

Prediction of Proteins Associated with COVID-19 Based Ligand Designing and Molecular Modeling

Majid Monajjemi^{1,*}, Rahim Esmkhani², Fatemeh Mollaamin¹ and Sara Shahriari³

¹Department of Chemical Engineering, Central Tehran Branch, Islamic Azad University, Tehran, 14987-54891, Iran

²Department of Chemistry, Khoy Branch, Islamic Azad University, Khoy, 58159-88838, Iran

³Department of Chemistry, Central Tehran Branch, Islamic Azad University, Tehran, 1498754891, Iran

*Corresponding Author: Majid Monajjemi. Email: maj.monajjemi@iauctb.ac.ir

Received: 14 July 2020; Accepted: 29 October 2020

Abstract: Current understanding about how the virus that causes COVID-19 spreads is largely based on what is known about similar coronaviruses. Some of the Natural products are suitable drugs against SARS-CoV-2 main protease. For recognizing a strong inhibitor, we have accomplished docking studies on the major virus protease with 4 natural product species as anti COVID-19 (SARS-CoV-2), namely “Vidarabine”, “Cytarabine”, “Gemcitabine” and “Matrine” which have been extracted from Gillan’s leaves plants. These are known as Chuchaq, Trshvash, Cote-Couto and Khlvash in Iran. Among these four studied compounds, Cytarabine appears as a suitable compound with high effectiveness inhibitors to this protease. Finally by this work we present a method on the Computational Prediction of Protein Structure Associated with COVID-19 Based Ligand Design and Molecular Modeling. By this investigation, auto dock software (iGEM-DOCK) has been used and via this tool, the suitable receptors can be distinguished in whole COVID-19 component structures for forming a complex. “iGEMDOCK” is suitable to define the binding site quickly. With docking simulation and NMR investigation, we have demonstrated these compounds exhibit a suitable binding energy around 9 Kcal/mol with various ligand proteins modes in the binding to COVID-19 viruses. However, these data need further evaluation for repurposing these drugs against COVID-19 viruses, in both vivo & vitro.

Keywords: COVID-19; receptor binding domain; Gillan’s leaves plants; angiotensin converting enzyme-2; protease domain

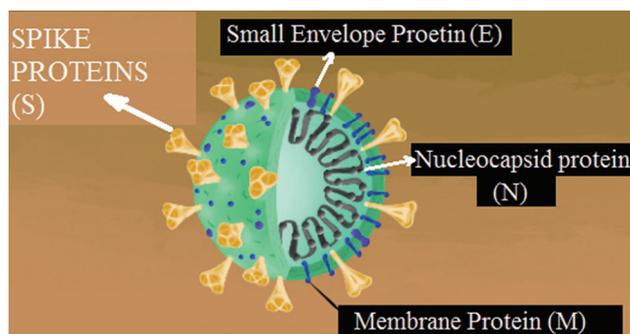
1 Introduction

One decade after SARS virus, the “Middle East Respiratory Syndromes” named MERS virus groups have infected human. The Mouse Hepatitis disease which is approximately depending to SARS and MERS corona virus, has long served as a model of study. This virus infects several of human host’s cells and also sometimes animal, that can be also, carry out their infection and replication. In addition, several proteins have a basic role in this replication mechanism. In these cases, there are necessary for understanding the definition of those proteins in view point



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

of related mechanisms [1]. These proteins consist of transcription/replication combinations of RNA, several proteins, and two structures of proteases. Those proteases represent major roles to incision the polyproteins from all of the functional components. The main part proteases of these viruses make most of these cuts [2–4]. The SARS-CoV-2 is currently posing the most risks of a dimer protein same as the serine proteases including Trypsin, Cysteine and Methionine amino acids. This dimerization has peptides-analogous inhibitor bound in center of active sites. These couple proteases in the SARS are the basic proteases that consist of several splits at 11 sites in the polyproteins. The codon of RNA in the COVID-19 is a positive layer and related units contain: Spike protein or (S), envelope protein or (E) and membrane structure (M) with nucleocapsid phosphoprotein [5–7] (Scheme 1). Transcribed non-structural proteins consist of: NSP1, NSP4, ORF1 (ab), ORF3 (a), ORF6, ORF7 (a), ORF9, ORF8 and ORF10 (Tab. 1). Although researchers have discovered the structure of COVID-19 proteins and also some non-structural components, the COVID-19 has strong genetic potential characteristics that a few of them are basically the cause of human death [8]. As instance, for the envelope protein ion channel can be signified in modulating virion release and COVID-19-host interaction [9]. Moreover, Spike, ORF3a and ORF8 proteins are completely different from other SARS-like proteins. Recently researchers indicated the mechanism of the COVID-19 diffusion into the epithelial cells via the s-protein. This interaction with the ACE2 receptors causes human disease. However structural evaluation of the S-protein from the COVID-19 and S-protein is due to the weakly binds of the ACE2 receptors. In addition, due to a lack of important experimental method the mechanism of viral proteins (Tab. 1) are still unknown [10]. Since the corona virus pandemic evolves, researchers have been competed for studding and understanding more about COVID-19. A few months ago in March, a novel study indicated that the S-protein interacted with the ACE2 receptor from COVID-19 and the reaction can be terminated via the enzyme TMPRSS2 [11] (Tab. 1).



Scheme 1: Spike protein (S), envelope protein (E), membrane protein (M), and nuclei-capsid phosphoprotein

Further research discovered 330 human proteins that interacted with the proteins of COVID-19 in that 68 of these proteins can be applied as targets for suitable drugs [12] (Tab. 1). SARS-CoV-2 strain is about 75% similar to SARS-COV strain and 35% close to the MERS-COV strain. Effects of COVID-19 can be occurring within 3 days or sometimes up to two weeks after exposure and there is an equal genomic system known as of β -coronaviruses containing ORF1 (ab), NSP1, S-protein from 5'-UTR (untranslated region) and E-protein, M-protein from 3'-UTR and also ORF6 (a), ORF7 (a), ORF (8), N-protein, ORF (10), and several other non-structural items.

Table 1: Transcribed non-structural proteins consist of: NSP1, NSP4, ORF1ab, ORF3a, ORF6, ORF7a, ORF9, ORF8 and ORF10

Human genome	COVID-19 genome	Biological system	Drug target	Protein system	Performance
PRIM1	Nsp1	DNA replication		DNA polymerase (α)	Induce host mRNA cleavage
PRIM2	Nsp1	DNA replication		DNA polymerase (α)	
POLA1	Nsp1	DNA replication		DNA polymerase (α)	Binds to PHBs 1, 2
POLA2	Nsp1	DNA replication		DNA polymerase (α)	
COLGALT1	Nsp1	DNA replication		DNA polymerase (α)	
PKP2	Nsp1	DNA replication		DNA polymerase (α)	
POLA1	Nsp1	DNA replication		DNA polymerase (α)	
POLA2	Nsp1	DNA replication		DNA polymerase (α)	
COLGALT1	Nsp1	DNA replication		DNA polymerase (α)	
PKP2	Nsp1	DNA replication		DNA polymerase (α)	
POLA1	Nsp1	DNA replication		DNA polymerase (α)	
POLA2	Nsp1	DNA replication		DNA polymerase (α)	
TIMM10	Nsp4	Mitochondria		TIM complex	Membrane rearrangement
TIMM10B	Nsp4	Mitochondria		TIM complex	
TIMM9	Nsp4	Mitochondria		TIM complex	
TIMM29	Nsp4	Mitochondria		TIM complex	
ACAD9	Orf9c	Mitochondria			SARS-CoV2 ORF9c, SARS-CoV-2 ORF9c, 9c
FAR2	Orf9c	Mitochondria			
WFS1	Orf9c	Mitochondria			
PIGO	Orf9c	Mitochondria			
RETREG3	Orf9c	Mitochondria			
UBXN8	Orf9c	Mitochondria			
NLRX1	Orf9c	Mitochondria			
TMEM97	Orf9c	Mitochondria	Potential drug candidate		
ERMP1	Orf9c	Mitochondria			
TAPT1	Orf9c	Mitochondria			
SLC30A6	Orf9c	Mitochondria			
TMED5	Orf9c	Mitochondria			
SCAP	Orf9c	Mitochondria			
BCS1L	Orf9c	Mitochondria			
NDFIP2	Orf9c	Mitochondria			
DPY19L1	Orf9c	Mitochondria			
F2RL1	Orf9c	Mitochondria	Potential drug candidate		
GHITM	Orf9c	Mitochondria			
ABCC1	Orf9c	Mitochondria	Potential drug candidate		
TMEM39B	Orf9c	Mitochondria			
ALG8	Orf9c	Mitochondria			
FBXL12	Orf8	Vesicle trafficking			
POGLUT3	Orf8	Vesicle trafficking			
PLEKHF2	Orf8	Vesicle trafficking			

