In-silico study of SARS-CoV-2 and SARS with special reference to intra-protein interactions, A plausible explanation for stability, divergency and severity of SARS-CoV-2

Debanjan Mitra
Raiganj University
Aditya K. Pal
Raiganj University
Pradeep Kumar Das Mohapatra (pkdmvu@gmail.com)
Raiganj University

Research Article

Keywords: COVID-19, evolution, cyclic salt bridge, intra-protein interactions, protein engineering

Posted Date: December 2nd, 2020

DOI: https://doi.org/10.21203/rs.3.rs-73762/v2

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 discloses in 2002, but not outspread worldwide. After 18 years, in 2020, it reemerges and outspread worldwide as SARS-CoV-2 (COVID 19), as the most treacherous virus creating disease in the world. Is it possible to create a favorable evolution within this (18 years) short time? 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? This study will find all those queries. Here, we took 4 types of protein sequences from SARS-CoV-2 and SARS are retrieved from the database to check their physicochemical and structural properties. Results showed that charged residues are playing a pivotal role in SARS-CoV-2 evolution. Those charged residues also contribute to helix stabilization of SARS-CoV-2. Formation of cyclic salt bridge and other intra-protein interactions also play crucial role in SAS-CoV-2. This comparative study will help to understand the evolution from SARS to SARS-CoV-2 and also helps in protein engineering.

Introduction

Disease caused by SARS-CoV-2 is 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 8,098 people and 774 deaths having a mortality rate of 11%. After that in 2012, MERS-CoV appeared in the Arabian Peninsula and had outspread into 27 countries, infecting a total of 2,494 individuals and took 858 lives with a mortality rate of 34%. Recently SARS-CoV-2 has been elevated in Wuhan city, Hubei province of China in December 2019. Till now, there are over five crore cases of COVID-19 and over a 1.3 million deaths (mortality rate around 2.40%) have been reported to globally affect 218 countries. On March 11, 2020, the World Health Organization announces the COVID-19 pandemic a public health emergency of global concern. All ages of people can catch this viral infection but immune-compromised aged people having co-morbidities are most vulnerable. Susceptibility of age, males with chronic diseases (like- diabetes, heart disease, cancer etc.) is higher than other groups of people[1]. This virus can be easily transmitted through the droplets generated at the time of coughing and sneezing by the infected people[2]. These infectious droplets can be spread up to 1–2 meters 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.[3]. 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 the one-week recovery started. It has been observed in patients that the progression of this disease increases the release of cytokine including interleukin (IL)-6 and IL-10 whereas the levels of CD4+T and CD8+T are reduced[4]. As of now, there is no approved treatment for COVID-19 but anti-viral drugs such as Remdesivir, Tocilizumab are in use for treatment[5].

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 nm to 140 nm[6]. There are four subtypes' such as alpha, beta, gamma, and delta type of coronavirus. Most of the highly pathogenic viruses; like- Severe acute respiratory syndrome coronavirus (SARS-CoV), Middle-East respiratory syndrome coronavirus (MERS-CoV), and SARS-CoV-2, all are a type of β-coronavirus[7]. 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[8-9]. Other ORFs of the genome near the 3'-terminus encodes the four main structural proteins, spike glycoproteins, membrane, envelope, and nucleocapsid proteins[10]. Genome analysis of SARS-CoV-2 revealed that there are 79.5% and 97% of similarity with the whole genome sequences of SARS-CoV and bat SARS-CoV respectively (Chen et al., 2020). SARS-CoV-2 enters the host respiratory mucosa by binding with the receptor of angiotensin-converting enzyme 2 (ACE2) with its spike glycoproteins[11]. A recent study has shown that SARS-CoV-2 binds with ACE2 with 10-fold higher affinity compared to SARS-CoV[12]. The basic reproduction number (R_0), which is the average number of secondary infections produced by patients, is between 2.47–2.86 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[13-15]. 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 coronaviruses (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[16]. SARS-CoV ORF 3b, ORF 6, and N proteins inhibit the expression of beta interferon (IFN-β)[17]. 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[18].

With the unavailability of specific vaccines and anti-viral drugs for nCoV, science demands sincere efforts in the field of drug design and discovery for COVID-19. Since 2002, SARS has present on this earth. But it creates a dangerous situation and makes a pandemic situation after 18 years. Why? Why is this virus so harmful to us? What are the basic differences between SARS-CoV-2 and SARS? How evolution makes them stronger than SARS? How can they gain stability for such extreme environments? Salt bridges and other intra-protein interactions are playing an essential role in protein stability to operate their physiological activity in an extreme environment[19-21]. Do they play a vital role in SARS-CoV-2? To find all those queries, all the 4 types (spike proteins, membrane proteins, nucleoproteins and ORF proteins) protein sequences of SARS-CoV-2 and SARS were extracted from the database for physicochemical and structural properties analysis. To check their stability salt bridges and other factors are also extracted.

Materials And Methods

Dataset

A detailed investigation of those sequences and structure of SARS-CoV-2 was performed with reference to the old SARS. Here we took 4 types of SARS-CoV-2 and SARS proteins i.e. spike proteins, membrane proteins, nucleoproteins and ORF proteins (ORF 3, ORF 6, ORF 7, ORF 8 and ORF 9). All protein sequences of SARS-CoV-2 and SARS were retrieved from UNIPROT[22] database (Table 1). The crystal structures of SARS-CoV-2 and SARS proteins were retrieved from the RCSB protein database (PDB)[23]. In structural comparison, we took the protease protein structure, cause it is heavily used as target in drug discovery.

Table 1. UniProtKB ID of SARS-CoV-2 and SARS proteins that were taken from database for *in-silico* study

Types of proteins	Uniprot ID	
	SARS-CoV-2	SARS
Spike proteins	P0DTC2	P59594
	A0A679G9E9	A0A1W6S788
Membrane proteins	P0DTC4	P59637
	P0DTC5	P59596
Nulceoproteins	P0DTC9	P59595
	A0A6C0T6Z7	A0A3S6GSS4
ORF Proteins	P0DTC3 (ORF 3a)	P59632 (ORF3a)
	P0DTC7 (ORF 7a)	P59633 (ORF3b)
	P0DTC6 (ORF 6)	Q7TFA1 (ORF 7b)
	P0DTD2 (ORF 9b)	P59634(ORF6)
	P0DTC8 (ORF 8)	P59635 (ORF7a)
		P59636 (ORF9b)

Physicochemical and evolutionary properties

All protein sequences were subject to multiple sequence alignment was done for all the sequences with the help of CLUSTAL Omega[24]. Both block and non-block FASTA[25]format of the sequences were analyzed. Block of the sequence was prepared by BLOCK Maker[26]. Non-block and block both formats were analyzed by ProtParam server[27] and ProtScale[28] server for calculation of physicochemical properties like 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, R, S), order forming residues (i.e. C, F, W, Y) are calculated from amino acid compositions[29-35]. Intrinsic disorder regions of protein were analyzed by DisEMBL[36].

