edited and stored in several molecule formats like sdf, mol, mol2 for the next steps.

3. RCSB pdb, PDBe and PDBj were the main database softwares used to obtain pdb format of the viral and host proteins according to the following IDs: 6lu7, 6m3m, 6m03, 6lvn, 6lxt, 5r7y, 5r7z, 5r80, 5r81, 5r82, 5r83, 5gwy, 5r84.

Preprocessing of COVID-19 papain-like protease

With Molsoft pro, the target protein is loaded and converted to PDB format with water deletion. The protein is analyzed for health, surface electropotential and reviewing all domains and cofactors and then a project of table drug surveying is conducted by creating the project directory. Receptor mapping and surface map. The test drugs table is loaded too for later trajectory. Adjusting the cut-off distances and selecting the candidate drugs table.

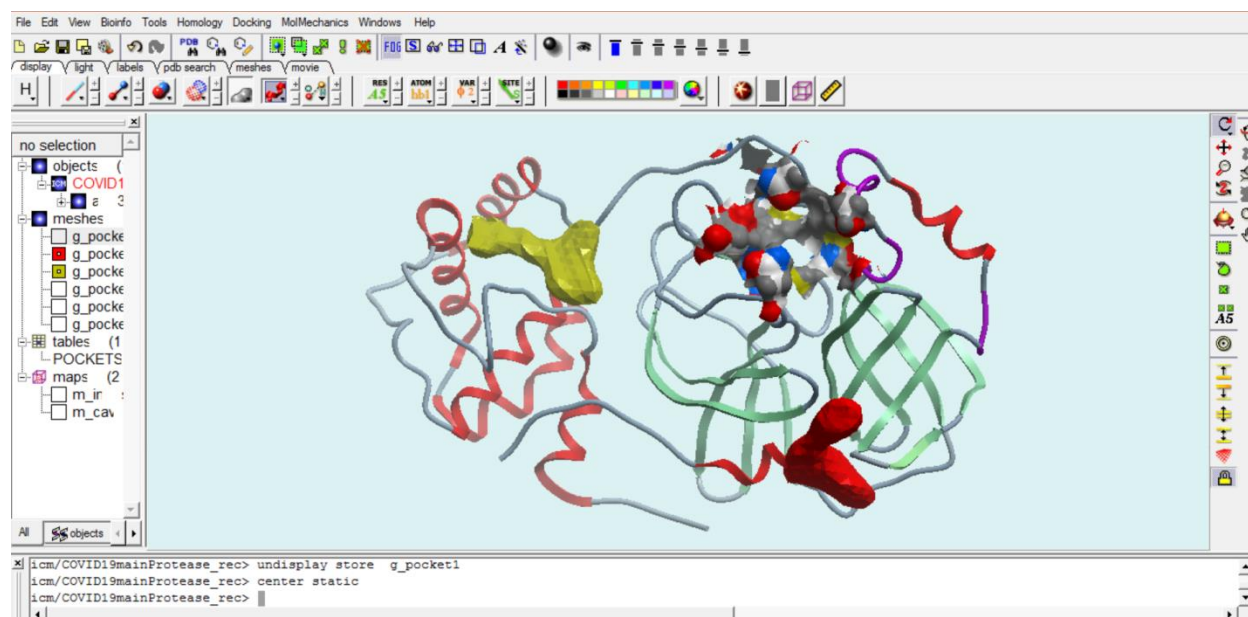


Figure (1) the main predicted binding scaffolds of PL protein at cut-off grid = 4.5

Preparing the table of candidate drugs for surveying for PL protease

The ligand based drug predictions were formulated by analyzing the ligand skeleton formula or functional group similarity in addition to the isomer and tautomer search were adopted. Further optimization of the candidate to maximally fit the binding site in the target protein then retrograde surveying of the present drug to be tested for their binding energies to the candidate target COVID-19 proteins. These steps are analyzed with different editing softwares like Molsoft pro.

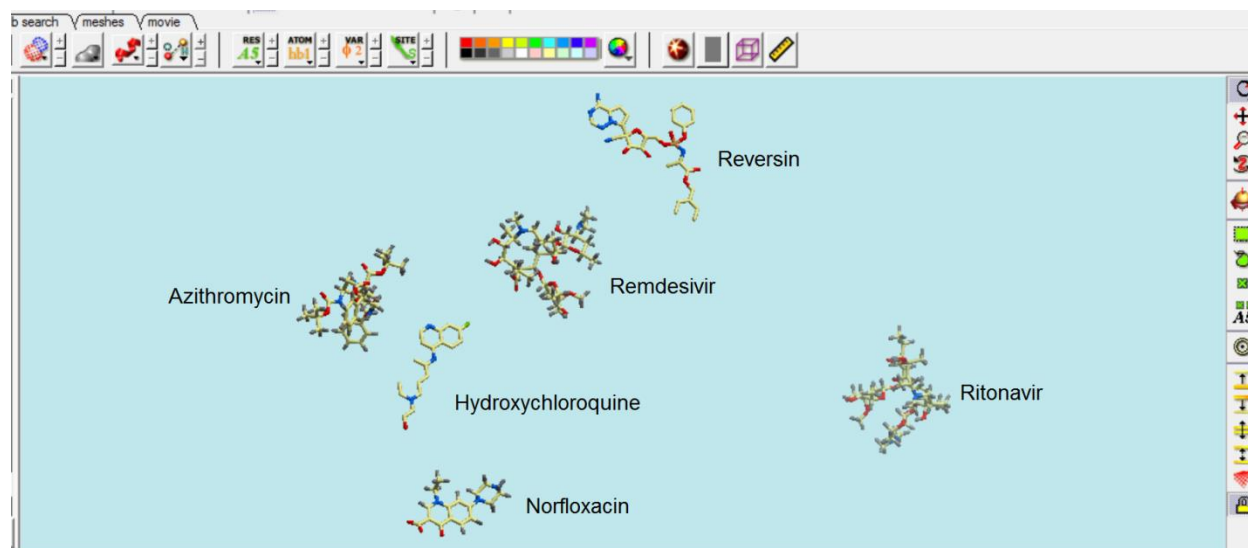


Figure (2) the drugs which were re-evaluated for their inhibitory efficacy against PL protease in Molsoft pro to analyze the trajectory, pockets, force field energy types and binding energy as indicators for the drug selectivity and affinity toward the binding site.

Preparing COVID-19 nucleocapsid protein N for ligand prediction

The same steps were followed in preparing the previous targets.

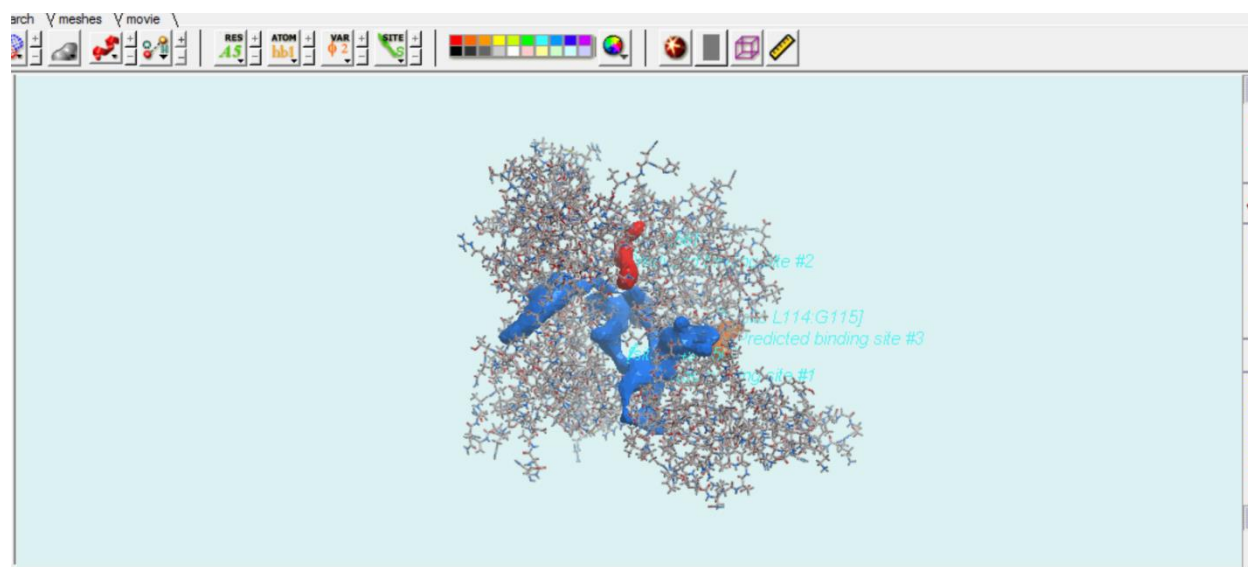


Figure (3) predicting and evaluating the main target sites in COVID-19 N protein for surveying inhibitor compounds and drugs.

Preprocessing of COVID-19 NSP12 (RDRP), PDB ID: 6nur.

Similar steps were followed in preparing the previous targets.

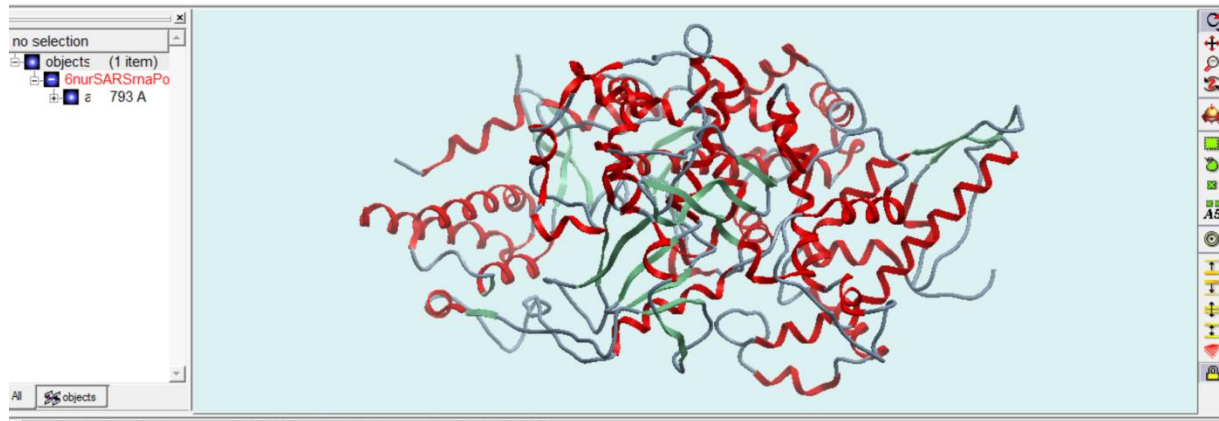


Figure (4) ribbon representation of COVID-19 NSP12 the main viral target is the RDRP before complexing with NSP7 and NSP8.

