DETERMINATION OF PROTEINS SPECIFICATION WITH SARS- COVID-19 BASED LIGAND DESIGNING

SARA SHAHRIARI [a], MAJID MONAJJEMI [b] AND FATEMEH MOLLAAMIN [c]*

[a]*Department of Chemistry, Central Tehran Branch, Islamic Azad University, Tehran, Iran.*
[b]*Department of Chemical Engineering, Central Tehran Branch, Islamic Azad University, Tehran, Iran.*
[c]*Department of Biomedical Engineering, Faculty of Engineering and Architecture, Kastamonu University, Kastamonu, Turkey.*

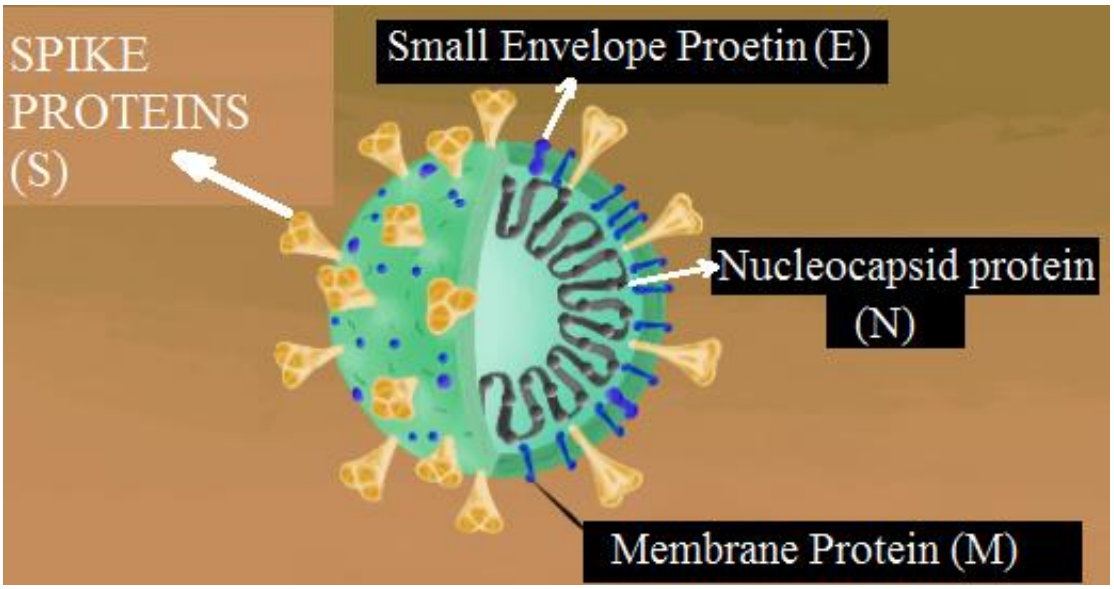
ABSTRACT

Towards to recognize a strong inhibitor we 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" from which are extractive from Gillan's leaves plants. These are known as Chuchaq, Trshvash, Cote-Couto and Khlvash in Iran. Among these four studied compounds, Cytarabine appear as suitable compound with high effectiveness inhibitors to this protease. With docking simulation and NMR investigation we 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 in both, vivo & vitro evaluation for repurposing these drugs against COVID-19 viruses.

Keywords: *Covid-19; RBD or Receptor binding domain; Gillan's leaves plants, ACE2 or Angiotensin Converting Enzyme-2;* PD or *protease domain*.

1. INTRODUCTION

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 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 that is currently posing most risk in "Wuhan" consist of a dimer proteins same as to serine proteases including Trypsin, Cysteine and Methionine amino acids. This dimerization has peptides- analogous inhibitor bound in center of active sites. This couple proteases in the SARS are the basic proteases same as the Wuhan's viruses 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 contains: Spike protein or (S), envelope protein or (E) and membrane structure (M) with nuclei-capsid phosphoprotein (Scheme 1) [5-7]. Transcribed non-structural proteins consist of: NSP1, NSP4, ORF1(ab), ORF3(a), ORF6, ORF7(a), ORF9, ORF8 and ORF10 (Table 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 (CoV EIC) can be signified in modulating virion release and Covid19-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 (Table 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 (Table 1) [11].



