




Intra-protein interactions of SARS-CoV-2 and SARS: a bioinformatic analysis for plausible explanation regarding stability, divergency, and severity

Debanjan Mitra¹ · Aditya K. Pal¹ · Pradeep Kr. Das Mohapatra¹ 

Received: 13 January 2022 / Revised: 27 February 2022 / Accepted: 28 February 2022 / Published online: 21 March 2022
© Jiangnan University 2022

Abstract

The current nightmare for the whole world is COVID-19. The occurrence of concentrated pneumonia cases in Wuhan city, Hubei province of China, was first reported on December 30, 2019. SARS-CoV first disclosed in 2002 but had not outspread worldwide. After 18 years, in 2020, it reemerged and outspread worldwide as SARS-CoV-2 (COVID-19), as the most dangerous virus-creating disease in the world. Is it possible to create a favorable evolution within the short time (18 years)? If possible, then what are those properties or factors that are changed in SARS-CoV-2 to make it undefeated? What are the fundamental differences between SARS-CoV-2 and SARS? The study is one of the initiatives to find out all those queries. Here, four types of protein sequences from SARS-CoV-2 and SARS were retrieved from the database to study their physicochemical and structural properties. Results showed that charged residues are playing a pivotal role in SARS-CoV-2 evolution and contribute to the helix stabilization. The formation of the cyclic salt bridge and other intra-protein interactions specially network aromatic–aromatic interaction also play the crucial role in SAS-CoV-2. This comparative study will help to understand the evolution from SARS to SARS-CoV-2 and helpful in protein engineering.

Keywords COVID-19 · Evolution · Cyclic salt bridge · Intra-protein interactions · Protein engineering

Introduction

Disease caused by SARS-CoV-2 has been recognized as Corona Virus Disease 2019 (COVID-19). SARS-CoV first came out in the Guangdong province of China in 2002 and had outspread into five countries infecting 8098 people and 774 deaths, having a mortality rate of 11% [1]. After that, in 2012, MERS-CoV appeared in the Arabian Peninsula and had outspread into 27 countries, infecting a total of 2494 individuals and took 858 lives with a mortality rate of 34% [2]. Recently SARS-CoV-2 has been elevated in Wuhan city, Hubei province of China, in December 2019. Till now (11.01.2022), there are over thirty core cases of COVID-19 and over 5.4 million deaths (mortality rate around 3.40%) have been reported to affect 222 countries globally. Currently, a new variant of COVID-19 named “Omicron” are

also reported in many countries with high transmission rate. On March 11, 2020, the World Health Organization announced the COVID-19 pandemic, a public health emergency of global concern. All age’s people can catch this viral infection, but immune-compromised people having co-morbidities are most vulnerable. Propensity of age, males with chronic diseases (like- diabetes, heart disease, cancer, etc.) are higher vulnerable than other groups of people [3]. This virus can be easily transmitted through the droplets generated when coughing and sneezing by the infected people [4]. These infectious droplets can be spread up to 1–2 m and stay on surfaces. This virus can survive on metal surfaces for several hours, even days, in favorable conditions but can be destroyed by disinfectants like hydrogen peroxide, sodium hypochlorite, etc. [5]. The incubation period varies from 2 to 14 days. Few common clinical symptoms are fever (except asymptomatic cases), dry cough, sore throat, fatigue, headache, breathlessness, sudden loss of smell and taste. Without proper treatment, this disease can cause pneumonia, respiratory failure and even death. Generally, after --week recovery started. It has been observed in patients that the progression of this disease increases the release of cytokines including

✉ Pradeep Kr. Das Mohapatra
pkdmvu@gmail.com

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

interleukin (IL)-6 and IL-10, whereas the levels of CD4⁺T and CD8⁺T are reduced [6]. There is no approved treatment for COVID-19 but anti-viral drugs such as Remdesivir, Tocilizumab are in use for treatment [7]. Also, many chemical compounds and bioactive compounds appear by molecular docking studies as a drug in treatment of COVID-19 [8, 9].

Coronavirus is an enveloped virus having a positive single-strand RNA genome, and they have spike proteins on the surface with a size of 60–140 nm [10]. There are four subtypes such as alpha, beta, gamma, and delta coronaviruses. Most of the highly pathogenic viruses are severe acute respiratory syndrome coronavirus (SARS-CoV), Middle East respiratory syndrome coronavirus (MERS-CoV), and SARS-CoV-2; and belongs to β -coronavirus [11]. Generally, the β -coronavirus genome contains six open reading frames (ORFs); first ORFs (ORF1a/b) are in two-thirds of the whole genome and encode 16 nonstructural proteins (nsps). There is one frameshift between ORF1a and ORF1b, which produces two polypeptides, pp1a and pp1ab. Main protease (M^{pro}) and chymotrypsin-like protease (3CL^{pro}) are involved in the processing of these polypeptides [12, 13]. Other ORFs of the genome near the 3'-terminus encode the four main structural proteins, spike glycoproteins, membrane, envelope, and nucleocapsid proteins [14]. Genome analysis of SARS-CoV-2 revealed that there are 79.5% and 97% of similarities with the whole genome sequences of SARS-CoV and bat SARS-CoV, respectively [3]. SARS-CoV-2 enters the host respiratory mucosa by binding with the receptor of angiotensin-converting enzyme 2 (ACE2) with its spike glycoproteins [15]. A recent study has shown that SARS-CoV-2 binds with ACE2 with a tenfold higher affinity compared to SARS-CoV [16]. The basic reproduction number (R_0), the average number of secondary infections produced by patients, is between 2.47 and 2.86 for SARS-CoV-2, whereas the R_0 value of SARS-CoV is 2.2–3.6, and 2.0–6.7 for MERS-CoV [17–19]. These results indicate that SARS-CoV-2 has comparatively high transmission ability than other coronaviruses. Sequence analysis of SARS-CoV 2, SARS-CoV, and other SARS-related coronavirus (SARSr-CoV) spike glycoproteins showed that four amino acids are inserted in the positions of 681–684 between S1 and S2 subunit of SARS-CoV-2 [20]. SARS-CoV ORF 3b, ORF 6, and N proteins inhibit the expression of beta interferon (IFN- β) [21]. The envelope (E) protein in coronavirus is a small membrane protein that has several functions in virion assembly and ion-channel activity, through which it can interact with the host [22].

With the unavailability of anti-viral drugs for nCoV, society demands sincere efforts in drug design and discovery for COVID-19 [23, 24]. Since 2002, SARS has present on this earth. But it creates a dangerous effect and makes a pandemic situation after 18 years. Why? Why is this virus so harmful to us? What are the fundamental differences

between SARS-CoV-2 and SARS? How evolution makes them stronger than SARS? How can they gain stability in such extreme environments? Do intra-protein interactions play a vital role in SARS-CoV-2? This study will help to find out all those questions.

