Systems Microbiology and Biomanufacturing
https://doi.org/10.1007/s43393-022-00110-x

ORIGINAL ARTICLE

# In silico comparative structural and compositional analysis of glycoproteins of RSV to study the nature of stability and transmissibility of RSV A

**Debanjan Mitra**[1] · **Pradeep K. Das Mohapatra**[1]

Received: 25 March 2022 / Revised: 7 May 2022 / Accepted: 8 May 2022
© Jiangnan University 2022

**Abstract**
The current scenario of COVID-19 makes us to think about the devastating diseases that kill so many people every year. Analysis of viral proteins contributes many things that are utterly useful in the evolution of therapeutic drugs and vaccines. In this study, sequence and structure of fusion glycoproteins and major surface glycoproteins of respiratory syncytial virus (RSV) were analysed to reveal the stability and transmission rate. RSV A has the highest abundance of aromatic residues. The Kyte–Doolittle scale indicates the hydrophilic nature of RSV A protein which leads to the higher transmission rate of this virus. Intra-protein interactions such as carbonyl interactions, cation–pi, and salt bridges were shown to be greater in RSV A compared to RSV B, which might lead to improved stability. This study discovered the presence of a network aromatic–sulphur interaction in viral proteins. Analysis of ligand binding pocket of RSV proteins indicated that drugs are performing better on RSV B than RSV A. It was also shown that increasing the number of tunnels in RSV A proteins boosts catalytic activity. This study will be helpful in drug discovery and vaccine development.

**Keywords** Respiratory syncytial virus · Stability · Intra-protein interactions · Tunnels · Drug discovery

## Introduction

Currently, the entire world is facing the appalling effect of a viral disease, i.e., COVID-19. Death processions are on the rise in every country. Wave after wave continues to show their destructing power. Researchers are pursuing many drugs for the treatment of COVID-19 [1, 2]. Although, vaccination of COVID-19 has started in many countries. The current scenario is forcing us to think about the deadly viruses that do not have any vaccine. Respiratory syncytial virus (RSV) is one of them [3–5]. It appears at a pre-pandemic level every year in the USA and UK [6]. The hideous effect of this virus is mainly reported in European countries [7, 8]. More than one lac children die every year throughout the earth due to this deadly virus [9]. Not only children but adults are also affected by RSV [10]. It causes bronchiolitis in infants, lethal respiratory disease in children, and usually cold in adults [11]. Two types of RSV have been reported till now, i.e., RSV A and RSV B [12]. It had disclosed that RSV A is far more dangerous and has high mortality rate than RSV B [13]. Not only virulence, but RSV A has a higher spreading rate than RSV B [14].

RSV is also known as the Human Respiratory syncytial virus placed under the family Pneumoviridae. It is a negative-sense enveloped virus with single-strand RNA. The linear genome of this virus contained a total of 15,000 nucleotides [15]. It forms a syncytial during the fusing of an infected cell. RSV has 10 genes that encode 11 types of protein [16]. Between them, major surface glycoprotein (G protein) and fusion glycoprotein (F protein) are playing an important role in the RSV life cycle. Major surface glycoprotein helps in the host–virus attachment, whereas fusion glycoprotein shows its activity in the fusion of cell membrane and formation of syncytial [12]. Host–virus cross-talking interactions had studied by researchers where only the unique host genes were discussed [13, 14]. Genomic comparison between RSV A and RSV B reveals their diversity and evolutionary rate. It also shows that the high substitution rate of G protein affects fearful infection of RSV [17].

✉ Pradeep K. Das Mohapatra
pkdmvu@gmail.com

[1] Department of Microbiology, Raiganj University, Raiganj, WB, India

Published online: 27 May 2022

Although in silico analysis of protein sequence has been done on N, G and F protein to check sequence variability. However, it is not sufficient enough to describe RSV strain variations [18]. Researchers are still trying to figure out why there is such a high level of stability, divergency, and severity of RSV A than RSV B. To find answers it is essential to analyze the structures of the proteins in addition to the sequence. Currently, the intra-protein interaction gives clues about protein stability and other properties [19]. Non-covalent interactions like aromatic–aromatic interactions, aromatic–sulfur interactions, cation–pi interactions, ionic interactions, etc., are crucial factors in protein stabilization. It plays a prodigious important role in extremophilic bacteria [20–24] and viruses [25]. Polar residues from protein sequences are mainly responsible for all those intra-protein interactions. Salt bridge, i.e., the interaction between acidic and basic amino acids are shown important role in extremophiles [26]. Mitra and Mohapatra have reported the effect of network aromatic–aromatic interaction and its effectiveness [27].

In silico study of RSV fusion glycoprotein and surface glycoprotein had not been done yet. In this study, sequences and structures of the major surface glycoprotein and the fusion glycoprotein from RSV A and RSV B have been analyzed. A comparative study between two subtypes describes the reasons behind the higher stability, divergency, and spreading rate of RSV A than RSV B. The drug protecting nature of RSV A will be revealed through their structural properties. This study will help to find better drugs and the development of vaccines.

## Materials and methods

### Dataset

All reviewed sequences of the major surface glycoprotein (G) and the fusion glycoprotein (F) of RSV were extracted from the protein database UniProt [28]. The Universal Protein Resource (UniProt) is a database that contains information on protein sequences and annotations. Crystal structures of the subtype RSV A major surface glycoprotein and fusion glycoprotein were retrieved from the largest protein database, RCSB PDB [29]. However, the structure of fusion glycoprotein of RSV B subtype was available in the database, but due to the lack of major surface glycoprotein structure, we have created its model through homology modeling.

### Homology modeling of RSV B surface glycoprotein

Due to the unavailability of the RSV B major surface glycoprotein structure, it was mandatory to construct that model by homology modeling to compare both structures of RSV subtypes. The model of RSV B major surface glycoprotein (O36633) was done by MODELLER V.9.4 [30]. The structure of the major surface glycoprotein from RSV A was considered as a template (P03423/5WNA). Target sequence showed 73.33% sequence identity with template sequence. Validation of model structure has been done by several online servers [31–33].

### Analysis of protein sequences

All reviewed sequences were subjected to multiple sequence alignment (MSA) with the help of Clustal Omega [34]. MSA was used to prepare a block of the protein sequence by Blockmaker [35]. Non-block format of sequences was used to calculate the amino acid compositions, and intrinsic disordered regions whereas a block of protein sequences were used to calculate Kyte–Doolittle hydrophobicity [36], Grantham polarity [37], etc. Amino acid compositions were calculated by the ProtParam server [38, 39]. Intrinsic disordered regions were identified by the DisEMBL server [40]. Hydrophobicity and polarity were calculated by Protscale [38]. Tm predictor was used to calculate the Tm index (http://tm.life.nthu.edu.tw/).

### Analysis of protein crystal structures

From the RCSB PDB database, 5EA4 (fusion glycoprotein) and 5WNA (C chain) (major surface glycoprotein) of RSV A were chosen for analysis. Structures of the protein had been minimized by UCSF Chimera in 1000 steps along with forcefield [41]. Protein minimization techniques have an immense role in protein structural analysis, and it also helps to get valid results [42]. Analysis of secondary structure was done by the CFSSP server [43]. Intra-protein interactions were identified from the PIC server [44] and Arpeggio server [45]. Analysis of pockets, tunnels, pores, and cavities was analyzed by several online servers [46–48].

## Result and discussion

### Study of sequence diversity

Amino acid compositions of a protein not only affect its properties but also have an enormous effect on protein structure and stability. It also affects the rate of mutations and its effect on protein structure [49]. Changing amino acid in a specific position or multiple positions can change protein stability. Generally, proteins that have higher stability show special variation in their amino acid compositions [49, 50]. Protein hydrolysis and functions are also influenced by their compositions [51, 52].

The fusion glycoprotein of RSV A showed a higher number of polar residues than RSV B (Fig. 1A). Although some hydrophobic residues (G, A, F, P, V) showed a slightly higher abundance in RSV A than RSV B. It is also reported that hydrophobic residues of the virus are playing an important role in the formation of the fusion loop in membrane fusion [53]. The increasing number of amino acids G and P in proteins of the RSV A subtype helps to increase protein flexibility. Amino acids N have a role in mediating metabolic crosstalk between the host and pathogen [54]. Charged residues are also played many important roles in the membrane fusion of viruses [55]. Increasing charged residues and some specific hydrophobic residue might affect RSV A host cell fusion. On the other hand, the major surface glycoprotein of RSV A contains a high number of uncharged polar residues (Fig. 1B) than RSV B. Polar residues on the surface of protein increase its stability [56]. Therefore, the abundance of high polar residues in the surface protein of RSV A makes it more stable than RSV B. Hydrophobic residues are almost equal in both subtypes of major surface glycoprotein. But surprisingly aromatic amino acids showed higher abundance in both types of proteins from RSV A. Those residues might play a role in RSV A protein stability. Both proteins (i.e., fusion glycoprotein and major surface glycoprotein) showed almost similar intrinsic disordered regions but there is some marked difference point indicating that the intrinsic disordered regions have a higher effect on RSV A than RSV B (Fig. 1C, D). Intrinsic disordered regions of proteins are playing a predominant role in protein folding, human disease, and cellular complexity [57]. It also affects cell signaling and protein regulations by assisting in the formation of signaling complexes [58]. At various levels, viral IDPs mediate effective infection and govern pathogenesis [59].