(Continued)

Table 1: (Continued)

Human genome	COVID-19 genome	Biological system	Drug target	Protein system	Performance
CISD3	Orf8	Vesicle trafficking			
INHBE	Orf8	Vesicle trafficking			
GDF15	Orf8	Vesicle trafficking			
SMOC1	Orf8	Vesicle trafficking			
NEU1	Orf8	Vesicle trafficking			
PLAT	Orf8	Vesicle trafficking			
POGLUT2	Orf8	Vesicle trafficking			SARS-CoV2 ORF8, SARS-CoV-2 ORF8, 8, NS7B_SARS2, PRO_0000 449655
STC2	Orf8	Vesicle trafficking			

Biochemical indexes of 100 patients with COVID-19 indicate an abnormal phenomenon due to hemoglobin-related biochemical phenomenon. By this disease, the neutrophil and hemoglobin amounts have reduced while the indexes of erythrocyte sedimentation rates, serum ferritins, C-reactive proteins, albumins, and lactate dehydrogenases increase significantly. This phenomenon indicates that by decreasing hemoglobin, HEM increases and too much harmful iron will accumulate in the body that will cause smart pain in the human and increase the albumin. Consequently cells produce, large scale of serum ferritin for reducing injury. Hemoglobin contains 4 subunits, 2-alpha and 2-beta, in which each subunit has iron- hemoglobin bonded [13].

Due to Fe^{2+} the CO_2 gases might be separated from hemoglobin and consequently capture O atoms and consequently Fe^{2+} is oxidized to Fe^{3+} . Hemoglobin can release oxygen atoms and capture CO_2 and then Fe^{3+} is reduced to Fe^{2+} . Because of no clinical trial-based vaccine present, vaccinations are a reliable effective way. So availability of genomic information and biological algorithms including immunological information might be help researchers for identifying the effective epitopes that can be used for developing vaccines [14,15]. The subunit vaccine consists of several segments of antigen which can be simulated the presence of the natural pathogen towards an immune situation of pathogen [15,16]. Basically, vaccinations designing against MERS, SARS, Ebola, ZIKA and Chikungunya viruses have produced promising results [17,18].

The modeling and simulation methods with genetic algorithms decrease the number of experimental activities and save expenditure and also increase the potential of a successful vaccination. A large amount of peptides involved in the multi-epitopes vaccine in which induce the activation of AIR (Adaptive Immune Response) are suitable and fast method for the viral-infections therapy and COVID-19 [19,20]. By this work, SARS-CoV-2 proteomes were explored for determining the antigenic proteins and various B&T-cells epitopes were foreteller with their main histocompatibility complexes characteristic. In addition conserved domain, homologies simulation and molecular docking methods were applied for analyzing the structure of virus-depended proteins.

This investigation exhibited that ORF9b and surfaces glycoproteins had a segment for combining with porphyrin to make the complexes structures, while ORF1 (ab), ORF10, ORF (3a) coordinately attack the hemoglobin on the $1 - \beta$ chain for dissociating the iron to form the porphyrin. Moreover the amino acid sequences of target proteins (ORF1, S-(protein), ORF3 (a), E-protein, M-protein, ORF6, ORF (7a), ORF8, ORF10 and ORF9b of SARS-CoV-2 were

extracted from the Gen-bank. This phenomenon of the coronavirus inhibited the usual metabolic pathways of hemoglobin, and made human show symptoms of these problems.

1.1 The Specified Model in Math for the Ligand Designing

For better modeling. Specifically, soft computing techniques from engineering sciences such as based on heuristic algorithms have been applied by this work [20,21]. The basics of MD simulations such as Monte Carlo sampling can be explained via considering a few major options. The idea of *energy surfaces* has a physical meaning due to conformational sampling of the energy landscape which is often has a computational approaches in both MD and docking. For proteins, there are 3 aspects of any dynamical process for considering which are included as: *First*, the *timescale* of the elementary process; *Second*, the *spatial extent* over which the event occurs; and *third* the *amplitude* of motion. The functional dynamics of bio macromolecules modulate its intra- and inter-macromolecular interactions and are of great physiological importance. via an MD *trajectory* we mean a list of positions and momenta of each particle in a system over time, as the system samples its phase space. These trajectories by approximating the equations of motion via numerical integration can be illustrated by the instantaneous force acting on each particle which is calculated by the formula as

$$F_i = -\nabla\mu_i, \quad (1)$$

where the forces are used to compute accelerations, and the accelerations are used to update particle velocities and positions. A force-field (FF) encapsulates all that we believe to be important about the physicochemical properties of the atomic interactions that govern molecular structure & dynamics. In this work the FF expresses molecular interactions quantitatively, using equations, free parameters, and estimates of parameter values. Those efforts were aimed at calculating primarily structural and stereo-chemical properties of small organic molecules conformational strain, geometry optimization, etc. In principle, computing the FF energy as a function of 3D structure, for all possible 3D conformations, would provide the complete potential energy surface of a molecule. In this classic MM approach, covalent interactions are taken as summations over several bonded terms, while non-bonded interactions are modeled pairwise, as sums over Lennard-Jones and Coulombic potentials as follows [22–24]:

$$U(r_i) = \sum_{bond} K_i^r (r - r_0)^2 + \sum_{angle} K_i^\theta (\theta - \theta_0)^2 + \sum_{torsions} K_i^\phi [1 + \text{Cos}(n_i\varphi_i - \delta_i) \\ + \sum_i \sum_{i \neq j} 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^6 \right] + \sum_i \sum_{i \neq j} \frac{q_i q_j}{\varepsilon r_{ij}} \quad (2)$$

The FF parameters, which may number well into the hundreds, list all the spring constants (k), reference bond lengths (r_0), and angles (θ_0), torsional angles (φ_i), multiplicities (n), and phases (δ_i), Lennard-Jones parameters (ε, δ), and partial charges (q). While bond lengths and angles are handled in a fairly straightforward and similar manner in different FFs, various biomolecular FFs treat torsional potentials and other terms in subtly different ways. For instance, AMBER and OPLS use specific scaling factors for vdW or electrostatic interactions between 1–4 atoms and some CHARMM FFs employ grid-based energy correction maps. Simulation analyses can range from routine and straightforward to highly sophisticated, and can be either highly generic or more specialized to the type of system/question at hand. An example of a generic type of trajectory analysis, applicable to any system, is computation of the root-mean-square deviation (RMSD) of coordinates over time. Though the RMSD is not always an ideal metric for assessing

equilibration and structural stability, an RMSD analysis is performed early on (within the first few ns) in virtually all atomistic simulation studies. The RMSD for two coordinate sets, S_x and S_y is readily defined as [25]:

$$RMSD = \sqrt{\frac{\sum_{I=1}^N [r_{s_x} - r_{s_y}]^2}{N}}$$

1.2 Research Significance for Any Further Research

Research priorities for the COVID-19 pandemic and beyond: a call to action for computer modeling in engineering science can be concluded as:

How do we increase adherence (and ability to adhere) to countries COVID-19-related instructions?

How do we promote maintenance of positive behavior changes and reverse negative behavior changes resulting from COVID-19-related lockdown?

How do we address the negative psychological impacts of the COVID-19 pandemic?

How do we maximize recovery from COVID-19 for those infected with the virus?

What is the impact of COVID-19-related stress on biological processes and health outcomes?

What makes people adhere to anti-COVID measures?

What are the bases of anti-social behaviors such as stockpiling?

How do mutual aid groups form and what makes them endure?

When does social cohesion give way to scapegoating, prejudice, and intergroup conflict?

What creates (or prevents) the potential for protests and collective disorder in the crisis?

What are the long-term mental health effects of COVID-19?

What coping mechanisms are useful in reducing mental health problems during a pandemic?

How do we provide beneficial remote psychological therapy and maintain therapeutic alliance?

Has discussion of mental health during the pandemic reduced stigma and discrimination in the community?

People detained in hospital under the Mental Health Act were discharged to free up beds—how was this possible?

What are the impacts of COVID-19 infection, treatment, and recovery on cognition, behavior, and the brain?

What are the drivers of COVID-19-related stress and its cognitive, neural, and physiological mechanisms and consequences?

