Briefings in Bioinformatics, 00(00), 2020, 1–15

doi: 10.1093/bib/bbaa378
Case Study

Identifying the natural polyphenol catechin as a multi-targeted agent against SARS-CoV-2 for the plausible therapy of COVID-19: an integrated computational approach

Chandra Bhushan Mishra†, Preeti Pandey†, Ravi Datta Sharma†, Md. Zubbair Malik, Raj Kumar Mongre, Andrew M. Lynn, Rajendra Prasad, Raok Jeon and Amresh Prakash

Corresponding authors: Amresh Prakash, Amity Institute of Integrative Sciences and Health (AIISH), Amity University Haryana, Gurgaon 122413, India. E-mail: amreshprakash@jnu.ac.in, aprakash@ggn.amity.edu; Raok Jeon, College of Pharmacy, Sookmyung Women's University, Cheongpa-ro 47-gil 100, Yongsan-gu, Seoul, 04310, Republic of Korea. E-mail: rjeon@sookmyung.ac.kr
†These authors contributed equally to this work.

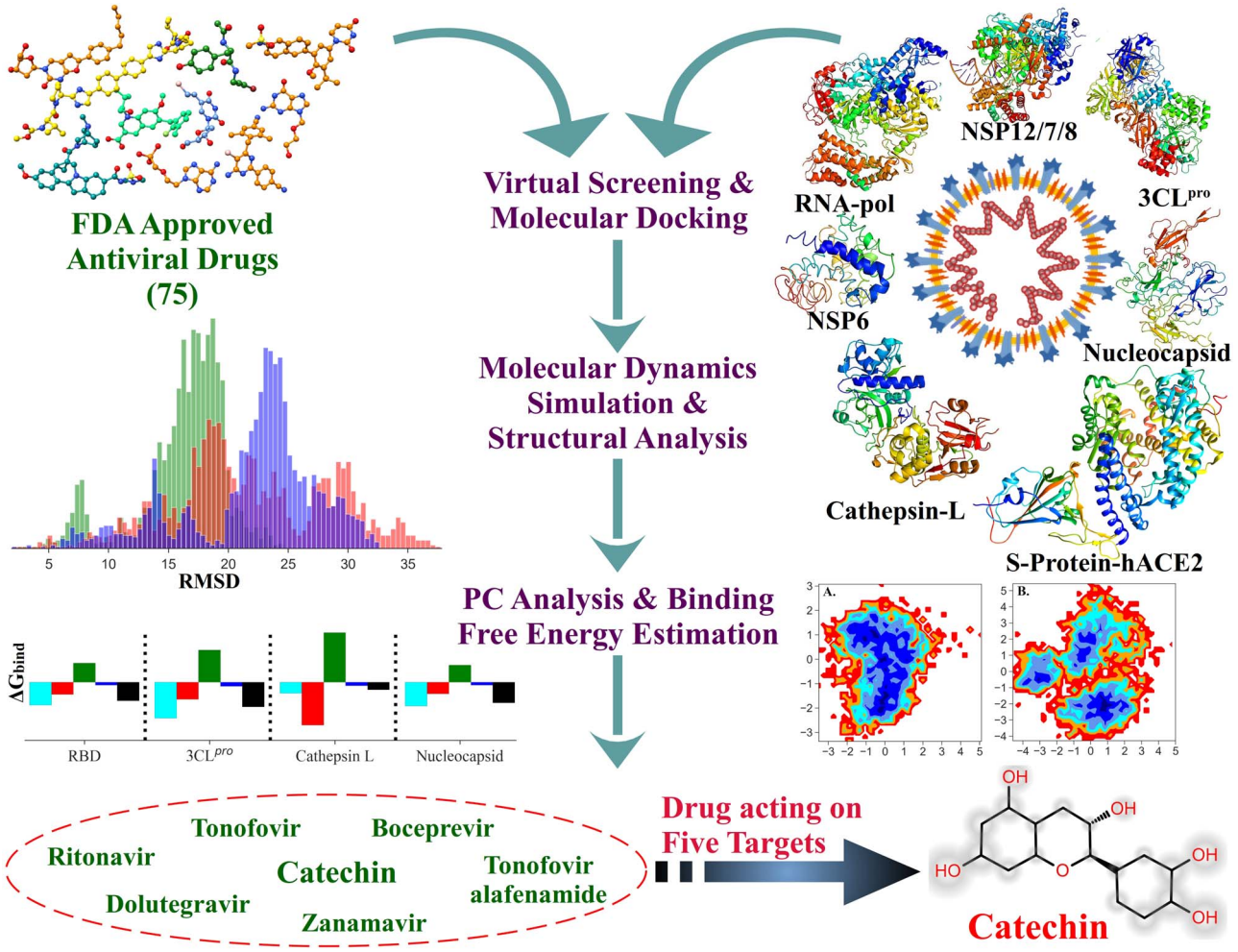
Abstract

The global pandemic crisis, coronavirus disease 2019 (COVID-19), caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has claimed the lives of millions of people across the world. Development and testing of anti-SARS-CoV-2 drugs or vaccines have not turned to be realistic within the timeframe needed to combat this pandemic. Here, we report a comprehensive computational approach to identify the multi-targeted drug molecules against the

Dr Chandra B. Mishra is working as a scientist at the College of Pharmacy, Sookmyung Women's University, Seoul, South Korea. His areas of expertise and research interests are development of selective carbonic anhydrase inhibitors, anti-epileptic agents, anti-Parkinsonian agents and anticancer agents. He is an experienced researcher in the field of drug discovery and development.
Dr Preeti Pandey has completed her PhD in computational biology and is currently working as a postdoctoral fellow at the Department of Chemistry & Biochemistry, University of Oklahoma, OK, USA. She has expertise in computer-aided drug designing and development of algorithm of the estimation of ligands binding free energy.
Dr Ravi D. Sharma is an assistant professor at the Amity Institute of Technology, Amity University, Haryana. His areas of expertise are computational biology, development of algorithm for proteomics (Big Data, RNA Seq) analysis.
Dr Md. Zubbair Malik is a young scientist at the School of Computational and Integrative Sciences, Jawaharlal Nehru University, New Delhi, India. His areas of expertise and research interests are various aspects of bioinformatics, network medicine, network biology, stochastic dynamics and modeling, nonlinear dynamics and complex network theory.
Dr Raj K. Mongre is working as a postdoctoral fellow at the College of Pharmacy, Sookmyung Women's University, Seoul, South Korea. After getting a PhD degree in animal biotechnology, he joined the College of Pharmacy as a postdoctoral fellow and is involved in the investigation of novel molecular pathway associated with cancer progression.
Prof. Andrew M. Lynn is a faculty at the School of Computational and Integrative Sciences, Jawaharlal Nehru University, New Delhi 110067, India. His areas of expertise are computational biology, development of algorithm for fragment-based ligand designing and estimation of ligand binding free energy.
Prof. (Dr) Rajendra Prasad is a director at the Amity Institute of Biotechnology and is the dean of Faculty of Science Engineering and Technology, Amity University Haryana, Haryana 122413, India. His area of expertise is biochemistry, and he has an experience of 50 years in clinical resistance to antifungal mechanism and regulation of resistances to new antifungal molecular mycology.
Prof. Raok Jeon is working as a professor of medicinal chemistry at the College of Pharmacy, Sookmyung Women's University, Seoul, South Korea. Her areas of expertise and research interests are development of kinase inhibitors and epigenetic modulators for the therapy of cancer, inflammation and autoimmune disease.
Dr Amresh Prakash is an assistant professor at Data Sciences, Amity Institute of Integrative Sciences and Health, Amity Institute of Integrative Sciences and Health, Amity University, Haryana. He is an expert in computational biophysics, having 8 years of postdoctoral experiences in conformational sampling and statistical analysis of proteins involved in neurodegenerative diseases and implication of machine leaning for identifying the structural ensembles which may trigger the disease state.
Submitted: 24 August 2020; **Received (in revised form):** 3 November 2020

SARS-CoV-2 proteins, whichare crucially involved in the viral–host interaction, replication of the virus inside the host, disease progression and transmission of coronavirus infection. Virtual screening of 75 FDA-approved potential antiviral drugs against the target proteins, spike (S) glycoprotein, human angiotensin-converting enzyme 2 (hACE2), 3-chymotrypsin-like cysteine protease ($3CL^{pro}$), cathepsin L (CTSL), nucleocapsid protein, RNA-dependent RNA polymerase (RdRp) and non-structural protein 6 (NSP6), resulted in the selection of seven drugs which preferentially bind to the target proteins. Further, the molecular interactions determined by molecular dynamics simulation revealed that among the 75 drug molecules, catechin can effectively bind to $3CL^{pro}$, CTSL, RBD of S protein, NSP6 and nucleocapsid protein. It is more conveniently involved in key molecular interactions, showing binding free energy (ΔG_{bind}) in the range of −5.09 kcal/mol (CTSL) to −26.09 kcal/mol (NSP6). At the binding pocket, catechin is majorly stabilized by the hydrophobic interactions, displays ΔE_{vdW} values: −7.59 to −37.39 kcal/mol. Thus, the structural insights of better binding affinity and favorable molecular interaction of catechin toward multiple target proteins signify that catechin can be potentially explored as a multi-targeted agent against COVID-19.

Graphical Abstract



Key words: SARS-CoV-2; COVID-19; multi-targeted drug; catechin; free energy landscape

Introduction

There are different members of Coronaviridae family of virus, which often cause mild, moderate to severe respiratory symptoms in humans [1]. Recently, the novel coronavirus [2019-nCoV or severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)] appears to be the highly infectious and contagious virus of this family [1, 2]. Although SARS-CoV-2 shares a high level of genetic similarity with SARS-CoV, the infection rate of SARS-CoV-2 is much higher than suspected [3]. The molecular and structural organization of the virus includes an envelope, non-segmented, positive-sense RNA which codes for several structural proteins such as spike (S) protein, envelope (E) protein, membrane (M) protein and the nucleocapsid (N) proteins and also 16 putative non-structural proteins [NSPs, encoded by replicase complex (orf1ab)] [1, 4–6]. During the last couple of months, dozens of coronavirus vaccines have already been tested, and approximately more than hundreds are either under development or in the preclinical investigation [5, 7]. However, the success of these efforts remains elusive. Therefore, the need of the hour is to identify novel and effective measures to control the pandemic [3, 7, 8].

The major setback in identifying drugs/vaccines against COVID-19 is the lack of knowledge about the mechanism of action of the virus, molecular targets and network of associated molecular interactions. Recently, several targets have already been discovered that interact with SARS-CoV-2, such as human angiotensin-converting enzyme 2 (hACE2), transmembrane protease serine 2 (TMPRSS2), phosphatidylinositol 3-phosphate 5-kinase (PIKfyve), two pore channel subtype 2 (TPC2) and cathepsin L (CTSL) [6, 9, 10]. At the same time, various research groups have identified the effective inhibitors against some of these targets, such as main protease M^{pro} ($3CL^{Pro}$) [11, 12], envelope (E) protein [13], RNA-dependent RNA polymerase (RdRp) [14] and spike (S) protein [15, 16].