The result of Kyte–Doolittle hydrophobicity from both proteins showed the hydrophilic nature of RSV A. That means proteins of RSV A are more interactive with aqueous media than RSV B. RSV generally transfers through the water droplet. Increasing hydrophilicity at the membrane can help RSV A to spread more easily than RSV B. Membrane hydrophilicity also increase many types of interactions like ion–dipole, dipole–dipole, hydrogen bonding between amino acids and water [60]. Increasing interactions of proteins increase its structural conformation and ultimately increase its stability [61]. In the case of polarity, both proteins are showing almost equal polarity in both subtypes, but RSV A showed some discernible difference points, which indicates RSV A has higher polarity than RSV B. It is already reported that proteins with high polarity attract aqueous solvents to increase their stability [62]. The average TM indexes of both proteins are low in RSV A than in RSV B which indicates that RSV A is more love a cold temperature. Therefore, it might be a factor for RSV A to present at a higher rate in winter than RSV B (Fig. 2).

## Evaluation of G glycoprotein model of RSV B

Homology modeling is a principal technique in structural biology and drug discovery [63, 64]. Due to the lack of structure in major surface glycoproteins from RSV B, it was necessary to create a model to analyze their structural properties. The structure of major surface glycoprotein (P03423/5WNA) from RSV A was considered as a template and sequence (O36633) from RSV B was taken as a target. After the preparation of the model, evaluation of the model is an important step to justify the model structure [65]. Global model quality estimation (GMQE) showed 0.3 in the case of model structure. GMQE value ranges between 0 and 1 in the model indicating its higher accuracy. QMEAN value of major surface glycoproteins OF RSV B structure is 0.78 which indicates that it is a good-quality model.

The overall quality factor (Table 1) of the model (100%) showed a higher value than the template (90%) that was used. *Z*-score RMS of the template is 1.197 and 1.075 in the case of model structure. *Z*-score RMS decrease when the resolution of structure and R-factor increase. The number of outliers is also low in the model than template (5WNA). Analysis of the Ramachandran plot (Fig. 3) showed the model allowed more residues in its favored regions than the template. Interestingly model has no residues in its disallowed regions, whereas the template has 3.8% residues in its disallowed region. The evaluation of the model showed that it is an acceptable model that can be used for further structural analysis.

## Secondary structure analysis

Analysis of secondary structure revealed the amino acid propensity in helix, coil, turn and sheet. The presence of amino acids in a specific position can influence protein stability. The higher number of amino acid residues showed the highest abundance in the sheet of both subtypes (Table 2). But interestingly, RSV A increases its amino acid propensity in coil, helix, and turn than RSV B. Previous report suggested that coil and helix have a massive role in protein stability [66, 67]. Uncharged polar amino acids showed higher abundance in the sheet of RSV. The charged polar amino acids show a higher propensity in the helix of both subtype proteins. Hydrophobic residues have a higher plethora in sheets. An increase of charged residues in the helix can increase protein stability [68]. In the case of fusion glycoprotein, both subtypes have 13 helix–helix interactions, 1psi-loop, 4 sheets, 10 beta-hairpins, etc. Between the sheets, 3 were antiparallel and 1 was mixed in type. However, the number of beta-turns was high in RSV B than RSV A subtype. The beta-turn conformation is a useful size for possible imitation by a drug-like small molecule and exhibits an energetically favorable configuration for a tetrapeptide segment. Within

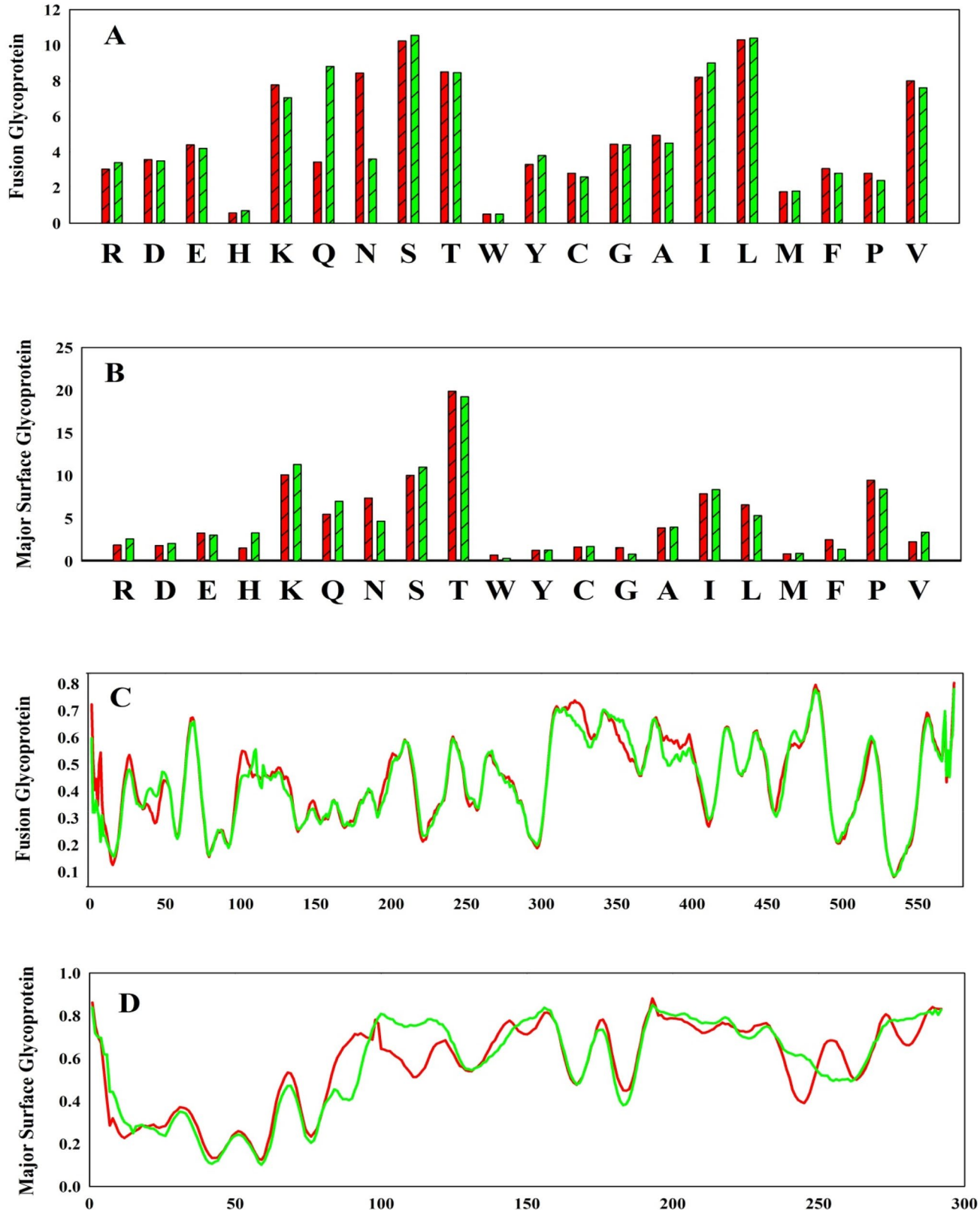


**Fig. 1** **A** Amino acid composition in fusion protein of RSV A and RSV B. **B** Amino acid composition in major surface glycoprotein protein of RSV A and RSV B. **C** Intrinsic disordered regions in fusion protein of RSV A and RSV B. **D** Intrinsic disordered regions in major surface glycoprotein protein of RSV A and RSV B