Scheme 1. A simple presentation of Spike Protein (S), envelope protein (E), membrane protein (M), and nuclei-capsid

1.1 Phosphoprotein

Human proteins that interacted with the proteins of covid-19 in 68 of these proteins can be applied as targets for suitable drugs (Table1) [12]. SARS-COV-2 strain is about 75% similar to SARS-COV strain and 35% close to the MERS-COV strain [13].

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.

*Corresponding author email: smollaamin@gmail.com

Table 1. 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
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)
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
TIMM10B	Nsp4	Mitochondria		TIM Complex
TIMM9	Nsp4	Mitochondria		TIM Complex
TIMM29	Nsp4	Mitochondria		TIM Complex
ACAD9	Orf9c	Mitochondria		
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		
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		

1.2. Neutrophil & Hemoglobin Specification:

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 [14]. Due to Fe2+ the CO2 gases might be separated from hemoglobin and consequently capture O atoms and consequently Fe2+ is oxidized to Fe3+. Hemoglobin can release

oxygen atoms and capture CO2 and then Fe3+ is reduced to Fe2+. 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 [15,16] .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 [17,18].

Basically, vaccinations designing against MERS, SARS, Ebola, ZIKA and Chikungunya viruses have produced promising results [19,20]. 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 [21,22]. 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-β 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.

2. THEORETICAL DETAILS

2.1. Simulation & Docking analysis

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 [23]. 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 all polypeptides of Wuhan novel covid-19 were applied for constructing 3-D structures via homology simulation.

IGEMDOCK 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 heme 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.

The Expassy-Protparam tools were applied for determining the chemical & physical properties of molecules [24]. For checking antigenicity of proteins, the related software (Vaxijen 2.0) was also applied [25] and for the predication of secondary polypeptides “Alignment self-optimized prediction method” (SOPMA) tools has been used [26]. In addition several additional online software or tools such as Swiss model [27], “Phyre2”, and “RaptorX” was applied for the tertiary polypeptide segment [27,28]. Model recovered was then refined via “Galaxy Refine Server” and accredited also via Ramachandran plots.

2.2. Genetic and Sequence Analysis

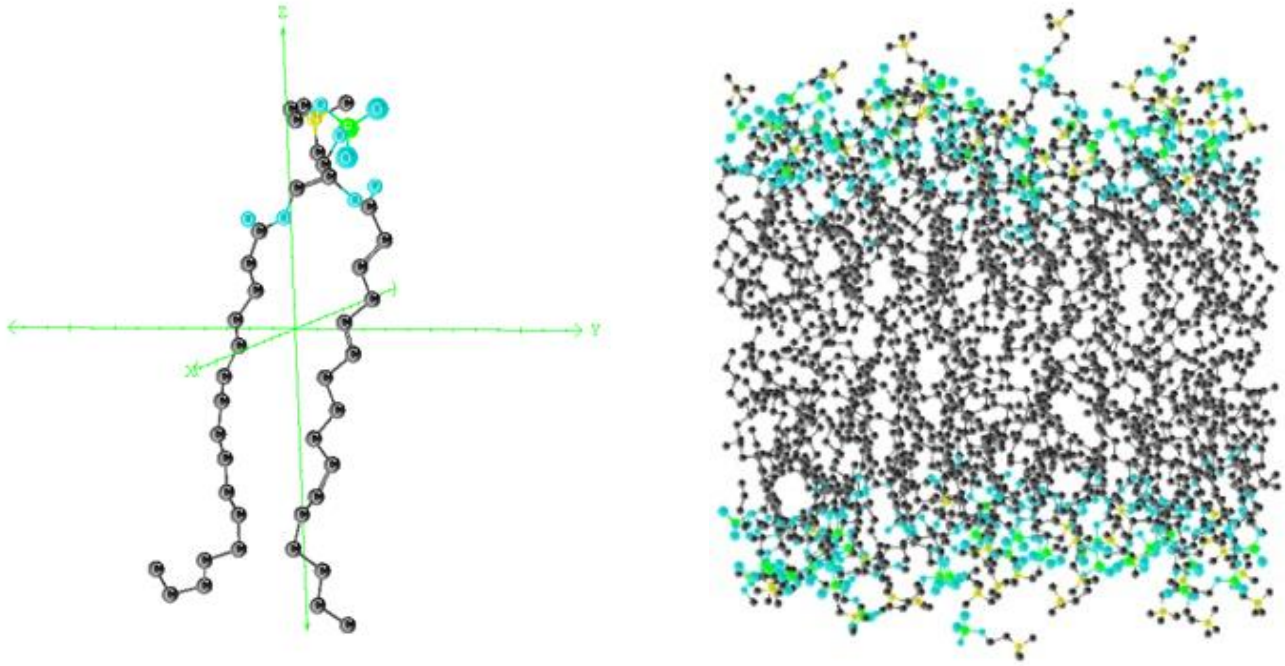
Conserved of viral proteins can be analyzed through MEME website [29-31]. These models are applied for predicting differences between the viral system and human proteins. The 3-D structures of viral proteins were structured through homology modeling of Swiss-model [32,33]. Sequence length extension and homology modeling should be adopted via applying molecular docking, 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.

