

# Comparative physicochemical and evolutionary study of SARS-CoV-2 and SARS with special reference to salt bridge and their microenvironment: A plausible explanation for divergency, stability and severity of SARS-CoV-2

Debanjan Mitra

Raiganj University

Aditya K. Pal

Raiganj University

Pradeep Kumar Das Mohapatra (✉ [pkdmvu@gmail.com](mailto:pkdmvu@gmail.com))

Raiganj University

---

## Research Article

**Keywords:** COVID-19, evolution, salt bridge, microenvironment, protein engineering

**Posted Date:** September 17th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-73762/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

The occurrence of concentrated pneumonia cases in Wuhan city, Hubei province of China was first reported on December 30, 2019. Currently, it is known as COVID 19 and now it is a nightmare for the whole world. SARS CoV first reported in 2002, but not spread worldwide. After 18 years, in 2020 it reappears and spread worldwide as SARS-CoV-2 (COVID 19), the most dangerous virus creating disease in the world. Is it possible to create a favorable evolution within this short time? If possible, then what are those properties that are changed in SARS-CoV-2 to make it undefeated? What are the basic differences between SARS-CoV-2 and SARS? This study will find all those queries. Here, all protein sequences of SARS-CoV-2 and SARS are retrieved from the database to check their physicochemical, evolutionary and structural properties. Results showed that, charged residues are playing a key role in SARS-CoV-2 evolution. SARS-CoV-2 increases its polarity by the help of charged residues, not by the polar residues. Their divergence is also strictly restricted. Induction of salt bridges with their high energies makes it very stable in any extreme conditions. Microenvironment residues also play a very crucial role in its stability. Mostly residues are favorable and contribute high energies. These microenvironment residues help in protein engineering to reduce its stability and make them weak. This comparative study will help to understand the evolution from SARS to SARS-CoV-2.

## 1. Introduction

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 [1]. 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  $\beta$ -coronavirus [2]. Disease caused by SARS-CoV-2 is known as Corona Virus Disease 2019 (COVID-19). SARS-CoV first emerged in the Guangdong province of China in 2002 and had spread into five countries infecting 8,098 people and 774 deaths having a mortality rate of 11%. After that in 2012, MERS-CoV emerged in the Arabian Peninsula and had spread 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 found in Wuhan city, Hubei province of China in December 2019. Till now there are over one crore cases of COVID-19 and over a half million deaths (mortality rate around 3%) have been reported to globally affect 218 countries. On March 11, 2020, the World Health Organization declared 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 [3]. This virus can be easily transmitted through the droplets generated at the time of coughing and sneezing by the infected people [4]. 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. [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 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 [6]. As of now, there is no approved treatment for COVID-19 but antiviral drugs such as Remdesivir, Tocilizumab are in use for treatment [7].

Generally, the  $\beta$ -coronavirus genome contains six open reading frames (ORFs), first ORFs (ORF1a/b) are in two-thirds of the whole genome and encodes 16 nonstructural proteins (nsps). There is a 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- $\beta$ ) [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 dangerous 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? To find all those queries, all 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, evolutionary and structural properties analysis. To check their stability salt bridges are also extracted with their microenvironment residues. Salt bridges are playing an important role in protein stability for operates their physiological activity in an extreme environment [19-21].

## 2. Materials And Methods

## 2.1 Dataset

A detailed analysis of those sequences and structure of SARS-CoV-2 was performed with reference to old SARS. All protein sequences of SARS-CoV-2 and SARS were retrieved from UNIPROT [22] database. Here we took 4 types of SARS-CoV-2 and SARS proteins i.e. spike proteins, membrane proteins, nucleoproteins and ORF proteins. The crystal structures of SARS-CoV-2 and SARS proteins were retrieved from the RCSB protein database (PDB) [23]. In structural comparison, we took protease which is a membrane protein.