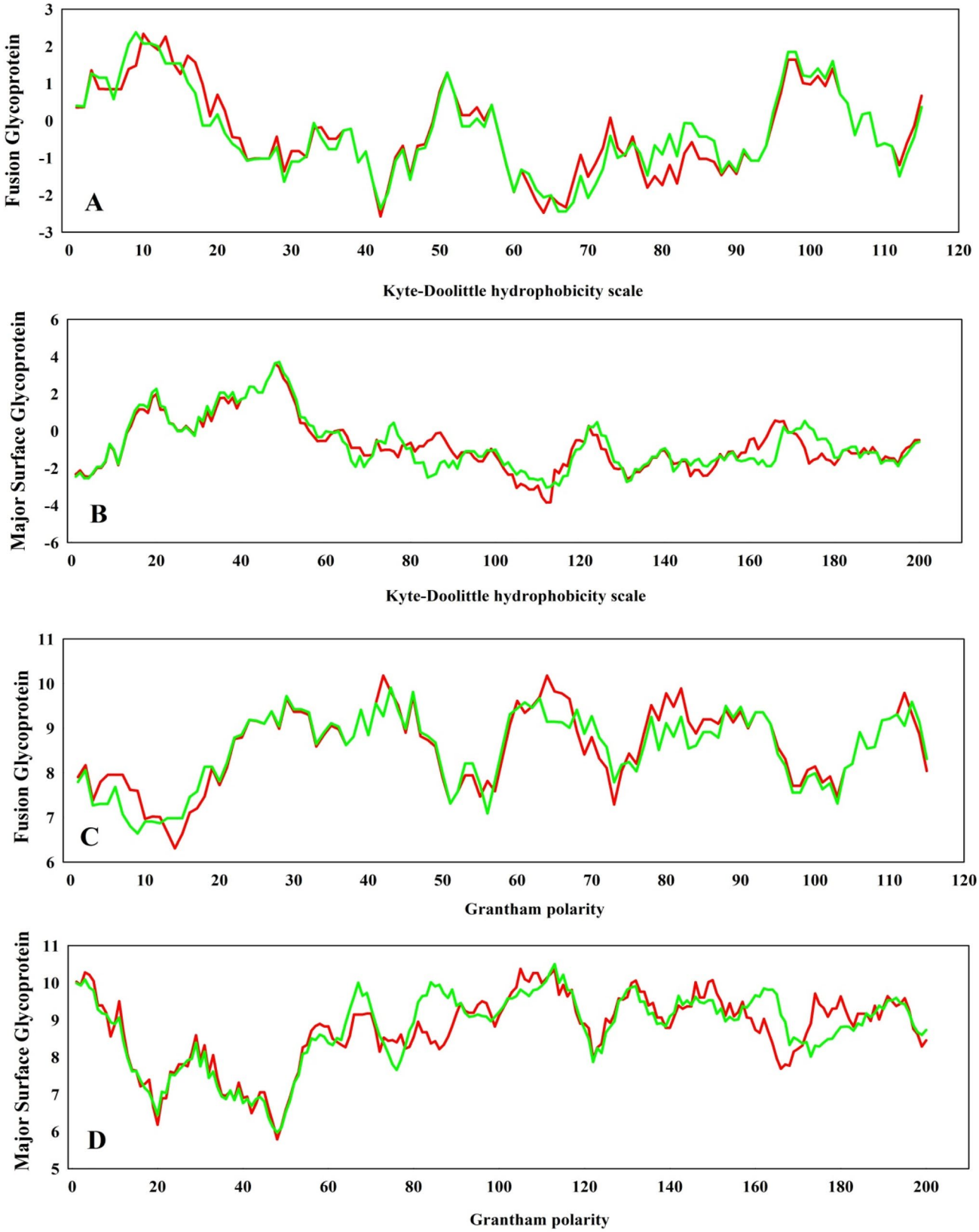


**Fig. 2** **A** Kyte–Doolittle hydrophobicity in fusion protein of RSV A and RSV B. **B** Kyte–Doolittle hydrophobicity in major surface glycoprotein protein of RSV A and RSV B. **C** Grantham polarity in fusion protein of RSV A and RSV B. **D** Grantham polarity in major surface glycoprotein protein of RSV A and RSV B

**Table 1** Properties for evaluation of model in compare to major surface glycoprotein of RSV A (5WNA)

| Properties | 5WNA | Model |
|---|---|---|
| Overall quality factor | 90 | 100 |
| *Z*-score RMS | 1.197 | 1.075 |
| Outliers (%) | 3.9 | 2.6 |
| Residues in most favoured regions | 88.5 | 96.2 |
| Residues in additional allowed regions | 7.7 | 3.8 |
| Residues in generously allowed regions | 0 | 0 |
| Residues in disallowed regions | 3.8 | 0 |

fusion glycoprotein, individual topological structure study of two distinct X-ray crystallographic structures (5EA4 and 6Q0S) revealed variability (Fig. 4). The creation of specific secondary structures is favored by different amino acids. It was also reported that an increasing amount of alanine in the sequence of a protein can form a stable helix during the formation of secondary structure [69]. RSV A specially designed its sequence to gain more fixity through the helix stabilization.

## Effect of intra-protein interactions in protein structures

The increasing aromatic residues in RSV A sequence showed their immeasurable role in increasing the stability of structures through different types of intra-protein interactions. Aromatic residues have participated in three types of interactions, i.e., aromatic–aromatic interactions, aromatic-sulfur interactions, and cation–pi interactions. Aromatic–aromatic interactions are the result of interactions between two aromatic ring distances at 4.5–7 Å [27]. In the case of aromatic–sulfur interactions, aromatic residues form a non-covalent interaction with sulfur-containing amino acids (MET and CYS).

Aromatic–aromatic interactions showed similar interactions in both subtypes in the case of both proteins (Table 3). But all dihedral angles have differed in both subtypes. Only a single network aromatic–aromatic interaction showed the difference in both subtypes from fusion glycoprotein. Fusion glycoprotein showed four network aromatic–aromatic interactions, and six isolated aromatic–aromatic interactions in RSV A and RSV B. Major surface glycoprotein showed only one isolated aromatic–aromatic interaction in each subtype. Major surface glycoprotein shows a small number of intra-protein interactions due to its small size. The database does not have the full-length structure of major surface glycoprotein. Most PHE amino acids have participated at a higher number in aromatic–aromatic interactions.

Aromatic-sulfur interaction showed a higher number in the fusion glycoprotein of RSV B than RSV A (Table 4). In the case of fusion glycoprotein, RSV A has five isolated, and two network aromatic–sulfur interactions, while RSV B has five isolated, and three network aromatic–sulfur interactions. The one extra network aromatic–sulfur interaction might give a slight advantage to the fusion glycoprotein of RSV B, but it does not have an edge in the case of major surface glycoprotein. The major surface glycoprotein of RSV A has one long network aromatic–sulfur interaction, which is furnished by one TRP and three CYS residues. RSV B has only one isolated aromatic–sulfur interaction in its major surface glycoprotein. Therefore, aromatic–sulfur interactions have an enormous role in the stability of RSV A major surface glycoprotein. However, this is the first report of a network aromatic-sulfur interaction in viral proteins.

Basic residues are enhancing the stability of RSV A protein by the formation of cation–pi interaction [24]. The fusion glycoprotein, RSV A contains five cation–pi interactions, whereas RSV B has three cation–pi interactions (Table 5). In the case of major surface glycoprotein from RSV A has two cation–pi interactions, and RSV B has only one cation–pi interaction. Most LYS residues have participated at a higher rate in cation–pi interactions. Increasing cation–pi interactions in RSV A might play a crucial role in its stability.

Non-covalent interactions between an acidic and a basic residue at a 4 Å distance indicate the formation of a salt bridge [24, 25]. When single pair amino acid residues are conforming this interaction, it is called the isolated salt bridge. Formations of network salt bridges of protein are the result of interaction between multiple pairs of isolated salt bridges. RSV A has 16 pairs of isolated salt bridges and four pairs of network salt bridges whereas RSV B has 17 pairs of isolated salt bridges and only a single pair of network salt bridges (Table 6). Salt bridges are the chief factor in the high stability of RSV A. Salt bridges increase protein stability and help to stabilize them in high temperatures, high salt concentrations, and highly acidic environments [24, 25, 27]. RSV A is stabling itself by increasing the number of network salt bridges to gain more stability over another subtype. Mostly GLU as acidic residues and LYS as basic residues have a higher abundance in salt bridge formation. In the case of basic residues, LYS is almost double the number of ARG. Those residues that participated in cation–pi interactions have also participated in salt bridge formation. Surprisingly, HIS is absent in the formation of salt bridges within 4A distance.

The number of carbonyl interactions in the fusion protein of RSV A is 57 and 53 in RSV B. Major surface glycoproteins of both subtypes showed two carbonyl interactions. It was reported that this type of short interaction can contribute energy to protein stability [70, 71]. Therefore, the high number of carbonyl interactions in RSV A has a role in its stability over RSV B.