2.3. Molecular Dynamic Modeling

Molecular dynamics modeling for polypeptide-ligand complexes were accomplished using the Desmond software [34]. 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.

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 (Scheme 2) [35] .



Scheme 2. Membrane simulation including 180 molecules of DPPC phospholipids.

2.4. Isotropy & Anisotropy of NMR Analysis

The anisotropy data of 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 ($\kappa = \frac{3\alpha}{\Omega}$) *and the skew is* $\kappa = \frac{3(\sigma_{1so}-\sigma_{22})}{\Omega}$ (13) ($-1 \leq \kappa \leq +1$). 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 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 Tables 2& 3).

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						
atom	**charge**	**σiso**	**S-NICS**	**η**	**Δδ**	**Ω**	**atom**	**charge**	**σiso**	**S-NICS**	**η**	**Δδ**	**Ω**
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						
atom	**charge**	**σiso**	**S-NICS**	**η**	**Δδ**	**Ω**	**atom**	**charge**	**σiso**	**S-NICS**	**η**	**Δδ**	**Ω**
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

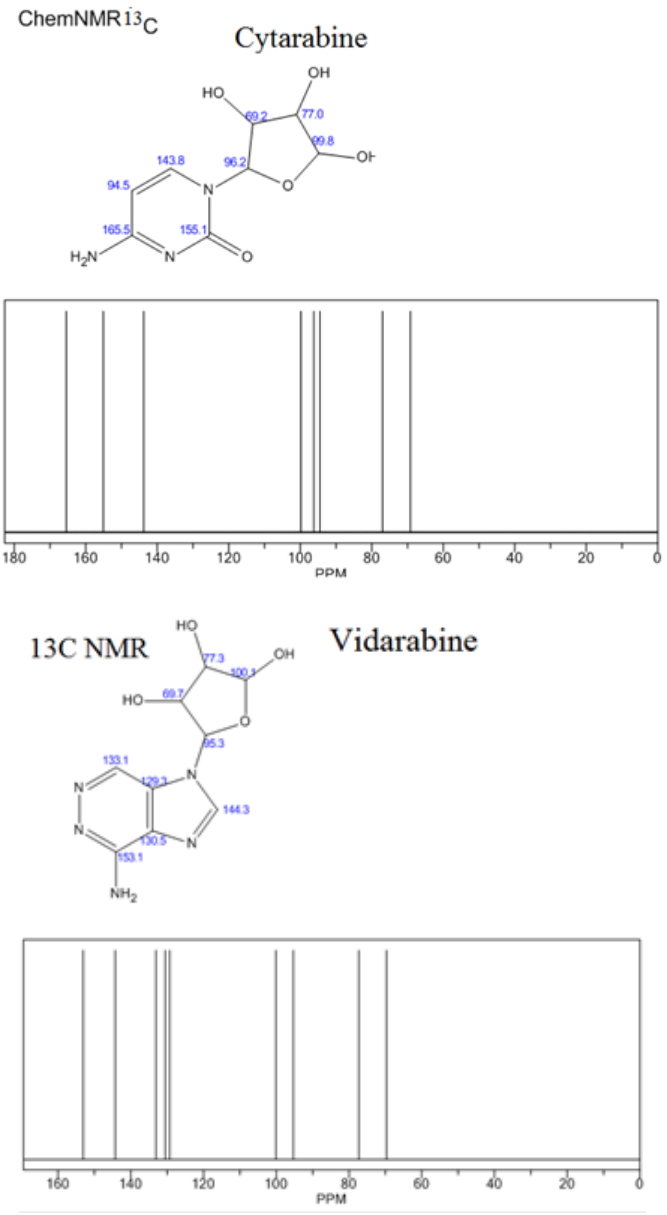


