Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
A hijack mechanism of Indian SARS-CoV-2 isolates for relapsing contemporary antiviral therapeutics

R. Prathiviraj, S. Saranya, M. Bharathi, P. Chellapandi*

Industrial Systems Biology Lab, Department of Bioinformatics, School of Life Sciences, Bharathidasan University, Tiruchirappalli, 620024, Tamil Nadu, India

ARTICLE INFO

Keywords:
COVID-19
SARS-CoV-2
Antivirals
Genome variation
Fold-rate
Phylogeography

ABSTRACT

Coronavirus disease (COVID-19) rapidly expands to a global pandemic and its impact on public health varies from country to country. It is caused by a new virus severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). It is imperative for relapsing current antiviral therapeutics owing to randomized genetic drift in global SARS-CoV-2 isolates. A molecular mechanism behind the emerging genomic variants is not yet understood for the prioritization of selective antivirals. The present computational study was aimed to repurpose existing antivirals for Indian SARS-CoV-2 isolates by uncovering a hijack mechanism based on structural and functional characteristics of protein variants. Forty-one protein mutations were identified in 12 Indian SARS-CoV-2 isolates by analysis of genome variations across 460 genome sequences obtained from 30 geographic sites in India. Two unique mutations such as W6152R and N5928H found in exonuclease of Surat (GBRC275b) and Gandhinagar (GBRC239) isolates. We report for the first time the impact of folding rate on stabilizing/retaining a sequence-structure-function-virulence link of emerging protein variants leading to accommodate hijack ability from current antivirals. Binding affinity analysis revealed the effect of point mutations on virus infectivity and the drug-escaping efficiency of Indian isolates. Emodin and artinemol suggested herein as repurposable antivirals for the treatment of COVID-19 patients infected with Indian isolates. Our study concludes that a protein folding rate is a key structural and evolutionary determinant to enhance the receptor-binding specificity and ensure hijack ability from the prevalent antiviral therapeutics.

1. Introduction

Coronavirus disease (COVID-19) is a global epidemic caused by a new virus severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The World Health Organization (WHO) declared this outbreak a global health emergency on March 11, 2020 [1]). COVID-19 is affecting 213 countries and territories around the world and 2 international conveyances. As of February 10, 2021, 107 M confirmed, and 2.33 M death cases have been reported worldwide. India is the 2nd most-affected 20 countries with COVID-19 with 8.41 M confirmed cases of which 125 K people have succumbed to the infection (https://www.worldometers.info/coronavirus/). The SARS-CoV-2 outbreak spread rapidly worldwide by human-to-human contact [2]. No drugs or vaccines have been approved by the U.S. Food and Drug Administration (FDA) for the prevention or treatment of COVID-19. The FDA authorized Remdesivir as an emergency antiviral for the treatment of severe COVID-19 patients on May 1, 2020 [3]. The WHO has suggested randomized standard care or active treatment (remdesivir, chloroquine/hydroxychloroquine plus azithromycin, lopinavir/ritonavir, or lopinavir/ritonavir plus interferon-β1a) to confirmed cases of COVID-19 [4]). Several systematic investigations on target-based antiviral discovery have been in progress for the development of effective therapies for COVID-19 infections [5].

SARS-CoV-2 is an enveloped single-positive-stranded RNA virus [1]. Have completed the first genome sequence (accession: NC_045512.2; Wuhan-Hu-1) of SARS-CoV-2 isolated from Wuhan, Hubei province, China. This reference genome consists of 29.9 Kb organized into specific genes encoding structural (spike, envelope, membrane, and nucleocapsid) and non-structural proteins (Nsp1-16). Spike glycoprotein plays an important role in viral infectivity and disease transmission by interacting with the human ACE-2 receptor [6]). Papain-like protease (PLPro), 3-chymotrypsin-like protease (3CLPro), RNA-dependent RNA polymerase (RdRp), helicase, exoribonuclease, and 2’-O-methyltransferase are promising therapeutic targets for antivirals development [5]. Besides, ORF3a, ORF6, ORF7a ORF8, and ORF10 genes are accessory proteins essential for virus assembly and infection [7]).
Enormous genome sequences of SARS-CoV-2 are available to the research community. Genomic structure and variations of SARS-CoV-2 are pro-vital for studying virulence, infectivity, and COVID-19 outbreak. SARS-CoV-2 isolates are distributed non-geographically with the prevalence of single nucleotide mutations as the major genomic variations. Genetic variant analyses revealed some subtypes, deleterious, and frequent mutation events in SARS-CoV-2 isolates [8–11]. Therefore, the emergence of new mutations in SARS-CoV-2 may force the subsequent establishment of a new escape or hijack mechanism for antivirals.

The viral hijack mechanism is an emphatic one to escape the host immune systems and antiviral therapies. Human immunodeficiency virus 1 (HIV-1) hijacks its genomic and structural perspective for survival advantage in a host [12], cell-to-cell transmission [13], and membrane traffic to promote infection [14]. A host cellular pathway is hijacked by SARS-CoV for supplying replication membranes for viral protein synthesis [15,16]. HIV-1 virion infectivity factor is found as a molecular determinant to escape from antiviral APOBEC3 (A3) proteins [17]. There is no report on the hijack ability of SARS-CoV-2 from antivirals.

The sequence-structure-function-conformational dynamic space-folding rate-virulence link revealed as a molecular determinant of some virulence proteins [18–27]. Accordingly, we attempted to know how SARS-CoV-2 protein variants hijack or escape from current antivirals and to increase host-specific infectivity in phylogenetic populations. A protein thermodynamic stability accounts for a large fraction of observed mutational effects in RNA viruses. Destabilizing mutations inflict a fitness penalty proportional to the reduction in folded protein [28,29]. A biophysical model developed from this hypothesis was used to study RNA viruses escaping antibody stress [30].

The folding kinetics of a protein depends on the reduced hydrophobic density and enhanced polar-polar repulsive interactions [31]. The folding rates of spike glycoprotein and membrane protein are essential in the formation of the viral envelope of coronaviruses. At the specific folding rates, both proteins are accumulated at the site of virus assembly (pre-Golgi membranes) [32]. Thus, viruses follow different blueprints of interference with protein folding to ensure their intracellular survival like bacteria [33]. Therefore, our study was intended to propose a hijack mechanism of Indian SARS-CoV-2 isolates for relapsing current antivirals by computing structural and functional determinants of identified protein variants. This study reports significant protein mutations from these isolates and stratifies them structurally. Our findings on selective variants would be useful for systematic repurposing of antivirals on a regional basis to mitigate the spreading of COVID-19.

