Indian Journal of Biochemistry & Biophysics
Vol. 59, November 2022, pp. 1088-1105
DOI: 10.56042/ijbb.v59i11.65886

Multi-level scoring approach to discover multi-targeting potency of medicinal plant phytochemicals against protein targets in SARS-CoV-2 and human ACE-2 receptor

N Arul Murugan[1], Jeyaraj Pandian Chitra[2], Jeyakanthan Jeyaraman[3]* & SM Rajendren[4]

[1]Department of Theoretical Chemistry and Biology, School of Engineering Sciences in Chemistry, Biotechnology and Health, KTH Royal Institute of Technology, S-106 91, Stockholm, Sweden

[2]Department of Biotechnology, Dr. Umayal Ramanathan College for Women, Karaikudi-630 003, India

[3]Department of Bioinformatics, Alagappa University, Karaikudi-630 003, Tamil Nadu, India

[4]CSIR-Central Electrochemical Research Institute, Karaikudi-630 003, Tamil Nadu

Received 25 August 2022; *revised 19 October 2022*

SARS-CoV-2 pandemic has become a major threat to human healthcare and world economy. Due to the rapid spreading and deadly nature of infection, we are in a situation to develop quick therapeutics to combat SARS-CoV-2. In this study, we have adopted a multi-level scoring approach to identify multi-targeting potency of bioactive compounds in selected medicinal plants and compared its efficacy with two reference drugs, Nafamostat and Acalabrutinib which are under clinical trials to treat SARS-CoV-2. In particular, we employ molecular docking and implicit solvent free energy calculations (as implemented in the Molecular Mechanics -Generalized Born Surface Area approach) and QM fragmentation approach for validating the potency of bioactive compounds from the selected medicinal plants against four different viral targets and one human receptor (Angiotensin-converting enzyme 2 -ACE-2) which facilitates the SARS-CoV-2 entry into the cell. The protein targets considered for the study are viral 3CL main protease (3CLpro), papain-like protease (PLpro), RNA dependent RNA polymerase (RdRp), and viral spike protein-human hACE-2 complex (Spike:hACE2) including human protein target (hACE-2). Herein, there liable multi-level scoring approach was used to validate the mechanism behind the multi-targeting potency of selected phytochemicals from medicinal plants. The present study evidenced that the phytochemicals Chebulagic acid, Stigmosterol, Repandusinic acid and Geranin exhibited efficient inhibitory activity against PLpro while Chebulagic acid was highly active against 3CLpro. Chebulagic acid and Geranin also showed excellent target specific activity against RdRp. Luteolin, Quercetin, Chrysoeriol and Repandusinic acid inhibited the interaction of viral spike protein with human ACE-2 receptor. Moreover, Piperlonguminine and Piperine displayed significant inhibitory activity against human ACE-2 receptor. Therefore, the identified compounds namely Chebulagic acid, Geranin and Repandusinic acid can serve as potent multi-targeting phytomedicine for treating COVID-19.

Keywords: 3CL Main protease, COVID-19, Molecular docking, Molecular mechanics-generalized born surface area approach, Papain-like protease, QM fragmentation scheme, RNA-directed RNA polymerase, SARS-CoV-2, Spike protein

Based on the pathophysiology of SARS-CoV-2, many existing antiviral drugs such as Ritonavir, Umifenovir, Favipiravir, Oseltamivir, Remdesivir *etc*. and other drugs including Tocilizumab, Azithromycin, Interferon β *etc*., are being tested in COVID-19 therapy[1]. Extensive researches are still in progress to discover effective therapeutics against SARS-CoV-2. Recently Ahmad *et al.* (2020)[2] have reported that the ATP binding site is located between palm and finger subdomains of RdRp and the inhibitors of RdRp of Hepatitis C virus can bind with this site. Considering the existing information about the plant mediated natural products as molecular frameworks for development of potent drugs, we believe that the medicinal plants as well as their secondary metabolites with anti-viral activity could shed light on the development and discovery of potent drugs or leads for COVID-19. Modern synthetic medicines emphasis only on killing the coronavirus but not on increasing the immunity of the host to recover and withstand severe infections[3]. However, phytochemicals in *Withania somnifera* and *Andrographis paniculata* showed both antiviral activity against chikungunya virus and immunomodulatory effects in human which can be used for all infected patients including immuno compromised individuals with minimum side effects. Medicinal plants are used for centuries in treatment of

*Correspondence:
E-mail: jjkanthan@gmail.com

various ailments including viral diseases at all stages of infection, however the computational and experimental validations of their therapeutic effect need to be established. Some medicinal plant phytochemicals are unique against viral infections with both anti-viral and immunomodulatory activities that prevent viral entry into host cells inhibit viral pathogenesis and boost host immune system. Anti-viral efficiency of *Andrographis paniculata* was documented for its therapeutic efficacy against Dengue and Chikungunya infections both in preclinical as well as clinical settings. Reinventing and repurposing of traditional medicinal plants antiviral compounds against SARS-CoV-2 infection is the inevitable approach in the current scenario as medical field is overwhelmed in addressing the current outbreak using synthetic antiviral therapeutics. Recently, Murugan *et al.* (2020)[4] have also demonstrated the antiviral activity of the compound neoandrographolide extracted from *Andrographis paniculata* against four main protein targets in SARS-CoV-2 namely Spike protein and three non-structural proteins such as 3CLpro, PLpro and RdRp which play active role in host cell recognition, viral replication and transcription. Christian *et al.* (2020)[5] have demonstrated that the lead molecule α-Copaene in *Justicia adhatoda* have higher binding affinity towards ACE-2 receptor-spike protein complex and S protein in SARS-CoV-2. Kiran *et al*. (2020)[6] have also reported that the bioactive compounds namely Magnoflorine, 5-Hydroxy-7,8-dimethoxyflavanone, Tinosponone, Cirsimaritin, Chrysoeriol, 6- Methoxygenkwanin, Vasicinone, Quercetin and Luteolin from medicinal plants *Sidaacuta*, *Andrographis paniculata*, *Tinospora cordifolia*, *Plectranthus amboinicus*, *Justicia adhatoda* and *Costus speciosus* displayed stronger binding affinity towards S protein of SARS-CoV-2 with good ADMET properties. These compounds are used in traditional medicines to treat phlegmatic fevers and other fevers with flu-like symptoms. Recent studies have evidenced these compounds could also act as an immuno-modulator by effecting signalling pathway involved in the production of Tumor Necrosis Factor (TNF) and thereby increases immunity and restores respiratory health[7]. Earlier studies have reported that leaves of *Piper betle* displayed both immunomodulatory effects as well as antiviral activity against influenza-A virus infections[8]. *Phyllanthus niruri* is another medicinal herb with high antiviral activity and is used to treat viral infections especially Hepatitis B virus[9]. Pharmacological properties of *Phyllanthus niruri* is due to the bioactive constituents namely Rutin, Phyllanthin, Beta-amylin, Beta-sitosterol, Caffeic acid, Geranin, Quercetin, Niruside and Repandusinic acid[10].