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 by using UCSF Chimera[37]. Analyses of the

secondary structure were done by CFSSP[38] 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 PIC server[40] and Arpeggio server[41]. Free solvation energy was calculated by ProWaVE server[42]. Surface area and volume was determined by CASTp server[43]. Phosphorylation sites of protein were identified by NetPhos server[44]. Protein mutations were analyzed by DUET[45].

Results

Effect of charged residues on SARS-CoV-2 sequence

Here D, E, H, R, K took as a charged residues and C, S, T, N, Q, Y, W took as uncharged polar residues. Amino acid compositions were calculated from the non-block format whereas block format was used to calculate disorder forming residues (Dis), order forming residues (Ofr), 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 that answer, we calculate all those physicochemical properties.

Spike proteins showed higher abundance (Fig. 1) of charged residues (except D) in SARS-CoV-2. Polar residues showed higher quantity (except T, W) in SARS-CoV-2. In nucleoproteins of SARS-CoV-2 D, K and R shows higher abundance and E, H shows lower abundance as charged amino acids. Polar residues in nucleoproteins also showed higher abundance (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. The 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. The higher number of disorder forming residues in SARS-CoV-2 indicates that it can easily create toxicity or disease in humans. The lower value of GRAVY (except nucleoproteins) indicates the hydrophilic nature of SARS-CoV-2. So, it can be easily mixed with aqueous or liquid medium. The aliphatic index is high in every SARS-CoV-2 proteins. High value of the aliphatic index in SARS-CoV-2 proved that SARS-CoV-2 is more thermally stable than SARS[47].

When we check the polarity of those proteins, it showed slightly high values in SARS-CoV-2 proteins 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 longer heating periods in hydrolysis[48]. They can tolerate heat better than SARS. The Kyte-Dollitle hydrophobicity scale indicates that the SARS-CoV-2 is hydrophilic in nature (Fig. 3). The hydrophilic nature of SARS-CoV-2 gives a clue that it can easily interact with water or aqueous medium and spread easily than SARS[49-50]. The intrinsic disorder regions are very much high in SARS-CoV-2 than SARS. High abundance of intrinsic disorder regions of SARS-CoV-2 indicates that it will more interact with other proteins than SARS[51-52].

Abundance of charged residues in helix of SARS-CoV-2

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

Table 2. 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

Charged residues showed higher abundance in every position (turn, helix, coil and sheet) of SARS-CoV-2 (Table 2) than SARS. Charged residues show higher abundance within the helix of both proteins. Introduction of high number charged residues in the helix results in proteins more resistant to the acidic environment or temperature denaturation and helps in increasing the stability[53-54]. Hydrophobic residues have higher abundance in SARS (except coil) than SARS-CoV-2. Polar residues also show higher abundance in SARS-CoV-2 than SARS.

Intra-protein interactions effect on stability of SARS-CoV-2

Salt bridges have a significant effect on protein stability[55-58]. Charged residues are participating in the formation of salt bridges. Normally 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 affect salt bridge formation to gain more stability. Other intra-protein interactions like, metal ion binding site[59], aromatic-aromatic interactions[60-61] are also helps in protein stabilization.

SARS-CoV-2 has large pocket area than SARS (Fig. 4), which gives it more protein-protein or protein-ligand interactions possibilities (Table 3). Volume of the protein is also high in SARS-CoV-2 than SARS. Protease from SARS-CoV-2 possess 7 isolated salt bridges and 1 network salt bridges, whereas SARS protease has 5 isolated and 1 network salt bridges. Result indicates that SARS-CoV-2 is highly stabilized by those salt bridges. Number of metal ion binding site is also high in SARS-CoV-2 than SARS. Free solvation energy is a thermodynamic factor that determines protein salvation or nature of denaturation[62]. By this property we can determine how fast proteins easily denature. Solvation free energy is also high in

SARS-CoV-2 than SARS which indicates the SARS-CoV-2 protein not easily denature in contact with solvent.

Table 3. 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})
5R80	669.47	1013.45	7	1	3	4786.55 Kcal/mol
2H2Z	228.27	835.26	5	1	2	3266.89 Kcal/mol

Aromatic-aromatic interactions show high number in SARS-CoV-2 than SARS (Table 4). Not only number, those residue (Phe8, Tyr37, Phe103, Tyr101, Phe150, Phe159) which participate in aromatic-aromatic interactions are forming a very long network, which is never been reported in any proteomics research. SARS-CoV-2 has 9 isolated and 1 network aromatic-aromatic interactions where as SARS has only 9 isolated aromatic-aromatic interactions.

Table 4. Aromatic-aromatic interactions of of SARS-CoV-2 (5R80) and SARS (2H2Z)

Protein	Position	Residue	Position	Residue	D(centroid-centroid)	Dihedral Angle
SARS-CoV-2 (5R80)	3	PHE	291	PHE	4.96	130.9
	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
	112	PHE	161	TYR	5.41	142.21
	126	TYR	140	PHE	6.38	64.71
	134	PHE	182	TYR	6.33	164.01
	150	PHE	159	PHE	6.43	56.09
	161	TYR	182	TYR	6.47	150.59
	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

The number of phosphorylation site (Fig. 5) in SARS-CoV-2 is 54, whereas the number of phosphorylation site in SARS is 45. That means SARS-CoV-2 has higher number of phosphorylation sites than SARS. The high number of phosphorylation site in SARS-CoV-2 increase the strength of protein-protein interactions and also helps in stability[63].