Figure 1. ^{13}C NMR and Chemical shifts of Cytarabine and Vidarabine.

3. RESULTS & DISCUSSION

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.

Table4. Peptide sequence of 3D structures for some proteins of covid-19

Covid-19	Peptide	3D Structure
ORFC	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	3D Structure for Q6UDL3
ORF8	SLQSCTQHQP YVVDDPCPIH FYSKWYIRVG ARKSAPLIEL CVDEAGSKSP IQYIDIGNYT VSCLPFTINC QEPKLGSLVV	

Other physicochemical characters were predicted from protparam (Fig.2 and Tables 4& 5). 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 5. Lipids sequences of SARS-CoV-2 include entries for the SARS-CoV-2 virus and SARS-CoV virus and human target proteins.

AC	ID	Length (amino acids)	short name	full name
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 ;

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 [36,37]. The compounds which exhibited suitable binding energies are listed in (Figs. 2, 3) [38-40].

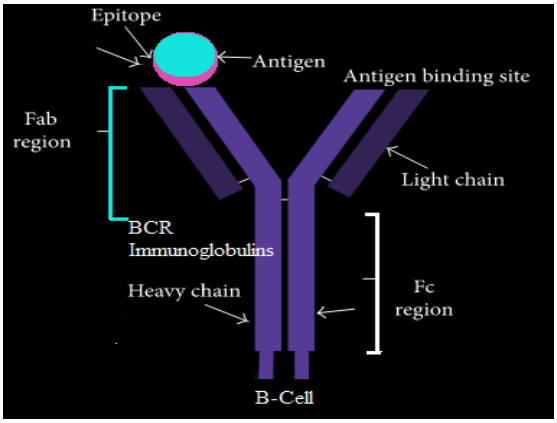


Figure 2. PROTEIN CONFORMATIONAL B-CELL EPITOPES 3D-STRUCTURE.

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 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 ([41,42].