Since the genes of RNA viruses (including SARS-CoV-2) are genetically variable [17–19], they can quickly accumulate genomic mutations through an error-prone viral reverse transcriptase, which advances their adaptation inside the human host. This further adds to the difficulty in designing active antiviral therapeutics against RNA viruses [9]. Moreover, most of the antiviral drugs today are single target drugs designed against specific viral enzymes, which are essential for viral interaction, replication or invasion. Therefore, the high rate of mutations in these single viral drug targets has been main cause for reduced susceptibility of currently available antiviral drugs [20].

Nevertheless, finding the compounds having efficacy for multiple molecular targets, remained a preferable approach in disorders caused by highly mutable pathogens [21, 22]. Targeting different molecular targets with a single drug is always a preferable approach over combination therapy to avoid unwanted drug interactions [9, 23]. Additionally, the drugs designed for multiple protein targets are extensively preferable for the treatment of infectious, inherited and complex diseases due to low treatment cost, less drug dosage and minimal side effect and drug–drug interactions. Therefore, with the immediate requirement of multi-targeted strategies against the novel coronavirus SARS-CoV-2, with either biologically active drug-like molecules or approved drugs are in pressing priority [6, 24, 25]. Recent advancements in the computational techniques have proven their efficiency for identifying the potential drug candidates [7, 26–28]. Considering the improvement, reliability and accuracy of computational methods, it has become a suitable choice to design structure-based drugs [21, 29–31]. Keeping these facts in mind, we adopted a multi-target drug discovery approach to hit various druggable targets of SARS-CoV-2, which may appear highly beneficial to strike this highly mutated virus. The repurposing of FDA-approved drug molecules is safe and significantly free from off-targets binding which warrants severe toxicity [32, 33].

Thereby, we screened 75 FDA-approved antiviral drugs against known targets of SARS-CoV-2 [34, 35]. We have taken both human proteins as well as viral targets as a strategy. The targets were chosen as the recently published structures of SARS-CoV-2 proteins complexed with drug molecules. Following seven targets were identified: (1) the hACE2 interacting with the (2) transmembrane viral spike (S) glycoprotein at receptor-binding domain (RBD), which forms homotrimers protruding from the viral surface [6, 7, 15, 16]; (3) the highly immunogenic, antigenic and abundantly expressed viral nucleocapsid (N) protein, which plays essential roles in viral genome packaging by formation of helical ribonucleoproteins [36]; (4) main protease M^{pro} ($3CL^{Pro}$), an essential viral enzyme for processing the polyprotein complexes that are translated from the viral RNA [37]; (5) the human endosomal cysteine protease CTSL required for viral entry [7, 38]; (6) non-structural viral protein NSP6 which dwells in the endoplasmic reticulum (ER) and has role in the generation of autophagosomes [39] and (7) the NSP12, along with the two other cofactors NSP7-NSP8 as a complex, aiding increased RdRp template binding and processivity. Firstly, the virtual screening of 75 FDA-approved antiviral drugs was performed against these targets to select a highly potent multi-targeted agent. The best binding poses of antiviral drugs with target proteins display a wide range of binding affinities. Among them, (+)-catechin (catechin) emerged as a multi-targeted agent that can effectively bind with five target proteins: RBD, CTSL, nucleocapsid protein, $3CL^{pro}$ and NSP6.

Catechin (flavan-3-ol) is a natural phenol and a major chemical component of sinecatechin, a first FDA-approved herbal drug for the treatment of external genital warts caused by human papilloma virus (HPV) infections. Topical ointment Veregen is a marketed medicine and it is a purified form of catechins, extracted from the leaves of Chinese green tea, which comprises 80% catechins. It is well recognized for the antiviral activity, anti-bacterial activity, anti-inflammatory and for the immunostimulatory actions [34]. Further, we have acquired several data contributing to the stable structural dynamics of the protein–ligand complex, including free energy landscape (FEL), which validates the binding efficacy of drug molecules using molecular dynamics (MD) simulation and molecular mechanics-Poisson–Boltzmann Surface Area (MM-PBSA). Thus, identifying the catechin as a novel multi-targeted agent may provide the structural basis for the designing strategy of potential drug molecules, targeting SARS-CoV-2 in the therapy of COVID-19.

Materials and methods

Protein structure retrieval

The three-dimensional coordinates of protein structures were taken from the Protein Data Bank (www.rcsb.org) [40]. The co-crystalized X-ray structure of SARS-CoV-2 spike receptor with ACE2 (PDB ID: 6M0J) [41], 3-chymotrypsin-like cysteine protease ($3CL^{pro}$) (PDB ID: 6M2N), CTSL (PDB ID: 6F06) [41], crystal structure of nucleocapsid protein (PDB ID: 6M3M) [36], RdRp (PDB ID: 6M71) [42], non-structural protein 6 (NSP6) [39, 43] and cryo–electron microscopy structure of RdRp enzyme with

remdesivir and NSP12-NSP7-NSP8 complex (PDB ID: 7BV2) [14] were taken as the targets for the virtual screening of selected FDA-approved antiviral agents as shown in the supporting information, Supplementary Figure S1 available online at https://academic.oup.com/bib.

Virtual screening and molecular Docking

The virtual screening of FDA-approved antiviral compounds against the selected SARS-CoV-2 proteins was performed using Glide, Schrodinger, LLC [44–46]. Glide involves three-step filtering methods, standard precision, extra precision and the selection of best docked compounds by integrating coulombic and van der Waals (vdW) interaction energies and Glide scoring function. CORINA v2.64 software package [47] was utilized to add the missing hydrogen atoms and optimize the sdf format structures of the ligands. The lowest energy three-dimensional structures of ligands were generated using Ligprep [48]. The ionization/tautomeric states of the selected compounds were taken care of by Epik and a maximum of up to 32 conformations was generated per ligand, using the Schrodinger protocol [49–51]. The molecular interactions of docked complexes were analyzed using PyMol [52] and LigPlot [53].

MD simulation

MD simulations were carried out on the coordinates of protein–ligand complexes using GROMACS-2018.1 [54], with the protein interactions approximated using CHARMM36 force field [55]. The ligand parameters were generated utilizing CGenFF server [56]. Each protein–ligand complex was placed in the center of a cubic simulation box with 10 Å distance to the edges and solvated with TIP3P water molecules. The counterions (Na^+ Cl^-), 0.15 M, were added to neutralize the system. The periodic boundary condition was defined in the x, y and z directions [57], and the electrostatic interactions were evaluated using particle-Ewald summation [57], and a cut-off of 10 Å was used for calculation of vdW interactions. The resulting systems were energy-minimized by steepest descent and conjugate gradient algorithms. Energy minimization was performed for 50 000 steps. Equilibration was first performed for 500 ps in an NVT ensemble and for the subsequent 500 ps in an NPT ensemble. Temperature and pressure were set at $T = 300$ K and 1 bar, which was controlled by a Parrinello–Danadio–Bussi thermostat [58] and Parrinello–Rahman pressure [59], respectively. The integration step of 2 fs was used. Each system was simulated for 200 ns and the snapshots were saved every 10 ps for further analysis. All production runs were performed on CUDA-enabled Tesla GPU machine (DELL T640 with V100 GPU) and OS Centos 7 [60, 61].

MD analysis

The obtained MD trajectories were analyzed using GROMACS utilities. The structural order parameters that we measured are: root-mean-square deviation (RMSD), the radius of gyration (R_g), solvent-accessible surface area (SASA), root-mean-square fluctuation (RMSF) and hydrogen bond (H-bond) interactions. H-bonds were defined by a distance cut-off of 3.5 Å between the donor and acceptor atoms and by an angle cut-off of 30°. Similarly, a hydrophobic interaction was defined by the condition that the distance between two residues (i and j, with $|i - j| > 3$) is less than 4.5 Å. Principal component analysis performed using the projection of principal components (PCs), PC1 and PC2, along the native structure [61, 62] and gmx-sham utilized for the FEL [63, 64].

Binding free energy estimation

The binding free energy of the protein–ligand complexes was evaluated using MM-PBSA, which describes the structural and molecular stability of the ligands in the active site of the protein [27, 28, 65, 66]. The binding free energy of a protein–ligand complex ($\Delta G_{binding}$) can be written as,

$$\Delta G_{binding} = \langle G_{complex} \rangle - \langle G_{receptor} \rangle - \langle G_{ligand} \rangle,$$

where $G_{complex}$ represents the free energy of the protein–ligand complex; $G_{receptor}$, the free energy of protein; G_{ligand} is the free energy of ligand and $< >$ represents the ensemble average.

Excluding the entropy term ($T\Delta S$), the above equation for the binding free energy can be approximately written as,

$$\Delta G_{binding} = \Delta E_{MM} + \Delta G_{solv},$$

where ΔE_{MM} is the change in the average molecular mechanics interaction energy (gas phase) upon ligand binding computed as the sum of the changes in the bonded and non-bonded (electrostatics and vdW) interactions upon ligand binding ($\Delta E_{MM} = \Delta E_{bonded} + \Delta E_{electrostatics} + \Delta E_{vdw}$). ΔG_{solv} is the change in solvation free energy upon ligand binding. Further, ΔG_{solv} can be written as,

$$\Delta G_{solv} = \Delta G_{POL} + \Delta G_{NP},$$

where ΔG_{POL} is the change in the polar part of the solvation free energy, and ΔG_{NP} is the change in the non-polar part of the solvation free energy as a result of ligand binding to the proteins. PB equation was used for the estimation of the polar part of the solvation free energy, and the non-polar part was estimated with a surface area-based approach. Binding free energy ($\Delta G_{binding}$) for the protein–ligand complex was estimated using the MMPBSA.py script of the AMBER Tools [67]. An ionic strength of 0.15 M and a solute dielectric constant value of 2 was used for the PBSA calculations. Considering the convergence issues associated with the MM-PBSA calculation, only last 50 ns data were used.

