The Binding Mechanism of Coronavirus disease 2019 with human Angiotensin Converting enzyme 2

Megha Lone[a] and Baliram Lone[b*]

[a]Department of conservative dentistry, Yashwantrao Chavan Dental College, Ahmednagar-414 003, Dist. Ahmednagar, Affiliated to Maharashtra University of Health Sciences, Maharashtra, India
[b*]Nanomaterials Research Laboratory, Department of Physics, Vinayakrao Patil Mahavidyalaya, Vaijapur-423701, Dist. Aurangabad, Maharashtra, India

*Corresponding author Email: baliram.lone@aggiemail.usu.edu

Abstract :

The outbreak across the globe due to coronavirus disease 2019 (COVID-19) has spread abruptly by infected humans worldwide. The continuous efforts by scientists is on way to understand how pandemic of COVID-19 resembles and differs from serve acute respiratory syndrome coronavirus (SARS-CoV) at transcriptomic and genomic level. The SARS-CoV and COVID-19 exploits the angiotensin converting enzyme 2 (ACE2) receptor to gain entry inside the cells.

We analyzed the entry COVID-19 into host cell due to receptor binding domain (RBD) of spike glycoprotein. The proposed simulation data shows similar ternary structures from two viruses shares approximately 80 percent identity in amino acid sequences.

Our molecular modeling investigation signifies that angiotensin converting enzyme 2 (ACE2) has stronger interaction with COVID -19 RBD. The Amino acid phenylaniline F486 LOOP plays vital role due to its penetration into hydrophobic pocket in ACE2.

The said investigation of S-Glycoprotein RBD of COVID-19 or SARS-CoV-2 via ACE2 provides post genome analysis of protein-protein interaction for rapid assessing transmission of infected patients by deadly CoVID-19. The scientific data extracted implies early guidance to control and viral prevention of CoVID-19.

Key wards: Coronavirus disease, COVID-19, SARS-CoV-2, ACE2, RBD,Molecular modelling

1. Introduction

Coronavirus (CoVs) is a single stranded RNA virus, from the coronavirinae family. The typical length of CoVs genome has in the range of 26 to 32 kilobases [1].Till date, there are six CoVs with low and high pathogenic each having three species respectively [2].Howerver,highly pathogenic CoVs are serve acute respiratory syndrome (SARS),middle east respiratory syndrome (MERS) which infects lower air ways and responsible for pneumonia [3].On last week of December 2019, received cluster of pneumonia cases similar to SARS like illness raised in the Hubei based province of Wuhan city, China. World Health Organization officially named it as 2019-nCoV and in march 2020,again renamed as COVID-19 [4].The origin of the COVID-19 epidemiologically linked to Huanan seafood wholesale market [5].However, transmission of COVID-19 from human-to human confirmed by the infection of fifteen health care practitioners when they came close contact with infected patient in Wuhan hospital [6].Today, April 14,2020 the number infected cases globally is 1844863 and 117021 deaths and total recovered patients 4,86,622 across 207 countries [7].

Basically, the coronavirus classified based on its functionality into four groups Viz.1.Alpha,2.Beta,3.Gamma and 4.Delta respectively.[8]. Beta coronavirus causes respiratory diseases including four common cold coronaviruses i.e.229E,NL63,OC43 and HKU1.

A new human coronavirus HCoV-NL-63 was associated with bronchiolitis [9] is alpha coronavirus. Whereas serve acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and middle East respiratory syndrome (MERS-CoV) beta coronavirus emerged in 2003 [10] and 2012 [5] respectively. The recent outbreak of pandemic COVID-19 remains major threat to mammals including human on the Earth and interventions strategies by means of proper vaccine or drug pursued rapidly.

Spike (S) glycoprotein responsible to enter coronavirus in host cells. S glycoprotein constitute (amino acids 1 to 14), an ectodomain indicated by amino acids 15 to 1190 ,amino acids 1191 to 1227 has a membrane spanning domain and amino acids 1227 to 1255 have a short intracellular tail [11].The ectodomain consists a receptor binding subunits S1 and S2 which is membrane fusion unit. The said virus to enter host cell S1 binds to cell surface receptor via RBD (Receptor Binding Domain). However, S2 fuses to the host cell with viral membrane which responsible the entry of viral genomes into host cells. If a cell or animal infected the specific Receptor Binding Domain (RBD)[12].Coronavirus invade human cells through binding with cell membrane receptor. These three membranes are DPP4,ACE2,and APN identified as entry receptors for human infecting coronaviruses. The ACE2 is associated with initial entry and infection of SARS-CoV.

However, we considered angiotensin converting enzymes 2(ACE2) as functional receptor for SARS-CoV[13].In present communication we performed molecular modeling of spike glycoprotein(S-GLYP) RBD of COVID-19 and reported its specific characteristics which can allow high affinity binding to ACE2 in human cells. Our investigation insights the mechanism involved in CoVID-19 virus entry associated with ACE2. Proposed structural molecular modeling of S-GLYP RBD of COVID-19 or SARS-CoV-2 via ACE2 ensures therapeutic targets and potential drugs treatment of humans suffered by transmission of CoVID-19 from human to human viral infection across the globe.

