



# IN-SILICO DRUG DISCOVERY FOR COVID19 BY TARGETING SPIKE GLYCOPROTEIN OF SARS COV - 2 (WUHAN CORONA VIRUS 2019 OUTBREAK) AGAINST THE DOCKING ANALYSIS WITH STRUCTURE PREDICTED HUMAN 'ACE2-FC REGION OF IgG1' FUSION PROTEIN AS A PROTEIN BASED DRUG.

<sup>1</sup> PANDIYA MUTHU KUMARAN B <sup>2</sup> KALIDAS VIJAYAKUMAR

<sup>1</sup>Research Associate, <sup>2</sup>Research Scholar,

<sup>1</sup>Bio Research Unit, <sup>1</sup>KBR Industrial solutions, Madurai, India. <sup>2</sup>School of Biotechnology, <sup>2</sup>MKU, Madurai, India.

**Abstract:** The globe was frightened with a newly isolated virus named as sars cov 2 formally called as novel corona virus 2019. This work is predicting a protein based drug for sars cov 2. A virus was outbreaks at wuhan city of china on December 2019. The RNA sequence of that virus was published in NCBI. (NCBI Ref Sequence: NC\_045512.2) We collected the surface glycoprotein sequences from NCBI and genbank and predicted the protein structure with swiss model and phyre 2 protein structure prediction tools. The predicted spike glycoprotein was have a docking function with the human ACE2 receptor. We predicted the 3D structure of ACE2-FC region of IgG1 fusion protein and docked with virus spike glycoprotein and found a good docking potential of fusion protein with Sars Cov 2. ACE2- FC region of IgG1 will be use as a protein drug against SarsCoV 2.

**Index Terms** – SARS CoV 2, Novel Corona Virus, ACE2-FC REGION IgG1, COVID19, INSILICO DRUG DISCOVERY.

## I. INTRODUCTION

AT present (2020) outbreak of lower respiratory tract infections with respiratory distress syndrome of an animal coronavirus to humans resulting in a major epidemic. The International Committee on Taxonomy of Viruses and The Coronavirus Study Group (CSG) is an international organization accountable for developing the official classification of viruses and taxonomy of the Coronaviridae family. This organization assessing the innovation of the newly isolated human viral pathogen named the novel coronavirus 2019. Studies on phylogenetic analysis, taxonomical and accepted official practice, the CSG formally recognizes this virus as a associate to severe acute respiratory syndrome coronaviruses (SARS-CoVs) of the species Severe acute respiratory syndrome-related coronavirus and nominate it as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The spectrum of clinical manifestations associated with SARS-CoV-2 infections in humans remains to be resolved. The independent infectious transmission of SARS-CoV and SARS-CoV-2 intimates the need for studying the entire (virus) species. Research focused on individual pathogenic viruses of immediate importance. This research will improve our understanding of virus-host interactions in a changing environment and increase our state of readiness for future outbreaks. (1)

The present status of COVID—19 is researched and the disease is forecasted. The statistical peak of the epidemic is different in three regions, Wuhan, Hubei province except form Wuhan, and mainland China except Hubei province. In Wuhan, the epidemic reached a peak but the reported number of newly infected patients oscillates largely. (2)

In 2003, a newly identified illness termed as severe acute respiratory syndrome (SARS) spread rapidly through the Globe. At that time a new coronavirus named SARS-CoV. And it was identified and isolated as the SARS viral pathogen which caused severe pneumonia and acute, often lethal, lung failure. The infected individuals are affected like influenza such as the Spanish flu and the emergence of new respiratory disease viruses that have caused high lethality developing from acute lung failure. In research on cell lines, the human angiotensin-converting enzyme 2 (ACE2) has been identified as a potential SARS-CoV receptor. (3) ACE2 is involved in cardiovascular and renal physiology, diabetes, lung disease, pregnancy and also ACE2 serves as a receptor for SARS and NL63

coronaviruses. (4)The angiotensin-converting enzyme (ACE)- related carboxypeptidase is called ACE2. It is a type I integral membrane protein have a sequence of 805 amino acids. It contains one HEXXH + E zinc-binding consensus sequence(5)

The Fc region is the tail region of an antibody and that region has a direct effect on cell surface receptors called Fc receptors and also has a reaction with few proteins of the complement system. Its Natural ability to allow antibodies to activate the immune system. In IgG, IgD and IgA antibody isotypes have the Fc region is self-possessed of two identical protein fragments these are obtained from the second and third changeless constant domains of the antibody's two heavy chains. But the IgM and IgE Fc regions contain three heavy chain constant domains - CH domains 2-4 in each polypeptide chain. This Fc regions of IgGs carry a highly preserved N-glycosylation site. The Glycosylation of the Fc fragment is more important for Fc receptor-mediated activity. The N-glycans bound to this site are mainly core-fucosylated biantennary structures of the complex type. Small amounts of these N-glycans also take bisecting GlcNAc and  $\alpha$ -2,6 linked sialic acid residues. (6)

making a fusion protein of ' Fc-ACE2 fusion protein' can use for therapeutic protein drug against sars Cov 2 Wuhan coronavirus by targeting its spike glycoprotein. And this Fusion complex protein may induce the immune response on the viral affected humans. This work is targeting a protein-based drug for COVID19 disease. By using bioinformatics tools.