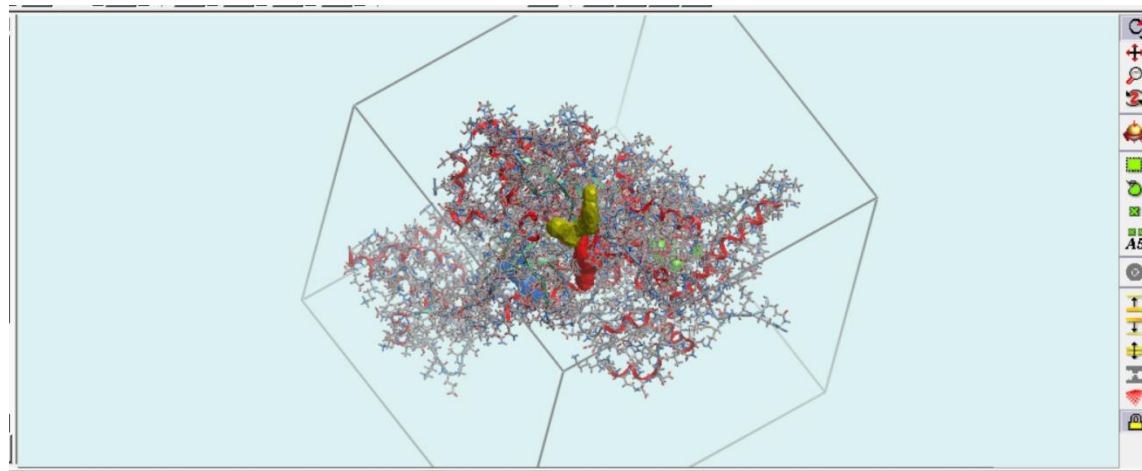


Figure (5) predicting and evaluating the target sites of RDRP

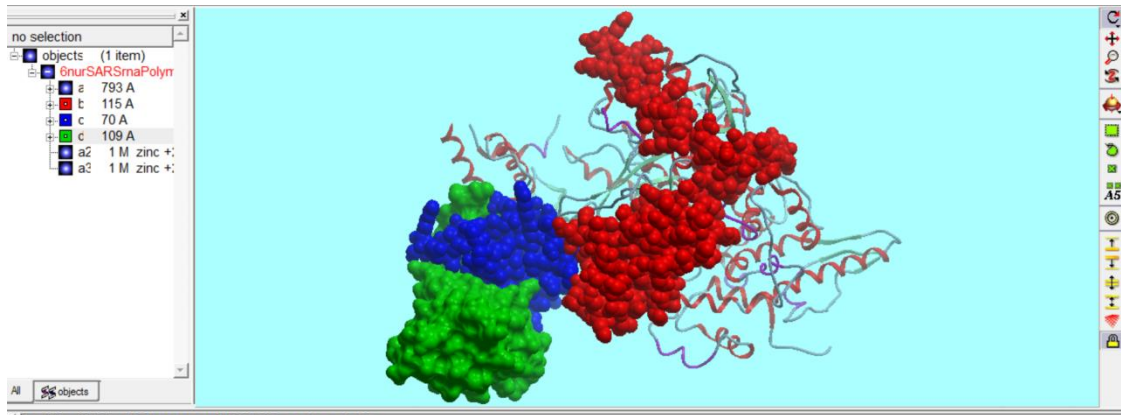


Figure (6) the complexed multidomain of COVID-19 NSP12 with NSP7, NSP8 and NSP9

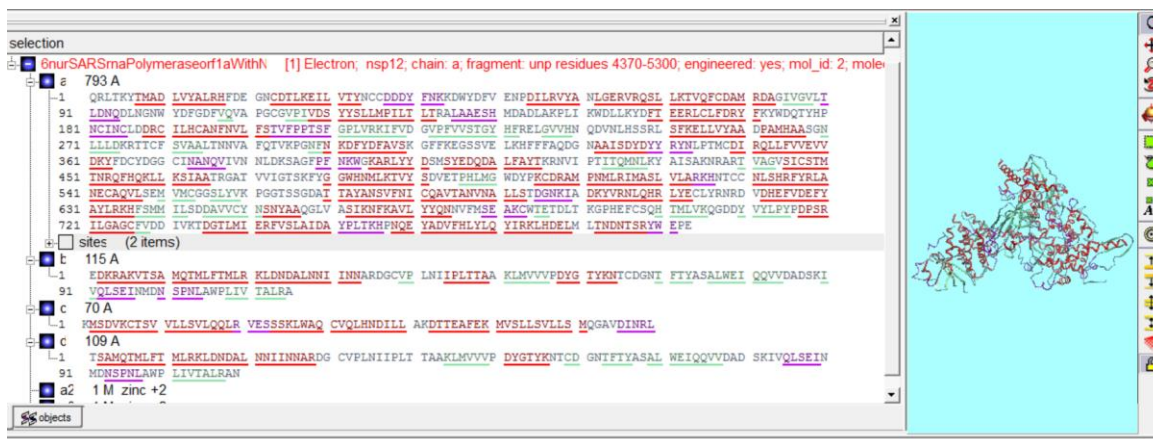


Figure (7) the FASTA format of the complexed multidomain of COVID-19 NSP12 with NSP7, NSP8 and NSP9

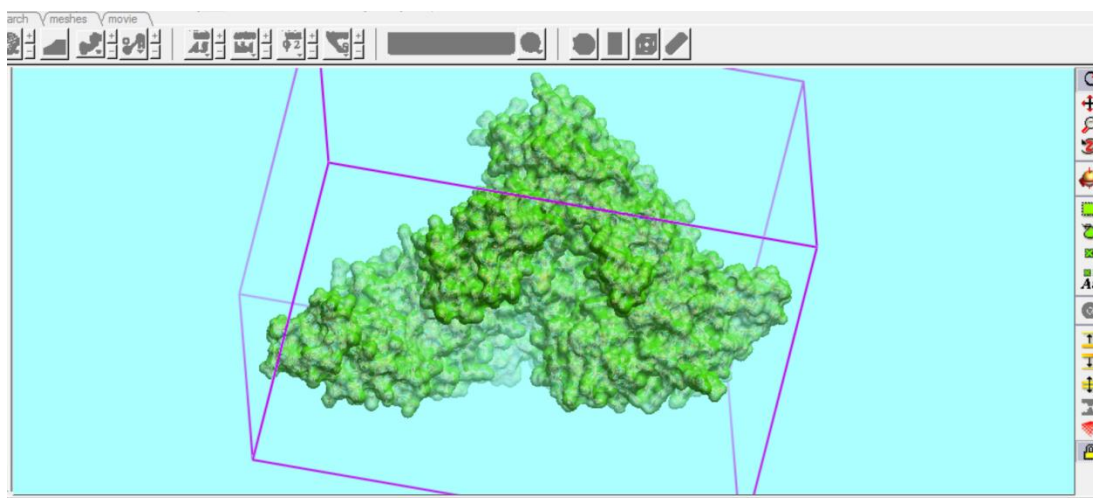


Figure (8) processing of the surface receptor of NSP12 complex

Processing of COVID-19 NSP9 the replicase associated protein

Similar steps were followed in preparing the previous targets.

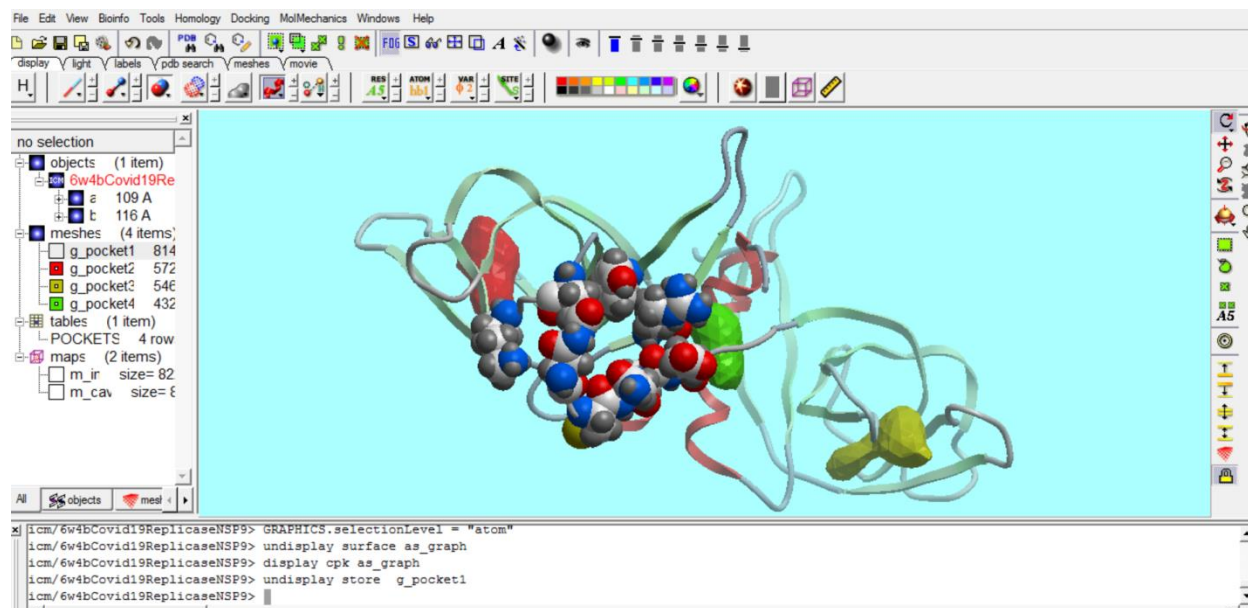


Figure (9) the 3D presentation of NSP9 (replicase associated protein) of COVID-19 with evaluating the target sites for drug design

Processing of the viral spike S protein

Similar steps were followed in preparing the previous targets.

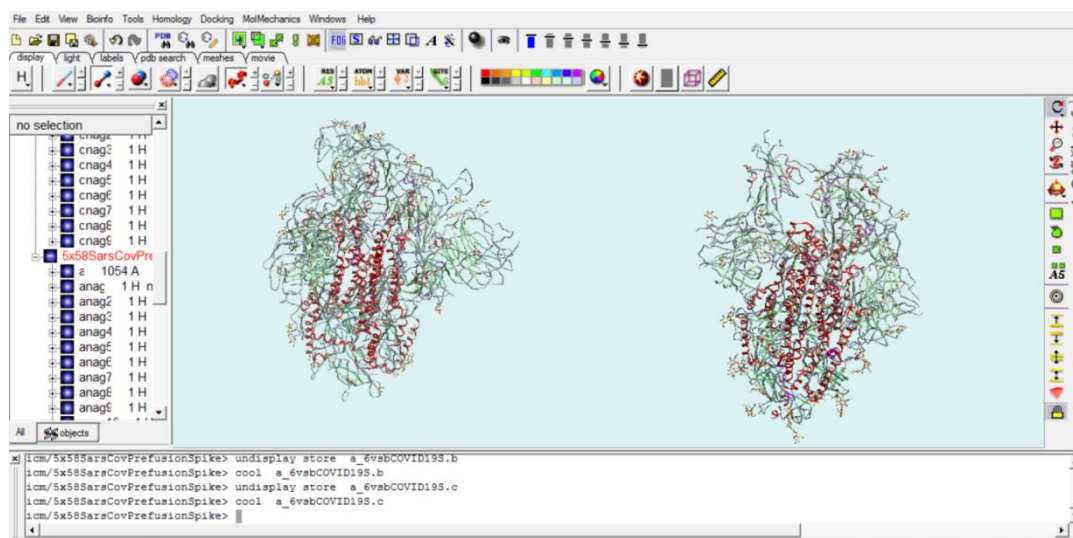
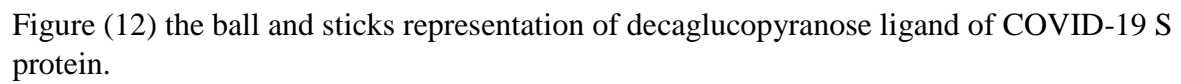
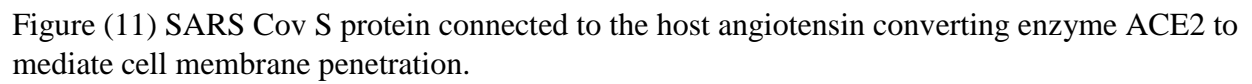


Figure (10) the 3D grid superimposition of the COVID-19 S protein (Rt) and Hcov S protein (Lt)



The 3D representation of the structural (S, E, M, N) and non-structural viral proteins were processed and viewed with Molsoft pro and RasMol as shown below.

