



IJCRR

Section: Healthcare

Sci. Journal Impact

Factor: 6.1 (2018)

ICV: 90.90 (2018)



Copyright@IJCRR

In-Silico Analysis of Interaction of Human Tetherin Protein with SARS-COV-2ORF7A Proteins and its Mutants

Kaveri Bisht¹, Kumud Pant², Navin Kumar³, Akshara Pande⁴, Bhasker Pant⁵, Devvret Verma⁶

¹Department of Life Sciences, Graphic Era (Deemed to be University), Dehradun, Uttarakhand, India; ^{2,3,6}Department of Biotechnology, Graphic Era (Deemed to be University), Dehradun, Uttarakhand, India; ⁴Department of Computer Sciences, Graphic Era Hill University, Dehradun, Uttarakhand, India; ⁵Department of Computer Sciences, Graphic Era (Deemed to be University), Dehradun, Uttarakhand, India.

ABSTRACT

Objective: One of the best ways to counter any pathogen including SARS-CoV-2 is through enhancement of self-immune response. Tetherin or BST2 is one such interferon-induced antiviral protein or immune regulatory protein of the human host that can restrict the release of virions of SARS-CoV-2. The ORF7a encoded protein of SARS-CoV-2 in-turn has been found to act as an antagonist to the human tetherin protein by interacting with key amino acid residues, thereby weakening the immune response.

Methods: In this research after consensus analysis of results from various servers mutations have been induced in selected residues of ORF7a coronavirus protein. Most conserved residues have been obtained through ConSurf server, ligand binding pockets, and active sites have been identified after combining results from multiple servers like CastP and 3DLigand Site.

Results: The most deleterious mutation has been identified through DUET and DynaMut servers and mutations induced thereafter have been tested for a decrease in binding affinity through the HawkDock server. The protein-protein interaction studies have been performed before and after the induction of mutation through the HawkDock server.

Conclusions: The decrease in binding affinity between human tetherin and ORF7a protein is supportive of the need to study the induction of mutation and its detrimental effect on viruses, so that effective combat study can be designed with the interplay of our native immune system and drugs from external sources. Moreover, other deterrents like anti-viral peptides can be designed against selected amino acids.

Key Words: ORF7a, Tetherin protein, Immunomodulation, In-silico, COVID-19, Coronavirus, SARS-COV-2

INTRODUCTION

There are 6 strains of novel coronavirus with a high mutation rate and therefore development of vaccine painstaking. The studies conducted to explore the structure of SARS-CoV-2 have revealed that the virus contains two polyproteins ORF1a and ORF1ab4 structural proteins, Spike, Envelope, Membrane, Nucleocapsid, and 8 accessory proteins¹. The majority of researches done on coronavirus have targeted spike protein and proteases. But accessory proteins of corona are yet to open accounts for studies as key players in disease control and drug design². Moreover, with the realization that it's the native immune system that will finally

ensure recovery of the individual as well as stoppage of viral spread, it has become very necessary to study the mutational changes in the accessory proteins of coronavirus so that the one responsible for weakening the interactions between host immune system and the pathogen can be identified and further worked upon³.

One such host protein is tetherin or bone marrow stromal antigen 2 (BST-2) that has been studied to have the ability to restrict SARS-CoV-2. Coronavirus genome in turn encodes potential modulators of tetherin protein by ORF7a⁴. ORF7a protein is a SARS-CoV-2 genome-encoded accessory protein that is composed of a type I transmembrane protein that

Corresponding Author:

Kumud Pant, PhD, Department of Biotechnology, Graphic Era (Deemed to be University), Dehradun, Uttarakhand, India.
Telephone: +919897529207; E-mail: pant.kumud@gmail.com

ISSN: 2231-2196 (Print)

ISSN: 0975-5241 (Online)

Received: 20.05.2020

Revised: 03.07.2020

Accepted: 29.08.2020

Published: 08.09.2020

localizes primarily to the Golgi apparatus but can be found on the cell surface. The interactions as depicted between several host and viral proteins in the Interactome of ViralZone Expaty suggest that ORF7a gene product of SARS-CoV-2 counteracts the antiviral activity of tetherin protein⁴⁻⁵. Susan Weiss, a microbiologist who studies coronaviruses at the University of Pennsylvania highlighted the importance of the study the effect of accessory proteins on the pathogenicity of SARS-CoV-2⁶.