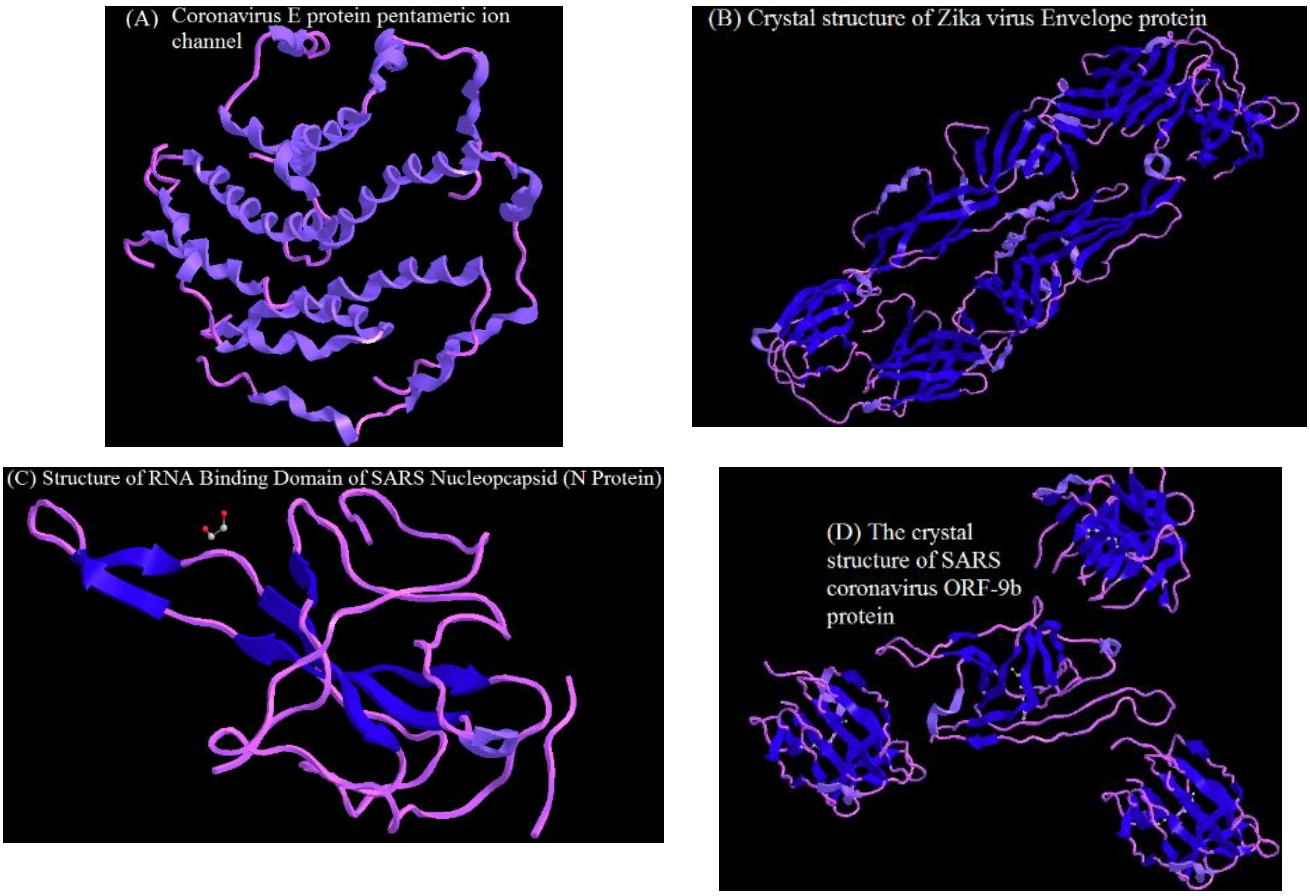


Figure 3. 3D structure schematics of the novel coronavirus proteins by the homology modeling. A. E glycoprotein of the surface glycoprotein. B. Envelope protein. C**.** nucleocapsid phosphoprotein, D. orf9b protein.

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 Matrine (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 Matrine were accomplished to make a certain results for the docking processing. As a result the binding site for SARSCoV-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.