## 2.2 Physicochemical and evolutionary properties

All protein sequences were subject to multiple sequence alignment was done for all sequences with the help of CLUSTAL Omega [24]. Both block and non-block format of the sequences were analyzed. Block of sequence was prepared by BLOCK by using ABPT [25]. Non-block and block both format were analyzed by PHYSICO2 [26] for calculation of physicochemical properties like amino acid composition, GRAVY, Zimmerman bulkiness, Grantham polarity, aliphatic index etc. The program works from UNIX like environment [27] including CYGWIN1 and is interpreted by AWK programming language [28]. It works on various forms of FASTA [29] files (Input). Mean relative abundance (MRA) was calculated from the mean value of SARS-CoV-2 proteins relative to SARS proteins to compare the data. Evolutionary properties like Maximum diverse residue (MDR), maximum conserve residue (MCR), dominant hetero pair (DHP), non conserve and conserve ratio, E value etc. were calculated by APBEST [25].

$$MRA = \frac{\text{Mean value of SARS 2 proteins} - \text{Mean value of SARS proteins}}{\text{Mean value of SARS proteins}}$$

## 2.3 Analysis of crystal structure

SARS-CoV-2 protease (4X81) and SARS protease (1PU8) were extracting from RCSB PDB [23] for structural comparison. All structured were minimized by AUTOMINv.1 [30]. Core and surface composition of those structures were identified by COSURIM [31]. Analyses of the secondary structure were done by PROPAB [32] to find the amino acid abundance in coil, helix and strand.

## 2.4 Analysis of salt bridges and their surrounding microenvironment

Salt bridges were extracted with their positions and types by SBION2 [33] and their energies were calculated by ADSBET2 [34]. Poisson-Boltzmann calculations were done with the help of APBS [35], PDB2PQR v1.9.0 [36] and CHARMM22 [37] force field. Four types of energy terms were calculated in this analysis; desolvation ( $\Delta\Delta G_{\text{dslv}}$ ), bridge ( $\Delta\Delta G_{\text{brd}}$ ), background ( $\Delta\Delta G_{\text{bac}}$ ), total net energy ( $\Delta\Delta G_{\text{net}}$ ). NUM

method [38] was used to calculate network salt bridge energies. Those salt bridge residues surrounded by charged, polar and hydrophobic residues which might affect those salt bridge energies. Those residues are indicating as microenvironment residues. It is a new concept in structural biology. Energies of microenvironments were calculated by an in house automated method [39].

## 3. Results

### 3.1 Divergence in SARS-CoV-2

To check the difference homologous positions of sequence of SARS-CoV-2 from SARS, we analyzed the BLOCK format [25] to observe the evolutionary properties like MCR, MDR, DHP and R ratio.

In case of maximum conserve ratio (MCR) (Table 1), it showed charged, uncharged polar and hydrophobic residues in SARS-CoV-2, whereas hydrophobic and neutral polar residues in SARS. In case of maximum diverse residue (MDR), it showed similar results like MCR. Dominant hetero pair also showed higher abundance of charged residues in SARS-CoV-2, but it is lacks in SARS. The R ratio (NCS:CS) showed lower value in every protein of SARS-CoV-2 than SARS.

### 3.2 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 polar uncharged residues. Amino acid compositions were calculated from the non-block format whereas block format was used to calculate disorder forming residues (Dis), Zimmerman bulkiness (ZB), aliphatic index (AI), Grantham polarity (GB) and order forming residues (Ofr) etc. 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. 1a) of charged residues (except D) in SARS-CoV-2. Polar residues showed higher abundance (except C, S) in SARS. Other proteins (membrane proteins, nucleoproteins, ORF proteins) showed almost similar results, where charged residues showed positive MRA in SARS-CoV-2 (Fig. 1c, e, g). Polar and hydrophobic residues showed negative MRA in SARS-CoV-2 which means those residues have higher abundance in SARS proteins. But when we check the polarity (Fig. 1b, d, f, h) of those proteins by Grantham polarity, it showed the high value in SARS-CoV-2 than SARS. The number of disorder forming residues (Dis) has higher abundance in SARS-CoV-2 than SARS. In the other hand, order forming residues (Orf) showed negative MRA in SARS-CoV-2. Due to the latter, Zimmerman bulkiness and aliphatic index are also higher in SARS-CoV-2

GRAVY (grand average of hydropathy) is calculated by adding the hydropathy value [40] for each residue and dividing by the length of the protein sequence. The lower value of GRAVY indicates the hydrophilic nature of SARS-CoV-2 (Table 2).

### 3.3 Charged residues on surface of SARS-CoV-2

The increasing of charged residues in SARS-CoV-2 protein sequences gives a clue that they might have effect on structure. To check, those structures of protease from SARS-CoV-2 (5R81) and SARS (1P9U) have been analyzed.