Processing and representing of the coronavirus membrane M protein

Similar steps were followed in preparing the previous targets.

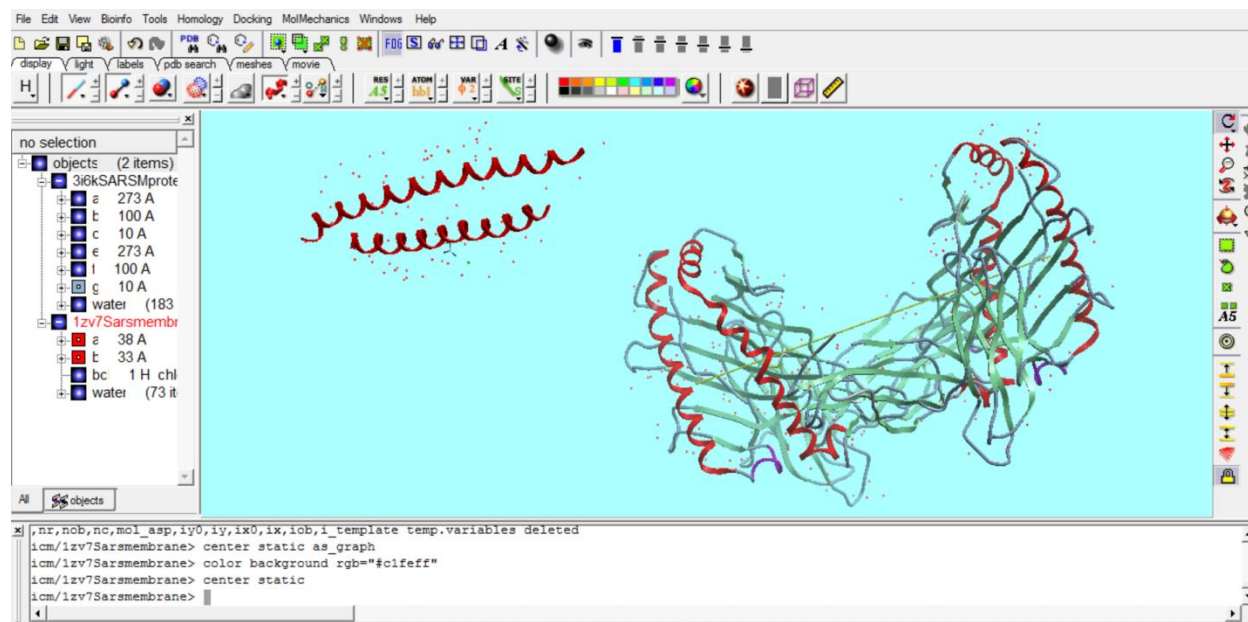


Figure (13) represent a prior assessment and visualizing the M protein related to SARS Cov linked to HLA (Rt) and free (Lt).

Laboratory assessment of the test antiviral compounds

To review drugs physicochemical, pharmacokinetics and toxicity properties in order to select the candidate of choice if it shows highly safe profile otherwise it is submitted to the next step. Steps of in silico and in vitro study were conducted in University of Kufa/ College of Medicine/ the research lab of the department of Pharmacology and Therapeutics in addition to some steps conducted in private lab.

Tis-phosphate buffer, alcohol 70%, norfloxacin pure powder, casein powder, papain-like protease powder, normal saline, distilled water, protease inhibitor powder were prepared.

Simulink image intensity analyzer was used to measure casein opalescence in response to the test papain-like proeae inhibiting drug, 1000X Olympus microscope, centrifuge, ultr-high speed mechanical stirrer, micropore filter and a mechanical shaker were used.

The artificial COVID-19 viral model

This model is aimed to test the bioavailability of the test drugs and their diffusion through the artificial viral membrane in addition to determining the inhibitory effects of those test drugs on papain-like protease by casein opalescence assay.

The model is made of an artificial membrane which is obtained from human RBCs to formulate RBCs vesicles RV as it has been described in the appendix I procedure. After preparing the artificial viral membrane, papain-like protease is then incorporated into the RV to form the artificial virus AV. Casein in a pure form and was formulated as described in procedure in appendix II.

Addition of 5 microg/ml of norfloxacin to the test well 1 in comparison to the control well and then the rate of casein opalescence change is measured with serial monitoring of the sample every one hour for 8 hours. Its inhibitory effects on viral protease with microscopic imaging and spectrophotometric analysis of absorbance in relation with casein hydrolysis. Data is compared with the in silico model to assess the consistency of the antiviral property of a test drug.

Data of the test and control wells are compared to assess norfloxacin inhibitory effect on viral papain-like protease. Such procedure could be repeated to assess other drugs for their protease inhibitory efficacy.

6. A drug repositioning step includes evaluating the level of similarity to an approved drug to be reassessed for antiviral use against that target protein.
7. Designing an in vitro model of antiviral assessment of the already preprocessed compounds and drugs. This model is formulated by incorporating the target COVID-19 papain-like protease into the lipid bilayer vesicles to evaluate the inhibitory efficacy of the experimental or approved drug on this enzyme function which causes casein opalescence.
8. Further surveying drug substructure and similarity to evaluate the optimal anti-COVID-19 drug.
9. Recommending further in vitro and in vivo study conduction and clinical study designs since the increase in antiviral treatment options is important for controlling viral infections
10. Data analysis with the compatible statistical tests for in silico and in vitro outcomes.

The results:

The outcomes of in vitro and in silico evaluation of compounds against COVID-19 were divided according the main pharmacological proteins targets related to this virus.

1-Findings of the inhibitory efficacy of test drugs repurposed against the COVID-19 main protease or the papain-like (PLP) protease.

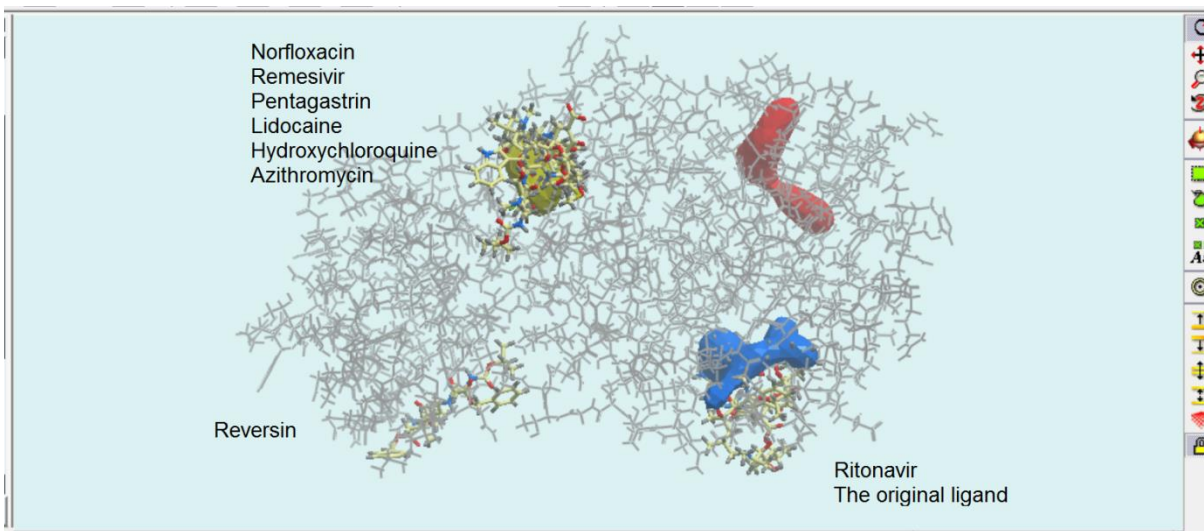


Figure (14) the binding sites of different tested drugs against COVID-19 protease PLP

Table (1) COVID-19 protease inhibition efficacy of the test drugs as a binding energy in kJ/mol as analyzed with Molsoft pro.

Test drugs against COVID-19 protease	The inhibitory efficacy In kJ/mol	Statistics Z score	Drug approval
1- Reversin	-137.30	>1.9	approved
2- Tetrazapentadecanoate	-129.57	>4.0	experimental
3- Remdesivir	-119.50	>1.6	approved
4- Pentagstrin	-118.82	>1,9	approved
5- Nitazoxanide	110.29	>1.9	approved
6- Norfloxacin	-103.70	>1.9	approved
7- Ritonavir	-95.22	>1.6	approved
8- Hydroxychloroquine	-86.42	>1.6	approved
9- Azithromycin	-85.78	>1.6	approved
10- Lidocaine	-80.69	>1.6	approved
Red = very potent			

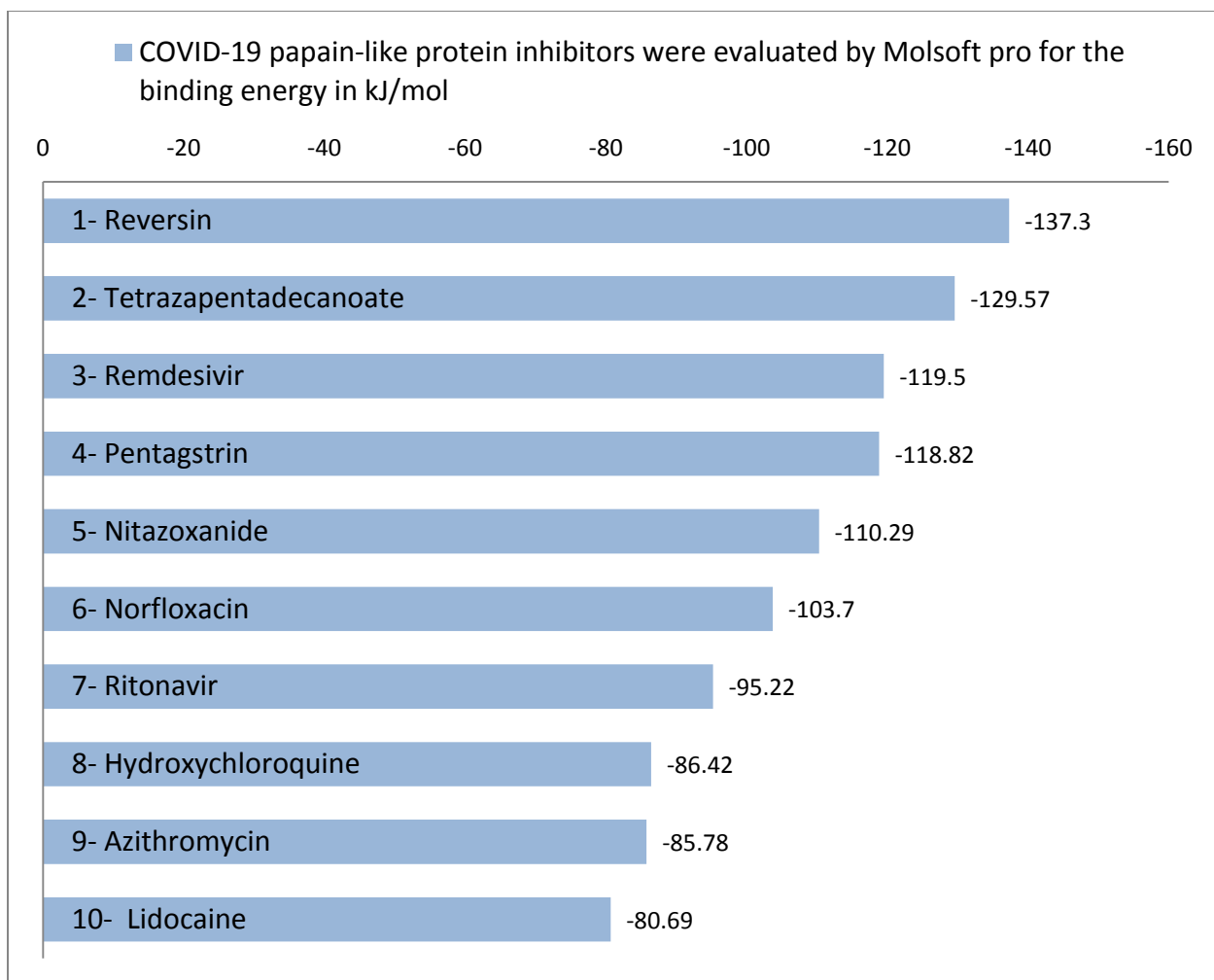


Figure (15) the comparative efficacy of drugs that were selected according to their skeletal superimposition to the predicted ligand of PL protease. The parameter was evaluated as the binding energy to PL protease in kJ/mol. Reversin, remdesivir, pentagastrin and norfloxacin showed enthalpy > 100 kJ/mol. The highest negative value the highest inhibitory efficacy.