## II.RESEARCH METHODOLOGY

### 2.1 Structure prediction for SARS CoV – 2 SPIKE GLYCOPROTEIN

The National center for biotechnology information (NCBI) provides biological information of the researched and collected data which includes genbank®.(7)From the NCBI sequences we have collected the gene of sars Cov 2 surface glycoprotein as a FASTA format. After that, we predicted the structure of that protein by using swiss model homology modeling tool (8)

### 2.2 STRUCTURE PREDICTION FOR ACE2

Again we go with ACE2 sequences from NCBI genbank® and predicted structure by using the swiss model and phyre 2 structure prediction tools. The predicted structure was saved as pdb format. (7)

### 2.3 DOCKING CONFIRMATION OF ACE2 AGAINST SARS CoV2 SPIKE GLYCO PROTEIN WITH HEX8.0 TOOL

Predicted Glycoprotein is taking as a receptor and docking studies with ACE2 receptor protein by using Hex 8.0 docking tool. Hex 8.0 is one of the best docking tool for analyzing protein-protein interactions.(9)

### 2.4 MAKING OF IN SILICO MODEL OF ACE2-FC REGION IgG1 FUSION PROTEIN

In modern days researchers are focused on the FC region IgG1 based fusion proteins as the drug against disease.Fc region IgG1 has a vital immunological response and it involves in the human immune system for curing the diseases. The proposed amino acid sequence of FC region IgG1 and human ACE2 receptor are taken and they are fused as amino acid sequence. FASTA format of the sequence was taken for protein structure prediction,phyre 2 protein structure prediction tool was used to predict the structure of that fusion protein. (10)

### 2.5Insilico docking studies with Fusion ACE2-FC region of IgG1 protein against SARS CoV 2 virus Glycoprotein

Again we run the docking analysis of the fusion protein and Sars CoV -2 spike Glycoprotein at Hex 8.0. The docking results are tabled and the interaction between ACE2-Fc region of IgG1 is the conformation of the fusion protein's potential ability against the virus activity of binding on host cell is monitored. (11)

## III.RESULTS AND DISCUSSION

### 3.1 Homology modeling and Structure prediction of spike glycoprotein of Sars CoV2.

The swiss model and phyre 2 tool used for predict the structure of spike glycoprotein from the surface glycoprotein amino acid sequence. (NCBI Ref Sequence: NC\_045512.2).The researchers are already found the nucleic acid sequences and amino acid sequences. We predicted protein structure by using the homology modeling and saved it by PDB file format. Swiss – model structure prediction was done and the predicted spike glycol protein was downloaded as a PDB file format for further docking studies. The genbank® Gene id for the amino acid sequence is 43740568.