Results and discussion

Targeting SARS-CoV-2 proteins for identifying multi-target agents

The development of novel molecules is not thought to be realistic in the timeframe needed to impact on the pandemic of SARS-CoV-2 [32, 35]. Thus, there is a clear need for new treatments that will reduce the mortality rate while we wait for a vaccine or the novel drug molecules. The repurposing of FDA-approved drugs often provides an advantage to identify the promising drug molecules in the shorter time which can be undertaken for the medical application without delay [68, 69]. Furthermore, it is more cost-effective and less risky as compared to the *de novo* drug discovery [32, 34, 35, 70]. However, the ability of a chemical compound to work as a drug lies in its competency to bind efficiently to a druggable target. While there exist plenty of methods to evaluate the binding affinity of a ligand toward a target starting from very reliable and accurate alchemical free energy methods (computationally very costly) to

less accurate docking methods, molecular docking and virtual docking remain the first choice to screen chemicals as they offer reasonable accuracy with modest computational efforts [26, 28, 29]. Considering this, we have first performed molecular docking study to screen the efficacy of 75 FDA-approved drugs against various important druggable targets of SARS-CoV-2 [34]. Our docking study revealed that some of the FDA-approved drugs have excellent interactions with particular target, displaying satisfactory docking scores. Docking scores of the top 10 antiviral drug molecules corresponding to each protein are enumerated in Table 1. Results demonstrate the best binding affinity of antiviral drugs: ritonavir, dolutegravir, tenofovir, tinofoviralafenamide, boceprevir, catechin and zanamivir toward the target proteins: $3CL^{pro}$, RdRp, ACE2, CTSL, NSP6, nucleocapsid protein and RBD, respectively (Supplementary Figures S2 and S3 available online at https://academic.oup.com/bib). Interestingly, the naturally derived polyhydroxy molecule catechin showed multi-targeted action against all seven targets; however, its promising binding capability (with the cut off range >5.0 kcal/mol) was noticed against the five important targets engaged in the invasion and survival of SARS-CoV-2 in to the human cells. We found that catechin effectively binds to $3CL^{pro}$, CTSL, NSP6, nucleocapsid protein and RBD of S protein, showing the docking scores in the range of −5.79 to −8.34 kcal/mol.

Molecular interaction of catechin as multi-targeted agent

The structure-based virtual screening protocol bestowed catechin as the effective multi-targeted agent which effectively hits $3CL^{pro}$, CTSL, NSP6, nucleocapsid protein and RBD (Figure 1 and supporting information, Supplementary Figure S3 available online at https://academic.oup.com/bib). The SARS-CoV-2 protein, $3CL^{pro}$ plays a critical role in the replication of the virus particles and is a potential target for anti-coronaviruses inhibitors screening. The active site of $3CL^{pro}$ consists of Cys-His catalytic dyad (Cys145 and His41), which is highly conserved in the CoVs family, also referred to as the main protease, M^{pro} [71]. The molecular docking result shows that catechin nicely fits in the active site of $3CL^{pro}$ with the highest docking score, −8.34 kcal/mol, among all selected five target proteins. It displays H-bonds with Thr26, Met49, Arg188 and Gln189, whereas Leu27, His41 and Leu58 are involved in hydrophobic interactions. Benzopyran moiety of catechin is oriented toward Met49, Arg188 and Gln189, noticeably, di-hydroxy phenyl occupied at the Cys145 and His41 catalytic dyad, which is crucial for ligand binding.

An excellent interaction of this compound was also seen with CTSL, a crucial human protease that promotes SARS-CoV-2 entry by S protein activation [7, 72]. Indeed, catechin occupied the active site of this protease, stabilizing through H-bond interactions as well as hydrophobic interaction. Phenolic hydroxy participate in H-bonding with Ser216, and hydroxy at benzopyran moiety showed H-bonding with Met161. Benzopyran ring was stabilized by the hydrophobic interaction with Trp26, His163 and Ala214.

In the CoVs family, nucleocapsid acts as a multifunctional RNA-binding protein and plays an indispensable role in regulating viral RNA transcription/replication and the modulation of host cell metabolism [73]. The recently solved N-terminal domain of SARS-CoV-2 nucleocapsid consists of antiparallel β-sheets at the core, protruding β-hairpin and short 3_{10}-helix [36]. The molecular interaction of catechin with nucleocapsid shows that phenolic hydroxy formed H-bond interaction with

Table 1. Molecular docking scores (kcal/mol) of the FDA-approved potential antiviral drug molecules (best 10 hit molecules) with SARS-CoV-2 proteins

S. No.	$3CL^{pro}$		RdRp		ACE2		CTSL		NSP6		Nucleocapsid protein		RBD (S protein)	
	Comps.	Score	Comps.	Score	Comps.	Score	Comps.	Score	Comps.	Score	Comps.	Score	Comps.	Score
01	Ritonavir	−8.79	Dolutegravir	−12.43	Tenofovir	−10.73	Tinofoviralafenamide	−9.03	Boceprevir	−8.29	Catechin	−6.23	Zanamivir	−8.78
02	Catechin	−8.34	Peramivir	−10.63	Abacavir	−10.07	Sofosbuvir	−8.77	Vidarabine	−7.90	Abacavir	−5.09	Ribavirin	−5.99
03	Nelfinavir	−8.16	Zanamivir	−10.43	Didanosine	−9.15	Ribavirin	−8.66	Tenofovir	−7.83	Ribavirin	−4.86	Didanosine	−5.72
04	Trifluridine	−8.10	Beclabuvir	−9.88	Zidovudine	−7.98	Indinavir	−8.28	Boceprevir	−7.50	Penciclovir	−4.72	Catechin	−5.79
05	Indinavir	−8.09	Laninamivir	−10.01	Entecavir	−7.13	Nelfinavir	−8.24	Zalcitabine	−7.42	Trifluridine	−4.61	Telbivudine	−5.66
06	Saquinavir	−8.01	Favipiravir	−9.05	Vidarabine	−6.11	Darunavir	−8.16	Podafilox	−7.35	Favipiravir	−4.56	Didanosine	−5.66
07	Boceprevir	−7.70	Peramivir	−9.72	Telbivudine	−5.47	Abacavir	−7.91	Catechin	−7.04	Brivudine	−4.54	Valganciclovir	−5.60
08	Abacavir	−7.63	Tenofovir	−8.58	Stavudine	−5.41	FV100	−7.70	Favipiravir	−6.99	Entecavir	−4.52	Indinavir	−5.42
09	Dolutegravir	−7.48	Adefovir	−8.51	Zalcitabine	−5.01	Catechin	−7.68	Famciclovir	−6.89	Stavudine	−4.44	Valaciclovir	−5.23
10	Sofosbuvir	−7.39	Saquinavir	−8.16	Catechin	−4.89	Stavudine	−7.55	Sofosbuvir	−6.83	Telbivudine	−4.37	Trifluridine	−5.19

Compounds (Comps)

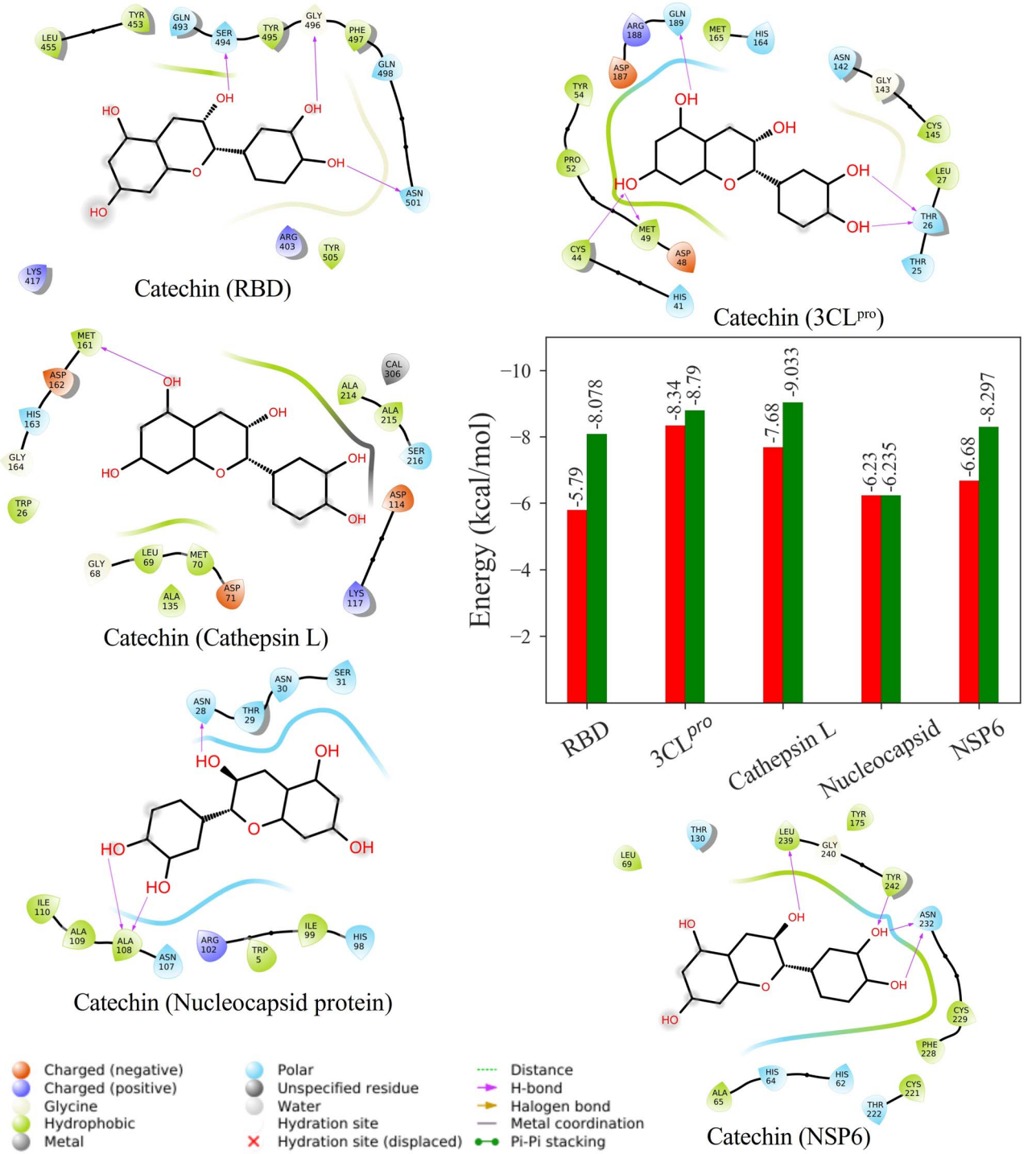


Figure 1. Molecular interactions of FDA-approved antiviral drug, catechin at the binding pocket of potential target proteins, RBD of S protein, 3CL^{pro}, CTSL, nucleocapsid and NSP6, using LigPlot. The bar chart showing the target proteins specific antiviral drugs binding affinity (kcal/mol) obtained through virtual screening (represented by green bars) and the comparative binding affinity (kcal/mol) with catechin shown with red bars.