Materials and methods

Dataset

A detailed investigation of the sequences and structures of SARS-CoV-2 was performed with reference to the old SARS. Four types of SARS-CoV-2 and SARS reviewed protein sequences, i.e., spike proteins, membrane proteins, nucleoproteins, and ORF proteins (ORF 3, ORF 6, ORF 7, ORF 8, and ORF 9) were considered in this study. All annotated protein sequences of SARS-CoV-2 and SARS were retrieved from the UNIPROT [25] database. The crystal structures of SARS-CoV-2 and SARS proteins were retrieved from the RCSB protein database (PDB) [26]. The structure was chosen based on some criteria of crystal structures.

Physicochemical properties

The protein sequences were subjected to multiple sequence alignment (MSA) with the help of CLUSTAL Omega [27]. Both block and non-block FASTA [28] formats of the sequences were analyzed. Block of the sequence was prepared by BLOCK Maker [29] from MSA. Both non-block and block formats were analyzed by ProtParam server [30–32] and ProtScale server [33] for calculation of physicochemical properties likes amino acid composition, GRAVY, aliphatic index, bulkiness, polarity, etc. The value of ORF protein analysis is the average of all ORF (ORF 3, ORF 6, ORF 7, ORF 8, and ORF 9). The total amount of disorder-forming residues (i.e., E, P, K, S) and order forming residues (i.e., I, F, W, Y) are calculated from amino acid compositions based on previous reports [34, 35]. Intrinsic disorder regions of protein were analyzed by DisEMBL [36] server.

Analysis of crystal structure

SARS-CoV-2 protease (5R80) and SARS protease (2H2Z) were extracting from RCSB PDB for structural comparison. All structured were minimized in 1000 steps using UCSF Chimera with forcefield [37]. Analyses of the secondary structure were done by CFSSP [38] server to find the amino acid abundance in coil, helix, sheet and turn. Number of salt bridges were extracted by WHAT IF server [39]. Intra-protein interactions were determined by Protein interaction calculator [40] and Arpeggio [41]. Free solvation energy

was calculated by ProWaVE server [42]. Surface area and volume were determined by the CASTp [43]. Phosphorylation sites of protein were identified by the NetPhos server [44]. Protein mutations were analyzed by the DUET [45].

Results and discussions

Effect of polar residues on SARS-CoV-2 sequence

Here D, E, H, R, K amino acids were considered as a charged residues and C, S, T, N, Q, Y, W as uncharged polar

residues. Amino acid compositions were calculated from the non-block format, whereas block format was used to calculate disorder-forming residues, order forming residues, bulkiness, aliphatic index (AI), and polarity. GRAVY (grand average of hydropathy) is calculated by adding the hydropathy value [46] for each residue and dividing by the length of the protein sequence. Is there a preference for amino acids in SARS-CoV-2 relative to SARS? To find out the answer, all physicochemical properties were calculated.

Spike proteins showed higher abundance (Fig. 1) of charged residues (except D) in SARS-CoV-2. Polar residues in spike proteins showed higher quantity (except T,

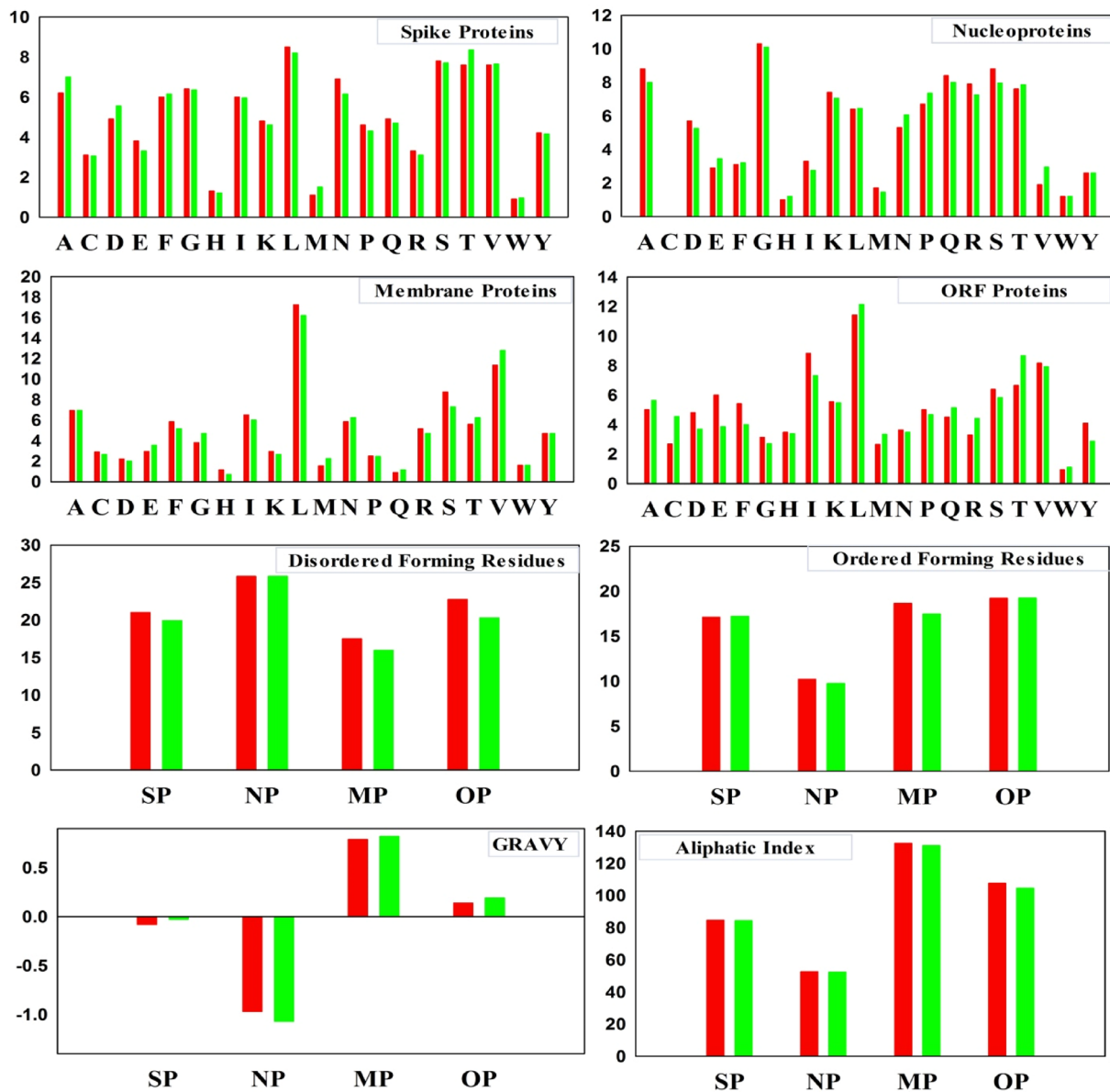


Fig. 1 Comparative analysis of physicochemical properties like amino acid compositions, disorder-forming residues, order forming residues, GRAVY, aliphatic index of spike proteins (SP), nucleopro-

teins (NP), membrane proteins (MP), ORF proteins (OP) from SARS-CoV-2 (red bar) and SARS (green bar)

W) in SARS-CoV-2. In nucleoproteins of SARS-CoV-2 D, K and R showed higher abundance and E, H showed lower abundance as charged amino acids. Polar residues in nucleoproteins also showed higher plenty (except T, N) in SARS-CoV-2. Surprisingly C is absent in both groups of sequence in nucleoproteins. Other proteins, i.e., membrane proteins and ORF proteins, showed almost similar abundance with those previous results. Polar residues also help proteins to tolerate temperature [47]. Number of disorder-forming residues has higher abundance in SARS-CoV-2 than SARS. The number of order forming residues has lower abundance in SARS-CoV-2 than SARS (in case of spike and ORF proteins). The higher number of disorder-forming residues in SARS-CoV-2 indicates that it can easily increase pathogenicity or virulence. Proline may give a preadaptive advantage by enhancing antioxidant defenses, which in the setting of disease would extend cell viability, raise colonization efficiencies, and enhance virulence [48]. It was also reported that disorder-forming residues like S and E, are responsible for increase pathogenicity [49, 50]. The aliphatic index is high in every SARS-CoV-2 protein. Increased value of the aliphatic index in SARS-CoV-2 proved that SARS-CoV-2 is more thermally stable than SARS [51].