What are the perceptual and cognitive demands of digital and other alternative forms of communication and how do they impact on work and social connectivity?

What factors influence the effectiveness of communication of scientific evidence and national guidance, and how do they influence behavior?

How do restrictions of movement, communication, and social support influence the cognitive, physical, and mental health of older individuals, and what factors lead to improved outcomes?

How has the COVID-19 pandemic affected parenting?

How has the COVID-19 pandemic affected children's development?

How has the COVID-19 pandemic affected family functioning?

Which factors moderate family members' response to the COVID-19 pandemic?

What support is most effective for families during the COVID-19 pandemic?

How do we assess biological markers of health and well-being remotely?

How can we use biological markers to facilitate people's return to work?

How do we link COVID-19-related biomarkers to existing population cohort databases?

How do we address the negative biological impacts of the COVID-19 virus on mental health?

What are the impacts of COVID-19 infection, treatment, and recovery on the brain?

How do school closures influence educational progress, and physical and mental health outcomes for all children and young people?

What 'homeschooling' practices are associated with positive educational and psychological outcomes?

What is the effect of social distancing on a range of social outcomes in children and young people?

What methods are used to track, monitor, and deliver local authority support services to vulnerable children and young people, families, and schools during lockdown, at transition back to school, and after return to school?

How are educational and psychological interventions allocated, structured, delivered, and evaluated for children and young people in need, after schools have reopened?

What is the impact of remote and flexible working arrangements on employee health, mental well-being, teamwork, performance, organizational productivity, and colleague/client relationships?

What is the impact of social distancing in the workplace on employee health, mental well-being, teamwork, performance, organizational productivity, and colleague/client relationships?

What managerial behaviors are most effective to manage remote working, possible mental health issues, job insecurity, and productivity?

What is the risk of longer-term mental ill health among frontline staff after the immediate crisis?

How can organizational resilience be developed to deal with the impact of COVID-19 whilst supporting employees and protecting jobs?

2 Computational Details

2.1 Docking Analysis

By this investigation, auto dock software (iGEM-DOCK) has been used and via this tool, the suitable receptors can be distinguished in whole COVID-19 component structures for forming a complex. "iGEMDOCK" is suitable to define the binding site quickly. Molecular docking simulation is a way of finding the best matching pattern among molecules via geometric and energies matching. Following items have been done in docking modeling and simulation: (a) Guess for a place of binding site on the ORF series for separation the iron for forming the porphyrin. (b) Browsing and selecting the protein file and in the initial step, the sequence of polypeptides

were extracted from gen-BANK [26,27]. Finally, whole of the sequences can be saved as FASTA format for any further evaluation. (c) Designing the binding sites as a bonded ligand. (d) Designing and preparing the center of the active site via proper ligands. (e) Setting the binding and active sites in view point of size and atoms by the extended radius from the choices ligands—proteins complex. These sequences were extracted from NCBI and also all COVID-19 proteins. (f) Hemoglobin-binding proteins, Hemoglobin-oxidase and peptide sequences were used for analyzing conserved chain. It is notable that some of the polypeptides of novel COVID-19 were applied for constructing 3-D structures via homology simulation. IGMDOCK produces an analysis environment including visualization with analysis tools for applicants that can be utilized the docked situation and related category via the system. The energies poses of each items would be outputted based on the position of “best: Pose.” That information is premeditated with analyzing of these poses. Through looking at the peptide bonded of various ligands, they might be selected through the checked box of ligands. If the co-crystallization ligands are saved on the active site structures, it will be predicted poses. Cluster analyses are the partitioning of the information set into subsets. This information in each subset will share some general aspects. Interaction specifications can be extracted via the protein-ligand intercalation and are accounted atomic kinds in several functional categories. The data in each subset will share some usual properties. Finally, the PDB files were obtained from the bioinformatics sites such as PDB database web sites. By this work, the main items which have been done via molecular docking by “Discovery-Studio” that can be summarized as: Preparation of a ligand perspective, Opening the ligands file, clicking “Prepare Ligands” in the “Dock Ligands” sub-menu of the “Receptor-Ligand Interactions” menu for generating the hem ligands. For the preparing our protein receptors, the protein’s pdb file has been generated via homology modeling by the “Dock Ligands” submenu of the “Receptor-Ligand Interactions” menu for generating the protein receptor models for our docking model. We set our docking data via the generating protein receptor system from the “Define and Edit Binding Site” submenus in the “Receptor-Ligand Interactions” menu. The binding energies were calculated via choosing the pose with the largest binding energies. After finishing docking several locations of ligand have been displayed and we choose the strongest binding energy and suitable stability of those complexes for any further discussion.

2.2 Sequence Retrieval

The Expassy-Protparam tools were applied for determining the chemical & physical properties of molecules (Walker, 2005). For checking antigenicity of proteins, the related software was also applied [27,28] and for the predication of secondary polypeptides “Alignment self-optimized prediction method” (SOPMA) tools has been used [28,29]. In addition several additional online software or tools such as Swiss model [30,31], “Phyre2,” and “RaptorX” was applied for the tertiary polypeptide segment [31,32]. Model recovered was then refined via “Galaxy Refine Server” and accredited also via Ramachandran plots.

2.3 Bioinformatics Analysis

Bioinformatics analyses were accomplished based on biological papers in viewpoint of protein sequences consist of conserved domains analysis, homology simulation, molecular docking and complex evaluation. Conserved of viral proteins can be analyzed through MEME website [32,33]. These models are applied for predicting differences between the viral system and human proteins. The 3-D structures of viral proteins were structured through homology modeling of Swiss-model. Sequence length extension and homology modeling should be adopted via applying molecular docking (Studio 2016), and also the ligand-receptor of viral proteins with porphyrins might

be simulated in real situation. Consequently, a life cycle system of the virus were structured and configured within related proteins of the COVID-19 that was proposed. Molecular docking simulation can be built on homology or 3-D molecular modeling. For the analysis of mentioned domains “MEME” suite websites where integrates several keys of predicting is suitable. By the MEME, the polypeptides sequences can be merged into a text file and then select the number of motifs which is needed before any clicking of the “Go” button. For the homology modeling SWISS-MODELS are completely automatic, which can be accessed via a web system through a few running steps such as entering into the Swiss-model, writing the sequence, and clicking “Search Template” for performing a simple template search. Second step is choosing a template for modeling and third is Building Model command that a template model is selected automatically. Although all models in PDB format can be visualized by the VMD software, SWISS-MODEL is used for the sequence lengths less than 6000 nt while discovery-Studio’s homology modeling tool can be used for more lengths of 5000 nt.

2.4 M.D Simulations

Molecular dynamics modeling for polypeptide-ligand complexes were accomplished using the Desmond software. The OPLS and charm force fields were applied for modeling the protein-small molecules interactions. Long-range electrostatic forces were estimated using the Particle-mesh Ewald (PME) software with a grid spacing of 0.75 Å. Nose–Hoover thermometry and Martyna–Tobias–Klein method were applied for maintaining the temperature and constant pressure, respectively. The formula of motion was considered using the multi run RESPA by 3.0 fs time step for bonded and non-bonded interactions within a low cutoff. An outer time step of 5.0 fs was used for non-bonded forces beyond the cutoff.

