



Characterization of SARS-CoV-2 Isolate (MZ558159) Reported from India for *in Silico* Drug Designing

Rajneesh Prajapat^{*1}, Suman Jain²

¹ Ph.D Pacific Institute of Medical Sciences, Sai Tirupati University, Udaipur, Rajasthan, India

² Ph.D Pacific Institute of Medical Sciences, Sai Tirupati University, Udaipur, Rajasthan, India

***Corresponding author:** Rajneesh Prajapat, **Address:** Department of Biochemistry, Pacific Institute of Medical Sciences, Sai Tirupati University, Umarda, Udaipur- 313003, Rajasthan, India, **Email:** rajneesh030041@gmail.com, **Tel:** +917976055027

Abstract

Background & Aims: There is inadequate information available about the genomics and proteomics characterization of SARS-CoV-2 isolates reported from India and other part of the globe. This characterization is important for the *in-silico* drug designing, as there are no approved medications available to treat SARS-CoV-2 infection. The present study based on the characterization of SARS-CoV-2 (MZ558159) isolate reported from India using homology modeling, validation, and *in silico* drug designing methods.

Materials & Methods: Genome sequence of SARS-CoV-2 (MZ558159) was retrieved from NCBI, and four protein sequences e.g., QXN18496, QXN18498, QXN18504, and QXN18497 were selected for the homology modeling, validation, and *in silico* drug designing. SWISS-MODEL and UCLA-DOE server were used for homology modeling. Validation for structure model performed using PROCHECK and molecular docking using MCULE-1-Click server.

Results: The surface glycoprotein (QXN18496) model corresponding to probability conformation with 93.6%, envelope protein (QXN18498) with 88.9%, nucleocapsid phosphoprotein (QXN18504) with 93.6%, and ORF3a protein (QXN18497) with 91.8% residues in core section of ϕ - ψ plot that specifies accuracy of prediction models. The corresponding ProSA Z-score score -12.67, -0.01, -4.4, and -2.87 indicates the good quality of the models. Molecular dynamic simulation and docking studies revealed that inhibitor binds effectively at the SARS-CoV-2 (MZ558159) proteins. Predicted inhibitor 2-acetamido-2-deoxy-beta-D-glucopyranose exhibited effective binding affinity against surface glycoprotein (QXN18496).

Conclusion: The results of this study established inhibitor 2-Acetamido-2-deoxy-beta-D-glucopyranose as valuable lead molecule with great potential for surface glycoprotein (QXN18496).

Keywords: SARS-CoV-2, COVID-19, ϕ - ψ plot, ProSA

Received 27 October 2022; accepted for publication 17 November 2022

Introduction

The respiratory infection SARS-associated coronavirus (SARS-CoV) is caused by the most viruses of *Coronaviridae* family (1). The first case of SARS-CoV-2 infection was reported from Wuhan, China in

December 2019, that causes severe acute respiratory syndrome and a global health pandemic (2). The disease associated with the infection of SARS-CoV-2 is termed as Coronavirus Disease 2019 (COVID-19) (3). The COVID-19 reported from 221 countries around the world and territories. Globally, more than 5.42 billion

people worldwide have received a Covid-19 vaccine yet, equal to about 70.6 percent of the world population. Estimated mortality risk is ~2% but more in the people with underlying disease, as is ~5.6% in diabetics (4).

Coronaviruses (CoVs) are family of RNA virus that divided into α , β , δ and γ genera. The genome sequencing result of SARS-CoV has polyadenylated RNA of 29.7 kb (4). The RNA genome of SARS-CoV-2 encodes 29 proteins and phylogenetic analysis suggests bat origin (4). The virus has four structural proteins, known as S (spike), E (envelope), M (membrane), and N (nucleocapsid) proteins. An envelope-anchored SARS-CoV-2 spike (S) glycoprotein facilitates coronavirus entry into host cells (5).

During clinical trial and after treatment with lopinavir/ ritonavir, little or no coronavirus titers were detected in infected patient (6). Chloroquine phosphate could be apparent efficacy against COVID-19 associated pneumonia in clinical trials (7).

The surface glycoprotein, envelope protein, nucleocapsid phosphoprotein, and ORF3a proteins of SARS-CoV-2 play an important role in its infection (1). The sequence analysis, structure modeling, and biological data management are possible using *in-silico* methods (8-11).

The aim of this study was to characterization of genome and proteome of SARS-CoV-2 (MZ558159) isolates reported from Lucknow, Uttar Pradesh, India. The RNA genome sequence size of SARS-CoV (MZ558159) is 29836 bp. The four protein CDS sequences of SARS-CoV-2 (MZ558159) e.g., surface glycoprotein (QXN18496), envelope protein (QXN18498), nucleocapsid phosphoprotein (QXN18504), and ORF3a protein (QXN18497) were selected for homology modeling based structural analysis validation, and as a target for *in silico* drug designing.

Materials & Methods

Retrieval of Genome Sequence:

The genome sequence of SARS-CoV-2 (MZ558159) received from NCBI database. The four CDS protein sequences of SARS-CoV-2 (MZ558159) e.g., surface

glycoprotein (QXN18496), envelope protein (QXN18498), nucleocapsid phosphoprotein (QXN18504), and ORF3a protein (QXN18497) were selected for homology modeling based structural analysis and validation.

Homology Modeling:

FASTA sequence of SARS-CoV-2 (MZ558159) CDS align separately with PDB database using NCBI-BLASTp. The significant alignments with maximum identity selected for homology modeling. All the homologous protein structures were downloaded from Protein Data Bank (PDB). The SWISS-MODEL server was used for homology-modeling. The pdb files of CDS were retrieved from SWISS-MODEL server and homologous from PDB were used for 3D structure designing, modeling, and validation.

Model Reputation:

The UCLA-DOE server used for the quality analysis of protein crystal structure (12), and PROCHECK server used for validation of structure model (13, 14). The results of analysis suggesting reliability of the model (15). The overall G-factor, residue positions in ϕ - ψ plot regions analysis, was used for the selection of suitable model (16, 17). The protein stability was analyzed using QMEAN (version 3.1.0) (18, 19) and ProSA (20) Z-score.

Molecular Docking:

Molecular docking is used to dock the binding of a peptide or ligand at the preferred site and orientation on a macromolecule. Docking programs sample, the conformations of the ligand and then rank these conformations using a scoring function (21, 35). Molecular docking techniques dock small molecules into the protein binding site (31). The MCULE-1-Click docking (<https://mcule.com>) and InterEvDock 2.0 (<https://bioserv.rpbs.univ-paris-diderot.fr/>) programs were used for docking calculations (32). 1-Click docking is an online server for drug discovery platforms that performs docking (33).