Ala108, and phenyl moiety imparted hydrophobic interaction with Trp5, Ala109 and Ile110. Hydroxy group at benzopyran moiety was stabilized by the H-bond interaction with Asn28 and Ser31. Additionally, this benzopyran moiety was stabilized through hydrophobic interactions with His98 and Ile99. Another important CoVs protein, non-structural viral protein 6 (NSP6), plays an essential role in viral RNA synthesis by sequestering the membrane of ER of the host cell [74]. A strong affinity of catechin with NSP6 was also discerned. Indeed, catechin contoured well in the active site of NSP6, displaying a satisfactory docking score (−6.68 kcal/mol). It was noticed that catechin forms tight interactions in the active site through three H-bond interactions and several hydrophobic interactions. Hydroxy groups of catechin are found to be involved in H-bonding with

Leu239, Tyr242 and Asn232. Phenyl group, as well as benzopyran moiety, was well stabilized by the hydrophobic interactions with His62, Ala65, Tyr175 and Phe228. The configuration and orientation of catechin was also found to be favorable, showing convincing interactions with the virus surface S glycoprotein which mediates entry into host cells, adhering at the host receptor (hACE2). Evidently, catechin also effectively inhibits the interaction of RBD of S protein with ACE2. Docking result indicates that the scaffold of catechin is well stabilized in the active site of RBD through four H-bonding interactions. These H-bonds are seen between hydroxy groups and active site amino acids such as Tyr453, Tyr495, Gly496 and Asn501. Additionally, hydrophobic interactions between catechin scaffold and amino acids, Arg403, Tyr453, Tyr495, Gly496 and Phe497, are also observed. Interestingly, catechin is also supported by π–π interactions, which can be noticed with Tyr453, Phe497 and Tyr495.

Thus, our molecular docking study evidently indicates that catechin strongly interacts with these five crucial targets associated with SARS-CoV-2, which clearly designates the multi-targeted action against SARS-CoV-2 [75]. However, considering the approximations made in molecular docking (lack of receptor flexibility and conformational entropy, lack of information about the number and free energy of water molecules in the binding site of the protein, etc.) to allow fast screening of chemicals, the dynamics of the protein–ligand interactions are overlooked, and therefore, might not explain the stability of the ligands in the active site of the protein. In the cellular system, biomolecular interactions are dynamic in nature, and the conformational flexibility is an intriguing property of proteins which triggers the biological functions and molecular recognitions [27, 30, 76]. A better understanding of protein–ligand interactions requires an accurate description of the spatial orientation of ligands at the active site of the protein, conformational dynamics which modulates the drug binding, interaction energy and molecular stability [77–79]. To understand the biomolecular interactions at atomic resolution, MD simulation is an efficient and well-established method which mimics the flexible nature of biomolecules, protein conformational changes, protein–ligand interactions, structural perturbation and provides a more realistic picture with atomic details in reference to time [61, 80, 81]. Thereby, to gain a deeper insight into the structural dynamics and stability of catechin binding with SARS-CoV-2 proteins, multiple MD simulations were performed for the period of 200 ns [82, 83]. Additionally, the energetic contribution of binding pocket residues to accommodate the drug molecule, catechin, is estimated using MM-PBSA [28, 84].

Conformational stability of protein–ligand complexes

We assessed the conformational stability of the protein–ligand complexes by measuring various structural order parameters like RMSD, R_g, SASA and RMSF, as shown in Figure 2. On comparing the Cα-RMSD of RBD, CTSL and nucleocapsid protein complex, we observed that catechin achieved stability in the active site of the protein very quickly (Figure 2A). These systems attained equilibrium in 0–5 ns and remained stable throughout the simulation time. The RMSD plot of $3CL^{pro}$ shows an initial rise in RMSD ~0.2 nm, which settles gradually, and a stable equilibrium can be seen up to 90 ns. We find a slight drop in the RMSD of ~0.1 nm around ~100 ns and the undisrupted trajectory is seen up to 200 ns, suggesting stable interaction of catechin in the binding pocket of CTSL. The trajectory of NSP6 with catechin shows slightly large deviations in RMSD during 0–80 ns; thereafter, a gradual drop in RMSD can be seen, which attains equilibrium around ~120 ns. Notably, the stable conformational dynamics of NSP6–catechin is observed up to 200 ns. Thus, the shorter equilibration time taken by RBD, CTSL, nucleocapsid and $3CL^{pro}$ to achieve a steady equilibrium suggests a better equilibrated and stabilized protein–ligand complex structure compared to NSP6. However, the stable trajectory of the NSP6–catechin complex during 120–200 ns signifies that the ligand is spatially well occupied and stabilized with the molecular interactions at the binding pocket of NSP6.

To further understand the structural stability of the protein–ligand complexes, we determined the compactness of the protein structure by computing the radius of gyration (R_g). The R_g plots represented in Figure 2B show that the structural dynamics of RBD, CTSL and nucleocapsid protein and $3CL^{pro}$ remain quite stable throughout the simulation time. The structural integrity of these four proteins was observed to be intact with the average R_g values, 1.83 ± 0.01 nm, 1.64 ± 0.01 nm, 1.50 ± 0.01 nm and 2.19 ± 0.01 nm, respectively. The slight deviations in the R_g plot of NSP6 can be seen during 0–80 ns; after that, the steady equilibrium is noted till the end of simulation at 200 ns, which signifies the stable structural dynamics of the NSP6–catechin complex with an average R_g value, 2.23 ± 0.03 nm. The initial perturbation in the R_g trajectory may indicate the spatial adjustment of the ligand in the binding site of NSP6.

Another important quantity that we measure and analyze to probe the conformational stability of the protein–ligand complex is the SASA. The solvent environment around the protein plays a key role in maintaining the protein fold and governs the protein–ligand interaction processes, orientation and stability. Interestingly, we find that the SASA plots of all five protein–ligand complexes (RBD, CTSL, nucleocapsid protein, $3CL^{pro}$ and NSP6) remain fairly equilibrated during the entire simulation period (0–200 ns) which provides clear evidence of the stable conformational dynamics of protein–ligand interactions (Figure 2C). The average values of the structural order parameters, RMSD, R_g and SASA are shown in the supporting information, Supplementary Table S2 available online at https://academic.oup.com/bib.

Next, we investigated the binding stability of the catechin at the active site of the respected proteins by monitoring the time evolution plots of the average distance from the center of binding pocket to the ligand, as shown in the supporting information, Supplementary Figure S4 available online at https://academic.oup.com/bib. During the period of 200 ns simulation, the average distance of catechin from the binding site of all five proteins ranges between 0.33 nm and 0.40 nm. Although, the peaks of sharp drifts appeared transiently at ~80 ns and 140 ns for RBD and nucleocapsid protein, respectively, the overall distance of catechin to the active site remains favorable for the stable molecular interaction. It is worth noting that the average distance plots of CTSL, $3CL^{pro}$ and NSP6 remain stable throughout simulation time. Thus, this analysis provides an elegant evidence of the spatially well-fitted catechin orientation in the binding sites of proteins.

We further performed RMSF analysis to evaluate the positional fluctuation of each amino acid around its average mean position (Figure 2D). This analysis provides a clue about the mobility of atomic fluctuations related to the structural stability of molecular interaction during the simulation. Usually, the higher values of RMSF are often associated with loops or may be the terminal residues, whereas the lower RMSF values indicate the rigid conformation of stable secondary structures

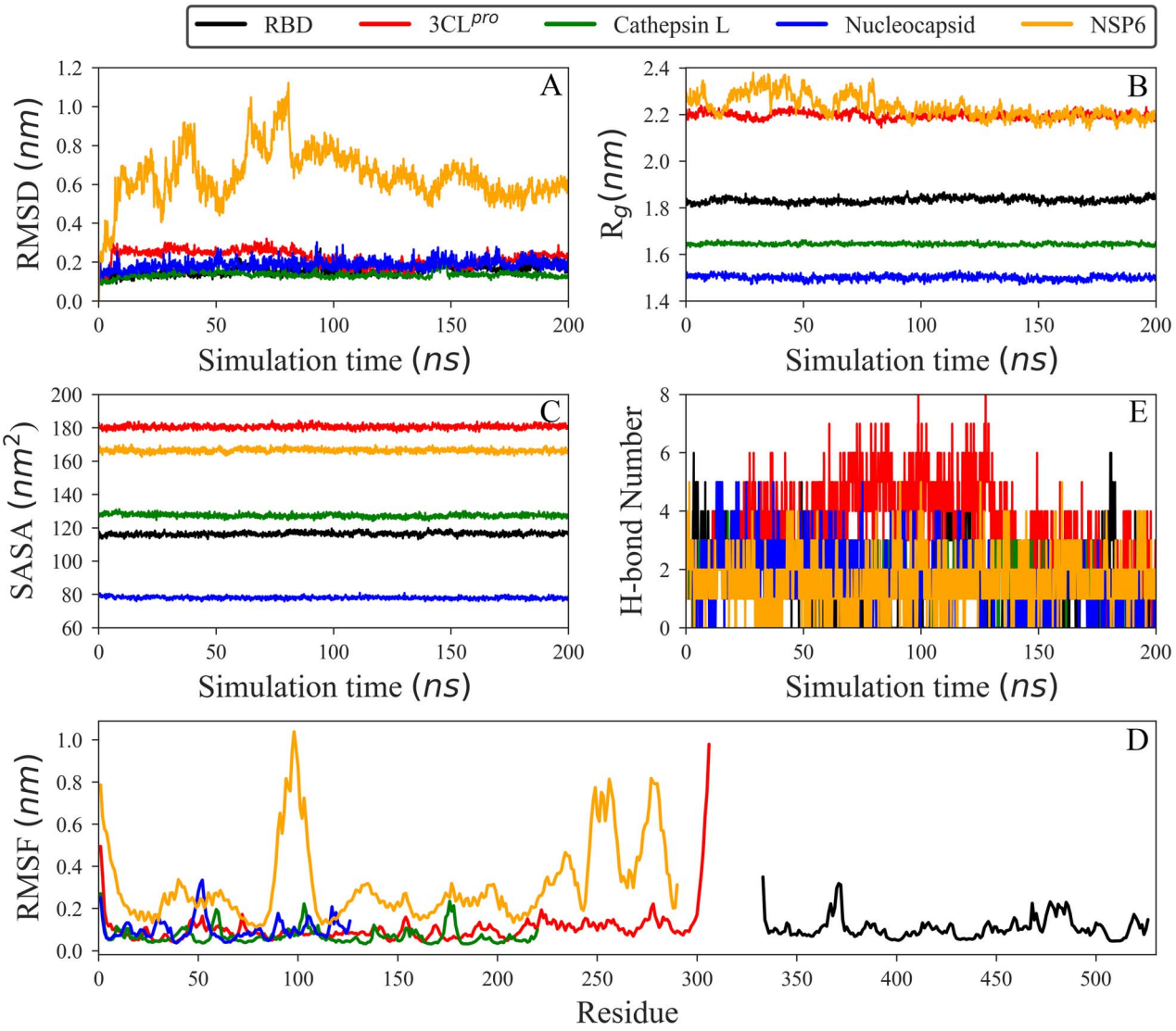


Figure 2. Time evolution plot of the structural order parameters of the antiviral drug, catechin-docked complex with target proteins: RBD, $3CL^{pro}$, CTSL, nucleocapsid and NSP6. (**A**) The RMSD of backbone C^{α}-atoms, (**B**) radius of gyration (R_g), (**C**) SASA plots, (**D**) RMSF showing the average fluctuation of amino acid residues and (**E**) the propensity of H-bonds interaction between the proteins and catechin during the period of simulation (200 ns) at 300 K.