2. Materials and methods

2.1. Dataset

In this study, a complete SARS-CoV-2 nucleotide sequence (accession: NC_045512.2; Wuhan-Hu-1 isolate) was downloaded as a reference genome from the National Centre for Biotechnology Information (NCBI)-SARS-CoV-2 resources (https://www.ncbi.nlm.nih.gov/sars-cov-2/). This resource consists of 59, 482 Sequence Read Archive runs, and 24,380 nucleotide records by 24, September 2020. The Sequence Read Archive runs were neglected in the phylogenomic study as poorly annotated coding regions from raw sequence data, the impact of recombination, and incomplete lineage sorting [34]. A total of 16, 557 SARS-CoV-2 nucleotide sequences (sequence length ranged from 29600 to 31000 bp) isolated from different countries were used in this study. Among six geographic origins, 11, 579 sequences isolated from Northeast and South America, 1424 from Oceania, 1571 from Asia, 460 from Europe, and 189 from Africa. The 460 genomes were sequenced in India of which 361 genomes sequenced from 277 confirmed COVID-19 cases (132 female and 229 male) by Gujarat Biotechnology Research Centre, Gandhi Nagar, India. The two genomes were deposited from Kerala State, three genomes from Una, Himachal Pradesh (Una), and one genome from Hyderabad, Telangana State. The number of ambiguous nucleotides was < 1% in the trimmed sequences. The dataset (accession, isolate name, country, region, collection date) used for the tree building is available in Supplementary file 1. The computational process for proposing a hijack mechanism of Indian SARS-CoV-2 isolates for relapsing current antiviral therapeutics is represented in Fig. 1.

2.2. Phylogeographic tree construction and analysis

We performed multiple sequence alignment of all SARS-CoV-2 sequences to a reference genome using MAFFT v7.0 (Multiple Alignment using Fast Fourier Transform) with truncating the bases in the 5′-UTR and 3′-UTR without losing key sites [35]. After alignment, the tree was generated for entire isolates based nucleotide variations with the NCBI-SARS-CoV-2 resources and Virus Pathogen Database and Analysis Resource (ViPR) [36]. The phylogenomic tree was constructed using the Neighbor-Joining algorithm [37] with the Jukes-Cantor evolutionary model [38] and 1000 bootstraps. The identical nodes were collapsed and then the tree was visualized by Interactive Tree Of Life (ITOL) v5.0 [39]). In this study, 95 SARS-CoV-2 sequences from different coronavirus hotspot countries and 177 sequences from Indian origin were selected and then aligned to a reference genome using MAFFT. The sampling trees were generated with the Bayesian Evolutionary Analysis by Sampling Tree (BEAST) v2.5 [40] with 1000 bootstrapping. Each tree was weighted proportional to its posterior probability. Evolutionary distances were computed between pairs of sequences based on mutational substitutions. Finally, phylogenomic tree nodes were annotated based on geographic regions using the ITOL tool.

2.3. Analysis of genomic variations

We computed the extent of variation at each position of nucleotide sequences uploaded to the ViPR genome variation analysis tool and visualized aligned nucleotide sequences using the ViPR interactive alignment viewer. Aligned sequences generated from the ViPR tool were further viewed in the BioEDIT v7.2 software [41] for the detection of mutation patterns compared to a reference genome. All nucleotide variants identified in the coding regions converted to encoded amino acid residues. The sequence similarity comparison was carried to find out protein variants from a reference genome by searching identical residues against a reference genome using the NCBI BLAST tool [42]. The optimal global protein sequence alignment was conducted between a reference genome and selected Indian SARS-CoV-2 isolates by the EMBOSS Needle tool (https://www.ebi.ac.uk/Tools/psa/emboss_needle). To identify cross conservation and to understand the nature of amino acid changes, we used the MicroGMT software, a mutation tracker for SARS-CoV-2 [43].

2.4. Structural and functional characterization of protein variants

The structural and functional characteristics of protein variants identified from selected Indian SARS-CoV-2 isolates were predicted by various bioinformatics tools. All predicted characteristics were compared with proteins from a reference genome (Wuhan-Hu-1). The structural flexibility deviation of protein variants from reference proteins was analyzed based on a library composed of structural prototypes using the PredyFlexy server [44]. The folding rate of identified protein variants was computed from the sequences with the FOLD-RATE based on a set of 49 diverse amino acid properties [45]. The average amino acid property (Pave) was calculated using the following equation.

\[ P(\text{ave}) = \frac{1}{N} \sum_{j=1}^{N} P(j) \]

where \( P(j) \) is the property value of jth residue and the summation is over
the total number of residues in a protein ($N$). Multiple regression techniques were used to correlate predicted folding rate related to experimental folding rate $\ln k_f(i)$. The protein folding kinetic order and folding rate of the variants were calculated from their 3D structures using K-FOLD based on a support vector machine [46]. The protein variants retaining any virulence property, information molecule, and involvement in the cellular metabolic process were predicted from amino acid and dipeptide composition. It was conducted with the VICMpred server using Matthew’s correlation coefficient [47].

2.5. Characterization of the binding specificity of protein variants

The FDA-approved antiviral drugs were retrieved from the DrugBank v5.0 [4,48]. A cryo-electron microscopic structure of SARS-CoV-2 binding human receptor ACE-2-B^AT1 (PDB ID: 6M17) was downloaded from the Protein Data Bank (PDB). The crystal structures of human neutralizing antibodies (anti-SARS-CoV-2 antibodies) CC12.1, CR3022 (PDB ID: 6XC3), m396 (PDB ID: 2G75), and 80R (PDB ID: 2GHW) were obtained from the PDB [49]. The 3-D structures of the proteins from both Wuhan-Hu-1 and selected Indian isolates were built by the
SWIS-MODEL [50] and I-TASSER [51]. The structural constraints in the modeled proteins were energetically minimized by the refineD server using machine learning-based restrained relaxation [52].