Realizing the importance of interaction study to unravel novel drug targets an effort has been made in this paper to understand the interactions, between human tetherin and viral ORF7a encoded protein and also deleterious mutations introduced in viral protein on host tetherin interactions using *in-silico* approaches.

MATERIALS AND METHODS

1. Variable and conserved residues of ORF7a with PDB ID: 6W37 were identified using ConSurf server⁷⁻⁹.
2. Protein-Protein docking was performed to identify binding pockets and interacting residues using the HawkDock server. It integrates the ATTRACT docking algorithm, the HawkRank scoring function, and the MM/GBSA free energy decomposition analysis for binding free energy calculations¹⁰.
3. Active sites and ligand binding sites were predicted using CastP and 3DLigand Site servers¹¹⁻¹². 3DLigand-Site uses predicted protein structures and the ligands present in homologous structures to predict ligand-binding sites.
4. To analyze and predict protein stability changes upon mutation various servers used are following.
 - a. DUET server predicts change in protein stability with a single mutation using two algorithms SDM (statistical potential function energy) and mCSM (structural pattern mining)¹³.
 - b. Dynamut server was further used since it calculates mutational stability analysis from 3 sources namely Bio3D, ENCoM, and DUET¹⁴.

RESULTS

ORF7a product is an accessory protein of coronavirus that may not be required for the replication of SARS-CoV but is responsible for invasion into the host innate immune system¹⁵. It interacts with host tetherin protein thereby preventing its immune-regulation activity as indicated in Viral Zone – ExPASy⁵. In this research work, the HawkDock server has been used to predict protein-protein interaction between tetherin protein with PDB ID 3NWH and Coronavirus ORF 7a protein with PDB ID 6W37^{10,16-18}. The interacting residues contributing to free energy of binding are tabulated from the

HawkDock server shown in Table 1. Only the top-scoring residues are tabulated. The interacting residues of both receptor and ligand are shown with red color in Figure 1.

The docking energies of top 05 models of wild type tetherin and SARS-CoV-2 7a protein as calculated by the HawkDock server are given in Table 2. These binding free energies are calculated as the sum total of contribution from various weak bonds mentioned in Table 1. The difference of 337.58914 (kcal/mol) is observed for the best model highlighting the disappearance of many bonds that contributed towards forming a stable association in wild type tetherin and ORF7a protein of the virus.

LIGAND BINDING SITE PREDICTION

Ligand binding sites were predicted using Server 3DLigandSite and CastP servers. The predicted binding site for Coronavirus ORF7a encoded protein from 3DLigandSite is shown in Figure 2. Also, the 3D picture with labeled atoms is shown in Figure 3.

Whereas for human tetherin or BST2 protein with PDB ID: 3NWH no ligand-binding site was predicted due to insufficient homologous structures with bound ligands.

With the CastP server the two best ligand-binding pockets predicted are shown by cartoon representation in Figure 4a for human tetherin protein and Figure 4b for ORF7a coronavirus protein.

The predicted pockets of human tetherin protein and ORF7a coronavirus protein with area and volume are shown in Table 3.

In Table 4 the various residues of the 4 peptide chains forming the two binding pockets viz. pocket 1 and pocket 2 are listed. The same is shown in Figure 5a for human tetherin protein PDB ID: 3NWH and 5b for ORF7a protein of Coronavirus PDB ID: 6W37.