Results

Protein Model Building:

The alignment between target and template was performed by using homology modeling (22). The 3D ribbon model surface glycoprotein (QXN18496),

envelope protein (QXN18498), nucleocapsid phosphoprotein (QXN18504), and ORF3a protein (QXN18497) generated using 3D structure server (<https://swissmodel.expasy.org/assess>) (Figure 1).

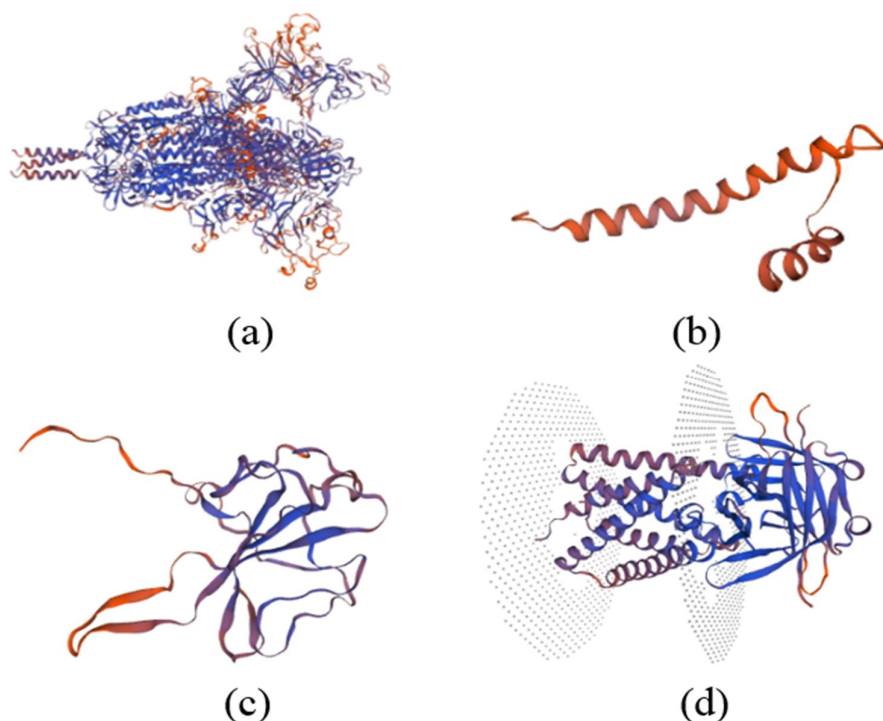


Fig. 1. 3D ribbon structure model of SARS-CoV-2 (MZ558159) proteins: [a] Surface glycoprotein (QXN18496), [b] envelope protein (QXN18498), [c] nucleocapsid phosphoprotein (QXN18504), and [d] ORF3a protein (QXN18497) generated using 3D assessment server.

Model Reputation:

SARS-CoV-2 surface glycoprotein (QXN18496) model corresponded to probability conformation with 2705 (90.4%) in favored, 252 (8.4%) in allowed and, 9 (0.3%) in generously allowed regions of ϕ - ψ plot. The envelope protein (QXN18498) model corresponding to probability conformation with 50 (92.6%) residue of core section, 4 (7.4%) of allowed section, and 0 (0.0%) residue of outer section in ϕ - ψ plot (23) (Fig. 2a and 2b; Table 1). These results indicate the reliability of protein models (24, 25).

The ϕ - ψ plot statistics of nucleocapsid phosphoprotein (QXN18504) and ORF3a protein (QXN18497) are explained in table 1.

Significant alignments using p-BLAST of surface glycoprotein (QXN18496) displayed maximum similarity with 7KRQ (99.05% identity) and QXN18498 and showed significant alignment with 2MM4 (91.38% identity). The pdb file of 7KRQ (SARS-CoV-2 spike protein) and 2MM4 (Envelope small membrane protein – SARS-CoV-2) were retrieved from PDB for homology modeling-based structure analysis and validation (Figure 3a and 3b; Table 2).

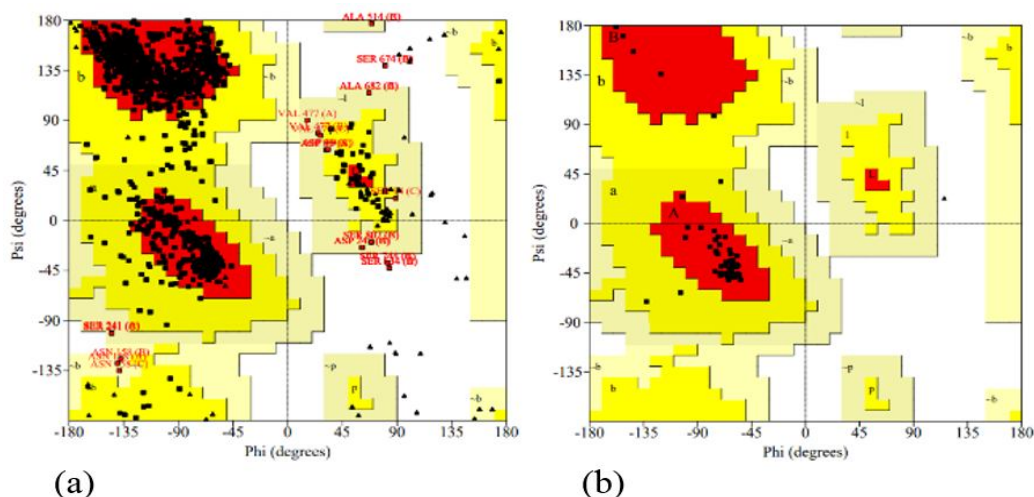


Fig. 2. (a) ϕ - ψ plot of SARS-CoV-2 surface glycoprotein (QXN18496): Total number of residues were 2705 (90.4%) in favored [A, B, L], 252 (8.4%) in allowed [a,b,l,p] and, 9 (0.3%) in generously allowed regions. (b) ϕ - ψ plot of envelope protein (QXN18498): Total number of residues were 50 (92.6%) in favored [A, B, L], 4 (7.4%) in allowed [a,b,l,p] and, 0 (0.0%) in generously allowed regions.