A protein-ligand docking was carried for the prediction of the binding specificity of antiviral drugs to the protein variants identified from Indian isolates using the PatchDock server [36]. The PatchDock algorithm is based on hot-spot rich surface patches matching and initial transformation with geometric hashing and poses clustering. The shape complementarity, interface shape, and sizes of the molecules determine the quality of docked models [36]). A rigid body protein-protein docking was conducted with clustering RMSD 4.0 for the prediction of the binding specificity of Spike protein to ACE-2-B01 and anti-SARS-CoV-2 vaccines by the PatchDock, followed by the FireDock [53]. Binding energies, hydrogen bonding, attractive and repulsive van der Waals were calculated from the resulted protein-protein complexes by optimizing side-chain conformations and rigid-body orientations. An iterative knowledge-based scoring function was used to calculate the ligand RMSD and docking score of protein-protein complexes by the HDOCK [6,54]. The scoring functions of all docking models were evaluated by the AutoDock Vina v1.1.2 software [55]. The docked models were visualized in the PyMOL v2.4 software [56]. The hydrogen and hydrophobic interactions in the drug-protein and protein-protein complexes were calculated and visualized using the LigPlot+ v2.2 software [57].

3. Results

3.1. Phylogenomic analysis

To identify the genetic variations that are potentially linked to the severity of COVID-19, we constructed a large-scale phylogenetic tree with 16,536 SARS-CoV-2 completely annotated genome sequences worldwide (Fig. S1). The majority of SARS-CoV-2 isolates from the USA and China were clustered within the country and some of them scattered across the world. As shown in Fig. 2A, the first Wuhan-Hu-1 isolate was closely clustered with isolates from Vietnam (VNM/nCoV-19-01S) and USA (USA/CA_2602, USA/NH_0004). This cluster is grouped with isolates from India (IND/Kerala/29), Pakistan (PAK/Manga1), and Nepal.
The second isolate (IND/Kerala/166) of India forms a separate sub-clade with isolates from China (CHN/Yunnan-01/2020, CHN/KMS1) and the USA (USA-AZ1). Our analysis suggests that many Gujarat isolates are genetically more related to the isolates from Israel, Columbia, and Malaysia. The results of this study indicate a phylogenetic disparity across the Indian isolates.

3.2. Phylogenetic analysis of Indian SARS-CoV-2 isolates

Besides, we constructed a phylogenetic tree with 3 clades to know the genome sequence similarity and dissimilarity across Indian SARS-CoV-2 isolates (Fig. 2B, Fig. S2). The isolates with identical sequence coverage were removed from the tree. As shown by our analysis, SARS-CoV-2 isolates obtained from Surat, Ahmadabad, and Vadodara have genetically diverged from each other. The first SARS-CoV-2 (IND/29) isolate had only 5 genomic variants compared to the Wuhan-Hu-1 isolate genetically diverged from each other. The first SARS-CoV-2 (IND/29) isolate was isolated from a patient of Kerala state on March 29, 2020 [58]. This isolate had only 5 genomic variants compared to the Wuhan-Hu-1 isolate. It has a distinct genetic variation with Kerala isolate had only 5 genomic variants compared to the Wuhan-Hu-1 isolate. The isolates with identical sequence divergence with multiple substitutions compared to the Gujarat isolates. Himachal Pradesh (GBRC279b, GBRC298) isolates are clustered with Jammagarr (GBRC1108, 106) and Kodinar (GBRC100) isolates whereas Kerala (IND/29) related to Surat (GBRC272a), Ahmadabad (GBRC157a) and Vadodara (GBRC125, 130) isolates. Telangana (GCMDK318) isolate is also grouped with Surat isolates. Phylogenetic studies of Indian SARS-CoV-2 isolates revealed that Surat isolates are the important ones to disperse genetic diversity leading to an increased virus transmission rate in India.

3.3. Analysis of genomic variations

Genome sequences of Indian SARS-CoV-2 isolates were compared with a reference genome (Wuhan-Hu-1) for the identification of gene mutation events (Supplementary file 2). A total of 291 gene mutation events were detected in 26 Indian SARS-CoV-2 isolates. Results suggested that the isolates obtained from Savali (GBRC225a), Surat (GBRC275b), and Chuda (GBRC250b) had more than 15 mutations compared to others. The number of nucleotide differences varies from a minimum of 5 nucleotides to a maximum of 18 nucleotides among Indian isolates. As shown by our analysis, the C-to T exchange has the most predominant numbers in all nucleotide change directions in Indian SARS-CoV-2 genomes. Generics variations C199T, A2250C, C833T, C2794T, C2995T, G4258T, C11041T, C14366T, C14529T, C18526T, C18835T, C19112T, G21682T, C22402T, A23361G, G25521T, C26693T, C28812T, C2995T, G4258T, C11041T, C14366T, C14529T, C18526T, C18835T, C19112T, G21682T, C22402T, A23361G, G25521T, C26693T, C28812T and C29708T identified from Indian isolates introducing mutations in corresponding proteins.

3.4. Analysis of protein variants

By our protein mutation analysis, 12 hotspot regions were identified in India of which 7 from Gujarat state and 5 from other Indian regions (Fig. 2C). We identified the above 9 protein mutations from Himachal Pradesh (GBRC279b, GBRC298) and Surat (GBRC275b) isolates. The mutation events frequently occur in spike protein, RdRp, Nsp2, and PLpro. A total of 171 D614G mutations are identified from 460 Indian isolates in which 59 mutations in Surat isolates. D614G is a frequent mutation in Surat isolates. D614G mutations W6152R and N5928H in exonuclease of Surat isolate (GBRC275b) and Gandhinagar isolate (GBRC239), respectively. There are no genetic variations in the Nsp7, Nsp8, Nsp9, Nsp11, endRNAs, orf6, and orf7a of Indian SARS-CoV-2 isolates. It suggests that these genes are highly conserved and identical to the Wuhan-Hu-1 isolate.