SARS-CoV-2 has cyclic salt bridge

Generally proteins have two types of salt bridges, isolated and network salt bridges. Both proteins have only one network salt bridge. But SARS-CoV-2 has special engineered salt bridge (Fig. 6) which forms a cyclic salt bridge (R131-E290,K137-E290,R131-D197,K137-D197,R131-D289), whereas SARS has normal network salt bridge. Novel cyclic salt bridge might have a great role in its stability.

Favorable point mutations of SARS-CoV-2

Result of MSA (Figure 7) of both structure shows some point mutations occur in SARS-CoV-2. So, we have analyzed their effect on SARS-CoV-2 protein stability.

Total 11 mutations have been identified between which 8 are favorable and 3 are unfavorable for SARS-CoV-2 protein stability (Table 5). Residue number 35 which was threonine of SARS substitute by valine in SARS-CoV-2 after mutation, contribute high energy i.e. -2.24 Kcal/mol in protein stability. By those specific point mutations SARS-CoV-2 ultimately got -7.46 kcal/mol energies which make them more stable than SARS.

Table 5. Effect of amino acid mutations in SARS-CoV-2 with their contributing energies

Residue in SASS-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

Conclusion

The comparative study between SARS-CoV-2 and SARS reveals that how favorable evolution makes SARS-CoV-2 more dangerous and stronger than SARS. Those acidic and basic residues play a major role in evolution. Charged residues also present in helix to increase the protein stability. Also the long network aromatic-aromatic interactions have an effect on its stability. This is the first report of cyclic salt bridge and long network aromatic-aromatic interaction in structural biology. Increasing of metal ion binding site, phosphorylation site also play crucial role in SARS-CoV-2 protein stability. So, the evolution of SARS-CoV-2 has a great role in its stability. Those point mutations show how SARS-CoV-2 engendered itself to gain more stability. It is also a clue for how to stop SARS-CoV-2 severity of the infection. Protein engineering helps us in this process. This study will also beneficial for drug or vaccine development against SARS-CoV-2.

Declarations

Authors contribution

D.M. and A.P. conceived and designed the project. P.K.D.M. conducted initial manual verifications. Protein sequence and structure were identified by D.M. Analysis of those results were done by D.M. Draft of the manuscript was prepared by D.M. and A.P. 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.

Conflict of interest

The authors declare that they have no conflict of interest.

Funding

None

References

1. 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. The Lancet. 2020, 395(10223), 507-513.
2. Bastola, A.; Sah, R.; Rodriguez-Morales, A.J.; Lal, B.K.; Jha, R.; Ojha, H.C.; Shrestha, B.; Chu, D.K.; Poon, L.L.; Costello, A.; Morita, K. The first 2019 novel coronavirus case in Nepal. Lancet Infec. Dis. 2020, 20(3), 279-280.
3. Kampf, G.; Todt, D.; Pfaender, S.; Steinmann, E. Persistence of coronaviruses on inanimate surfaces and their inactivation with biocidal agents. J Hos. Infec. 2020, 104(3), 246-51.
4. 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.

5. Hung, I.F.; Lung, K.C.; Tso, E.Y.; Liu, R.; Chung, T.W.; Chu, M.Y.; Ng, Y.Y.; Lo, J.; Chan, J.; Tam, A.R.; Shum, H.P. 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. The Lancet. 2020, 395(10238), 1695-704.
6. Cascella, M.; Rajnik, M.; Cuomo, A.; Dulebohn, S.C.; Di Napoli, R. Features; evaluation and treatment coronavirus (COVID-19). InStatpearls [internet] 2020. StatPearls Publishing.
7. Chan, J.F.W.; To, K.K.W.; Tse, H.; Jin, D.Y.; Yuen, K.Y. Interspecies transmission and emergence of novel viruses, lessons from bats and birds. Trends Microbiol. 2013, 21(10), 544-555.
8. Ziebuhr, J.; Snijder, E.J.; Gorbalenya, A.E. Virus-encoded proteinases and proteolytic processing in the Nidovirales. Journal of General Virology. 2000, 81(4), 853-879.
9. Masters, P.S. The molecular biology of coronaviruses. Virus. Res. 2006, 66, 193-292.
10. Yoshimoto, F.K. 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(1), 198-216.
11. Hoffmann, M.; Kleine-Weber, H.; Krüger, N.; Mueller, M.A.; 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.
12. Wrapp, D.; Wang, N.; Corbett, K.S.; Goldsmith, J.A.; Hsieh, C.L.; Abiona, O.; Graham, B.S.; McLellan, J.S. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science. 2020, 367(6483), 1260-1263.
13. Remais, J. Modelling environmentally-mediated infectious diseases of humans, transmission dynamics of schistosomiasis in China. InModelling parasite transmission and control 2010 (pp. 79-98). Springer; New York; NY.
14. Wu, J.T.; Leung, K.; Leung, G.M. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan; China, a modelling study. The Lancet. 2020, 395(10225), 689-697.
15. Lipsitch, M.; Cohen, T.; Cooper, B.; Robins, J.M.; Ma, S.; James, L.; Gopalakrishna, G.; Chew, S.K.; Tan, C.C.; Samore, M.H.; Fisman D. Transmission dynamics and control of severe acute respiratory syndrome. Science. 2003, 300(5627), 1966-1970.
16. Walls, A.C.; Park, Y.J.; Tortorici, M.A.; Wall, A.; McGuire, A.T.; Veesler, D. Structure; function; and antigenicity of the SARS-CoV-2 spike glycoprotein. Cell. 2020, 181(2), 281-292.
17. Kopecky-Bromberg, S.A.; Martínez-Sobrido, L.; Frieman, M.; Baric, R.A.; 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), 548-557.
18. Ruch, T.R.; Machamer, C.E. The coronavirus E protein, assembly and beyond. Viruses. 2012, 4(3), 363-382.
19. Kumar, S.; Nussinov, R. Salt bridge stability in monomeric proteins. J Mol. Bio. 1999, 293(5), 1241-1255.