2.5 Simulation of Lipid Bilayers & Membrane Protein (M)

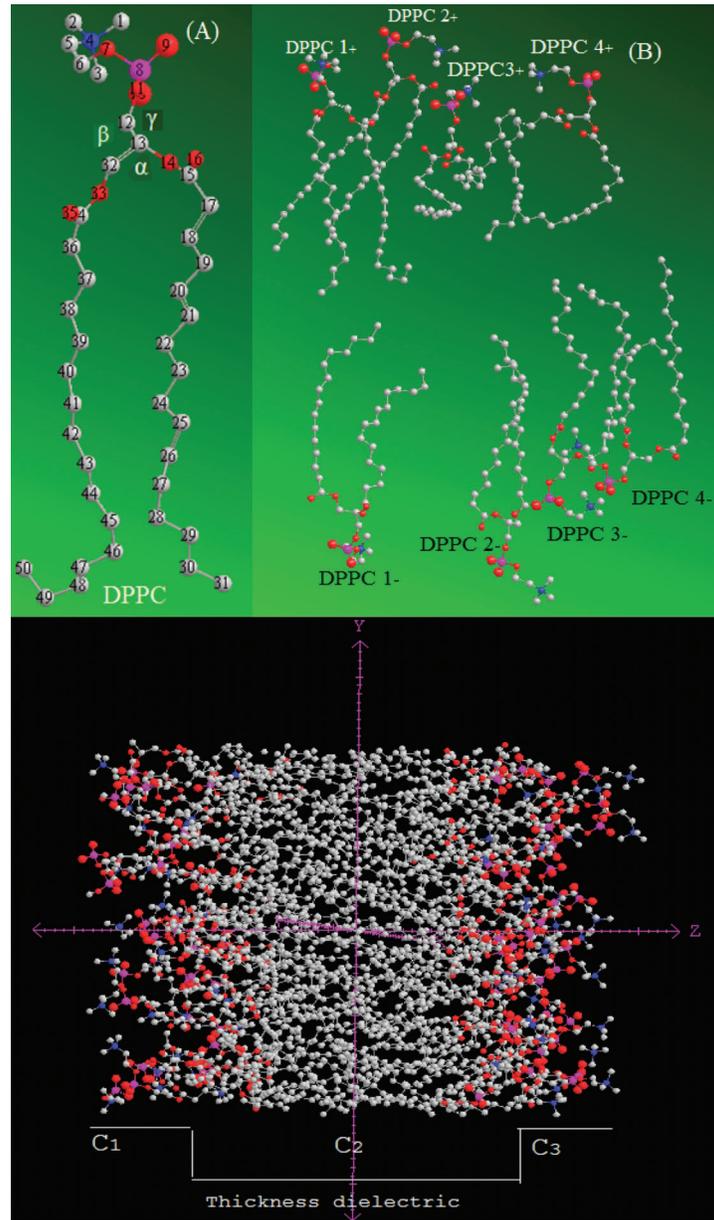
Although, accurate structures of the phosphor lipid’s bilayers which are in biochemistry pertaining fluid medium are not possible for getting experimental data, fluctuations of these types of dynamic bilayers demonstrate correct structures. Quantum mechanics and Molecular dynamics simulation (QM/MM) are strong tools for tabulating and guiding the interpretation of these experimental parts. The validity of modeling may be measured in contrast existing experimental results. There are several and various techniques such as Deuterium NMR quadrupol splitting which can give certain results of physical & chemical properties. Membrane electrostatics area per lipid and membrane thickness and acyl parameters are also important to study for any further simulation. The absence of experimental data and results are reversed in molecular modeling of membrane protein (M) due to several force fields parameterization. Tight level AB-initio estimations are needed for parameterization of these types’ force fields and presently allow evaluation of the heavy atoms for gaining accurate results. Moreover, there are some limitations in weak QM calculations due to London’s dispersion of non-bonded interactions for such molecules. We simulated our model based on our previous works [34–37] (Scheme 2) .

2.6 NMR Shielding

The anisotropy data of related parameters for shielding and non-shielding spaces of the hetero rings in all antibiotics, (σ_{11} , σ_{22} , σ_{33}), are labeled based on IUPAC instruction. Moreover, σ_{33} indicates the direction of minimum shielding, with the highest frequency, while σ_{11} exhibited a direction of maximum shielding, with the lowest one. In addition, the orientations of the

asymmetry tensors are given by $\left(\kappa = \frac{3a}{\Omega}\right)$ and the skew is

$$\kappa = \frac{3(\sigma_{1so} - \sigma_{22})}{\Omega} \quad (-1 \leq \kappa \leq +1). \quad (3)$$



Scheme 2: Optimized of DPPC and membrane simulation including 120 molecules of DPPC phospholipids

In our calculations of various halogenated antibiotic's rings, (κ) is basically positive, and the negative values are related to some critical or boundary points. The stabilities of the isotropies are tightly affected on suitable places in the shielding area spaces and are dependent on the configuration of aromatic rings. Therefore, using this method of isotropy can be calculated as an aromaticity criterion for any further simulation of COVID-19 due to hydrogen bonds between "Vidarabine," "Cytarabine," "Gemcitabine" and "Matrine" and COVID-19 proteins. Obviously that structural function cause changes in the magnetic field experienced through the nuclei and change the resonant frequencies. So, the chemical shielding and other properties such as hydrogen bonding and magnetic anisotropy of π -systems might be changed due to the electrons around the proton which produce a magnetic field, countering the applied fields. Consequently reduces the field experienced at the nucleus. In other words, electrons are said to shield the proton, an effect that is exactly dependent on the distance of the center. By this simulation we have calculated those parameters for any binding of those natural products to the COVID-19 components (Fig. 1 and Tabs. 2, 3).

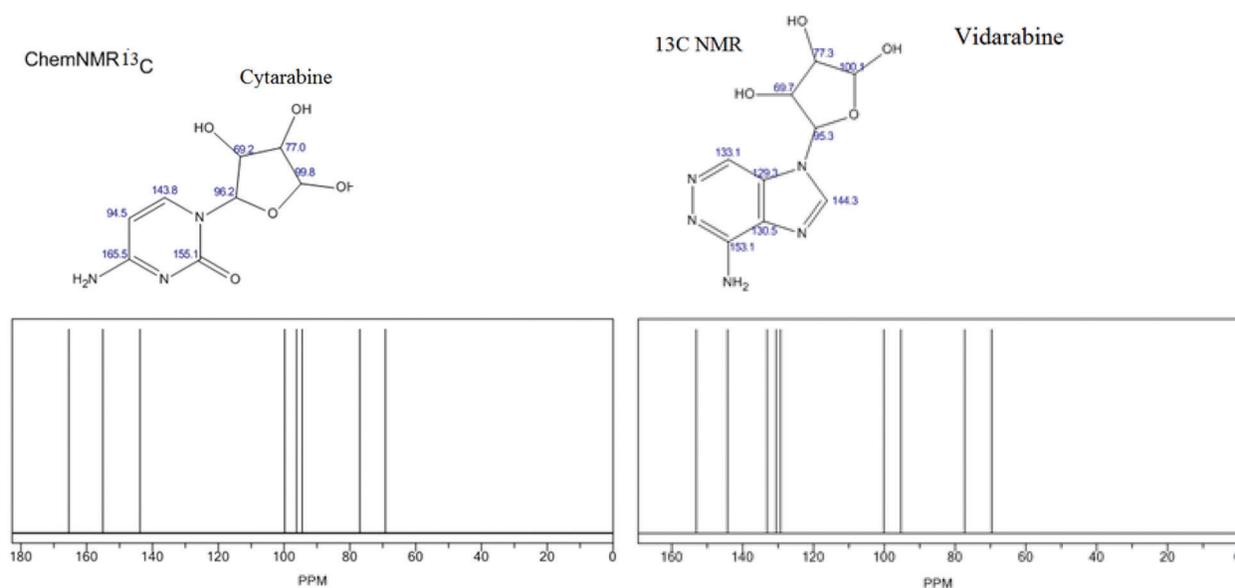


Figure 1: ¹³C NMR and chemical shifts of cytarabine and vidarabine

3 Results & Discussion

3.1 Design Analysis

The amino acid sequences of target protein DNA binding ORF8, ORF3a gene [(Porcine transmissible gastroenteritis coronavirus (strain Purdue) (TGEV)], ORF10 (Beluga whale coronavirus SW1), ORFC (recombinant protein), ORF9B_SARS2 (P0DTD2), Spike glycoprotein (P0DTC2), E protein (P0DTC4), M protein (P0DTC5), ORF6 and ORF7 (a) were extracted from Gen/bank in FASTA format. Vaxijen has been applied for checking the antigenicity of those related proteins. Several highly proteins as antigens were distinguished where the most of them have been founded by the S, E, M, ORF10, ORF6, ORF (7a), and ORF8.