3.5. Impact of folding rate on mutated proteins

The folding rate is a key measure to determine the protein multiplication, virulence, host specificity, and infectivity of a pathogen. The influence of folding rates of different structural classes of mutated proteins was studied herein (Table 1). It indicates the kinetic folding rates of all mutated proteins are shown slightly faster than reference proteins. We predicted a fast-folding rate in all β-classes followed by all α-classes compared to reference proteins. It suggests that there are no drastic changes in secondary structural elements in the mutated proteins. A slow folding rate is predicted in mixed structural classes of mutated proteins. The folding rates are influenced by the unknown classes of mutated proteins, particularly in coil and turns (Fig. S3). It indicates the structure-function stabilization points to be found in such folding regions. We also analyzed the folding rate of mutated proteins from Indian isolates to know the effect on the structure-function-virulence link. The fast folding rates are predicted in the Orf1ab protein of Telangana and Kerala isolates and spike, PLPro, Nsp2, Orf8, and RdRp of many Indian isolates. A slow folding rate is noted in all structural classes of helicase from Kerala isolates, α-class of spike protein from Surat (GBRC275b) isolate, and nucleocapsid protein of many isolates. There is a drastic change in the folding rate of exonuclease of Indian isolates as the results of W6152R and N5928H mutations. The results of this study suggest the folding rate can determine the structure-function relationships of mutated proteins, which in turn increase viral load inside hosts via a fitted post-translation process.

3.6. Impact of virulence property of mutated proteins

A virulence property of a protein can be determined by its amino acid composition and physicochemical characteristics of substituted amino acids. Our results show that selected Indian isolates have no major involvement in information processing but some isolates contributed to the metabolic process (Table 2). The virulence properties are predicted in PLpro from Kerala (IND/29), Gacihal Pradesh (GBRC98) isolates, and 3CLpro from Himatnagar (GBRC242) isolate. Himachal Pradesh (GBRC297b), Surat (GBRC275b), and Gandhinagar (GBRC239) isolates have lost virulence but Wuhan-Hu-1 isolates not retained so. Himachal Pradesh (GBRC297b) isolate had virulence property in the helicase. Interestingly, we found a new functional shift from virulence to metabolic process upon T5538I mutation in helicase of Kerala (IND/166) isolate, suggested that its survivability and adaptability in host systems.

3.7. Impact of binding specificity to the FDA-approved antiviral drugs

Antiviral drugs were docked with respective SARS-CoV-2 proteins and their binding affinity of them calculated to investigate drug escaping mechanisms (Table 3). Molecular docking studies show that antiviral drugs have low binding energies to many Indian isolates compared to Wuhan-Hu-1. It points out that current antivirals could not interact with proteins from Indian isolates having deleterious mutation events. Remdesivir has shown a low binding specificity to the RdRp of SARS-CoV-2 isolates reported in India, suggestive of relapsing current antiviral treatment. Emodin is a proficient antiviral targeting spike protein, which shows more binding specificity to Indian isolates except for the D614G variant. The drug specificity is predicted to decrease in Surat isolate (GBRC275b) due to the newly identified mutation (W6152R) in exonuclease. However, there is no radical change in the drug specificity of a new mutation (N5928H) in Gandhinagar isolate (GBRC239) except ribavirin. Artenimol also shows a more binding affinity to helicase in Indian isolates. Remdesivir, ribavirin, and azacitidine are no longer interactions with respective targets whereas emodin and arteminol are suggested as suitable antivirals for the treatment of patients infected with these isolates.
Table 1

Comparative analysis of folding rate and folding mechanism of protein variants identified from Indian SARS-CoV-2 isolates.

| Genotype | Isolate | Protein | Mutation | Folding classes | All-α | All-β | α+β/αβ | Unknown |
|----------|---------|---------|----------|----------------|-------|-------|---------|---------|
|          |         |         |          | Ref. Var. | Ref. Var. | Ref. Var. | Ref. Var. | Ref. Var. |
| IND/29 (Kerala) | Orf1ab | Nsp2 | I671T | 2.3 | 6.1 | 0.5 | 1.7 | 6.4 | 0.1 | 1.7 |
|            | PLPro | P2144S | 6.2 | 6.9 | 1.9 | 1.7 | 6.4 | 0.1 | 1.7 |
|            | RdRp | A4798V | 13.9 | 13.6 | 21.7 | 20.6 | 13.3 | 5.2 | 1.7 |
|            | Spike | R407I | 3.4 | 4.76 | 21.5 | 23.7 | 11 | 9.1 | 1.7 |
| IND/166 (Kerala) | Orf1ab | Nsp2 | I476V | 11.4 | 13 | 21.3 | 15.8 | 28.7 | 17.1 | 6.9 | 0.1 | 1.7 |
|            | Helicase | T5538I | 3.56 | 3.84 | 17.8 | 16.4 | 3.33 | 2.1 | 1.7 |
|            | Spike | A930V | 3.4 | 4.38 | 21.5 | 20.4 | 11 | 5.9 | 1.7 |
|            | Orf8 | L84S | 0.827 | 0.037 | 42.2 | 42.6 | 44.6 | 14.4 | 5.9 | 1.7 |
| GMCKN318 (Telangana: Hyderabad) | Orf1ab | Nsp2 | L204F | 11.4 | 12 | 12.3 | 12.3 | 12.3 | 17.4 | 11.2 | 1.7 | 1.7 |
|            | PLPro | K4489V | 13.9 | 13.6 | 21.7 | 20.6 | 13.3 | 5.2 | 1.7 |
|            | RdRp | A1812D | 6.26 | 6.67 | 13.9 | 12.1 | 11 | 5.9 | 1.7 |
| GBRC242 (Gujarat: Himatnagar) | Orf1ab | Nsp6 | L3606F | 6.6 | 7.5 | 63.8 | 66.9 | 17.9 | 3.5 | 1.7 |
|            | Spike | G519S | 11.4 | 12.3 | 12.3 | 17.4 | 11 | 5.9 | 1.7 |
| GBRC97b (Himachal Pradesh: Una) | Orf1ab | Nsp2 | L3606F | 6.6 | 7.5 | 63.8 | 66.9 | 17.9 | 3.5 | 1.7 |
|            | PLPro | L204F | 11.4 | 12 | 12.3 | 12.3 | 12.3 | 17.4 | 11.2 | 1.7 | 1.7 |
|            | RdRp | A1812D | 6.26 | 6.67 | 13.9 | 12.1 | 11 | 5.9 | 1.7 |
| GBRC273b (Gujarat: Surat) | Orf1ab | Nsp6 | L3606F | 6.6 | 7.5 | 63.8 | 66.9 | 17.9 | 3.5 | 1.7 |
|            | Spike | G519S | 11.4 | 12 | 12.3 | 12.3 | 12.3 | 17.4 | 11.2 | 1.7 | 1.7 |
| GBRC98 (Himachal Pradesh: Una) | Orf1ab | Nsp2 | L3606F | 6.6 | 7.5 | 63.8 | 66.9 | 17.9 | 3.5 | 1.7 |
|            | PLPro | K4489V | 13.9 | 13.6 | 21.7 | 20.6 | 13.3 | 5.2 | 1.7 |
|            | RdRp | A1812D | 6.26 | 6.67 | 13.9 | 12.1 | 11 | 5.9 | 1.7 |
| GBRC275b (Gujarat: Surat) | Orf1ab | Nsp6 | L3606F | 6.6 | 7.5 | 63.8 | 66.9 | 17.9 | 3.5 | 1.7 |
|            | Spike | G519S | 11.4 | 12 | 12.3 | 12.3 | 12.3 | 17.4 | 11.2 | 1.7 | 1.7 |

The multi-state kinetic folding mechanism was predicted for both reference and variant proteins.