2. Methods:

Modeling of COVID-19 S-Glycoprotein RBD structure: The amino acid sequence COVID -19 spike glycoprotein RBD selected from first genome sequence of Serve acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1,complete genome deposited in GenBank 18-march-2020 (GenBank No.MN908947.3) [14].The BLAST code used for comparing DNA and protein

sequence. Clustal Omega code applied to alignment of multiple sequence. The Cn3D code used to analyze 3D structure. The Swiss Model code found at https://swissmodel.expasy.org used to perform protein structure simulation. The cocrystal structure of human ACE2 with SARS-CoV-spike glycoprotein-RBD (PDB 2AJF) as modeling templet.ACE2 and RBD interaction analyzed by molecular docking using PatchDock and fire dock code respectively. RamaChandran Plot were plotted using MolProbity [15].The Rosetta all atom energy function 2015 used to evaluate conformal energy[16].The standard protocol applied all atoms are:

```
-relax.linuxgccrelease /
-database $database_path /
-s $start_model.pdb /
-relax:through /
-nstructure $model_number
```

The protein-protein docking and all atom refinement of the lowest energy conformation performed as used in [17].
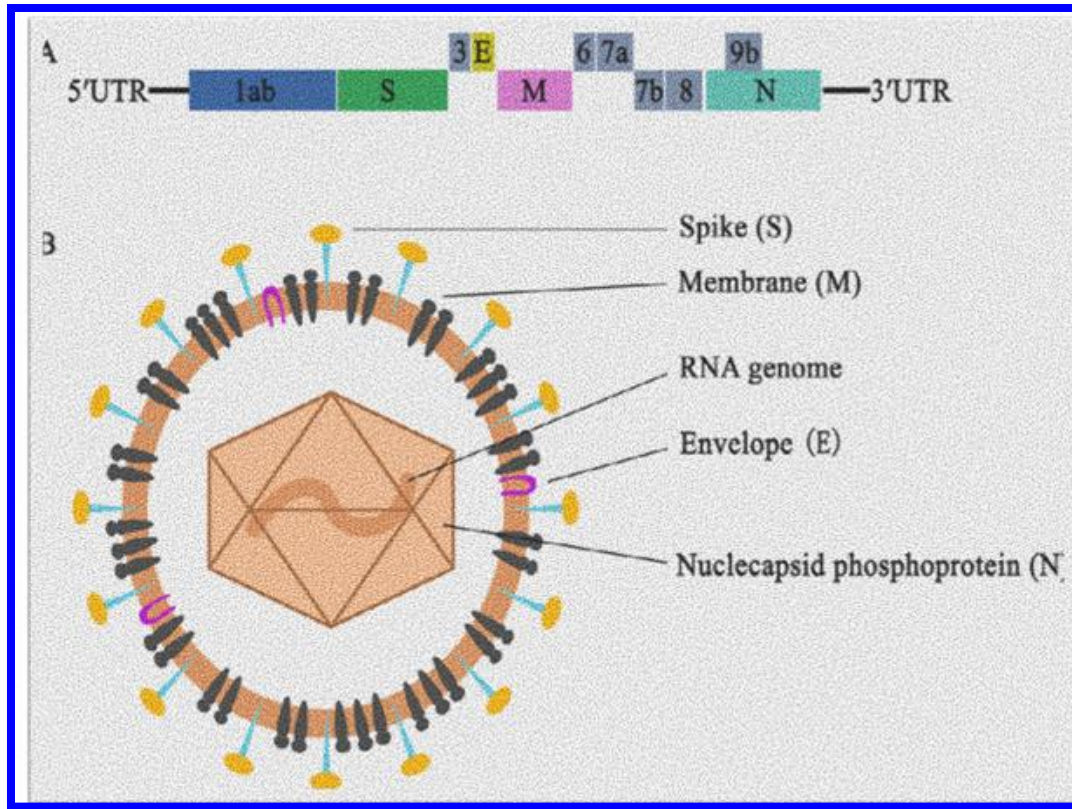
Also,retrived from https://www.rcsb.org/structure/6ACC

3. Results and Discussion

3.1 Genomic organization with virus structure of COVID-19

Figure 1 A. Describes the genomic organization with COVID-19 **19** genome consists of 5 untranslated region (5 UTR) including 5 leader sequence, open reading frame (ORF) 1a/b, envelop, membrane and nucleoprotein, accessory proteins such as 3, 6,7a, 7b, 8 and 9b and 3 untranslated region (3 UTR) in sequence.

Figure 1 B. The COVID-19 structure consists of single strand, positive sense RNA as genetic material surrounded by nucleocapsid protein in the core and an envelope constitutes proteins: Spike protein, Envelope protein, Membrane protein.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 May 2020   doi:10.20944/preprints202005.0186.v1

Figure 1. A. Genomic organization with virus structure of COVID-19. (A) COVID-19 genome (B) COVID- 19 structure with single strand, positive sense RNA.