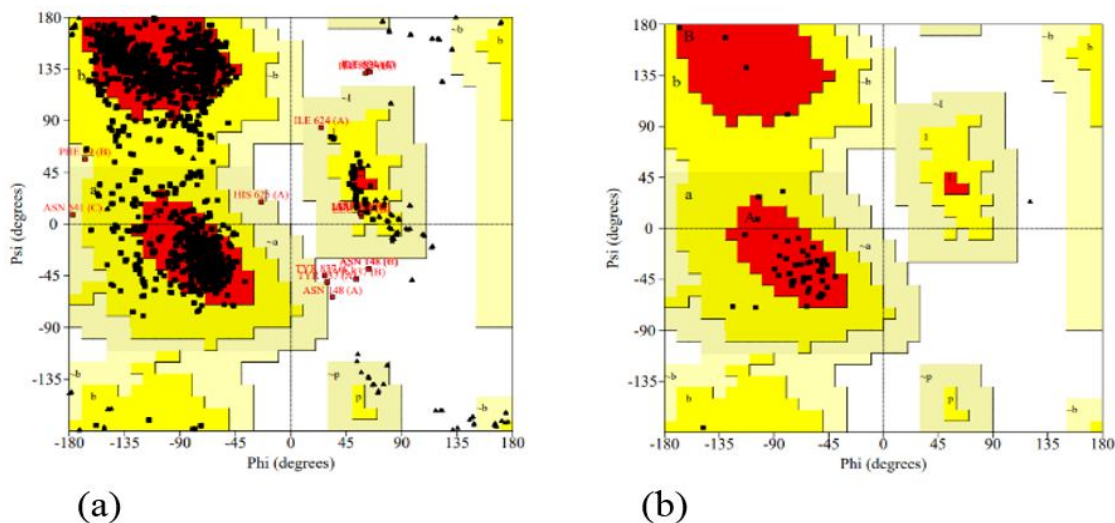


Fig. 3. (a) ϕ - ψ plot of SARS-CoV-2 7KRQ (SARS-CoV-2 spike protein): Total number of residues were 2566 (86.5%) in favored [A, B, L], 383 (12.9%) in allowed [a,b,l,p] and, 10 (0.3%) in generously allowed regions. (b) ϕ - ψ plot of 2MM4 (Envelope small membrane protein – SARS-CoV-2): Total number of residues were 50 (92.6%) in favored [A, B, L], 4 (7.4%) in allowed [a,b,l,p] and, 0 (0.0%) in generously allowed regions.

SARS-CoV-2 nucleocapsid phosphoprotein (QXN18504) model corresponding to probability conformation with 88 (93.6%) in favored, 5 (5.3%) in allowed and, 1 (1.1%) in generously allowed regions of ϕ - ψ plot. The ORF3a protein (QXN18497) model

corresponding to probability conformation with 334 (91.8%) residue of core section, 28 (7.7%) of allowed section and 0.0 % residue of outer section in ϕ - ψ plot (23) (Figure 4a and 4b; Table 1). The above results indicate the reliability of the protein model (24).

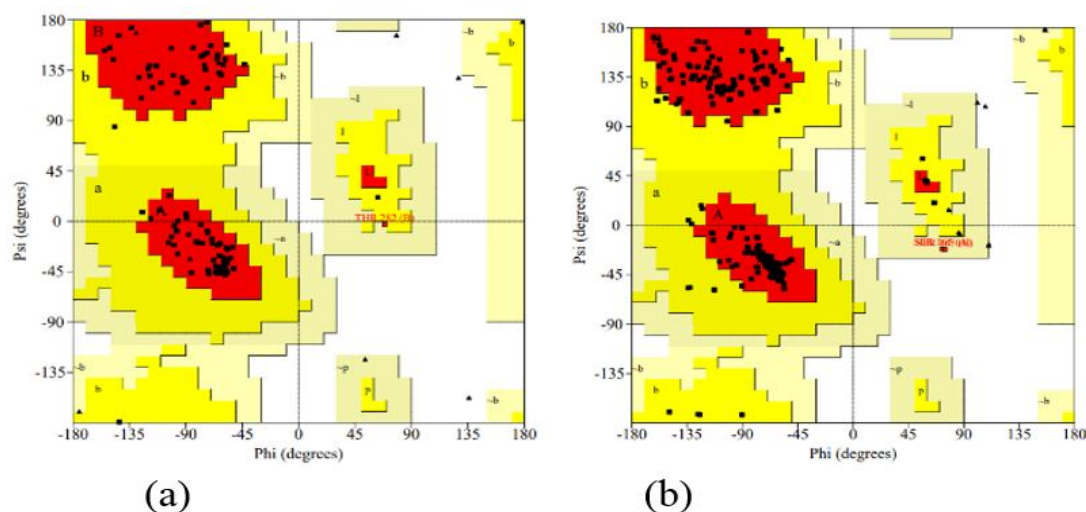


Fig. 4: (a) ϕ - ψ plot of SARS-CoV-2 nucleocapsid phosphoprotein (QXN18504): Total number of residues were 88 (93.6%) in favored [A, B, L], 5 (5.3%) in allowed [a,b,l,p], and 1 (1.1%) in generously allowed regions. (b) ϕ - ψ plot of ORF3a protein (QXN18497): Total number of residues were 334 (91.8%) in favored [A, B, L], 28 (7.7%) in allowed [a,b,l,p], and 2 (0.5%) in generously allowed regions.

Significant alignments using p-BLAST of QXN18504 against PDB displayed maximum similarity with 7CR5 (99.25% identity) and QXN18497 show significant alignment with 6XDC (99.64% identity). The pdb file of 7CR5 (SARS-CoV-2 nucleocapsid

protein) and 6XDC (SARS-CoV-2 ORF3a) were retrieved from PDB for homology modeling-based structure analysis and validation (Figure 5a and 5b; Table 2).

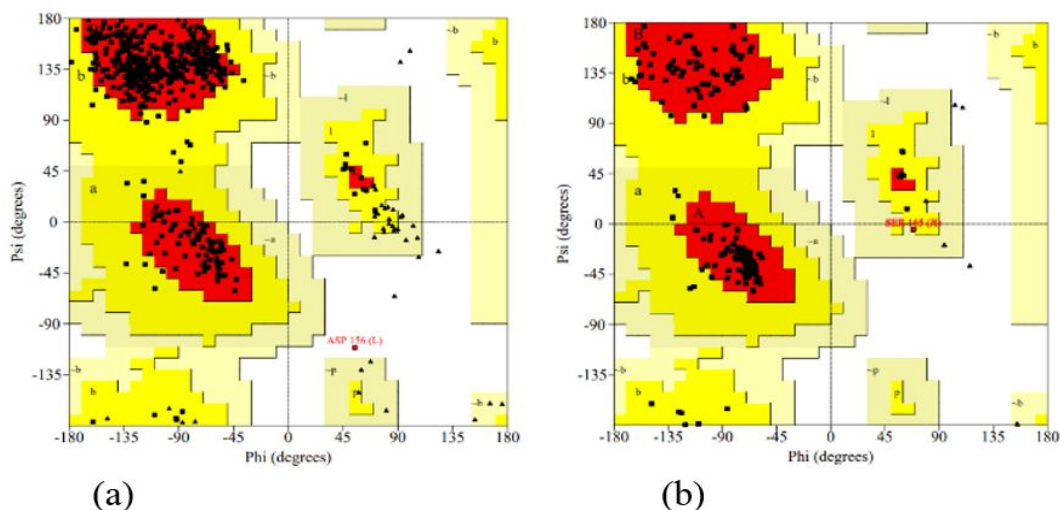


Fig. 5: (a) ϕ - ψ plot of SARS-CoV-2 7CR5 (SARS-CoV-2 nucleocapsid protein): Total number of residues were 422 (91.5%) in favored [A, B, L], 38 (8.2%) in allowed [a,b,l,p], and 0 (0.0%) in generously allowed regions. (b) ϕ - ψ plot of 6XDC (SARS-CoV-2 ORF3a): Total number of residues were 314 (89.2%) in favored [A, B, L], 36 (10.2%) in allowed [a,b,l,p], and 2 (0.6%) in generously allowed regions.