of α-helices and β-sheets. The result indicates that all the complexes show equilibrium fluctuations, except NSP6. The plot shows that the values of RMSF significantly vary for all residues of NSP6 in comparison to the other proteins, RBD, CTSL, nucleocapsid protein and $3CL^{pro}$, respectively. The RMSF plots of these proteins show an average atomic fluctuation <0.15 Å for amino acid residues, which belong to the stable secondary structure, and the regions which displayed high fluctuations represented the atomic flexibility of loops. The structure of NSP6 consists of a helical structure at the core, capped by antiparallel β-sheets and two small helices. The binding pocket is characterized by α-helix-2-4 and α-helix-6-7 and antiparallel β-sheets (β1 and β2) which are enclosed by a hydrophobic loop (Phe235-Tyr242) connecting helices and β-sheets. The RMSF plot shows on average, high fluctuations for the residues belonging to longest loop (Val84-Leu110) connecting α-helix-3 and α-helix-4, loops (Gln257-Ser262 and Leu275-Pro282) connecting to small two helices (α8 and α9) with β-sheets, respectively. The average residual fluctuations observed were reasonably lower for terminal residues of α-helix-2 (Phe42-Phe59), α-helix-3 (Lys63-Met86), α-helix-4 (Lys109-Arg129), α-helix-6 (Ala157-Thr196), α-helix-7(Tyr175-Tyr196) and N-terminal of connecting loop (Phe235-Tyr242) and β1 (Asp243-Val246) which are actively involved in molecular interactions with drug molecules. Further, we also established the conformational stability by analyzing the secondary structural contents, which were observed intact during the simulation (supporting information, Supplementary Figure S5 available online at https://academic.oup.com/bib). The structural snapshots of protein–ligand interactions captured at the time interval of 20 ns described in the supporting information Supplementary Figure S6 available online at https://academic.oup.com/bib. Thus, the overall results clearly indicate the stable conformational dynamics of the target proteins complexed with antiviral drug, catechin.

Another parameter, H-bonds interaction is one of the major players in governing the ligand stability at the active site of the

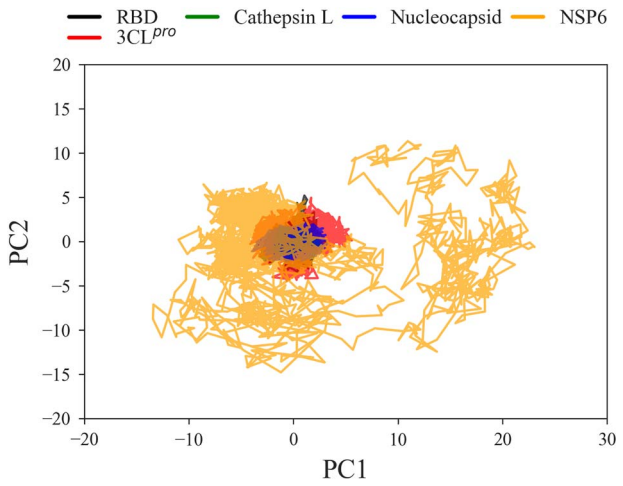


Figure 3. ED plots. The conformational landscape of target proteins complexed with the FDA-approved antiviral drug, catechin. The projection of the collective motion of proteins in the essential subspace along the PCs, PC1 and PC2. The target proteins are represented with different color codes as shown in the schema.

protein. Thus, we further investigated the time evolution plots of H-bonds involved in the molecular interaction of catechin with proteins (Figure 2E). The average occupancy of H-bonds (donor and acceptor), excluding ionic interactions were examined during the last 50 ns of simulation time and are summarized in the supporting information, Supplementary Table S3 available online at https://academic.oup.com/bib. We found the maximum occupancy of five H-bonds between the catechin and RBD; however, four H-bonds remained consistent up to ∼50 ns. Out of these, two H-bonds vanished at ∼60 ns, which were regained at ∼100 ns. Nevertheless, three H-bonds were observed to be stable during the last 50 ns of simulation. The result shows that catechin formed H-bond interaction with residues, Tyr451 (OH) with ligand (H7 and H8) and Asp442 (O) with ligand (H7). 3CL^{pro} showed six H-bond interactions with catechin which could be seen up to ∼150 ns, but only three remained stable during the last 50 ns, which were formed between the residues, Asp187 (O)—ligand (H7), His164 (HE2)—ligand (O2) and Thr26 (O)—ligand (H14). CTSL formed three H-bonds, which were observed consistent up to 170 ns. It showed maximum occupancy with Asp162 (OD2)—ligand (H7), Asp160 (OD1)—ligand (H8) and Asp160 (OD2)—ligand (H8). Three H-bonds were formed between nucleocapsid–catechin, which were observed with Arg46 (HN)—ligand (O1), Asp56 (OD1)—ligand (H7) and Thr44 (HG1)—ligand (O2). NSP6 showed the possibility of three-four H-bonds, however, there were two H-bonds between Thr238 (O)—ligand (H14) and Thr130 (HG1)—ligand (O1), having maximum occupancy during the last 50 ns of simulation. These results indicate that catechin is stabilized by an average of two-three H-bonds at the binding pocket of proteins.

Essential dynamics

Protein function is regulated by switching between various conformations. The modular nature of proteins to switch between various states is governed by the collective motion of protein, which is intrinsic to many biological processes and plays a crucial role in the transmission of biological signals. For a protein to be functional, a reasonable amount of flexibility, as well as rigidity is required, specifically for the residues in the binding site. Essentially, a tighter interaction would restrict the motion of the protein, thereby not allowing it to sample some conformations required for its activity. Therefore, in order to understand the collective motion of protein occupied in the conformational space during the simulation, we applied the dimension reduction method, essential dynamics (ED) analysis by the projection of the first two PCs, PC1 and PC2. The PC1 and PC2 were calculated by diagonalizing the covariance matrix of eigenvectors to define the essential subspace in which most of the protein dynamics occur. The dynamic motion of proteins obtained through the projection of PC1 and PC2 are shown in Figure 3. It is apparent from these plots that the collective motion of proteins, RBD, CTSL and nucleocapsid protein is localized in a small conformational space in comparison to 3CL^{pro} and NSP6, which revealed consistent results corresponding to the structural order analyses, RMSD, R_g and SASA, as described in Figure 2A–C. The well-defined small clusters of RBD, CTSL and nucleocapsid protein clearly indicate the reliability and stability of the complex structure with catechin. The ED plot of 3CL^{pro} displays a slight increase in the conformational phase space, which can be seen along the PC2, which suggests that protein navigated the broad conformational space before achieving the ensemble of the dynamically equilibrated state. Contrary to this, NSP6 experiences a wide region of phase space. In fact, it explored a large conformational space in comparison to the other four proteins, which represent the overall higher flexibility of the protein. Thus, we observed a significantly compact structure of RBD, CTSL and nucleocapsid protein and 3CL^{pro} as compared to NSP6, which may facilitate the vital interactions with catechin.

Free energy landscape

FEL provides an accurate description of the minimum energy conformational ensembles of biomolecules, which is undoubtedly essential to understand the conformational transition underlying protein–ligand interactions [85]. Thus, FEL plot is constructed using Boltzmann inversion ($F = -RT \ln P$), where P is the two-dimensional probability distribution of the first two PCs, PC1 and PC2, as reaction coordinates. Figure 4 shows that the binding of catechin with proteins occurs through the minimum free energy pathway. The structural ensemble derived from FEL shows that the catechin-bound complex with RBD navigated the broad conformational space, clustered in the different energy basins, distributed along the PC1. However, these energy minima separated through the low transition barrier <2.0 kcal/mol indicates that with the small excursion, the ensemble states of RBD can easily move out from one energy basin to another. This may be the reason we observed a small and consolidated cluster of stable populations in the ED analysis. The FEL plot of 3CL^{pro} shows the appearance of two distinct populations confined to two different energy basins, separated with high transition barrier >4.0 kcal/mol, which signifies the population of loosely and tightly ligand bound conformations of the protein. The conformational ensemble occupying the small energy basin represented the population of the equilibration phase, which readily achieved a stable equilibrium. These equilibrated ensembles of stable complex transverse to a broad and deep energy basin. The complex with CTSL shows single but elongated energy minima, which depicts the heterogeneous population of different sub-states, but the very less transition barrier <1.0 kcal/mol between the ensemble states suggested the stable conformation of the protein–ligand complex confined to energy basin interplays between the

