ORIGINAL PAPER

Computational modelling studies on in silico missenses in COVID-19 proteins and their effects on ligand–protein interactions

Laxmi Sule[1] · Swagata Gupta[2] · Nilanjana Jain[3] · Nitin S. Sapre[4]

Abstract

The paper presents the incorporation of in silico missenses and studies the effect of missenses to understand its effect on the ligand–protein interactions, of COVID-19 protein. In silico protein–ligand interaction, studies are being used to understand and investigate the drug-likeness of various molecules. 19 novel COVID-19 proteins are designed by inducing in silico missenses by mutating N691 amino acid residue in 7bv2 protein, the only residue forming H-bond with the ligand molecule in the parent protein. The work illustrates the effects of in silico-induced mutation on various interactions such as H-Bond, VDW, π-alkyl interactions, and changes in the number and type of surrounding amino acid residues. The results have suggested a common pattern of behaviour on mutation with T, V, W, and Y. Further, it is observed that the number and type of amino acid residues increase on mutation, suggesting the effect of mutation on the ligand–protein binding.

Keywords COVID-19 · In silico missenses · De novo design · COVID-19 Protein

---

✉ Nitin S. Sapre
   sukusap@yahoo.com; nsapre@sgsits.ac.in

   Laxmi Sule
   lakshmisule3@gmail.com

   Swagata Gupta
   swagataagupta@yahoo.co.uk; swagatagupta28@gmail.com

   Nilanjana Jain
   nilanjanamjain@gmail.com

[1] Department of Applied Chemistry, SGSITS, Indore, Madhya Pradesh, India

[2] Department of Chemistry, Govt. Holkar (Model Autonomous) Science College, Indore, Madhya Pradesh, India

[3] Department of Chemistry, Govt. College, Udaynagar, Dewas, Madhya Pradesh, India

[4] Department of Chemistry, SGSITS, Indore, Madhya Pradesh, India

1 Introduction

A recent pandemic, which erupted in 2019, caused by Corona Virus affecting the respiratory system, causing severe acute respiratory syndrome (SARS) has caused havoc in the world [1–8].

It primarily affected the lungs as in pneumonia. The earlier nomenclature of the virus was done as "Wuhan Virus" and "2019 n-CoV" which was later renamed as SARS-CoV-2 due to its resemblance to the SARS-CoV RNA genome [6, 9–11]. Belonging to the Coronaviridae family and Orthocoronavirinae, the virus has four genera [6, 9, 12–15]

The spread of the virus was unprecedented and was like wildfire. As of today, there is no medication available to treat the disease effectively [16, 17].

Remdesivir which is used in treatment, is a derivative of adenosine and inhibits RNA-dependent RNA polymerase enzyme (RsRp) [18–21]. Immunity booster vaccines such as Covishield, Covaxin, and Sputnik have been used as a possible defence against infection [2, 22–25]

Mutation in species (protein or genetic material), usually results in a reduction or change in the effect of the drug. Some of the reasons causing this are mode of attachment, change in binding affinity, and hydrophobic interactions to list a few [26–29].

Alterations occurring in the genetic material or the character are the causes of the mutation [30]. Mutation can be heritable or non-heritable depending on the tissue of the origin. When mutation effects are observed in the next generation they are referred to as the germinal mutation while, when no such changes are observed in the next generation, then it is classified as the somatic mutation. The term mutant is used to represent the organism in which changes are observed while mutation-causing agents are referred to as mutagen [31–33].

Mutation(s) can occur in chromosomes, genes, or proteins. Mutations of all three types are lethal, but the chromosomal mutation is more lethal compared to the other two.

2 Protein mutations

Synonymous and non-synonymous mutations refer to the replacement of an amino acid residue by another and a change in the amino acid sequence respectively. Non-synonymous mutation can lead to a change in the structure of the protein or premature termination of the protein. Protein mutation can occur naturally or can be induced in vitro or in silico [34–36].

In this work single point non-synonymous mutations are performed on the COVID-19 protein, 7BV2 [37]. Docking of remdesivir was performed with the parent and designed proteins in which missenses were introduced, to understand various interactions and the effect of mutations on these interactions [38–40].

3 Methods

The present work deals with the inducing in silico missenses in COVID-19 followed by interaction studies of remdesivir with COVID-19 (7Bv2) and the designed proteins obtained by such in silico-*induced* single-point mutation.

3.1 Protein preparation

The structure of the protein (PDB ID-7bv2) is procured from the RCSB-protein data bank in the PDB file format. Single point mutation was induced in the parent protein molecule on Asparagine-691(N691) using Molegro Virtual Docker (MVD) to design nineteen (19) new protein molecules [41]. The newly designed proteins are exported as pdb files. The reason for taking N691 residue to perform mutation was that it was the only amino acid residue exhibiting H-bond with the ligand in the parent protein in the 6 Å vicinity.

The Secondary structure of the protein (pdb: 7bv2) is presented in Fig. 1.

The ligand and protein are used for performing the docking process and all the files (Ligand and Protein) were converted to pdbqt file format, where 'q' and 't' are the charges and the type of atoms respectively.

Chain A: RNA-directed RNA polymerase (Sequence Length: 951), the Part Red & Orange Colour is Uracil, cytosine and other RNA parts
Chain B: Non-structural protein (Sequence Length: 207)
Chain C : Non-structural protein ((Sequence Length: 92)

Fig. 1 Structure of 7bv2 protein (cartoon form) with different chains

3.2 Ligand preparation

The 2-D structure of remdesivir was drawn in ChemBioDraw ultra 14.0 and geometry optimization was done using MM2 force field in Chem3D pro-14. The mol file of the ligand was converted into pdbqt file format using Open Babel [42] This was then imported into AutoDock4.2 for further use [43].

The 2D and 3D structures of remdesivir are presented in Figs. 2A and B.

3.3 Docking

Remdesivir was docked with the 7bv2 protein as well as with the 19 mutated protein structures using AutoDock.

4 Results and discussion

Comparative analyses of the effect of mutation of amino acid residues (AARs) on docking results were studied considering the surrounding amino acids residues (SAARs), number of H-bonds, AARs exhibiting H-bonds in all the proteins, and AA residues exhibiting Van der Waals interactions.