3.1 Comparative study of COVID-19 -SARS-CoV

We searched the genome sequence MN908947.3 in BLAST search at NCBI data base accession

NC_045512 version NC_045512.2 ,30-MAR-2020.The six inputs of this virus with identical

sequence accession NC_045512.2,MN908947.3,MN975262.1,MN985325.1,MN988713.1 and

MN938384.1.The genome sequence MG772933.1 from bat homolog of COVID-19 is SARS like

coronavirus
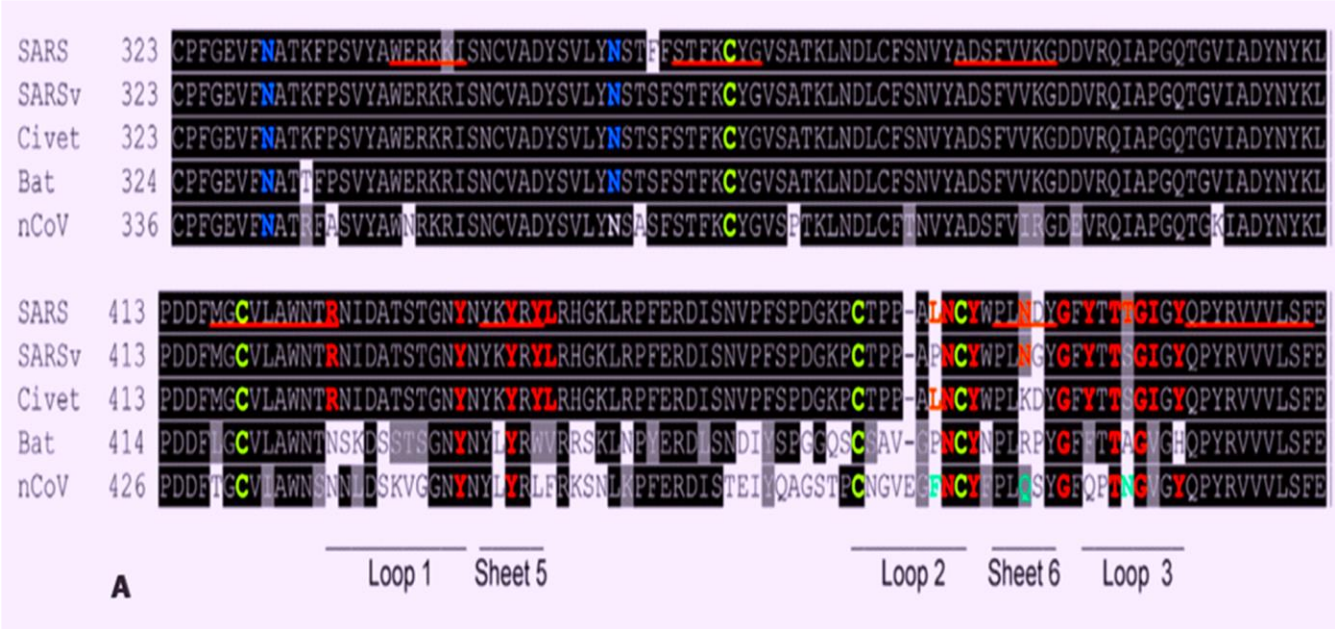
Figure 2. A. Amino acid sequence alignment for RBD of S-glycoprotein from five closely related CoV.

The above figure 2 A briefly summarized as: The genome sequence of SARS (SARS-CoV isolated during the initial 2002-2003 SARS outbreak [18],from GenBank accession numbers AY274119, AY525636. The genome sequence of SARSv (SARS-CoV isolated during the mild 2003-2004 outbreak [19], accession numbers AY304486. The genome sequence of civet (SARS-CoV-like virus isolated from palm civets [20] with accession number AGZ48806.1 The bat (SARS-CoV-like coronavirus RsSHC014 isolated from bat [21] and MN908947.1 for COVID-19.

The residue of amino acid interacting with ACE2 in SARS- CoV RBD are highlighted in RED. The secondary structures are underlined in RED. The main thrust of our study altered amino acid of COVID -19 highlighted in green. The N linked glycosylation sites and cysteine residues forming disulphide bonds are shown by yellow and cyan colours.

3.2 Crystal structure of SARS-CoV-2 or Coronavirus 2019 (COVID-19)

Figure 3. shows the crystal structure of SARS-CoV-2 protease with complex PCM -

0102578 resolution at 1.88 Angstrom unit (A.U.).
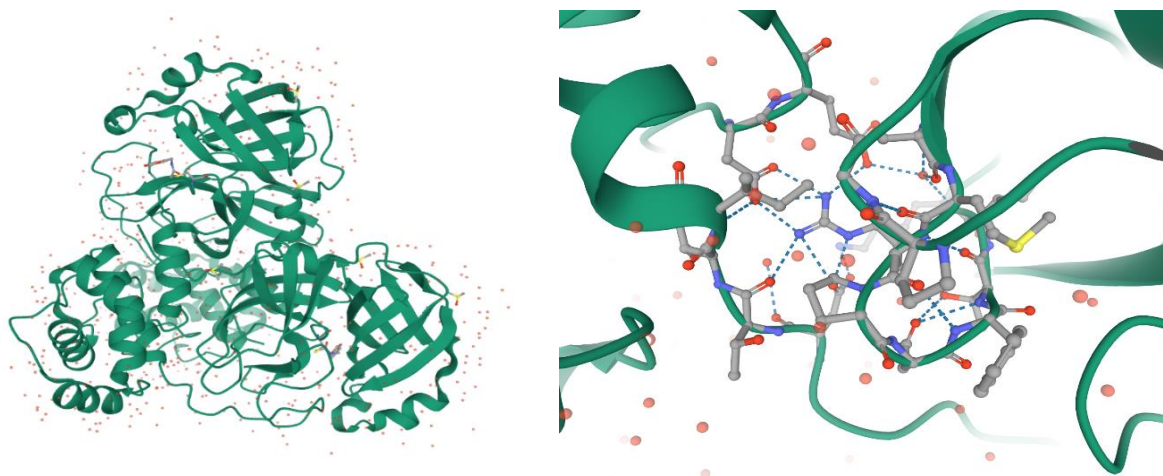
Figure 3. Crystal Structure of SARS-CoV-2 main protease in complex with PCM-0102578
Resolution at 1.88 Å.