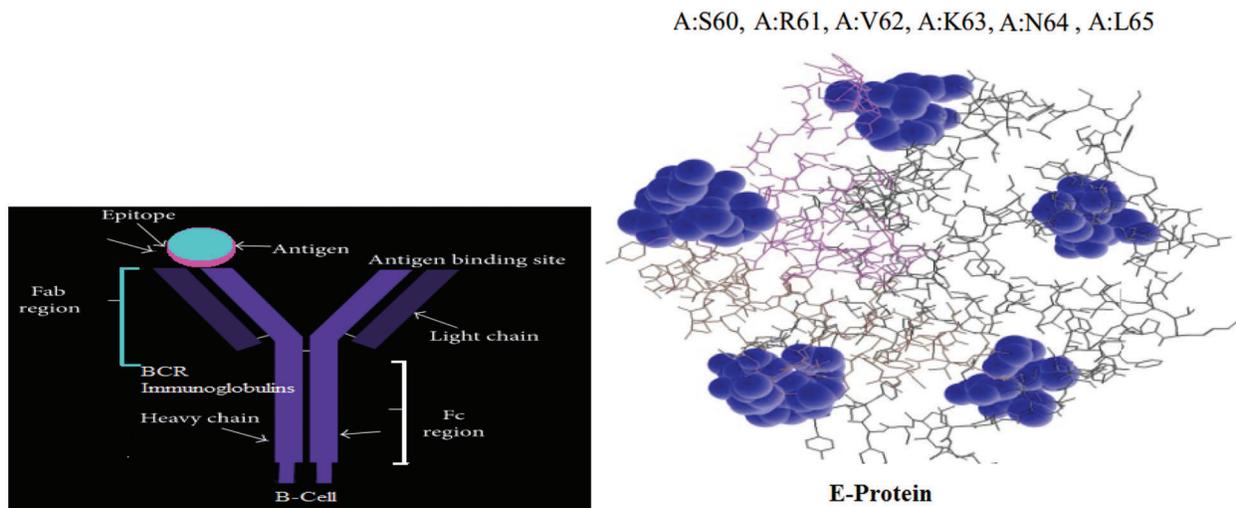
Other physicochemical characters were predicted from protparam (Fig. 2 and Tabs. 4, 5).

Table 2: The isotropy and SNICS as an aromaticity criterion

Matrine with S-protein in gas phase							Matrine with S-protein in water pH = 7						
Ω	$\delta\Delta$	η	S-NICS	σ_{iso}	Charge	Atom	Ω	$\delta\Delta$	η	S-NICS	σ_{iso}	Charge	Atom
3 C	-0.109	144.6	142.7	0.5679	-20.60	17.52	3 C	-0.109	149.34	143.9	0.986	-25.56	-17.041
4 N	-0.217	146.9	144.9	0.2451	-10.16	10.88	4 N	-0.214	147.43	146.6	0.500	-11.64	10.345
5 C	0.156	47.00	46.7	0.7037	-134.3	105.1	5 C	0.167	44.451	44.8	0.647	-136.3	110.36
6 N	0.098	55.50	55.4	0.9760	-114.1	77.01	6 N	0.096	56.897	57.5	0.992	-115.3	-76.923
7 C	-0.149	60.68	62.1	0.9963	-114.3	-76.2	7 C	-0.147	60.482	61.8	0.984	-113.8	76.539
8 O	0.191	47.84	46.8	0.8919	-139.1	-92.77	8 O	0.191	48.078	49.7	0.884	-138.8	-92.595
1 H	-0.33	160.1	169.8	0.334	-13.2	13.22	1 H	-0.340	159.0	158.7	0.291	-13.4	13.91
2 C	-0.294	164.0	163.5	0.2341	-4.497	4.858	2 C	-0.294	164.70	163.1	0.517	-5.98	5.2589

Table 3: The Isotropy and SNICS as an aromaticity criterion

Cytarabine with S-protein in gas phase							Cytarabine with S-protein in water pH = 7						
Ω	$\delta\Delta$	η	S-NICS	σ_{iso}	Charge	Atom	Ω	$\delta\Delta$	η	S-NICS	σ_{iso}	Charge	Atom
4 N	-0.193	147.65	145.8	0.511	-12.45	10.989	4 N	-0.196	147.8	141.9	0.485	-11.06	9.934
5 O	0.1262	46.492	149.6	0.741	-143.8	110.15	5 O	0.122	47.31	46.5	0.757	-143.6	108.95
6 C	-0.061	72.119	74.9	0.747	-120.1	-80.08	6 C	-0.060	70.69	72.4	0.760	-121.3	-80.882
7 N	-0.178	64.33	65.8	0.370	-139.3	-92.91	7 N	-0.174	60.01	61.8	0.544	-139.7	-93.18
8 C	0.1905	57.44	58.9	0.982	154.3	103.80	8 C	0.143	56.22	57.6	0.997	-151.3	101.05
9 O	-0.353	160.51	164.0	0.495	-15.0	13.43	9 O	-0.35	163.4	164.5	0.348	-11.73	11.608
10 N	-0.33	161.4	163.3	0.483	-15.8	-10.59	10 N	-0.340	162.0	161.3	0.452	-16.14	-10.76
1 O	-0.334	166.51	167.3	0.624	-10.8	8.936	1 O	-0.334	166.6	165.8	0.612	-10.87	8.993
2 C	-0.287	165.98	164.7	0.955	-6.175	-4.1172	2 C	-0.287	165.8	164.5	0.875	-6.230	-4.1534
3 O	-0.090	144.94	142.9	0.987	-23.8	-15.88	3 O	-0.091	145.0	143.7	0.961	-24.1	-16.087

**Figure 2:** Protein conformational B-cell epitopes 3D-structure

In order for evaluating the performance of those models and in the first step, the regression model for Mpro inhibitors has been constructed using the several compounds such as E-Protein, M-Protein, S-Protein and several ORFX (X = 10, 8, 7, 3) in ChEMBL and using Pubchem fingerprint with default neuron network parameters.

Table 4: Peptide sequence of 3D structures for some proteins of COVID-19

Protein	Peptide	3D Structure
ORFC	MNKILGLRRA KSAPLVPGAE KGKEKSVEET GFMTLAGRLR RGMQRLSRRG YGDNRRSRGS ENNEQDPQPG DKIASPQRRD YTKSEASCRP GSGKTSPCGS SGTPCSDDAG GGRNGQENSG TRDTPCWMYK DSKSRYRVGV TPDLIPTIFG VSEVAASGLP RCRDKAAKRQ PQSLLSPGVE ALLVTIAESL ETNGKRVSGR TAGKLWSWRV RDKAPERDYR NVTPTMFEGS CFGKPIRAGV FNAPRAYLDD LLGDHYFVPY LRRLPRDFTR EETLSLRVAT EAAVFANMLW EARHKNNFGA GVSYYPGALA SATGTAPDFT DRGRSSLDSP YFASTFLPGI FVILPPGELP IDFMRLAVLL AVSAIETCVT TV	ModBase 3D for Q6UDL3
ORF8	MKFLVFLGII TTVAAFHQEC SLQSQTHQHP YVVDPCPIH FYKWKYIRVG ARKSAPLIEL CVDEAGSKSP IQYIDIGNYT VSCLPFTINC QEPKLGSLVV RCSFYEDFLE YHDVRRVLDLFI	

Our model exhibits an acceptable performance. The structures of all compounds were submitted to Deep Screening and predicted a suitable model via Molecular Docking of those compounds similar to several previous works. The compounds which exhibited suitable binding energies are listed in (Tab. 6 and Figs. 2, 3).

Therefore, for the treatment of disease, some drugs have been simulated and modeled by this work that can treat the disease and prevent it to be spread. In this regard, drug repurposing may help us for treating and preventing infections associated with COVID-19 or SARS-CoV-2 (Figs. 3, 4).

In this study, we simulated several drugs from Drug bank database against the target Main protease (Mpro) for the treatment of COVID-19. Among several drugs, a few best (Tab. 6) drugs were selected, that had better binding energies as compared to the reference molecules. Based on the Binding energies scores, we can suggest that the identified drug might be considered for therapeutic development against the COVID-19. This research will help to get new drugs against COVID-19 and help humans against this pandemic disease (Fig. 4). Among the herbarium drugs we selected four molecules to be examined as anti COVID-19, namely Matrine, Cytarabine, Gemcitabine and Vidarabine from which are extracted from Gillan's plants such as Trshvash, Chuchaq, Cote D' Couto and Khlvash in Iran (Monajjemi, 2020 issue 3; Monajjemi, 2020 issue 5).