The polarity of those proteins showed slightly higher values in SARS-CoV-2 than SARS (Fig. 2). Due to the latter, bulkiness is also high in SARS-CoV-2 than SARS. The high value of bulkiness in SARS-CoV-2 indicates that they need more extended heating periods in hydrolysis [52]. They can tolerate heat better than SARS. The Kyte–Doolittle hydrophobicity scale suggests that the SARS-CoV-2 is hydrophilic in nature (Fig. 3). The lower value of GRAVY (except nucleoproteins) indicates the hydrophilic nature of SARS-CoV-2. The hydrophilic nature of SARS-CoV-2 gives a clue that it can quickly interact with water or aqueous medium and spread easily than SARS [53, 54]. The intrinsic disorder regions are very much high in SARS-CoV-2 than SARS. A high abundance of intrinsic disorder regions of SARS-CoV-2 indicates that it helps in protein folding of SARS-CoV-2 and will interact more with other proteins than SARS. Many intrinsically disordered proteins (IDPs) have been found to undergo a disorder-to-order transition, implying that their folding processes are inherently distinct from those seen in globular proteins. After binding to natural partners, certain IDPs can fold into a unique 3D form. Many IDPs/IDPRs can fold when they engage with their binding partners and have various binding specificities, allowing them to participate in one-to-many and many-to-one interactions [34, 55–60]. At various levels, viral IDPs mediate successful infection and govern pathogenesis. Because of their widespread engagement in host–pathogen mediated regulators and great prevalence in viral proteomes, virus IDPs are being investigated as possible therapeutic targets [61].

Analysis of secondary structure of SARS-CoV-2 and SARS

The building blocks of proteins, i.e., amino acids, are found in four positions of secondary structure, i.e., coil, helix, sheet, and turn.

Charged residues showed higher abundance in every position (turn, helix, coil, and sheet) of SARS-CoV-2 (Table 1) than SARS. Charged residues showed higher abundance within the helix of both proteins. The introduction of higher number of charged residues in the helix, resulted in more resistant proteins to the acidic environment or temperature denaturation which helps in increasing the stability [62, 63]. Hydrophobic residues have higher abundance in SARS (except coil) than SARS-CoV-2. Polar residues also showed higher abundance in every secondary structure position of SARS-CoV-2 than SARS. It was already proved that polar amino acids on the surface can influence helix formation and increase its stability [64]. However, the highest abundance of polar residues was found in sheet of both SARS-CoV-2 and SARS. More than 50% of residues were present in sheet of SARS, whereas SARS-CoV-2 have 39.33% and 31.54% residues on sheet and helix. So, SARS-CoV-2 increase amino acids propensity in helix to increase its stability.

Effect of intra-protein interactions on SARS-CoV-2 and SARS

Salt bridges have a significant effect on protein stability [65–68] Charged residues are participating in the formation of salt bridges. Usually, two types of salt bridges are found in proteins, i.e., isolated salt bridge and network salt bridge. The increasing number of charged residues of SARS-CoV-2 indicates that charged residues might enhance salt bridge formation to gain more stability. Other intra-protein interactions like, metal ion binding site [69], aromatic-aromatic interactions [70–72] also help in protein stabilization.

SARS-CoV-2 has large pocket area than SARS (Fig. 4A, B), which gives it more protein–protein or protein–ligand interactions possibilities (Table 2). The volume of the protein is also high in SARS-CoV-2 than in SARS. Protease from SARS-CoV-2 possess 9 isolated salt bridges and 1 network salt bridge, whereas SARS protease has 8 isolated and 1 network salt bridge. The result indicated that SARS-CoV-2 is highly stabilized by the help of salt bridges. Though SARS-CoV-2 and SARS proteins have only one type of network salt bridge, but SARS-CoV-2 has gained a special engineered salt bridge (Fig. 4E, F), which is cyclic in nature (R131-E290, K137-E290, R131-D197, K137-D197, R131-D289). Residue number 131R participated maximum time to form this cyclic salt bridge. Novel cyclic salt bridge might have a great role in its protein stability [68, 72].