Reversin showed the highest inhibitory efficacy against COVID-19 papain-like protease as indicated by the ligand-PLP binding energy with Molsoft pro and BioXLab analysis. The calculated inhibitory binding was -137.30 kJ/mol $z > 1.9$. as compared with the tetrazapentadecanoate -129.57 kJ/mol $z = 4.0$, whereas remdesivir, pentagastrin, nitazoxanide and norfloxacin had a moderate antiPLP activity (>- 100 kJ/mol).

2- Findings of the in vitro study design of COVID-19 papain-like protease

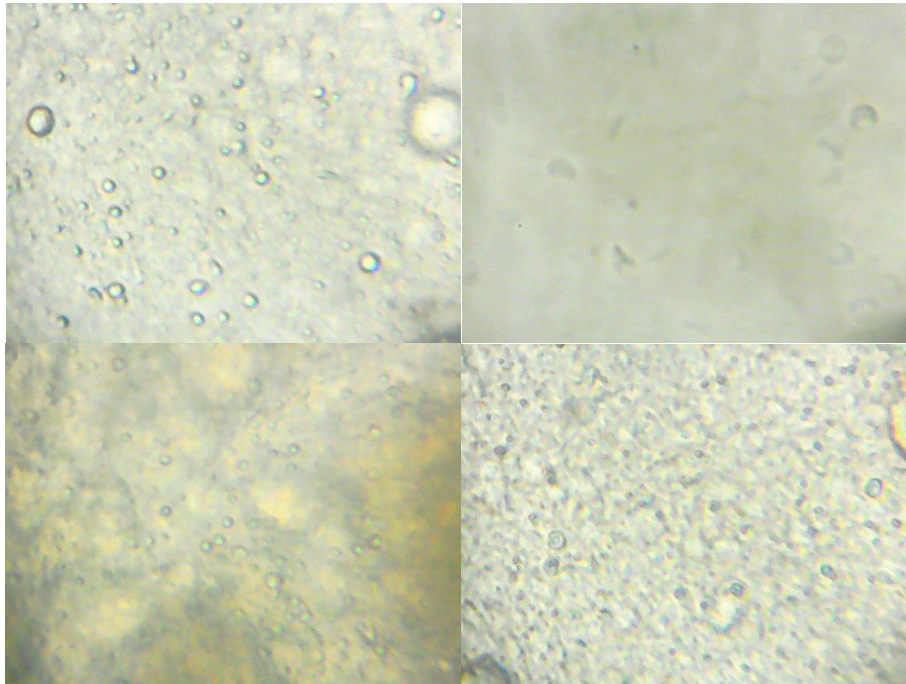


Figure (16) in vitro antiviral model under 1000X microscopic image showing the casein micelles of the blank sample (upper left) and the artificial papain-like protease containing vesicles AV (upper right). The lower left sample represents casein sample mixed with AV and the lower right sample shows the casein sample mixed with AV and norfloxacin to assess its effect on papain-like protease.

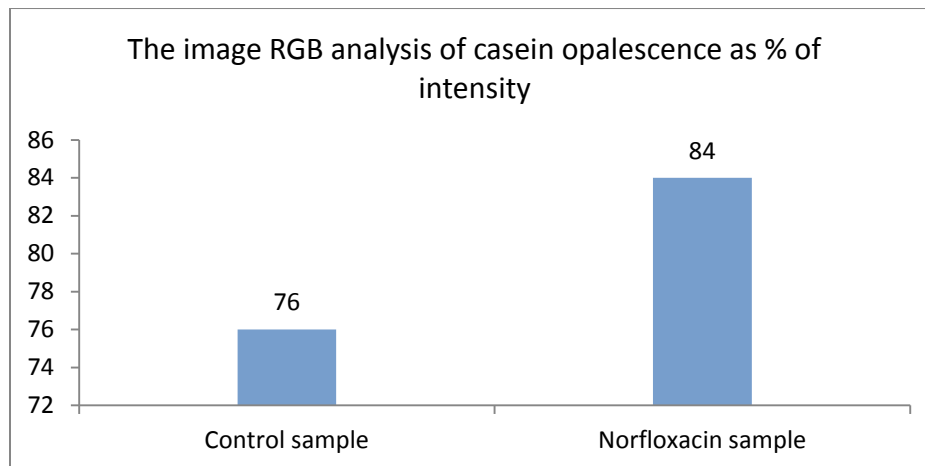


Figure (17) the relative protection with norfloxacin against casein decrease in opalescence. This decrease in opalescence is due to the effect of papain-like protease in hydrolyzing casein at cysteine residues. It was obvious that norfloxacin had a protective effect which indicates its papain-like protease inhibitory action. Casein opalescence was measured with Simulink image analysis to determine the relative loss in image intensity.

3- Findings of COVID-19 nucleocapsid N inhibiting ligands

Table (2) COVID-19 nucleocapsid N inhibiting ligands evaluated and represented by Molsoft pro and iGemdock

COVID-19 nucleocapsid N inhibitor drugs	Anti-N binding energy in kJ\mol	Satistics Z score	Drug approval
Phenanthridine glycenamide	-102.10	4.2	experimental
Morpholin quinoline	-90.70	4.0	experimental

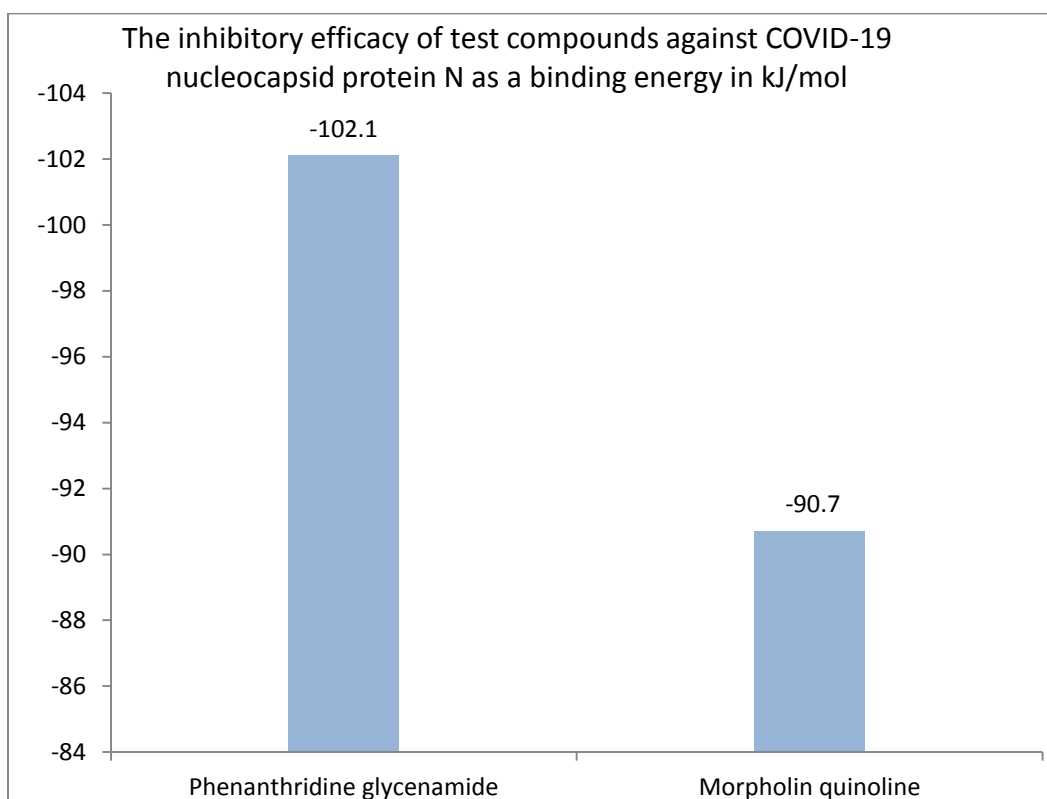


Figure (18) the main COVID-19 nucleocapsid N binding ligands and their binding energy in kJ/mol

Table (3) the inhibitory efficacy of a group of drugs evaluated for their NSP12 (RDRP) binding energy with Molsoft pro

COVID-19 NSP12 inhibiting drugs	Anti-RNA polymerase Efficacy in kJ/mol	Statistics Z score	Change
1- Benzylglutathione	-129.28	>1.9	experimental
2- Lopinavir	-118.80	2.2	approved
3. Hydroxymethylglutathione	-112.00	>1.9	experimental
4- Glutathione	-105.72	>1.9	approved
5- Favonavir	-93.10	2.2	approved
6- Ribavirin	-91.50	2.2	approved
red = highly potent			

Assessment of the drugs activity against NSP12 (RDRp) included selected ligand analogues benzylglutathione, hydroxylmethylglutathione, lopinavir, glutathione, favonavir and ribavirin showed promising results against COVID-19 infection. Benzylglutathione had the highest inhibitory efficacy against COVID-19 RDRp with a binding energy of -129.28 kJ/mol $z > 1.9$ and lopinavir came second in efficacy with a binding energy of -118.80 kJ/mol, $z = 2.2$. Glutathione had a moderate RDRp binding energy (> -105.72 kJ/mol).