20. Makwana, K.M.; Mahalakshmi, R. Implications of aromatic–aromatic interactions, From protein structures to peptide models. Prot. Sci. 2015 De, 24(12), 1920-1933.
21. Gallivan JP; Dougherty DA. A computational study of cation− π interactions vs salt bridges in aqueous media, implications for protein engineering. J Am. Chem. 2000, 122(5), 870-874.
22. Consortium, U. UniProt, a hub for protein information. Nucleic Acids Res. 2015, 43(D1), D204-212.
23. Rose, P.W.; Bi, C.; Bluhm, W.F.; Christie, C.H.; Dimitropoulos, D.; Dutta, S.; Green, R.K.; Goodsell, D.S.; Prlić, A.; Quesada, M.; Quinn, G.B. The RCSB Protein Data Bank, new resources for research and education. Nucleic Acids Res. 2012, 41(D1), D475-D82.
24. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T.J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; Thompson, J.D. Fast; scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 2011, 7(1), 539.
25. Pearson, W.R. Using the FASTA program to search protein and DNA sequence databases. InComputer Analysis of Sequence Data 1994 (pp. 307-331). Humana Press.
26. Henikoff, S.; Henikoff, J.G.; Alford, W.J.; Pietrokovski, S. Automated construction and graphical presentation of protein blocks from unaligned sequences. Gene. 1995, 163(2), GC17-GC26.
27. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Wilkins, M.R.; Appel, R.D.; Bairoch, A. Protein identification and analysis tools on the ExPASy server. InThe proteomics protocols handbook 2005 (pp. 571-607). Humana press.
28. Gasteiger, E.; Gattiker, A.; Hoogland, C.; Ivanyi, I.; Appel, R.D.; Bairoch, A. ExPASy, the proteomics server for in-depth protein knowledge and analysis. Nucleic acids Res. 2003, 31(13), 3784-3788.
29. Hansen, J.C.; Lu, X.; Ross, E.D.; Woody, R.W. Intrinsic protein disorder; amino acid composition; and histone terminal domains. J Bio. Chem. 2006, 281(4), 1853-1856.
30. Campen, A.; Williams, R.M.; Brown, C.J.; Meng, J.; Uversky, V.N.; Dunker, A.K. TOP-IDP-scale, a new amino acid scale measuring propensity for intrinsic disorder. Prot. Pept. Lett. 2008 Sep 1;15(9),956-63.
31. He, B.; Wang, K.; Liu, Y.; Xue, B.; Uversky, V.N.; Dunker, A.K. Predicting intrinsic disorder in proteins, an overview. Cell Res. 2009, 19(8), 929-949.
32. Uversky, V.N. The alphabet of intrinsic disorder, II. Various roles of glutamic acid in ordered and intrinsically disordered proteins. Intrinsically Disord. Prot. 2013, 1(1), e24684.
33. Lieutaud, P.; Ferron, F.; Uversky, A.V.; Kurgan, L.; Uversky, V.N.; 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.
34. Hamdi, K.; Salladini, E.; O’Brien, D.P.; 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.
35. Biswas, I.; Mitra, D.; Bandyopadhyay, A.K.; Mohapatra, P.K. Contributions of protein microenvironment in tannase industrial applicability, An in-silico comparative study of pathogenic and non-pathogenic bacterial tannase. Heliyon. 2020, 6(11), e05359.

36. Linding, R.; Jensen, L.J.; Diella, F.; Bork, P.; Gibson, T.J.; Russell, R.B. Protein disorder prediction, implications for structural proteomics. Structure. 2003, 11(11), 1453-1459.
37. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera—a visualization system for exploratory research and analysis. J Comput. Chem. 2004, 25(13), 1605-1612.
38. Kumar, T.A. CFSSP, Chou and Fasman secondary structure prediction server. Wide Spectrum. 2013, 1(9), 15-19.
39. Vriend, G. WHAT IF, a molecular modeling and drug design program. J Mol. Graph. 1990, 8(1), 52-56.
40. Tina, K.G.; Bhadra, R.; Srinivasan, N. PIC, protein interactions calculator. Nucleic acids Res. 2007, 35(suppl_2), W473-W476.
41. Jubb, H.C.; Higueruelo, A.P.; Ochoa-Montaño, B.; Pitt, W.R.; Ascher, D.B.; Blundell, T.L. Arpeggio, a web server for calculating and visualising interatomic interactions in protein structures. Journal of molecular biology. 2017, 429(3), 365-371.
42. Chong, S.H.; 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, 33(18), 7075-7083.
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), W363-W367.
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), 1633-1649.
45. Pires, D.E.; Ascher, D.B.; Blundell, T.L. DUET, a server for predicting effects of mutations on protein stability using an integrated computational approach. Nucleic acids Res. 2014, 42(W1), W314-W319.
46. Kyte, J.; Doolittle, R.F. A simple method for displaying the hydropathic character of a protein. J Mol. Bio. 1982, 157(1), 105-132.
47. Ikai, A. Thermostability and aliphatic index of globular proteins. J Biochem. 1980, 88(6), 1895-1898.
48. Cho, M.K.; Kim, H.Y.; Bernado, P.; Fernandez, C.O.; 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), 3032-3033.
49. Zhang, L.; Hermans, J. Hydrophilicity of cavities in proteins. Proteins. 1996, 24(4), 433-438.
50. Kuhn, L.A.; Swanson, C.A.; Pique, M.E.; Tainer, J.A.; Getzoff, E.D. Atomic and residue hydrophilicity in the context of folded protein structures. Proteins. 1995, 23(4), 536-547.
51. Xie, H.; Vucetic, S.; Iakoucheva, L.M.; Oldfield, C.J.; Dunker, A.K.; Uversky, V.N.; Obradovic, Z. Functional anthology of intrinsic disorder. 1. Biological processes and functions of proteins with long disordered regions. J Proteome Res. 2007, 6(5), 1882-1898.
52. Oldfield, C.J.; Xue, B.; Van, Y.Y.; Ulrich, E.L.; Markley, J.L.; Dunker, A.K.; Uversky, V.N. Utilization of protein intrinsic disorder knowledge in structural proteomics. Biochimica et Biophysica Acta (BBA)-