3.3 Interaction of 2019-nCoV spike glycoprotein- receptor-binding domain (RBD)

The crystal structure of SARS -CoV RBD in complex with human ACE2 [22].We successfully

performed molecular modeling and achieved RBD of 2019 -nCoV spike glycoprotein structure

with the model templet: 6vsb.1A and obtaining sequence identity about 86% as well.

However,we successfully explored the interaction of 2019-nCoV spike glycoprotein with single

receptor-binding domain (RBD). The simulation is shown in figure 4.
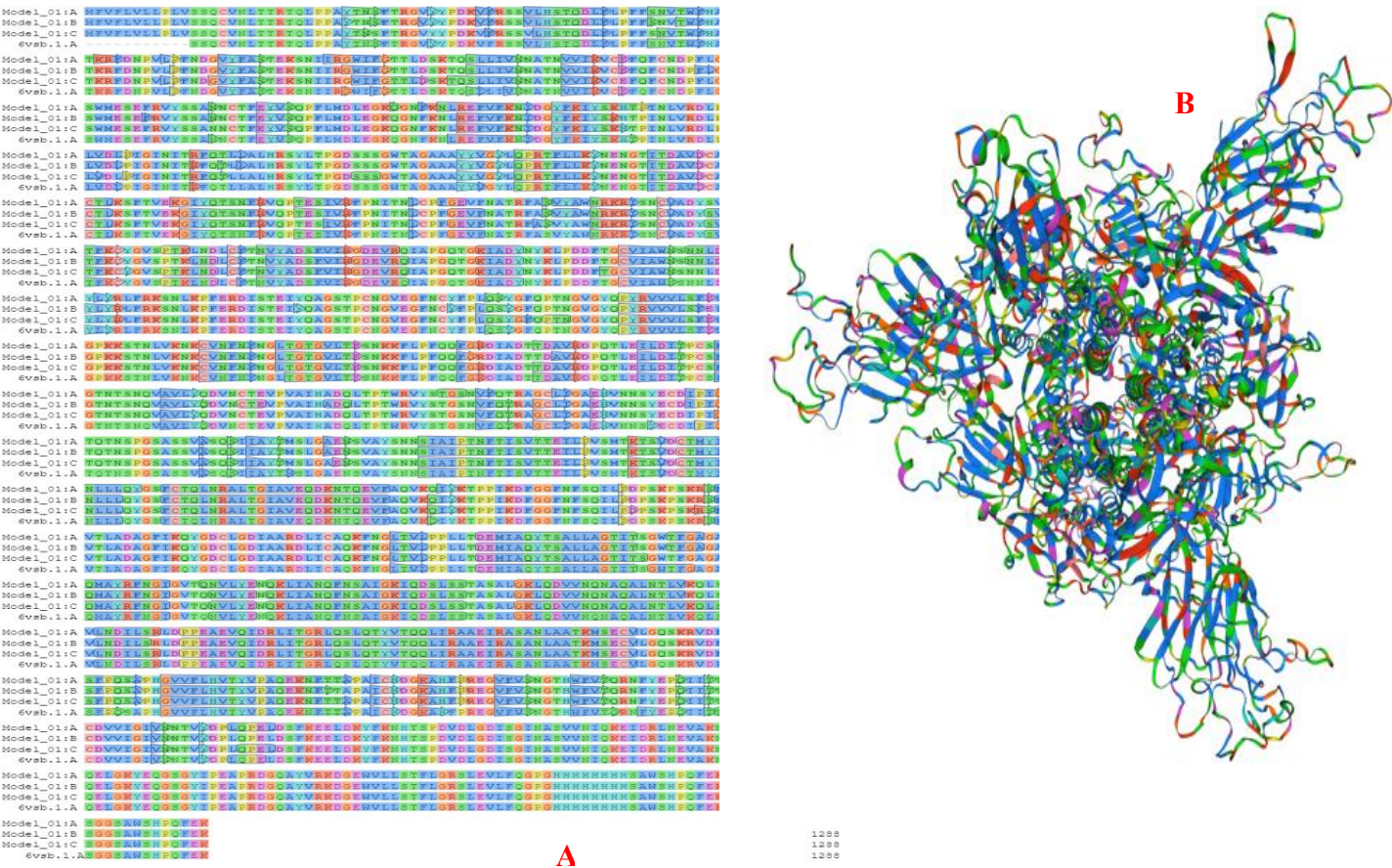
Figure 4. The model templet: 6vsb.1A. sequence identity 86% : 2019-nCoV spike glycoprotein with a single receptor-binding domain (RBD) ups: alignment clustal.