Based on this rationale, we have selected 41 phytochemicals from 16 medicinal plants such as *Zingiber officinale, Piper longum, Syzygium aromaticum, Tragia involucrate, Anacyclus pyrethrum, Cyperus rotundus, Hygrophila auriculate, Sidaacuta, Terminalia chebula, Justicia adhatoda, Costusspeciosus, Coleus aromaticus, Tinospora cordifolia, Clerodendron serratum, Piper betle and Phyllanthus niruri* to evaluate its effects against viral protein targets (PLpro, 3CLpro and RdRp) including binding with human ACE-2 receptor and with the interfacial region of spike protein-ACE-2 receptor complex.

The present study includes homology modelling, molecular docking and molecular dynamics simulation, implicit solvent binding free energy calculation and QM fragmentation approach-based calculations to predict the potent small molecule inhibitors in selected medicinal plants against SARS-CoV-2 targets. Moreover, the results of our study could further provide scientific evidence that the pharmaceutical active compounds in medicinal plants can be used for better therapeutic management of SARS-CoV-2 infections. The study also included selected drugs used under clinical trials for COVID-19 therapy such as Dexamethasone, Acalabrutinib and Nafamostat as reference compounds. Our study will also provide the possible novel strategies for drug repositioning to cure SARS-CoV-2 infections.

Computational methods

In this study we have used multi-level scoring functions involving molecular docking, MM-GBSA and QM fragmentation scheme. Since the force-field based scoring functions as implemented in the molecular docking software such as autodock4.0 for ranking potency of compounds may yield many false positives, we aimed at testing the performance of such a multi-layered scoring approach.[33] In this scheme, the three approaches were employed sequentially *i.e*. molecular docking, molecular mechanics - Generalized Born surface area approach (MM-GBSA) and Quantum mechanics (QM) fragmentation scheme[11,12]. The binding modes were found using molecular docking and the stability of viral targets: ligand complex structures has been studied using

molecular dynamics approach. The trajectories from minimization run and low temperature simulations were used for carrying out MM-GBSA based free energy calculations. Further the representative configurations from the low temperature MD were as well used for computing the binding energies using QM fragmentation scheme.

Retrieval of 3D structures of phytochemicals

The three-dimensional structure for all the phytochemicals (Sesquiphellandrene, Bisabolene, Geranial, Piperine, Piperlonguminine, Eugenol, Caryophyllene, Squalene, Sitosterol, HDF, Lupeol, Betulin, Chebulagic acid, Gallic acid, Vasicinone, Carvacrol, Cirsimaritin, Chrysoeriol, Stigmosterol, Luteolin, Costunolide, Elemol, Tinosponone, Bharangin, Scutellarein, Magnoflorine, Cycleanine, Cyperene, Beta-selinene, Vasicine, Quercetin, Ursolic acid, DDDO, Chevibetol, Allylpyrocatechol, Beta-amylin, Beta-sitosterol, Caffeic acid, Geranin and Phyllanthin) were obtained from Pubchem[13]. The 3D structure of selected standard drugs under clinical trials such as Dexamethasone, Acalabrutinib and Nafamostat) were obtained from the Drugbank[14]. However, for certain compounds only two-dimensional structures are available. Wherever, the 3D structure is not available, we have used Openbabel[15] to generate structures which has routines to generate 3D structure once a SMILES datum or 2D structure is available (in mol format). Before proceeding for molecular docking, all the structures were geometry optimized using semiempirical PM7 level of theory as implemented in Gaussian09[16,17].

Molecular docking

Four different targets from SARS-CoV-2 namely S protein, 3CLpro, PLpro and RdRp and human ACE-2 protein were considered for the docking studies using autodock4.0 software[18]. Except for the PLpro target the three dimensional structures for the remaining targets are available in protein databank[19]. The structures for S protein and 3CLpro were reported very earlier while the structure for RdRp has been reported recently[19]. In particular, for the spike protein the structures for both open form and closed form and in complexation with hACE-2 enzyme have been reported from cryogenic-electron microscopy experiments[20–22]. In the current study, we have used those structures of 3CLpro, spike-protein: ACE-2 complex and RdRp and the corresponding PBD ids are 6LU7, 6LZG and 6M71 respectively. From the complex structure of spike-protein: ACE-2, the model for hACE-2 has been isolated and used for molecular docking studies. For the PLpro, the three-dimensional structure has been proposed based on homology modeling using SWISSMODEL webserver[23]. The template structure for this is from the PLpro of SARS-CoV as in the PDB structure 5Y3E which has a sequence identity of about 83% to that of SARS-CoV-2. In all cases, the binding sites were selected carefully and then the center as well as the size of the gridbox were selected accordingly. In the case of ACE-2 enzyme as well the peptidase domain which is interacting with the spike protein has been considered for the docking and the grid box dimensions were restricted to this region only. In the case of spike protein: ACE-2 complex, the interfacial region was chosen as the binding site for docking study. There are two domains namely S1 and S2 in the spike protein as we discussed above and we have selected the region between receptor binding domain of S1 and mammalian ACE-2 receptor as our target binding site[24,25]. In the case of 3CLpro, the binding site was selected based on the location of inhibitor, N3 which has been co-crystallized with the target protein as in 6LU7. The active site in this target is made of the residues THR24, THR26, PHE140, ASN142, GLY143, CYS145, HIS163, HIS164, GLU166 and HIS172[26]. Further, we have used the apo form of the 3CLpro for carrying out molecular docking. In the case of PLpro, the binding site location has been chosen based on the PLpro-ligand complex of SARS CoV (PDB id is 4OW0)[27]. The binding site in RdRp has been chosen based on the information available in the literature[28]. A recent paper evidenced that the active site of RdRp is highly conserved in many RNA viruses (Hepatitis C and Zika) and in human coronaviruses such as 229E, NL63, OC43, HKU1, SARS and MERS[29]. In particular, the binding site of RdRp was characterized by two residues D760 and D761[30]. In the molecular docking using autodock4.0, the protein was kept rigid while the ligand was fully flexible and the position, orientation and conformational states were varied to identify the least energy binding modes and poses. In all the docking studies, up to 100 least energy binding modes and binding poses were stored for further analysis. The relative binding affinities of phytochemicals along with the trial compounds with the four viral targets and hACE2 target were analysed.