Proteins and Proteomics. 2013, 1834(2), 487-498.
53. Villegas, V.; Viguera, A.R.; Avilés, F.X.; Serrano, L. Stabilization of proteins by rational design of α-helix stability using helix/coil transition theory. Fold. Des. 1996, 1(1), 29-34.
54. Facchiano, A.M.; Colonna, G.; Ragone, R. Helix stabilizing factors and stabilization of thermophilic proteins, an X-ray based study. Protein Eng. 1998, 11(9), 753-760.
55. Jelesarov, I.; Karshikoff, A. Defining the role of salt bridges in protein stability. InProtein Structure; Stability; and Interactions 2009 (pp. 227-260). Humana Press.
56. Basu, S.; Mukharjee, D. Salt-bridge networks within globular and disordered proteins, characterizing trends for designable interactions. Journal of Molecular Modeling. 2017, 23(7), 206.
57. Panja, A.S.; 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.
58. Sindelar, C.V.; Hendsch, Z.S.; Tidor, B. Effects of salt bridges on protein structure and design. Prot. Sci. 1998, 7(9), 1898-1914.
59. Quintanar, L.; Rivillas-Acevedo, L. Studying metal ion–protein interactions, electronic absorption; circular dichroism; and electron paramagnetic resonance. InProtein-Ligand Interactions 2013 (pp. 267-297). Humana Press; Totowa; NJ.
60. Burley, S.K.; Petsko, G.A. Aromatic-aromatic interaction, a mechanism of protein structure stabilization. Science. 1985, 229(4708), 23-28.
61. Burley, S.K.; Petsko, G.A. Electrostatic interactions in aromatic oligopeptides contribute to protein stability. Trends Biotechnol. 1989, 7(12), 354-359.
62. Chiche, L.; Gregoret, L.M.; Cohen, F.E.; Kollman, P.A. Protein model structure evaluation using the solvation free energy of folding. Natl. Acad. Sci. U. S. A. 1990, 87(8), 3240-3243.
63. Nishi, H.; Hashimoto, K.; Panchenko, A.R. Phosphorylation in protein-protein binding, effect on stability and function. Structure. 2011, 19(12), 1807-1815.

Figures

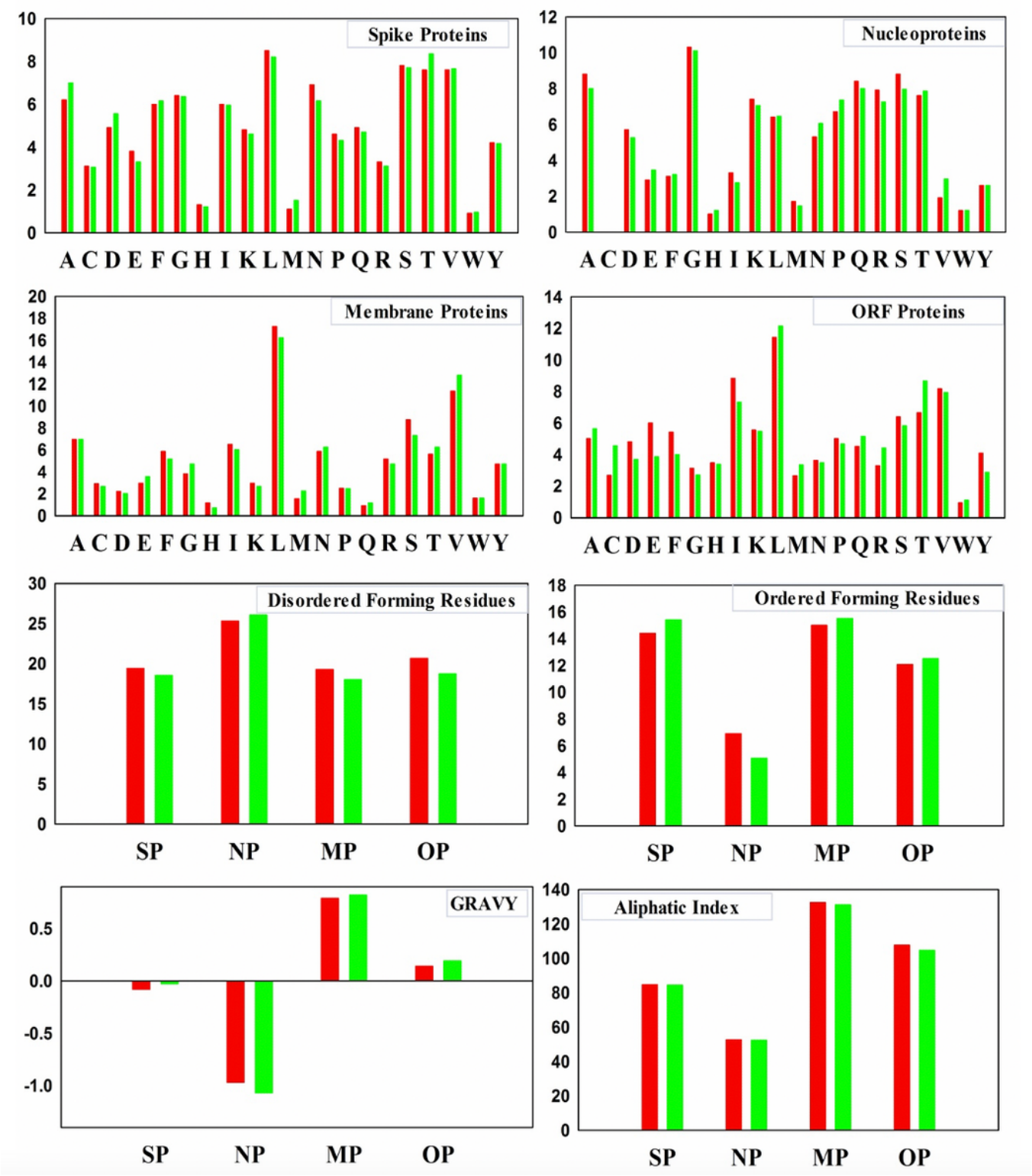

Figure 1

Comparative analysis of physicochemical properties like amino acid compositions, disorder forming residues, order forming residues, GRAVY, aliphatic index of spike proteins (SP), nucleoproteins (NP), membrane proteins (MP), ORF proteins (OP) from SARS-CoV-2 (red bar) and SARS (green bar).