The details of results obtained from the docking of Remdesivir with the parent and the 19 designed (mutated) protein molecules are discussed herewith.

4.1 Surrounding amino acid residues (SAARs)

The SAARs are the AA residues which are identified as those in the vicinity of 6 Å from the centre of the ligand (remdesivir) obtained after redocking in the parent
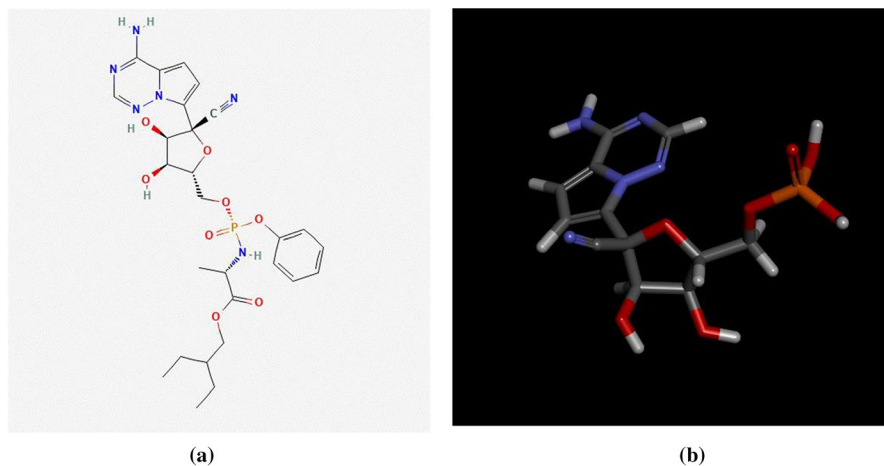

| (a) | (b) |

Fig. 2 **A** and **B**: 2D and 3D Structures of Remdesivir

protein (PP) and docking in the designed proteins. Table 1 presents the parent protein (PP), the mutation of N691 by specific AAs, numbers, and single alphabet code (with their positions) of SAARs observed within 6 Å vicinity of remdesivir.

Table 1 presents the list containing, the parent protein (PP) and all the proteins which are obtained on mutation with a specific amino acid. The column 2 depicts the AARs used for mutating N691, representing the designed protein (mutated). The column 3 depicts the number of AARs surrounding the ligand that are in the vicinity of 6 Å from the ligand. The last column details the AARs which are observed surrounding the ligand molecule and are in the vicinity of 6 Å.

From Table 1, it is observed that the 14 SAARs are observed in the parent protein molecule. Further, it is observed that a higher number of SAARs, 19, 20, 20, and 18, is obtained in designed proteins when N691 (chain A) is mutated with T, V, W, and Y. This may be due to increased ligand–protein intermolecular attractive interactions. Further, the orientation of protein might be occurring in such a way that the intermolecular attractive forces are overpowering the repulsive forces. In other cases, the number of SAARs obtained is lesser than compared to 7BV2.

Table 2 presents all the SAARs which are observed in the vicinity of 6 Å and the number of proteins in which they are observed including parent and the designed.

Table 2 further presents the number of designed proteins in which a specific AAR is present in the 6 Å vicinity of the ligand, i.e. the table gives details of the occurrence of amino acid which is in 6 Å vicinity of the ligand.

From Table 2 it can be observed that only 41 of the total 951 AARs of chain-A are observed in this vicinity of 6 Å. The number of SAARs conserved in the designed proteins and the mutation-causing AA residues are presented in Table 3.

Out of these 41 SAARs, ten of the AARs are present in more than 10 designed proteins. R555 AAR is present in all the designed proteins, R545 is present in 18 designed proteins, whereas A547, I548, K551, R553 each in 15, A554, S549 in 14 and N552, S814 in 11 designed proteins. The rest of the AARs occurrence is less than 10 in the remaining designed proteins. From the Table 2, it can be observed that R555 never moved away from the ligand and is even present in the parent protein. This suggests that R555 has more affinity for the ligand than the other SAARs.

From Table 3 it is observed that R555 residue is the most conserved SAAR i.e., in 18 designed proteins except the one mutated with 'Y', while K545 is conserved in 17 proteins except those mutated with 'R' and 'S'.

From Table 3 it is further observed that out of the 14 SAARs in the parent protein, nine SAARs namely K545(in 17), A547(in 14), R555 (in 18), V557(in 4), C662(in 3), D623(in 2), T680(in 4), S681(in 4), and S682(in 4) are conserved, while T687, A688, N691, S759, D760 are not conserved in any of the designed protein.

Further, it is observed that SAAR numbers Y545, V557, T680, S681, and S682 are common in parents as well as the de novo-designed proteins which showed a higher number of SAARs than the parent protein.

A total of 250 SAARs are observed in the vicinity of 6 Å from the centre of the ligand (remdesivir) in all the 20 proteins (1 parent and 19 mutated) and these are presented in Table 4.

From Table 4, it is evident that 14 of the 20 AARs are showing a total of 250 SAARs. It is further observed that 'A', 'K' and 'R' (at varied positions) are the

Table 1 Parent protein (7bv2), mutated proteins, and surrounding amino acid residues (SAARs)