Acidic residues showed similar abundance in the core of SARS-CoV-2 and SARS, but it showed higher abundance in the surface of SARS-CoV-2 (Table 3). Polar residues also showed higher abundance in the surface of SARS-CoV-2. But hydrophobic residues showed slightly low abundance in the surface of SARS-CoV-2 than SARS.

### 3.4 Charged residues in helix of SARS-CoV-2

Charged residues showed higher abundance in the helix of SARS-CoV-2 (Table 4). SARS-CoV-2 significantly increased the stability by stabilizing its helix structure. In case of SARS, charged residues are mostly present in coil rather than the helix. Polar residues also present in high number at the coil of SARS, whereas in SARS-CoV-2 polar residues are mostly present in strand. Hydrophobic residues also have higher abundance in the helix of SARS-CoV-2 and in coil of SARS.

### 3.5 Stability of SARS-CoV-2

Salt bridges have a huge effect on protein stability [41, 42]. Charged residues are participating in the formation of salt bridges. Normally two types of salt bridges are found in proteins [43]. The increasing number of charged residues of SARS-CoV-2 indicates that charged residues might have an effect on salt bridge formation to gain more stability.

Results of salt bridges (Table 5) showed how it plays an important role in SARS-CoV-2 stability. The number of isolated salt bridges is high in SARS, but the energies of those salt bridges lower than SARS-CoV-2. In SARS-CoV-2, the salt bridge contributes -8.47 kcal/mol/SB energy, whereas it is -3.95 kcal/mol/SB energy in SARS. The numbers of network salt bridges are same in both SARS-CoV-2 and SARS. Both possess only one network salt bridge. But a huge energy difference identified between them. It showed -9.56 kcal/mol energy in SARS and -21.19 kcal/mol in SARS-CoV-2. It is very interesting that a single salt bridge contributes such high energy. Further investigation showed that the network salt bridge of SARS-CoV-2 is not a normal network salt bridge; it is a cyclic network salt bridge (R131-E290:K137-E290:R131-D197:K137-D197:R131-D289). ; We know that the desolvation ( $\Delta\Delta G_{\text{dslv}}$ ) energy is always positive and the bridge ( $\Delta\Delta G_{\text{brd}}$ ) energy is always negative. They almost nullified each other. So, the background ( $\Delta\Delta G_{\text{bac}}$ ) energy play an important role in gaining the total net energy ( $\Delta\Delta G_{\text{net}}$ ). The higher value of background energy in SARS-CoV-2 indicates that the background or surrounding environment of those salt bridges has an effect on the stability of salt bridges. So, the microenvironment of isolated and network salt bridges has been analyzed.

## 3.6 Effect of microenvironment residues in SARS-CoV-2

Due to the high number of isolated salt bridges (ISB) in SARS, they have more surrounding residues than SARS-CoV-2 (Table 6). But the total energies of the ISB microenvironment are high in SARS-CoV-2. SARS-CoV-2 has engineered its microenvironment for more stability. Thr 292 in SARS-CoV-2 contributes the highest energy i.e. -24.39 kJ/mol. In case of network salt bridge microenvironment, SARS-CoV-2 has high number of residues and energies. Most of the microenvironment residues are present in coil followed by strand and helix.

## 4. Discussion

The comparative study between SARS-CoV-2 and SARS reveals that how favorable evolution makes SARS-CoV-2 more dangerous and stronger than SARS. The R ratio (Table 1) indicates that their divergence must be restricted [44]. Although polar residues have lower abundance in SARS-CoV-2, it means polarity increased in SARS-CoV-2 due to higher abundance of charged residue, not by the polar residues (Fig. 1). Those acidic and basic residues play a major role in such alteration [39]. The higher number of disorder forming residues in SARS-CoV-2 indicates that it can easily create toxicity or disease in humans [45, 46]. High value of Zimmerman bulkiness in SARS-CoV-2 clearly indicates that they need longer heating periods in hydrolysis [47]. They can tolerate heat better than SARS. High value of aliphatic index in SARS-CoV-2 proved that SARS-CoV-2 is more thermally stable than SARS [48]. The hydrophilic nature of SARS-CoV-2 (Table 2) gives a clue that it can easily interact with water or aqueous medium and spread easily than SARS.

SARS-CoV-2 increases hydrophilic residues (Table 3) on the surface to interact easily with the aqueous medium, which helps it in stabilization [49-50]. In other hand, helix has a great effect on protein stability. Introduction of high number charged residues in the helix results in proteins more resistant to acidic environment or temperature denaturation and helps in increasing the stability [51-52].