Molecular dynamics simulation and MM-GBSA free energy calculations

The input configurations of the protein-ligand complexes for carrying out molecular dynamics

simulation are based on the least energy (or the one with larger binding affinity) binding mode of the ligands. The ligands included here are the selected phytochemicals and the selected drugs under clinical trial. MD simulations were carried out to study the finite temperature effect and stability of the complexes formed in the ambient condition and this requires the charges and force-fields available for all the subsystems in the complexes to be studied. Therefore, the electrostatic potential fitted charges for the ligands were computed by employing CHELPG approach[31,32] and B3LYP/6-31G* level of theory as implemented in Gaussian09 software. Further, the GAFF force-field[33] has been used to describe the ligand interaction with other subsystems like protein and solvents. Similarly, the FF99SB force-field has been adopted for proteins and for water solvent, TIP3P force-field has been employed. In all the cases, sufficient number of counter-ions were added to neutralize the system. In case of any hot spots present in the protein or in the complex, minimization run was performed, followed by molecular dynamics simulation in a constant volume ensemble. Further simulations in the isothermal isobaric ensemble have been carried out to study the system stability in low temperature (30 K and 1 atm pressure) and ambient condition (300 K and 1 atm pressure). Initially, a low temperature simulation at 30 K was carried out followed by ambient temperature simulation. The Langevin thermostat and Barendsen barostat were adopted to maintain the temperature and pressure in the simulation[34,35]. The time step for solving Newton's equation of motion was set to 2 fs. The simulations were carried out for a total time scale of 10 ns. For those compounds showed high affinity binding the simulations were carried for a time scale of 100 ns (in two sets) and the binding free energies at 300 K and 1atm were also calculated for these compounds. In particular, top 4-6 high affinity compounds for each targets studied were subjected to such long time scale simulations to illustrate the stability of the complexes. All simulations were carried out using Amber16 software[36]. Time evolution of various properties such as density and energies were checked to ensure that the simulations were completed successfully. Also, the root mean square displacements (RMSDs) and root mean square fluctuations (RMSFs) were analysed to show the stability of the protein ligand complexes and to study the ligand-induced structural changes in targets.

The configurations from trajectory corresponding to 5ns timescale after the production runs were used for calculating binding free energies using molecular mechanics- Generalized Born surface area approach[37]. It has been discussed in the literature that a free energy calculation using the longer trajectories does not improve the accuracy of the binding free energies. It has been observed that instead of computing free energies as an average over long time scale trajectory, the binding free energies obtained as an average over many short time scale MDs yielded better accuracy[38,39]. Further the free energies obtained for the minimum energy configurations are comparable to those obtained as an average over trajectories corresponding to finite temperature simulations. Considering these reports, binding free energies were computed for the minimum energy configurations of protein-ligand complex and using the trajectories corresponding to low temperature simulation. At low temperature, the entropic contributions are insignificant and so the binding enthalpies can be directly compared to the binding free energies. Since the entropic contributions are computationally demanding we have used the binding enthalpies for carrying out the analysis on relative binding affinities of the studied ligands. In the MM-GBSA approach, the binding free energies were computed using implicit solvent method[40] and were estimated as the differences between the free energies of complex and individual subsystems (*i.e*., ligand and proteins when they are not bound). The free energy for each system is generally computed as the sum of van der Waals, electrostatic, polar and non-polar solvation energies as well as change in the entropic contributions due to complex formation is added. Since the calculations of entropies are computationally demanding and for assessing the relative stabilities of different complexes, the entropic contributions are not significant, they were not included in this study. Further by choosing the trajectories from low temperature simulation, where the entropies are less significant, we have avoided the explicit calculation of entropies.

For the selected high affinity compounds the binding free energies were computed using finite temperature MD trajectories to show about the effect of such entropic effects. It is worth recalling that in certain systems the binding free energies computed at ambient conditions showed larger fluctuations and non-convergence and so we used the binding

free energies corresponding to low temperature simulation for our discussion[41]. Since the computed binding free energies are not absolute, the relative free energy differences are used to classify a compound as an inhibitor or non-inhibitor. In order to estimate the relative potencies of phytochemicals compounds with various viral targets, we have also included selected compounds (such as Dexamethasone, Acalabrutinib and Nafamostat) which are being considered in various stages of drug clinical trials against COVID-19.

QM fragmentation scheme for computing the protein-ligand interaction energies

There are many reports in the literature showing the success of free energy calculation methods such as the MM-GBSA or molecular mechanics-Poisson Boltzmann surface area (MM-PBSA) approaches and limited reports are available on the failure of such approaches[42]. So, when the experimental binding affinity data are not available it is recommended to compute the binding free energies with more than one computational approach to further validate the predicted results. Moreover, the relative binding affinities of ligands in different biomolecular targets are usually very difficult to predict because the accuracy in binding free energy required for reliable prediction should be within a few kcal/mol. Since the binding affinities of the phytochemicals and compounds considered for clinical trials with various viral targets are not available from experimental studies, we aimed to validate the binding free energy data with a QM fragmentation-based approach. We have developed an in-house fragmentation scheme, which can fragment the whole protein into individual amino acids with the total protein-ligand interaction energies computed as the sum over the fragment contributions[44,45]. In this scheme, the proteins/ enzymes are cut along the peptide bonds and then each individual amino acid is capped either with hydrogens or with NH-CH3 and CO-CH3 groups. Moreover, it is possible to compute the interaction of dipeptide fragments with ligand so that one can also obtain three body contributions to the interaction energies. Further, the water-ligand interaction energies can be computed as well with an explicit treatment of the solvent. In the case of MM-GBSA approach the solvation energies are computed using an implicit solvent model. The interactions between ligand and residues of protein targets can be computed using different levels of theory including dispersion corrected density functional theory[46], M06-2X and MP2 and it is worth mentioning that these methods are better suited for describing the weak dispersion interactions effectively. In the present study we have computed the interactions of viral targets and ligands using the M06-2X/6-31+G* level of theory. Further the individual residue-wise contributions from each amino acid in viral tar-gets are available and this data have been used to validate the residue-wise decomposition of binding free energies as obtained using MM-GBSA approach.