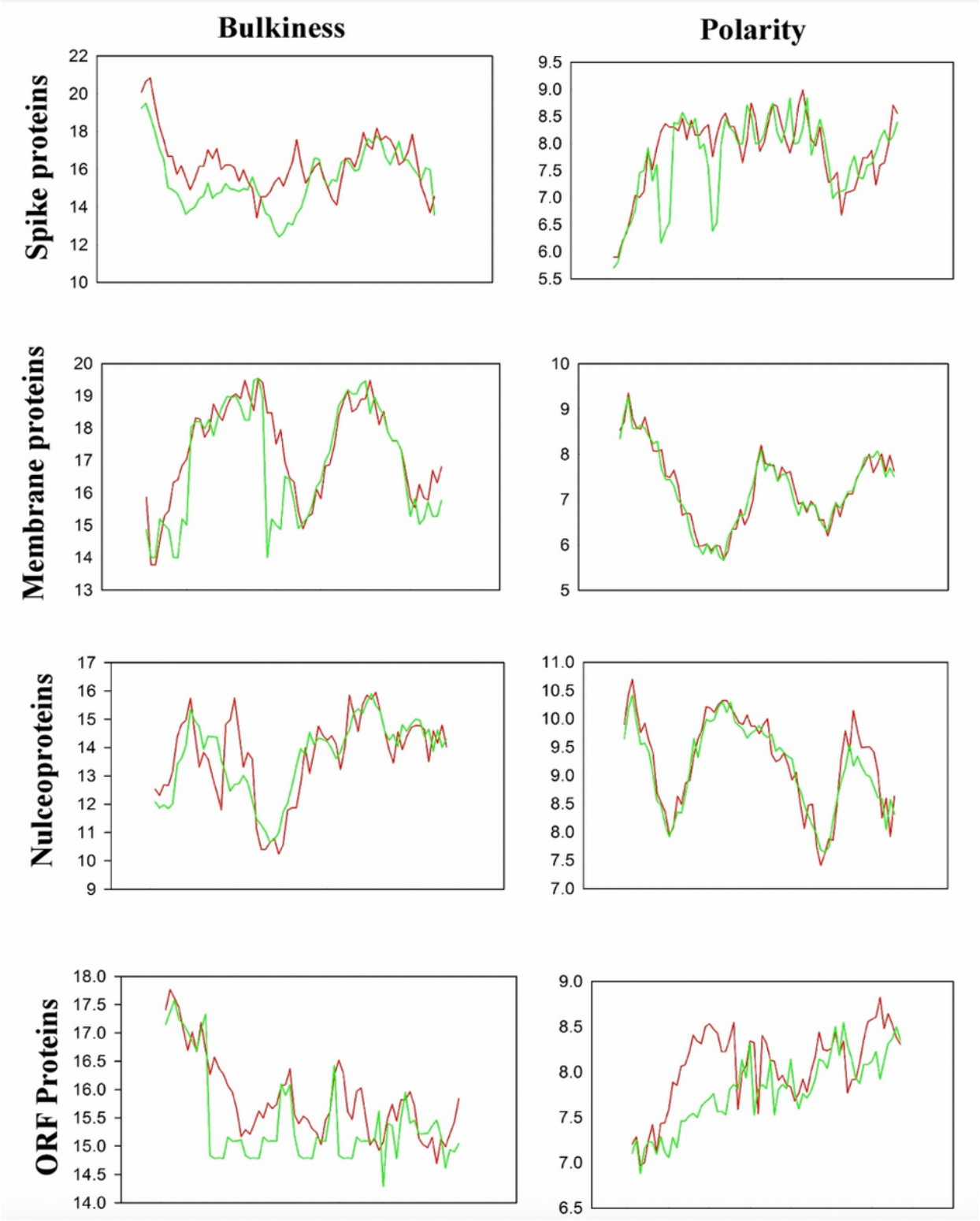

Figure 2

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).