High numbers of salt bridges contribute high energies in SARS-CoV-2 stability for extreme environments (Table 5). SARS-CoV-2 significantly increases its total salt bridges energies by lowering the desolvation cost. Formation of the cyclic salt bridge is also a great evolution with respect to protein stability. Cyclic salt bridge always has a high contribution in protein stability [53]. SARS-CoV-2 has specially designed salt bridge to gain more stability in extreme environments [19-21].

The analysis of microenvironment residues of those salt bridges revealed that they have played an important role in salt bridge energies and overall protein stability (Table 6). The polar residues and the charged residues that are not forming any salt bridge have contributed in the microenvironment to gain more stability [39]. So, the evolution of SARS-CoV-2 has a great role in its stability. It is also a clue for how to stop SARS-CoV-2 severity of infection. Protein engineering helps us in this process. Unfavorable mutation of those favorable microenvironment residues can decrease the stability of SARS-CoV-2. The concept of the microenvironment is the key factor to identify those residues.

## 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.

## Acknowledgment

We are thankful to Dr. Amal Kumar Bandyopadhyay, Department of Biotechnology, University of Burdwan for his help.

## Conflict of interest

The authors declare that they have no conflict of interest

## Funding

None

## References

1. Cascella M, Rajnik M, Cuomo A, Dulebohn SC, Di Napoli R (2020) Features, evaluation and treatment coronavirus (COVID-19). In Statpearls [internet]. StatPearls Publishing.
2. Chan JFW, To KKW, Tse H, Jin DY, Yuen KY (2013) Interspecies transmission and emergence of novel viruses: lessons from bats and birds. *Trends Microbiol.* 21(10):544-555.
3. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, Qiu Y, Wang J, Liu Y, Wei Y, Yu T (2020) Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *The Lancet*, 395:507-513.
4. Bastola A, Sah R, Rodriguez-Morales AJ, Lal BK, Jha R, Ojha HC, Shrestha B, Chu DK, Poon LL, Costello A, Morita K (2020) The first 2019 novel coronavirus case in Nepal. *Lancet Infect. Dis.* 20:279-280.
5. Kampf G, Todt D, Pfaender S, Steinmann E (2020) Persistence of coronaviruses on inanimate surfaces and their inactivation with biocidal agents. *J Hosp Infect.* 104:246-251.

6. Wan S, Yi Q, Fan S, Lv J, Zhang X, Guo L, Lang C, Xiao Q, Xiao K, Yi Z, Qiang M (2020) Characteristics of lymphocyte subsets and cytokines in peripheral blood of 123 hospitalized patients with 2019 novel coronavirus pneumonia (NCP). MedRxiv.
7. Hung IFN, Lung KC, Tso EYK, Liu R, Chung TWH, Chu MY, Ng YY, Lo J, Chan J, Tam AR and Shum HP (2020) 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*, 395:1695-1704.
8. Ziebuhr J, Snijder EJ, Gorbalenya AE (2000) Virus-encoded proteinases and proteolytic processing in the Nidovirales. *Journal of General Virology*, 81:853-879. Ikai A (1980) Thermostability and aliphatic index of globular proteins. *J Biochem.*, 88(6):1895-1898.
9. Masters PS (2006) The molecular biology of coronaviruses. *Virus. Res.* 66:193-292.
10. Yoshimoto FK (2020) The Proteins of Severe Acute Respiratory Syndrome Coronavirus-2 (SARS CoV-2 or n-COV19), the Cause of COVID-19. *Protein J.* 39:198-216.
11. Hoffmann M, Kleine-Weber H, Krüger N, Mueller MA, Drosten C, Pöhlmann S (2020) The novel coronavirus 2019 (2019-nCoV) uses the SARS-coronavirus receptor ACE2 and the cellular protease TMPRSS2 for entry into target cells. *BioRxiv*.
12. Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh CL, Abiona O, Graham BS, McLellan JS (2020) Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*, 367:1260-1263.
13. Remais J (2010) Modelling environmentally-mediated infectious diseases of humans: transmission dynamics of schistosomiasis in China. In *Modelling parasite transmission and control* (pp. 79-98). Springer, New York, NY.
14. Wu JT, Leung K, Leung GM (2020) Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *The Lancet*, 395:689-697.
15. Lipsitch M, Cohen T, Cooper B, Robins JM, Ma S, James L, Gopalakrishna G, Chew SK, Tan CC, Samore MH, Fisman D (2003) Transmission dynamics and control of severe acute respiratory syndrome. *Science*, 300:1966-1970.
16. Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Velesler D (2020) Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell*.
17. Kopecky-Bromberg SA, Martínez-Sobrido L, Frieman M, Baric RA, Palese P (2007) Severe acute respiratory syndrome coronavirus open reading frame (ORF) 3b, ORF 6, and nucleocapsid proteins function as interferon antagonists. *J Virol.* 81:548-557.
18. Ruch TR, Machamer CE (2012) The coronavirus E protein: assembly and beyond. *Viruses*, 4(3):363-382.
19. Kumar S, Nussinov R (1999) Salt bridge stability in monomeric proteins. *Mol. Biol.* 293:1241-1255.
20. Nayek A, Gupta PSS, Banerjee S, Mondal B, Bandyopadhyay AK (2014) Salt-bridge energetics in halophilic proteins. *Plos one*, 9, e93862.