For protein encoded by ORF7a of Coronavirus, two binding pockets with maximum area and volume were selected for further study as shown in Figure 4. The dimensions of pockets predicted in ORF7a protein of Coronavirus are shown in Table 3. Also, the amino acids involved in forming the pockets are enlisted in Table 5.

The prediction of conserved residues in ORF7a protein with PDB ID: 6W37 has been made with the ConSurf server. The conserved and non-conserved residues with color codes are shown for human tetherin in Figure 6a and for Coronavirus in 6b.

As can be seen in Figure 6b, the most conserved residues for coronavirus ORF7a protein are Histidine (H) at position 4, Threonine (T) at position 12 and position 13, Valine (V) at 14, Threonine (T) at 24, Asparagine (N) at 28, Serine (S) at 29, and Asparagine (N) at 37. The results are obtained on

the basis of conserved residue predicted through multiple sequence alignment of homologous sequences. The prediction of conserved residues for Human tetherin protein with PDBID 2NWH is shown in Figure 6a. Since the number of sequences homologous to human tetherin is large therefore only color-coded conserved residues for tetherin are shown in Figure 5b.

From Figure 6b the most conserved residues in human tetherin protein were found to be Serine (S-10 position), Cysteine (C-13 pos), Alanine (A-19 pos), Threonine (T-27 pos), Leucine (L-34 pos), Threonine (T-54), Leucine (L-62 and 87 pos).

Prediction of effects of the mutation on selected proteins

The residues have been selected for analysis on the basis of consensus obtained in results of servers for prediction of 1) Sequence conservation using ConSurf 2) Protein interaction prediction using HawkDock server 3) Mutation effect prediction using DUET server and Dynamut server.

The following amino acids have been found to be consensus residues.

1. Threonine at position 24 and Histidine at position 4 have been found making ligand-binding pockets using the castP server. They are also the most conserved residue as predicted by ConSurf server
2. Lysine at position 57 and Leucine at position 16 in Coronavirus ORF7a protein with PDB ID 6W37 have been chosen since are predicted to be making ligand-binding pockets, using the castP server. They are also involved in protein interaction prediction using the HawkDock server

According to DUET and DynaMut server, the most destabilizing interactions are shown with negative (-) or change in folding free energy on the introduction of mutation designated as $\Delta\Delta G$ in kcal/mol and are tabulated in Table 6.

After the introduction of these mutated residues in the wild protein encoded by ORF7a, protein interaction analysis was again performed with the HawkDock server to observe the change in interaction and interaction stability upon mutation. The change in binding free energy for interaction between wild type receptor and ligand and the mutated ligand is shown in Table2.

DISCUSSION

The mutational analysis previously has been done for SARS-CoV-2 spike protein for studying its interaction with the human ACE2 receptor¹⁹. But mutational analysis of ORF7a protein has not been done yet. The following are the significant findings from the study.

From Table 1, a decrease in interaction energy is quite evident. The highest shift in this energy is found in the interaction of R-A-VAL-13 from -0.33 to -0.19, R-A-VAL-316 from -0.57 to -0.3. A new interaction has been found to have arisen after the incorporation of mutation namely R-A-VAL-217 with interaction energy of -0.1. The appearance of newer interaction and lowering of interaction energy highlights the fact that many major changes have taken place in the binding site of coronavirus.

The significant decrease in interaction energy as revealed through the analysis of results obtained on docking with mutated viral protein highlights the importance of amino acids mentioned in Table 7 viz. K57, H4, T24, and L16 in modulating immune responses. The mentioned amino acids can prove significant drug targets since they are involved in protein-protein interaction with tetherin. Binding of designed ligands with these amino acids will result in loss of function therefore the viral protein will not be able to hinder the tetherin activity. Anti-viral peptides can be designed against ORF7a proteins, therefore, sequestering and preventing viral proteins from binding with host tetherin.