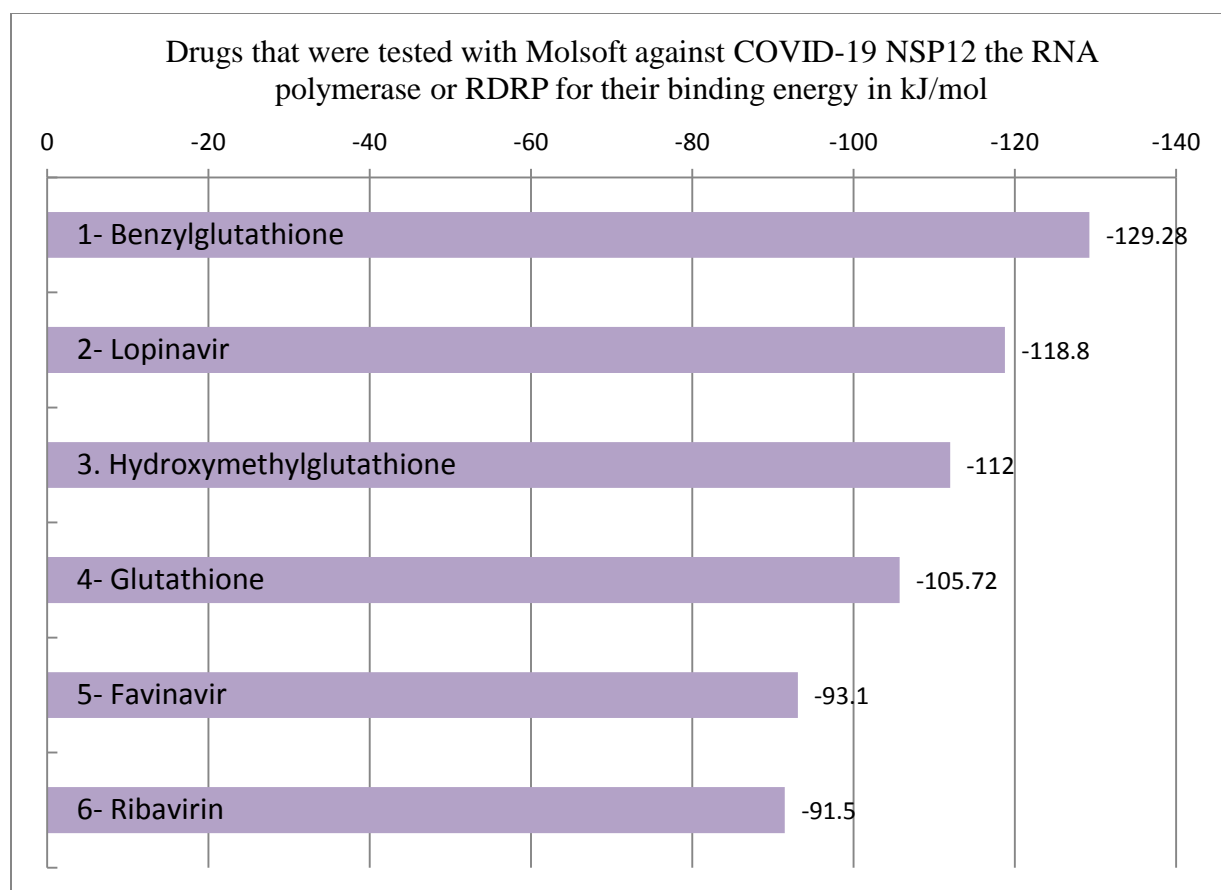


Figure (19) the comparative efficacy of test drugs against COVID-19 RDRP, the parameter was evaluated as the binding energy to RDRP in kJ/mol. Benzylglutathion, lopinavir, hydroxymethylchloroquine and glutathione had a binding enthalpy > 100 kJ/mol and z score > 1.9 .

Discussion

Although viruses like COVID-19 having limited number of protein, however these proteins contain multiple pockets and binding sites to guide the rational for developing many antiviral compounds. (25)

Among the COVID-19 genome set of proteins, the current study concerned with papain-like protease (PLpro) and RNA-dependent RNA polymerase (RdRp) as the main targets. The near future steps will include more precise optimization of the ligands against other COVID-19 proteins. (26)(27)(28)

Ten drugs were assessed by in silico model of studying the PLP binding energy in kJ/mol. These drugs included reversin, pentagastrin, the original ligand tetrapentadecanoate, remdesivir, nitazoxanide, norfloxacin, hydroxychloroquine, ritonavir, azithromycin and lidocaine. The anti-resistance reversin showed the highest inhibitory efficacy against COVID-19 papain-like protease as indicated by the ligand-PLP binding energy. The calculated inhibitory binding energy was -137.30 kJ/mol $z > 1.9$ as compared with the tetrapentadecanoate -129.57 kJ/mol $z = 4.0$, whereas remdesivir, pentagastrin, nitazoxanide and norfloxacin had a moderate PLP binding energy (> -100 kJ/mol). In vitro PLP inhibiting activity for norfloxacin was slightly consistent with the in silico outcomes. The designed in vitro model for COVID-19 was not highly reliable due to the limited facilities under the current epidemic, however adopting a more sophisticated in vitro models against risky viral infections is the key of development of new antiviral drugs because this model is accessible for a wider number of researchers. (29) (30) (31)

Other tested drugs against PLP showed just a weak binding energy (-80 to -95 kJ/mol). These drugs included hydroxychloroquine, azithromycin, ritonavir and lidocaine. Antiviral activities of remdesivir, ritonavir, nitazoxanide, quinolones and some oligopeptides were confirmed by different studies against viral infections other than COVID-19. (32)

Assessment of the drugs activity against NSP12 (RdRp) included selected ligand analogues benzylglutathione, hydroxymethylglutathione, lopinavir, glutathione, favinavir and ribavirin showed promising results against COVID-19 infection. Benzylglutathione had the highest

inhibitory efficacy against COVID-19 RDRp with a binding energy of -129.28 kJ/mol $z > 1.9$ and lopinavir came second in efficacy with a binding energy of -118.80 kJ/mol, $z = 2.2$. Glutathione had a moderate RDRp binding energy (> -105.72 kJ/mol). Glutathione, lopinavir, favinavir and ribavirin had also antiviral effects on other viral infections (33) (34)

Comparing the 3D conformational binding was also evaluated and showed that some of the evaluated drugs had different docking sites while others had the same docking sites with different binding residues and energies.

Evaluation of the ligands against NSP9 COVID-19 replicase associated protein

The predicted ligands binding to NSP9 were evaluated using Molsoft pro and they showed no significant similarity in their structure or formula although the similarity cut-off was set at 0.4.

Evaluation of the ligands against COVID-19 spike S protein.

This protein was relatively large in size and it performs a structural unit and cell penetration mechanisms which mean it has a macromolecular surface of binding site so that direct inhibitory actions of small drugs may have a limited efficacy against S protein. (35)(36)(37). One ligand was predicted to be of applied value which was the decaglucofuranose.

The COVID-19 relevant host components furin and ACE2

Furin is an essential housekeeping enzyme for activating many metabolic and cellular proteins it needs for a highly selective mechanism of modifying its action in order to spare the physiologic effects (38) On the other hand, angiotensin converting enzyme ACE2 is an essential cytoprotective enzyme although it's a one mediator of coronaviruses penetration into the cell.(39)(40)

Other host components like INF gamma and ILs and vaccines

Immunotherapy is critically important in treating and controlling viral diseases. Vaccines may comprise the versatile health care measures against future viral epidemics; however several weak points are correspondent with immunotherapy in the future of viral infections. Of these drawbacks of immunotherapy is its expiry of protection since most of pathogens have the virulent strategy to change their antigens, moreover, host immune response has its own duration

of action which may extend from weeks to few years, however immunoprotection is uncommonly to be lifelong. (41)

As it was confirmed by bioinformatics of different databases, another critical point in viral immunotherapy is that most of the viral infections will eventually exaggerate immune system. This response is at most the cytotoxic and and pathogenic event that in many instances gives rise to the seriousness of even simple viral infection so that immunotherapy will remain risky and in many times it is cautious. (42-46)

For all these reasons the antiviral drug therapy is highly promising to combat the epidemic infections.

Conclusion

From the overall results quinolones, antiviral drugs, glutathione and peptides like reversin and pentagastrin had a promising inhibitory efficacy against COVID-19 protein targets so that developing COVID-19 proteins blockers from the approved drugs is an accessible approach and could provide rapid and safe therapeutic option against the risky viral epidemics.

Statement of novelty:

Quinolones, antiviral drugs, glutathione and the small peptide compounds are promising inhibitors for different COVID-19 proteins and provide the rationale to be repurposed in further studies against coronavirus infections.

Acknowledgement:

A high respect and gratitude for the accessible and helpful scientific databases offered by NIH components that facilitate researches to speed the drug design and development processes against the global health challenges like the viral epidemics. Special thanks to NCBI, GenBank, BLAST, Pubmed and PubChem and special thanks and gratitude to PDB, NeXtProt, Genomix, GeneCard and Viral genome for their websites databases that helped us in conducting updated antiviral research. Thanks and respect to Uniprot, DrugBank, ChemSpider for their help in cheminformatics. Deep gratitude and respect to the offered editor softwares for drug design and survey namely Molsoft pro, BioXLab and SwissDock.

References

- 1- Firas A. Rabi¹, Mazhar S. Al Zoubi, Ghena A. Kasasbeh³, Dunia M. Salameh and Amjad D. Al-Nasser. SARS-CoV-2 and Coronavirus Disease 2019: What We Know So Far by *Pathogens* 2020, 9(3), 231.
- 2- Canrong Wu, Yang Liu, Yueying Yang, Peng Zhang, Wu Zhong, Yali Wang, Qiqi Wang, Yang Xu, Mingxue Li, Xingzhou Li, Mengzhu Zheng, Lixia Chen, Hua Li P. Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods II: S2211-3835(20)30299-9.
- 3- Drug targets for corona virus: A systematic review . Manisha Prajapat, Phulen Sarma, Nishant Shekhar, Pramod Avti, Shweta Sinha, Hardeep Kau, Subodh Kumar, Anusuya Bhattacharyya, Harish Kumar¹, Seema Bansal¹, Bikash Medhi¹ Year : 2020 | Volume : 52 | Issue : 1 | Page : 56-65
- 4- Drug targets for corona virus: A systematic review. Prajapat M, Sarma P, Shekhar N, et al. *Indian J Pharmacol.* 2020;52(1):56–65.
- 5- Fehr, A.R.; Perlman, S. Coronaviruses: An overview of their replication and pathogenesis. *Methods Mol. Biol.* **2015**, 1282, 1–23.
- 6- Masters PS. The molecular biology of coronaviruses. *Adv Virus Res.* 2006;66:193–292.
- 7- Schoeman, D., & Fielding, B. C. (2019). Coronavirus envelope protein: current knowledge. *Virology Journal*, 16(1). doi:10.1186/s12985-019-1182-0.
- 8- Yang, S., Fu, C., Lian, X., Dong, X. & Zhang, Z. Understanding human-virus protein-protein interactions using a human protein complex-based analysis framework. *mSystems* **4**, e00303 (2019).
- 9- Dyal, J. et al. Repurposing of clinically developed drugs for treatment of Middle East respiratory syndrome coronavirus infection. *Antimicrob. Agents Chemother.* **58**, 4885–4893 (2014)
- 10- Wink, M. (2015). Modes of Action of Herbal Medicines and Plant Secondary Metabolites. *Medicines*, 2(3), 251–286. doi:10.3390/medicines2030251
11. Ge XY, Li JL, Yang XL, Chmura AA, Zhu G, Epstein JH, et al. Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature* 2013; 503: 535-8.

12. Han DP, Penn-Nicholson A, Cho MW. Identification of critical determinants on ACE2 for SARS-CoV

entry and development of a potent entry inhibitor. *Virology* 2006; 350: 15-25.

13. Li W, Moore MJ, Vasilieva N, Sui J, Wong SK, Berne MA, et al. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* 2003; 426: 450-4. RRR-8

14-David Baxter. Active and passive immunity, vaccine types, excipients and licensing *Occupational Medicine*, Volume 57, Issue 8, December 2007, Pages 552–556, <https://doi.org/10.1093/occmed/kqm110>

15-Eakachai Prompetchara,1,2,3 Chutitorn Ketloy,1,2 Tanapat Palaga4,5 Immune responses in COVID-19 and potential vaccines: Lessons learned from SARS and MERS epidemic . *Asian Pac J Allergy Immunol* DOI 10.12932/AP-200220-0772

16-Afzal O, Kumar S, Kumar R, Firoz A, Jaggi M, Bawa S. Docking based virtual screening and molecular dynamics study to identify potential monoacylglycerol lipase inhibitors. *Bioorg Med Chem Lett*. 2014;24:3986–96.