3.8. Impact of binding specificity to human receptor

The binding specificity of Indian SARS-CoV-2 isolates was studied by predicting molecular interactions between spike protein and human ACE-2 receptor; spike protein and human ACE-2-BβAT complex (Fig. 3; Fig. S3; Tables S1 and S2). Surat (GBRC275b) isolate is found to be effective but its interactions with this complex are predicted to be lower. The binding affinity of Wuhan-Hu-1 with ACE-2 is somewhat similar to Indian isolates and some variants show a reasonable binding affinity to ACE-2. Spike protein from Kerala (IND/29) isolate has a more binding affinity to ACE-2-BβAT complex than that of Kerala (IND/166) isolate. The results of ligand RMSD indicates the binding poses of Telangana and Himachal Pradesh (GBRC97b) isolates are more favorable to ACE-2 and ACE-2-B AT1 complex. Telangana and Himachal Pradesh (GBRC97b) isolates have shown a considerable binding affinity to ACE-2, reflecting these isolates could increase the host specificity and infectivity.

3.9. Impact of binding specificity to WHO-approved anti-SARS-CoV-2 vaccines

The crystal structures of human neutralizing antibodies were used to dock with spike proteins of Indian isolates for the prediction of their binding specificities to relapse current anti-SARS-CoV-2 vaccines. The binding poses observed for the spike protein of these isolates have been shown to destabilize the H-bonding patterns to the CR3022 antibody to Kerala (IND/166) and Telangana isolates. The binding specificity of the CR3022 antibody to Kerala (IND/166) and Telangana isolates are relatively high to that of Wuhan-Hu-1. CC12.1 shows to have more interactions to Surat (GBRC275b) and Kerala isolates whereas CC12.1 and
4. Discussion

RNA viruses evolve rapidly because their genomes can mutate one million times higher than the host genome. The genetic diversity of RNA viruses is determined by selective pressure on viral and host-dependent processes [59]. Accordingly, a phylogenetic disparity was found across the Gujarat and Kerala isolates. The individuals from Kerala states have acquired SARS-CoV-2 at least from two sources [58]. Our study revealed a phylogenetic discrepancy between two Kerala isolates as both isolates have typical mutations in spike (R407I/A930V; Nsp2: I671T/I476V), PLPro, RdRp, helicase, and Orf8. The mutational events might have reflected on the genetic diversity of Kerala isolates. Similarly, we found a phylogenetic discrepancy within Surat and Himachal Pradesh isolates. Jazz at the start of the year 2021, by which time the number of SARS-CoV-2 cases had peaked, suggests current anti-SARS-CoV-2 vaccines are still effective against viruses from different parts of the world and within other hotspots in India.
| Genotype | Isolate | Gene | Mutant | Drug name | DrugBank/PMID | Score | Area | ACE |
|----------|---------|------|--------|-----------|---------------|-------|------|-----|
| 1        | IND/29 (Kerala) | Nsp2 | I671T  | Ribostamycin sulfate | DB03615 31443266 5334 5334 687.7 687.7 | 72.19 72.19 |
|          |         |      |        | Emodin  | DB06090 18678486 6956 6956 944.4 944.4 | 306.29 306.29 |
|          | PIPro   |      |        | Simeprevir | DB02145 24326 4606 537.8 | 286.28 286.28 |
|          | RdRp    |      |        | Remdesivir | DB14761 32020029 6472 6460 858.6 | 297.68 298.62 |
|          |         |      |        | Hydroxychloroquine | DB16111 32020029 5546 5602 642.2 | 200.07 34.74 |
|          | Spike   |      |        | Emodin | DB07715 32194980 4022 4242 518.4 | 159.29 169.22 |
| 2        | IND/166 (Kerala) | Nsp2 | I476V  | Ribostamycin sulfate | DB06315 31443266 5334 5334 687.7 687.7 | 72.19 72.19 |
|          |         |      |        | Emodin | DB06090 18678486 6956 6956 944.4 944.4 | 306.29 306.29 |
|          | Spike   |      |        | Emodin | DB07715 32194980 4022 4242 518.4 | 159.29 169.22 |
| 3        | GMCKN318 (Telangana: Hyderabad) | Spike | AA2244 | Artenimol | DB11638 32696720 4102 4190 515.9 | 250.26 77.86 |
|          |         |      |        | Ethanesulfonate | DB01196 31443266 5334 5334 687.7 687.7 | 291.20 72.29 |
|          |         |      |        | Hydroxychloroquine | DB01611 32020029 5546 5646 642.2 | 200.07 198.71 |
|          | Spike   |      |        | Emodin | DB07715 32194980 4022 4220 507.4 | 532.9 157.1 |
| 4        | GBRC97b (Himachal Pradesh: Una) | Nsp2 | L204F  | Ribostamycin sulfate | DB03615 31443266 5334 5334 687.7 687.7 | 72.19 72.19 |
|          |         |      |        | Emodin | DB06090 18678486 6956 6956 944.4 944.4 | 306.29 306.29 |
|          | Spike   |      |        | Emodin | DB07715 32194980 4022 4242 518.4 | 159.29 150.58 |
| 5        | GBRC98 (Himachal Pradesh: Una) | Nsp2 | G519S  | Ribostamycin sulfate | DB03615 31443266 5334 5334 687.7 687.7 | 72.19 72.19 |
|          |         |      |        | Emodin | DB06090 18678486 6956 6956 944.4 944.4 | 306.29 306.29 |
|          | Spike   |      |        | Emodin | DB07715 32194980 4022 4242 518.4 | 159.29 150.58 |
| 6        | GBRC275b (Gujarat: Surat) | Exonuclease | E6126R | Ribostamycin sulfate | DB03615 31443266 5334 5334 687.7 687.7 | 72.19 72.19 |
|          |         |      |        | Emodin | DB06090 18678486 6956 6956 944.4 944.4 | 306.29 306.29 |
|          | Spike   |      |        | Emodin | DB07715 32194980 4022 4242 518.4 | 159.29 150.58 |
| 7        | GBRC258b (Gujarat: Gondal) | Nsp6 | L360F  | Loratadine | DB04455 25318072 3896 3896 518.3 | -216.58 216.58 |
|          |         |      |        | Hydroxychloroquine | DB16111 32020029 5546 5534 642.2 | 200.07 29.61 |
|          | Spike   |      |        | Emodin | DB07715 32194980 4022 4242 518.4 | 159.29 150.90 |
| 8        | GBRC245 (Gujarat: Modasa) | Nsp2 | S2015R | Simeprevir | DB03615 31443266 5334 5334 687.7 687.7 | 72.19 72.19 |
|          |         |      |        | Emodin | DB06090 18678486 6956 6956 944.4 944.4 | 306.29 306.29 |
|          | Spike   |      |        | Emodin | DB07715 32194980 4022 4242 518.4 | 159.29 150.90 |
| 9        | GBRC273b (Gujarat: Choryasi) | Nsp6 | S2015R | Simeprevir | DB03615 31443266 5334 5334 687.7 687.7 | 72.19 72.19 |
|          |         |      |        | Emodin | DB06090 18678486 6956 6956 944.4 944.4 | 306.29 306.29 |
|          | Spike   |      |        | Emodin | DB07715 32194980 4022 4242 518.4 | 159.29 150.90 |
| 10       | GBRC242 (Gujarat: Himatnagar) | Nsp6 | S2015R | Simeprevir | DB03615 31443266 5334 5334 687.7 687.7 | 72.19 72.19 |
|          |         |      |        | Emodin | DB06090 18678486 6956 6956 944.4 944.4 | 306.29 306.29 |
|          | Spike   |      |        | Emodin | DB07715 32194980 4022 4242 518.4 | 159.29 150.90 |
| 11       | GBRC239 (Gujarat: Gandhinagar) | Nsp6 | S2015R | Simeprevir | DB03615 31443266 5334 5334 687.7 687.7 | 72.19 72.19 |
|          |         |      |        | Emodin | DB06090 18678486 6956 6956 944.4 944.4 | 306.29 306.29 |
|          | Spike   |      |        | Emodin | DB07715 32194980 4022 4242 518.4 | 159.29 150.90 |
| 12       | GBRC244 (Gujarat: Ahmadabad) | Nsp6 | S2015R | Simeprevir | DB03615 31443266 5334 5334 687.7 687.7 | 72.19 72.19 |
|          |         |      |        | Emodin | DB06090 18678486 6956 6956 944.4 944.4 | 306.29 306.29 |
|          | Spike   |      |        | Emodin | DB07715 32194980 4022 4242 518.4 | 159.29 150.90 |
isolates have two unique mutations W6152R and N5928H in exonuclease, respectively. Mutations in this enzyme can confer resistance to ribavirin and other nucleotide analogs via an error-correcting mechanism [63]). This enzyme also interacts with RdRp and associated Nsp7 and Nsp8 to perform proofreading activity [64]. It is a particularly challenging task to current antivirals targeting not only exonuclease and also RdRp in Surat and Gandhinagar isolates, as reported earlier on other coronaviruses [65].