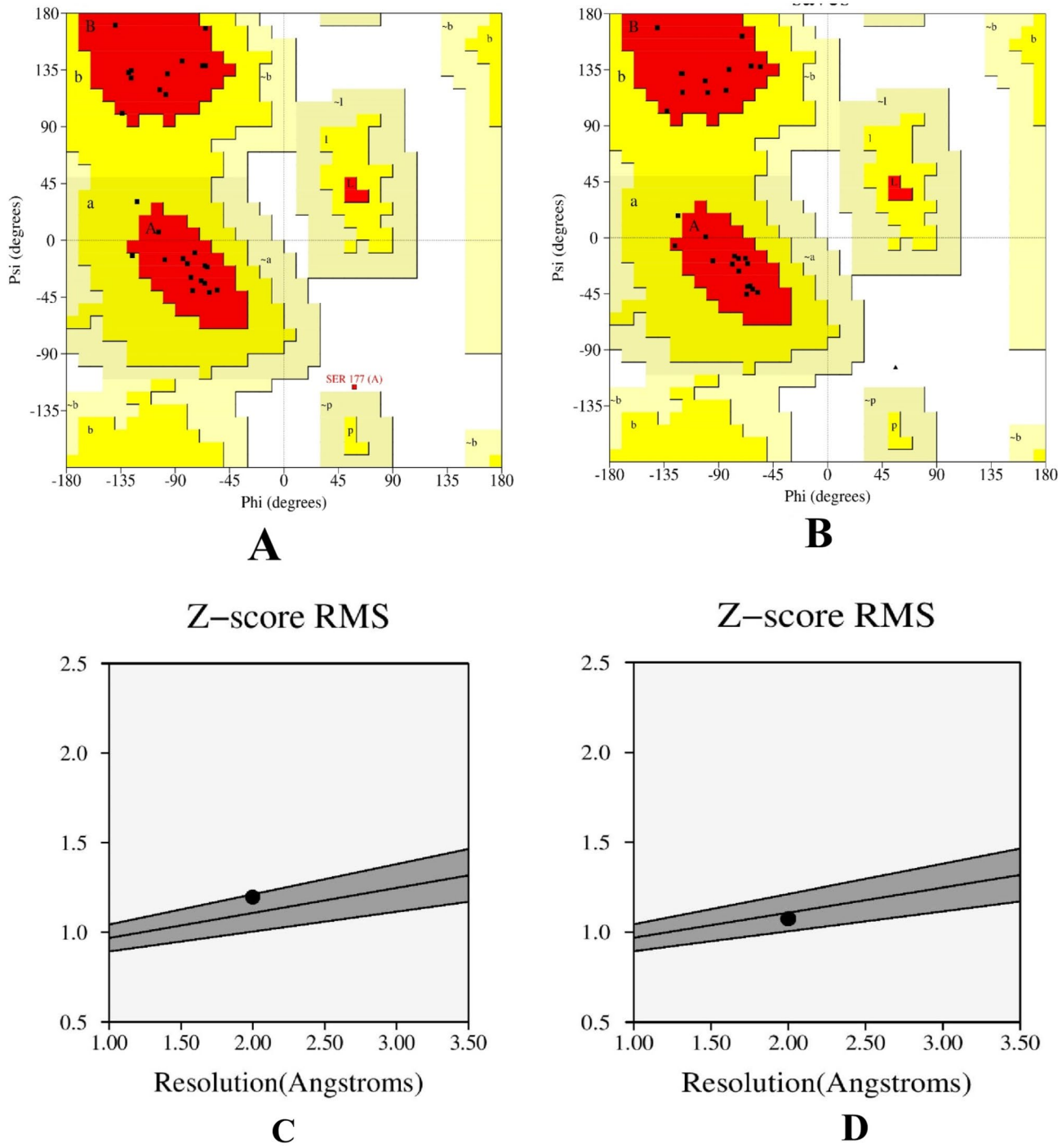


**Fig. 3** **A** Ramachandran plot of major surface glycoprotein (5WNA) from RSV A. **B** Ramachandran plot of major surface glycoprotein model structure from RSV B. **C** *Z*-score RMS of major surface glycoprotein (5WNA) from RSV A. **D** *Z*-score RMS of major surface glycoprotein model structure from RSV B.

## Analysis of pocket, tunnels, cavities of RSV proteins

Pockets are mainly present on the surface of proteins and play an essential role in the attachment of drugs or small molecules [72]. Identification of pockets is very important for the discovery of new drugs by molecular docking [73]. The number of pockets in fusion glycoprotein and major surface glycoprotein shows a higher number in RSV B than in RSV A (Table 7). In the case of fusion glycoprotein, RSV A has 58 pockets, and RSV B

**Table 2** Secondary structure analysis of fusion glycoprotein and major surface glycoprotein to revealed amino acid positions in coil, sheet, helix and turn

| Position | RSV A | | | RSV B | | |
|---|---|---|---|---|---|---|
| | Uncharged polar | Charged | Hydrophobic | Uncharged polar | Charged | Hydrophobic |
| Fusion glycoprotein | | | | | | |
| COIL | 5.75 | 2.44 | 4.88 | 1.57 | 1.39 | 2.79 |
| SHEET | 15.85 | 2.79 | 23.52 | 25.44 | 5.05 | 30.66 |
| HELIX | 11.67 | 11.67 | 14.11 | 6.79 | 10.45 | 9.06 |
| TURN | 4.36 | 2.26 | 0.70 | 4.18 | 1.92 | 0.70 |
| Major surface glycoprotein | | | | | | |
| COIL | 8.39 | 3.02 | 5.70 | 7.05 | 5.03 | 4.03 |
| SHEET | 28.52 | 5.37 | 19.80 | 31.54 | 6.71 | 18.79 |
| HELIX | 4.03 | 7.38 | 6.71 | 5.37 | 7.05 | 4.03 |
| TURN | 5.70 | 2.35 | 3.02 | 4.03 | 3.69 | 3.02 |

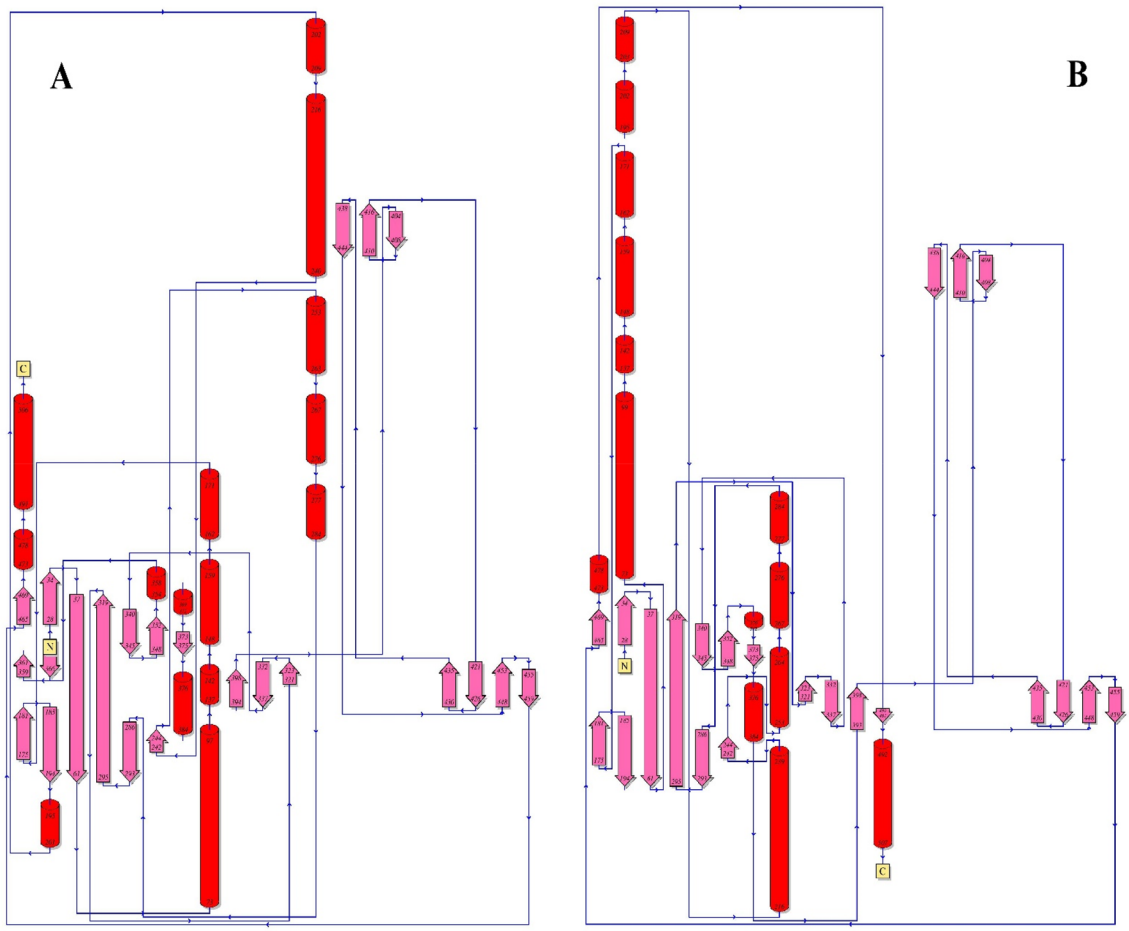


**Fig. 4** Topology structure of fusion protein from **A** RSV A (5EA4) and **B** RSV B (6Q0S) showed difference in their secondary structures

has 72 pockets. On the flip side, the major surface glycoprotein of RSV A has two pockets, and RSV B has three pockets. That means RSV B can deal more with drugs or ligand molecules. The pocket area of the major surface glycoprotein of RSV A is 13.46 and 18.97 $Å^2$ in the case of RSV B. However, the pocket area of fusion glycoprotein in RSV A is higher than in RSV B. A similar effect showed on the volume of pockets. But intriguingly, the