S.No	AAR	Number of SAARs	Position of surrounding amino acid residues (SAARs)
1	PP* (N)	14	K545, A547, R555, V557, C622, D623, T680, S681, S682, T687, A688, N691, S759, D760
2	A	9	K545, A547, I548, S549, K551, N552, A554, R553, R555
3	C	9	K545, A547, I548, S549, K551, A554, R555, S814, R836
4	D	10	K545, I548, S549, K551, N552, R553, A554, R555, S814, R836
5	E	13	F442, Q444, K545, A547, I548, S549, A550, K551, N552, R553, A554, R555, S814
6	F	11	K545, A547, I548, S549, A550, K551, N552, R553, A554, R555, R836
7	G	10	K545, A547, I548, S549, A550, K551, R553, R555, S814, R836
8	H	13	F442, Q444, K545, A547, I548, S549, A550, K551, N552, R553, A554, R555, S814
9	I	9	K545, A547, I548, S549, K551, R553, A554, R555, S814
10	K	9	K545, Y546, A547, I548, K551, R553, A554, R555, S814
11	L	12	F442, K545, A547, I548, S549, A550, K551, N552, R553, A554, R555, R836
12	M	12	F441, K545, A547, I548, S549, K551, N552, R553, A554, R555, S814, R836
13	P	11	F442, K545, A547, I548, S549, K551, N552, R553, A554, R555, R836
14	Q	9	K545, A547, I548, S549, K551, A554, R555, S814, R836
15	R	11	F442, A547, I548, S549, A550, K551, N552, R553, A554, R555, S814
16	S	11	F442, Q444, A547, I548, S549, K551, N552, R553, A554, R555, S814
17	T	19	Y453, Y455, Y456, T540, K545, R555, T556, V557, A558, K621, C622, D623, R624, V667, K676, T680, S681, S682, G683
18	V	20	Y455, Y456, T540, M542, Y543, K545, R553, R555, T556, V557, A558, K621, C622, R624, V667, K676, T680, S681, G683, S682
19	W	20	Y453, Y455, Y456, T540, K545, R553, R555, T556, V557, A558, K621, C622, D623, R624, V667, K676, T680, S681, S682, G683
20	Y	18	Y453, T455, Y456, T540, M542, K545, N552, T556, V557, A558, K621, R624, V667, K676, T680, S681, S682, G683

Table 2 Surrounding amino acid residues (SAARs)

S.No	SAAR	Number of proteins
1	R555	19
2	K545	18
3	A547	15
4	I548	15
5	K551	15
6	R553	15
7	A554	14
8	S549	14
9	N552	11
10	S814	11
11	R836	8
12	A550	6
13	F442	6
14	S681	5
15	S682	5
16	T680	5
17	V557	5
18	A558	4
19	C622	4
20	G683	4
21	K621	4
22	K676	4
23	R624	4
24	T540	4
25	T556	4
26	V667	4
27	Y456	4
28	D623	3
29	Q444	3
30	Y453	3
31	Y455	3
32	M542	2
33	A688	1
34	D760	1
35	F441	1
36	N691	1
37	S759	1
38	T455	1
39	T687	1
40	Y543	1
41	Y546	1

Table 3 Number of SAARs, number of times conserved and mutating AAR

SAAR number	Number of times Conserved	Mutated by
K545	17	A, C, D, E, F, G, H, I, K, L, M, P, Q, T, V, W, Y
A547	14	A, C, E, F, G, H, I, K, L, M, P, Q, R, S
R555	18	A, C, D, E, F, G, H, I, K, L, M, P, Q, R, S, T, V, W
V557	4	T, V, W, Y
C622	3	T, V, W
D623	2	T, W
T680	4	T, V, W, Y
S681	4	T, V, W, Y
S682	4	T, V, W, Y

*T687, A688, N691, S759, D760 are not conserved in any designed proteins

Table 4 SAARs as observed for all the 20 proteins

| S.No	SAARs	Number of SAARs
(1	A	40
2	C	4
3	D	4
4	F	7
5	G	4
6	I	15
7	K	41
8	N	14
9	Q	3
10	R	46
11	S	36
12	T	15
13	V	9
14	Y	12
	Total	250

three SAARs which are observed the most. 'Q' is observed only in 3, while 'C', 'C', and 'G' (varied positions) are the SAARs observed in the vicinity of 6 Å from the centre of the ligand.

Figures 3a and b depict the bar diagram of the number of times a specific SAAR, presented in Table 3, is conserved and in designed proteins and the SAARs causing mutation as observed in all the proteins respectively. From the table it is evident that Fig. 3b depicts the number of times Surrounding Amino Acid Residues (SAAR) are observed in the designed proteins.
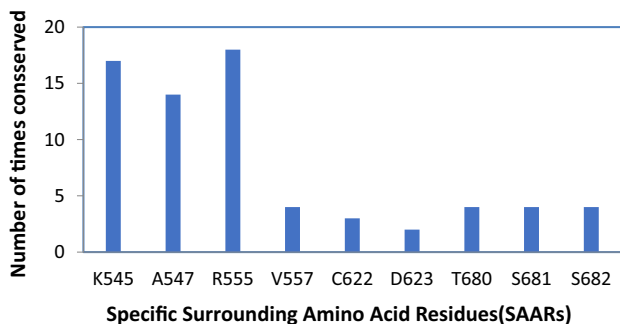
(a) **Specific Surrounding Amino Acid Residues (SAARs) in designed proteins**
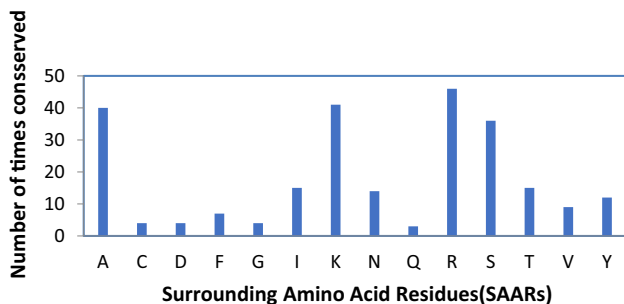
(b) **Surrounding Amino Acid Residues (SAARs) observed in designed proteins**

Fig. 3 a: Specific Surrounding Amino Acid Residues (SAARs) in designed proteins. **b**: Surrounding Amino Acid Residues (SAARs) observed in designed proteins

4.2 Hydrogen bond interactions

The number of hydrogen bonds observed between the ligand (remdesivir) and the parent protein (7bv2) and designed proteins is presented in Table 5. The table also presents the SAARs forming hydrogen bonds with the ligand moiety.

From Table 5, it can be noted that in the parent protein, only one ligand–protein H-bond is observed between N691 and remdesivir. It was surprising that in the designed proteins no H-bond was observed between N691 and the ligand, suggesting drastic outward movement in protein to push N691 beyond 6 Å distance from the ligand. Further, it can be observed that in 13 designed proteins there is an increase in the number of H-bond interactions, while the remaining 6 showed only one H-bond.