The sequence identity of 2019-nCoV spike glycoprotein with a single receptor-binding domain (RBD) achieved about 86% using alignment clustal as depicted in above figure 4.

3.4 The SARS-CoV-spike glycoprotein docking- ACE2 complex

Figure 5.shows that starting complex of SARS-CoV-spike glycoprotein docking (A)- ACE2 complex Templet 6acd.1   (B) ,sequence identity 86.92% We selected SARS-CoV S-RBD as templet, The homology modeling method was applied to predict 3D model [23].
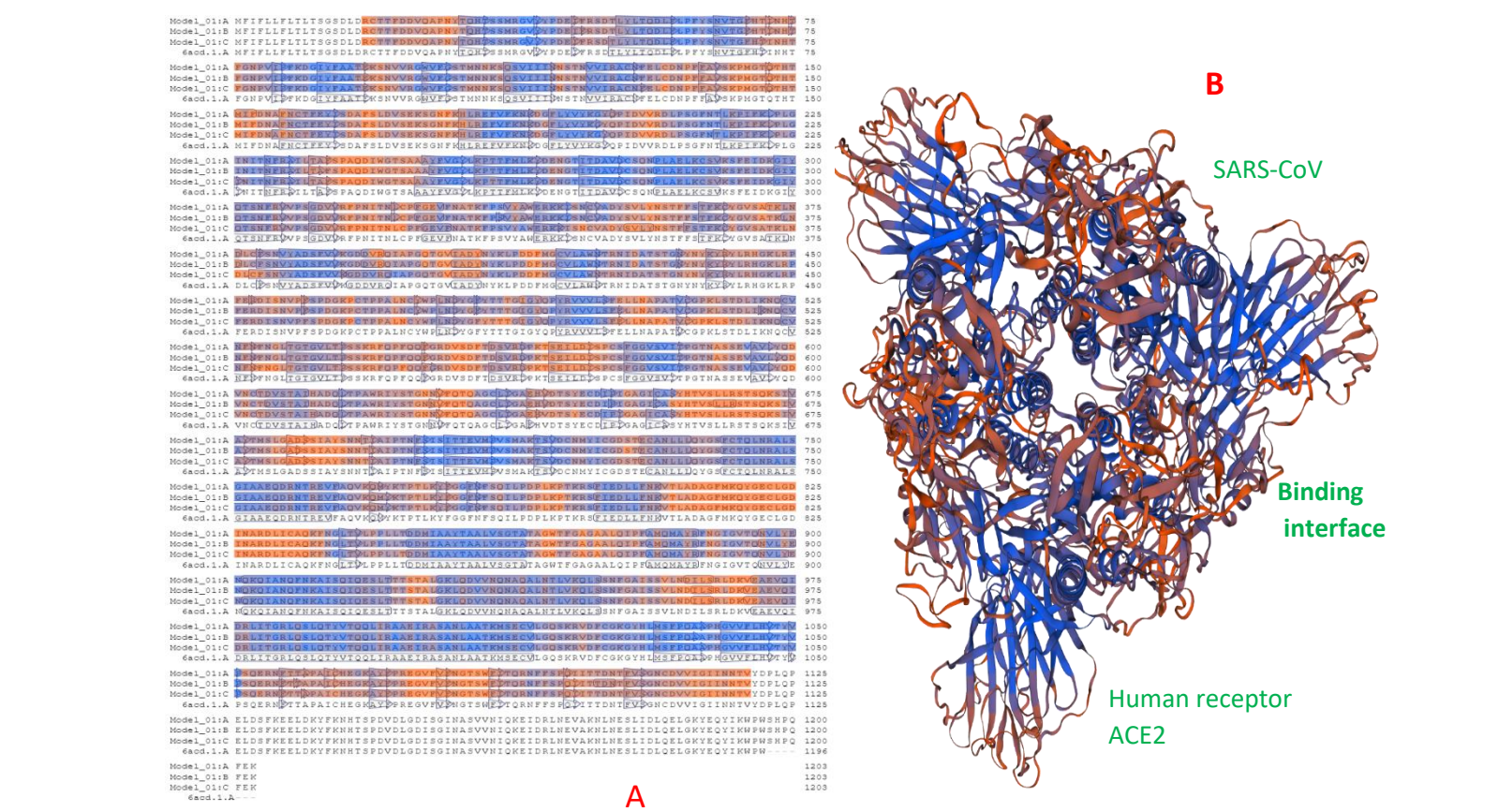
Figure 5. Starting complex of SARS-CoV-spike glycoprotein docking (A)- ACE2 complex Templet 6acd.1 (B) ,sequence identity 86.92%

In short, such residue distribution could responsible to different binding affinity from SARS-CoV Spike glycoprotein docking-RBD binding to ACE2 at sequence identity 86.92%