**Table 3** Network (Colour pair) and isolated (black colour) aromatic–aromatic interactions in fusion glycoprotein and major surface glycoprotein of RSV A and RSV B

| Protein name | 5EA4 | | | 6Q0S | | |
|---|---|---|---|---|---|---|
| | Position | D(centroid-centroid) | Dihedral Angle | Position | D(centroid-centroid) | Dihedral Angle |
| Fusion glycoprotein | Y441-Y468 | 6.5 | 116.08 | Y441-Y468 | 6.54 | 62.03 |
| | F32-Y441 | 5.15 | 12.8 | F32-Y441 | 5.17 | 169.78 |
| | F32-Y468 | 5.08 | 115.73 | F32-Y468 | 5.08 | 117.36 |
| | Y86-F223 | 6.86 | 52.06 | Y86-F223 | 6.89 | 49.12 |
| | Y198-F223 | 5.96 | 153.67 | Y198-F223 | 5.93 | 156.04 |
| | F137-F488 | 6.45 | 85.53 | F137-F140 | 6.75 | 102.84 |
| | F140-F488 | 5.87 | 134.63 | F137-F488 | 5.02 | 39.91 |
| | W314-Y342 | 5.83 | 176.88 | W314-Y342 | 5.77 | 175.28 |
| | Y342-F351 | 4.65 | 30.11 | Y342-F351 | 4.56 | 28.4 |
| | Y44-W341 | 4.82 | 86.88 | Y44-W341 | 4.87 | 96.39 |
| | W52-Y286 | 5.83 | 112.41 | W52-Y286 | 5.85 | 118.72 |
| | F237-Y299 | 5.81 | 109.44 | F237-Y299 | 5.67 | 103.6 |
| | Y306-F366 | 5.46 | 72.85 | Y306-F366 | 5.28 | 69.65 |
| | F387-F468 | 6.21 | 50.9 | F387-F468 | 6.54 | 40.91 |
| | Y478-F483 | 5.38 | 71.18 | Y478-F483 | 5.03 | 49.95 |
| | 5WNA | | | Model | | |
| Major surface glycoprotein | F163-F168 | 5.35 | 146.48 | F163-F168 | 6.41 | 144.78 |

**Table 4** Network (Colour pair) and isolated (black colour) aromatic–sulphur interactions in fusion glycoprotein and major surface glycoprotein of RSV A and RSV B

| Protein name | 5EA4 | | | 6Q0S | | |
|---|---|---|---|---|---|---|
| | Position | D(centroid-centroid) | Dihedral Angle | Position | D(centroid-centroid) | Dihedral Angle |
| Fusion glycoprotein | F32-C439 | 5.27 | 94.62 | F32-C439 | 5.23 | 93.56 |
| | Y53-M264 | 4.9 | 134.01 | Y53-M264 | 4.83 | 141.24 |
| | F137-M396 | 4.74 | 23.4 | Y198-M226 | 5.25 | 146.79 |
| | Y391-C382 | 4.9 | 65.48 | Y391-C382 | 4.85 | 66.5 |
| | Y417-C322 | 4.06 | 26.15 | Y417-C322 | 4.14 | 27.52 |
| | F237-M251 | 4.12 | 144.41 | F137-M396 | 4.61 | 49.92 |
| | F237-M289 | 4.29 | 41.24 | F488-M396 | 5.21 | 32.43 |
| | Y299-M251 | 4.43 | 23.47 | F237-M251 | 4.01 | 148.89 |
| | F352-C358 | 4.26 | 149.5 | F237-M289 | 4.97 | 35.33 |
| | F352-C367 | 3.83 | 168.66 | Y299-M251 | 4.46 | 30.09 |
| | | | | F352-C358 | 4.05 | 155.95 |
| | | | | F352-C367 | 3.88 | 173.74 |
| | 5WNA | | | Model | | |
| Major surface glycoprotein | W183-C173 | 4.73 | 130.05 | F165-C182 | 4.45 | 50.62 |
| | W183-C176 | 5.09 | 153.47 | | | |
| | W183-C186 | 4.35 | 133.09 | | | |

**Table 5** Cation–pi interactions in fusion glycoprotein and major surface glycoprotein of RSV A and RSV B

| Protein name | 5EA4 | | | 6Q0S | | |
|---|---|---|---|---|---|---|
| | Position | D (centroid–centroid) | Dihedral angle | Position | D (centroid–centroid) | Dihedral angle |
| Fusion glycoprotein | W52-R49 | 4.12 | 147.41 | W52-R49 | 4.04 | 144.45 |
| | F137-R339 | 3.76 | 146.01 | F190-K176 | 4.46 | 49.01 |
| | F190-K176 | 3.97 | 155.4 | F435-K433 | 5.17 | 136.24 |
| | Y198-K226 | 4.96 | 137.84 | | | |
| | F435-K433 | 4.99 | 133.15 | | | |
| | 5WNA | | | Model | | |
| Major surface glycoprotein | F170-K187 | 5.82 | 170.93 | F170-K187 | 5.89 | 137.2 |
| | W183-R188 | 5.01 | 70.7 | | | |

**Table 6** Isolated and network (colour pair) salt bridge interactions in fusion glycoprotein of RSV A (5EA4) and RSV B (6Q0S)

| 5EA4 | | 6Q0S | |
|---|---|---|---|
| **Isolated** | **Network** | **Isolated** | **Network** |
| E31-K42 | E60-K168 | E31-R42 | E60-K191 |
| R49-D368 | E60-K191 | R49-D368 | E60-K196 |
| E60-K168 | E161-K293 | E66-K80 | |
| E60-K191 | K293-E294 | E82-K85 | |
| E82-K85 | K196-D200 | D84-K87 | |
| D84-K87 | K196-E295 | E161-K293 | |
| E163-K166 | K427-D448 | K168-E294 | |
| K176-D263 | D448-K461 | K176-D263 | |
| E222-K226 | | R202-E222 | |
| R229-E256 | | R229-E256 | |
| E232-R235 | | E232-R235 | |
| D310-R364 | | D310-R364 | |
| R336-D338 | | R336-D338 | |
| K394-D489 | | K394-D489 | |
| K433-D440 | | K433-D440 | |
| E487-K498 | | K445-E463 | |
| | | D448-K461 | |

**Table 7** No. of pockets, area and volume of pocket, tunnels, cavities, voids of major surface glycoprotein and fusion glycoprotein from RSV A (5WNA and 5EA4) and RSV B (Model and 6Q0S)

| Properties | Major surface glycoprotein | | Fusion glycoprotein | |
|---|---|---|---|---|
| | 5WNA | Model | 5EA4 | 6Q0S |
| No. of pockets | 2 | 3 | 58 | 72 |
| Pocket area ($Å^2$) | 13.46 | 18.97 | 811.449 | 686.478 |
| Volume ($Å^3$) | 3.303 | 4.402 | 290.221 | 201.48 |
| Tunnels | 0 | 0 | 24 | 21 |
| Cavities | 1 | 1 | 19 | 20 |
| Voids | 0 | 0 | 8 | 4 |

fusion protein of RSV B has a larger single pocket than RSV A. It can be targeted by any drug molecule to inhibit its function. Therefore, RSV B can be smoothly interplayed with drugs or ligand molecules. In another word, fewer pockets and the small size of individual pockets increase anti-drug properties in RSV A.

Tunnels are connecting the internal spaces of proteins with the exterior portion, by which substrate, co-factors, and ions move towards enzymes' active sites. Not only transport but also tunnels are playing a major role in enzyme catalytic activity [74]. The fusion protein of RSV A (5EA4) showed a higher number of tunnels than RSV B (Fig. 5). Those tunnels might increase the activity of RSV A fusion protein by avoiding intermediate side reactions that are premature or undesired. Pores are also a type of channel that passes through the protein from one surface side to another surface side. Pores of RSV A are also larger than RSV B. Cavities are described as a deeper type of pocket [75]. It also showed similar results with pockets. The number of cavities is high in RSV B than in RSV A (Fig. 6). RSV A subtype showed 19 cavities in its fusion protein, whereas subtype B showed 20 cavities. The volume of individual cavities is also quantitatively dependent