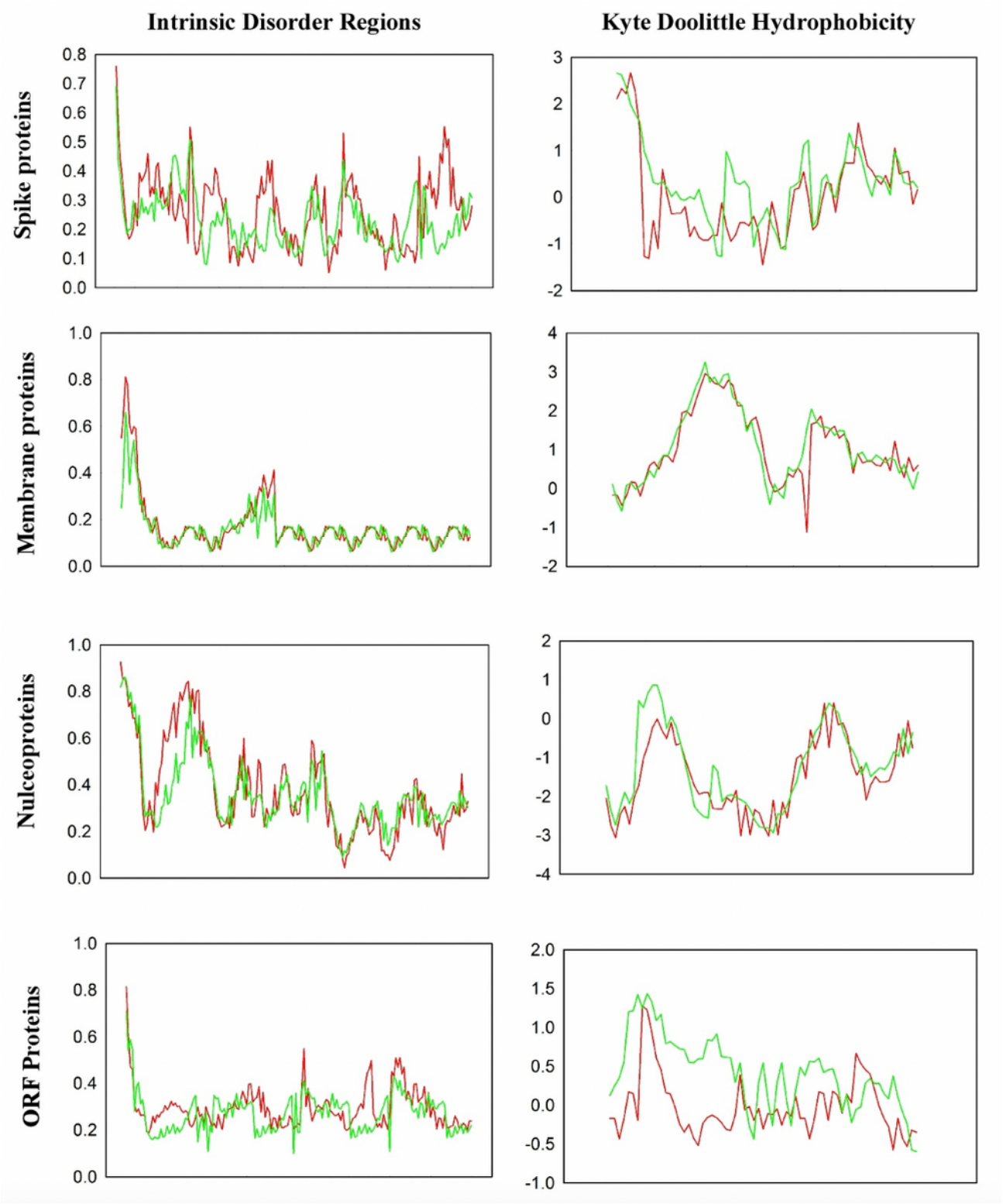

Figure 3

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

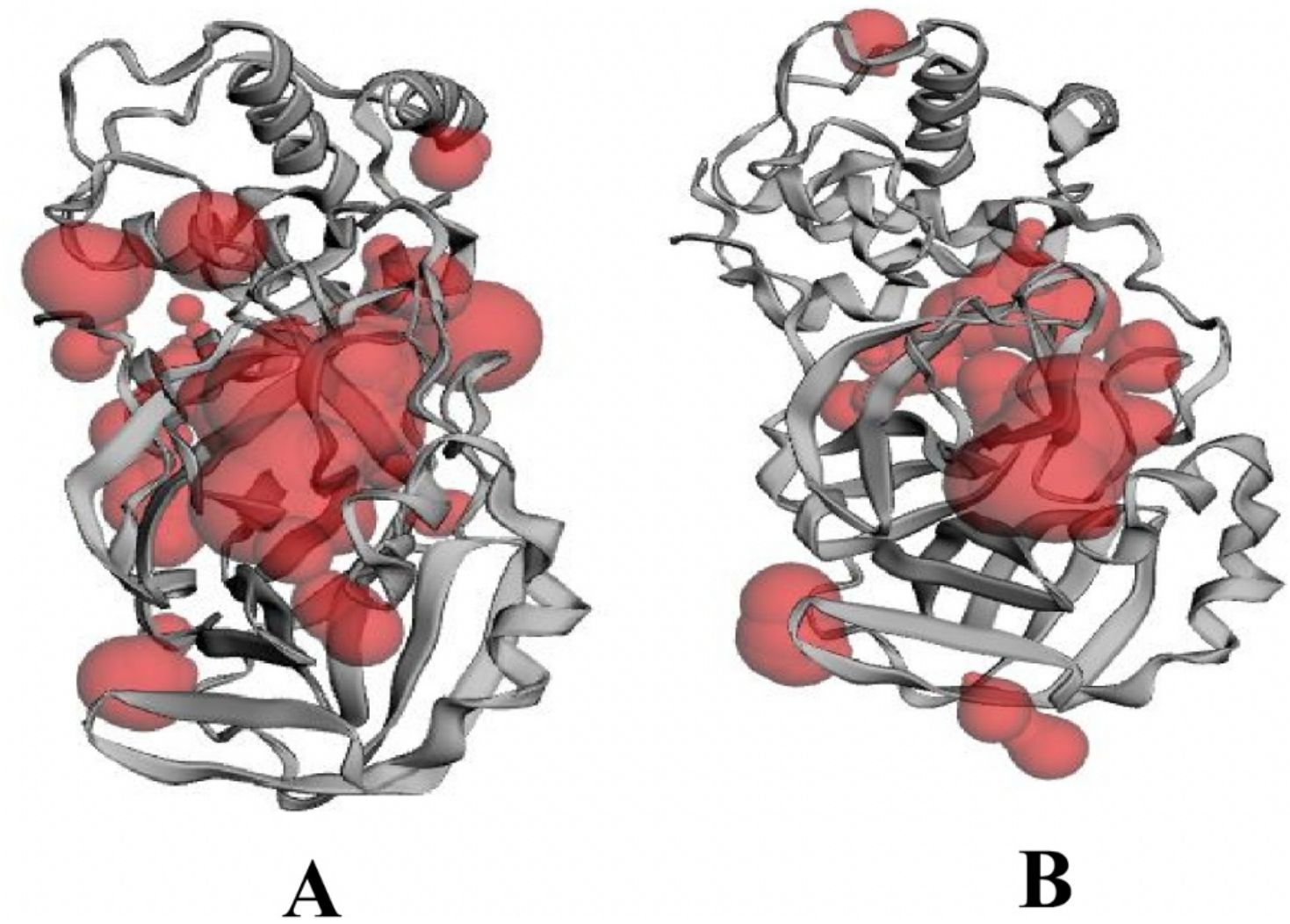

Figure 4

Large pocket area (red zone) in 3D protein structures of SARS-CoV-2 (A) than SARS (B).