21. Gallivan JP, Dougherty DA, (2000) A computational study of cation-  $\pi$  interactions vs salt bridges in aqueous media: implications for protein engineering. *Am. Chem. Soc.* 122:870-874.
22. UniProt Consortium, 2015. UniProt: a hub for protein information. *Nucleic Acids Res.* 43:204-212.
23. Rose PW, Bi C, Bluhm WF, Christie CH, Dimitropoulos D, Dutta S, Green RK, Goodsell DS, Prlić A, Quesada M, Quinn GB (2012) The RCSB Protein Data Bank: new resources for research and education. *Nucleic Acids Res.* 41:D475-D482.
24. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Syst. Biol.* 7:539.
25. Gupta PSS, Banerjee S, Islam RNU, Sur VP, Bandyopadhyay AK (2017) Substitutional Analysis of Orthologous Protein Families Using BLOCKS. *Bioinformatics*, 13:1-7.
26. Banerjee S, Gupta PSS, Nayek A, Das S, Sur VP, Seth P, Islam RNU, Bandyopadhyay AK (2015) PHYSICO2: an UNIX based standalone procedure for computation of physicochemical, window-dependent and substitution based evolutionary properties of protein sequences along with automated block preparation tool, version 2. *Bioinformatics*, 11:366-368.
27. Kernighan BW, Mashey JR (1984) Unix programming environment. In *Interact Program Environ* (pp. 175-197). McGraw-Hill Book Co.
28. Aho AV, Kernighan BW, Weinberger PJ (1987). *The AWK programming language*. Addison-Wesley Longman Publishing Co., Inc.
29. Pearson WR (1994) Using the FASTA program to search protein and DNA sequence databases. In *Computer Analysis of Sequence Data* (pp. 307-331). Humana Press.
30. Islam RNU, Mitra D, Gupta PSS, Banerjee S, Mondal B, Bandyopadhyay AK (2018) AUTOMINv1. 0: an automation for minimization of Protein Data Bank files and its usage. *Bioinformatics*, 14:525- 529.
31. Islam RN, Yasmeen S (2017) Cosurim: A program for automated computation of composition of core, rim and surface of protein data bank files and its applications.
32. Islam RNU, Roy C, Gupta PSS, Banerjee S, Mitra D, Banerjee S, Bandyopadhyay AK (2018) PROPAB: Computation of Propensities and Other Properties from Segments of 3D structure of Proteins. *Bioinformatics*, 14:190-193.
33. Gupta PSS, Nayek A, Banerjee S, Seth P, Das S, Sur VP, Roy C, Bandyopadhyay AK (2015) SBION2: Analyses of Salt Bridges from Multiple Structure Files, Version 2. *Bioinformatics*, 11:39-42.
34. Nayek A, Gupta, PSS, Banerjee S, Sur VP, Seth P, Das S, Islam RNU, Bandyopadhyay AK (2015) ADSBET2: Automated Determination of Salt-Bridge Energy-Terms version 2. *Bioinformatics*, 11:413-415.
35. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Natl. Acad. Sci.* 98:10037-10041.
36. Dolinsky TJ, Czodrowski P, Li H, Nielsen JE, Jensen JH, Klebe G, Baker NA (2007) PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular

- simulations. *Nucleic Acids Res.* 35:W522-W525.
37. Buck M, Bouguet-Bonnet S, Pastor RW, MacKerell Jr AD (2006) Importance of the CMAP correction to the CHARMM22 protein force field: dynamics of hen lysozyme. *J.* 90:L36-L38.
  38. Bandyopadhyay AK, Islam RNU, Mitra D, Banerjee S, Goswami A (2019) Stability of buried and networked salt-bridges (BNSB) in thermophilic proteins. *Bioinformation*, 15:61-67.
  39. Mitra D, Bandyopadhyay AK, Islam RNU, Banerjee S, Yasmeen S, Mohapatra PKD (2019) November. Structure, salt-bridge's energetics and microenvironments of nucleoside diphosphate kinase from halophilic, thermophilic and mesophilic microbes. In *Biotechnology and Biological Sciences: Proceedings of the 3rd International Conference of Biotechnology and Biological Sciences (BIOSPECTRUM 2019)*, August 8-10, 2019, Kolkata, India (p. 107-113). CRC Press.
  40. Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. *Mol. Biol.* 157:105-132.
  41. Jelesarov I, Karshikoff A (2009) Defining the role of salt bridges in protein stability. In *Protein Structure, Stability, and Interactions* (pp. 227-260). Humana Press.
  42. Panja AS, Maiti S, Bandyopadhyay B (2020) Protein stability governed by its structural plasticity is inferred by physicochemical factors and salt bridges. *Scientific reports*, 10:1-9.
  43. Basu S, Mukharjee D (2017) Salt-bridge networks within globular and disordered proteins: characterizing trends for designable interactions. *Mol. Model.* 23:206.
  44. Bandyopadhyay AK, Islam RNU, Mitra D, Banerjee S, Yasmeen S, Goswami A (2019) Insights from the salt bridge analysis of malate dehydrogenase from *H. salinarum* and *E. coli*. *Bioinformation*, 15:95-103.
  45. Lobanov MY, Galzitskaya OV (2015) How common is disorder? Occurrence of disordered residues in four domains of life. *J. Mol. Sci.* 16:19490-19507.
  46. Uversky VN (2019) Intrinsically disordered proteins and their "mysterious"(meta) physics. *Phys.* 7:10.
  47. Gromiha MM, Oobatake M, Kono H, Uedaira H, Sarai A (1999) Role of structural and sequence information in the prediction of protein stability changes: comparison between buried and partially buried mutations. *Protein Eng.* 12:549-555.
  48. Ikai A (1980) Thermostability and aliphatic index of globular proteins. *J Biochem.* 88:1895-1898.
  49. Strickler SS, Gribenko AV, Gribenko AV, Keiffer TR, Tomlinson J, Reihle T, Loladze VV, Makhatadze GI (2006) Protein stability and surface electrostatics: a charged relationship. *Biochemistry*, 45:2761-2766.
  50. Grimaldi J, Radhakrishna M, Kumar SK, Belfort G (2015) Stability of proteins on hydrophilic surfaces. *Langmuir*, 31:1005-1010.
  51. Facchiano AM, Colonna G, Ragone R (1998) Helix stabilizing factors and stabilization of thermophilic proteins: an X-ray based study. *Protein Eng.* 11:753-760.
  52. Villegas V, Viguera AR, Avilés FX, Serrano L (1996) Stabilization of proteins by rational design of  $\alpha$ -helix stability using helix/coil transition theory. *Fold. Desi.* 1:29-34.

53. Sindelar CV, Hendsch ZS, Tidor B (1998) Effects of salt bridges on protein structure and design. Protein Sci. 7(9):1898-1914.

## Tables

**Table 1-** Evolutionary properties i.e. Maximum diverse residue (MDR), maximum conserve residue (MCR), dominant hetero pair (DHP), non conserve and conserve ratio (R ratio) shows difference between SARS-CoV-2 and SARS

Proteins name	Source	MCR	MDR	DHP	R ratio
Spike proteins	SARS-CoV-2	L>V>N	E>L>D	EL>YF>VL	47.2
	SARS	F>L>T	T>S>V	IV>TS>YF	59.85
Membrane proteins	SARS-CoV-2	L>S>E	L>V>R	LV>YF>VA	60.53
	SARS	L>V>A	L>V>F	LV>LI>LA	67.57
Nucleoproteins	SARS-CoV-2	G>A>Q	E>D>N	QE>DN>VA	61.2
	SARS	G>R>Q	S>T>G	TS>ED>QP	72.73
ORF proteins	SARS-CoV-2	I>E>F	L>T>I	LI>YF>LE	78.2
	SARS	L>I>E	L>S>I	LI>SL>IV	30.73