The folding rate is a molecular measure to understand the sequence-structure-function-virulence link of a protein [18–27]. The population of properly folded viral proteins is directly proportional to the level of viral spread and the lethality of the infection [32,66]. However, the kinetic rate of folding directly relies on the types of linear amino acids in a protein, the reduced hydrophobic density, and enhanced polar-polar repulsive interactions [31]. The substituted branched-chain amino acids were shown to interfere with the folding rate, shape, or conformation of a protein, which agreed to the previous studies [28,30,67]; Liji, 2020; [29]. We observed altered drug specificity in the spike protein of Surat isolates when glycine was replaced with valine similar to West Bengal isolates [68]. G14S mutation can influence the catalytic activity and folding rate of PLPro worldwide [69]. Our study found P2144S, A1812D, and S2015R mutations in PLPro affecting the structural

Fig. 3. Comparative analysis of binding specificities of protein variants from Indian SARS-CoV-2 isolates with a human receptor (ACE-2).
characteristics of this enzyme in Kerala (IND/29) and Himachal Pradesh isolates. HIV relies on a host-encoded peptidyl-prolyl cis-trans isomerase and cyclophilin A for the maturation of its coat proteins [70]. SARS-CoV-2 proteins similarly mediate folding and maturation processes with aid of host chaperones to hijack from many antivirals interfaces.

Virulence is tightly coupled to virus load in the host leading to an increase in the transmission of this virus [71]. The virulence and metabolic properties of SARS-CoV-2 increased at a fast-folding rate of a virus as well as host proteins. The peptidyl-prolyl cis-trans isomerase is a protein-folding catalyst that may process virulence determinants of SARS-CoV-2 for cell-to-cell spread [72] and antiviral resistance development [73]. Mutations can assist in the understanding mechanism of pathogenesis and virulence of SARS-CoV-2. Interestingly, a new functional shift from virulence to the metabolic process was predicted when a neutral polar threonine was exchanged with hydrophobic isoleucine in helicase of Kerala (IND/166) isolate. It is known that pathogens produce macromolecules to scavenge for nutrients to feed metabolism boosts virulence [74–76]. According to the ‘survival of the fittest’ theories of evolution, virulence, and metabolism of SARS-CoV-2 might be intimately linked to survival in a certain niche, particularly in host respiratory tracts.