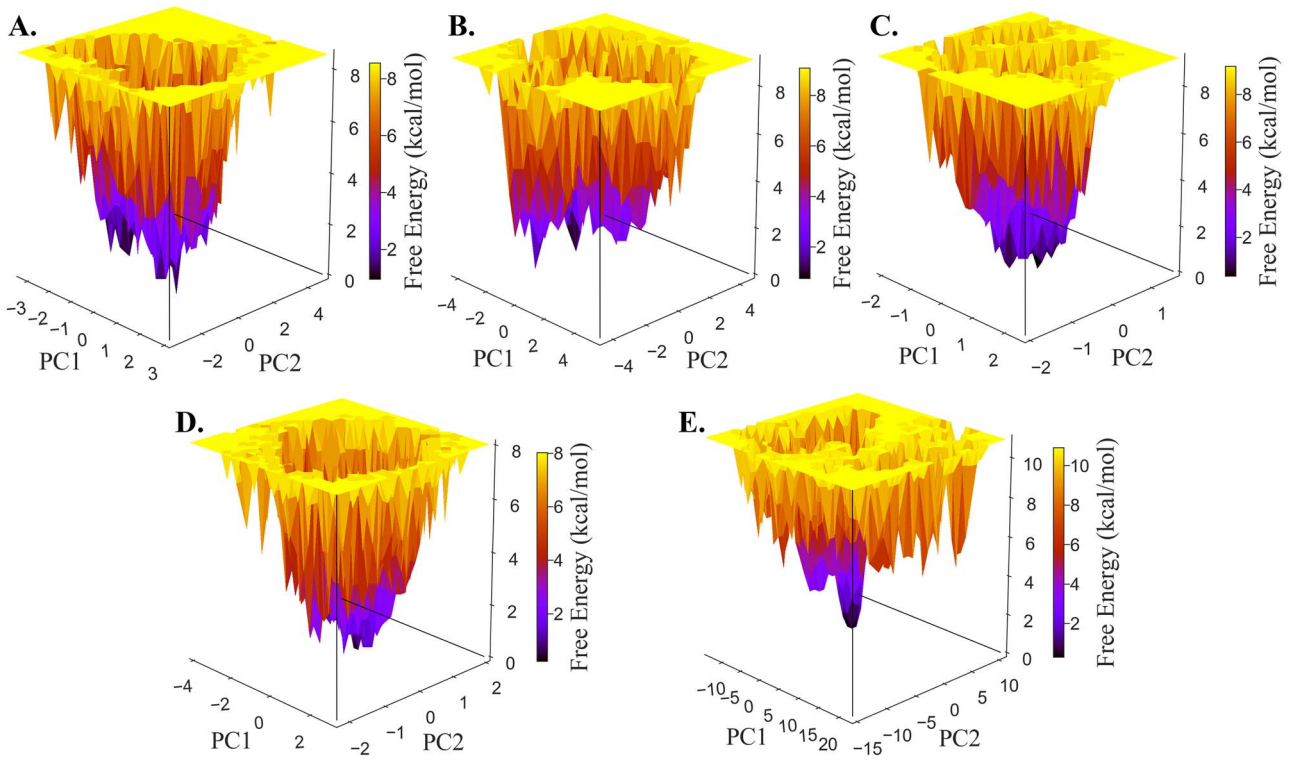


Figure 4. FEL of target proteins complexed with catechin. **(A)** RBD, **(B)** $3CL^{pro}$, **(C)** CTSL, **(D)** nucleocapsid and **(E)** NSP6. The free energy is given in kcal/mol and indicated by the color code, from lower to higher energy in the right panel.

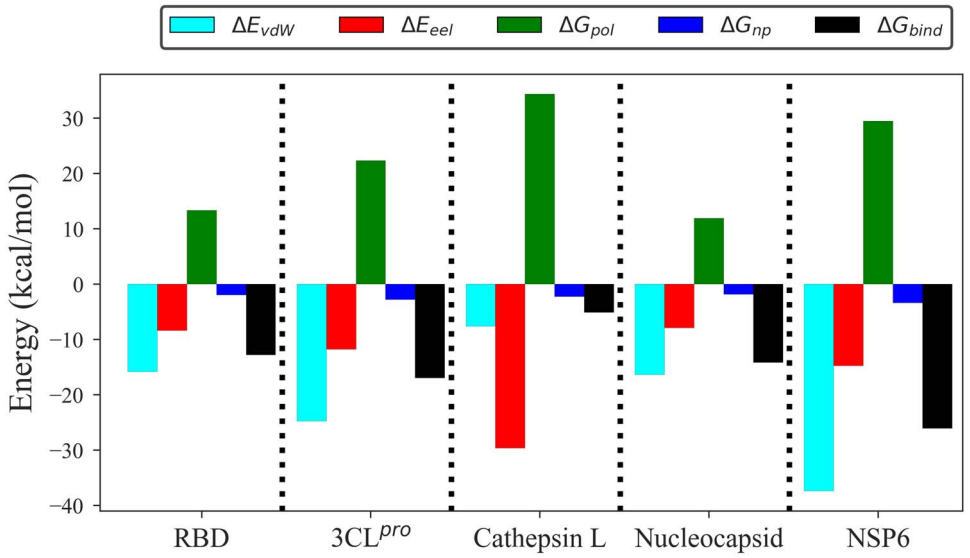


Figure 5. The binding free energy terms obtained from the MM-PBSA calculations relative to the binding of catechin, with five target proteins as labeled in the plot and the color codes for different energy components as shown in the schema.

subspace. Whereas, the rugged FEL with segmented small energy minima of nucleocapsid protein suggested the population of loosely bound complex [86]. The low transition barriers (~1.5 kcal/mol) between the small energy basins indicate a more prolonged equilibration phase of complex structure. During the progression of simulation, the protein underwent structural modifications to accommodate the ligand and adopted a stable conformation; thus, the equilibrated ensemble smoothly shifted to broad and deep energy minima. Remarkably, The FEL of NSP6 shows that the stably bound conformation of the protein–ligand complex is widely populated to a single consolidated energy minimum, which provides the elegant evidence of interactions inducing the stable conformational transition of NSP6–catechin complex. Thus, the comparison of FEL results indicates the different binding stabilities of the catechin-bound protein–ligand complexes.

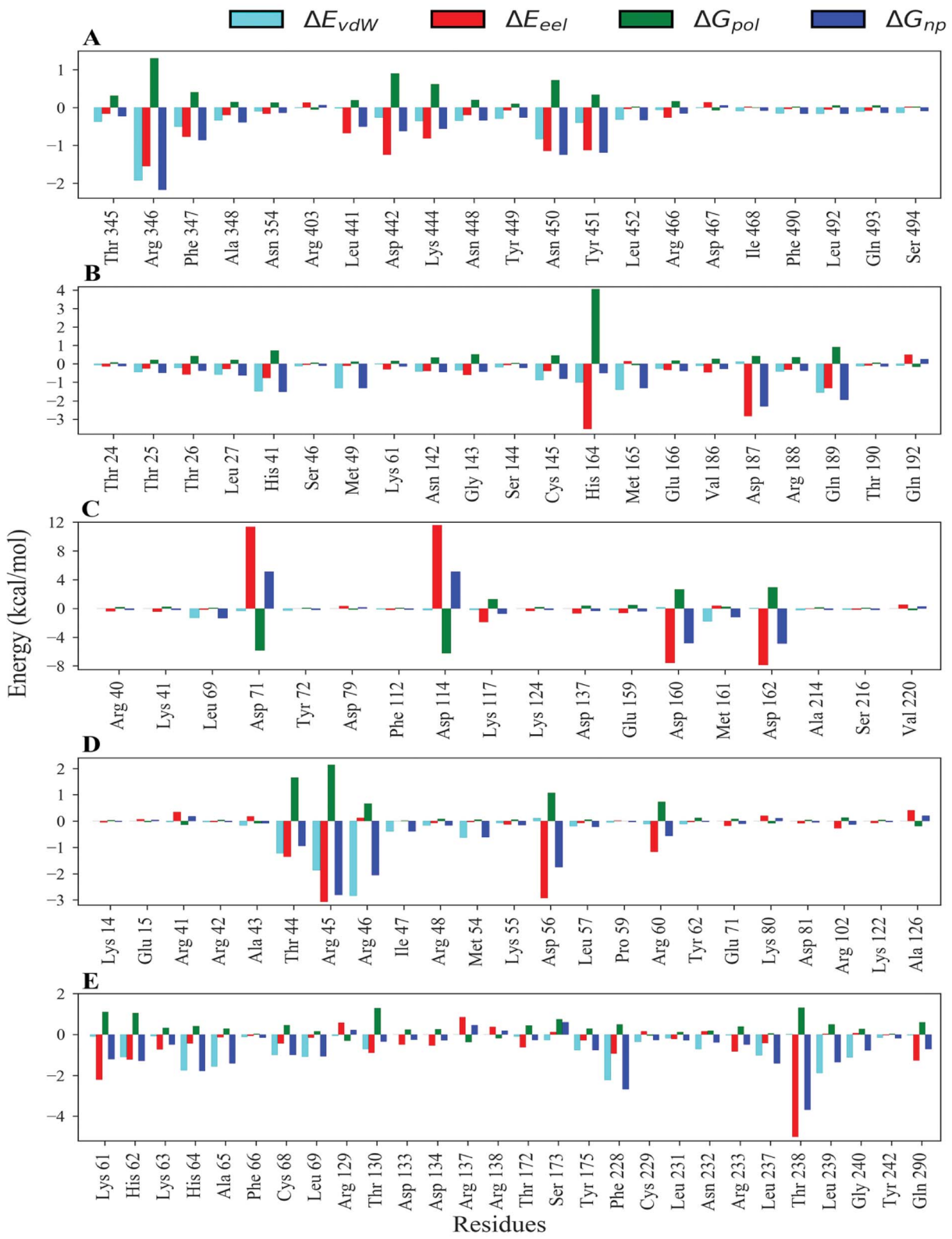


Figure 6. The residue decomposition plot (MM-PBSA) representing the binding energy contribution of the active site residues of five target proteins energetically stabilizing the catechin at binding pockets. **(A)** RBD, **(B)** $3CL^{pro}$, **(C)** CTSL, **(D)** nucleocapsid and **(E)** NSP6.

Binding free energy and ligand–residue interaction decomposition

In order to understand the molecular interaction and stability associated with the binding of catechin to five different proteins of SARS-CoV-2, a detailed analysis of the binding free energy is executed through the MM-PBSA. MM-PBSA provides the best prediction accuracy in terms of energy components of bonded, polar and non-polar solvation free energy, electrostatic and vdW interactions. In addition, it supplies the residue decomposition plot, which helps to probe the contribution of amino acid residues involved in the spatial interaction to stabilize ligands at the binding pocket of the protein. This analysis was performed on the fully converged trajectory of the last 50 ns with a solute dielectric constant value of 2 and an ionic strength of 0.15 M (supporting information, Supplementary Figure S7 available online at https://academic.oup.com/bib). Results show that catechin favorably binds to all five proteins, however, it shows a wide range of the total binding free energies (ΔG_{bind}) as enumerated in the supporting information, Supplementary Table S4 available online at https://academic.oup.com/bib. As illustrated in Figure 5, catechin possesses the highest binding affinity toward NSP6, with a maximum value of ΔG_{bind} = −26.09 kcal/mol, whereas the lowest was toward CTSL (ΔG_{bind} = −5.09 kcal/mol). Results show relatively more favorable contribution of vdW energies (ΔE_{vdW}) –37.39 kcal/mol and −24.76 kcal/mol for NSP6 and $3CL^{pro}$ as compared to nucleocapsid protein (−16.34 kcal/mol) and RBD (−15.8 kcal/mol), respectively. Whereas the less contribution of vdW interaction (−7.59 kcal/mol) is observed for CTSL. Another binding energy component, electrostatic (ΔE_{eel}) energy, which describes ligand–protein interactions, is a critical factor in determining the binding stability of ligand. The interaction of catechin with CTSL shows the major contribution of electrostatic energy −29.61 kcal/mol; however, the lowest values of ΔE_{vdW} = −7.59 kcal/mol and ΔG_{bind} = −5.09 kcal/mol signify that the electrostatic energy contributes relatively less as compared to the other energies in the binding stability of catechin. Thus, the binding free energy analysis revealed that the binding of catechin at the active site of proteins are predominantly stabilized by hydrophobic interactions.