Results and Discussion

Firstly, we will discuss the stability of the formed protein-ligand complexes and ligand-induced structural changes in the target proteins. Even thoughthese analyses have been carried out for all five targets, we only discuss for the two selected targets. The root mean square displacements (RMSDs) for top five high affinity compounds for both targets 3CLPro and PLPro are given in (Fig. 1A & C). The convergence in RMSDs evidence that the ligands are bound to protein throughout the simulations. However, in the case of certain ligands such as Piperine and Scutellarein, RMSD shows stepwise activities which has to be attributed to two different conformational states of the ligands within the target. In certain cases, the RMSDs do not display any such activity which means that ligand conformations do not change significantly within the target binding sites. Of course, such activity will also be reflected in the binding free energies where the conformationally flexible ligands will show larger fluctuations. The root mean square fluctuations (RMSFs) which is proportional to thermal factor (or B factor) experimentally measurable quantity for the 3CLpro and PLpro in their free forms and when they are bound to ligands (only the high affinity compounds) are shown in (Fig. 1B & D). In the case of 3CLpro, the ligand binding reduces the conformational flexibility of residues in the range 185-197. Similarly, except in the case of piperine, the conformational flexibility of residues in the range 44-54 is reduced further. Notably, these are the residues in the binding site as we will see when discuss about the protein-ligand interaction diagram for selected ligands in 3CLpro target. In the case of PLpro target, the changes in RMSFs due to ligand binding are not so significant. Exceptions were observed in the case of two ligands, Acalabrutinib and Nafamostat where

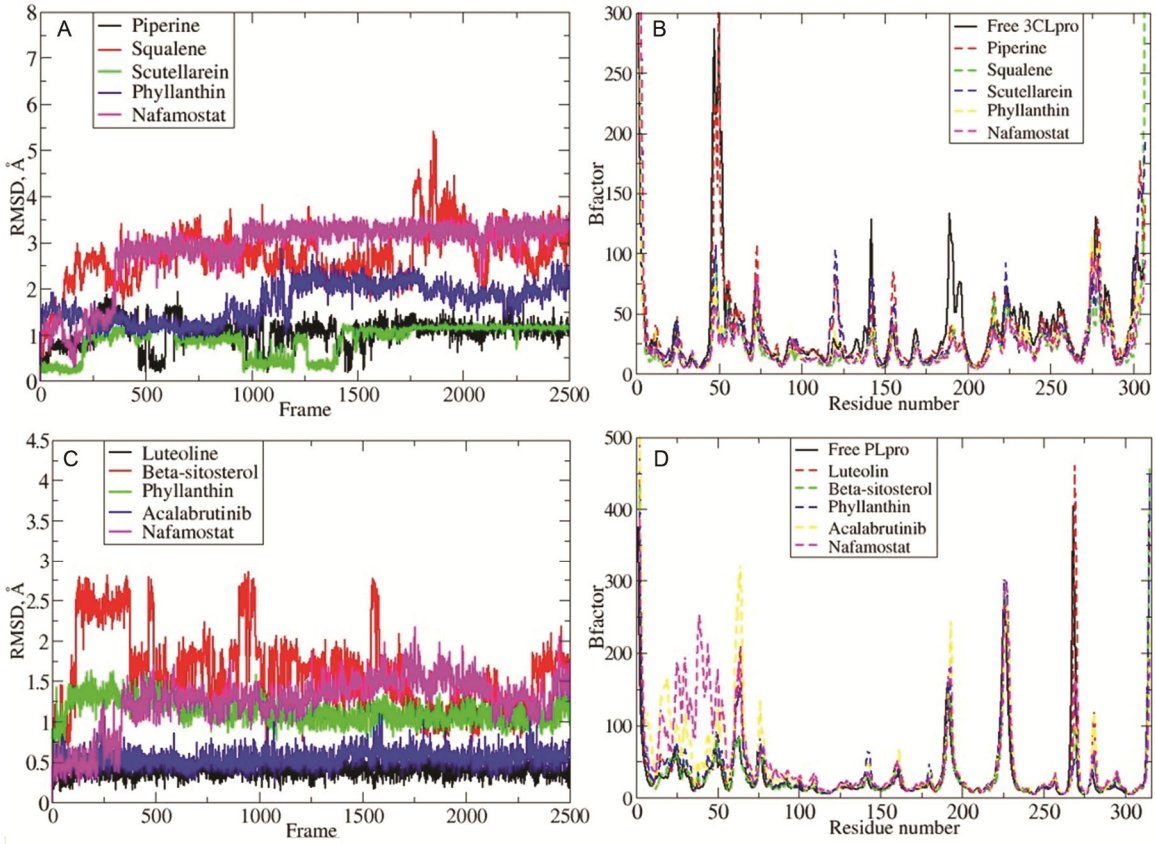

Fig. 1 — (A) RMSDs computed for top five high affinity compounds in 3CLpro target; (B) RMSFs computed for free 3CLpro and when it is bound to selected high affinity compounds; (C) RMSDs computed for top five high affinity compounds in PLpro target; and (D) RMSFs computed for free PLpro and when it is bound to selected high affinity compounds

RMSFs for the residues in the range, 12-71 were significantly increased.

We have also analysed the binding modes and binding affinities computed for various phytochemicals and selected compounds under clinical trials based on molecular docking using autodock4.0 software. There are generally multiple binding sites available in a biomolecular target and for the therapeutic purposes it is essential to target specific binding sites which are involved in the characteristic catalytic functionality. Since, SARS-CoV-2 has a larger genetic overlap to SARS-CoV-1 and MERS viruses, the information about binding sites and the role of different targets is largely available. Therefore, we have also made use of these details in carrying out the molecular docking studies. For example, in the case of 3CLpro, the substrate binding site is available in a cleft between the domains I and II where domain I is made of residues 8-111 and domain II is made of the residues in the range 102-184[47]. The substrate binding site as reported similar to other coronavirus consists of catalytic dyad made of Cys145 and His41 residues[47]. Other key residues in the catalytic site are Phe140, Leu141, Asn142, Gly143, Ser144, Cys145, Met165, Glu166, Gln189 and Thr190.[48] A recent study on the apo form of 3CLpro shows that due to the ligand binding the residues Met49, Leu50, and Met165 were shown to change their conformations.[49] We have also chosen the grid box for molecular docking to include this catalytic site in domain II. The molecular docking study showed that all the ligands studied were able to bind to this target site with binding affinities in the range from μM to pM. The binding free energies computed for the ligands are presented in (Table 1). In this case, the binding free energies are sum of van der Waals, electrostatic, hydrogen bonding and desolvation energies along with the entropic contributions due to rotational bonds. The free energies also include the change in internal energy of the ligand due to binding to the target. The molecular docking has predicted Chebulagic acid as the top-most compound with fM binding affinity (corresponding to the free energy of binding equals to −17.05 kcal/mol) and other high affinity compounds were found to be