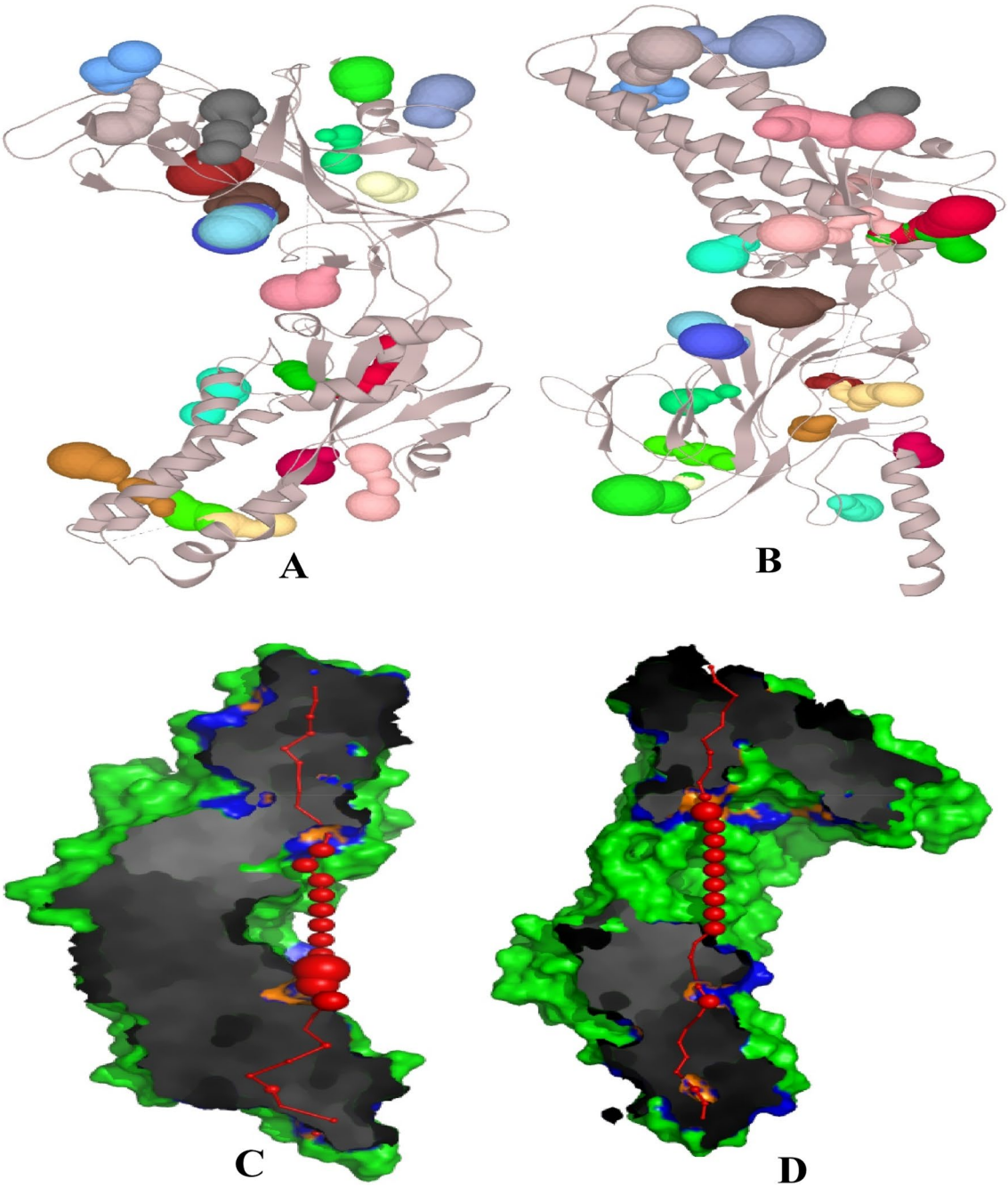


**Fig. 5** **A** Tunnels in fusion glycoprotein (5EA4) of RSV A. **B** Tunnels in fusion glycoprotein (6Q0S) of RSV B. **C** Pore in fusion glycoprotein (5EA4) of RSV A. **D** Pore in fusion glycoprotein (6Q0S) of RSV B

on the number of amino acids that were lining the cavity. The best druggable cavity might be targeted for RSV B drug discovery using pharmacophores. At the molecular level, protein cavities play a crucial role in biological processes such as ligand binding. By decreasing the free energy of the natural state for the protein, a reduction in cavity size improves protein stability [76]. Voids are one type of cavities that are situated inside the protein and not occupied by protein atoms [49]. The number of voids showed a higher abundance in RSV A than RSV B. However, a large void space was observed in RSV B which was 131 $Å^3$. Subtype A showed the highest void volume 110 $Å^3$.

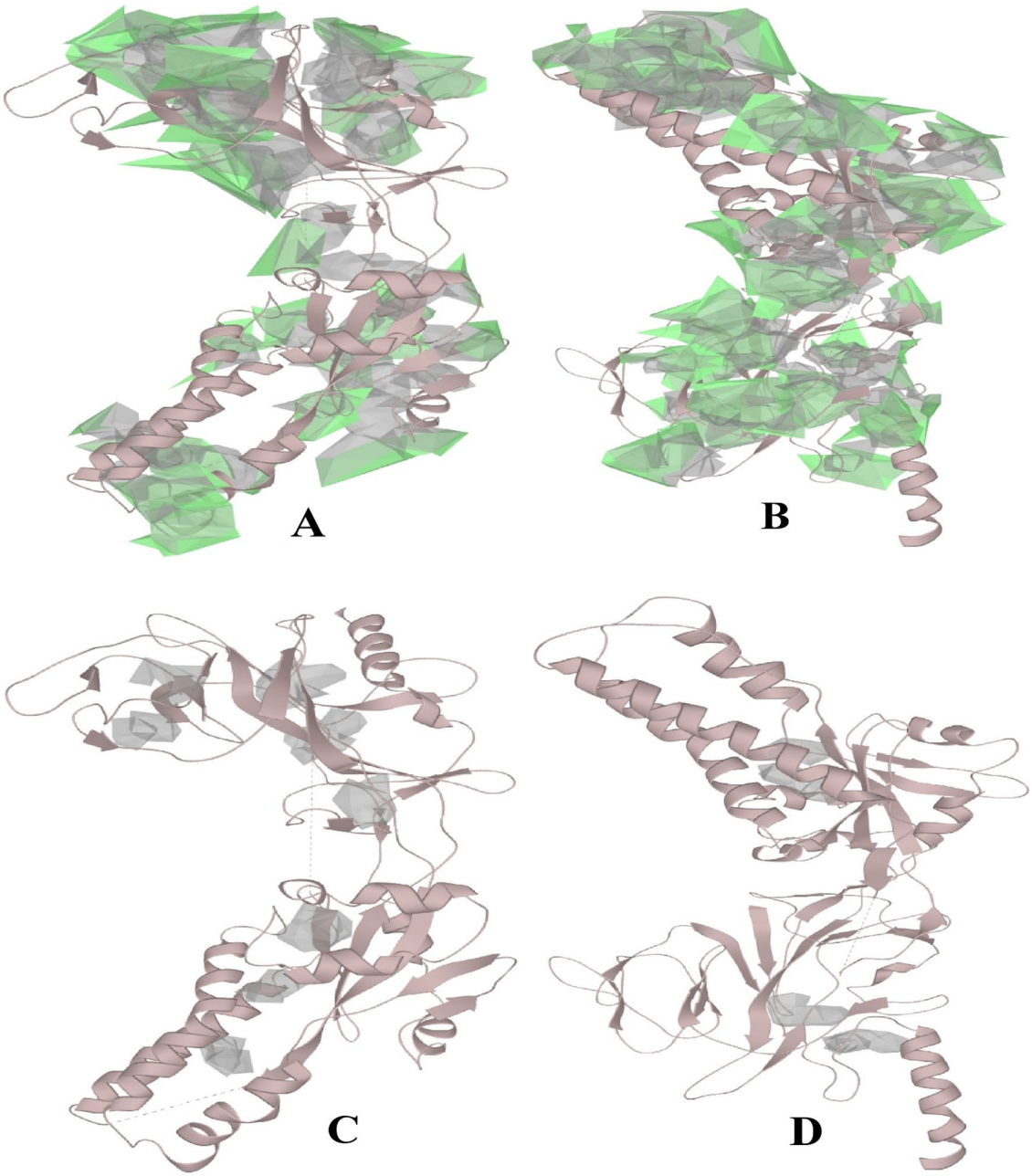


**Fig. 6** **A** Cavities in fusion glycoprotein (5EA4) of RSV A. **B** Cavities in fusion glycoprotein (6Q0S) of RSV B. **C** Voids in fusion glycoprotein (5EA4) of RSV A. **D** Voids in fusion glycoprotein (6Q0S) of RSV B

## Conclusion

The present study is very much helpful to understand the plausible explanation behind the stability and higher transmission possibility of RSV A. Amino acid compositions revealed that RSV A has incorporated more polar residues mainly aromatic residues in its sequences. Improving hydrophilicity helps it to transmit quickly through water droplets and also intensifies its stability by increasing its polarity. High intrinsic disordered regions help in protein folding and increase RSV A functionality. The higher propensity of charged residues in helix gives more stability to RSV A. Intra-protein interactions are giving the edge to elevate the stability of RSV A. Presence of network aromatic–sulfur interaction in viral protein has been revealed earliest in this study. Reduce the number of ligand binding

pockets to protect RSV A from drug–protein interaction. The increasing number of tunnels is increasing its catalytic activity. The study of fusion glycoprotein and major surface glycoprotein of the RSV helps to understand the reason for the dominant nature of RSV A and also helps in drug discoveries.

**Authors' contribution** D.M. conceived and designed the project. P.K.D.M. conducted initial manual verifications. Viral protein sequence and structure were identified by D.M. Analysis of results was done by D.M. Draft of the manuscript was prepared by D.M. Final version of the manuscript was prepared by P.K.D.M. The whole work was done under the supervision of P.K.D.M.