Based on the analysis of 118 structures of resolution of at least 2.0 Å and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favored regions. All selected sequences had

more than 90% residues in Favored regions [A, B, L] of ϕ - ψ plot, which indicated good quality of models (Table 1).

Table 1. The ϕ - ψ plot regions of surface glycoprotein (QXN18496), envelope protein (QXN18498), nucleocapsid phosphoprotein (QXN18504), and ORF3a protein (QXN18497).

ϕ - ψ Plot regions	QXN18496		QXN18498		QXN18504		QXN18497	
	Residues	%	Residues	%	Residues	%	Residues	%
Favored regions [A, B, L]	2705	90.4	50	92.6	88	93.6	334	91.8
Allowed regions [a, b, l, p]	252	8.4	4	7.4		5 5.3	28	7.7
Generously allowed regions [~a, ~b, ~l, ~p]	25	0.8	0	0.0	1	1.1	2	0.5
Disallowed regions	9	0.3	0	0.0	0	0.0	0	0.0
No. of end residues (excl. Gly and Pro)	9	-	2	-	1	-	4	-
Glycine residues	219	-	1	-	9	-	20	-
Proline residues	159	-	1	-	7	-	10	-

Table 2. The ϕ - ψ plot regions of 7KRQ (SARS-CoV-2 spike protein), 2MM4 (Envelope small membrane protein – SARS-CoV-2), 7CR5 (SARS-CoV-2 nucleocapsid protein), and 6XDC (SARS-CoV-2 ORF3a).

ϕ - ψ Plot regions	7KRQ		2MM4		7CR5		6XDC	
	Residues	%	Residues	%	Residues	%	Residues	%
Favored regions [A, B, L]	2566	86.5	50	92.6	422	91.5	314	89.2
Allowed regions [a, b, l, p]	383	12.9	4	7.4	38	8.2	36	10.2
Generously allowed regions [~a, ~b, ~l, ~p]	10	0.3	0	0.0	0	0.0	2	0.6
Disallowed regions	8	0.3	0	0.0	1	0.2	0	0.0
No. of end residues (excl. Gly and Pro)	21	-	2	-	6	-	6	-
Glycine residues	216	-	1	-	53	-	20	-
Proline residues	159	-	1	-	34	-	8	-

Validation of Model:

ProSA was used to Figure out potential errors in 3D model of SARS-CoV-2 (MZ558159) CDS. The archived ProSA Z-score score -12.67, -0.01, -0.01 and -4.4 for surface glycoprotein (QXN18496), envelope protein (QXN18498), nucleocapsid phosphoprotein

(QXN18504), and ORF3a protein (QXN18497) respectively indicates two aspects: overall model quality and energy deviation (Fig. 6 to 9). The predicted values of Z-score indicated less erroneous structures (20, 25). Reliability of projected model based on scoring function of QMEAN that stated as 'Z-score' (Figure 6-9) (10, 26).

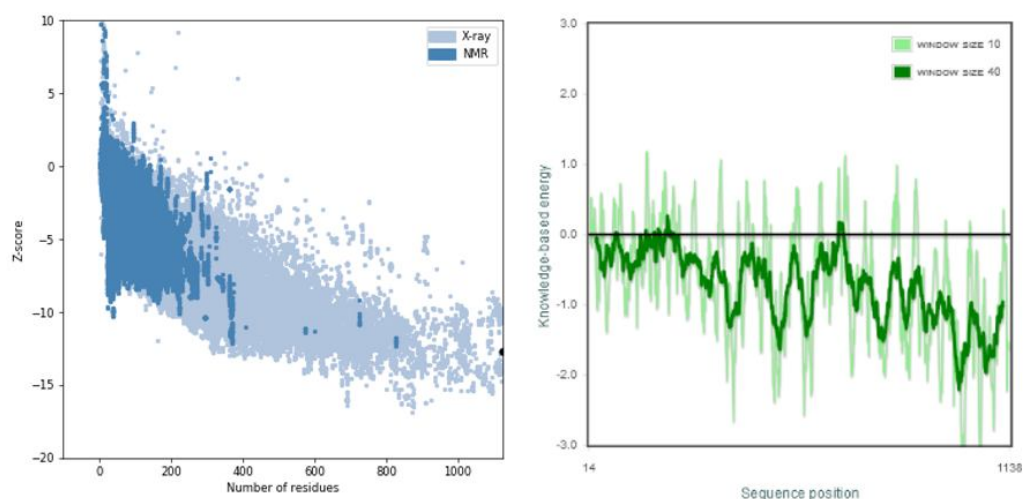


Fig. 6. ProSA service examination of surface glycoprotein (QXN18496) chain A (1125 aa) overall model quality (a) and local model quality (b) Z-Score -12.67

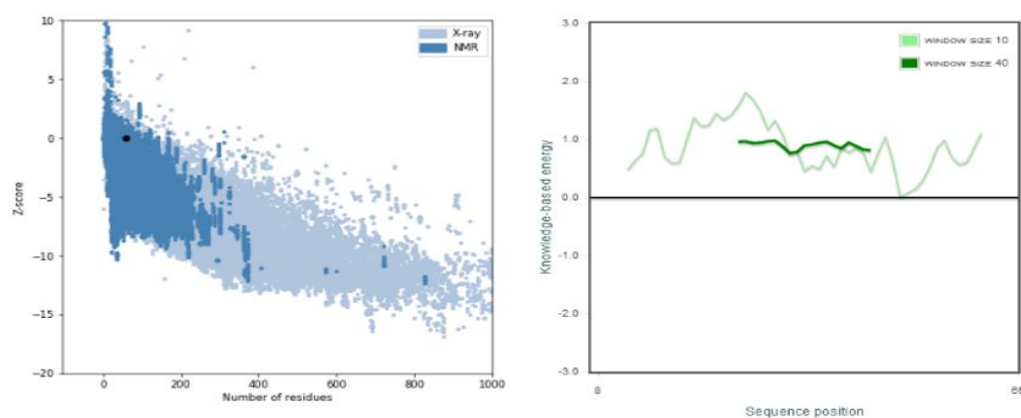


Fig. 7. ProSA service examination of envelope protein (QXN18498) chain A (58 aa) overall model quality (a) and local model quality (b) Z-Score -0.01.