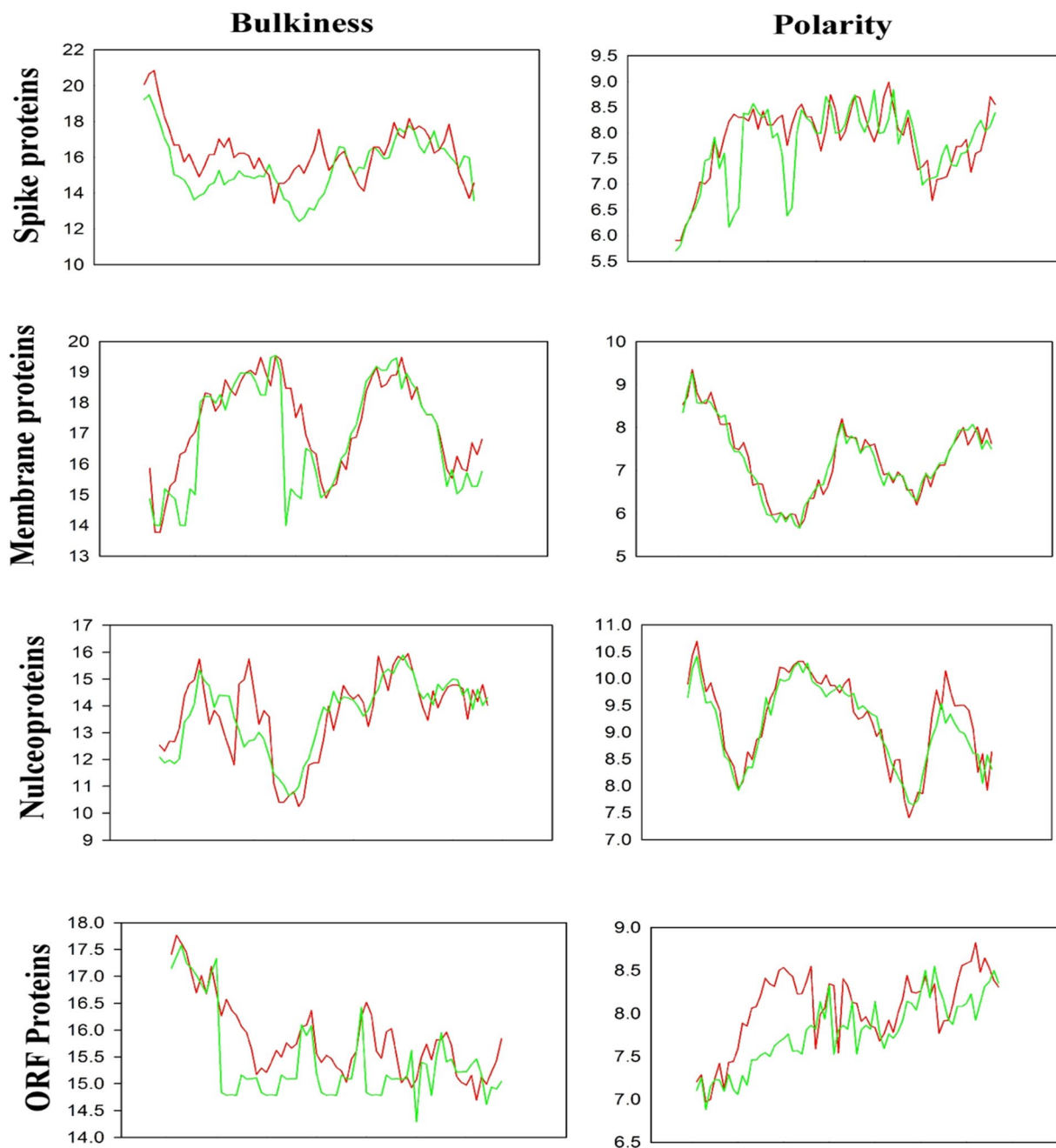


Fig. 2 Comparative analysis of bulkiness and polarity of spike proteins (SP), nucleoproteins (NP), membrane proteins (MP), ORF proteins (OP) from SARS-CoV-2 (red line), and SARS (green line)

Number of metal ion binding sites is also high in SARS-CoV-2 than SARS. These 3 metal ion binding sites contain dimethyl sulfoxide in COVID-19 virus. Free solvation energy is a thermodynamic factor that determines protein solvation or the nature of denaturation [73]. By this property, the rate of proteins denaturation can be determined. Solvation free energy is also high in SARS-CoV-2 than SARS which indicates that, the SARS-CoV-2 protein cannot be easily denatured in contact with the solvent.

Aromatic-aromatic interactions showed high number in SARS-CoV-2 than SARS (Table 3). Not only number, but some of the residues are participated in aromatic-aromatic interactions are forming a very long network, which has never been reported in any viral proteins. SARS-CoV-2 has 3 isolated and 2 network aromatic-aromatic interactions whereas SARS has only 9 isolated aromatic-aromatic interactions.

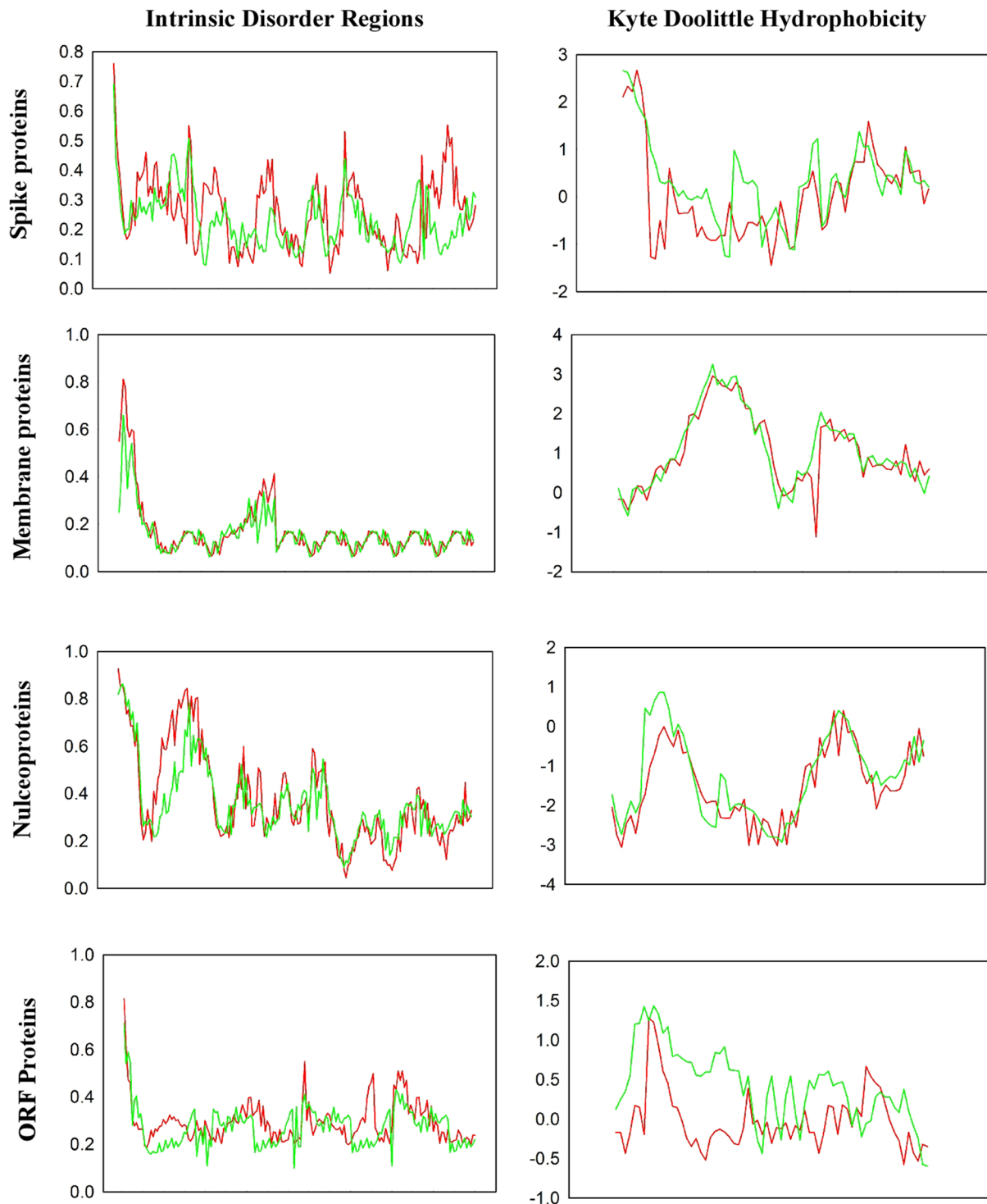


Fig. 3 Comparative study of intrinsic disorder regions and Kyte–Doolittle hydrophobic scale of spike proteins (SP), nucleoproteins (NP), membrane proteins (MP), ORF proteins (OP) from SARS-CoV-2 (red line) and SARS (green line)

The number of phosphorylation sites (Fig. 4C, D) in SARS-CoV-2 is 54, whereas the number of phosphorylation sites in SARS is 45. That means SARS-CoV-2 has

higher number of phosphorylation sites than SARS. The high numbers of phosphorylation sites in SARS-CoV-2

Table 1 Amino acid abundance (%) in protein secondary structures (turn, helix, coil, and sheet) of SARS-CoV-2 (5R80) and SARS (2H2Z)

	SARS-CoV-2 (5R80)			SARS (2H2Z)		
	Charged	Polar	Hydrophobic	Charged	Polar	Hydrophobic
Turn	1.63	2.45	2.86	0.90	2.28	3.59
Helix	11.47	5.32	14.75	8.16	3.92	14.70
Coil	2.04	8.60	11.47	1.63	5.88	8.16
Sheet	9.01	19.67	10.65	6.53	16.01	28.10

increase the strength of protein–protein interactions and helps in stability [74].