Five of the designed proteins (obtained on mutation by M, T, V, W, and Y) showed a higher number of H-bonds than the parent protein. Out of these, four are the same ones which showed a higher number of SAARs (mutated by T, V, W, and Y). AARs Y456, R624, and S682 have been found to be common to all these four designed proteins suggesting their more attractive interactions with the ligand.

Five designed proteins showed 3 H-bonds, while three showed 2 H-bonds. Six designed proteins exhibiting one H-bond are the ones, wherein N691 is mutated

Table 5 Number of H-bonds observed and the SAAR

S.No	Protein	No. of H-bond	H-bond interactions
1	PP	1	N691
2	A	1	R553
3	C	1	R555
4	D	2	I548, R553
5	E	2	S549, R553
6	F	3	S549, R553, R555
7	G	3	I548, S549, R553
8	H	3	K551, R553, R555
9	I	1	R553
10	K	1	R553
11	L	3	S549, R553, R555
12	M	4	I548, S549, R553 R555
13	P	1	R553
14	Q	1	R555
15	R	3	K551, R553, R555
16	S	2	S549, R553
17	T	4	Y456, R555, R624, S682
18	V	6	Y456, R553, R555, A558, R624, S682
19	W	6	Y456, R553, R555, A558, R624, S682
20	Y	5	Y456, R552, A558, R624, S682

by A, C, I, K, P, and Q. Even though they exhibit the same number of H-bond as the parent protein, but with different AARs (R553 and R555).

The number of H-bonds exhibited by various AARs and all AARs with the ligand (Remdesivir) are presented in Tables 6 and 7 respectively.

Table 6 Number of H-bonds formed by specific AARs

S.No	AARs	Number of H-bonds
1	Y456	4
2	I548	3
3	S549	6
4	K551	2
5	R552	1
6	R553	15
7	R555	10
8	A558	3
9	R624	4
10	S682	4
11	N691	1
	Total	53

Table 7 Number of H-bonds formed by AARs

S.No	AAR	Total H-bonds	Positions of AARs
1	A	3	558
2	I	3	548
3	K	2	551
4	R	30	552, 553, 555, 624
5	S	10	549, 682
6	Y	4	456

Table 6 presents the number of H-bonds formed by specific AARs. R553 forms the maximum number of H-bonds (15), followed by R555 (10 H-bonds). R552 and N691 form just one H-bond.

None of the AARs exhibiting H-bond interaction in Table 6 are conserved in the mutated proteins suggesting a drastic change in the conformation of the protein on mutations.

Table 7 shows the total number of H-bonds formed by various AARs causing mutation.

From Table 7 it is further observed that in 15 of the designed proteins R553 forms H-bond, while R555 forms 10, and out of these 7 are common to both. Other than these, S549 forms 6, Y456, R624, and S682 forms 4, I548 and A558 forms 3, K551 forms 2, and R552 forms 1 H-bond. An increase in H-bonding could be considered for understanding ligand–protein complex stability.

Figures 4A and B show the number of H-bonds formed by specific and all AARs causing mutation, respectively with the ligand (Remdesivir).
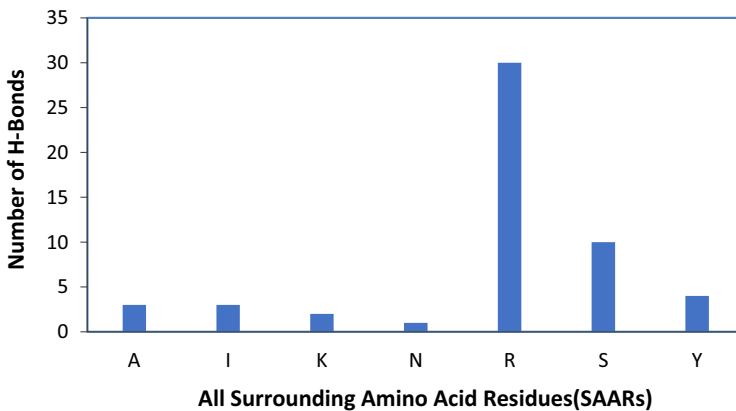
4.3 Van der Waals interactions

Van der Waals interactions between the ligand and protein are observed and presented in Table 8. The table presents the number and the AARs exhibiting van der Waals interactions.

From the Table 8, it is observed that proteins designed by mutating with T, V, W, and Y show a higher number of VDW interactions, suggesting the similarity in pattern as observed for SAARs and H-Bond. Proteins obtained by mutating with W showed equal numbers, i.e., 12, and those with T, V and Y though slightly less i.e., 11, 11, and 10 VDW interactions respectively, than the parent protein. AARs K545 and S681 are conserved in the parent as well as the four designed proteins (designed by mutating with T, V, W, and Y).

The number of VDW interactions exhibited by specific AARs causing mutation is presented in Table 9. The table demonstrates the number and positions of AARs exhibiting VDW interactions in the mutated proteins. Out of the 34 AARs, K545, I548, K551, N552, A554, S814 show more than 10 van der Waals interactions, whereas F441, Y543, V557, T680, S682, T687, A688, S759, and D760 have just
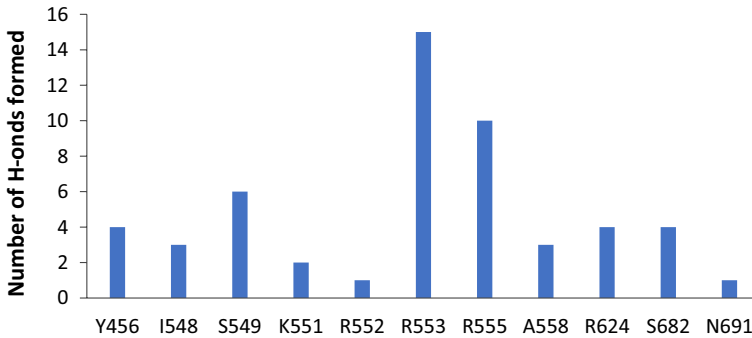
(a) Specific AARs in designed proteins forming H-bond.

(b) Number of H-bonds formed by AARs in designed proteins

Fig. 4 **A**: Specific AARs in designed proteins forming H-bond. **B**: Number of H-bonds formed by AARs in designed proteins

single van der Waals interaction. The higher number of van der Waal interactions are the parameter for stronger protein–ligand interaction.