Table 1 — The binding free energies computed using molecular docking software, autodock4.0 for phytochemicals and trial compounds with four vital targets of COVID-19 virus and hACE2 receptor. HDF refers to Hydroxy-7, 8-dimethoxy flavanone and DDDO refers to 3-(2,4-dimethoxyphenyl)-6,7- dimethoxy-2,3- dihydrochromen-4-one

Compound	3CLpro	PLpro	RdRp	Spike:hACE2	hACE2
Sesquiphellandrene	–6.55	–6.13	–5.13	–7.84	–4.94
Bisabolene	–6.46	–5.93	–5.16	–6.97	–5.36
Geranial	–4.95	–5.14	–4.17	–5.54	–4.26
Piperine	–7.56	–7.41	–5.90	–7.85	–5.60
Piperlonguminine	–6.92	–6.76	–5.80	–7.71	–5.71
Eugenol	–5.30	–5.39	–4.78	–6.09	–4.31
Caryophyllene	–5.55	–5.78	–4.79	–6.27	–4.82
Squalene	–8.15	–7.88	–5.57	–7.83	–5.60
Sitosterol	–9.50	–10.13	–7.86	–10.66	–6.35
HDF	–7.41	–6.77	–6.27	–7.09	–5.14
Lupeol	–7.72	–7.96	–7.75	–6.67	–6.02
Betulin	–7.50	–7.85	–7.43	–6.98	–5.48
Chebulagic acid	–17.05	–14.68	–15.37	–3.55	–11.88
Gallic acid	–5.79	–6.25	–6.55	–6.37	–4.50
Vasicinone	–5.39	–6.28	–5.10	–6.22	–4.72
Carvacrol	–5.31	–5.33	–4.38	–5.66	–4.37
Cirsimaritin	–7.56	–7.08	–6.46	–8.12	–4.47
Chrysoeriol	–8.64	–7.81	–7.61	–8.15	–5.52
Stigmosterol	–10.01	–11.19	–8.64	–9.40	–6.93
Luteolin	–8.93	–8.10	–8.08	–8.46	–6.05
Costunolide	–6.72	–6.87	–5.41	–6.67	–5.01
Elemol	–5.83	–6.07	–4.97	–7.37	–4.94
Tinosponone	–7.80	–7.23	–6.53	–7.82	–5.32
Bharangin	–7.83	–7.21	–6.63	–8.01	–6.28
Scutellarein	–9.55	–7.53	–7.83	–7.91	–5.34
Magnoflorine	–8.46	–7.24	–7.09	–7.74	–6.09
Cycleanine	–9.48	–8.33	–6.65	–6.68	–6.12
Cyperene	–6.07	–5.91	–4.77	–7.10	–5.17
Beta-selinene	–6.11	–6.47	–4.80	–6.72	–5.16
Vasicine	–5.58	–6.03	–5.48	–6.05	–4.51
Quercetin	–8.84	–8.34	–8.13	–8.91	–5.98
Ursolic acid	–7.99	–7.59	–7.29	–6.88	–5.37
DDDO	–6.85	–6.98	–6.38	–8.26	–4.82
Chevibetol	–5.31	–5.38	–5.04	–6.44	–4.12
Allylpyrocatechol	–5.55	–6.04	–5.75	–6.94	–5.48
Beta–amylin	–7.37	–8.17	–7.58	–7.27	–6.05
Beta–sitosterol	–9.84	–10.20	–7.96	–9.97	–6.91
Caffeic acid	–6.06	–6.24	–7.36	–6.59	–4.79
Geranin	–13.76	–13.62	–13.18	+7.22	–12.29
Phyllanthin	–5.99	–5.68	–5.10	–6.52	–3.78
Repandusinic acid	–15.28	–17.24	–20.62	–0.37	–13.57
Dexamethasone	–8.53	–8.49	–7.22	–8.02	–6.23
Acalabrutinib	–7.21	–8.51	–6.56	–8.83	–5.72
Nafamostat	–9.99	–9.74	–9.27	–10.32	–6.79

Repandusinicacid, Stigmosterol, Geranin and Nafamostat. In the case of PLpro, the binding site has been chosen appropriately and the computed binding affinities range were similar to that of the target, 3CLpro (i.e. μM to pM). In this case, Repandusinic acid has been predicted to be the topmost inhibitor and other high affinity compounds were found to be Chebulagic acid, Stigmosterol, Geranin, Beta-sitosterol. For the RdRp target, the target site was chosen as the ATP binding site which had key residues such as Gly616, Trp617, Asp618, Tyr619, Leu7[58], Ser759, Asp760, Asp761, Ala762, Lys798, Tys799, Trp800, Glu811, Phe812, Cis813 and Ser814.[50]

For this target, Repandusinic acid was found to be the best inhibitor followed by other high affinity compounds such as Chebulagic acid, Geranin and

Nafamostat. In the case of spike protein-hACE2 complex, the interfacial region has been targeted as the binding site. There are other binding sites available in the HR1 domain of spike protein which inhibits the HB-6 formation necessary for membrane fusion[51]. However, we aimed at the interfacial site as the binding site in this site will directly modulate the protein-protein interaction between the spike protein and hACE-2 receptor and can have favourable therapeutic effect. The severity of infection is measured by the protein-protein interaction between these two biomolecules[52]. Therefore, any small molecules which can modulate in particular weaken this interaction will reduce the infection. Since the protein-protein interaction between these two targets is the first step in the infection process, the molecules that could bind to this interfacial site will directly modulate this interaction and hence this site is the potential therapeutic target site for COVID-19. With the above considerations, we have also studied the interaction of various phytochemicals in this specific binding site. Except Geranin, all the other phytochemicals showed binding with significant affinity. In this case, the Sitosterol was found to be the topmost inhibitor followed by Nafamostat, beta-sitosterol and Stigmosterol. Even for the hACE-2 receptor, we have targeted the peptidase domain responsible for binding to spike protein in the molecular docking. Many phytochemicals were found to be bind to this domain and remaining compounds bind to the region beneath the interfacial region. For the hACE-2 target, Repandusinic acid showed the superior binding affinity followed by other high affinity compounds like Geranin and Chebulagic acid.