To further quantify the contribution of binding pocket residues to the molecular interaction of catechin with five different proteins, the free energy decomposition per residue was employed (Figure 6). The plot of free energy decomposition analysis shows that the active site residues, Agr346, Phe347, Leu441, Asp442, Lys444, Tyr449, Asn450 and Try451 energetically favor the binding stability of catechin to RBD. Remarkably, it is noted that Agr346 contributed the highest binding free energy, ΔE_{vdW} (−2.10 kcal/mol), ΔE_{eel} (−1.73 kcal/mol) and ΔG_{bind} (−2.10 kcal/mol), which indicated the favorable electrostatic and vdW interactions with catechin (Figure 6A). The protonated (-$NH3^{+}$) Agr346 shows the electrostatic interactions, whereas the side chain guanidinium [-$C(NH_2)_2$] facilitated the hydrophobic interaction with ligand. The binding interaction with $3CL^{pro}$ shows that the amino acid residues, Leu27, His47, Ser46, Met49, His164, Met165, Asp187 and Gln189, contributed the most to the total ΔG_{bind} (−16.98 kcal/mol). Although the vdW interaction primarily stabilizes the catechin at the binding pocket of $3CL^{pro}$, the electrostatic interaction also contributes toward the observed stability by His164 (−3.80 kcal/mol) and Asp187 (−3.10 kcal/mol), respectively (Figure 6B). Indeed, catechin is predominantly stabilized in the binding pocket of CTSL through the electrostatic interaction, which is mostly contributed by the residues Asp71 (−11.31 kcal/mol), Asp114 (−11.36 kcal/mol), Asp160 (−7.61 kcal/mol) and Asp162 (−7.89 kcal/mol), respectively (Figure 6C). Figure 6D showing the free energy decomposition plot of nucleocapsid protein indicates the substantial contribution of amino acids, Thr44, Arg45, Arg46, Asp56, Arg60 and Tyr62 to energetically hold catechin at the binding pocket. Surprisingly, it is noted that Arg45 contributed to both electrostatic (−3.07 kcal/mol) and vdW interaction (−1.88 kcal/mol), but Arg46 contributed only vdW interaction (−2.83 kcal/mol). This may be the reason we observed a moderate range of total binding energy (ΔG_{bind}) of value: −14.15 kcal/mol. The favorable binding of catechin with NSP6 shows the significant contribution of residues, Lys61, His62, His64, Ala65, Asp133, Asp134, Phe228, Leu237, Thr238, Leu239 and Gln290 (Figure 6E). Interestingly, it is noted that Thr238 contributed the higher electrostatic energy (−5.0 kcal/mol), whereas the maximum vdW energy (−2.22 kcal/mol) was contributed by Phe228. However, the binding pocket of NSP6 mostly consists of hydrophobic residues; thus, we observed the major collective contribution of vdW energy for stabilizing the ligand interaction.

Conclusion

In summary, using molecular docking and classical MD simulation, we have explored the possibility of 75 FDA-approved antiviral drugs for their potential of being used as an effective therapeutic strategy to control SARS-CoV-2 infections. The virtual screening results showed that seven therapeutic agents, ritonavir, dolutegravir, tenofovir, tinofoviralafenamide, boceprevir, catechin and zanamivir, could efficiently bind to the SARS-CoV-2 proteins: $3CL^{pro}$, RdRp, ACE2, CTSL, NSP6, nucleocapsid protein and RBD of spike (S) protein, respectively. Of these, catechin has the potential to act as a multi-targeted agent, as it has the highest binding affinity toward the five crucial proteins of the virus, RBD, CTSL, nucleocapsid protein, $3CL^{pro}$ and NSP6, which are essential for the invasion and infection of the host cell. Further, MD simulation, FELs and binding free estimation of catechin with the five target proteins explained the stable interactions of catechin with the critical residues in terms of occupancy of H-bonds and residue contributions to the binding free energy. Thus, our investigation bestowed promising multi-targeted agent catechin, which can be explored as an effective therapeutic agent against the SARS-CoV-2 virus to control the COVID-19 pandemic.

Key Points

- Virtual screening-based repurposing of FDA-approved antiviral drugs for identifying the multi-targeted agent against SARS-CoV-2.
- Catechin is identified as an effective multi-targeted agent.
- Out of seven target proteins, it shows higher binding affinity with $3CL^{pro}$, CTSL, RBD of S protein, NSP6 and nucleocapsid protein.
- Molecular interactions were evaluated through MD simulations, FELs and binding free energy estimations.

Supplementary Data

Supplementary data are available online at *Briefings in Bioinformatics*.

Funding

Science and Engineering Research Board (SERB), Government of India (YSS/2015/000228/LS); National Research Foundation (NRF) of Korea (2019R1F1A1057386 and 2019H1D3A2A0210 2142), supported by Korean government, The Ministry of Science, ICT and Future Planning (MSIP).

Conflict of interest

Authors declare there is no conflict of interest.

Author's Contributions

Conceptualization was by C.B.M., A.P. and R.J.; methodology and analysis were taken care of by A.P., C.B.M., P.P., R.D.S., R.K.M. and R.J.; funding acquisition was by R.J.; supervision was done by A.P., R.J. and C.B.M. and writing, review and editing were by A.P., C.B.M., R.J., P.P., R.D.S., R.P. and A.M.L.

References

1. Zhu N, Zhang D, Wang W, *et al*. A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med* 2020;**382**:727–33.
2. Gorbalenya AE, Baker SC, Baric RS, *et al*. The species severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat Microbiol* 2020;**5**:536–44.
3. World Health O. *Coronavirus disease 2019 (COVID-19): situation report, 72*., Geneva: World Health Organization, 2020.
4. Ul Qamar MT, Alqahtani SM, Alamri MA, *et al*. Structural basis of SARS-CoV-2 3CL(pro) and anti-COVID-19 drug discovery from medicinal plants. *J Pharm Anal* 2020;**10**:313–9.
5. Chan JF, Kok KH, Zhu Z, *et al*. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. *Emerg Microbes Infect* 2020;**9**:221–36.
6. Hoffmann M, Kleine-Weber H, Schroeder S, *et al*. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* 2020;**181**:271–280.e278.
7. Ou X, Liu Y, Lei X, *et al*. Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nat Commun* 2020;**11**:1620.
8. Cascella M, Rajnik M, Cuomo A, *et al*. *Features, Evaluation, and Treatment of Coronavirus*. StatPearls Publishing 2020.
9. Naik B, Gupta N, Ojha R, *et al*. High throughput virtual screening reveals SARS-CoV-2 multi-target binding natural compounds to lead instant therapy for COVID-19 treatment. *Int J Biol Macromol* 2020;**160**:1–17.
10. Bhowmik D, Nandi R, Jagadeesan R, *et al*. Identification of potential inhibitors against SARS-CoV-2 by targeting proteins responsible for envelope formation and virion assembly using docking based virtual screening, and pharmacokinetics approaches. *Infect Genet Evol* 2020;**84**:104451.
11. Dai W, Zhang B, Jiang X-M, *et al*. Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. *Science* 2020;**368**:1331–5.
12. Alamri MA, Tahir ul Qamar M, Mirza MU, *et al*. Pharmacoinformatics and molecular dynamics simulation studies reveal potential covalent and FDA-approved inhibitors of SARS-CoV-2 main protease 3CLpro. *J Biomol Struct Dyn* 2020;1–13.
13. Gupta MK, Vemula S, Donde R, *et al*. In-silico approaches to detect inhibitors of the human severe acute respiratory syndrome coronavirus envelope protein ion channel. *J Biomol Struct Dyn* 2020;1–11.
14. Yin W, Mao C, Luan X, *et al*. Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. *Science* 2020;**368**:1499–504.
15. Barh D, Tiwari S, Andrade BS, *et al*. Potential chimeric peptides to block the SARS-CoV-2 spike receptor-binding domain. *F1000Research* 2020;**9**:576.
16. Tai W, He L, Zhang X, *et al*. Characterization of the receptor-binding domain (RBD) of 2019 novel coronavirus: implication for development of RBD protein as a viral attachment inhibitor and vaccine. *Cell Mol Immunol* 2020;**17**:613–20.
17. Phan T. Genetic diversity and evolution of SARS-CoV-2. *Infect Genet Evol* 2020;**81**:104260.
18. Cao Y, Li L, Feng Z, *et al*. Comparative genetic analysis of the novel coronavirus (2019-nCoV/SARS-CoV-2) receptor ACE2 in different populations. *Cell Discov* 2020;**6**:1–4.
19. Lu R, Zhao X, Li J, *et al*. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* 2020;**395**:565–74.
20. Joshi RS, Jagdale SS, Bansode SB, *et al*. Discovery of potential multi-target-directed ligands by targeting host-specific SARS-CoV-2 structurally conserved main protease. *J Biomol Struct Dyn* 2020;1–16.
21. Levitzki A, Klein S. My journey from tyrosine phosphorylation inhibitors to targeted immune therapy as strategies to combat cancer. *Proc Natl Acad Sci U S A* 2019;**116**:11579–86.
22. Ismail HM, Barton V, Phanchana M, *et al*. Artemisinin activity-based probes identify multiple molecular targets within the asexual stage of the malaria parasites plasmodium falciparum 3D7. *Proc Natl Acad Sci U S A* 2016;**113**:2080–5.
23. Kumari S, Mishra CB, Tiwari M. Polypharmacological drugs in the treatment of epilepsy: the comprehensive review of marketed and new emerging molecules. *Curr Pharm Des* 2016;**22**:3212–25.
24. Gupta AK, Kumar A, Rajput A, *et al*. NipahVR: a resource of multi-targeted putative therapeutics and epitopes for the Nipah virus. *Database (Oxford)* 2020;**2020**.
25. Shi Y, Zhang X, Mu K, *et al*. D3Targets-2019-nCoV: a webserver for predicting drug targets and for multi-target and multi-site based virtual screening against COVID-19. *Acta Pharm Sin B* 2020;**10**:1239–48.
26. Chodera JD, Mobley DL, Shirts MR, *et al*. Alchemical free energy methods for drug discovery: progress and challenges. *Curr Opin Struct Biol* 2011;**21**:150–60.
27. Prakash A, Luthra PM. Insilico study of the A(2A)R-D (2)R kinetics and interfacial contact surface for heteromerization. *Amino Acids* 2012;**43**:1451–64.
28. Wang E, Sun H, Wang J, *et al*. End-point binding free energy calculation with MM/PBSA and MM/GBSA: strategies and applications in drug design. *Chem Rev* 2019;**119**:9478–508.
29. Williams-Noonan BJ, Yuriev E, Chalmers DK. Free energy methods in drug design: prospects of 'alchemical perturbation' in medicinal chemistry. *J Med Chem* 2018;**61**:638–49.
30. Luthra PM, Kumar R, Prakash A. Demethoxycurcumin induces Bcl-2 mediated G2/M arrest and apoptosis in human glioma U87 cells. *Biochem Biophys Res Commun* 2009;**384**:420–5.
31. Luthra PM, Prakash A, Barodia SK, *et al*. In silico study of naphtha [1, 2-d] thiazol-2-amine with adenosine a 2A