The total number of VDW interactions exhibited by various mutations causing AARs are presented in Table 10. The table demonstrates the number and positions of AARs exhibiting VDW interactions in the mutated proteins.

From the Table 10, it is evident that 14 of the 20 AARs are showing a total of 165, VDW interactions. It is observed that more than 50% of the contribution to VDW interactions is shown by Lysine (K-36), Alanine (A-25,) and Serine (S-23) located at various positions in the protein chain-A.

The highest number (15) of van der Walls interactions is shown by K545. A554, K551, I548, S814, and N552 are showing 14, 13, 12, 11, and 10 VDW interactions respectively. K545 in 14, R555 in 4, C622 in 3, D623 in 1, and S681

Table 8 Vander Waal interactions in 7bv2 and designed proteins

S.No	Protein	No. of VWI	AARs showing VWI
1	PP	12	K545, R555, V557, C622, D623, T680, S681, S682, T687, A688, S759, D760
2	A	4	I548, K551, N552, A554
3	C	8	K545, A547, I548, S549, K551, A554, S814, R836
4	D	8	K545, S549, K551, N552, A554, R555, S814, R836
5	E	9	F442, Q444, I548, A550, K551, N552, A554, R555, S814
6	F	7	K545, I548, A550, K551, N552, A554, R836
7	G	6	K545, A550, K551, R555, S814, R836
8	H	9	F442, Q444, K545, I548, S549, A550, N552, A554, S814
9	I	5	K545, I548, K551, A554, S814
10	K	5	Y546, I548, K551, A554, S814
11	L	8	F442, K545, I548, A550, K551, N552, A554, R836
12	M	7	F441, K545, K551, N552, A554, S814, R836
13	P	9	F442, K545, I548, S549, K551, N552, A554, R555, R836
14	Q	8	K545, A547, I548, S549, K551, A554, S814, R836,
15	R	7	F442, A547, I548, A550, N552, A554, S814
16	S	9	F442, Q444, A547, I548, K551, N552, A554, S814, R836
17	T	11	Y453, Y455, T540, K545, T556, K621, C622, V667, K676, S681, G683
18	V	11	Y455, T540, Y543, K545, T556, K621, C622, V667, K676, S681, G683
19	W	12	Y453, Y455, T540, K545, T556, K621, C622, D623, V667, K676, S681, G683
20	Y	10	Y453, Y455, T540, K545, T556, K621, V667, K676, S681, G683

in 4 are conserved, while T680, S682, T687, A688, S759 and D760 are not conserved in any mutated protein.

Figures 5A and B show the number of VDW interactions shown by specific AARs and the contribution of mutation-causing AARs, respectively, with the ligand (Remdesivir).

4.4 π-alkyl interactions

Number of π-alkyl interactions are presented in Table 11. Further, the positions of AARs forming π-alkyl interactions are also shown in the table.

From Table 11, it is evident that only one (A547) amino acid residue is showing interactions with the parent compound, which is conserved in 11 mutated proteins. Proteins mutated with C, D and S are not showing any π-alkyl interactions. While G, H, M, P, R, T and W are showing the same number of π-alkyl interactions as parent proteins but in some cases with different amino acid residues (K551 and A558).

**Table 9** Number of VDW interactions exhibited by specific AARs

S.No	AARs	Number of VDW Interactions
1	F441	1
2	F442	6
3	Q444	3
4	Y453	3
5	Y455	4
6	T540	4
7	Y543	1
8	K545	15
9	Y546	1
10	A547	4
11	I548	12
12	S549	5
13	A550	6
14	K551	13
15	N552	10
16	A554	14
17	R555	5
18	T556	4
19	V557	1
20	K621	4
21	C622	4
22	D623	2
23	V667	4
24	K676	4
25	T680	1
26	S681	5
27	S682	1
28	G683	4
29	T687	1
30	A688	1
31	S759	1
32	D760	1
33	S814	11
34	R836	9
	Total	165

Other than these 9 such mutated proteins (A, E, F, I, K, M, Q, V and Y) are there which are showing a higher number of interactions than the parent protein.

M542 is showing 2 ligand protein interactions (V and Y), K545 in E and Q, K551 in R, R555 in A, F, I, K and L, and A558 in T, V, W and Y, while the highest

Table 10 AARs, number of VDW interactions, and positions of AARs

S.No	AARs	No. of VDW interactions	Positions of AARs
1	A	25	547, 550, 554, 688
2	C	4	622
3	D	3	623, 760
4	F	7	441, 442
5	G	4	683
6	I	12	548
7	K	36	545, 551, 621, 676
8	N	10	552
9	Q	3	444
10	R	14	555, 836
11	S	23	549, 681, 682, 759, 814
12	T	10	540, 556, 680
13	V	5	667
14	Y	9	453, 455, 543, 546
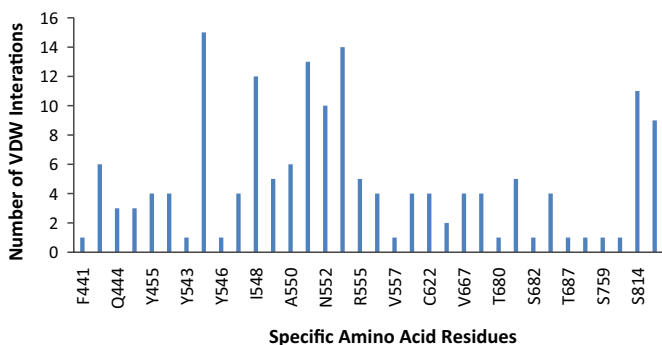
(a) Bar chart showing the specific AARs exhibiting van der Waal interactions.
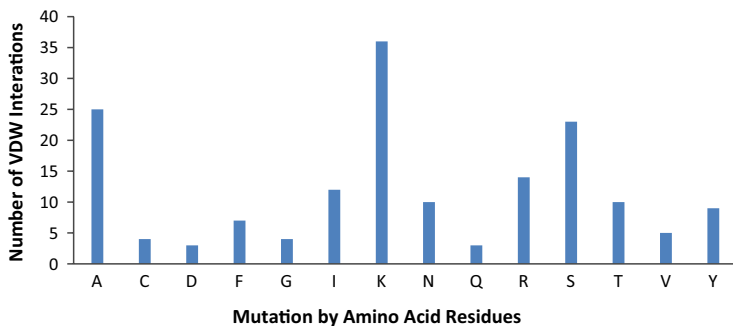
(b) Bar chart showing van der Waal interactions in designed proteins.