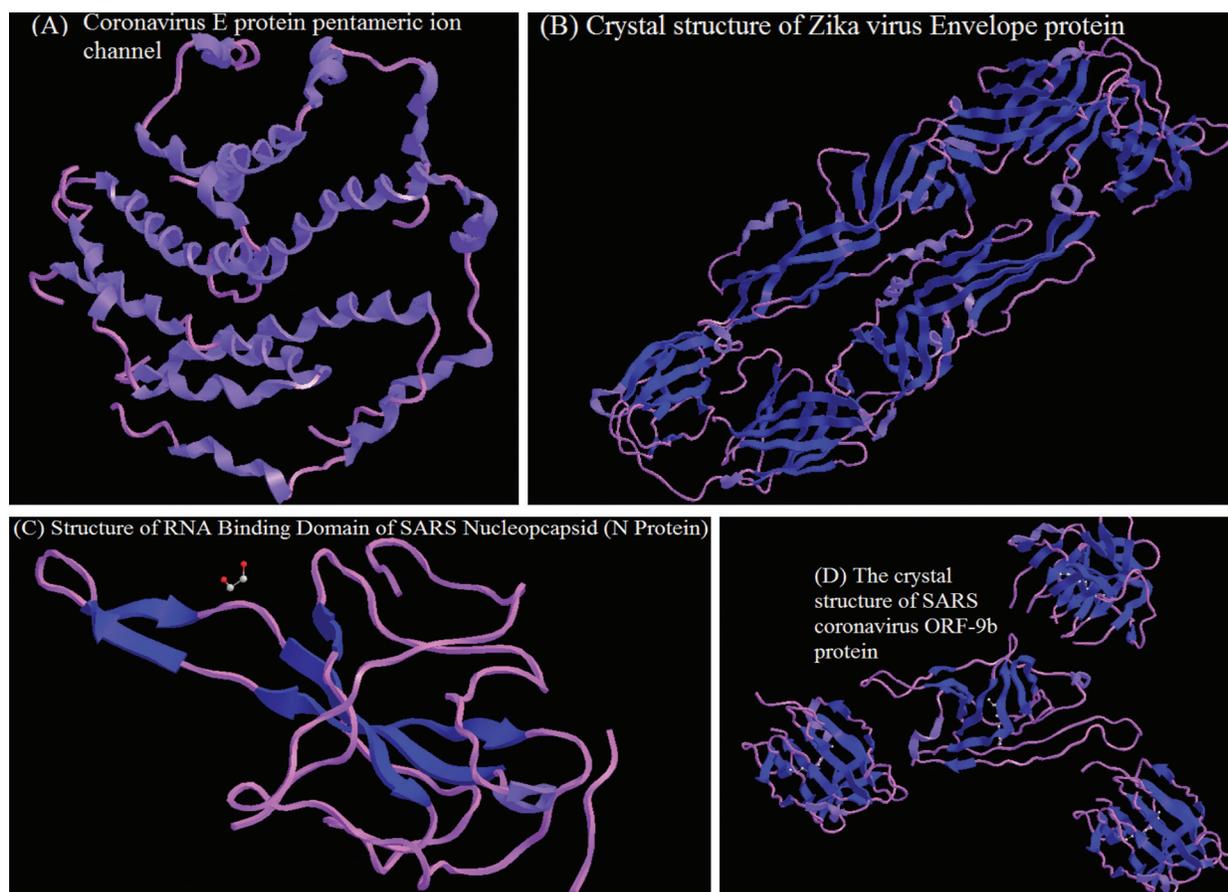
Table 5: Amino acid sequences of SARS-CoV-2 include entries for the SARS-CoV-2 virus (proteome UP000464024) and SARS-CoV virus (proteome UP000000354) and human target proteins

AC	ID	Length(amino acids)	Short name	Full name
P0DTC3	AP3A_SARS2	275		Protein 3a
P0DTC9	NCAP_SARS2	419	NC; Protein N;	Nucleoprotein
P0DTC6	NS6_SARS2	61	ns6;	Non-structural protein 6
P0DTC7	NS7A_SARS2	121		Protein 7a;
P0DTD8	NS7B_SARS2	43	ns7b;	Protein non-structural 7b
P0DTC8	NS8_SARS2	121	ns8;	Non-structural protein 8
P0DTC2	SPIKE_SARS2	1273	S glycoprotein;	Spike glycoprotein
P0DTC4	VEMP_SARS2	75	E protein; sM protein	Envelope small protein
P0DTC5	VME1_SARS2	222	M protein;	Membrane protein; E1 glycoprotein; Matrix glycoprotein membrane
P0DTD3	Y14_SARS2	73		Uncharacterized protein 14
P59632	AP3A_CVHSA	274		protein 3a;
P59595	NCAP_CVHSA	422	NC; protein N;	Nucleoprotein;
P59633	NS3B_CVHSA	154	ns3b;	Non-structural protein 3b
P59634	NS6_CVHSA	63	ns6;	Non-structural protein 6
P59635	NS7A_CVHSA	122		Protein 7a
Q7TFA1	NS7B_CVHSA	44	ns7b;	Protein non-structural 7b
Q7TFA0	NS8A_CVHSA	39	ns8a;	Protein non-structural 8a
Q80H93	NS8B_CVHSA	84	ns8b;	Non-structural protein 8b
P59636	ORF9B_CVHSA	98		Protein 9b
P0C6X7	R1AB_CVHSA	7073	pp1ab;	Replicase polyprotein 1ab; ORF1ab polyprotein;
P0C6U8	R1A_CVHSA	4382	pp1a;	Replicase polyprotein 1a; ORF1a polyprotein;
P59594	SPIKE_CVHSA	1255	S glycoprotein;	spike glycoprotein; E2
P59637	VEMP_CVHSA	76	E protein; sM protein	Envelope small protein
P59596	VME1_CVHSA	221	M protein;	Membrane protein; E1 glycoprotein; matrix glycoprotein; membrane glycoprotein;
Q7TLC7	Y14_CVHSA	70		Uncharacterized protein 14

This information with grate efficiency selected for anti-SARS-CoV-NSPs that confirm our molecular simulation & modelling. By this work it has been shown that Cytarabine molecule (in Chuchaq), and Matriline (from Trshvash), have lower binding energies compared to the respected reference compounds. (Figs. 4, 5). The grids between 15–20 Å were generated over the peptide-like inhibitors of all proteins and as well as for small-molecule inhibitors. Re-docking of the Cytarabine and Matriline were accomplished to make a certain results for the docking processing. As a result the binding site for SARS-CoV-2 is restricted with hydrophilic residues with mines charged (ASP, GLU) and one also positively charged (Arg). Besides those residues, ASN-955, and VAL-951 residues also interact with the ligands. Therefore the Cytarabine is a suitable inhibitor against SARS-CoV-2.

Table 6: Binding energy of drug compounds against M^{pro} receptor peptide-like and small-molecule inhibitors against COVID-19

Plant source & compound ID	Ligand	M protein (P0DTC5)	Spike glycol protein (P0DTC2)	RBD-ACE2 (6VW1)	E glycol protein	Envelope protein	ORF9B_SARS2 (P0DTD2)	Nucleocapsid phosphoprotein
<i>Chuchaq</i>	<i>Cytarabine</i>	-8.13	-9.13	-8.35	-9.23	-9.32	-8.45	-7.56
<i>Trshvash</i>	<i>Matrine</i>	-7.72	-7.45	-7.15	-8.32	-8.43	-4.99	-9.66
<i>Cote D' Couto</i>	<i>Gemcitabine</i>	-6.93	-8.13	-6.40	-7.53	-7.54	-7.44	-7.56
<i>Khlvash</i>	<i>Vidarabine</i>	-8.15	-7.92	-6.22	-6.34	-6.54	-8.42	-6.55

**Figure 3:** The 3D schematics of the novel coronavirus proteins by the homology modeling including (A) E glycoprotein of the surface glycoprotein, (B) envelope protein, (C) nucleo-capsid phosphoprotein, (D) orf9b protein

3.2 Free Energies of Docking

The grids between 15–20 Å were generated over the peptide-like inhibitors of all proteins and as well as for small-molecule inhibitors. Re-docking of the compounds was accomplished to validate the docking protocols. The docked molecules were superimposed to the original crystal structures for calculating the root mean square deviation (RMSD). The re-docking of peptide-like

structures and small molecule inhibitors reproduces the original pose with 1.20 Å and 0.80 Å RMSD, respectively. Lower RMSD exhibits that our docking simulation is adequate and can be utilized to search small molecule inhibitors. Docking calculations were carried out in three different modes, virtual screening followed by standard-precision (SP) and extra-precision (XP) docking using the Glide program. After XP docking, compounds were re-scored using prime MM-GBSA free energy calculations (Schrodinger, 2020). SARS-CoV-2 main protease ligand-binding pocket could be divided into 4 sub pockets. The interaction map and surface diagram of SARS-CoV-2 are depicted in Fig. 4. The binding site for SARS-CoV-2 is surrounded with hydrophilic as well as hydrophobic residues with two negatively charged (ASP, GLU) and one positively charged (Arg) residues, Fig. 4. Besides those residues, ASN-955, and VAL-951 residues also interact with the ligands. An interaction map of the Matrine, Cytarabine, Gemcitabine and Vidarabine compounds are depicted in Fig. 5. As the docking scores were not able to distinguish between the molecules, we utilized MM-GBSA based binding free energy (DG-bind) values for selecting the best complexes for MD simulations. FDA approved drugs were docked inside the SARS-CoV-2 main protease. The Cytarabine is a suitable inhibitor against SARS-CoV-2.