The RdRP is vital for the replication machinery of RNA viruses. The P4715L mutation was located in a hydrophobic cleft of this protein providing the ability to hijack from remdesivir by the earlier work [77]. Our study also detected consecutive alterations of R203K and G204R in the viral nucleocapsid protein. It was found to bring a change in the local structural environment of this protein to escape from antivirals [78]. The ACE-2-B0AT1 complex is assembled with a collectrin-like domain of ACE-2 that is recognized by a receptor-binding domain of spike glycoprotein in SARS-CoV-2 [6]. Mutations on this complex or ACE-2 alone can alter the binding specificity to drugs, vaccines, and ACE-2. The folding rate is a key measure to determine a sequence-structure-function-virulence link of the SARS-CoV-2 proteome, which in turn increases the viral load in a host via the rapid copying of matured (folded) proteins. The new mutations identified in a proofreading exonuclease can alter the drug specificity of not only this enzyme but also the RdRp complex, suggesting the failure of ribavirin and remdesivir to Indian isolates. Mutations in antigenic determinants of spike protein confer the escaping ability to Surat isolate from human neutralizing antibodies. Most antivirals and vaccines are still more effective for Indian isolates. However, emodin and artenimol are shown better molecular interactions with the respective targets, suggesting repurposable antivirals for the treatment of COVID-19 in India. Finally,
we conclude that further in vitro and in vivo studies are required to evaluate the efficacy of current antiviral therapies on SARS-CoV-2 infectivity.

Declaration of competing interest
The authors confirm that this article’s content has no Competing interest.

Acknowledgments
The author would like to thank the Science and Engineering Research Board, Department of Science and Technology (MSC/2020/000438), New Delhi, India, for financial support.

Appendix A. Supplementary data
Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2021.104315.

Authors’ contributions
CP conceived of the presented idea and verified the analytical methods and manuscript. PR collected the data and performed the analysis. SS contributed to data collection and analysis. BM contributed to the final manuscript. All authors discussed the results and contributed to the manuscript.

Data availability
Supplementary file 1.

List of genomic variations identified from Indian SARS-CoV-2 isolates.

Ethical declaration
The present study does not contain any studies with human participants or animals performed by any of the authors.

For this type of study formal consent is not required.

References
[1] F. Wu, S. Zhao, B. Yu, Y.M. Chen, W. Wang, Z.G. Song, Y.H. Tao, J.H. Tian, Y.Y. Pei, L.M. Yuan, Y.L. Zhang, F.H. Dai, Y. Liu, Q.M. Wang, J.J. Zheng, L. Xu, E. Holmes, Y.Z. Zhang, A new coronavirus associated with human respiratory disease in China, Nature 579 (2020) 265–269.
[2] R. Ralph, J. Lew, T. Zeng, M. Francis, B. Xue, M. Roux, A. Toloue Ostdaghavai, S. Rubinio, N.J. Dawe, M.N. Al-Ahmad, D.J. Kelvin, C.D. Richardson, J. Kindrachuk, D. Falzarano, A.A. Kelvin, 2019 nCoV (Wuhan virus), a novel coronavirus: human-to-human transmission, travel-related cases, and vaccine readiness, J Infect Dev Ctries 14 (2020) 3–17.
[3] J.H. Beigel, K.M. Tomashkin, L.E. Dodd, A.K. Mehta, B.S. Zingman, A.C. Kalil, E. Hohmann, H.Y. Chu, A. Luetkemeyer, S. Kline, D. Lopez de Castilla, R. Trenker, J.M. Young, D. Falzarano, A.A. Kelvin, The structure of SARS-CoV-2 reveals a common origin for viral entry with MERS-CoV and HCoV-229E, bioRxiv Preprint (2020), https://doi.org/10.1101/2020.07.07.197095.
[4] D. Mercatelli, F.M. Giorgi, Geographic and genomic distribution of SARS-CoV-2 mutations, Front. Microbiol. 11 (2020) 1800.
[5] R. Kumar, H. Verma, N. Singhvi, U. Sood, V. Gupta, M. Singh, R. Kumari, P. Hira, S. Nagar, C. Talwar, N. Nayyar, A. Anand, C.D. Rawat, M. Verma, R.K. Negi, Y. Singh, R. Lal, Comparative genomic analysis of rapidly evolving SARS-CoV-2 reveals mosaic pattern of phylogenographical distribution, mSystems 5 (2020) e00505–e00520.
[6] R. Choppa, Z. Woodman, K. Misana, M. Miotshwa, D.P. Martin, C. Seoige, F. Treurnicht, D.A. de Rosa, W. Hide, S.A. Karim, C.M. Gray, C. Williamson, Transmission of HIV-1 CTL escape variants provides HLA-mismatched recipients with a survival advantage, PLoS Pathog. 4 (2008) e100033.
[7] L. Brocz, M. Xie, S. Benichou, J. Bouchet, Mechanisms for cell-to-cell transmission of HIV-1, Front. Immunol. 9 (2018) 260.
[8] C.Z. Bufalo, T. Iwamoto, J.H. Hurley, X. Ren, How HIV Nef proteins hijack membrane traffic to promote infection, J. Virol. 93 (2019) e01322–19.
[9] M.P. Robertson, H. Igel, R. Baertsch, C.D. Richardson, The structure of a rigorously conserved RNA element within the SARS virus genome, PLoS Biol. 3 (2005) e5.
[10] R. Kumar, H. Verma, N. Singhvi, U. Sood, V. Gupta, M. Singh, R. Kumari, P. Hira, S. Nagar, C. Talwar, N. Nayyar, A. Anand, C.D. Rawat, M. Verma, R.K. Negi, Y. Singh, R. Lal, Comparative genomic analysis of rapidly evolving SARS-CoV-2 reveals mosaic pattern of phylogenographical distribution, mSystems 5 (2020) e00505–e00520.
[11] R. Kumar, H. Verma, N. Singhvi, U. Sood, V. Gupta, M. Singh, R. Kumari, P. Hira, S. Nagar, C. Talwar, N. Nayyar, A. Anand, C.D. Rawat, M. Verma, R.K. Negi, Y. Singh, R. Lal, Comparative genomic analysis of rapidly evolving SARS-CoV-2 reveals mosaic pattern of phylogenographical distribution, mSystems 5 (2020) e00505–e00520.
[12] P. Forster, L. Forster, C. Renfrew, M. Forster, Phylogenetic network analysis of SARS-CoV-2 genomes, Proc. Natl. Acad. Sci. U.S.A. 117 (2020) 9241–9243.
[13] M. Joshi, A. Puvur, D. Kumar, A. Ansari, M. Pandya, J. Raval, Z. Shokat, K. Shokat, N.J. Krogan, A SARS-CoV-2 S protein interaction map reveals targets for drug repurposing, Nature 566 (2019) 459–466.
[14] M. Joshi, A. Puvur, D. Kumar, A. Ansari, M. Pandya, J. Raval, Z. Patel, P. Trividli, M. Gandhi, L. Pandya, K. Patel, N. Savaliya, S. Bagatharia, S. Kumar, C. Joshi, Genomic variants in SARS-CoV-2 genomes from Gujarat: underlining role of variants in disease epidemiology, bioRxiv Preprint (2020), https://doi.org/10.1101/2020.07.10.197095.
[15] C.Z. Bufalo, T. Iwamoto, J.H. Hurley, X. Ren, How HIV Nef proteins hijack membrane traffic to promote infection, J. Virol. 93 (2019) e01322–19.
[16] M.P. Robertson, H. Igel, R. Baertsch, C.D. Richardson, The structure of a rigorously conserved RNA element within the SARS virus genome, PLoS Biol. 3 (2005) e5.
[17] R. Kumar, H. Verma, N. Singhvi, U. Sood, V. Gupta, M. Singh, R. Kumari, P. Hira, S. Nagar, C. Talwar, N. Nayyar, A. Anand, C.D. Rawat, M. Verma, R.K. Negi, Y. Singh, R. Lal, Comparative genomic analysis of rapidly evolving SARS-CoV-2 reveals mosaic pattern of phylogenographical distribution, mSystems 5 (2020) e00505–e00520.
[18] P. Forster, L. Forster, C. Renfrew, M. Forster, Phylogenetic network analysis of SARS-CoV-2 genomes, Proc. Natl. Acad. Sci. U.S.A. 117 (2020) 9241–9243.
[19] M. Joshi, A. Puvur, D. Kumar, A. Ansari, M. Pandya, J. Raval, Z. Patel, P. Trividli, M. Gandhi, L. Pandya, K. Patel, N. Savaliya, S. Bagatharia, S. Kumar, C. Joshi, Genomic variants in SARS-CoV-2 genomes from Gujarat: underlining role of variants in disease epidemiology, bioRxiv Preprint (2020), https://doi.org/10.1101/2020.07.10.197095.
S. Yuan, H.C.S. Chan, S. Filipek, H. Vogel, PyMOL and Inkscape bridge the data and K. Katoh, J. Rozewicki, K.D. Yamada, MAFFT online service: multiple sequence A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F. A.G. de Brevern, A. Bornot, P. Craveur, C. Etchebest, J.C. Gelly, PredyFlexy: W. Zheng, C. Zhang, E.W. Bell, Y. Zhang, I-TASSER gateway: a protein structure A. Gavryushkina, J. Heled, G. Jones, D. Kühnert, N. De Maio, M. Matschiner, F. K. Mendes, N.F. Müller, H.A. Ogilvie, L. du Plessis, A. Popinga, A. Rambaut, D. Rassmussen, I. Siveronri, M.A. Suchard, C.H. Wu, D. Xie, C. Zhang, T. Stadler, A. J Drummond, Beast 2.5, an advanced software platform for Bayesian evolutionary analysis, PLoS Comput. Biol. 15 (2019), e1006650.