17-Karthick, V., Nagasundaram, N., Doss, C. G. P., Chakraborty, C., Siva, R., Lu, A., Zhu, H. (2016). Virtual screening of the inhibitors targeting at the viral protein 40 of Ebola virus. *Infectious Diseases of Poverty*, 5(1). doi:10.1186/s40249-016-0105-1

18- Leidner, F., Kurt Yilmaz, N., & Schiffer, C. A. (2019). Target-Specific Prediction of Ligand Affinity with Structure-Based Interaction Fingerprints. *Journal of Chemical Information and Modeling*.

19- Zou, X.; Sun, Y.; Kuntz, I. D. Inclusion of solvation in ligand binding free energy calculations using the generalized-born model. *J. Am. Chem. Soc.* 1999, 121, 8033–8043.

20- Wang, R.; Lai, L.; Wang, S. Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *J. Comput.-Aided Mol. Des.* 2002, 16, 11–26.

21- Totura, A. L., & Bavari, S. (2019). Broad-spectrum coronavirus antiviral drug discovery. *Expert Opinion on Drug Discovery*, 1–16. doi:10.1080/17460441.2019.1581171

22- Janosko K, Holbrook MR, Adams R, et al. Safety Precautions and Operating Procedures in an (A)BSL-4 Laboratory: 1. Biosafety Level 4 Suit Laboratory Suite Entry and Exit Procedures. *J Vis Exp*. 2016;(116):52317. Published 2016 Oct 3. doi:10.3791/52317 –

23- Wimmer E, Mueller S, Tumpey TM, Taubenberger JK. Synthetic viruses: a new opportunity to understand and prevent viral disease. *Nat Biotechnol*. 2009;27(12):1163–1172. doi:10.1038/nbt.1593

- 24- Siti Khaerunnisa ,Hendra Kurniawan, Rizki Awaluddin, Suhartati Suhartati, Soetjipto Soetjipto. Potential Inhibitor of COVID-19 Main Protease (Mpro) from Several Medicinal Plant Compounds by Molecular Docking Study , Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 March 2020. doi: 10.20944/preprints202003.0226.v1
- 25- Fehr AR1, Perlman S 1- Coronaviruses: an overview of their replication and pathogenesis. *Methods Mol Biol.* 2015;1282:1-23.
- 26- Baez-Santos, Y. M., St John, S. E. & Mesecar, A. D. The SARS-coronavirus papain-like protease: structure, function and inhibition by designed antiviral compounds. *Antiviral Res.* **115**, 21–38 (2015).
- 27- Lu, A. et al. Attenuation of SARS coronavirus by a short hairpin RNA expression plasmid targeting RNA- dependent RNA polymerase. *Virology* **324**, 84–89 (2004).
- 28- Wang, Z. et al. Inhibition of severe acute respiratory syndrome virus replication by small interfering RNAs in mammalian cells. *J. Virol.* **78**, 7523–7527 (2004).
- 29- Liu, C., Zhou, Q., Li, Y., Garner, L. V., Watkins, S. P., Carter, L. J., ... Albaiu, D. (2020). Research and Development on Therapeutic Agents and Vaccines for COVID-19 and Related Human Coronavirus Diseases. *ACS Central Science*.
- 30- Arya, Rimanshee; Das, Amit; Prashar, Vishal; Kumar, Mukesh (2020): Potential inhibitors against papain-like protease of novel coronavirus (SARS-CoV-2) from FDA approved drugs. *ChemRxiv*. Preprint.
- 31- Pizzorno, A., Terrier, O., Nicolas de Lamballerie, C., Julien, T., Padey, B., Traversier, A., Rosa-Calatrava, M. (2019). Repurposing of Drugs as Novel Influenza Inhibitors From Clinical Gene Expression Infection Signatures. *Frontiers in Immunology*, 10.
- 32-Tim Smith, PharmD, BCPS; Jennifer Bushek, PharmD; Tony Prosser, PharmD Clinical Drug Information | Clinical Solutions COVID-19 Drug Therapy – Potential Options Updated March 26, 2020 Elsevier © 2020
- 33- Te Velhuis, A. J. W., Arnold, J. J., Cameron, C. E., van den Worm, S. H. E., & Snijder, E. J. (2011). The RNA polymerase activity of SARS-coronavirus nsp12 is primer dependent. *Nucleic Acids Research*, 39(21), 9458–9458.
- 34- Xu X, Liu Y, Weiss S, Arnold E, Sarafianos SG, Ding J .Molecular model of SARS coronavirus polymerase: implications for biochemical functions and drug design. *Nucleic Acids Res.* 2003 Dec 15; 31(24):7117-30
- 35- Sutton, G., Fry, E., Carter, L., Sainsbury, S., Walter, T., Nettleship, J., Berrow, N., Owens, R., Gilbert, R., Davidson, A., et al., 2004. The nsp9 replicase protein of SARS-coronavirus, structure and functional insights. *Structure* 12, 341–353.

- 36- Bavan, S., Sherman, B., Luetje, C. W., & Abaffy, T. (2014). Discovery of Novel Ligands for Mouse Olfactory Receptor MOR42-3 Using an In Silico Screening Approach and In Vitro Validation. *PLoS ONE*, 9(3), e92064. doi:10.1371/journal.pone.0092064
- 37- Smith, R. D., Engdahl, A. L., Dunbar, J. B., & Carlson, H. A. (2012). Biophysical Limits of Protein–Ligand Binding. *Journal of Chemical Information and Modeling*, 52(8), 2098–2106.
- 38-Thomas, G. (2002). Furin at the cutting edge: From protein traffic to embryogenesis and disease. *Nature Reviews Molecular Cell Biology*, 3(10), 753–766.
- 39-Wong, S.K., Li, W., Moore, M.J., Choe, H., Farzan, M., 2004. A 193-amino acid fragment of the SARS coronavirus S protein efficiently binds angiotensin-converting enzyme 2. *Journal of Biological Chemistry* 279, 3197–3201.
- 40- Xiao, X., Chakraborti, S., Dimitrov, A.S., Gramatikoff, K., Dimitrov, D.S., 2003. The sars-cov s glycoprotein: expression and functional characterization. *Biochemical and Biophysical Research Communications* 312, 1159–1164.
- 41- Naran, K., Nundalall, T., Chetty, S., & Barth, S. (2018). Principles of Immunotherapy: Implications for Treatment Strategies in Cancer and Infectious Diseases. *Frontiers in Microbiology*, 9.
- 42-Troy, N. M., & Bosco, A. (2016). Respiratory viral infections and host responses; insights from genomics. *Respiratory Research*, 17(1). doi:10.1186/s12931-016-0474-9.
43. <https://pubchem.ncbi.nlm.nih.gov/substrate/178103463>:29.3.2020.
44. <https://www.uniprot.org/uniprot/P0C6X7>.
45. <https://www.rcsb.org/structure/5U98>:29.3.2020.
46. <https://www.drugbank.ca/drugs>.
- 47- Qing Xia†, YongtaiZhang†, ZheLi,XuefengHou,NianpingFeng. Red bloodcellmembrane-camouflaged nanoparticles:anoveldrugdeliverysystemfor antitumorapplication. *Acta Pharmaceutica Sinica B*.2019.9(4):675-689.
48. Umeo Yoshino. Studies on the Opalescence in Sodium-caseinate Solution Developed by the Milk Coagulating Enzymes. *Journal of the Agricultural Chemical Society of Japan*. 29 Jul 2014.

Appendices

Appendix I

The red blood cell membrane camouflaged nanoparticle

Appendix II

Studies on the opalescence in sodium-caseinate solution developed by the milk coagulating enzymes

Appendix III

List of abbreviations

Appendix IV

NCBI viral genome and proteome data

Appendix I

The red blood cell membrane camouflaged nanoparticle

Preparation methods of RBCM-NPs Various methods including physical and chemical property-based techniques have been reported for encapsulating drugs or other bioactive agents in erythrocytes, such as hypotonic hemolysis, hypotonic dilution, hypotonicdialysis hypotonic preswelling and osmotic pulse or chemical perturbation of the membrane along with electrical breakdown. In addition, endocytosis, lipid fusion, and intrinsic uptake of substances by erythrocytes are used to encapsulate different compounds. To achieve successful covering, the encapsulated compounds may require a considerable degree of water solubility as well as erythrocyte inactivity; i.e., lack of physical and chemical interactions with erythrocyte membranes to avoid leakage from the loaded RBCs, which could result in toxicological problems. Therefore, drugs are commonly prepared into a nano preparation with less toxicity and higher stability as the core, followed by the use of RBCMs to disguise the nano-preparations to avoid identification by immune systems. Preparation of RBCM-derived vesicles (RVs) In general, the optimized and common preparation of RBCM-NPs can be divided into two parts: membrane-derived vesicles from RBCs and vesicle-particle fusion. RVs are obtained by combining two steps, hypotonic treatment and sequential extrusion. Fresh whole blood obtained from the organism (e.g., mouse) is centrifuged at 4 °C to maintain protein activity, and then the serum and buffycoat are removed to collect erythrocytes. The resulting RBCs are repeatedly washed with phosphate buffered saline (PBS) and re-collected by centrifugation to remove residual plasma and other unwanted cells. RBC ghosts are then acquired by hypotonic treatment, where in the washed RBCs are gently mixed with an excess of 0.25% PBS and held to release the intracellular RBC components. Following high-speed centrifugation to remove hemoglobin, the RBC ghosts comprising the resulting pink precipitate are sonicated in a bath sonicator, and then passed through different pore size polycarbonate porous membranes using an Avanti mini-extruder to obtain the target size of RVs. To keep the membrane bio-active, protease inhibitors are usually added to the samples and the samples are stored at 4 °C. (47)

Appendix II

Studies on the opalescence in sodium-caseinate solution developed by the milk coagulating enzymes

Casein was prepared from unpasteurized cow's skim milk by the method of Hipp. 2) α -Casein was prepared by the method of Warner S) and β -casein by the method of Hipp 4) with the use of urea, specifically. Both of them were electrophoretically pure at pH 7.6.