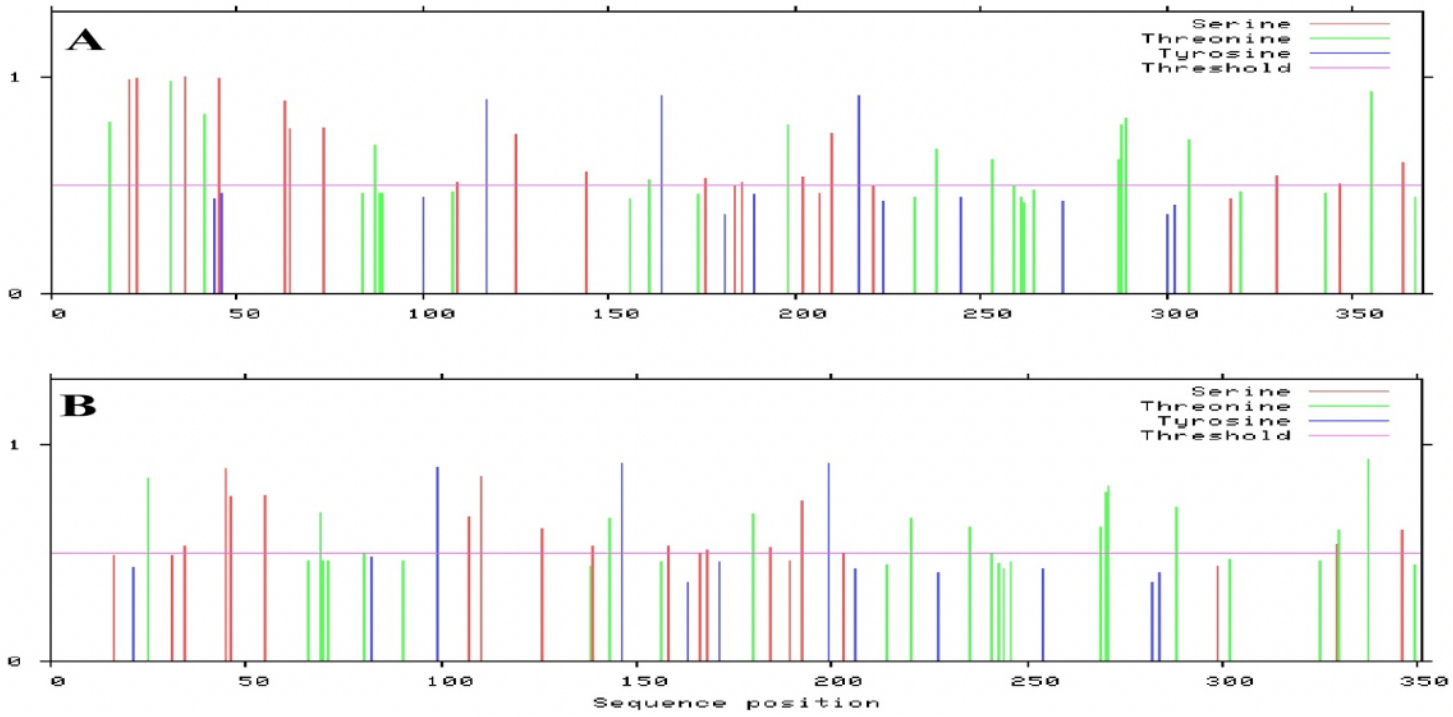

Figure 5

Phosphorylation sites of SARS-CoV-2 (5R80) and SARS (2H2Z)

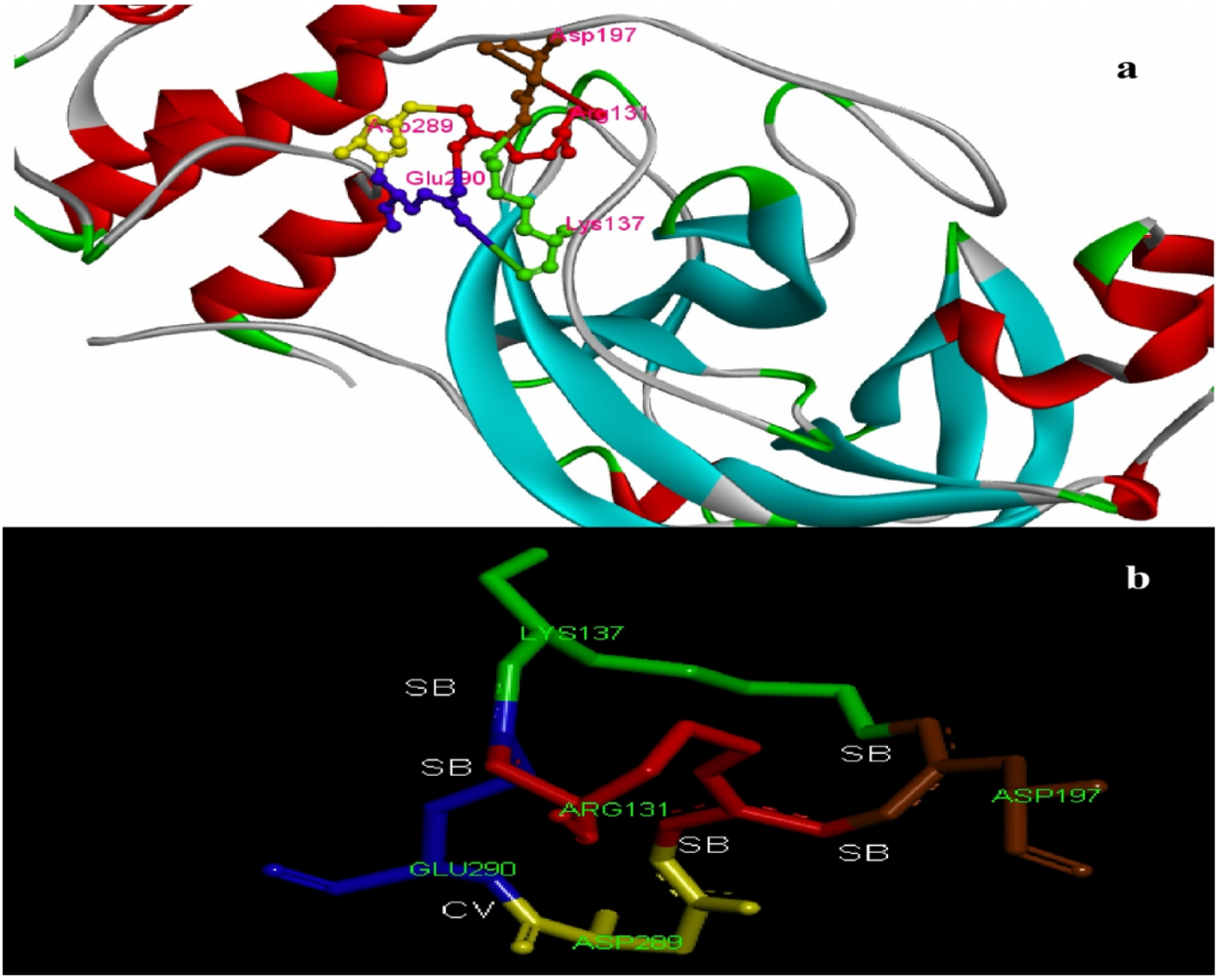

Figure 6

3D protein structure of SARS-CoV-2 with cyclic salt bridge, b. 3D view of cyclic salt bridge formation of SARS-CoV-2 with salt bridge (SB) and a covalent bond (CV) interaction

Figure 7

MSA of SARS-CoV-2 (5R80) and SARS (2H2Z) shows favorable mutation (green) and unfavorable mutation (red)