In-silico techniques can aid in multifaceted analyses therefore the ability for launching multifarious attacks on pathogens is enhanced multifold. The only requirement is the availability of a laboratory testing facilities that can provide impetus to existing research on anti-microbial agents.

CONCLUSION

The *in-silico* study of ORF7a protein has revealed many important amino acids that are required for the interaction of viruses with the host. Therefore targeting these residues can be an effective strategy for combat against any pathogen. Although researches throughout the globe are being done for improving protection and disinfection strategies against emerging viruses like Coronavirus²⁰ still there is a requirement for more. The search for newer strategies has become imperative due to ever-mutating viral and bacterial forms. *In-silico* tools and techniques have provided quicker and less costly means of data analysis but the requirement of validation of results and establishing the authenticity of treatment modalities through laboratory experimentations still remains.

Acknowledgement: We are highly thankful to Graphic Era (Deemed to be) University, Dehradun, Uttarakhand for providing us with the infrastructure facility to carry out our research work.

Funding: No funding has been obtained from anywhere for conduction of this research.

Conflict of Interest: We have no conflict of interest with anyone.

REFERENCES

- Francis K. Yoshimoto. The Proteins of Severe Acute Respiratory Syndrome Coronavirus-2 (SARS CoV-2 or n-COV19), the cause of COVID-19. *Protein J.* 2020; 39(3): 198–216.
- Kristopher M. Curtis, Boyd Yount, and Ralph S. Baric. Heterologous Gene Expression from Transmissible Gastroenteritis Virus Replicon Particles. *J Virol.* 2002 Feb;76(3):1422-34.
- Li et al., Virus-Host Interactome and Proteomic Survey Reveal Potential Virulence Factors Influencing SARS-CoV-2 Pathogenesis, *Med* (2020). Published: July 21, 2020.
- Taylor JK, Coleman CM, Postel S, Sisk JM, Bernbaum JG, Venkataraman T, Sundberg EJ, Frieman MB. 2015. Severe acute respiratory syndrome coronavirus ORF7a inhibits bone marrow stromal antigen 2 virion tethering through a novel mechanism of glycosylation interference. *J Virol* 89:11820 –11833. doi:10.1128/JVI.02274-15.
- Chantal Hulo , Edouard de Castro , Patrick Masson , Lydie Bougueleret , Amos Bairoch, Ioannis Xenarios and Philippe Le Mercier, *Viral Zone: a knowledge resource to understand virus diversity* . D576–D582 *Nucleic Acids Research*, 2011, Vol. 39, Database issue
- What do we know about the novel coronavirus's proteins? These biomolecules could hold clues to why the virus is so infectious and to how to stop it. *C&EN*: updated on April 13, 2020
- Haim Ashkenazy, Shiran Abadi, Eric Martz, Ofer Chay, Itay Mayrose, Tal Pupko and Nir Ben-Tal. *ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules* *Nucleic Acids Research*, 2016, Vol. 44, Web Server issue W344–W350 .
- Gershon Celniker, Guy Nimrod, Haim Ashkenazy Fabian Glaser Eric Martz, Itay Mayrose, Tal Pupko and Nir Ben-Tal. *ConSurf: Using Evolutionary Data to Raise Testable Hypotheses about Protein Function.* *Isr. J. Chem.* 2013 March 10
- Haim Ashkenazy, Elana Erez, Eric Martz , Tal Pupko and Nir Ben-Tal. *ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids* . *Nucleic Acids Research*, 2010, Vol. 38, Web Server issue W529–W533 .