Fig. 5 A: Bar chart showing the specific AARs exhibiting van der Waal interactions. **B**: Bar chart showing van der Waal interactions in designed proteins

Table 11 π-alkyl interactions table

S.No	Protein	No. of π-alkyl-π-alkyl interactions	Positions of AARs
1	PP	1	A547
2	A	2	A547 R555
3	C	–	–
4	D	–	- π-alkyl interactions between the ligand and protein molecule
5	E	2	K545 A547
6	F	2	A547 R555
7	G	1	A547
8	H	1	A547
9	I	2	A547 R555
10	K	2	A547 R555
11	L	2	A547 R555
12	M	1	A547
13	P	1	A547
14	Q	2	K545A547
15	R	1	K551
16	S	–	–
17	T	1	A558
18	V	2	M542 A558
19	W	1	A558
20	Y	2	M542 A558

Table 12 Mutation causing AARs, number of π-alkyl interactions and positions of AARs

S.No	AARs	No. of π-alkyl interactions	Positions of AARs
1	A	15	547 558
2	R	5	555
3	K	3	545, 551
4	M	2	542

number of interactions are shown by A547 in 11 mutated proteins (A, E, F, G, H, I, K, L, M, P and Q) and 1 with parent compound.

The total number of π-alkyl interactions exhibited by various mutation-causing AARs is presented in Table 12. The table demonstrates the number and positions of AARs exhibiting VDW interactions in the mutated proteins.

The graphical representations contribution of specific and mutation-causing amino acids in π-alkyl interactions are presented in Figs. 6A and B respectively. A547 shows the maximum number of π-alkyl interactions.
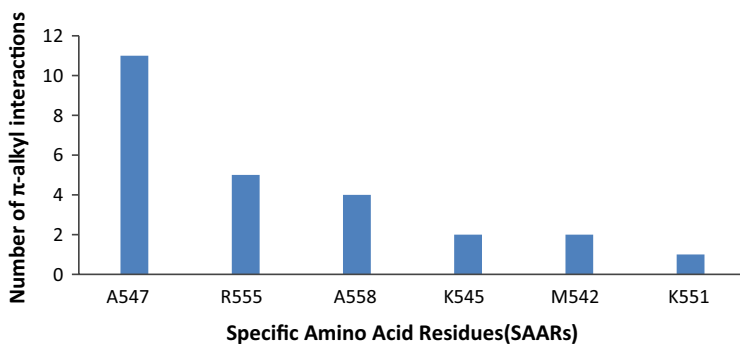
(a) Bar chart showing the specific AARs exhibiting π-alkyl interactions
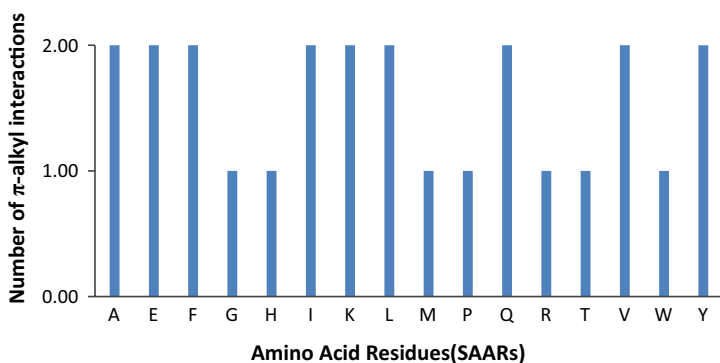
(b) Bar chart showing the specific AARs exhibiting π-alkyl interactions

Fig. 6 **A**: Bar chart showing the specific AARs exhibiting π-alkyl interactions

5 Electrostatic interactions (Energy)

The proportion and distribution of polar and charged residues govern the electrostatic properties of a protein. The electrostatic properties in the protein are controlled by these residues by means of forming short-range interactions and by defining the overall electrostatic environment. The mechanisms of formation of protein–protein, molecular recognitions, conformational adaptabilities, protein movements, etc. are some of the properties in which electrostatics play an important role.

The electrostatic interactions observed in all the proteins (parent and mutated or designed) are presented in Table 13.

From Table 13 it can be observed that higher values are obtained for the proteins designed by mutating with T, V, W, and Y. A bar chart depicting electrostatic interactions for 7bv2 and designed protein is presented in Fig. 7.

Fig. 8a–s, show the summary of all the types of interactions thus observed in the parent and the designed protein in a combined image table form
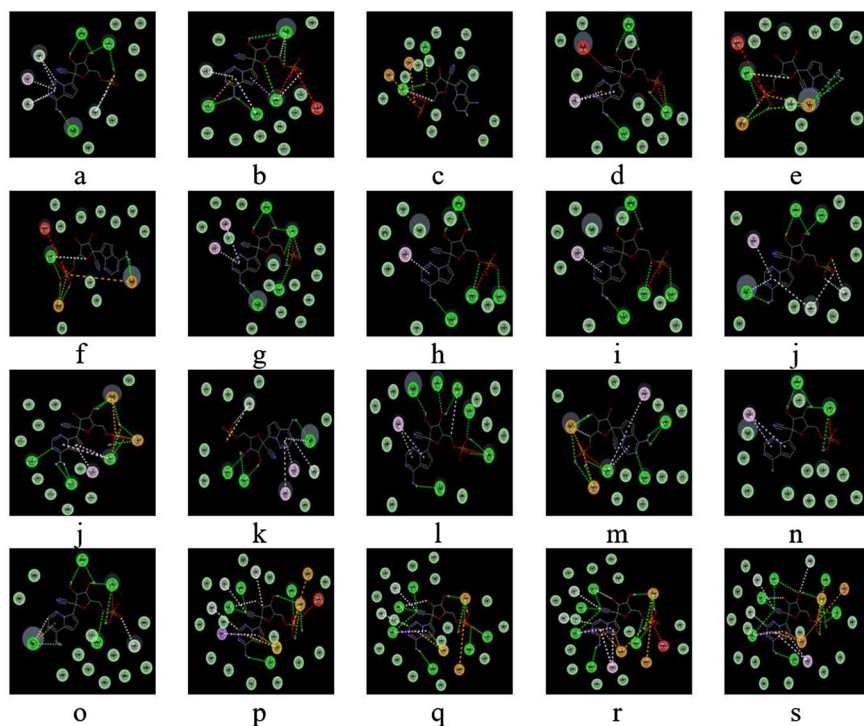
The colour coding of various interactions is given herewith.