**Table 2-** MRA of GRAVY (grand average of hydropathy) of SARS-CoV-2 and SARS

Proteins name	GRAVY		
	SARS 2	SARS	MRA
Spike	-0.22	-0.15	-0.47
Membrane	-0.59	0.25	-3.36
Nucleoprotein	0.26	0.57	-0.54
ORF	-0.25	0.34	-1.74

**Table 3-** Percentage of abundance of all amino acids in core and surface of SARS-CoV-2 (5R81) and SARS (1P9U)

Amino acids	5R81 (SARS-CoV-2)		1P9U (SARS)	
	core	surface	core	surface
VAL	6.9	2.0	6.4	3.7
ILE	2.3	1.0	3.3	1.0
LEU	6.6	3.0	6.7	1.7
MET	2.3	1.0	1.7	1.7
PHE	3.9	1.0	3.7	0.7
ALA	3.0	2.6	2.7	1.7
PRO	1.6	2.6	1.0	1.7
GLY	2.3	6.2	3.3	7.0
ASP	1.3	4.3	1.0	1.3
GLU	0.7	2.3	1.3	5.0
ARG	0.7	3.3	0.7	3.0
LYS	0	3.9	0.3	3.7
HIS	1.3	1.0	1.3	0.7
CYS	3.6	0.3	1.7	0.3
ASN	1.3	5.6	3.3	4.7
GLN	1.0	3.3	1.3	1.0
SER	1.6	3.6	3.3	7.0
THR	3.0	4.9	1.3	4.0
TRP	0.7	0.3	0.7	0.3
TYR	2.0	1.6	3.3	1.3

**Table 4-** Amino acid position (coil, helix, strand) in protein secondary structures of SARS-CoV-2 (5R81) and SARS (1P9U)

Position	5R81 (SARS-CoV-2)	1P9U (SARS)
COIL	F, D, E, N, R, T, P, G	F, M, C, D, E, N, Q, R, S, T, P, G
HELIX	V, A, I, L, D, E, N, Q, K, H, S, W	A, I, L, M, D, E, H, T, W, Y
STRAND	V, F, M, C, Q, K, H, S, W, Y	V, F, I, L, C, K, H, W, Y, P

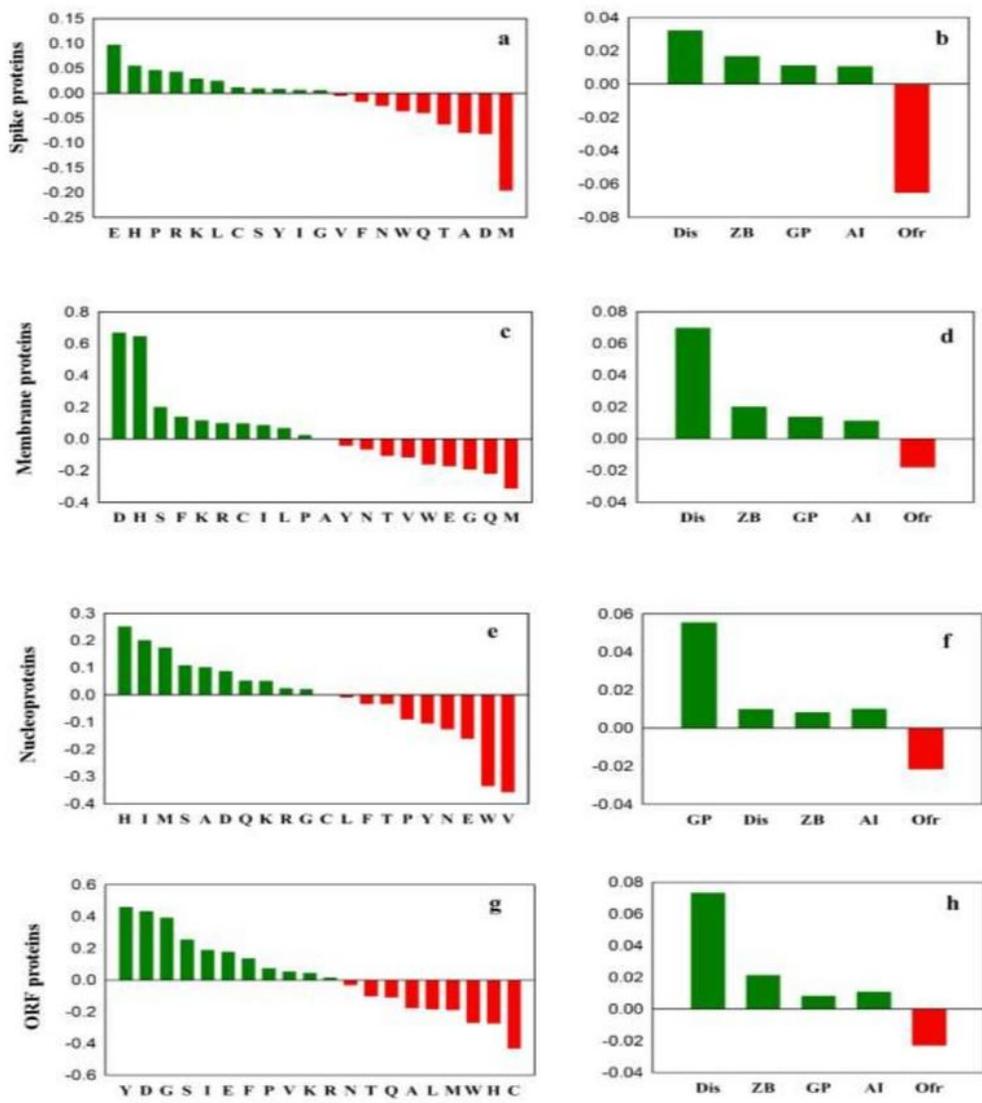
**Table 5-** Isolated and network salt bridges (SB) with their desolvation ( $\Delta\Delta G_{\text{dslv}}$ ), background ( $\Delta\Delta G_{\text{bac}}$ ), bridge ( $\Delta\Delta G_{\text{brd}}$ ) and total net ( $\Delta\Delta G_{\text{net}}$ ) energies of SARS-CoV-2 (5R81) and SARS (1P9U)