The enzymes used were Hansen's Rennet Tablet and Mikuni's pepsin preparation. Hydrolysis was carried out at pH 6.5 on 2.5% ; ; sodium caseinate solution at 35°C. (48)

Appendix III

List of abbreviations

NSP: non-structural protein

PLP: papain-like protease

Orf: open file read of gene

N: nucleocapsid protein

M: membrane glycoprotein

S: spike viral glycoprotein

E: envelope protein

RV: red blood cell derived vesicles

AV: artificial viral vesicle

SARS: Severe acute respiratory syndrome coronavirus

COVID-19: coronavirus disease 2019

HLA: human lymphocytes antigen

PDB: protein data bank

ICM: internal coordinate mechanics of Molsoft pro

ACE2: angiotensin converting enzyme 2

RBCs: red blood cells

RDRp: RNA dependent RNA polymerase

Appendix IV

Whole genome FASTA of COVID-19 from NCBI database

```
source      1..29903
            /organism="Severe acute respiratory syndrome coronavirus
            2"
            /mol_type="genomic RNA"
            /isolate="Wuhan-Hu-1"
            /host="Homo sapiens"
            /db_xref="taxon:2697049"
            /country="China"
            /collection_date="Dec-2019"
5'UTR      1..265
gene       266..21555
            /gene="orf1ab"
CDS        join(266..13468,13468..21555)
            /gene="orf1ab"
            /ribosomal_slippage
            /note="pp1ab; translated by -1 ribosomal frameshift"
            /codon_start=1
            /product="orf1ab polyprotein"
            /protein_id="QHD43415.1"
            /translation="MESLVPGFNEKTHVQLSLPVLQVRDVLVRGFGDSVEEVLSEARQ
            HLKDGTCGLVEVEKGVLPQLEQPYVFIKRS DARTAPHGHVMVELVAELEGIQYGRSGE
```

TLGVLVPHVGEIPVAYRKVLLRKNGNKGAGGHSYGADLKSFDLGDELGTDPYEDFQEN
WNTKHSSGVTRMLRELNGGAYTRYVDNNFCGPDGYPLECIKDLLARAGKASCTLSEQ
LDFIDTKRGVYCCREHEHEIAWYTERSEKSYELQTPFEIKLAKKFDTFNCECPNFVFP
LNSIIKTIQPRVEKKKLDGFMGRIRSVYPVASPNECNQMCLSTLMKCDHCGETSWQTG
DFVKATCEFCGTENLTKEGATTCGYLPQNAVVKIYCPACHNSEVGPEHSLAEYHNESG
LKTILRKGGRTIAFGGCVFSYVGCHNKCAYWVPRASANIGCNHTGVVGEGSEGLNDNL
LEILQKEKVNINIVGDFKLNEEIAILASFSASTSAFVETVKGLDYKAFKQIVESCGN
FKVTKGKAKKGAWNIGEQKSILSPLYAFASEAARVVRISFRTLETAQNSVRVLQKAA
ITILDGISQYSLRLIDAMMFTSDLATNNLVVMAYITGGVVQLTSQWLTNIFGTVYEKL
KPVLDWLEEKFKEGVEFLRDGWEIVKFISTCACEIVGGQIVTCAKEIKESVQTFFKL
NKFLALCADSIIIGGAKLKALNLGETFVTHSKGLYRKCVKSREETGLLMPLKAPKEII
FLEGETLPTEVLTEEVVLKTGDLQPLEQPTSEAVEAPLVGTPVCINGLMMLLEIKDTEK
YCALAPNMMVTNNTFTLKGGAFTKVTFGDDTVIEVQGYKSVNITFELDERIDKVLNEK
CSAYTVELGTEVNEFACVVADAVIKTLQPVSELLTPLGIDLDEWSMATYYLFDESGEF
KLASHMYCSFYPPDEDEEEGDCEEEEFEPSTQYEGTEDDYQGKPLEFGATSAALQPE
EEQEEDWLDDDSQQTVGQQDGSQDQTTTIQTTIVEVQPQLEMELTPVVQTIEVNSFSG
YLKLTDNVYIKNADIVEEAKKVKPTVVVNAANVYLKHGGGVAGALNKATNNAMQVESD
DYIATNGPLKVGGSCLVSGHNLAKHCLHVVGPNVNKGEDIQLLKSAYENFNQHEVLLA
PLLSAGIFGADPIHSLRVCVDTVRTNVYLAVFDKNLYDKLVSSFLEMKSEKQVEQKIA
EIPKEEVKPFITESKPSVEQRKQDDKKIKACVEEVTTTLEETKFLTENLLLYIDINGN
LHPDSATLVSDIDITFLKKDAPYIVGDVVQEGVLTAVVIPTKKAGGTTEMLAKALRKV
PTDNYITTYPGQGLNGYTVVEAKTVLKKCKSAFYILPSIISNEKQEILGTVSWNLREM
LAHAEETRKLMPVCVETKAIVSTIQRKYKGIKIQEGVVDYGARFYFYTSKTTVASLIN
TLNDLNETLVTMPLGYVTHGLNLEEAAARYMRSLKVPATVSVSSPDAVTAYNGYLTSSS
KTPEEHFIETISLAGSYKDWSYSGQSTQLGIEFLKRGDKSVYYTSNPTTFHLDGEVIT
FDNLKTLLSLREVRTIKVFTTVDNINLHTQVVDMSMTYGGQFGPTYLDGADVTKIKPH
NSHEGKTFYVLPNDDTLRVEAFEYYHTTDPFLGRYMSALNHTKKWKYPQVNGLTSLIK
WADNNCYLATALLTLQQIELKFNPPALQDAYYRARAGEAANFCALILAYCNKTVGELG

DVRETMSYLFQHANLDSCKRVLNVVCKTCGQQQTTLKGVEAVMYMGTLSEYQFKKGVQ
IPCTCGKQATKYL VQQESPFVMMSAPPAQYELKHGTFTCASEYTGNYQCGHYKHITSK
ETLYCIDGALLTKSSEYKGPITDVIFYKENSYTTTIKPVTYKLDGVVCTEIDPKLDNYY
KKDNSYFTEQPIDLVPNQYPYNASFDNFKFVCDNIKFADDLNQLTGYKKPASRELKVT
FFPD LN GDVVAIDYKH YTPSF KKGAKLLHKPIVWHVNNATNKATYKPNTWCIRCLWST
KPVETSNSFDVLKSEDAQGMDNLACEDLKPVSEEVVENPTIQKDVLECNVKTTEVVGD
IILKPANNSLKITEEVGHTDLMAAYVDNSSLTIKKPNELSRVLGLKTLATHGLAAVNS
VPWDTIANYAKPFLNKVVSTTTNIVTRCLNRVCTNYMPYFFTLLLQLCTFTRSTNSRI
KASMP TTI AKNTVKS V GKFCLEASFNYL KSPNFSKLINIIWFLLLSVCLGSLIYSTA
ALGVLMSNLGMPSYCTGYREGYLNSTNVTIATYCTGSIPCSVCLSGLDSDTYP SLET
IQITISSFKWDLTAFGLVAEWFLAYILFTRFFYVLGLAAIMQLFFSYFAVHFISNSWL
MWLIINLVQMAPISAMVRMYIFFASFYVWKS YVHVVDGCNSSTCMMC YKRN RATRVE
CTTIVNGVRRSFYVYANGGKGFKLHNWNCVNCDTFCAGSTFISDEVARDLSLQFKRP
INPTDQSSYIVDSVTVKNGSIHLYFDKAGQKTYERHSLSHFVNLDNLRANNTKGSLPI
NVIVFDGKSKCEESSAKSASVYYSQLMCQPILLDQALVSDVGDSA EVAVKMF DAYVN
TFSSTFNVPMEKLKTLVATAEAE LAKNVSLDNVLSTFISAARQGFVDS DVETKD VVEC
LKL SHQSDIEVTGDSCNNYMLTYNKVENMTPRDLGACIDCSARHINAQVAKSHNIALI
WNVKDFMSLSEQLRKQIRSAAKKNNLPFKLTCATTRQVVNVVTTKIALKGKIVNNWL
KQLIKVTLVFLFVAAIFYLITPVHVMSKHTDFSSEIIGYKAIDGGVTRDIASTDT CFA
NKHADFDTWFSQRGGSYTNDKACPLIAAVITREVG FVVPGLPGTILRTTNGDFLHFLP
RVFSAVGNICYTPSKLIEYTD FATSACVLAAECTIFKDASGKVPYCYDTNVLEGSVA
YESLRPDTRYVLMDGSIIQFPNTYLEG SVRVVTTFDSEYCRHGTCERSEAGVCVSTSG
RWVLNNDY YRSLPGVFCGVDAVNLLTNMFTPLIQIGALDISASIVAGGIVAIVVTCL
A Y YFMRFRRAFGEYSHVVAFN TLLFLMSFTVLCLTPVYSFLPGVYSVIYLYLTFYLTN
DVSFLAHIQWMVMFTPLVPFWITIA YIICISTKH FYWFFSNY LKRRVVFNGVSFSTFE
EAALCTFLLNKEMYLKLRSDVLLPLTQYNRYLALYNKYKYFSGAMDTTSYREAAACCHL
AKALNDFSNSGSDVLYQPPQTSITS AVLQSGFRKMAFPSGKVEGCMVQVTCGTTTLNG
LWLDDVVYCPRHVICTSEDMLNPNYEDLLIRKSNHNFLVQAGNVQLRVIGHSMQNCVL

KLKVDTANPKTPKYKFVRIQPGQTFSVLACYNGSPSGVYQCAMRPNFTIKGSFLNGSC
GSVGFNIDYDCVSFCYMHMELPTGVHAGTDLEGNFYGPFVDRQTAQAAGTDTTITVN
VLAWLYAAVINGDRWFLNRFTTTLNDFNLVAMKYNYEPLTQDHVDILGPLSAQTGIAV
LDMCASLKELLQNGMNGRITILGSALLEDEFTPFDVVRQCSGVTFQSAVKRTIKGTHHW
LLLILTSLVLVQSTQWSLFFFLYENAFLPFAMGIIAMSAFAMMFVKHKHAFLCLFL
LPSLATVAYFNMVYMPASWVMRIMTWLDMVDTSLSGFKLKDCVMYASAVVLLILMTAR
TVYDDGARRVWTLMNVLTLVYKVYYGNALDQAISMWALIISVTSNYSGVVTTVMFLAR
GIVFMCVEYCPIFFITGNTLQCIMLVYCFLGYFCTCYFGLFCLLNRYFRLTLGVYDYL
VSTQEFRYMNSQGLLPPKNSIDAFKLNKLLGVGGKPCIKVATVQSKMSDVKCTSVVL
LSVLQQLRVESSSKLWAQCVQLHNDILLAKDTTEAFEKMSVLLSVLLSMQGAVDINKL
CEEMLDNRATLQAIASEFSSLPSYAAFATAQEAYEQAVANGDSEVVLLKKLKKSLNVAK
SEFDRDAAMQRKLEKMADQAMTQMYKQARSEDKRAKVTSAMQTMFTMLRKLDNDALN
NIINNARDGCVPLNIPLTTAAKLMVIPDYNTYKNTCDGTTFTYASALWEIQQVVDA
DSKIVQLSEISMDNSPNLAWPLIVTALRANSVVKLQNNELSPVALRQMSCAAGTTQTA
CTDDNALAYYNTTKGGRFVLALLSDLQDLKWARFPKSDGTGTIYTELEPPCRFVTDTP
KGPKVKYLYFIKGLNNLNRGMVLGSLAATVRLQAGNATEVPANSTVLSFCAFAVDAAK
AYKDYLASGGQPITNCVKMLCTHTGTGQAITVTPEANMDQESFGGASCCLYCRCHIDH
PNPKGFCDLKGKYVQIPTTCANDPVGFTLKNTVCTVCGMWKGYGCSCDQLREPMLQSA
DAQSFLNRVCGVSAARLTPCGTGTSTDVVYRAFDIYNDKVAGFAKFLKTNCCRFQEKD
EDDNLIDSYFVVKRHTFSNYQHEETIYNLLKDCPAVAKHDDFFKFRIDGDMVPHISRQR
LTKYTMADLVYALRHFDEGNCDTLKEILVTYNCCDDDYFNKKDWYDFVENPDILRVYA
NLGERVRQALLKTVQFCDAMRNAGIVGVLTLDNQDLNGNWYDFGDFIQTTPGSGVPVV
DSYYSLMPILTLTRALTAESHVDTDLTKPYIKWDLKYDFTEERLKLFDYFKYWDQ
TYHPNCVNCLDDRCILHCANFNVLFSTVFPPTSFGPLVRKIFVDGVPFVVSTGYHFRE
LGVVHNQDVNLHSSRLSFKELLVYAADPAMHAASGNLLLDKRTTCFSVAALTNNVAFQ
TVKPGNFNKDFYDFAVSKGFFKEGSSVELKHFFFAQDGNAAISDYDYRYNLPTMCDI
RQLLFVVEVVDKYFDCYDGGCINANQVIVNNLDKSAGFPFNKWGKARLYYDSMSYEDQ
DALFAYTKRNVIPITITQMNLYAISAKNRARTVAGVSICSTMTNRQFHQKLLKSIAAT