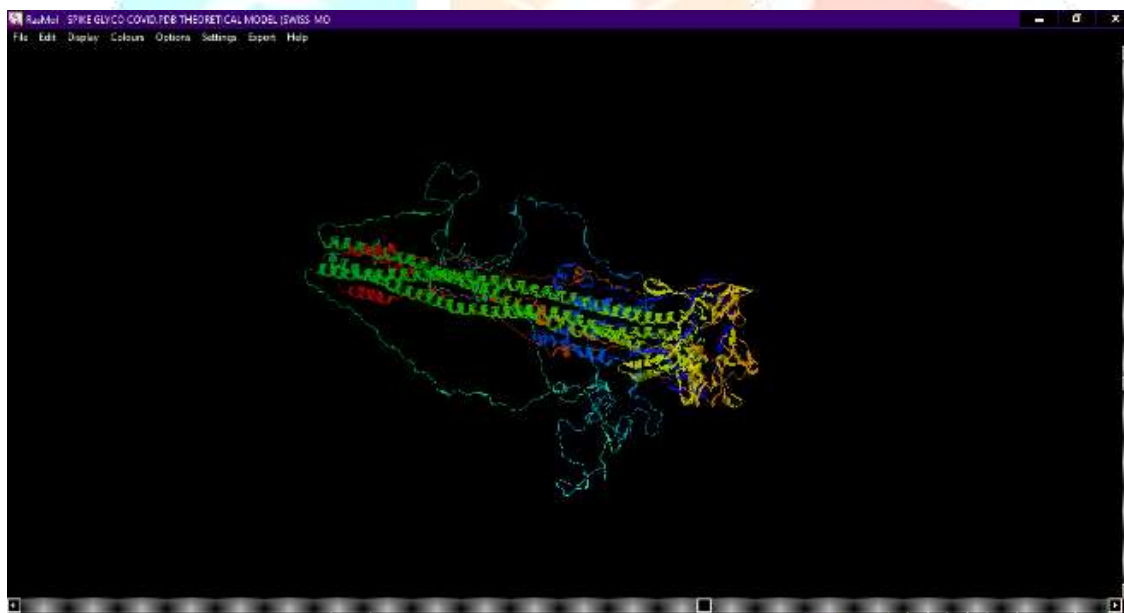
Fig 1: NCBI Published Amino acid sequence of spike glycoprotein of SarsCoV 2 (NCBI Ref sequence :NC\_045512.2)

```

CDS
/locus_tag="bu200_gp04"
/db_xref="GeneID:43748568"
21563..25384
/genes="S"
/locus_tag="GU280_gp02"
/note="structural protein; spike protein"
/codon_start=1
/product="surface glycoprotein"
/protein_id="YP_009724398.1"
/db_xref="GeneID:43748568"
/translation="MPVFLVLLPLVSSQCVMLTTRTQLPPAYTNSFTRGVVYYPQKVFRR
SSVLHSTQQLFLPFFSNWTFHAIHVSGMTKRFDNPVLPFNDGVVYFASTKSNIIIR
GWIFGTLDLSTQSLIVMNAIVVIVKVECFQCNDFLGVVYHKMKSNHSEFRVY
SSAMNCTFEYVSQPLNDLEGGQGNFKNLREFVFNKIDGVFKIYSKHTPILNLRDLQ
GFSALEPLVDLPIGINITRFQTLALHRSYLTPGSSSGHTAGAAAYVGYLQPRFLL
LKYVENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITN
LCPFGVEFNATRFASVYAMNRKRISNCVADYSVLYNSASFSTFKCYGVSPKLNLDLCLF
TINVYDSFVIRGDEVQIAPGQTGKIADYNYKLPDDFTGCVIAMSMLDSKVGGMN
YLYRFRKSNLKPFRDISEIVQAGSTPCNGVEGFNCFPLQSYGFQPTNGVGYQPY
RVVWLSFELHAPATVCGPKKSTNLVKNKCVNFNGLTGTGLTESNKKFLPFQDFG
RDIADTTDAVRDQQLLELDITPCSFGGVSITPGTNTSNQAVLYQDMCTEVPVAI
HADQLPTWRVYSTGNSVFTQTRAGCLIGAEHVMNSYECDIPIGAGICASYQTQNSPR
RARSVASQSIAYTHSLGAENSVAYSMMISAIPTNFTISVTEILPVSHTKTSNOCIM
YICGDSTECNLLLYGGSFCTQNRALTGIAVEQDKNTQEVFAQVQIYKTPPIKDFG
GFNFSQILPDPSPKSKRSFIEDLLFNKVTIADAGFIKQVGDCLGDIARDLICAQKFN
GLTVLPPLLTDEHIAQYTSALLAGTITSGHTFGAGAALQIPFAMQVAYRFNGIGVTON
VLYENQKLIANQVFNIAIGKIQDSLSSTASALGKLDVNIQAQALNTLVKQLSSNFGA
ISSVNDITLSRLDKVEAEVQIDRLITGRQLSQTIVYVQQLIRAAEIRASAILAATKHS
ECVLGQSKRVDFCGKGYHLMSFRQSAPHGWFLHVTYVPAQEKNFITAPAICHGKAH
FPREGVFSNGTHIVTQQRNFYEPQIITDNTFVSGNCDVIGIWMNTVYDPLQPELD
SFKEELDKYFKNHTSPVDLGDISGINASVWNIQKEIDRLNEVAKNLESILDLQELG
KYEYIKMHWIHLGFIAGLIAIVPMTIIMLCCHTSCCSCLKGCCSCGCKFDEDDSE
PVLKGVKLVHT"

```

Fig 2:Swiss model Predicted structure of spike glycoprotein of SarsCoV 2.



### 3.2 Structure prediction of ACE2

Similarly, we found the results for in silico structure of homology modeled ACE2 protein. (NCBI Reference Sequence: NC\_000023.11) and genbank® GeneID:59272.

Fig 3: NCBI genbank® provided the amino acid sequence of the human ACE2 receptor.3.2 Structure prediction of ACE2  
 Similarly, we found the results for in silico structure of homology modeled ACE2 protein. (NCBI Reference Sequence: NC\_000023.11) and genbank® GeneID:59272.