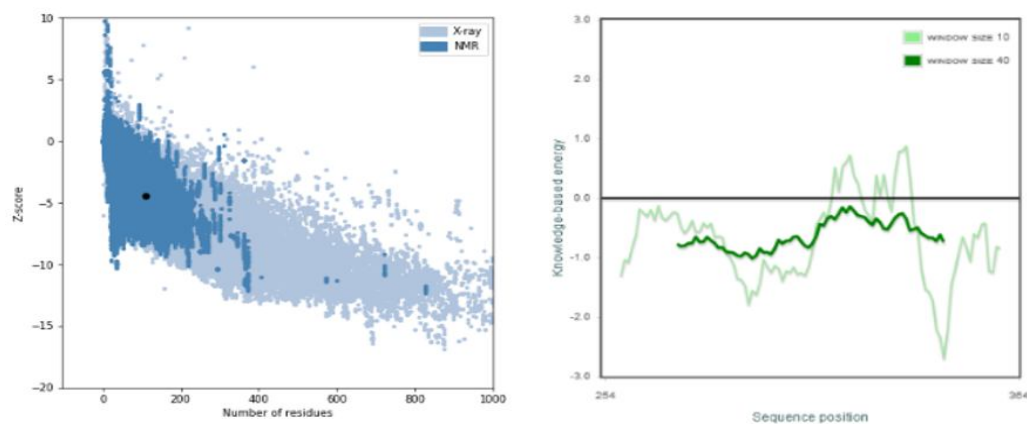


Fig. 8. ProSA service examination of nucleocapsid phosphoprotein (QXN18504) chain B (111 aa) overall model quality (a) and local model quality (b) Z-Score -4.4.

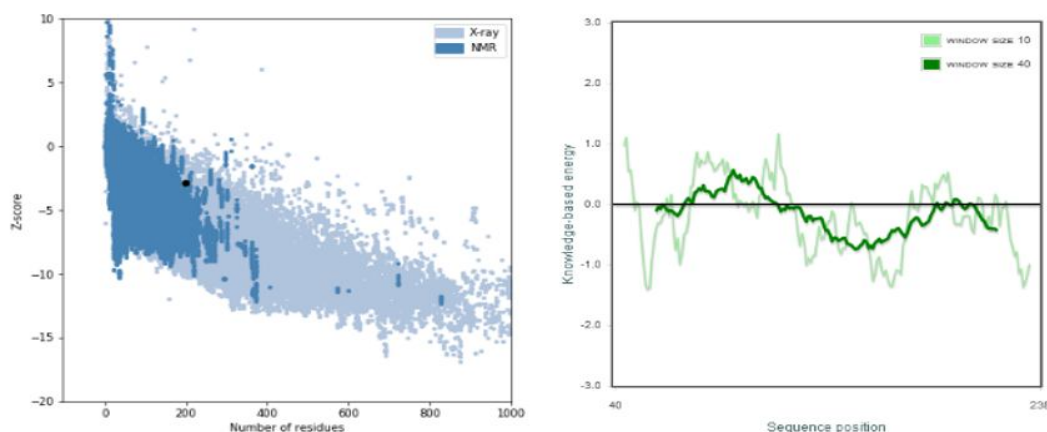


Fig. 9. ProSA service examination of ORF3a protein (QXN18497) chain A (199 aa) overall model quality (a) and local model quality (b) Z-Score -2.87.

The QMEAN scores -2.48, -3.40, 0.06 and -2.84 for biological unit reference set of for surface glycoprotein (QXN18496), envelope protein (QXN18498), nucleocapsid phosphoprotein (QXN18504), and ORF3a protein (QXN18497) respectively, which was very close to 0 and its illustrations acceptable value (28). Assessed validity of model was predictable among 0 and 1, which could be concluded from the density plot locus set for QMEAN score (Figure 10-13).

Figures 10 to 13 illustrated QMEAN scores for biological unit references set that were used as a tool for oligomeric protein assessment. We have applied QMEAN Z-scores to experimental structures from the PDB database (27). Table 3 shows the Z-scores analysis

of two experimental structures solved by X-ray diffraction of SARS-CoV-2 (MZ558159) CDS. The QMEAN Z-score of the QXN18504 was -0.06, *i.e.*, the score of the structure was clearly within the expected quality range as it deviates less than 1 standard deviation from the mean score in similar sized high-quality proteins from the reference dataset. The structure of the QXN18496, QXN18498, and the QXN18497 had QMEAN scores deviating by more than 3 standard deviations, indicating that there was clearly something wrong with this structure. Both the composite QMEAN score as well as all individual terms deviated strongly from expected values. Indeed, these structures, had been identified as fabricated and have been retracted.

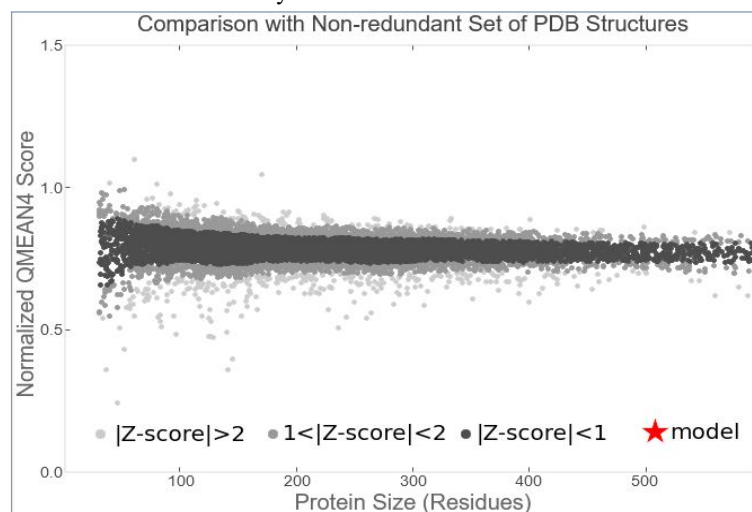


Fig. 10. QMEAN scores (-2.48) for biological unit reference set of surface glycoprotein (QXN18496). (a) Plot showing Z-score.

Higher QMEAN Z-scores in a pairwise comparison with their homologous underline the significance of the QMEAN Z-score as an estimate of protein stability (28).

Table 3. Z-score analysis of the QXN18496, QXN18498, QXN18504 and the QXN18497.

PDB	QMEAN	C- β	All-atom	Solvation	Torsion
QXN18496	-2.48	-0.71	-0.98	-0.92	-1.99
QXN18498	-3.40	-2.20	-2.36	-1.94	-2.28
QXN18504	-0.06	-0.18	-0.68	-1.31	-0.62
QXN18497	-2.84	-3.97	-0.73	-1.05	-1.97

The QMEAN Z-scores as well as the Z-scores of individual statistical potential terms were reported. The structure had been retracted from the PDB. The QMEAN Z-score of -2.48 for QXN18496, -3.40 for QXN18498 and -2.84 for QXN18497 in isolation were

unfavorable, especially the solvation and the C- β interaction terms exhibit large differences between the Z-score of the isolates (Table 3) (28).