Molecular docking approach has been extensively used for screening small molecules against various targets. There were many success stories reported on the identification of lead compounds for various diseases using this approach[53]. However, it cannot be overlooked that the method has been reported to be underperforming in ranking many complexes and so the reliability of the results needs to be tested using a more accurate binding free energy calculation method. In this way, the molecular mechanics and Generalized Born surface area approach or molecular mechanics Boltzmann-Poisson surface area approach and free energy perturbation approach or quantum mechanics-based QM clustering model, QM/MM approach or QM fragmentation approaches are found to be suitable approaches. In this particular study, we have employed MM-GBSA and QM fragmentation approach to further validate the molecular docking-based scoring.

The binding free energies computed for all the phytochemicals and selected compounds being considered for clinical trials with five different targets are presented in (Table 2). It is evident from the Table 2 that the top most inhibitors predicted using the MM-GBSA approach were different from those set of compounds depicted from the molecular docking approach. In the case of 3CLpro, Nafamostat was the top most inhibitor and Squaline, Phyllanthin, Piperine, Scutellarein and Acalabrutinib were the compounds having the least binding free energies next to this compound. It is very striking to see that the two compounds considered for clinical trial are in the group of topmost inhibitors suggesting that the MM-GBSA method is performing well in identifying the compounds having the binding affinity towards 3CLpro target. However, it is experimentally not established on the targets to which these compounds are actually binding to exert the therapeutic effect. Here, in this study we have shown that the two compounds under clinical trial are the topmost inhibitors of 3CLpro which is one of the essential enzymes required by the virus for the replication. In particular, it is responsible for cleaving the two polyproteins pp1a and pp1b in 11 different sites and the cleaving site is recognized by the sequence, Leu-Gln-Ser-Ala-Gly. The binding site for Nafamostat is shown in (Fig. 2A) which shows that the compound targets the same binding site as the N3, a peptide based irreversible inhibitor. The binding site for the high affinity phytochemical compound, squalene is shown in (Fig. 2B) and it can be deduced that this compound also targets the same substrate binding site as Nafamostat and N3. The protein-ligand interaction diagram computed for Nafamostat and N3 are shown in (Fig 3A & B). It is evidenced that in both cases the residues PHE140, LEU141, ASN142, CYS165, HIS163, HIS164, MET165 and GLN192 of 3CLpro were involved in the interaction with the ligands which suggest again the same that the ligand, Nafamostat targets the same substrate binding site. One striking difference is that the N3 inhibitor had interaction with severalresidues in domain I (in particular with the residues HIS41, THR24, THR25 and THR26) when compared to Nafamostat.

Now we will consider about the inhibition of the enzyme PLpro which is reported to be essential for

Table 2 — The binding free energies computed using MM-GBSA approach for phytochemicals and trial compounds with four vital targets of COVID-19 virus and hACE2 receptor

Compound	3CLpro	PLpro	RdRp	Spike:hACE2	hACE2
Sesquiphellandrene	–25.22	–20.88	–9.38	–7.75	–12.84
Bisabolene	–23.79	–21.07	–15.54	–21.55	–15.55
Geranial	–17.25	–17.59	–12.75	–25.87	–11.52
Piperine	–35.85	–28.17	–24.15	–38.87	–25.80
Piperlonguminine	–25.66	–15.61	–22.20	–26.83	–27.82
Eugenol	–14.10	–17.06	–16.90	–33.51	–11.71
Caryophyllene	–12.94	–12.80	–11.20	> 0	–14.89
Squalene	–42.64	–30.07	–18.15	–9.79	–23.89
Sitosterol	–25.17	–25.61	–1.67	–31.21	–12.05
HDF	–26.79	–22.90	–14.93	–34.86	–23.91
Lupeol	–7.42	–15.49	> 0	> 0	–7.91
Betulin	–10.63	–25.23	> 0	> 0	–9.06
Chebulagic acid	–32.41	–44.82	–30.93	–22.20	–19.71
Gallic acid	–18.76	–17.11	–25.67	–26.46	–7.51
Vasicinone	–16.05	–21.15	–8.51	> 0	–13.23
Carvacrol	–18.23	–16.52	–17.47	–16.95	–9.61
Cirsimaritin	–31.77	–28.73	–18.40	–30.96	–16.10
Chrysoeriol	–33.90	–25.26	–23.10	–42.35	–20.27
Stigmosterol	–29.94	–28.87	> 0	> 0	–16.93
Luteolin	–29.54	–32.80	–19.72	–45.93	–19.25
Costunolide	–24.60	–16.94	–8.03	–25.15	–14.93
Elemol	–14.37	–16.75	–7.55	–13.65	–13.84
Tinosponone	–27.05	–21.17	–4.87	–13.13	–13.00
Bharangin	–21.37	–19.39	–10.80	–37.67	–14.34
Scutellarein	–34.67	–23.10	–18.43	–28.99	–11.09
Magnoflorine	–26.57	–18.75	–10.91	–20.76	–12.74
Cycleanine	–32.67	–18.75	–22.53	–5.71	–11.64
Cyperene	–13.25	–14.65	–11.34	> 0	–8.47
Beta–selinene	–17.35	–18.14	–4.62	–14.51	–17.36
Vasicine	–15.11	–13.29	–8.59	> 0	–6.51
Quercetin	–28.84	–22.60	–27.61	–45.89	–23.11
Ursolic acid	–15.63	–27.29	> 0	> 0	–6.50
DDDO	–24.33	–28.06	–22.43	–20.52	–19.36
Chevibetol	–17.89	–17.92	–21.55	–8.09	–14.57
Allylpyrocatechol	–17.87	–19.40	–19.14	–16.57	–17.73
Beta–amylin	–14.33	–15.01	> 0	> 0	–5.71
Beta–sitosterol	–20.58	–31.56	–7.14	–19.98	–17.85
Caffeic acid	–22.46	–11.54	–19.87	–37.44	–15.67
Geranin	–30.31	–29.23	–32.91	–21.35	–18.92
Phyllanthin	–37.03	–32.74	–27.56	–27.46	–13.10
Repandusinic acid	–28.66	–21.01	–22.98	–58.65	–19.65
Dexamethasone	–20.15	–15.06	–0.59	> 0	–9.75
Acalabrutinib	–34.27	–41.78	–36.46	–37.48	–25.14
Nafamostat	–47.84	–58.73	–34.26	–50.56	–28.77

the cleavage of non-structural proteins 1-3 and is associated with the suppression of host's innate immune response. The results of MM-GBSA binding free energies as reported in (Table 2) suggest that Nafamostat displays the superior inhibition potency (having the least binding free energy of –58.7 kcal/mol) when compared to other studied compounds for the PLpro target. Other high affinity compounds for this target are Chebulagic acid and Acalabrutinib. Again, except Chebulagic acid, other compounds were not predicted to be high affinity compounds using docking energies as computed from Autodock4 for scoring. The binding site for the top two high affinity compounds namely Nafamostat and