**Funding** None.

## Declarations

**Conflict of interest** None.

## References


1. Mitra D, Paul M, Thatoi H, Mohapatra PK. Study of potentiality of dexamethasone and its derivatives against Covid-19. J Biomol Struct Dyn. 2021. https://doi.org/10.1080/07391102.2021.1942210.
2. Mitra D, Bose A. Remarkable effect of natural compounds that have therapeutic effect to stop COVID-19. In: Recent advances in pharmaceutical sciences. Bhopal: Innovare Academic Sciences Pvt Ltd.; 2020. p. 115–26.
3. Higgins D, Trujillo C, Keech C. Advances in RSV vaccine research and development–A global agenda. Vaccine. 2016;34(26):2870–5. https://doi.org/10.1016/j.vaccine.2016.03.109.
4. Graham BS. Vaccine development for respiratory syncytial virus. Curr Opin Virol. 2017;23:107–12. https://doi.org/10.1016/j.coviro.2017.03.012.
5. Fries L, Shinde V, Stoddard JJ, Thomas DN, Kpamegan E, Lu H, Smith G, Hickman SP, Piedra P, Glenn GM. Immunogenicity and safety of a respiratory syncytial virus fusion protein (RSV F) nanoparticle vaccine in older adults. Immun Ageing. 2017;14(1):1–4. https://doi.org/10.1186/s12979-017-0090-7.
6. Moudy RM, Sullender WM, Wertz GW. Variations in intergenic region sequences of Human respiratory syncytial virus clinical isolates: analysis of effects on transcriptional regulation. Virology. 2004;327(1):121–33. https://doi.org/10.1016/j.virol.2004.06.013.
7. Brüggmann D, Köster C, Klingelhöfer D, Bauer J, Ohlendorf D, Bundschuh M, Groneberg DA. Respiratory syncytial virus: a systematic scientometric analysis of the global publication output and the gender distribution of publishing authors. BMJ Open. 2017;7(7):e013615. https://doi.org/10.1136/bmjopen-2016-013615.
8. Broberg EK, Waris M, Johansen K, Snacken R, Penttinen P. Seasonality and geographical spread of respiratory syncytial virus epidemics in 15 European countries, 2010 to 2016. Eurosurveillance. 2018;23(5):17–00284. https://doi.org/10.2807/1560-7917.ES.2018.23.5.17-00284.
9. Geoghegan S, Erviti A, Caballero MT, Vallone F, Zanone SM, Losada JV, Bianchi A, Acosta PL, Talarico LB, Ferretti A, Grimaldi LA. Mortality due to respiratory syncytial virus. Burden and risk factors. Am J Respir Crit Care Med. 2017;195(1):96–103. https://doi.org/10.1164/rccm.201603-0658OC.
10. Falsey AR, Walsh EE. Respiratory syncytial virus infection in adults. Clin Microbiol Rev. 2000;13(3):371–84. https://doi.org/10.1128/CMR.13.3.371.
11. Coultas JA, Smyth R, Openshaw PJ. Respiratory syncytial virus (RSV): a scourge from infancy to old age. Thorax. 2019;74(10):986–93. https://doi.org/10.1136/thoraxjnl-2018-212212.
12. Mufson MA, Örvell C, Rafnar B, Norrby E. Two distinct subtypes of human respiratory syncytial virus. J Gen Virol. 1985;66(10):2111–24. https://doi.org/10.1099/0022-1317-66-10-2111.
13. Walsh EE, McConnochie KM, Long CE, Hall CB. Severity of respiratory syncytial virus infection is related to virus strain. J Infect Dis. 1997;175(4):814–20. https://doi.org/10.1086/513976.
14. Laham FR, Mansbach JM, Piedra PA, Hasegawa K, Sullivan AF, Espinola JA, Camargo CA Jr. Clinical profiles of respiratory syncytial virus subtypes A and B among children hospitalized with bronchiolitis. Pediatr Infect Dis J. 2017;36(8):808. https://doi.org/10.1097/INF.0000000000001596.
15. Rouka E, Hatzoglou C, Gourgoulianis KI, Zarogiannis SG. Interactome networks between the human respiratory syncytial virus (HRSV), the human metapneumovirus (HMPV), and their host: In silico investigation and comparative functional enrichment analysis. Microb Pathog. 2020;141: 104000. https://doi.org/10.1016/j.micpath.2020.104000.
16. Mei S, Zhang K. In silico unravelling pathogen-host signaling cross-talks via pathogen mimicry and human protein-protein interaction networks. Comput Struct Biotechnol J. 2020;18:100–13. https://doi.org/10.1016/j.csbj.2019.12.008.
17. Tan L, Coenjaerts FE, Houspie L, Viveen MC, van Bleek GM, Wiertz EJ, Martin DP, Lemey P. The comparative genomics of human respiratory syncytial virus subgroups A and B: genetic variability and molecular evolutionary dynamics. J Virol. 2013;87(14):8213–8000026. https://doi.org/10.1128/JVI.03278-12.
18. Souza C, Zanchin NI, Krieger MA, Ludwig A. In silico analysis of amino acid variation in human respiratory syncytial virus: insights into immunodiagnostics. Mem Inst Oswaldo Cruz. 2017;112(10):655–63. https://doi.org/10.1590/0074-02760170013.
19. Yang S, Guo Q, Wu F, Chu Y, Wang Y, Zhou M, Ding CF. Investigation of noncovalent interactions between peptides with potential intrinsic sequence patterns by mass spectrometry. Rapid Commun Mass Spectrom. 2020;34(10): e8736. https://doi.org/10.1002/rcm.8736.
20. Kumar S, Nussinov R. Salt bridge stability in monomeric proteins. J Mol Biol. 1999;293(5):1241–55. https://doi.org/10.1006/jmbi.1999.3218.
21. Makwana KM, Mahalakshmi R. Implications of aromatic–aromatic interactions: from protein structures to peptide models. Protein Sci. 2015;24(12):1920–33. https://doi.org/10.1002/pro.2814.
22. Zauhar RJ, Colbert CL, Morgan RS, Welsh WJ. Evidence for a strong sulfur–aromatic interaction derived from crystallographic data. Biopolymers. 2000;53(3):233–48. https://doi.org/10.1002/(SICI)1097-0282.
23. Zhang H, Li C, Yang F, Su J, Tan J, Zhang X, Wang C. Cation-pi interactions at non-redundant protein-RNA interfaces. Biochem. 2014;79(7):643–52. https://doi.org/10.1134/S0006297914070062.
24. Mitra D, Mohapatra PK. Discovery of novel cyclic salt bridge in thermophilic bacterial protease and study of its sequence and structure. Appl Biochem Biotechnol. 2021;193(6):1688–700. https://doi.org/10.1007/s12010-021-03547-3.