Higher QMEAN Z-scores in a pairwise comparison with their homologous underline the significance of the QMEAN Z-score as an estimate of protein stability (28).

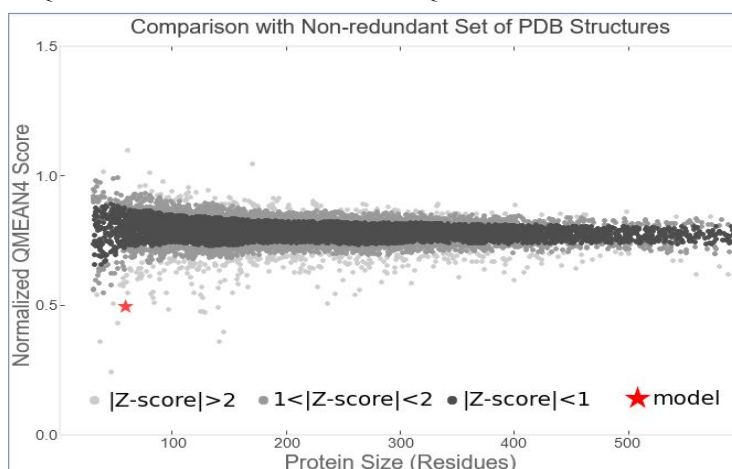


Fig. 11. QMEAN scores (-3.40) for biological unit reference set of envelope protein (QXN18498). Plot shows Z-score (a).

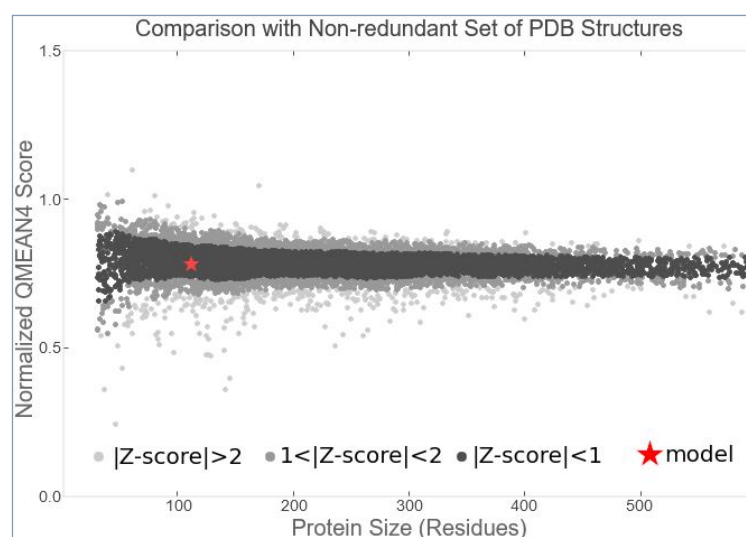


Fig. 12: QMEAN scores (0.06) for biological unit reference set of nucleocapsid phosphoprotein (QXN18504). Plot shows Z-score (a).

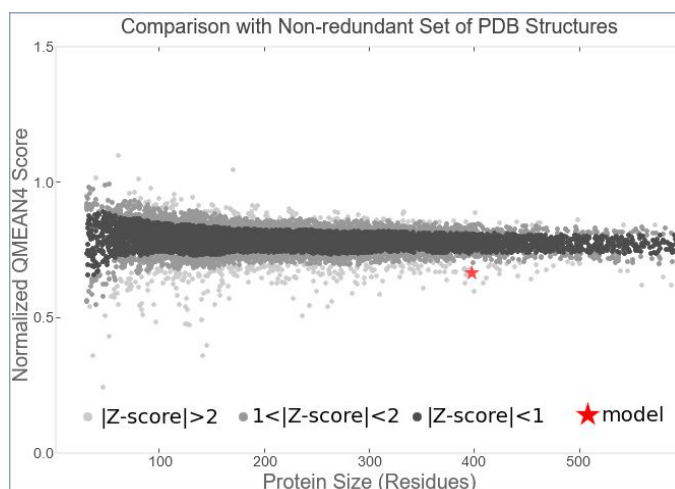


Fig. 13. QMEAN scores (-2.84) for biological unit reference set of ORF3a protein (QXN18497). Plot shows Z-score (a).

The QMEAN value comparison with the non-redundant protein collection revealed different set of Z-values for different parameters. The diversion of total energy of protein was measured by using Z-score (29).

Molecular Docking:

The *in-silico* approaches utilized for the prediction of binding pockets using *in-silico* docking study. MCULE-1-Click docking (<https://mcule.com>) and InterEvDock-2.0 server were employed to explore the binding of ligands to the respective protein. The top five docking models of binding pockets of SARS-CoV-2 (MZ558159) CDS were identified and ranked based on the energy. More negative docking scores indicate higher binding affinity.

The binding pocket and interacting residues of the selected inhibitor (2-Acetamido-2-deoxy-beta-D-glucopyranose) were analyzed in 3D using both servers (Figure 14, 15). The binding residues of the cavities were explored for the fruitful binding of novel ligands. The energy range of predicted cavities also elaborates on the efficacy of pockets. The mutational study of binding residues suggested that these residues could be used as a clinical prospectus against the effective treatment of COVID-19. The predicted binding residues lead to the drug designing of lead compounds against SARS-CoV-2 surface glycoprotein (QXN18496).

Lead Compounds	Target Proteins
2-Acetamido-2-deoxy-beta-D-glucopyranose	Surface glycoprotein (QXN18496)

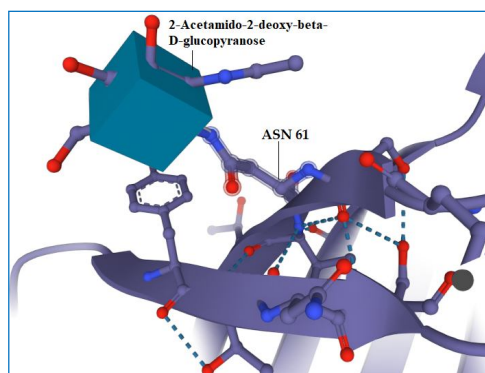


Fig. 14: Binding pocket and interacting residues of inhibitor 2-Acetamido-2-deoxy-beta-D-glucopyranose with surface glycoprotein (QXN18496) using 1-click docking server.

The ligand 2-Acetamido-2-deoxy-beta-D-glucopyranose [PubChem CID: 828139] (Table 4) retrieved for docking with surface glycoprotein (QXN18496). The docking results revealed the ASN 61 of residue exhibit good binding interactions with inhibitors and mutational studies of these residues could be highly effective in further studies.