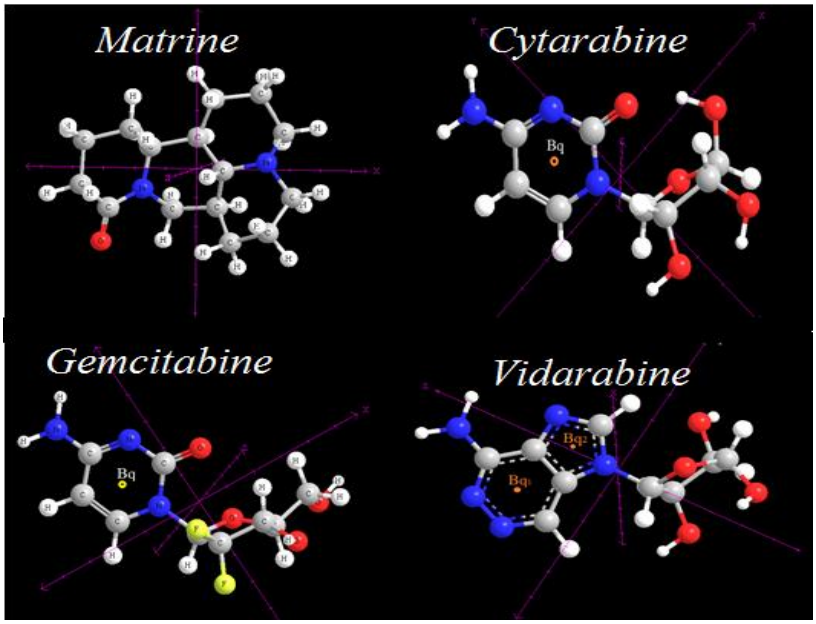


Figure 4. Chemical optimized structures of Matrine, Cytarabine, Gemcitabine and Vidarabine with M06 and m06-L (DFT) functional /cc-pvdz & cc-pvtz basis sets including NMR=CSGT including Pop=ChelpG.

CONCLUSION

Some of the Natural products are suitable drugs against covid-19. This study exhibited an analyzing information of docking simulation that exhibited those compounds tend to form hydrogen bonds. Therefore they can be widely being explored for overcoming the current outbreak of SARS-Cov-2. Although, A few natural product molecules were introduced, but further experimental studies are needed for any further confirmation of these natural products.

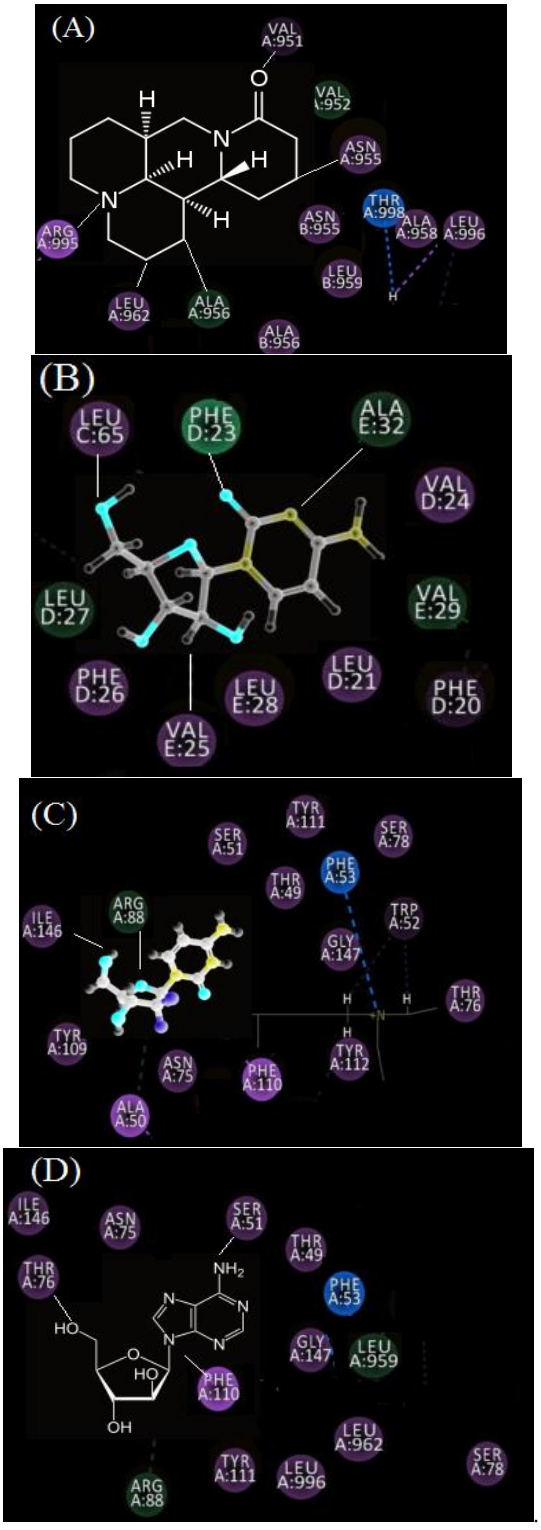


Figure 5. Molecular docking results of viral structure proteins and (A)= *Matrine, B= Cytarabine, C= Gemcitabine, (D)= Vidarabine with* SARS-COVID-19.