- Gaoqi Weng, Ercheng Wang, Zhe Wang, Hui Liu, Feng Zhu, Dan Li and Tingjun Hou. *HawkDock: a web server to predict and analyze the protein–protein complex based on computational docking and MM/GBSA* . *Nucleic Acids Research*, 2019, Vol. 47, Web Server issue W322–W330
- Wei Tian, Chang Chen¹, Xue Lei¹, Jieliang Zhao and Jie Liang¹. *CASTp 3.0: computed atlas of surface topography of proteins* . *Nucleic Acids Research*, 2018, Vol. 46, Web Server issue W363–W367
- Mark N. Wass, Lawrence A. Kelley and Michael J. E. Sternberg. *3DLigandSite: predicting ligand-binding sites using similar structures.* *Nucleic Acids Research*, 2010, Vol. 38, Web Server issue W469–W473
- Douglas E.V. Pires¹, David B. Ascher and Tom L. Blundell¹. *DUET: a server for predicting effects of mutations on protein stability using an integrated computational approach.* *Nucleic Acids Research*, 2014, Vol. 42, Web Server issue W314–W319
- Carlos H.M. Rodrigues, Douglas E.V. Pires and David B. Ascher. *DynaMut: predicting the impact of mutations on protein conformation, flexibility and stability.* *Nucleic Acids Res.* 2018 Jul 2;46(W1):W350-W355
- Ding Xiang Liu , To Sing Fung , Kelvin Kian-Long Chong , Aditi Shukla , Rolf Hilgenfeld . *Accessory proteins of SARS-CoV and other coronaviruses.* *Antiviral Res.* 2014 Sep;109:97-109.
- Helen M. Berman, John Westbrook, Zukang Feng, Gary Gilliland, T. N. Bhat, Helge Weissig, Ilya N. Shindyalov, Philip E. Bourne, *The Protein Data Bank, Nucleic Acids Research*, Volume 28, Issue 1, 1 January 2000, Pages 235–242
- Heidi L. Schubert, Qianting Zhai, Virginie Sandrin, Debra M. Eckert, Mitla Garcia-Maya, Louise Saul, Wesley I. Sundquist, Roberto A. Steiner, and Christopher P. Hill. *Structural and functional studies on the extracellular domain of BST2/tetherin in reduced and oxidized conformations* . *Proc Natl Acad Sci U S A.* 2010 Oct 19;107(42):17951-6
- Nelson, C.A., Minasov, G., Shuvalova, L., Fremont, D.H., *Center for Structural Genomics of Infectious Diseases (CSGID). Structure of the SARS-CoV-2 ORF7A encoded accessory protein.* *PDB Entry - 6W37* . *RDT - 2020/03/09*.
- Joseph Thomas Ortega , Maria Luisa Serrano , Flor Helene Pujol , Hector Rafael Rangel. *Role of changes in SARS-COV-2 spike protein in the interaction with the human ACE2 receptor: An in silico analysis* . *Excli J.* 2020; 19: 410–417.
- Sanjay M, Md. Irfan A, Bhoomika A, Ashish G. *Use of Newer Protective and Disinfection Strategies: A Simple Tool Guide for the Dentists During the COVID-19 Pandemic.* *International Journal of Current Research and Review.* 2020 12;16:151-154.

Table 1: Interacting residues of Receptor (R), (human tetherin 3NWH) and Ligand (L), (ORF7a encoded Coronavirus protein 6W37). The yellow highlighted is the changed interaction energy after mutation

RESIDUE_ID	VDW	ELE	GB	SA	TOTAL	RESIDUE_ID	VDW	ELE	GB	SA	TOTAL
R-A-ARG-11	-0.55	-14.83	11.63	-0.43	-4.18	R-A-VAL-321	-1.19	-0.98	1.2	-0.02	-0.99
					-4.2						-1.00
R-A-ALA-12	-0.2	-0.02	-0.04	-0.01	-0.26	R-A-HIE-323	-1.17	0.63	0.6	-0.21	-0.15
					-0.24						-0.23
R-A-VAL-13	-0.06	-6.29	6.02	0	-0.33	R-A-LEU-324	-3.99	-2.14	2.97	-0.6	-3.76
					-0.19						-3.79
R-A-CYS-15	-0.14	-0.42	0.44	0	-0.12	R-A-CYS-318	-0.21	-0.7	0.69	0	-0.22
					-0.11						-0.23
R-A-VAL-18	-1.21	0.77	-0.49	-0.08	-1.01	R-A-ASN-320	-3.81	1.18	2.54	-0.64	-0.72
					-0.94						-0.66
R-A-VAL-217	-0.01	-5.95	5.86	0	-0.1	R-A-ALA-315	-0.38	-0.99	0.65	0	-0.72
											-0.71
R-A-LEU-325	-0.23	-1.46	1.35	0	-0.33	R-A-VAL-316	-0.54	-5.39	5.42	-0.06	-0.57
					-0.32						-0.3
R-A-GLN-327	-1.65	-9.37	9.02	-0.42	-2.42						
					-2.63						