T.A. Hall, BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT, Nucleic Acids Symp. Ser. 41 (1999) 35–38.

S.F. Altschul, T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, D. J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402.

V. Xiong, X. Li, X. Gao, Q. Dong, MicroG/M: a mutation tracker for SARS-CoV-2 and other microbial genome sequences, Front. Microbiol. 11 (2020) 1502.

AG. de Brevem, A. Bornot, P. Craveur, C. Etchebest, J.C. Gelly, PredyFlexy: flexibility and local structure prediction from sequence, Nucleic Acids Res. 40 (2012) W317–W322.

M.M. Gromiha, A.M. Thangakani, S. Selvaraj, FOLD-RATE: prediction of protein kinetic order and rate, Bioinformatics 23 (2007) 385–386.

S. Saha, G.P.S. Raghava, VICMpred: SVM-based method for predicting functional domains of gram-negative bacteria using amino acid patterns and composition, Dev. Reprod. Biol. 4 (2006) 42–47.

D.S. Wistuba, Y.D. Feumang, A.C. Guo, E.J. Lo, A. Marcu, J.R. Grant, T. Sajed, D. Johnson, C. Li, Z. Sayeeda, N. Assempre, I. Ilynnkar, Y. Liu, A. Maciejewski, N. Gale, A. Wilson, L. Chin, R. Cummings, D. Le, A. Pon, C. Knox, M. Wilson, Drugbank5.0: a major update to the Drugbank database for 2018, Nucleic Acids Res. 46 (2018) D1074–D1082.

M. Yuan, H. Liu, N.C. Wu, C.D. Lee, X. Zhuo, F. Zhao, D. Huang, W. Yu, Y. Hu, H. Tien, T.F. Rogers, E. Landais, D. Sok, J.G. Jardine, D.R. Burton, L.A. Wilson, Structural basis of a viral antibody response to SARS-CoV-2: bioInsight Preprint. http://doi.org/10.1101/2020.06.08.141267, 2020.

A. Waterhouse, M. Bertiini, S. Bienert, G. Studer, G. Tauriello, G. Gremi, F. T. Heer, T.A.P. de Beer, C. Remperier, L. Bordoli, R. Lepore, T. Schwede, SWISS-MODEL: homology modelling of protein structures and complexes, Nucleic Acids Res. 46 (2018) W296–W303.

W. Zheng, C. Zhang, E.W. Bell, Y. Zhang, I-TASSER gateway: a protein structure and function prediction server powered by XSEDE, Future Generat. Comput. Syst. 99 (2019) 73–85.

D. Bhattacharya, refined: improved protein structure refinement using machine learning based restrained relaxation, Bioinformatics 35 (2019) 3320–3328.

E. Mashiali, D. Schneider-Duhamon, N. Andrustras, R. Nustinov, H.L. Wolfson, FireDock: a web server for fast interaction prediction in molecular docking, Nucleic Acids Res. 36 (2008) W229–W232.

S.-Y. Huang, Z. Zou, An iterative knowledge-based scoring function for protein–protein recognition, Proteins 72 (2008) 557–579.

T. Ostroumova, N. Sotirov, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multitreading, J. Comput. Chem. 31 (2010) 455–461.

S. Yuan, H.C.S. Chan, S. Filipek, H. Vogel, PyMOL and Inkscape bridge the data and the data visualization landscape 24 (2016) 2041–2042.

R.A. Laskowski, M.B. Swindells, LigPilot: multiple ligand-protein interaction diagrams for drug discovery, J. Chem. Inf. Model. 51 (2011) 2778–2786.