Favorable point mutations of SARS-CoV-2

Result of MSA of both structures showed some point mutations in SARS-CoV-2. So, their effect on protein stability has been analyzed. Total 11 mutations have been identified, among which 8 are favorable and 3 are unfavorable in SARS-CoV-2 (Table 4). Residue number 35, which was threonine of SARS substituted by valine in SARS-CoV-2 after mutation, contributes highest energy, i.e., -2.24 kcal/mol. Residue S63N mutation contributed 2nd highest energy to SARS-CoV-2, i.e., -1.16 kcal/mol. However, mutation on A46S, K180N and I286L showed destabilization in SARS-CoV-2 protein stability. 6 polar and 5 non-polar amino acids of SARS were mutated to 5 polar and 6 non-polar amino acids in SARS-CoV-2. The point mutations predicted in SARS-CoV-2 contributed about the total energy level of -7.46 kcal/mol, which is the main driving force in more stability as compared to SARS.

Conclusion

The acidic and basic residues are playing the significant role in evolution. The presence of charged residues in the helix region contributed increasing in protein stability. Increasing hydrophilicity helps SARS-CoV-2 to spread easily through air droplets. Disorder forming residues increase SARS-CoV-2 pathogenicity. High bulkiness of SARS-CoV-2 make them heat tolerate. The long network aromatic-aromatic interactions are the added advantage in protein stability. It is the first report of the presence of cyclic salt bridge and long network aromatic-aromatic interaction in viral protein. Increasing of metal ion binding sites and phosphorylation sites are also playing a crucial role in SARS-CoV-2 protein stability. The point mutations showed, how SARS-CoV-2 engendered itself to gain more stability. It is also a clue to stop SARS-CoV-2 infection severity by deleting those favorable mutant amino acid residues. Protein engineering helps us in this process. The findings of the present investigation contributed many more things, which are essential in drug and vaccine development against SARS-CoV-2.

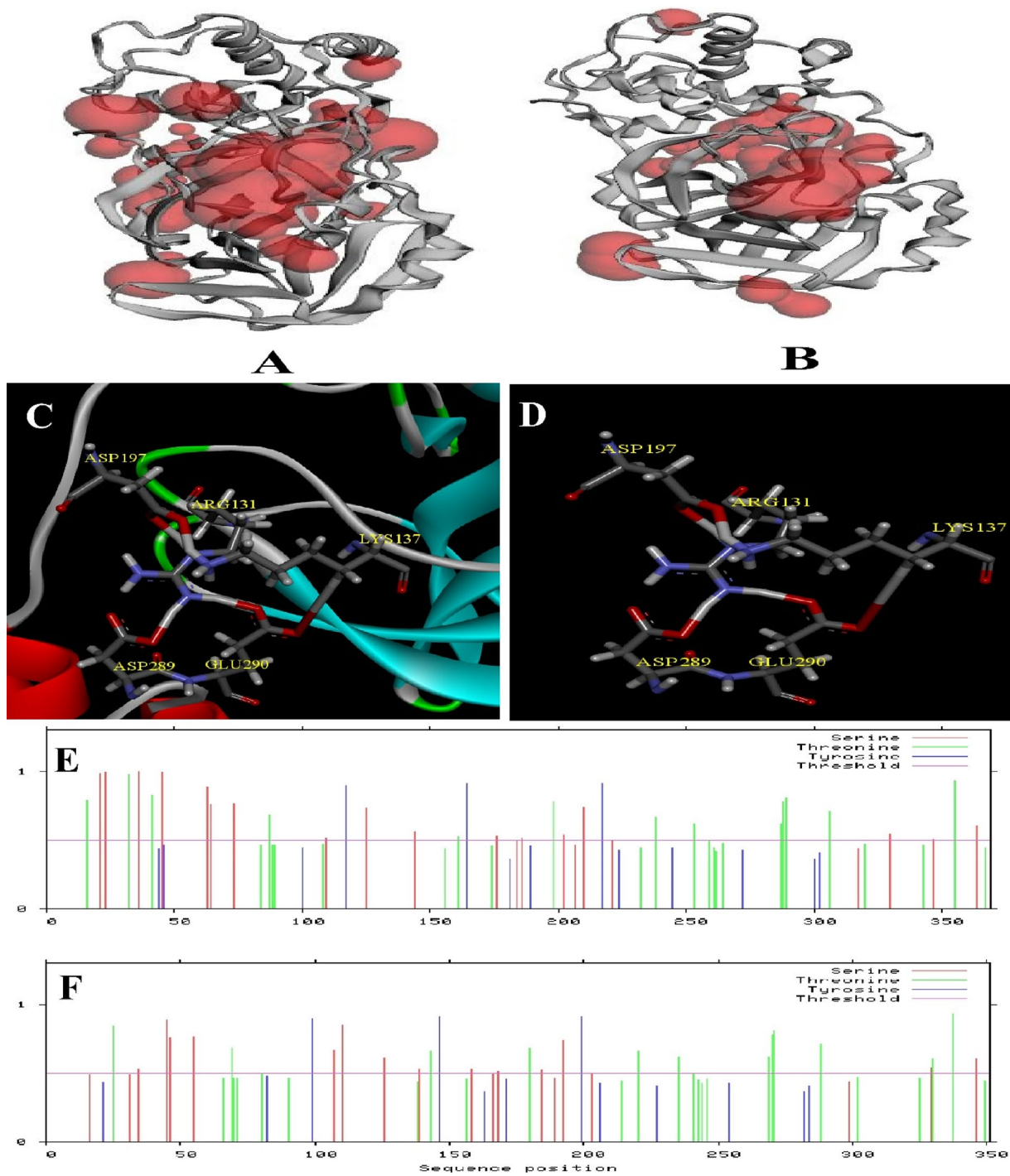


Fig. 4 Large pocket area (red zone) in 3D protein structures of SARS-CoV-2 (**A**) than SARS (**B**). 3D protein structure of SARS-CoV-2 with cyclic salt bridge (**C**). 3D view of only cyclic salt bridge

formation of SARS-CoV-2 (**D**). Phosphorylation sites of SARS-CoV-2 (**E**) and SARS (**F**)

Table 2 Volume, pocket area, isolated salt bridges (ISB), network salt bridges (NSB), metal binding site (MBS), and solvation free energy (ΔG_{solv}) of SARS-CoV-2 (5R80) and SARS (2H2Z)