3.4 Binding mechanism of COVID-19 with human ACE2

Figure 6 signifies the complex model approximately 3500 independent dockings runs were performed by virtue of thermodynamical probability to evaluate low binding conformations.
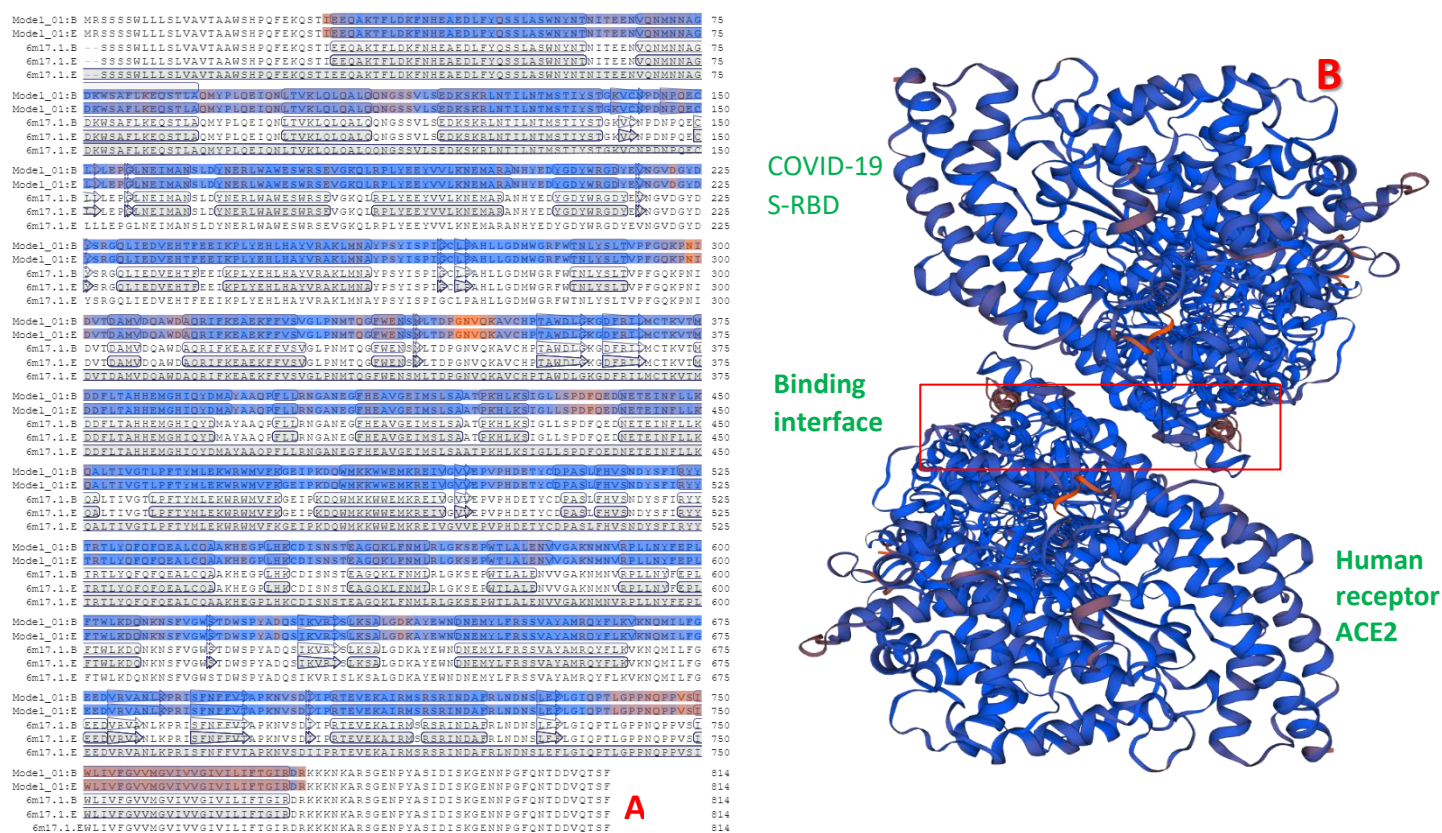
Figure 6. Starting complex of COVID-19 S-RBD-ACE2 for docking or SARS-CoV-2- human Angiotensin-converting enzyme 2 (ACE2);sequence identity 86%

The simulated data depicted in above Fig.6 exhibits COVID-19 S-RBD-ACE2 for docking or

SARS-CoV-2- human Angiotensin-converting enzyme 2 (ACE2). We performed docking simulations

[24] in energetic funnel within the 0.7Å .

Our docking process successfully enable to identify native like interacting conformations for

COVID-19 S-RBD binding to ACE2. Let us assume lowest energy conformations at the bottom

of energetic funnel those being engaged in receptor binding function.

Already, sampled large number of thermodynamically probable low binding energy formed

conformations, the results represent existence of ACE2 binding conformation for COVID-19 with