Chebulagic acid in PLpro are shown in (Fig. 6A & B). It can be observed that both compounds could bind to the same binding site in PLpro. The PLpro catalytic site comprised of catalytic triad made of residues, Cys111, His272 and Asp286. In the case of Nafamostat, the residues were not involved in the interaction as it is shown in protein-ligand diagram, (Fig. 4C). However, it is evident that number of residues in the region between 264 and 273 were involved in the interaction with this ligand. Further, the Nafamostat has showed hydrogen bonding interaction with at least two residues in the substrate binding site namely GLY163 and ASP302

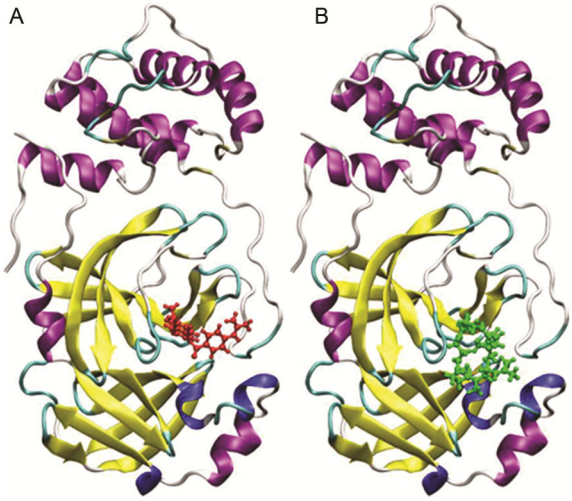

Fig. 2 — (A) Binding sites for Nafamostat compound within 3CLpro; and (B) Binding sites for topmost phytochemical inhibitor within 3CLpro

which could be responsible for the high affinity binding of this ligand.

In the case of RdRp, the MM-GBSA based scoring predicted Acalabrutinib as the top-most inhibitor and the compounds such as Nafamostat, Geranin and Chebulagic acid were the subsequent high affinity compounds for the same target. Here, the results were slightly different when compared to docking-energy based scoring which predicted Repandusinic acid, Chebulagic acid, Geranin, Stigmosterol and Sitosterol as the top high affinity compounds for this target. Figure 7A & B show the binding site for the two high affinity inhibitors namely Acalabrutinib and Geranin which target the Adenosine triphophate (ATP) binding site in the RdRp. The usual mechanism of RdRp inhibitors is that they bind to the same ATP binding site. This enzyme is responsible for the replication and transcription of the genome and utilizes the ATP for this purpose. The occupation of this ATP binding site with any inhibitors could intervene with the RNA synthesis. The ATP was reported to target the binding site made of the three aspartic acid residues namely D618, D760 and D761 and the mutation of the residues at places 760 and 761 to N760 and N761 made this enzyme inactive which establishes the role of these residues for the enzyme activity[54]. In order to show that the compounds are targeting the right binding site, we have computed the protein-ligand diagram for a representative compound, Acalabrutinib and is shown in (Fig. 5C). It can be deduced that the ligand was known to interact

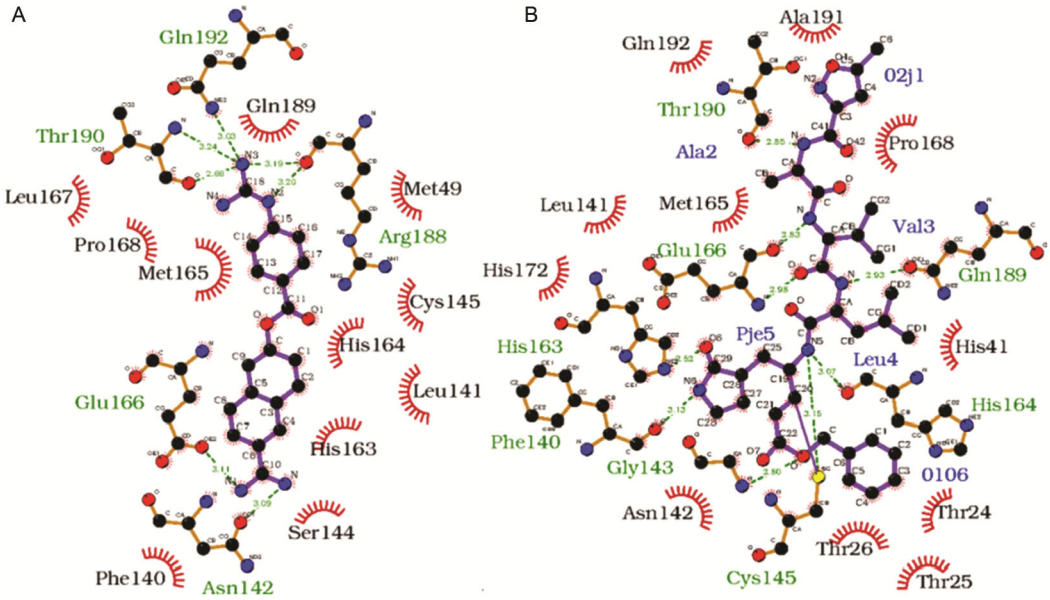

Fig. 3 — (A) Protein-ligand interaction diagram computed for 3CLpro and Nafamostat complex; and (B) Protein-ligand interaction diagram computed for 3CLpro-N3 inhibitor complex based on the crystal structure reported in PDB with reference id 6LU7

with many residues located in this binding site including those key residues namely D618 and D760. Moreover, the compound was engaged in hydrogen bonding interaction with residues ARG624 and THR680 which could be responsible for the high affinity of this compound towards this target.