25. Mitra D, Pal AK, Mohapatra PK. Intra-protein interactions of SARS-CoV-2 and SARS: a bioinformatic analysis for plausible explanation regarding stability, divergency, and severity. Syst Microbiol Biomanuf. 2022. https://doi.org/10.1007/s43393-022-00091-x.
26. Sharma R, Sastry GN. Deciphering the dynamics of non-covalent interactions affecting thermal stability of a protein: Molecular dynamics study on point mutant of Thermus thermophilus isopropylmalate dehydrogenase. PLoS ONE. 2015;10(12): e0144294. https://doi.org/10.1371/journal.pone.0144294.
27. Mitra D, Mohapatra PK. Cold adaptation strategy of psychrophilic bacteria: an in-silico analysis of isocitrate dehydrogenase. Syst Microbiol Biomanuf. 2021;1(4):483–93. https://doi.org/10.1007/s43393-021-00041-z.
28. UniProt Consortium. UniProt: a hub for protein information. Nucleic acids Res. 2015;43:D204–12. https://doi.org/10.1093/nar/gku989.
29. Rose PW, Prlić A, Altunkaya A, Bi C, Bradley AR, Christie CH, Costanzo LD, Duarte JM, Dutta S, Feng Z, Green RK. The RCSB protein data bank: integrative view of protein, gene and 3D structural information. Nucleic Acids Res. 2016;45(D1):D271–81. https://doi.org/10.1093/nar/gkw1000.
30. Webb B, Sali A. Comparative protein structure modeling using MODELLER. Curr Protocol Bioinform. 2016;54(1):5–6. https://doi.org/10.1002/cpbi.3.
31. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr. 1993;26(2):283–91. https://doi.org/10.1107/S0021889892009944.
32. Eisenberg D, Lüthy R, Bowie JU. VERIFY3D: assessment of protein models with three-dimensional profiles. Methods Enzymol. 1997;277:396–404. https://doi.org/10.1016/S0076-6879(97)77022-8.
33. Pontius J, Richelle J, Wodak SJ. Deviations from standard atomic volumes as a quality measure for protein crystal structures. J Mol Biol. 1996;264(1):121–36. https://doi.org/10.1006/jmbi.1996.0628.
34. Sievers F, Higgins DG. Clustal omega. Curr Protoc Bioinform. 2014;48(1):3–13. https://doi.org/10.1002/0471250953.bi0313s48.
35. Henikoff S, Henikoff JG, Alford WJ, Pietrokovski S. Automated construction and graphical presentation of protein blocks from unaligned sequences. Gene. 1995;163(2):GC17–26. https://doi.org/10.1016/0378-1119(95)00486-P.
36. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J Mol Biol. 1982;157(1):105–32. https://doi.org/10.1016/0022-2836(82)90515-0.
37. Grantham R. Amino acid difference formula to help explain protein evolution. Science. 1974;185(4154):862–4. https://doi.org/10.1126/science.185.4154.862.
38. Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch A. Protein identification and analysis tools on the ExPASy server. In: Walker JM, editor. The proteomics protocols handbook. Totowa: Humana Press; 2005. p. 571–607.
39. Mitra D, Dey A, Biswas I, Das Mohapatra PK. Bioactive compounds as a potential inhibitor of colorectal cancer; an insilico study of Gallic acid and Pyrogallol. Anna Colorec Res. 2021;9(1):32–9. https://doi.org/10.30476/acrr.2021.89642.1080.
40. Linding R, Jensen LJ, Diella F, Bork P, Gibson TJ, Russell RB. Protein disorder prediction: implications for structural proteomics. Structure. 2003;11:1453–9. https://doi.org/10.1016/j.str.2003.10.002.
41. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera—a visualization system for exploratory research and analysis. J Comput Chem. 2004;25(13):1605–12. https://doi.org/10.1002/jcc.20084.
42. Nedwidek MN, Hecht MH. Minimized protein structures: a little goes a long way. Proc Natl Acad Sci USA. 1997;94(19):10010–1. https://doi.org/10.1073/pnas.94.19.10010.
43. Kumar TA. CFSSP: Chou and Fasman secondary structure prediction server. Wide Spectrum. 2013;1:15–9. https://doi.org/10.5281/zenodo.50733.
44. Tina KG, Bhadra R, Srinivasan N. PIC: protein interactions calculator. Nucleic Acids Res. 2007;35(Suppl_2):W473–6. https://doi.org/10.1093/nar/gkm423.
45. Jubb HC, Higueruelo AP, Ochoa-Montaño B, Pitt WR, Ascher DB, Blundell TL. Arpeggio: a web server for calculating and visualising interatomic interactions in protein structures. J Mol Biol. 2017;429:365–71. https://doi.org/10.1016/j.jmb.2016.12.004.
46. Binkowski TA, Naghibzadeh S, Liang J. CASTp: computed atlas of surface topography of proteins. Nucleic Acids Res. 2003;31(13):3352–5. https://doi.org/10.1093/nar/gkg512.
47. Berka K, Hanák O, Sehnal D, Banáš P, Navratilova V, Jaiswal D, Ionescu CM, Svobodová Vařeková R, Koča J, Otyepka M. MOLE online 2.0: interactive web-based analysis of biomacromolecular channels. Nucleic Acids Res. 2012;40(W1):W222–7. https://doi.org/10.1093/nar/gks363.
48. Pellegrini-Calace M, Maiwald T, Thornton JM. PoreWalker: a novel tool for the identification and characterization of channels in transmembrane proteins from their three-dimensional structure. PLoS Comput Biol. 2009;5(7): e1000440. https://doi.org/10.1371/journal.pcbi.1000440.
49. Hormoz S. Amino acid composition of proteins reduces deleterious impact of mutations. Sci Rep. 2013;3(1):1–10. https://doi.org/10.1038/srep02919.
50. Kreil DP, Ouzounis CA. Identification of thermophilic species by the amino acid compositions deduced from their genomes. Nucleic Acids Res. 2001;29(7):1608–15. https://doi.org/10.1093/nar/29.7.1608.
51. Weiss M, Manneberg M, Juranville JF, Lahm HW, Fountoulakis M. Effect of the hydrolysis method on the determination of the amino acid composition of proteins. J Chromatogr A. 1998;795(2):263–75. https://doi.org/10.1016/S0021-9673(97)00983-7.
52. Chou KC. Prediction of protein cellular attributes using pseudo-amino acid composition. Proteins. 2001;43(3):246–55. https://doi.org/10.1002/prot.1035.
53. Backovic M, Jardetzky TS, Longnecker R. Hydrophobic residues that form putative fusion loops of Epstein-Barr virus glycoprotein B are critical for fusion activity. J Virol. 2007;81:9596–600. https://doi.org/10.1128/JVI.00758-07.
54. Trevino SR, Schaefer S, Scholtz JM, Pace CN. Increasing protein conformational stability by optimizing β-turn sequence. J Mol Biol. 2007;373(1):211–8. https://doi.org/10.1016/j.jmb.2007.07.061.
55. Carneiro FA, Vandenbussche G, Juliano MA, Juliano L, Ruysschaert JM, Da Poian AT. Charged residues are involved in membrane fusion mediated by a hydrophilic peptide located in vesicular stomatitis virus G protein. Mol Membr Biol. 2006;23(5):396–406. https://doi.org/10.1080/09687860600780892.
56. Pokkuluri PR, Raffen R, Dieckman L, Boogaard C, Stevens FJ, Schiffer M. Increasing protein stability by polar surface residues: domain-wide consequences of interactions within a loop. Biophy J. 2002;82(1):391–8. https://doi.org/10.1016/S0006-3495(02)75403-9.
57. Babu MM. The contribution of intrinsically disordered regions to protein function, cellular complexity, and human disease. Biochem Soc Trans. 2016;44:1185–200. https://doi.org/10.1042/BST20160172.

58. Wright PE, Dyson HJ. Intrinsically disordered proteins in cellular signalling and regulation. Nat Rev Mol Cell Biol. 2015;16:18–29. https://doi.org/10.1038/nrm3920.
59. Mishra PM, Verma NC, Rao C, Uversky VN, Nandi CK. Intrinsically disordered proteins of viruses: Involvement in the mechanism of cell regulation and pathogenesis. Prog Mol Biol Transl Sci. 2020;174:1–78. https://doi.org/10.1016/bs.pmbts.2020.03.001.
60. Semenova SI, Ohya H, Soontarapa K. Hydrophilic membranes for pervaporation: an analytical review. Desalination. 1997;110(3):251–86. https://doi.org/10.1016/S0011-9164(97)00103-3.
61. Levy Y, Onuchic JN. Water and proteins: a love–hate relationship. Proc Natl Acad Sci USA. 2004;101(10):3325–6. https://doi.org/10.1073/pnas.0400157101.
62. Panja AS, Maiti S, Bandyopadhyay B. Protein stability governed by its structural plasticity is inferred by physicochemical factors and salt bridges. Sci Rep. 2020;10:1–9. https://doi.org/10.1038/s41598-020-58825-7.
63. Wallner B, Elofsson A. Can correct protein models be identified? Prot Sci. 2003;12(5):1073–86. https://doi.org/10.1110/ps.0236803.
64. Villegas V, Viguera AR, Avilés FX, Serrano L. Stabilization of proteins by rational design of α-helix stability using helix/coil transition theory. Fold Des. 1996;1(1):29–34. https://doi.org/10.1016/S1359-0278(96)00009-0.
65. Xiang Z. Advances in homology protein structure modeling. Curr Protein Pept Sci. 2006;7:217–27. https://doi.org/10.2174/138920306777452312.
66. Muhammed MT, Aki-Yalcin E. Homology modeling in drug discovery: overview, current applications, and future perspectives. Chem Biol Drug Des. 2019;93(1):12–20. https://doi.org/10.1111/cbdd.13388.
67. Neumaier S, Büttner M, Bachmann A, Kiefhaber T. Transition state and ground state properties of the helix–coil transition in peptides deduced from high-pressure studies. Proc Natl Acad Sci USA. 2013;110(52):20988–93. https://doi.org/10.1073/pnas.1317973110.
68. Armstrong KM, Baldwin RL. Charged histidine affects alpha-helix stability at all positions in the helix by interacting with the backbone charges. Proc Natl Acad Sci USA. 1993;90(23):11337–40. https://doi.org/10.1073/pnas.90.23.11337.
69. Padmanabhan S, Marqusee S, Ridgeway T, Laue TM, Baldwin RL. Relative helix-forming tendencies of nonpolar amino acids. Nature. 1990;344(6263):268–70. https://doi.org/10.1038/344268a0.
70. Choudhary A, Gandla D, Krow GR, Raines RT. Nature of amide carbonyl− carbonyl interactions in proteins. J Am Chem Soc. 2009;131(21):7244–6. https://doi.org/10.1021/ja901188y.
71. Rahim A, Saha P, Jha KK, Sukumar N, Sarma BK. Reciprocal carbonyl–carbonyl interactions in small molecules and proteins. Nat Commun. 2017;8(1):1–3. https://doi.org/10.1038/s41467-017-00081-x.
72. Liang J, Woodward C, Edelsbrunner H. Anatomy of protein pockets and cavities: measurement of binding site geometry and implications for ligand design. Prot Sci. 1998;7(9):1884–97. https://doi.org/10.1002/pro.5560070905.
73. Zheng X, Gan L, Wang E, Wang J. Pocket-based drug design: exploring pocket space. AAPS J. 2013;15(1):228–41. https://doi.org/10.1208/s12248-012-9426-6.
74. Kaushik S, Marques SM, Khirsariya P, Paruch K, Libichova L, Brezovsky J, Prokop Z, Chaloupkova R, Damborsky J. Impact of the access tunnel engineering on catalysis is strictly ligand-specific. FEBS J. 2018;285:1456–76. https://doi.org/10.1111/febs.14418.
75. Guo R, Cang Z, Yao J, Kim M, Deans E, Wei G, Kang SG, Hong H. Structural cavities are critical to balancing stability and activity of a membrane-integral enzyme. Proc Natl Acad Sci USA. 2020;117(36):22146–221456. https://doi.org/10.1073/pnas.1917770117.
76. Saito M, Kono H, Morii H, Uedaira H, Tahirov TH, Ogata K, Sarai A. Cavity-filling mutations enhance protein stability by lowering the free energy of native state. J Phys Chem B. 2000;104(15):3705–11. https://doi.org/10.1021/jp991717f.