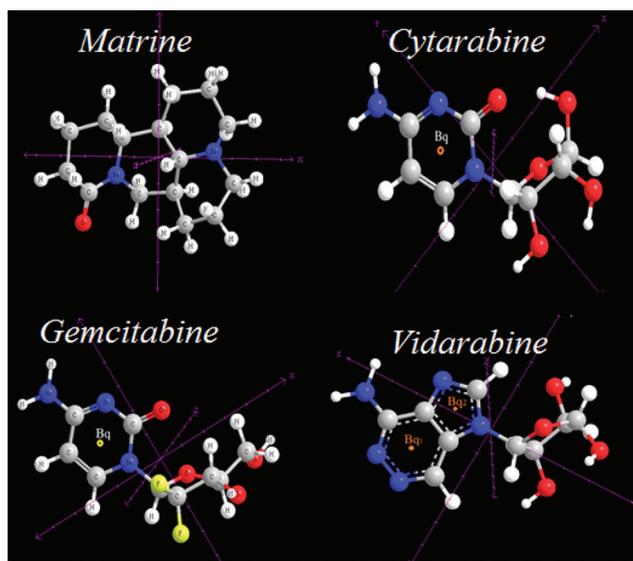


Figure 4: Chemical optimized structures of Matrine, Cytarabine, Gemcitabine and Vidarabine with M06 and m06-L (DFT) functional/cc-pvdz & cc-pvtz basis sets, including NMR = CSGT including Pop = ChelpG

Finally by this work we present a method on the Computational Prediction of Protein Structure Associated with COVID-19 Based Ligand Design and Molecular Modeling. Specifically, we, based on docking simulation and NMR investigation demonstrated, a protease with 4 natural product species as anti-COVID-19 (SARS-CoV-2) exhibit suitable binding energy around 9 Kcal/mol with various ligand proteins modes in the binding to COVID-19 viruses.

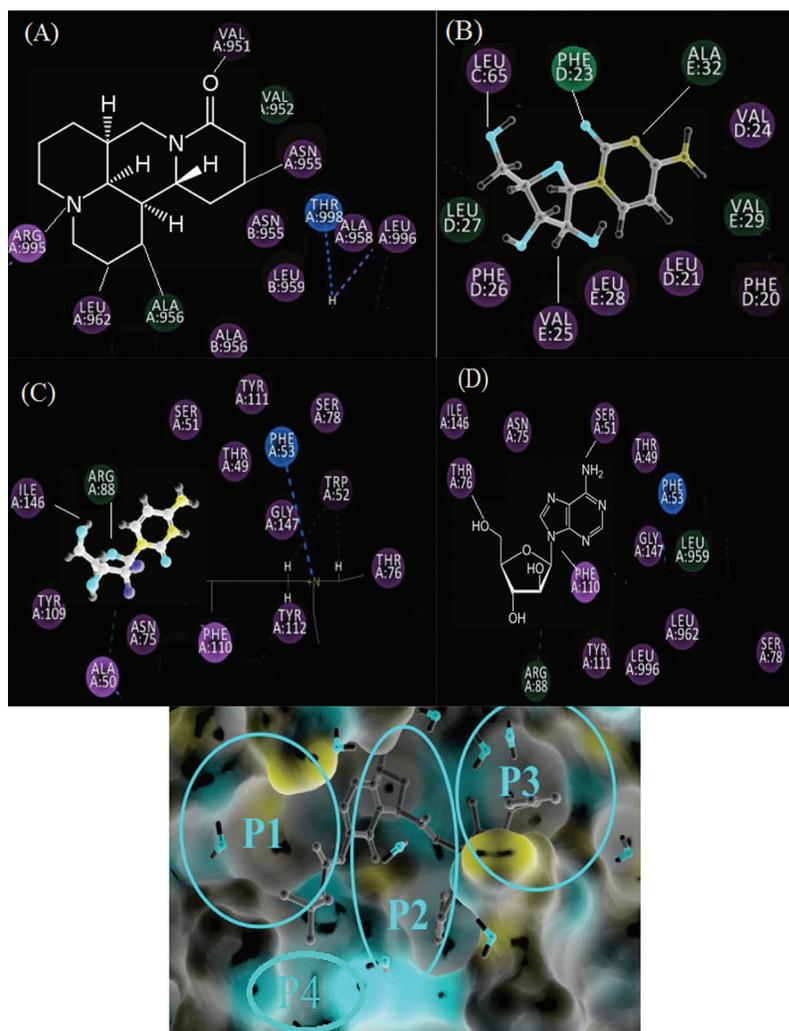


Figure 5: Molecular docking results of viral structure proteins and (A) = *Matrine*, (B) = *Cytarabine*, (C) = *Gemcitabine*, (D) = *Vidarabine* with ORF9B_SARS2 (P0DTD2)

4 Conclusions

Peptide like molecules provides a basic pharmacophore for the design of SARS-CoV-2 main protease inhibitors. Natural product might be an excellent alternative for general inhibitors, as they are easy to synthesize and less toxic when compared to those inhibitors. This study provides a detailed analysis of essential residues and ligand-receptor interactions for the development of peptide-like structures as SARS-CoV-2 main protease inhibitors. Interactions of four mentioned structures were further validated utilizing MD simulations. The chosen compounds showed strong binding affinities with residues inside the binding site and formed the strong H-bonding. Docking analysis suggests that those compounds which are hydrophilic and tend to form hydrogen bonds. Drug repurposing techniques are widely being explored to overcome the current outbreak of SARS-CoV-2. Although, A few natural product compounds were identified and some of

them are currently under clinical trials, further experimental studies are necessary for validating our findings.

Current understanding about how the virus that causes COVID-19 spreads is largely based on what is known about similar coronaviruses. The virus is thought to spread mainly from person to person: Between people who are in close contact with one another (within about 6 feet). Through respiratory droplets produced when an infected person coughs or sneezes. These droplets can land in the mouths or noses of people who are nearby or possibly be inhaled into the lungs. COVID-19 may be spread by people who are not showing any symptoms. The virus may also be spread through surfaces: By a person touching a surface or object that has virus on it and then touching their own mouth, nose, or possibly their eyes. This is not thought to be the main way the virus spreads, but we are still learning more about how this virus spreads. In addition based on this work constructive advice as possible, as well as possible treatment options of COVID-19 might be concluded as: Wash hands often for 20 s and encourage others to do the same. Use hand sanitizer with at least 60% alcohol if soap and water are unavailable. Wear a cloth face covering in public and during large gatherings. Cover coughs and sneezes with a tissue, and then throw the tissue away. Avoid touching your eyes, nose, and mouth with unwashed hands. Disinfect surfaces, buttons, handles, knobs, and other places touched often. Stay six feet apart from others. Avoid close contact with people who are sick.

Funding Statement: The author thanks the Islamic Azad university, for providing only the software and computer equipment without funding.

Conflicts of Interest: The Author declares that they have is no conflict of interests to report regarding the present study.

References

1. Báez-Santos, Y. M., Barraza, S. J., Wilson, M. W., Agius, M. P., Mielech, A. M. et al. (2014). X-ray structural and biological evaluation of a series of potent and highly selective inhibitors of human coronavirus papain-like proteases. *Journal of Medicinal Chemistry*, *57*, 2393–2412.
2. Hilgenfeld, R. (2014). From SARS to MERS: Crystallographic studies on coronaviral proteases enable antiviral drug design. *FEBS Journal*, *281(18)*, 4085–4096. DOI 10.1111/febs.12936.
3. Lehmann, K. C., Gulyaeva, A., Zevenhoven-Dobbe, J. C., Janssen, G. M., Ruben, M. et al. (2015). Discovery of an essential nucleotidylating activity associated with a newly delineated conserved domain in the RNA polymerase-containing protein of all nidoviruses. *Nucleic Acids Research*, *43(17)*, 8416–8434. DOI 10.1093/nar/gkv838.
4. Sevajol, M., Subissi, L., Decroly, E., Canard, B., Imbert, I. (2014). Insights into RNA synthesis, capping, and proofreading mechanisms of SARS-coronavirus. *Virus Research*, *194*, 90–99. DOI 10.1016/j.virusres.2014.10.008.
5. Zhang, C., Zheng, W., Huang, X., Bell, E. W., Zhou, X. et al. (2020). Protein structure and sequence reanalysis of 2019-nCoV genome refutes snakes as its intermediate host and the unique similarity between its spike protein insertions and HIV-1. *Journal of Proteome Research*, *19(4)*, 1351–1360. DOI 10.1021/acs.jproteome.0c00129.
6. Eisenberg, D., McLachlan, A. D. (1986). Solvation energy in protein folding and binding. *Nature*, *319*, 199–203.
7. John, S. E., Tomar, S., Stauffer, S. R., Mesecar, A. D. (2015). Targeting zoonotic viruses: Structure-based inhibition of the 3C-like protease from bat coronavirus HKU4-the likely reservoir host to the human coronavirus that causes Middle East Respiratory Syndrome (MERS). *Bioorganic & Medical Chemistry*, *23(17)*, 6036–6048. DOI 10.1016/j.bmc.2015.06.039.