energy value <-39.20 kcal.mol^{-1} is limited. Therefore, binding strength of COVID -19 S-RBD to
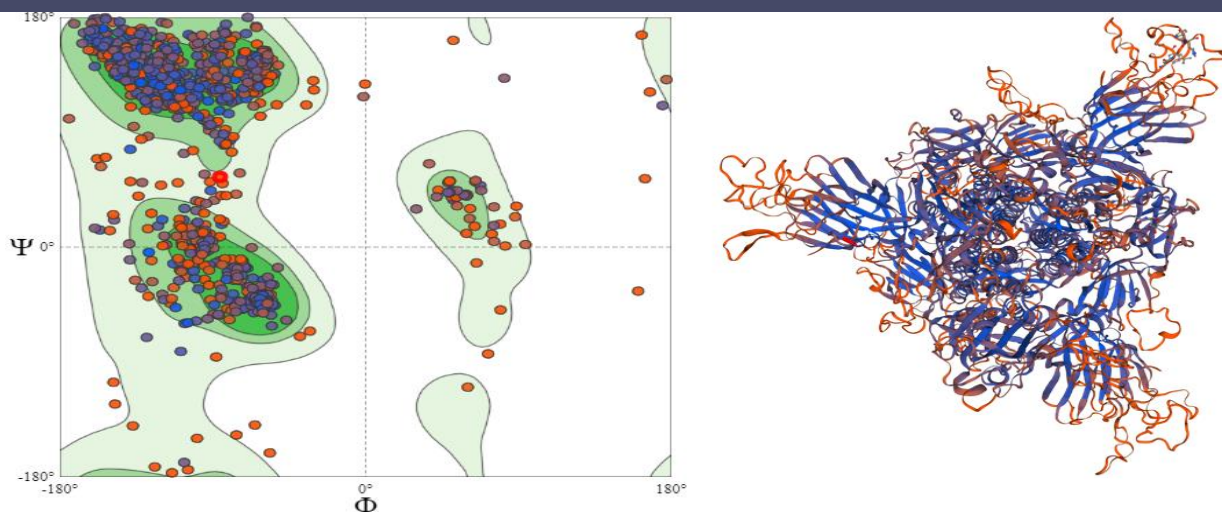
ACE2 is lower than SARS -CoV S-RBD.

Figure 7.Ramachandran plot using MolProbity version 4.4 Results obtained of complex system of 2019-nCoV spike glycoprotein -RBD , Created: Wed 29 Apr 2020, 11:07
https://swissmodel.expasy.org/assess/7zSfLH/01

To cross verify the complex system of SARS-CoV Spike glycoprotein docking-RBD binding to ACE2 at sequence identity 86.92% ,we performed Ramachandran plot to know the interaction level. The plot confirms the interaction of SARS-CoV Spike glycoprotein with human ACE2 .

3.4 Simulation of COVID-19-SARS-CoV to check human receptor binding capability

After identification of native like binding conformations of S-RBD with ACE2 it enables to compare receptor binding capability of COVID-19 with SARS -CoV. The evaluated energy scores experimentally determined structure of SARS-CoV which defines binding free energy scale, on this basis we could enable to calculate relative receptor binding strength of COVID-19

The standard Monte Carlo docking simulation was applied within statistical thermodynamics, we considered sampled binding conformations. The thermodynamically probable in the binding process of S-glycoprotein to ACE2.The binding affinity (binding free energy) is thermodynamic average over all possible binding modes two interacting active components.

Therefore, binding free energy is approximated by thermodynamic average of the binding energies of all the sampled low energy conformations can be written as:

11

$$< \Delta L > \approx \frac{\sum_i^N \Delta L_i \exp\left(-\frac{\Delta L_i}{RT}\right)}{\sum_i^N \exp\left(-\frac{\Delta L_i}{RT}\right)} \qquad [1]$$

Where N total number of low energy conformations performed in docking simulation within Monte Carlo simulation method. R-ideal gas constant-absolute temperature 298 K for room temperature

From above equation [1] and all sampled binding conformations in fig.5.A binding free energy of SARS-CoV S-RBD to ACE2 extracted is:

$$< \Delta L^{SARS} >= -39.20 \text{ kcal} \cdot \text{mole}^{-1} \qquad [2]$$

Similarly, we computed values COVID-19 S-RBD to ACE2 is

$$< \Delta L^{COVID-19} >= -28.66 \text{ kcal} \cdot \text{mole}^{-1} \qquad [3]$$

After comparing extracted values, we able to predict that COVID-19 S-RBD reaches about ~80 % of binding strength of SARS CoV S-RBD to ACE2.When receptor binding of S glycoprotein plays a vital role for viral transmission to human host cell. However, such binding strength signifies that COVID-19 able to infect humans as observed in SARS-CoV. The modeling of COVID-19 [25] performed to investigate binding affinity of S-RBD to ACE2.They successfully simulated two single complex conformations and evaluated the free binding energies of two CoVs (including crystal and predicted structure).The affinity of COVID-19 S-RBD to ACE2 found in the order of ~65% for SARS-CoV and Huang [17] predicted the binding strength reaches ~ 73% SARS-CoV S-RBD to ACE2.Our modeling data i.e. extracted free binding energies (~80%) surely enhanced with difference of 15% and 07% .

Surely, our extracted values from low energy binding conformations were obtained by performing number of thermodynamical probable .The further investigation is necessary on experimental measurements to verify accuracy of performed simulation method.

4. Conclusions

The newly emerged Coronavirus-19 (COVID-19) originated from Hunan seafood market at Wuhan, China. The source of COVID-19 or SARS-CoV-2 is not confirmed.However,our sequence based analysis suggests that bats [Fig.2A] as the key reservoir. The DNA recombination was found involved at S-glycoprotein which assorted SARS-CoV with RBD of another β CoV.This could be the reason for cross species transmission responsible for rapid infection.