REFERENCES

1. Báez-Santos, Y.M., Barraza, S.J., Wilson, M.W.,Agius, M.P., Mielech, A.M., Davis, N.M., Baker, S.C., Larsen, S.D., Mesecar, A.D. (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, http://dx.doi.org/10.1021/jm401712t.
2. Hilgenfeld, R., (2014), From SARS to MERS: crystallographic studies on coronaviral proteases enable antiviral drug design. *FEBS J*., *281*,4085-4096, https://doi.org/10.1111/febs.12936.
3. Sevajol, M.,Subissi, L., Decroly, E., Canard, B., Imbert, I. (2014) Insights into RNA synthesis, capping, and proofreading mechanisms of SARS-coronavirus. *Virus Res*. *194*, 90–99, https://doi.org/10.1016 /j.virusres. 2014.10.008.
4. Lehmann, K.C., Gulyaeva, A., Zevenhoven-Dobbe, J.C., Janssen, G.M., Ruben, M., Overkleeft, H.S., van Veelen, P.A., Samborskiy, D.V., Kravchenko, A.A., Leontovich, A.M., Sidorov, I.A., Snijder, E.J., Posthuma, C.C., Gorbalenya, A.E. (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 Res*, *43*, 8416-8434, https://doi.org/10.1093/nar/gkv838.
5. Zhang, C.; Zheng, W.; Huang, X.; Bell, E.W.; Zhou, X.; Zhang, Y. (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*, 1351-1360, https://doi.org/10.1021/ acs.jproteome.0c00129.
6. 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). *Bioorg. Med. Chem. 23*, 6036-6048, https://doi.org/10.1016/ j.bmc.2015.06.039.
7. 7.Santos, Y.M., Barraza, S.J., Wilson, M.W., Agius, M.P., Mielech, A.M., Davis, N.M., Baker, S.C.,Larsen, S.D., Mesecar, A.D. (2014) X-ray Structural and Biological Evaluation of a Series of Potent and Highly Selective Inhibitors of Human Coronavirus Papain-like Proteases. *J.Med.Chem. 57*, 2393-2412, https://doi.org/10.1021/jm401712t.
8. Paraskevis, D. *et al.* (2020) Full-genome evolutionary analysis of the novel corona virus (2019-nCoV) rejects thehypothesis of emergence as a result of a recent recombination event. *Infection, Genetics and Evolution*,104212.
9. Li, S. *et al.* (2020) Regulation of the ER Stress Response by the Ion Channel Activity of the Infectious BronchitisCoronavirus Envelope Protein Modulates Virion Release, Apoptosis, Viral Fitness, and Pathogenesis.*Frontiers in Microbiology* **10**, 3022.
10. Rothe, C. *et al.* (2020) Transmission of 2019-nCoV infection from an asymptomatic contact in Germany. *NewEngland Journal of Medicine*
11. Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S.,, & Pöhlmann, S. (2020).SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. Cell. doi: https://doi.org/10.1016/j.cell.2020.02.052.
12. Gordon, D.E., Jang, G.M., Bouhaddou, M., Xu, J., Obernier, K., O'Meara, M.J.., & Krogan, N.J. (2020). A SARS-CoV-2-Human Protein-Protein Interaction Map Reveals Drug Targets and Potential Drug-Repurposing. bioRxiv doi: https://doi.org/10.1101/2020.03.22.002386
13. Báez-Santos, Y.M., Barraza, S.J., Wilson, M.W.,Agius, M.P., Mielech, A.M., Davis, N.M., Baker, S.C., Larsen, S.D., Mesecar, A.D. (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, http://dx.doi.org/10.1021/jm401712t.
14. Kazazian Jr, H. H., Woodhead, A. P. (1973) Hemoglobin A synthesis in the developing fetus. *New England Journal of Medicine* **289**, 58-62
15. De Gregorio, E., Rappuoli, R., (2012) Vaccines for the future: learning from human immunology. *Microbial biotechnology* **5**, 149-155.
16. Patronov, A., Doytchinova, (2013) I. T-cell epitope vaccine design by immunoinformatics. *Open biology* **3**, 120139.
17. Saadi, M., Karkhah, A.,Nouri, H. R.(2017) Development of a multi-epitope peptide vaccine inducing robust T cell responses against brucellosis using immunoinformatics based approaches. *Infection, Genetics and Evolution* **51**, 227-234
18. 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. *bioRxiv*, 795146
19. Tahir ul Qamar, M. *et al.* (2018) Peptide vaccine against chikungunya virus: immuno- informatics combined with molecular docking approach. *Journal of translational medicine* **16**, 298