8. Paraskevis, D., Kostaki, E. G., Magiorkinis, G., Panayiotakopoulos, G., Sourvinous, G. (2020). Full-genome evolutionary analysis of the novel corona virus (2019-nCoV) rejects the hypothesis of emergence as a result of a recent recombination event. *Infection, Genetics and Evolution*, 79, 104212.
9. Li, S., Yuan, L., Dai, G., Chen, R. A., Liu, D. X. (2020). Regulation of the ER stress response by the ion channel activity of the infectious bronchitis coronavirus envelope protein modulates virion release, apoptosis, viral fitness, and pathogenesis. *Frontiers in Microbiology*, 10, 52. DOI 10.3389/fmicb.2019.03022.
10. Rothe, C., Schunk, M., Sothmann, P., Bretzel, G., Froeschl, G. et al. (2020). Transmission of 2019-nCoV infection from an asymptomatic contact in Germany. *New England Journal of Medicine*, 382(10), 970–971. DOI 10.1056/NEJMc2001468.
11. Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T. et al. (2020). SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell*, 181(2), 271–280.
12. Gordon, D. E., Jang, G. M., Bouhaddou, M., Xu, J., Obernier, K. et al. (2020). A SARS-CoV-2-human protein—protein interaction map reveals drug targets and potential drug-repurposing. *Nature*, 583, 459–468.
13. Kazazian, H. H. Jr., Woodhead, A. P. (1973). Hemoglobin a synthesis in the developing fetus. *New England Journal of Medicine*, 289(2), 58–62. DOI 10.1056/NEJM197307122890202.
14. De Gregorio, E., Rappuoli, R. (2012). Vaccines for the future: Learning from human immunology. *Microbial Biotechnology*, 5(2), 149–155. DOI 10.1111/j.1751-7915.2011.00276.x.
15. Patronov, A., Doytchinova, I. (2013). T-cell epitope vaccine design by immunoinformatics. *Open Biology*, 3(1), 120139. DOI 10.1098/rsob.120139.
16. Nain, Z., Karim, M. M., Sen, M. K., Adhikari, U. K. (2019). *Structural basis and designing of peptide vaccine using PE-PGRS family protein of mycobacterium ulcerans-An integrated vaccinomics approach*, Cold Spring Harbor Laboratory, <https://www.biorxiv.org/content/10.1101/795146v1>.
17. Tahir ul Qamar, M., Bari, A., Adeel, M. M., Maryam, A., Ashfaq, U. A. et al. (2018). Peptide vaccine against chikungunya virus: Immuno-informatics combined with molecular docking approach. *Journal of translational medicine*, 16(1), e1430. DOI 10.1186/s12967-018-1672-7.
18. Shahid, F., Ashfaq, U. A., Javaid, A., Khalid, H. (2020). Immunoinformatics guided rational design of a next generation multi epitope based peptide (MEBP) vaccine by exploring Zika virus proteome. *Infection, Genetics and Evolution*, 80, 104199. DOI 10.1016/j.meegid.2020.104199.
19. Usman Mirza, M., Rafique, S., Ali, A., Munir, M., Ikram, N. et al. (2016). towards peptide vaccines against Zika virus: Immunoinformatics combined with molecular dynamics simulations to predict antigenic epitopes of Zika viral proteins. *Scientific Reports*, 6(1), 23. DOI 10.1038/srep37313.
20. Khan, A., Junaid, M., Kaushik, A. C., Ali, A., Ali, S. S. et al. (2018). Computational identification, characterization and validation of potential antigenic peptide vaccines from hrHPVs E6 proteins using immunoinformatics and computational systems biology approaches. *PLoS One*, 13, 1–25.
21. Asteris, P. G., Douvika, M. G., Karamani, C. A., Skentou, A. D., Chlichlia, K. et al. (2020). A novel heuristic algorithm for the modeling and risk assessment of the COVID-19 pandemic phenomenon. *Computer Modeling in Engineering & Sciences*, 125(2), 815–828. DOI 10.32604/cmescs.2020.013280.
22. Alamri, M. A., Tahir ul Qamar, M., Alqahtani, S. M. (2020). Pharmaco informatics and molecular dynamic simulation studies reveal potential covalent and FDA-approved inhibitors of SARS-CoV-2 main protease 3CLpro. *Journal of Biomolecular Structure & Dynamics*, 1–13.
23. Steinbrecher, T., Elstner, M. (2013). QM and QM/MM simulations of proteins. *Methods in Molecular Biology*, 924, 91–124.
24. Mackerell, A. D., Jr. (2004). Empirical force fields for biological macromolecules: Overview and issues. *Journal of Computational Chemistry*, 25(13), 1584–1604. DOI 10.1002/jcc.20082.
25. Saxena, A., Wong, D., Diraviyam, K., Sept, D. (2009). The basic concepts of molecular modeling. *Methods in Enzymology: Computer Methods, Part B*, 467, 307–334.

26. Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E. et al. (2009). MEME SUITE: Tools for motif discovery and searching. *Nucleic Acids Research*, 37(Web Server), W202–W208. DOI 10.1093/nar/gkp335.
27. Doytchinova, I. A., Flower, D. R. (2007). VaxiJen: A server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics*, 8(4), 1–7. DOI 10.1186/1471-2105-8-4.
28. Deléage, G. (2017). ALIGNSEC: Viewing protein secondary structure predictions within large multiple sequence alignments. *Bioinformatics*, 33, 24, 15, 3991–3992.
29. Pandey, R. K., Bhatt, T. K., Prajapati, V. K. (2018). Novel immunoinformatics approaches to design multi-epitope subunit vaccine for malaria by investigating anopheles salivary protein. *Scientific Reports*, 8, 1–11. DOI 10.1038/s41598-017-17765-5.
30. Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., Sayers, E. W. (2008). GenBank. 2. *Nucleic Acids Research*, 37, D26–D31.
31. Bailey, T. L., Johnson, J., Grant, C. E., Noble, W. S. (2015). The MEME SUITE. *Nucleic Acids Research*, 43(W1), W39–W49. DOI 10.1093/nar/gkv416.
32. Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G. (2014). SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Research*, 42(W1), W252–W258. DOI 10.1093/nar/gku340.
33. Desmond Molecular Dynamics System, D. E. Shaw Research (2020). *Maestro-desmond interoperability tools*. New York, NY: Schrodinger, ECT Register. (2020), <https://www.clinicaltrialsregister.eu/ctrsearch/search?query=COVID-19>.
34. Monajjemi, M., Shahriari, S., Mollaamin, F. (2020). Evaluation of coronavirus families & COVID-19 proteins: Molecular modeling study. *Biointerface Research in Applied Chemistry*, 10(5), 6039–6057. DOI 10.33263/BRIAC105.60396057.
35. Monajjemi, M., Mollaamin, F., Shojaei, S. (2020). An overview on coronaviruses family from past to COVID-19: Introduce some inhibitors as antiviruses from Gillan's plants. *Biointerface Research in Applied Chemistry*, 10(3), 5575–5585. DOI 10.33263/BRIAC103.575585.
36. Monajjemi, M. (2015). Cell membrane causes the lipid bilayers to behave as variable capacitors: A resonance with self-induction of helical proteins. *Biophysical Chemistry*, 207, 114–127. DOI 10.1016/j.bpc.2015.10.003.
37. Monajjemi, M., Mollaamin, M., Shekarabi, A. S., Ghadami, A. (2021). Variety of spike protein in COVID-19 mutation: Stability. *Effectiveness and Outbreak Rate as a Target for Vaccine and Therapeutic Development*, 11(3), 10016–10026, (in press).