This finding confirm that COVID-19 reaches about 86% of ACE2 receptor binding strength of SARS-CoV.The COVID-19 S-glycoprotein encoded by its genome has processed human receptor binding ability approximately close to SARS-CoV. The COVID-19 virus likely to bind to human ACE2 and in due course it transmits from human to human.

The said investigation of S-Glycoprotein RBD of COVID-19 or SARS-CoV-2 via ACE2 provides post genome analysis of protein-protein interaction for rapid assessing transmission of infected patients by deadly CoVID-19. The scientific data extracted implies early guidance to control and viral prevention of CoVID-19.To improve our computational predications, we expect experimental validation for deeply understanding human transmission of COVID-19 via receptor binding interactions.

Declaration of competing Interest

Additional information:

The investigated related data will be made available on request to baliram.lone@aggiemail.usu.edu

References:

1.   Fehr, A.R. and S. Perlman, *Coronaviruses: An Overview of Their Replication and Pathogenesis*, in *Coronaviruses: Methods and Protocols*, H.J. Maier, E. Bickerton, and P. Britton, Editors. 2015, Springer New York: New York, NY. p. 1-23.
2.   Cordes, A.K. and A. Heim, *Rapid random access detection of the novel SARS-coronavirus-2 (SARS-CoV-2, previously 2019-nCoV) using an open access protocol for the Panther Fusion.* J Clin Virol, 2020. **125**: p. 104305.
3.   Gross, M., *Virus outbreak crosses boundaries.* Current Biology, 2020. **30**(5): p. R191-R194.
4.   Chan, J.F.-W., et al., *Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan, Emerging Microbes & Infections*. Vol. 9:1. 2020.
5.   Huang, C., et al., *Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China.* The Lancet, 2020. **395**(10223): p. 497-506.
6.   Wang, R., et al., *Emergence of SARS-like coronavirus poses new challenge in China.* Journal of Infection, 2020. **80**(3): p. 350-371.
7.   situation, C.d.C.-o., *https://www.who.int/emergencies/diseases/novel-coronavirus-2019.* 3 April 2020, 23:17 GMT+5:30.
8.   Li, F., *Structure, Function, and Evolution of Coronavirus Spike Proteins.* Annual Review of Virology, 2016. **3**(1): p. 237-261.
9.   Arden, K.E., et al., *New human coronavirus, HCoV-NL63, associated with severe lower respiratory tract disease in Australia.* Journal of Medical Virology, 2005. **75**(3): p. 455-462.
10.   Assiri, A., et al., *Epidemiological, demographic, and clinical characteristics of 47 cases of Middle East respiratory syndrome coronavirus disease from Saudi Arabia: a descriptive study.* Lancet Infect Dis, 2013. **13**(9): p. 752-61.

11.   Babcock, G.J., et al., *Amino acids 270 to 510 of the severe acute respiratory syndrome coronavirus spike protein are required for interaction with receptor.* Journal of virology, 2004. **78**(9): p. 4552-4560.

12.   Chen, Y., et al., *Structure analysis of the receptor binding of 2019-nCoV.* Biochem Biophys Res Commun, 2020.

13.   Xiao, X., et al., *The SARS-CoV S glycoprotein: expression and functional characterization.* Biochemical and Biophysical Research Communications, 2003. **312**(4): p. 1159-1164.

14.   Wu, F., Zhao, S., Yu, B., Chen, Y.-M., Wang, W., Hu, Y., Song, Z.-G., Tao, Z.-W., Tian, J.-H., Pei, Y.-Y., et al, *Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1, complete genome.* 2020.

15.   Chen, V.B., et al., *MolProbity: all-atom structure validation for macromolecular crystallography.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 1): p. 12-21.

16.   Alford, R.F., et al., *The Rosetta All-Atom Energy Function for Macromolecular Modeling and Design.* J Chem Theory Comput, 2017. **13**(6): p. 3031-3048.

17.   Huang, Q. and A. Herrmann, *Fast assessment of human receptor-binding capability of 2019 novel coronavirus (2019-nCoV).* bioRxiv, 2020: p. 2020.02.01.930537.

18.   He, R., et al., *Analysis of multimerization of the SARS coronavirus nucleocapsid protein.* Biochemical and Biophysical Research Communications, 2004. **316**(2): p. 476-483.

19.   Song, H.-D., et al., *Cross-host evolution of severe acute respiratory syndrome coronavirus in palm civet and human.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(7): p. 2430-2435.

20.   Guan, Y., et al., *Isolation and Characterization of Viruses Related to the SARS Coronavirus from Animals in Southern China.* Science, 2003. **302**(5643): p. 276.

21.   Ge, X.-Y., et al., *Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor.* Nature, 2013. **503**(7477): p. 535-538.

22.   Li, F., et al., *Structure of SARS Coronavirus Spike Receptor-Binding Domain Complexed with Receptor.* Science, 2005. **309**(5742): p. 1864.

23.   Qian, B., et al., *High-resolution structure prediction and the crystallographic phase problem.* Nature, 2007. **450**(7167): p. 259-264.

24.   Chaudhury, S., et al., *Benchmarking and analysis of protein docking performance in Rosetta v3.2.* PLoS One, 2011. **6**(8): p. e22477.

25.   Xu, X., et al., *Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission.* Science China Life Sciences, 2020. **63**(3): p. 457-460.