Table 13 Designed proteins and the electrostatic interactions

S.No	AARs	Electrostatic interactions [KJ/Mol]
1	PP	− 0.45
2	A	− 0.5
3	C	− 0.05
4	D	− 0.45
5	E	− 0.42
6	F	− 0.49
7	G	− 0.49
8	H	− 0.48
9	I	− 0.38
10	K	− 0.24
11	L	− 0.45
12	M	− 0.41
13	P	− 0.51
14	Q	− 0.1
15	R	− 0.34
16	S	− 0.04
17	T	− 0.94
18	V	− 0.81
19	W	− 0.87
20	Y	− 1.09
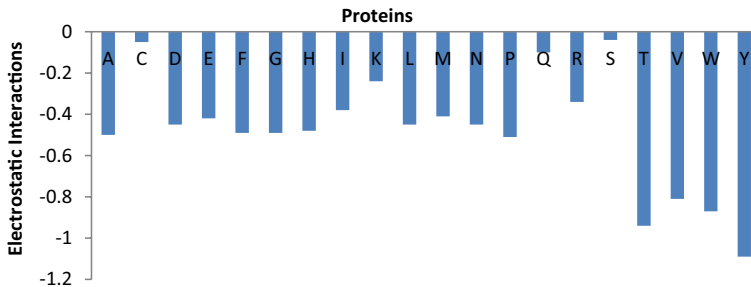
Fig. 7 Electrostatic interactions for 7bv2 and designed proteins

In Fig. 8a, the single H-bond between N691 and remdesivir is observed, while the highest number of H-bonds can be visualized in Fig. 8p–s. Maximum atoms of ligand structure are in bonding with the protein amino acids. From the figure it should be noted that O, H, and N of the ligand are participating in Hydrogen bonding. Various interactions are depicted using different colours. The details of all the interactions namely, van der Waals, π-alkyl interactions and all

The colour coding of various interactions is given herewith

Fig. 8 a–s: Combined images showing all interactions of parent and designed proteins with Remdesivir

other relevant information, presented in Fig. 8a–s have already been depicted in Tables 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 and 13.

6 Conclusions

The present work discusses the effect of in silico-induced mutation on various interactions and the associated ligand–protein interactions. From the results thus obtained it can be concluded that:

The parent protein pdb (7bv2) showed only one H-bond suggesting weak binding of the ligand with the protein. The results are supported by the results obtained for SAARs and VDW interactions. On the other hand, the designed proteins obtained on

mutation with T, V, W, and Y show a similar trend in properties used in the present investigation. All of these have behaved in a common manner and have shown (a) a Higher number of SAARs than the parent protein, (b) a Higher number of H-Bonds than the parent protein, and (c) a Higher number of VDW interactions than the parent protein. The AARs A, R, K, and M are the common residues present in all the proteins designed by inducing mutation by T, V, W, and Y. The work suggests that theoretically induced mutations lead to some interesting results. The missense mutations using the interaction between drug & receptor, which is relatively, a new method, may help to design better compounds and leverage the mutant structures of viruses like Corona or HIV which change faster than others. It discusses the effect of in silico- induced mutation on various interactions and the associated ligand–protein interactions. These can be used for further investigation of other ligand–protein interaction studies and explore their behaviour to extract better binding and active molecules. Analyzing in silico missense mutations and their impact on Ligand–Protein interactions in COVID-19 protein can facilitate drug discovery for combating the virus. The study revealed that the binding of remdesivir with mutated proteins is notably stronger than with the parent protein. This suggests remdesivir's potential efficacy against similar mutations which can be further investigated. These findings hold significant importance for repurposing small molecules to combat COVID infection, serving as valuable insights for forthcoming experimental investigations.

As observed in a work binding of remdesivir with all the mutated protein is comparatively better than a parent protein. From this work we can predict that remdesivir can be helpful if such types of mutations are observed.

Declarations

Competing interests Not Applicable.

Ethical approval Not Applicable.

References

1. A. Chhetri, S. Chettri, P. Rai, B. Sinha, D. Brahman, J. Mol. Struct. (2021). https://doi.org/10.1016/j.molstruc.2020.129178
2. T. Ben Hadda, M. Berredjem, F.A. Almalki, V. Rastija, J. Jamalis, T. Bin Emran, T. Abu-Izneid, E. Esharkawy, L.C. Rodriguez, A.M. Alqahtani, J. Biomol. Struct. Dyn. **40**, 9429 (2022)
3. C. Carlsten, M. Gulati, S. Hines, C. Rose, K. Scott, S.M. Tarlo, K. Torén, A. Sood, R.E. Hoz, Am. J. Ind. Med. **64**, 227 (2021)
4. B. Hu, H. Guo, P. Zhou, Z.-L. Shi, Nat. Rev. Microbiol. **19**, 141 (2021)
5. S. Baloch, M.A. Baloch, T. Zheng, X. Pei, Tohoku J. Exp. Med. **250**, 271 (2020)