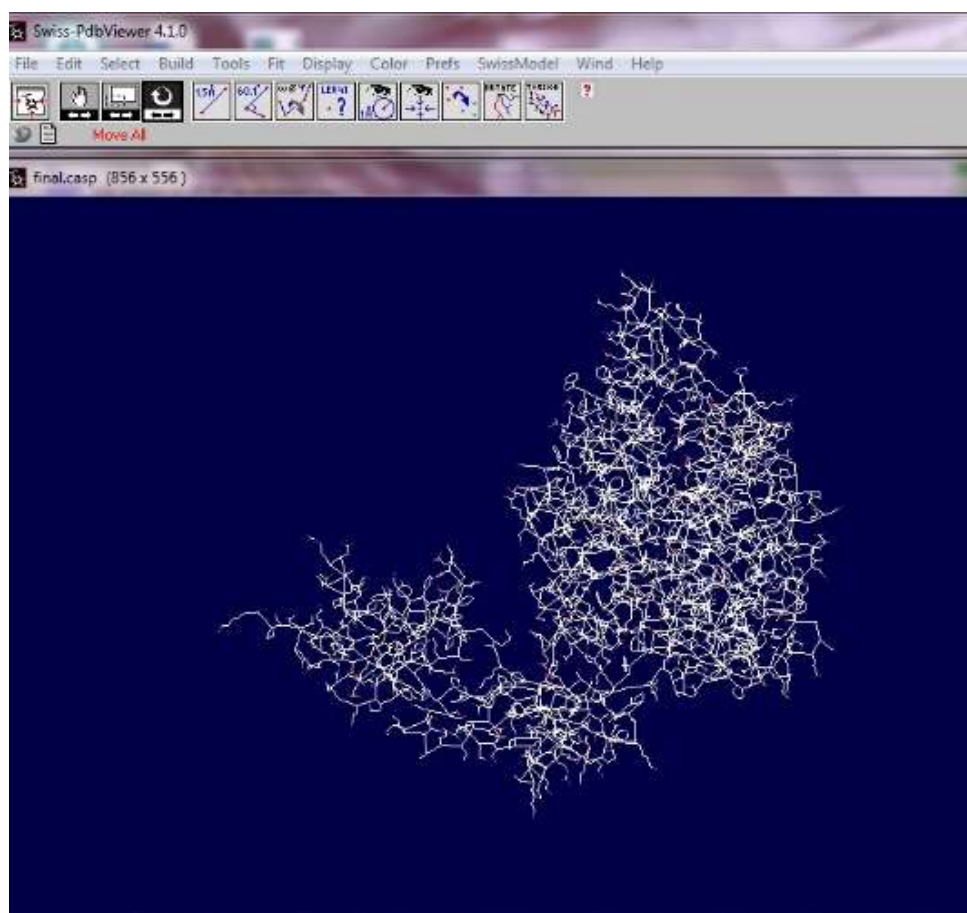
Fig 3 : NCBI genbank® provided the amino acid sequence of the human ACE2 receptor.

```

>human ACE2 sequence
1  mssswllls  lvavtaaqst  ieeqaktfld  kfnheaedlf  yqsslaswny  ntniteenvg
61 nmnnagdkws  aflkegstla  qmyplqeiqn  ltvklqlqal  qqngssvlse  dkskrlnlil
121 ntmstiyystg  kvcpndnpqe  clilepglne  imansldyne  rlwaweswrs  evgkqlrply
181 eeyvvlknem  aranyedyg  dywrgdyevn  gvdgydysrg  qliedvehtf  eeikplyeh1
241 hayvraklmn  aypsyispig  clpahllgdm  wgrfwtlnys  ltvpgqkpn  idvtdamvdq
301 awdaqrifke  aekffvsvgl  prmtqgfwen  smldpgnvg  kavchptawd  lgkgdfri1m
361 ctkvtmddf1  tahhemghiq  ydmayaaqpf  llrnganegf  heavgeimsl  saatpkh1ks
421 igllspdfqe  dneteinf1l  kqaltivgt1  pftymlekwr  wmvfkgeipk  dqwmkkwem
481 kreivgvvpe  vphdetycdp  aslfhvsndy  sfiryrt1  yqgfgealc  qaakhegplh
541 kcdinstea  gqklfnmlr1  gksepwt1a1  envvgaknmn  vrpllnyfef  lftwlkdqnk
601 nsfvgwstdw  spyadqsikv  rislksalgd  kayewndnem  y1frssvaya  mrqyflkvkn
661 qmilfgeedv  rvanlkpris  fnffvtapkn  vsdiiprtev  ekairmsrsr  indafrlndn
721 sleflgiqpt  lgppnqppvs  iwlvfvgvm  gvivvgivil  iftgirdrkk  knkarsgenp
781 yasadiskge  nngfqntdd  vqtsf

```

Fig 4:PHYRE 2Predicted structure of ACE2 receptor



### 3.3 In silico docking studies of ACE2 with SARS CoV 2 spike glycoprotein

Hex8.0 is a potential tool for identifying the docking solution between protein and it predicts the protein-protein interactions. Fig 5: Docking analysis with spike glycol protein of SarsCoV 2 and ACE2