Table 2: Interaction energy or binding free energy (kcal/mol) top 5 models (wild type and mutated) predicted by HawkDock server

Top 05 Models developed	Binding free energy (kcal/mol) Interaction between wild type receptor and ligand	Binding free energy (kcal/mol) Interaction between wild type receptor and mutated ligand	Difference in binding energy (kcal/mol)
1	4494.779140	-4157.19	337.58914
2	4289.967622	-3889.44	400.52
3	3930.096650	-3785.03	145.06665
4	3914.106540	-3781.63	132.47654
5	3864.791166	-3780.87	83.921166

Table 3: The dimension and volume of pockets predicted with CastP server for human tetherin protein with PDBID 3NWH and ORF7a protein of Coronavirus PDBID 6W37

Poc ID 3NWH	AREA	VOLUME	Poc ID 6W37	AREA	VOLUME
1	235.769	193.067	1	235.769	193.067
2	255.134	110.897	2	255.134	110.897

Table 4: Amino acid residues of four (4) polypeptide chains involved in forming pocket 1 and pocket 2 of human tetherin protein with PDBID 3NWH

Poc ID	Chain	Seq ID	AA	Poc ID	Chain	Seq ID	AA	Poc ID	Chain	Seq ID	AA
1	B	105	GLU	1	B	112	LYS	1	D	108	GLN
1	B	106	LYS	1	B	113	VAL	1	D	109	GLY
1	B	108	GLN	1	B	116	LEU	1	D	110	GLN
1	B	109	GLY	1	D	105	GLU	1	D	112	LYS
1	B	110	GLN	1	D	106	LYS	1	D	113	VAL
1	D	117	GLU	2	B	84	VAL	2	B	91	CYS
2	A	49	ASN	2	B	85	GLU	2	B	92	ASN
2	A	50	SER	2	B	87	GLN	2	C	49	ASN
2	A	52	ALA	2	B	88	ALA	2	C	50	SER
2	A	53	CYS	2	B	89	ALA	2	C	53	CYS
2	D	84	VAL	2	D	91	CYS	2	D	87	GLN
2	D	85	GLU	2	D	92	ASN	2	D	88	ALA

Table 5: The various amino acid residues of ORF7a protein of Coronavirus forming two (2) binding pockets PDBID 6W37

Poc ID	Chain	Seq ID	AA	Poc ID	Chain	Seq ID	AA	Poc ID	Chain	Seq ID	AA
1	A	24	THR	1	A	51	ALA	2	A	6	GLN
1	A	26	GLU	1	A	57	LYS	2	A	16	LEU
1	A	49	ALA	2	A	4	HIS	2	A	48	PHE
1	A	50	PHE	2	A	5	TYR	2	A	60	TYR
2	A	60	TYR	2	A	61	GLN	2	A	62	LEU

Table 6: Prediction of most destabilizing mutation using DUET and DynaMut server for target amino acids