6. C.B. Jackson, M. Farzan, B. Chen, H. Choe, Nat. Rev. Mol. Cell Biol. **23**, 3 (2022)
7. D.S.N.B.K. Prasanth, M. Murahari, V. Chandramohan, S.P. Panda, L.R. Atmakuri, C. Guntupalli, J. Biomol. Struct. Dyn. **39**, 4618 (2021)
8. A. Sadeghi Dousari, M. Taati Moghadam, N. Satarzadeh, Infect Drug Resist **13**, 2819 (2020)
9. D. Mishra, R.R. Maurya, K. Kumar, N.S. Munjal, V. Bahadur, S. Sharma, P. Singh, I. Bahadur, J. Mol. Liq. **335**, 116185 (2021)
10. G. Pascarella, A. Strumia, C. Piliego, F. Bruno, R. Del Buono, F. Costa, S. Scarlata, F.E. Agrò, J. Intern. Med. **288**, 192 (2020)
11. M. Ciotti, M. Ciccozzi, A. Terrinoni, W.-C. Jiang, C.-B. Wang, S. Bernardini, Crit. Rev. Clin. Lab. Sci. **57**, 365 (2020)
12. D. Ferraz, E.B. Mariano, P.R. Manzine, H.F. Moralles, P.C. Morceiro, B.G. Torres, M.R. Almeida, J.C. Soares de Mello, D.A. Nrebelatto, Soc. Indic. Res. **158**, 197 (2021)
13. R. Yan, Y. Zhang, Y. Li, L. Xia, Y. Guo, Q. Zhou, Science **367**, 1444 (2020)
14. A.T. Jamiu, C.H. Pohl, S. Bello, T. Adedoja, S. Sabiu, All Life **14**, 1100 (2021)
15. D.C. Dinesh, D. Chalupska, J. Silhan, E. Koutna, R. Nencka, V. Veverka, E. Boura, PLoS Pathog. **16**, e1009100 (2020)
16. M. Phadke, S. Saunik, Drug. Dev. Res. **81**, 541 (2020)
17. R. Keni, A. Alexander, P.G. Nayak, J. Mudgal, K. Nandakumar, Front. Public Health **8**, 216 (2020)
18. K. Guruprasad, Curr. Res. Struct. Biol. **4**, 41 (2022)
19. A.A. Elfiky, J. Biomol. Struct. Dyn. **1**, 3204–3212 (2020)
20. J.F.-W. Chan, C.C.-Y. Yip, K.K.-W. To, T.H.-C. Tang, S.C.-Y. Wong, K.-H. Leung, A.Y.-F. Fung, A.C.-K. Ng, Z. Zou, H.-W. Tsoi, G.K.-Y. Choi, A.R. Tam, V.C.-C. Cheng, K.-H. Chan, O.T.-Y. Tsang, K.-Y. Yuen, J. Clin. Microbiol. (2020). https://doi.org/10.3390/ijms21072574
21. A.K. Banerjee, N. Arora, Curr. Top Med. Chem. **20**, 1434 (2020)
22. S. Kashte, A. Gulbake, S.F. El-Amin III., A. Gupta, Hum. Cell **34**, 711 (2021)
23. D. Ndwandwe, C.S. Wiysonge, Curr. Opin. Immunol. **71**, 111 (2021)
24. J.S. Tregoning, K.E. Flight, S.L. Higham, Z. Wang, B.F. Pierce, Nat. Rev. Immunol. **21**, 626 (2021)
25. N. Chaudhary, D. Weissman, K.A. Whitehead, Nat. Rev. Drug Discov. **20**, 817 (2021)
26. A.V. Nesta, D. Tafur, C.R. Beck, Trends Genet. **37**, 717 (2021)
27. R. Sanjuán, P. Domingo-Calap, Cell. Mol. Life Sci. **73**, 4433 (2016)
28. L. van Dorp, M. Acman, D. Richard, L.P. Shaw, C.E. Ford, L. Ormond, C.J. Owen, J. Pang, C.C.S. Tan, F.A.T. Boshier, A.T. Ortiz, F. Balloux, Infection. Genetics Evolut. **83**, 104351 (2020)
29. M. Pachetti, B. Marini, F. Benedetti, F. Giudici, E. Mauro, P. Storici, C. Masciovecchio, S. Angeletti, M. Ciccozzi, R.C. Gallo, D. Zella, R. Ippodrino, J. Transl. Med. **18**, 179 (2020)
30. C.L.D.C. Badua, K.A.T. Baldo, P.M.B. Medina, J. Med. Virol. **93**, 1702 (2021)
31. R. Wang, Y. Hozumi, C. Yin, G.-W. Wei, Genomics **112**, 5204 (2020)
32. E. Nowakowska, J. Michalska, S.S. Michalak, A. Paczkowska, Acta Poloniae Pharmaceutica - Drug Res. **78**, 291 (2021)
33. C.J. Hemmer, M. Löbermann, E.C. Reisinger, Radiologe **61**, 880 (2021)
34. L. Guruprasad, Proteins: Struct. Function Bioinfo. **89**, 569 (2021)
35. R. Sasidharan, C. Chothia, Proc. Natl. Acad. Sci. **104**, 10080 (2007)
36. S. Wu, C. Tian, P. Liu, D. Guo, W. Zheng, X. Huang, Y. Zhang, L. Liu, J Med Virol **93**, 2132 (2021)
37. W. Yin, C. Mao, X. Luan, D.-D. Shen, Q. Shen, H. Su, X. Wang, F. Zhou, W. Zhao, M. Gao, S. Chang, Y.-C. Xie, G. Tian, H.-W. Jiang, S.-C. Tao, J. Shen, Y. Jiang, H. Jiang, Y. Xu, S. Zhang, Y. Zhang, H.E. Xu, Science **368**, 1499 (2020)
38. S. Singh, H. Florez, F1000Res **9**, 502 (2020)
39. Y. Wang, P. Li, S. Rajpoot, U. Saqib, P. Yu, Y. Li, Y. Li, Z. Ma, M.S. Baig, Q. Pan, Sci Rep **11**, 23465 (2021)
40. S. Saikia, M. Bordoloi, Curr Drug Targets **20**, 501 (2019)
41. Molegro Virtual Docker (Trial Version)
42. N.M. O'Boyle, M. Banck, C.A. James, C. Morley, T. Vandermeersch, G.R. Hutchison, J Cheminform **3**, 33 (2011)
43. G.M. Morris, R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell, A.J. Olson, J Comput Chem **30**, 2785 (2009)