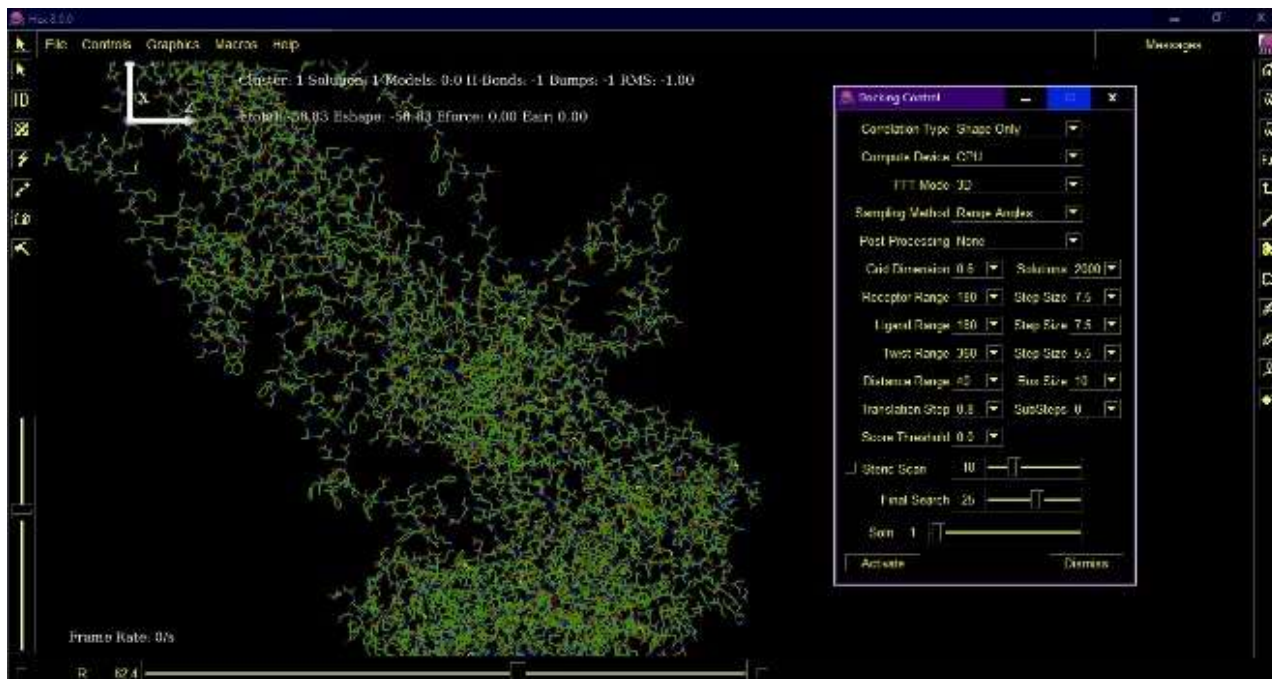
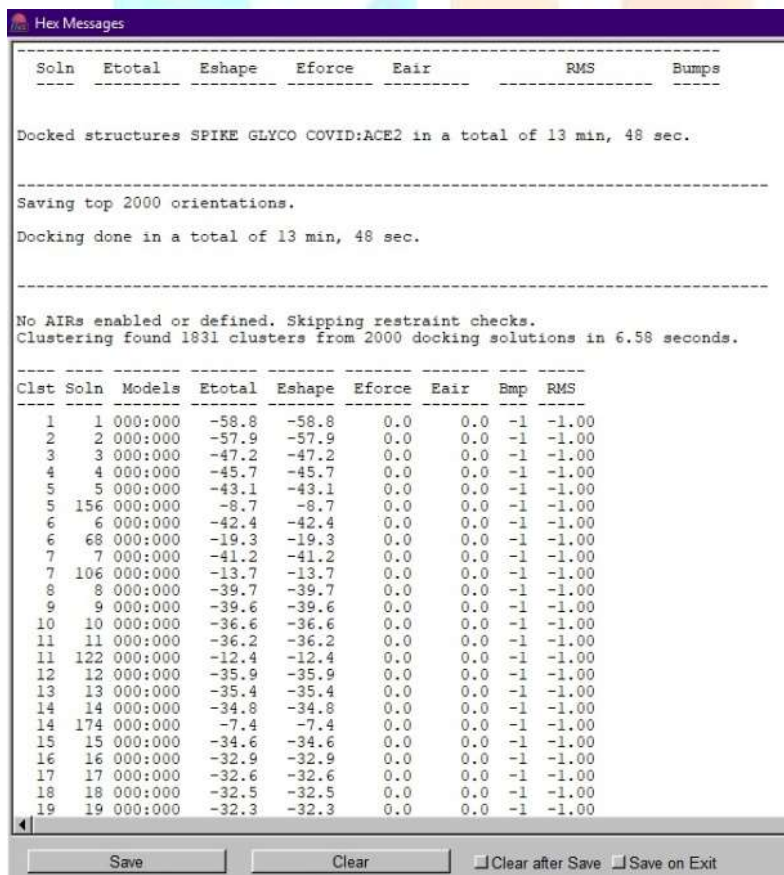


Fig 6: Hex8.0 message result of spike glycoprotein with the ACE2 receptor.



We got docking results between these proteins. As a result, we found the 1831 clusters in 2000 docking solutions.

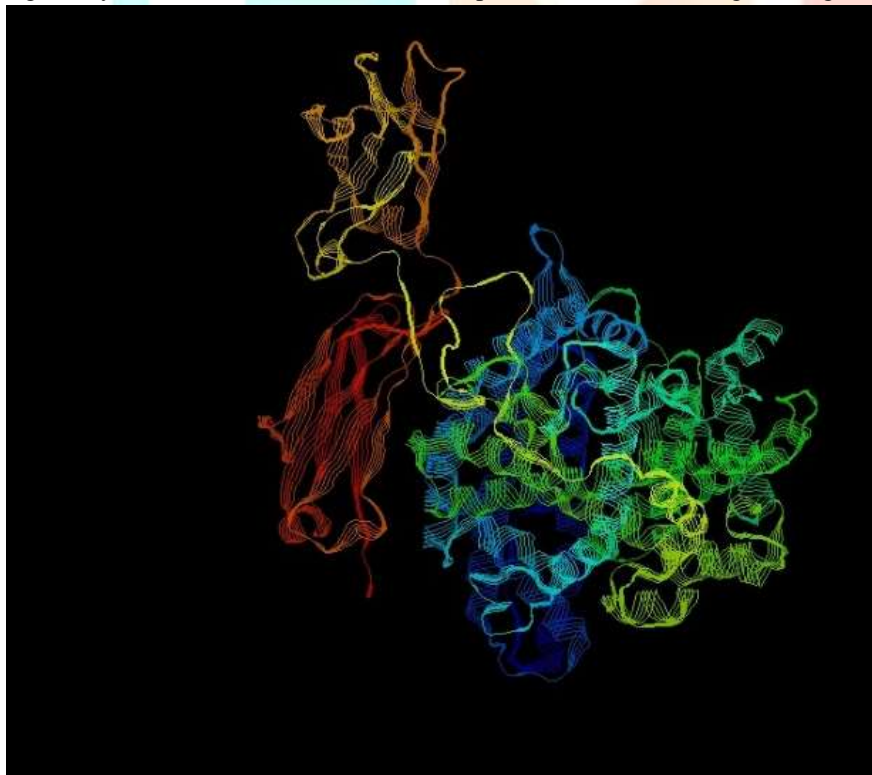
### 3.4 Structure prediction of ACE2-FC region Fusion protein