Type of salt bridge	Proteins name	Number of SB	$\Delta\Delta G_{\text{dslv}}$ (kcal/mol)	$\Delta\Delta G_{\text{bac}}$ (kcal/mol)	$\Delta\Delta G_{\text{brd}}$ (kcal/mol)	$\Delta\Delta G_{\text{net}}$ (kcal/mol)
Isolated salt bridge	SARS-CoV-2 (5R81)	8	26.23	-22.79	-71.2	-67.76
	SARS (1P9U)	10	74.34	-16.75	-97.15	-39.56
Network salt bridge	SARS-CoV-2 (5R81)	1	23.31	-2.34	-42.16	-21.19
	SARS (1P9U)	1	10.49	-0.92	-19.12	-9.56

**Table 6-** Isolated and network salt bridges microenvironment with total residue number, highest contribution by a single residue, positions and their types

Type of microenvironment	Proteins name	Number of residues	Total energies (kj/mol)	Highest contribution by a single residues	Position	Type
Isolated salt bridge	SARS-CoV-2 (5R81)	49	-90.19	Thr 292 (-24.39 kj/mol)	Coil= 46.93% Helix= 20.40 % Strand= 32.65 %	Acid= 22.44% Base= 29.57 % Polar= 36.73% Hydrophobic= 10.20%
	SARS (1P9U)	55	-76.65	Ser 110 (-22.36 kj/mol)	Coil= 44.95% Helix= 18.32 % Strand= 36.73 %	Acid= 23.63% Base= 32.72% Polar= 38.18% Hydrophobic= 5%
Network salt bridge	SARS-CoV-2 (5R81)	17	-10.15	Lys 5 (-9.61 kj/mol)	Coil= 47.05% Helix= 29.41 % Strand= 23.52 %	Acid= 17.64% Base= 35.29% Polar= 35.29% Hydrophobic= 5.88%
	SARS (1P9U)	1	-1.72	His 100 (-1.72 kj/mol)	Coil= 100% Helix= 0 % Strand=0 %	Acid= 0% Base= 100% Polar= 0% Hydrophobic= 0%

## Figures



**Figure 1**

MRA of amino acid compositions, disorder forming residues (Dis), Zimmerman bulkiness (ZB), aliphatic index (AI), Grantham polarity (GP) and order forming residues (Ofr) of spike proteins (Fig. 1a, 1b), membrane proteins (Fig. 1c, 1d), nucleoproteins (Fig. 1e, 1f), ORF proteins (Fig. 1g, 1h) from SARS-CoV-2.