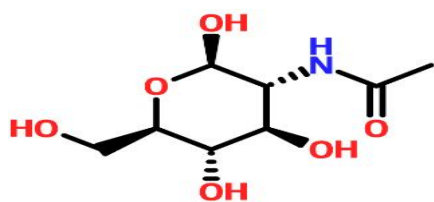


Fig. 15. Structures of inhibitor 2-Acetamido-2-deoxy-beta-D-glucopyranose.

The simulated ^1H spectrum with explicit hydrogens and the list of peaks modules were linked. The results highlighted the chemical shifts, scalar couplings and peak multiplicity, and corresponding atom in the structure and the relevant peak in the spectrum (Figure 16) (30).

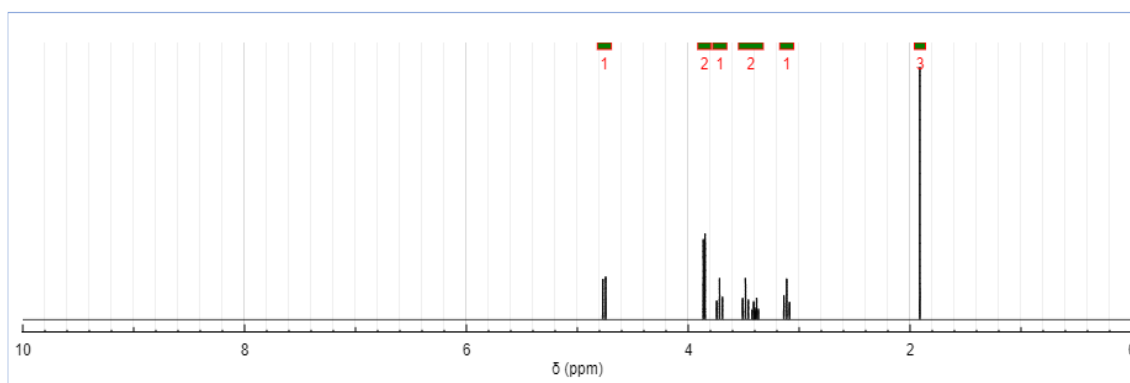


Fig. 16. ^1H NMR Spectra of 2-Acetamido-2-deoxy-beta-D-glucopyranose. List of ^1H NMR Signals: ^1H NMR: δ 1.91 (3H, s), 3.11 (1H, t, $J = 10.2$ Hz), 3.32-3.55 (2H, 3.40 (dt, $J = 10.3, 6.4$ Hz), 3.48 (t, $J = 10.2$ Hz)), 3.72 (1H, t, $J = 10.2$ Hz), 3.80-3.91 (2H, 3.85 (d, $J = 6.4$ Hz), 3.85 (d, $J = 6.4$ Hz)), 4.76 (1H, d, $J = 10.3$ Hz).

Table 4. Inhibitor 2-Acetamido-2-deoxy-beta-D-glucopyranose (PubChem CID: 24139) properties and binding residue ³⁴

Ligand Properties	24139
Molecular Weight (g/mol)	221.21
Component type	Non-polymer
Hydrogen Bond Donor Count	05
Hydrogen Bond Donor Count	06
Rotatable Bonds Count	02
Topological Polar Surface Area	119 Å ²
Heavy Atom Count	15
Formal Charge	0
Boiling Point	595.35 °C
Melting Point	205 °C
Binding residue	ASN 61

It was observed that the inhibitor binds at the binding residues ASN 16 (Table 4). The predicted structural and docking model described in this study may be further used for finding interactions with other SARS-CoV-2 proteins to identify new anti-coronavirus targets.

Discussion

Till now, there is no approved therapies available for treatment of coronavirus infection. In the present study, the four CDS protein sequences of SARS-CoV-2

(MZ558159) were selected for homology modeling based structural analysis and *in silico* drug designing.

SARS-CoV-2 surface glycoprotein (QXN18496) model correspond to probability conformation with 2705 (90.4%) in favored, 252 (8.4%) in allowed, and 9 (0.3%) in generously allowed regions of ϕ - ψ plot. The envelope protein (QXN18498) model correspond to probability conformation with 50 (92.6%) residue of core section, 4 (7.4%) of allowed section, and 0 (0.0%) residue of outer section in ϕ - ψ plot. The above results indicated the reliability of protein models.

ProSA was used to Figure out potential errors in 3D model of SARS-CoV-2 (MZ558159) CDS. The archived ProSA Z-score score -12.67, -0.01, -0.01, and -4.4 for selected CDS protein sequences of SARS-CoV-2 (MZ558159) indicated two aspects: overall model quality and energy deviation. The predicted values of Z-score indicated less erroneous structures. Reliability of projected model based on scoring function of QMEAN, stated as 'Z-score'.

Molecular dynamic simulation and docking studies revealed that the inhibitors 2-Acetamido-2-deoxy-beta-D-glucopyranose binds effectively at the binding site ASN 61 of surface glycoprotein (QXN18496). The docking results revealed the ASN 61 of residue exhibit good binding interactions with inhibitors. Hence, the proposed inhibitors could potently inhibit SARS-CoV-2 fusion and entry by using a pseudo-typed-virus system. These results lay the groundwork for future inhibitory peptide drug design and future clinical trials.

Conclusion

The functional characteristics of CDS protein sequences of SARS-CoV-2 (MZ558159) could be predicted by the generated modes for structure analysis and validation, and a designed inhibitor could potently inhibit SARS-CoV-2 fusion and entry by using a pseudo-typed-virus system using *in silico* modeling approach. Methods ProSA, QMEAN, and PROCHECK build model reliability. The 1-click docking server used for template-based and template-free protein-protein docking. Molecular dynamic simulation and docking studies revealed that the inhibitors 2-Acetamido-2-

deoxy-beta-D-glucopyranose binds effectively at the binding site ASN 61 of surface glycoprotein (QXN18496). These findings lay the groundwork for future inhibitory peptide drug design and clinical trials.

Acknowledgments

The authors are grateful to Dr. Indrajeet Singhvi (Vice Chancellor, Sai Tirupati University, Udaipur, Rajasthan, India) for precious support and guidance during research work and to biochemistry and bioinformatics research group members for technical support

Ethical statement

The research project was approved by the Ethics Committee (Registration no. STU/2019/5006) of Pacific Institute of Medical Sciences, Sai Tirupati University, Udaipur, Rajasthan, India.

Data availability

The raw data supporting the conclusions of this article are available from the authors upon reasonable request.