receptor and its role in antagonism of haloperidol-induced motor impairments in mice. *Neurosci Lett* 2009;**463**: 215–8.
32. Sadegh S, Matschinske J, Blumenthal DB, *et al*. Exploring the SARS-CoV-2 virus-host-drug interactome for drug repurposing. *Nat Commun* 2020;**11**:3518.
33. Levin JM, Oprea TI, Davidovich S, *et al*. Artificial intelligence, drug repurposing and peer review. *Nat Biotechnol* 2020;**38**:1127–31.
34. De Clercq E, Li G. Approved antiviral drugs over the past 50 years. *Clin Microbiol Rev* 2016;**29**:695–747.
35. Gordon DE, Jang GM, Bouhaddou M, *et al*. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* 2020;**583**:459–68.
36. Kang S, Yang M, Hong Z, *et al*. Crystal structure of SARS-CoV-2 nucleocapsid protein RNA binding domain reveals potential unique drug targeting sites. *Acta Pharm Sin B* 2020;**10**:1228–38.
37. Zhang L, Lin D, Sun X, *et al*. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. *Science* 2020;**368**:409–12.
38. Liu T, Luo S, Libby P, *et al*. Cathepsin L-selective inhibitors: a potentially promising treatment for COVID-19 patients. *Pharmacol Ther* 2020;**107587**.
39. Benvenuto D, Angeletti S, Giovanetti M, *et al*. Evolutionary analysis of SARS-CoV-2: how mutation of non-structural protein 6 (NSP6) could affect viral autophagy. *J Infect* 2020;**81**:e24–7.
40. Burley SK, Berman HM, Bhikadiya C, *et al*. RCSB protein data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. *Nucleic Acids Res* 2019;**47**:D464–74.
41. Giroud M, Dietzel U, Anselm L, *et al*. Repurposing a library of human cathepsin L ligands: identification of macrocyclic lactams as potent rhodesain and trypanosoma brucei inhibitors. *J Med Chem* 2018;**61**:3350–69.
42. Gao Y, Yan L, Huang Y, *et al*. Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Science* 2020;**368**:779–82.
43. Pandey P, Prasad K, Prakash A, *et al*. Insights into the biased activity of dextromethorphan and haloperidol towards SARS-CoV-2 NSP6: in silico binding mechanistic analysis. *J Mol Med (Berl)* 2020;1–15.
44. Halgren TA, Murphy RB, Friesner RA, *et al*. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J Med Chem* 2004;**47**(7):1750–9.
45. Friesner RA, Banks JL, Murphy RB, *et al*. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J Med Chem* 2004;**47**:1739–49.
46. Kumar N, Srivastava R, Prakash A, *et al*. Structure-based virtual screening, molecular dynamics simulation and MM-PBSA toward identifying the inhibitors for two-component regulatory system protein NarL of mycobacterium tuberculosis. *J Biomol Struct Dyn* 2019;1–15.
47. Sadowski J, Gasteiger J, Klebe G. Comparison of automatic three-dimensional model builders using 639 X-ray structures. *J Chem Inf Comput Sci* 1994;**34**(4):1000–8.
48. Guillemette G, Poitras M, Boulay G. Two Ca2+ transport systems are distinguished on the basis of their Mg2+ dependency in a post-nuclear particulate fraction of bovine adrenal cortex. *Cell Calcium* 1991;**12**:51–60.
49. Shelley JC, Cholleti A, Frye LL, *et al*. Epik: a software program for pK(a) prediction and protonation state generation for drug-like molecules. *J Comput Aided Mol Des* 2007;**21**: 681–91.
50. Jorgensen WL, Tirado-Rives J. Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *Proc Natl Acad Sci U S A* 2005;**102**:6665–70.
51. Jorgensen WL, Maxwell DS, Tirado-Rives J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc* 1996;**118**:11225–36.
52. Mooers BHM. Shortcuts for faster image creation in PyMOL. *Protein Sci* 2020;**29**:268–76.
53. Laskowski RA, Swindells MB. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model* 2011;**51**:2778–86.
54. Kutzner C, Pall S, Fechner M, *et al*. More bang for your buck: improved use of GPU nodes for GROMACS 2018. *J Comput Chem* 2019;**40**:2418–31.
55. Huang J, Rauscher S, Nawrocki G, *et al*. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat Methods* 2017;**14**:71–3.
56. Vanommeslaeghe K, Raman EP, MacKerell AD, Jr. Automation of the CHARMM general force field (CGenFF) II: assignment of bonded parameters and partial atomic charges. *J Chem Inf Model* 2012;**52**:3155–68.
57. Darden T, York D, Pedersen L. Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems. *J Chem Phys* 1993;**98**:10089–92.
58. Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. *J Chem Phys* 2007;**126**: 014101.
59. Parrinello M, Rahman A. Crystal structure and pair potentials: a molecular-dynamics study. *Phys Rev Lett* 1980;**45**:1196–9.
60. Singh R, Meena NK, Das T, *et al*. Delineating the conformational dynamics of intermediate structures on the unfolding pathway of beta-lactoglobulin in aqueous urea and dimethyl sulfoxide. *J Biomol Struct Dyn* 2019;1–10.
61. Prakash A, Dixit G, Meena NK, *et al*. Elucidation of stable intermediates in urea-induced unfolding pathway of human carbonic anhydrase IX. *J Biomol Struct Dyn* 2018;**36**(9):2391–406.
62. Laberge M, Yonetani T. Molecular dynamics simulations of hemoglobin A in different states and bound to DPG: effector-linked perturbation of tertiary conformations and HbA concerted dynamics. *Biophys J* 2008;**94**:2737–51.
63. Prakash A, Kumar V, Banerjee A, *et al*. Structural heterogeneity in RNA recognition motif 2 (RRM2) of TAR DNA-binding protein 43 (TDP-43): clue to amyotrophic lateral sclerosis. *J Biomol Struct Dyn* 2020;1–11.
64. Prakash A, Kumar V, Meena NK, *et al*. Elucidation of the structural stability and dynamics of heterogeneous intermediate ensembles in unfolding pathway of the N-terminal domain of TDP-43. *RSC Adv* 2018;**8**(35):19835–45.
65. Batt SM, Jabeen T, Bhowruth V, *et al*. Structural basis of inhibition of mycobacterium tuberculosis DprE1 by benzothiazinone inhibitors. *Proc Natl Acad Sci U S A* 2012;**109**: 11354–9.
66. Sastry GM, Adzhigirey M, Day T, *et al*. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J Comput Aided Mol Des* 2013;**27**:221–34.

67. Wang C, Nguyen PH, Pham K, *et al*. Calculating protein-ligand binding affinities with MMPBSA: method and error analysis. *J Comput Chem* 2016;**37**:2436–46.
68. Oprea TI, Bauman JE, Bologa CG, *et al*. Drug repurposing from an academic perspective. *Drug Discov Today Ther Strateg* 2011;**8**:61–9.
69. Talevi A, Bellera CL. Challenges and opportunities with drug repurposing: finding strategies to find alternative uses of therapeutics. *Expert Opin Drug Discovery* 2020;**15**: 397–401.
70. Pushpakom S, Iorio F, Eyers PA, *et al*. Drug repurposing: progress, challenges and recommendations. *Nat Rev Drug Discov* 2019;**18**:41–58.
71. Muramatsu T, Takemoto C, Kim YT, *et al*. SARS-CoV 3CL protease cleaves its C-terminal autoprocessing site by novel subsite cooperativity. *Proc Natl Acad Sci U S A* 2016;**113**: 12997–3002.
72. Jaimes J, Millet J, Whittaker G. Proteolytic cleavage of the SARS-CoV-2 spike protein and the role of the novel S1/S2 site. *SSRN* 2020;3581359.
73. Tan YW, Fang S, Fan H, *et al*. Amino acid residues critical for RNA-binding in the N-terminal domain of the nucleocapsid protein are essential determinants for the infectivity of coronavirus in cultured cells. *Nucleic Acids Res* 2006;**34**(17): 4816–25.
74. Baliji S, Cammer SA, Sobral B, *et al*. Detection of nonstructural protein 6 in murine coronavirus-infected cells and analysis of the transmembrane topology by using bioinformatics and molecular approaches. *J Virol* 2009;**83**: 6957–62.
75. Ramsay RR, Popovic-Nikolic MR, Nikolic K, *et al*. A perspective on multi-target drug discovery and design for complex diseases. *Clin Transl Med* 2018;**7**:3–3.
76. Amaral M, Kokh DB, Bomke J, *et al*. Protein conformational flexibility modulates kinetics and thermodynamics of drug binding. *Nat Commun* 2017;**8**:2276.
77. Horoiwa S, Yokoi T, Masumoto S, *et al*. Structure-based virtual screening for insect ecdysone receptor ligands using MM/PBSA. *Bioorg Med Chem* 2019;**27**:1065–75.
78. Prakash A, Kumar K, Islam A, *et al*. Receptor chemoprint derived pharmacophore model for development of CAIX inhibitors. *J Carcinogen Mutagen* 2013;**5**:1–9.
79. Lionta E, Spyrou G, Vassilatis DK, *et al*. Structure-based virtual screening for drug discovery: principles, applications and recent advances. *Curr Top Med Chem* 2014;**14**:1923–38.
80. Koul A, Arnoult E, Lounis N, *et al*. The challenge of new drug discovery for tuberculosis. *Nature* 2011;**469**:483–90.
81. Wang C, Greene D, Xiao L, *et al*. Recent developments and applications of the MMPBSA method. *Front Mol Biosci* 2017;**4**:87.
82. Xu WW, Huang ZH, Liao L, *et al*. Direct targeting of CREB1 with imperatorin inhibits TGFbeta2-ERK signaling to suppress esophageal cancer metastasis. *Adv Sci (Weinh)* 2020;**7**:2000925.
83. Panda PK, Arul MN, Patel P, *et al*. Structure-based drug designing and immunoinformatics approach for SARS-CoV-2. *Sci Adv* 2020;**6**:eabb8097.
84. Pandey P, Lynn AM, Bandyopadhyay P. Identification of inhibitors against alpha-isopropylmalate synthase of mycobacterium tuberculosis using docking-MM/PBSA hybrid approach. *Bioinformation* 2017;**13**:144–8.
85. Boehr DD, Nussinov R, Wright PE. The role of dynamic conformational ensembles in biomolecular recognition. *Nat Chem Biol* 2009;**5**:789–96.
86. Fu H, Shao X, Cai W, *et al*. Taming rugged free energy landscapes using an average force. *Acc Chem Res* 2019;**52**:3254–64.