Protein	Volume	Area	ISB	NSB	MBS	Solvation free energy (ΔG_{solv}) (kcal/mol)
5R80	669.47 Å ³	1013.45 Å ²	9	1	3	4786.55
2H2Z	228.27 Å ³	835.26 Å ²	8	1	2	3266.89

Table 3 Isolated and network (bold and italic) aromatic–aromatic interactions of SARS-CoV-2 (5R80) and SARS (2H2Z)

Protein	Position	Residue	Position	Residue	D(centroid–centroid)	Dihedral angle
SARS-CoV-2 (5R80)	8	PHE	150	PHE	6.77	45.38
	37	TYR	103	PHE	5.27	88.61
	101	TYR	103	PHE	5.7	133.64
	101	TYR	159	PHE	6.78	118.46
	103	PHE	159	PHE	6.44	74.2
	150	PHE	159	PHE	6.43	56.09
	<i>112</i>	<i>PHE</i>	<i>161</i>	<i>TYR</i>	<i>5.41</i>	<i>142.21</i>
	<i>134</i>	<i>PHE</i>	<i>182</i>	<i>TYR</i>	<i>6.33</i>	<i>164.01</i>
	<i>161</i>	<i>TYR</i>	<i>182</i>	<i>TYR</i>	<i>6.47</i>	<i>150.59</i>
	3	PHE	291	PHE	4.96	130.9
126	TYR	140	PHE	6.38	64.71	
218	TRP	219	PHE	5.96	104.81	
SARS (2H2Z)	3	PHE	300	CYS	4.72	116
	54	TYR	44	CYS	4.29	52.78
	66	PHE	22	CYS	4.61	29.56
	112	PHE	160	CYS	4.18	149.49
	126	TYR	128	CYS	4.58	86.27
	181	PHE	85	CYS	4.87	85.18
	182	TYR	130	MET	4.84	149.68
	209	TYR	264	MET	4.98	8.08
230	PHE	265	CYS	4.58	166.08	

Table 4 Effect of amino acid mutations in SARS-CoV-2 with their contributing energies

Residue in SARS-CoV-2 (5R80)	Residue number	Residue in SARS (2H2Z)	Contributing energy in stability of SARS-CoV-2 (Kcal/mol)	Result of mutations in stability of SARS-CoV-2
V	35	T	– 2.24	Stabilized
S	46	A	0.08	Destabilized
N	63	S	– 1.17	Stabilized
V	86	L	– 1.08	Stabilized
K	88	R	– 0.59	Stabilized
A	94	S	– 1.14	Stabilized
F	134	H	– 0.98	Stabilized
N	180	K	0.08	Destabilized
V	202	L	– 0.33	Stabilized
S	267	A	– 0.13	Stabilized
L	286	I	0.04	Destabilized

Author contributions DM and AP conceived and designed the project. PKDM conducted initial manual verifications. Protein sequence and structure were identified by DM. Analysis of results was done by DM. Draft of the manuscript was prepared by DM and AP. Final version of the manuscript was prepared by PKDM. The whole work was done under the supervision of PKDM.

Funding None.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

- Park SE. Epidemiology, virology, and clinical features of severe acute respiratory syndrome-coronavirus-2 (SARS-CoV-2; Coronavirus Disease-19). *Clin Exp Pediatr*. 2020;63(4):119–24. <https://doi.org/10.3345/cep.2020.00493>.
- Singhal T. A review of coronavirus disease-2019 (COVID-19). *Indian J Pediatr*. 2020;87(4):281–6. <https://doi.org/10.1007/s12098-020-03263-6>.
- Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, Qiu Y, Wang J, Liu Y, Wei Y, Yu T. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet*. 2020;395(10223):507–13. [https://doi.org/10.1016/S0140-6736\(20\)30211-7](https://doi.org/10.1016/S0140-6736(20)30211-7).
- Bastola A, Sah R, Rodriguez-Morales AJ, Lal BK, Jha R, Ojha HC, Shrestha B, Chu DK, Poon LL, Costello A, Morita K. The first 2019 novel coronavirus case in Nepal. *Lancet Infect Dis*. 2020;20(3):279–80. [https://doi.org/10.1016/S1473-3099\(20\)30067-0](https://doi.org/10.1016/S1473-3099(20)30067-0).
- Kampf G, Todt D, Pfaender S, Steinmann E. Persistence of coronaviruses on inanimate surfaces and their inactivation with biocidal agents. *J Hos Infect*. 2020;104(3):246–51. <https://doi.org/10.1016/j.jhin.2020.01.022>.
- Wan S, Yi Q, Fan S, Lv J, Zhang X, Guo L, Lang C, Xiao Q, Xiao K, Yi Z, Qiang M. Characteristics of lymphocyte subsets and cytokines in peripheral blood of 123 hospitalized patients with 2019 novel coronavirus pneumonia (NCP). *MedRxiv*. 2020. <https://doi.org/10.1101/2020.02.10.20021832>.
- Hung IF, Lung KC, Tso EY, Liu R, Chung TW, Chu MY, Ng YY, Lo J, Chan J, Tam AR, Shum HP. Triple combination of interferon beta-1b, lopinavir–ritonavir, and ribavirin in the treatment of patients admitted to hospital with COVID-19: an open-label, randomised, phase 2 trial. *Lancet*. 2020;395(10238):1695–704. [https://doi.org/10.1016/S0140-6736\(20\)31042-4](https://doi.org/10.1016/S0140-6736(20)31042-4).
- Mitra D, Paul M, Thatoi H, Mohapatra PK. Study of potentiality of dexamethasone and its derivatives against Covid-19. *J Biomol Struct Dyn*. 2021;16:1–11. <https://doi.org/10.1080/07391102.2021.1942210>.
- Mitra D, Bose A. Remarkable effect of natural compounds that have therapeutic effect to stop COVID-19, in *Recent Advances in Pharmaceutical Sciences, Innovare academic sciences 2021*; 115–126.
- Cascella M, Rajnik M, Aleem A, Dulebohn S, Di Napoli R. Features, evaluation and treatment coronavirus (COVID-19), *StatPearls Publishing*, 2020.
- Chan JF, To KK, Tse H, Jin DY, Yuen KY. Interspecies transmission and emergence of novel viruses: lessons from bats and birds. *Trends Microbiol*. 2013;21(10):544–55. <https://doi.org/10.1016/j.tim.2013.05.005>.
- Ziebuhr J, Snijder EJ, Gorbalenya AE. Virus-encoded proteinases and proteolytic processing in the Nidovirales. *J Gen Virol*. 2000;81(4):853–79. <https://doi.org/10.1099/0022-1317-81-4-853>.
- Masters PS. The molecular biology of coronaviruses. *Adv Virus Res*. 2006;66:193–292. [https://doi.org/10.1016/S0065-3527\(06\)66005-3](https://doi.org/10.1016/S0065-3527(06)66005-3).
- Yoshimoto FK. The proteins of severe acute respiratory syndrome coronavirus-2 (SARS CoV-2 or n-COV19), the cause of COVID-19. *The Prot J*. 2020;39:198–216. <https://doi.org/10.1007/s10930-020-09901-4>.
- Hoffmann M, Kleine-Weber H, Krüger N, Müller M, Drosten C, Pöhlmann S. The novel coronavirus 2019 (2019-nCoV) uses the SARS-coronavirus receptor ACE2 and the cellular protease TMPRSS2 for entry into target cells. *BioRxiv*. 2020. <https://doi.org/10.1101/2020.01.31.929042>.
- Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh CL, Abiona O, Graham BS, McLellan JS. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*. 2020;367(6483):1260–3. <https://doi.org/10.1126/science.abb2507>.
- Remais J. Modelling environmentally-mediated infectious diseases of humans, transmission dynamics of schistosomiasis in China. In: *Modelling parasite transmission and control*. Springer, New York 2010; 79–98. https://doi.org/10.1007/978-1-4419-6064-1_6
- Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *Lancet*. 2020;395(10225):689–97. [https://doi.org/10.1016/S0140-6736\(20\)30260-9](https://doi.org/10.1016/S0140-6736(20)30260-9).
- Lipsitch M, Cohen T, Cooper B, Robins JM, Ma S, James L, Gopalakrishna G, Chew SK, Tan CC, Samore MH, Fisman D. Transmission dynamics and control of severe acute respiratory syndrome. *Science*. 2003;300(5627):1966–70. <https://doi.org/10.1126/science.1086616>.
- Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell*. 2020;181(2):281–92. <https://doi.org/10.1016/j.cell.2020.02.058>.
- Kopecky-Bromberg SA, Martínez-Sobrido L, Frieman M, Baric RA, Palese P. Severe acute respiratory syndrome coronavirus open reading frame (ORF) 3b, ORF 6, and nucleocapsid proteins function as interferon antagonists. *J Virol*. 2007;81(2):548–57. <https://doi.org/10.1128/JVI.01782-06>.
- Ruch TR, Machamer CE. The coronavirus E protein: assembly and beyond. *Viruses*. 2012;4(3):363–82. <https://doi.org/10.3390/v4030363>.
- Mitra D, Das Mohapatra PK. Inhibition of SARS-CoV-2 protein by bioactive compounds of edible mushroom; a bioinformatics insight. *Int J Adv Sci Eng Technol*. 2021;9(2):84–8.
- Acharya A, Agarwal R, Baker MB, Baudry J, Bhowmik D, Boehm S, Byler KG, Chen SY, Coates L, Cooper CJ, Demerdash O. Supercomputer-based ensemble docking drug discovery pipeline with application to COVID-19. *J Chem Inf Model*. 2020;60(12):5832–52. <https://doi.org/10.1021/acs.jcim.0c01010>.
- Consortium U. UniProt, a hub for protein information. *Nucleic Acids Res*. 2015;43(D1):D204–12. <https://doi.org/10.1093/nar/gku989>.
- Rose PW, Bi C, Bluhm WF, Christie CH, Dimitropoulos D, Dutta S, Green RK, Goodsell DS, Plić A, Quesada M, Quinn GB. The RCSB Protein Data Bank: new resources for research and education. *Nucleic Acids Res*. 2012;41(D1):D475–82. <https://doi.org/10.1093/nar/gks1200>.
- Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson

- JD. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol.* 2011;7(1):539. <https://doi.org/10.1038/msb.2011.75>.
28. Pearson WR. Using the FASTA program to search protein and DNA sequence databases, in *Computer Analysis of Sequence Data*, Humana Press, 1994; pp. 307–331. <https://doi.org/10.1385/0-89603-246-9:307>
 29. 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](https://doi.org/10.1016/0378-1119(95)00486-P).
 30. Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch A. Protein identification and analysis tools on the ExPASy server. *The proteomics protocols handbook.* 2005;571–607. <https://doi.org/10.1385/1-59259-890-0:571>
 31. 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. *Ann Colorec Res.* 2021;9(1):32–9. <https://doi.org/10.30476/acrr.2021.89642.1080>.
 32. Mitra D, Das Mohapatra PK. Cold adaptation strategy of psychrophilic bacteria: an in-silico analysis of isocitrate dehydrogenase. *Syst Microbiol Biomanufac.* 2021;1(4):483–93. <https://doi.org/10.1007/s43393-021-00041-z>.
 33. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic acids Res.* 2003;31(13):3784–8. <https://doi.org/10.1093/nar/gkg563>.
 34. Hansen JC, Lu X, Ross ED, Woody RW. Intrinsic protein disorder, amino acid composition, and histone terminal domains. *J Bio Chem.* 2006;281(4):1853–6. <https://doi.org/10.1074/jbc.R500022200>.
 35. Lieutaud P, Ferron F, Uversky AV, Kurgan L, Uversky VN, Longhi S. How disordered is my protein and what is its disorder for? A guide through the “dark side” of the protein universe. *Intrinsically Disord Prot.* 2016;4(1):e1259708. <https://doi.org/10.1080/21690707.2016.1259708>.
 36. Lindring R, Jensen LJ, Diella F, Bork P, Gibson TJ, Russell RB. Protein disorder prediction: implications for structural proteomics. *Structure.* 2003;11(11):1453–9. <https://doi.org/10.1016/j.str.2003.10.002>.
 37. 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>.
 38. Kumar TA. CFSSP: Chou and Fasman secondary structure prediction server. *Wide Spectrum.* 2013;1(9):15–9. <https://doi.org/10.5281/zenodo.50733>.
 39. Vriend G. WHAT IF: a molecular modeling and drug design program. *J Mol Graph.* 1990;8(1):52–6. [https://doi.org/10.1016/0263-7855\(90\)80070-V](https://doi.org/10.1016/0263-7855(90)80070-V).
 40. Tina KG, Bhadra R, Srinivasan N. PIC: protein interactions calculator. *Nucleic Acids Res.* 2007;35(2):W473–6. <https://doi.org/10.1093/nar/gkm423>.
 41. Jubb HC, Higuero AP, Ochoa-Montaña 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(3):365–71. <https://doi.org/10.1016/j.jmb.2016.12.004>.
 42. Chong SH, Lee C, Kang G, Park M, Ham S. Structural and thermodynamic investigations on the aggregation and folding of acylphosphatase by molecular dynamics simulations and solvation free energy analysis. *J Am Chem.* 2011;133(18):7075–83. <https://doi.org/10.1021/ja1116233>.
 43. Tian W, Chen C, Lei X, Zhao J, Liang J. CASTp 3.0: computed atlas of surface topography of proteins. *Nucleic acids Res.* 2018;46(W1):363–7. <https://doi.org/10.1093/nar/gky473>.
 44. Blom N, Sicheritz-Pontén T, Gupta R, Gammeltoft S, Brunak S. Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics.* 2004;4(6):1633–49. <https://doi.org/10.1002/pmic.200300771>.
 45. Pires DE, Ascher DB, Blundell TL. DUET: a server for predicting effects of mutations on protein stability using an integrated computational approach. *Nucleic Acids Res.* 2014;42(W1):W314–9. <https://doi.org/10.1093/nar/gku411>.
 46. Kyte J, Doolittle RF. A simple method for displaying the hydrophobic character of a protein. *J Mol Bio.* 1982;157(1):105–32. [https://doi.org/10.1016/0022-2836\(82\)90515-0](https://doi.org/10.1016/0022-2836(82)90515-0).
 47. Rath SL, Kumar K. Investigation of the effect of temperature on the structure of SARS-Cov-2 spike protein by molecular dynamics simulations. *Front Mol Biosci.* 2020. <https://doi.org/10.3389/fmolb.2020.583523>.
 48. Christgen SL, Becker DF. Role of proline in pathogen and host interactions. *Antioxid Redox Signal.* 2019;30(4):683–709. <https://doi.org/10.1089/ars.2017.7335>.
 49. Faber M, Faber ML, Papaneri A, Bette M, Weihe E, Dietzschold B, Schnell MJ. A single amino acid change in rabies virus glycoprotein increases virus spread and enhances virus pathogenicity. *J Virol.* 2005;79(22):14141–8. <https://doi.org/10.1128/JVI.79.22.14141-14148.2005>.
 50. Askarian F, Uchiyama S, Valderrama JA, Ajayi C, Sollid JU, Van Sorge NM, Nizet V, van Strijp JA, Johannessen M. Serine-aspartate repeat protein D increases *Staphylococcus aureus* virulence and survival in blood. *Infect Immun.* 2017;85(1):e00559-e616. <https://doi.org/10.1128/IAI.00559-16>.
 51. Ikai A. Thermostability and aliphatic index of globular proteins. *J Biochem.* 1980;88(6):1895–8. <https://doi.org/10.1093/oxfordjournals.jbchem.a133168>.
 52. Cho MK, Kim HY, Bernado P, Fernandez CO, Blackledge M, Zweckstetter M. Amino acid bulkiness defines the local conformations and dynamics of natively unfolded α -synuclein and tau. *J Am Chem.* 2007;129(11):3032–3. <https://doi.org/10.1021/ja067482k>.
 53. Zhang L, Hermans J. Hydrophilicity of cavities in proteins. *Proteins Struct Funct Bioinform.* 1996;24(4):433–8. [https://doi.org/10.1002/\(SICI\)1097-0134](https://doi.org/10.1002/(SICI)1097-0134).
 54. Kuhn LA, Swanson CA, Pique ME, Tainer JA, Getzoff ED. Atomic and residue hydrophilicity in the context of folded protein structures. *Proteins Struct Funct Bioinform.* 1995;23(4):536–47. <https://doi.org/10.1002/prot.340230408>.
 55. He B, Wang K, Liu Y, Xue B, Uversky VN, Dunker AK. Predicting intrinsic disorder in proteins: an overview. *Cell Res.* 2009;19(8):929–49. <https://doi.org/10.1038/cr.2009.87>.
 56. Uversky VN. The alphabet of intrinsic disorder: II Various roles of glutamic acid in ordered and intrinsically disordered proteins. *Intrinsically Disord Prot.* 2013;1(1):e24684. <https://doi.org/10.4161/idp.24684>.
 57. Hamdi K, Salladini E, O'Brien DP, Brier S, Chenal A, Yacoubi I, Longhi S. Structural disorder and induced folding within two cereal, ABA stress and ripening (ASR) proteins. *Sci Rep.* 2017;7(1):1–21. <https://doi.org/10.1038/s41598-017-15299-4>.
 58. Uversky VN. Intrinsic disorder, protein–protein interactions, and disease. *Adv Protein Chem Struct Biol.* 2018;1(110):85–121. <https://doi.org/10.1016/bs.apcsb.2017.06.005>.
 59. Xie H, Vucetic S, Iakoucheva LM, Oldfield CJ, Dunker AK, Uversky VN, Obradovic Z. Functional anthology of intrinsic disorder. 1. Biological processes and functions of proteins with long disordered regions. *J Proteome Res.* 2007;6(5):1882–98. <https://doi.org/10.1021/pr060392u>.
 60. Oldfield CJ, Xue B, Van YY, Ulrich EL, Markley JL, Dunker AK, Uversky VN. Utilization of protein intrinsic disorder knowledge in structural proteomics. *BBA Proteins Proteoms.*