By formulating the new fusion protein by using the Fc region of IgG1 amino acid sequence, signal peptides and amino acid sequence of ACE2 the sequence makes as a FASTA format and taken into the homology modeling by using phyre2 structure prediction tool. The unique job identifier phyre2 id is a78d9c9c7b8a2ef4.

Fig 7:FASTA FORMAT FOR ACE2-Fc Fusion protein

```
>FUSIONMODEL
MDWIWRILFLVGAATGAHS
QSTIEEQAKTFLDKFNHEAEDLFYQSSLASWNYNTNITEENVQNMNAGDKWSAFLKEQSTLA
QMYPLQEIQLTVLQLQALQQNGSSVLSEDKSKRLNTILNTMSTIYSTGKVCNPDNPQE
CLLLEPGLNEIMANSLDYNERLWAWESWRSEVVKQLRPLYEEYVVLKNEMARANHYEDYG
DYWRGDYEVNGVDGYDYSRGLIEDVEHTFEEIKPLYEHLHAYVRAKLMNAYPSYISPIG
CLPAHLLGDMWGRFWTNLYSLTVPPFGQKPNIDVTDAMVDQAWDAQRIFKEAEKFFVSVGL
PNMTQGFWENSMLTDPGNVQKAVCHPTAWDLGKGDFRILMCTKVTMDDFLTAHHEMGHIQ
YDMAYAAQPFLLRNGANEGFHEAVGEIMSLSAATPKHLKSIKGLLSPDFQEDNETEINFL
KQALTIIVGTLPTTYMLEKWRWVFKGEIPKDQWMKKWEMKREIVGVVEPVPHDETYCDP
ASLFHVSNDYSFIRYYTRTL YQFQFQALCQAAKHEGPLHKCDISNSTEAGQKLFNMLRL
GKSEPWTLAENVVGAKNMNVRPLLNYFEPLFTWLKDQNKNSFVGWSTDWSPYAD
EPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDITLMIKSRTEPEVTCVVVDV
HEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGK
EYCKKVSNAKALPAPIEKTIISKAKGQPREPQVYTLPPSRDELTKNQVSLTCL
LVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRW
QQGNVVFSCSVMHEALHNHYTQKSLSLSPGK
```

Fig 6: Phyre2 Predicted structure of Fusion protein of ACE2 - Fc region of IgG1.



### 3.5 In silico docking studies of ACE2-Fc region of IgG1 fusion protein with SARS CoV 2 spike glycoprotein

Fig 7: Fusion protein (ACE2 – Fc region of IgG1) docked with spike glycoprotein of SarsCoV 2.