RGATVVIGTSKFYGGWHNMLKTVYSDVENPHLMGWDYPKCDRAMPNMLRIMASLVLAR
KHTTCCSLSHRFYRLANECAQVLSEMVMCGGSLYVKPGGTSSGDATTAYANSVFNICQ
AVTANVNALLSTDGNKIADKYVRNLQHRLYECLYRNRDVTDFVNEFYAYLRKHFSMM
ILSDDAVVCFNSTYASQGLVASIKNFKSVLYYQNNVFMSEAKCWTETDLTKGPHEFCS
QHTMLVKQGDDYVYLPYPDPSRILGAGCFVDDIVKTDGTLMIERFVSLAIDAYPLTKH
PNQEYADVHLYLQYIRKLHDELTHGMLDMYSVMLTNDNTSRYWEPEFYEAMYTPHTV
LQAVGACVLCNSQTSLRGACIRRPFLCCKCCYDHVISTSHKLVLSVNPYVCNAPGCD
VTDVTQLYLGGMSYYCKSHKPPISFPLCANGQVFGLYKNTCVGSDNVTDFNAIATCDW
TNAGDYILANTCTERLKLFAAETLKATEETFKLSYGIATVREVLSRELHLSWEVGKP
RPPLNRNYVFTGYRVTKNSKVQIGEYTFEKGDYGDVAVYRGTTTTYKLVNGDYFVLTS
H
TVMPLSAPTLVPQEHYVRITGLYPTLNISDEFSSNVANYQKVGMQKYSTLQGPPTGK
SHFAIGLALYYPSARIVYTACSHAAVDALCEKALKYLPIDKCSRIIPARARVECFDKF
KVNSTLEQYVFCTVNALPETTADIVVFDEISMATNYDLSVVNARLRAKHVYIGDPAQ
LPAPRTLLTKGTLEPEYFNSVCRLMKTIGPDMFLGTCRRCPAEIVDTVSALVYDNKLG
AHKDKSAQCFKMFYKGVITHDVSSAINRPQIGVVREFLTRNPAWRKAVFISPYNSQNA
VASKILGLPTQTVDSSQGSEYDYVIFTQTTETAHSCNVNRFNVAITRAKVILCIMS
RDLYDKLQFTSLEIPRRNVATLQAENVGTGLFKDCSKVITGLHPTQAPTHLSVDTKFKT
EGLCVDIPGIPKDMTYRRLISMMGFKMNYQVNGYPNMFITREEAIRHVRAWIGFDVEG
CHATREAVGTNLPLQLGFSTGVNLVAVPTGYVDTPNNTDFSRVSAKPPPGDQFKHLIP
LMYKGLPWNVVRKIVQMLSDTLKNLSDRVVFVLWAHGFELTSMKYFVKIGPERTCCL
CDRRATCFSTASDTYACWHHSIGFDYVYNPFMIDVQQWGFTGNLQSNHDLYCQVHGNA
HVASCDAIMTRCLAVHECFVKRVDWTIEYPIIGDELKINAACRQVQHMVKAALLADK
FPVLHDIGNPKAIKCVQADVEWKFYDAQPCSDKAYKIEELFYATHSDKFTDGVCL
FWNCNVDRYPANSIVCRFDTRVLSNLNLPGLDGGSLYVKNHAFHTPAFDKSAFVNLKQ
LPFFYYSDSPCESHGKQVVSIDYVPLKSATCITRCNLGGAVCRHHANEYRLYLDAYN
MMISAGFSLWVYKQFDTYNLWNTFTRLQSLNVAFNVVNKGHFDGQQGEVPVSIINNT
VYTKVDGVDVELFENKTTLPVNVAFELWAKRNIKPVPEVKILNNLGVDAANTVIWDY
KRDAPAHISTIGVCSMTDIAKKPTETICAPLTVFFDGRVDGQVDLFRNARNGVLITEG

SVKGLQPSVGPQASLNGVTLIGEAVKTQFNYYKKVDGVVQQLPETYFTQSRNLQEFK
PRSQMEIDFLELAMDEFIERYKLEGYAFEHIVYGDFSHSQLGGLHLLIGLAKRFKESP
FELEDFIPMDSTVKNYFITDAQTGSSKVCVCSVIDLLLDDFVEIIKSQDLSVVSKVVKV
TIDYTEISFMLWCKDGHVETFYPKLQSSQAWQPGVAMPNLYKMQRMLLEKCDLQNYGD
SATLPKGIMMNVAKYTQLCQYLNTLTLAVPYNMRVIHFGAGSDKGVAPGTAVLRQWLP
TGTLVDSDLNDFVSDADSTLIGDCATVHTANKWDLISDMYDPKTKNVTKENDSKEG
FFTYICGFIQQLALGGSVAIKITEHSWNADLYKLMGHFAWWTAFVTNVNASSSEAF
IGCNYLKGKPREQIDGYVMHANYIFWRNTNPIQLSSYSLFDMSKFPLKLRGTAVMSLKE
GQINDMILSLLSKGRLIIRENNRVVISSDVLVNN"

gene 21563..25384

/gene="S"

CDS 21563..25384

/gene="S"

/note="structural protein"

/codon_start=1

/product="surface glycoprotein"

/protein_id="QHD43416.1"

/translation="MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFR

SSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIR

GWIFGTTLDSKTQSLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVY

SSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPO

GFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRFTL

LKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITN

LCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCF

TNVYADSFVIRGDEVQRQIAPGQTGKIADYNYKLDDFTGCVIAWNSNNLDSKVGGNYN

YLYRLFRKSNLKPFERDISTEIQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPY

RVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFG

RDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAI

HADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSECDIPIGAGICASYQTQTNSPR
RARSVASQSIIAYTMSLGAENSVAYSNNNSIAIPTNFTISVTTEILPVSMTKTSVDCTM
YICGDSTECNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFG
GFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFN
GLTVLPPLLTDemiaQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQN
VLYENQKLIANQFNsAIGKIQDSLSTASALGKLQDVVNQNAQALNTLVKQLSSNFGA
ISSVLNDILSRDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMS
ECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAH
FPREGVFVSNGTHWFVTQRNFYEPQIITDNTFVSGNCDVVIGIVNNTVYDPLQPELD
SFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELG
KYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCCLKGCCSCGSCCKFDEDDSE
PVLKGVKLHYT"

gene 25393..26220

/gene="ORF3a"

CDS 25393..26220

/gene="ORF3a"

/codon_start=1

/product="ORF3a protein"

/protein_id="QHD43417.1"

/translation="MDLFMRIFTIGTVTLKQGEIKDATPSDFVRATATIPIQASLPFG

WLIVGVALLAVFQSASKIITLKKRWQLALSKGVHFVCNLLLLFVTVYSHLLLVAAGLE

APFLYLYALVYFLQSINFVRIIMRLWLCWKCRSKNPLLYDANYFLCWHTNCYDYCIPY

NSVTSSIVITSGDGTTSPISEHDYQIGGYTEKWESGVKDCVVLHSYFTSDYYQLYSTQ

LSTDTGVEHVTFFIYNKIVDEPEEHVQIHTIDGSSGVVNPVMEPIYDEPTTTTTSVPL"

gene 26245..26472

/gene="E"

CDS 26245..26472

/gene="E"

/note="structural protein; E protein"

/codon_start=1

/product="envelope protein"

/protein_id="QHD43418.1"

/translation="MYSFVSEETGTLIVNSVLLFLAFVVFLVTLAILTALRLCAYCC

NIVNVSLVKPSFYVYSRVKNLNSSRVPDLLV"

gene 26523..27191

/gene="M"

CDS 26523..27191

/gene="M"

/note="structural protein"

/codon_start=1

/product="membrane glycoprotein"

/protein_id="QHD43419.1"

/translation="MADSNGTITVEELKKLLEQWNLVIGFLFTWICLLQFAYANRNR

FLYIIKLIFLWLLWPVTLACFVLAADVIRINWITGGIAIAMACLVGLMWLSYFIASFRL

FARTRSMWSFNPETNILLNVPLHGTILTRPLLESELVIGAVILRGHLRIAGHHLGRCD

IKDLPKEITVATSRTLSYYKLGASQRVAGDSGFAAYSRYRIGNYKLNTHSSSSDNIA

LLVQ"

gene 27202..27387

/gene="ORF6"

CDS 27202..27387

/gene="ORF6"

/codon_start=1

/product="ORF6 protein"

/protein_id="QHD43420.1"

/translation="MFHLVDFQVTIAEILLIIMRTFKVSIWNLDYIINLIKNLSKSL

TENKYSQLDEEQPMEID"

gene 27394..27759
 /gene="ORF7a"

CDS 27394..27759
 /gene="ORF7a"
 /codon_start=1
 /product="ORF7a protein"
 /protein_id="QHD43421.1"
 /translation="MKIILFLALITLATCELYHYQECVRGTTVLLKEPCSSGTYEGNS
 PFHPLADNKFALTCFSTQFAFACPDGVKHHVYQLRARSVSPKLFIRQEEVQELYSPIFL
 IVAAIVFITLCFTLKRKTE"

gene 27894..28259
 /gene="ORF8"

CDS 27894..28259
 /gene="ORF8"
 /codon_start=1
 /product="ORF8 protein"
 /protein_id="QHD43422.1"
 /translation="MKFLVFLGIITTVA AFHQECSLQSQCTQHQPYYVDDPCPIHFYSK
 WYIRVGARKSAPLIELCVDEAGSKSPIQYIDIGNYTVSCLPFTINCQEPKLGSLVVRC
 SFYEDFLEYHDVRVVLDFI"

gene 28274..29533
 /gene="N"

CDS 28274..29533
 /gene="N"
 /note="structural protein"
 /codon_start=1
 /product="nucleocapsid phosphoprotein"
 /protein_id="QHD43423.2"

/translation="MSDNGPQNQRNAPRITFGGPSDSTGSNQNGERSGARSKQRRPQG
LPNNTASWFTALTQHGKEDLKFPRGQGVPIINTNSSPDDQIGYYRRATRRIRGGDGKMK
DLSPRWYFYYLGTGPEAGLPYGANKDGIIWVATEGALNTPKDHIGTRNPANNAIVLQ
LPQGTTLPGKFYAEGSRGGSQASSRSSSRNSSRNSTPGSSRGTS ParmagnggdAA
LALLLLDRLNQLESKMSGKGQQQGGQTVTKKSAAEASKKPRQKRTATKAYNVTQAFGR
RGPEQTQGNFGDQELIRQGTDYKHWPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTYT
GAIKLDDKDPNFKDQVILLNKHIDAYKTFPPTPEPKDKKKKKADETQALPQRQKKQQTV
TLLPAADLDDFSKQLQQSMSSADSTQA"

gene 29558..29674

/gene="ORF10"

CDS 29558..29674

/gene="ORF10"

/codon_start=1

/product="ORF10 protein"

/protein_id="QHI42199.1"

/translation="MGYINVFAFPFTIYSLLLCRMNSRNYIAQVDVVNFNLT"

3'UTR 29675..29903