Conflict of interest

None declared

References

1. Peiris JS. Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet* 2003; 361:1319–25.
2. Jaiswal G, Kumar V. In-silico design of a potential inhibitor of SARS-CoV-2 S protein. *PLoS ONE* 2020; 15(10): e0240004.
3. Drosten C. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N Engl J Med* 2003; 348:1967–76.
4. Novosad P, Jain R, Campion A, Asher S. COVID-19 mortality effects of underlying health conditions in India: a modeling study. *BMJ Open* 2020;10 (12): e043165.
5. Li F. Structure, Function, and Evolution of Coronavirus Spike Proteins. *Annu Rev Virol* 2016 29; 3(1):237-61
6. Lim J, Jeon S, Shin HY, Kim MJ, Seong YM, Lee WJ, Choe KW, Kang YM, Lee B, Park SJ. Case of the Index

- Patient Who Caused Tertiary Transmission of COVID-19 Infection in Korea: The Application of Lopinavir/Ritonavir for the Treatment of COVID-19 Infected Pneumonia Monitored by Quantitative RT-PCR. *J Korean Med Sci* 2020; 35(6): e79.
7. Gao J, Tian Z, Yang X. Breakthrough: Chloroquine phosphate has shown apparent efficacy in treatment of COVID-19 associated pneumonia in clinical studies. *Biosci Trends* 2020;14(1):72-3. doi: 10.5582/bst.2020.01047.
 8. Naulaerts S, Meysman P, Bittremieux W, Vu TN, WV Berghe, Goethals B, Laukens K. A primer to frequent itemset mining for bioinformatics. *Briefings Bioinform* 2015;16:216-31.
 9. Rasouli H, Fazeli-Nasab B. Structural validation and homology modeling of lea 2 protein in bread wheat. *Am Eurasian J Agric Environ Sci* 2014; 14:1044-8.
 10. Prajapat R, Jain S, Vaishnav MK, Sogani S. Structural Modeling and Validation of Growth/Differentiation Factor 15 [NP_004855] Associated with Pregnancy Complication- Hyperemesis Gravidarum. *J Krishna Instit Med Sci* 2020; 9(3):40-7.
 11. Prajapat R, Marwal A, Shaikh Z, Gaur RK. Geminivirus Database (GVDB): First database of family geminiviridae and its genera Begomovirus. *Pak J Biol Sci* 2012;15:702-6.
 12. Luthy R, Bowie JU, Eisenberg D. Assessment of protein models with three-dimensional profiles. *Nature* 1992;356:83-5.
 13. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: A program to check the stereochemical quality of protein structures. *J Applied Cryst* 1993;26:283-91.
 14. Vriend G. WHAT IF: A molecular modeling and drug design program. *J Mol Graphics* 1990;8:52-6.
 15. Sehgal SA, Tahir RA, Shafique S, Hassan M, Rashid S. Molecular modeling and docking analysis of CYP1A1 associated with head and neck cancer to explore its binding regions. *J Theoret Comput Sci* 2014;1(3):1-6. doi: 10.4172/2376-130X.1000112
 16. Agrawal P, Thakur Z, Kulharia M. Homology modeling and structural validation of tissue factor pathway inhibitor. *Bioinformation* 2013; 9: 808-12.
 17. Benkert P, Tosatto SCE, Schomburg D. QMEAN: A comprehensive scoring function for model quality assessment. *Proteins: Struct Funct Bioinform* 2008; 71:261-77.
 18. Benkert P, Kunzli M, Schwede T. QMEAN server for protein model quality estimation. *Nucleic Acids Res* 2009; 37:W510-W4.
 19. Novotny WF, Girard TJ, Miletich JP, Broze GJ. Platelets secrete a coagulation inhibitor functionally and antigenically similar to the lipoprotein associated coagulation inhibitor. *Blood* 1988;72: 2020-5.
 20. Wiederstein M, Sippl MJ. ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res* 2007;35:W407-W10.
 21. Jaiswal G, Kumar V. In-silico design of a potential inhibitor of SARS-CoV-2 S protein. *PLoS ONE* 2020;15(10):e0240004.
 22. Prajapat R, Marwal A, Gaur RK. Recognition of errors in the refinement and validation of three-dimensional structures of AC1 proteins of begomovirus strains by using ProSA-web. *J Viruses* 2014; 6. doi:10.1155/2014/752656.
 23. Bowie JU, Luthy R, Eisenberg D. A method to identify protein sequences that fold into a known three-dimensional structure. *Science* 1991;253:164-70.
 24. Mustufa MMA, Chandra S, Wajid S. Homology modeling and molecular docking analysis of human RAC-alpha serine/threonine protein kinase. *Int J Pharma Bio Sci* 2014; 5:1033-42.
 25. Rekik I, Chaabene Z, Grubb CD, Drira N, Cheour F, Elleuch A. In silico characterization and molecular modeling of double-strand break repair protein MRE11 from *Phoenix dactylifera* v deglet nour. *Theor. Biol Med Model* 2015; 12:23. doi: 10.1186/s12976-015-0013-2.
 26. Prajapat R, Jain S, Vaishnav MK, Sogani S. In Silico Characterization of Surface Glycoprotein [QHD43416] of SARS-Coronavirus. *Chinese J Med Res* 2020; 3(2): 32-6. doi: 10.37515/cjmr.091X.3201
 27. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank, *Nucleic Acids Res* 2000; 28: 235-42.

28. Benkert P, Biasini M, Schwede T. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* 2011; 27(3):343–50.
29. Studer G, Rempfer C, Waterhouse AM, Gumienny G, Haas J, Schwede T. QMEANDisCo - distance constraints applied on model quality estimation. *Bioinformatics* 2020;36:1765-71.
30. Aires-de-Sousa M, Hemmer JG. Prediction of ¹H NMR Chemical Shifts Using Neural Networks. *Anal Chem* 2002;74(1):80-90.
31. Wiederstein M, Sippl MJ. Protein sequence randomization: Efficient estimation of protein stability using knowledge-based potentials. *J Mol Biol* 2005;345:1199-212.
32. Gao J, Lu G, Qi J, Li Y, Wu Y, Deng Y, Geng H, Li H, Wang Q, Xiao H, Tan W, Yan J, Gao GF. Structure of the fusion core and inhibition of fusion by a heptad repeat peptide derived from the S protein of Middle East respiratory syndrome coronavirus. *J Virol* 2013; 87(24):13134-40. doi: 10.1128/JVI.02433-13.
33. Ramachandra SC, Prashant A, Vishwanath P. COVID-19 Induced Cytokine Storm and the Impact of Obesity and Vitamin D Deficiency. *J Krishna Instit Med Sci* 2021;10(1):1-14.
34. National Center for Biotechnology Information (2022). PubChem Compound Summary for CID 24139. <https://pubchem.ncbi.nlm.nih.gov/compound/2-acetamido-2-deoxy-beta-D-glucopyranose>.
35. Lokhande KB, Doiphode S, Vyas R, Swamy KV. Molecular docking and simulation studies on SARS-CoV-2 Mpro reveals Mitoxantrone, Leucovorin, Birinapant, and Dynasore as potent drugs against COVID-19. *J Biomol Struct Dyn* 2020; 1-12. doi:10.1080/07391102.2020.1805019.

This is an open-access article distributed under the terms of the [Creative Commons Attribution-noncommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/) which permits copy and redistribute the material just in noncommercial usages, as long as the original work is properly cited.