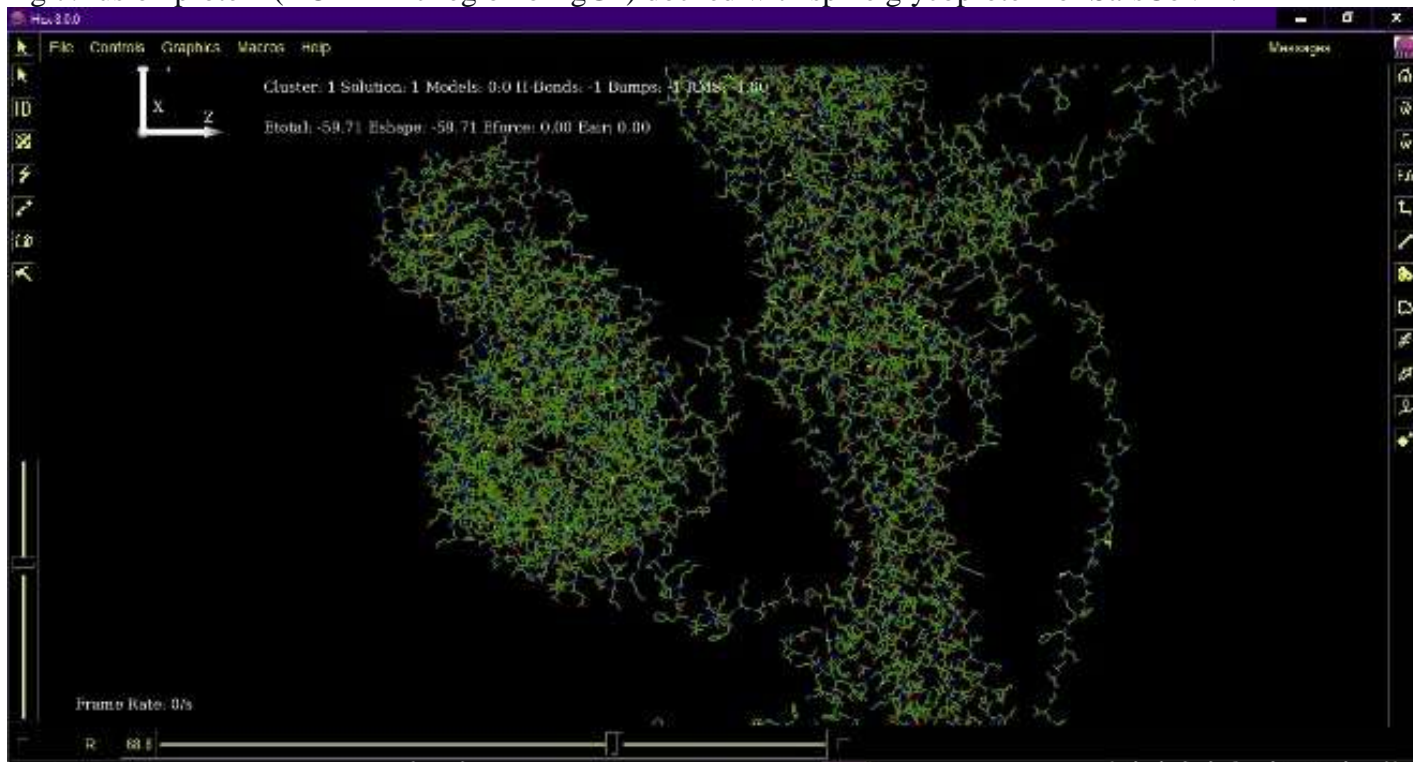
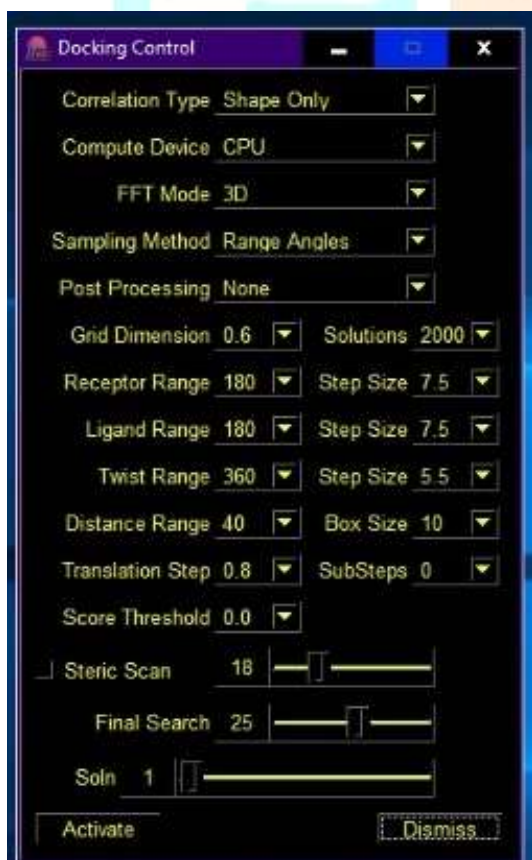


Fig 8: Control parameters for docking of fusion with spike protein.



The image shows a screenshot of the Docking Control panel. The panel is titled "Docking Control" and contains various parameters for docking. The parameters are as follows:

Parameter	Value
Correlation Type	Shape Only
Compute Device	CPU
FFT Mode	3D
Sampling Method	Range Angles
Post Processing	None
Grid Dimension	0.6
Solutions	2000
Receptor Range	180
Step Size	7.5
Ligand Range	180
Step Size	7.5
Twist Range	360
Step Size	5.5
Distance Range	40
Box Size	10
Translation Step	0.8
SubSteps	0
Score Threshold	0.0
Steric Scan	18
Final Search	25
Soln	1