20. 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
21. Mirza, M. U. *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**, 37313
22. Khan, A. *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** (2018).
23. Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J. & Sayers, E. W. (2008) GenBank. 2 *Nucleic acids research* **37**, D26-D31
24. Walker, J. M. *The proteomics protocols handbook*. (Springer, 2005).
25. 25.Doytchinova, I. A., Flower, D. R. VaxiJen: (2007)a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC bioinformatics* **8**, 4
26. Deléage, G. (2017) ALIGNSEC: viewing protein secondary structure predictions within large multiple sequence alignments. *Bioinformatics*
27. Waterhouse, A. *et al.* (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic acids research* **46**, W296-W303
28. 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
29. Bailey, T. L., Williams, N., Misleh, C. & Li, W. W. (2006) MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic acids research* **34**, W369-W373 .
30. Bailey, T. L. *et al.* (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic acids research* **37**, W202-W208 .
31. Bailey, T. L., Johnson, J., Grant, C. E. & Noble, W. S. (2015) the MEME suite. *Nucleic acids research* **43**, W39-W49 (2015).
32. Schwede, T., Kopp, J., Guex, N. & Peitsch, M. C. (2003) SWISS-MODEL: an automated proteinhomology-modeling server. *Nucleic acids research* **31**, 3381-3385 (2003).
33. Biasini, M. *et al.* (2014) SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic acids research* **42**, W252-W258 (2014).
34. Desmond Molecular Dynamics System, D. E. Shaw Research. (2020). Maestro-Desmond interoperability tools. New York, NY: Schrodinger. ECT Register. (2020, April 10). https://www.clinicaltrialsregister.eu/ctrsearch/search? query=covid-19
35. 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
36. Meier, C., Aricescu, A.R., Assenberg, R., Aplin, R.T., Gilbert, R.J.C., Grimes, J.M., Stuart, D.I.(2006) The Crystal Structure of Orf-9B, a Lipid Binding Protein from the Sars Coronavirus., *Structure* 14: 1157
37. Saikatendu, K.S., Joseph, J.S., Subramanian, V., Neuman, B.W., Buchmeier, M.J., Stevens, R.C., Kuhn, P. (2007) Ribonucleocapsid formation of severe acute respiratory syndrome coronavirus through molecular action of the N-terminal domain of N protein. J Virol 81: 3913-3921
38. Zhenming Jin, Xiaoyu Du, […] Haitao Yang , (2020), Structure of M^{pro} from SARS-CoV-2 and discovery of its inhibitors , *Nature* 582, 289-293.
39. Surya, W., Li, Y., Torres, J. (2018), Structural model of the SARS coronavirus E channel in LMPG micelles, Biochim Biophys Acta , 1860: 1309-1317
40. Dai, L., Song, J., Lu, X., Deng, Y.Q., Musyoki, A.M., Cheng, H., Zhang, Y., Yuan, Y., Song, H., Haywood, ., Xiao, H., Yan, J., Shi, Y., Qin, C.F., Qi, J., Gao, G.F.(2016) Structures of the Zika Virus Envelope Protein and Its Complex with a Flavivirus Broadly Protective Antibody. Cell Host Microbe , 19: 696-704
41. Monajjemi, M., Mollaamin, F., Shojae, S., (2020) An overview on Coronaviruses family from past to Covid-19: introduce some inhibitors as antiviruses from Gillan's plants, *Biointerface Res. Appl. Chem*., Volume 10, Issue 3,
42. Monajjemi, M., Shahriari, S., Mollaamin, F., (2020) Evaluation of Coronavirus Families & Covid-19 Proteins: Molecular Modeling Study, *Biointerface Res. Appl. Chem*., Volume 10, Issue 5.