- 2013;1834(2):487–98. <https://doi.org/10.1016/j.bbapap.2012.12.003>.
61. 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;1(174):1–78. <https://doi.org/10.1016/bs.pmbts.2020.03.001>.
 62. Shi Z, Olson CA, Bell AJ Jr, Kallenbach NR. Stabilization of α -helix structure by polar side-chain interactions: complex salt bridges, cation– π interactions, and C-H... OH-bonds. *Peptide Sci Original Res Biomol.* 2001;60(5):366–80. [https://doi.org/10.1002/1097-0282\(2001\)60:5%3C366](https://doi.org/10.1002/1097-0282(2001)60:5%3C366).
 63. 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](https://doi.org/10.1016/S1359-0278(96)00009-0).
 64. Facchiano AM, Colonna G, Ragone R. Helix stabilizing factors and stabilization of thermophilic proteins: an X-ray based study. *Protein Eng.* 1998;11(9):753–60. <https://doi.org/10.1093/protein/11.9.753>.
 65. Jelesarov I, Karshikoff A. Defining the role of salt bridges in protein stability. In *Protein Structure, Stability, and Interactions 2009* (pp. 227–260). Humana Press. https://doi.org/10.1007/978-1-59745-367-7_10
 66. Basu S, Mukharjee D. Salt-bridge networks within globular and disordered proteins: characterizing trends for designable interactions. *J Mol Model.* 2017;23(7):1–17. <https://doi.org/10.1007/s00894-017-3376-y>.
 67. 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):1–9. <https://doi.org/10.1038/s41598-020-58825-7>.
 68. Sindelar CV, Hendsch ZS, Tidor B. Effects of salt bridges on protein structure and design. *Prot Sci.* 1998;7(9):1898–914. <https://doi.org/10.1002/pro.5560070906>.
 69. Quintanar L, Rivillas-Acevedo L. Studying metal ion–protein interactions, electronic absorption; circular dichroism; and electron paramagnetic resonance, in *Protein-Ligand Interactions*, Humana Press, Totowa, NJ, 2013, pp. 267–297. https://doi.org/10.1007/978-1-62703-398-5_10
 70. Burley SK, Petsko GA. Aromatic-aromatic interaction: a mechanism of protein structure stabilization. *Science.* 1985;229(4708):23–8. <https://doi.org/10.1126/science.3892686>.
 71. Burley SK, Petsko GA. Electrostatic interactions in aromatic oligopeptides contribute to protein stability. *Trends Biotechnol.* 1989;7(12):354–9. [https://doi.org/10.1016/0167-7799\(89\)90036-X](https://doi.org/10.1016/0167-7799(89)90036-X).
 72. 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>.
 73. Chiche L, Gregoret LM, Cohen FE, Kollman PA. Protein model structure evaluation using the solvation free energy of folding. *Proc Natl Acad Sci USA.* 1990;87(8):3240–3. <https://doi.org/10.1073/pnas.87.8.3240>.
 74. Nishi H, Hashimoto K, Panchenko AR. Phosphorylation in protein-protein binding: effect on stability and function. *Structure.* 2011;19(12):1807–15. <https://doi.org/10.1016/j.str.2011.09.021>.