Buttons: Activate, Dismiss

Fig 9: Final docking results with HEX Message.

```

Hex Messages
Solution buffer reached 33216/200000 = 16.6% occupancy with no culling.
Starting guess not found in top 33216 solutions.
Emin = -59.71, Emax = -562.54

Docking correlation summary by RMS deviation and steric clashes
-----
Soin  Etotol  Eshape  Eforce  Eair  RMS  Bumps
-----

Docked structures FUSION_1 :SPIKE GLYCOmodel_04 in a total of 8 min, 32 sec.

Saving top 2000 orientations.

Docking done in a total of 8 min, 32 sec.

No AIRs enabled or defined. Skipping restraint checks.
Clustering found 1875 clusters from 2000 docking solutions in 16.63 seconds.

Clst  Soin  Models  Etotol  Eshape  Eforce  Eair  Bmp  RMS
-----
1 1 000:000  -59.7  -59.7  0.0  0.0  -1  -1.00
2 2 000:000  -59.2  -59.2  0.0  0.0  -1  -1.00
3 3 000:000  -56.7  -56.7  0.0  0.0  -1  -1.00
4 4 000:000  -56.1  -56.1  0.0  0.0  -1  -1.00
5 5 000:000  -53.7  -53.7  0.0  0.0  -1  -1.00
6 6 000:000  -50.9  -50.9  0.0  0.0  -1  -1.00
7 7 000:000  -50.4  -50.4  0.0  0.0  -1  -1.00
8 8 000:000  -50.1  -50.1  0.0  0.0  -1  -1.00
9 9 000:000  -49.4  -49.4  0.0  0.0  -1  -1.00
10 10 000:000  -48.6  -48.6  0.0  0.0  -1  -1.00
11 11 000:000  -48.5  -48.5  0.0  0.0  -1  -1.00
12 12 000:000  -46.3  -46.3  0.0  0.0  -1  -1.00
13 13 000:000  -46.2  -46.2  0.0  0.0  -1  -1.00
14 14 000:000  -46.0  -46.0  0.0  0.0  -1  -1.00
15 15 000:000  -45.8  -45.8  0.0  0.0  -1  -1.00
16 16 000:000  -45.5  -45.5  0.0  0.0  -1  -1.00
17 17 000:000  -45.4  -45.4  0.0  0.0  -1  -1.00
18 18 000:000  -45.1  -45.1  0.0  0.0  -1  -1.00
19 30 000:000  -40.6  -40.6  0.0  0.0  -1  -1.00

```

As per the protein-protein docking studies with the Hex8.0 docking tool, the fusion protein ACE@ - Fc region of IgG1 docked with 1875 clusters from 2000 docking solutions in 16.63 seconds. With the good E-total of -59.7.

SARS CoV 2 SPIKE GLYCOPROTEIN and ACE2-FC region IgG1 both have a bonding and docking ability. This protein complex as a potential neutralization drug against SARS CoV 2 VIRUS. This research will help to produce future therapeutic solutions for this kind of future Coronavirus outbreaks.

### Acknowledgment

We sincerely thank Dr.Sankaranyar P from Madurai Kamarajar University for his support to making this research work.

### REFERENCES

- [1] Gorbalenya, A. E., Baker, S. C., & Baric, R. S. 2020. Severe acute respiratory syndrome-related coronavirus: the species and its viruses—a statement of the coronavirus study group. *bioRxiv preprint first posted online February 11, 2020*. 2020: 2020.2002.2007.937862. doi: 10.1101/2020.02.07.937862. Accessed February, 12.
- [2] Tomie, T. 2020. Understanding the present status and forecasting of COVID—19 in Wuhan. *medRxiv*.
- [3] Kuba, K., Imai, Y., Rao, S., Gao, H., Guo, F., Guan, B., ... & Bao, L. 2005. A crucial role of angiotensin converting enzyme 2 (ACE2) in SARS coronavirus-induced lung injury. *Nature medicine*, 11(8), 875-879.
- [4] Hamming, I., Cooper, M. E., Haagmans, B. L., Hooper, N. M., Korstanje, R., Osterhaus, A. D., ... & van Goor, H. 2007. The emerging role of ACE2 in physiology and disease. *The Journal of Pathology: A Journal of the Pathological Society of Great Britain and Ireland*, 212(1), 1-11.
- [5] Towler, P., Staker, B., Prasad, S. G., Menon, S., Tang, J., Parsons, T., ... & Patane, M. A. 2004. ACE2 X-ray structures reveal a large hinge-bending motion important for inhibitor binding and catalysis. *Journal of Biological Chemistry*, 279(17), 17996-18007.
- [6] Peipp, M., Lammerts van Bueren, J. J., Schneider-Merck, T., Bleeker, W. W., Dechant, M., Beyer, T., ... & Parren, P. W. 2008. Antibody fucosylation differentially impacts cytotoxicity mediated by NK and PMN effector cells. *Blood, The Journal of the American Society of Hematology*, 112(6), 2390-2399.
- [7] Sayers, E. W., Beck, J., Brister, J. R., Bolton, E. E., Canese, K., Comeau, D. C., ... & Kitts, P. A. 2020. Database resources of the national center for biotechnology information. *Nucleic acids research*, 48(D1), D9.
- [8] Zhang, Y. Z., Wu, F., Chen, Y. M., Pei, Y. Y., Xu, L., Wang, W., ... & Song, Z. G. 1988. Nucleotide Accession No. MN908947. 1, Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1, complete genome. Bethesda (MD): National Library of Medicine (US). National Center for Biotechnology Information.



[9] Öchsner, A. 2020. Three-Dimensional Elements. In Computational Statics and Dynamics (pp. 385-415). Springer, Singapore.

[10] Wang, Z., Zhu, J., & Lu, H. 2020. Antibody glycosylation: impact on antibody drug characteristics and quality control. Applied Microbiology and Biotechnology, 1-10.

[11] Dehghani, B., Hashempour, T., & Hasanshahi, Z. 2020. Interaction of human herpesvirus 8 viral interleukin-6 with human interleukin-6 receptor using in silico approach: the potential role in HHV-8 pathogenesis. Current Proteomics, 17(2), 107-116.