Wild Residue	Residue Position	Mutant Residue	mCSM predicted $\Delta\Delta G$	SDM predicted $\Delta\Delta G$	DUET predicted $\Delta\Delta G$	Total $\Delta\Delta G$	Dynamut prediction $\Delta\Delta G$
K	57	S	-1.553	-0.55	-1.422	-3.525	-0.793 kcal/mol (Destabilizing)
H	4	D	-1.139	-1.21	-1.152	-3.501	-0.409 kcal/mol (Destabilizing)
T	24	P	-0.232	-1.01	-0.23	-1.47	-0.069 kcal/mol (Destabilizing)
L	16	S	-2.517	-1.2	-2.431	-6.418	

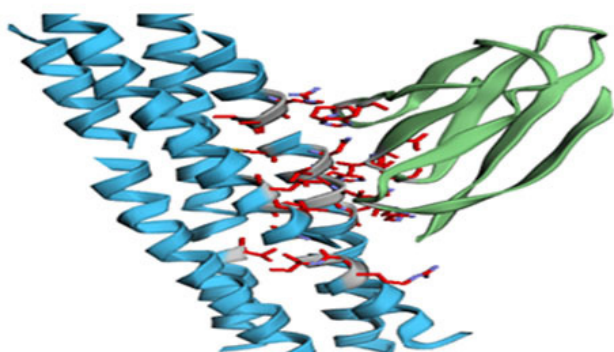


Figure 1: Binding or interaction prediction using HawkDock server.

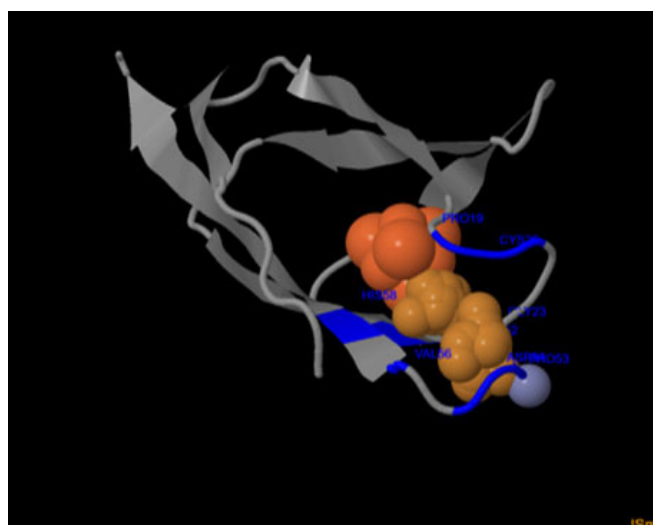


Figure 3: Binding site in Coronavirus ORF7a encoded protein with PDBID 6W37.

Predicted Binding Site				
Residue	Amino acid	contact	av distance	JS divergence
19	PRO	33	0.42	
20	CYS	47	0.09	
23	GLY	19	0.50	
50	PHE	35	0.03	
51	ALA	50	0.07	
52	CYS	53	0.01	
53	PRO	18	0.00	
54	ASP	17	0.09	
56	VAL	21	0.37	
58	HIS	27	0.35	

Figure 2: Binding site residues from 3DLigandSite server.

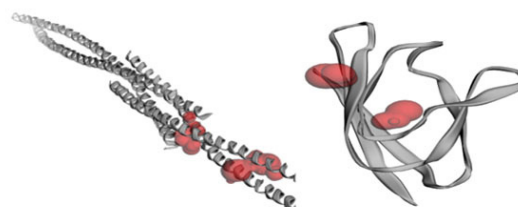


Figure 4: CastP server predicted two binding pockets of human tetherin protein (4a) and ORF7a protein (4b).



Figure 5: The amino acid residues involved in forming pocket 1 and 2 in different polypeptide chains viz. A, B, C and D are highlighted in blue in tetherin protein PDBID 3NWH (5a) and pocket 1 and 2 in ORF7a protein PDBID 6W37 (5b).



Figure 6: Predicted conserved residues in human tetherin (6a), Coronavirus (6b).