Table 2 provides the details about potency of molecules that can modulate the interaction between the spike protein and human ACE-2 receptor. Once the hACE-2 receptor and receptor binding domain of the spike protein come in contact, the later protein divides into two subunits. The S1 subunit is made of 14–685 residues and the S2 subunit is made of 686–1273 residues. The S1 subunit is responsible for binding to receptor while the S2 is involved in the host membrane fusion of the virus. In order for the S1 unit to bind to hACE-2, the spike protein needs to be cleaved. The cleavage of spike protein is carried out by proteases such as including cathepsins, transmembrane protease serine protease (TMPRSS2, TMPRSS4) or human airway trypsin-like protease which allows the S1 unit to interact with peptidase domain of ACE2 making the membrane fusion domain exposed. The S1 subunit itself is made of N-terminal domain (14-305 residues) and a receptor-binding domain (319–541 residues). Overall, the RBD of S1 subunit interaction with hACE2 followed by the interaction of HR2 and HR1 domains of S2 subunit to form 6-HB are the major steps involved in the viral infection. So, drugs can be designed to inhibit any of these two steps. One can design compounds that can weaken the interaction between the RBD of S1 unit interactions with hACE-2 or one can develop compounds that will inhibit the formation 6-HB through the intervention of the association process of H1 and H2 of S2 subunit. Here, we have studied how

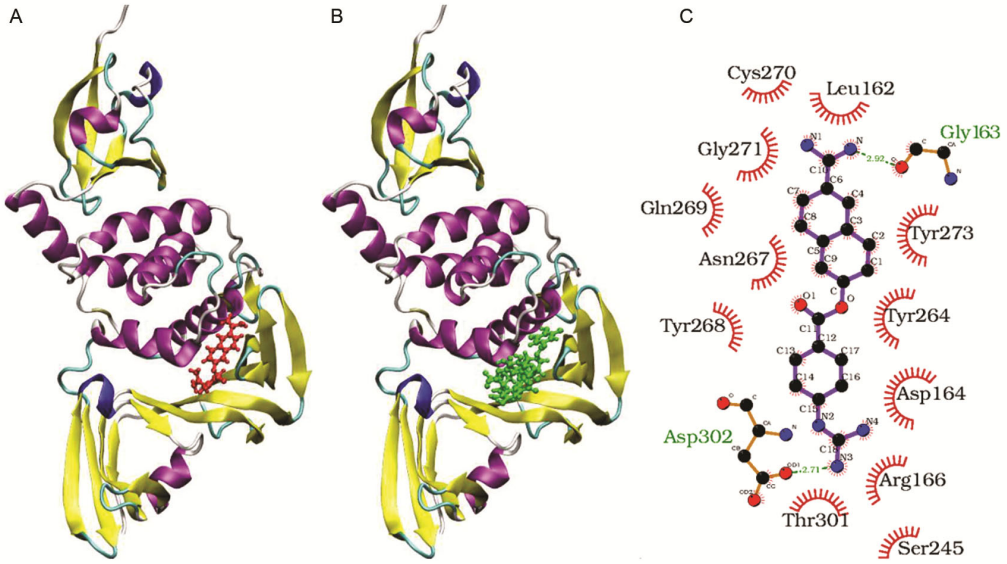

Fig. 4 — (A) Binding sites for Nafamostat within PLpro; (B) Binding sites for topmost phytochemical inhibitor within PLpro; and (C) Protein-ligand interaction diagram computed for PLpro and Nafamostat complex

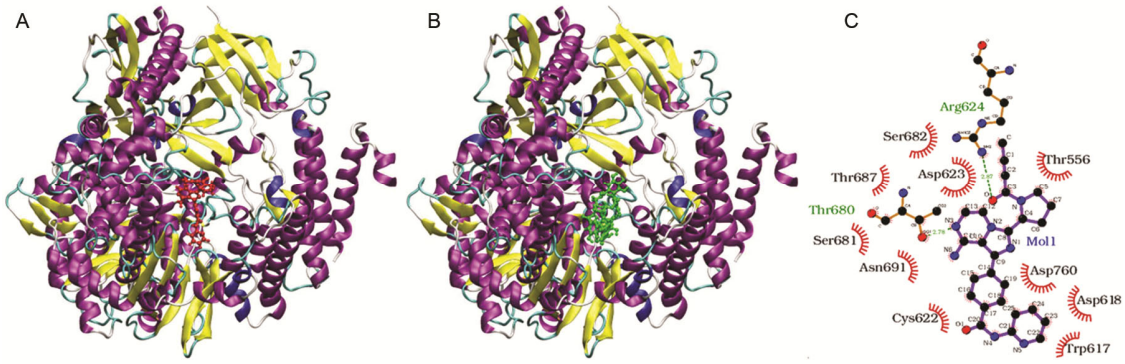

Fig. 5 — (A) Binding sites for Acalabrutinib within RdRp; (B) Binding sites for topmost phytochemical inhibitor, Geranin within RdRp; and (C) Protein-ligand interaction diagram computed for RdRp and Acalabrutinib complex

strong the compounds can bind to the interfacial region of the RBD and hACE-2 protein-protein complex. The interfacial region has shown surface area of 1700 A^2 which stands as a challenge for small organic molecules to span through and to weaken the protein-protein interactions. Previous studies have reported that the dissociation constant (k_D) of RBD and hACE-2 receptor complex was 14.7 nM in SARS-CoV-2 while in the case of SARS-CoV it was only 325.8 nM and this clearly explains why the former class of coronavirus is more sensitive to hACE-2 receptors[44]. So, modulating the protein-protein interaction using small molecules and peptides can lead to active therapeutics development to inhibit viral infection. Based on this, we have estimated the binding free energies of phytochemicals and clinical trial compounds in the binding site located in the interfacial region of the protein-protein complex. The compounds such as Repandusinic acid, Nafamostat, Luteolin, Quercetin and Chrysoeriol showed high affinity in targeting this interfacial binding site. The binding sites for the two high affinity compounds namely Nafamostat and Repandusinic acid are shown in (Fig. 6A & B), respectively. Both the compounds are bound to the interfacial region between the spike protein and hACE-2 receptor and the binding free energies are −50.6 and −58.7 kcal/mol respectively. The protein-ligand interaction diagram for Nafamostat bound to the spike protein: hACE2 complex is shown in (Fig. 6C). It is evident from the results that the ligand interacts with many residues in receptor binding domain of spike protein (such as Lys353, Glu406, Gly496) and a few residues in the peptidase domain of hACE-2 receptor (His34, Glu35, Glu37, Asp38). It is likely that such an uneven interaction with one biomolecule when compared to other could lead to weakening of the protein-protein interaction which is essential to display favourable therapeutic effect. This could be the mechanism behind Nafamostat showing favourable therapeutic effect in inhibiting the viral infection.

Finally, we will discuss about the relative binding affinities of various phytochemicals to hACE-2 receptor. So far, the targets we have discussed are localized on the virus. One can as well target certain biomolecules in the host system and hACE-2 is one such key target to inhibit the viral infection. As it is involved in the recognition of RBD domain of spike protein, targeting this receptor could lead to intervention of viral infection[55]. Since the peptidase domain of the hACE-2 receptor is responsible for binding to spike protein, this region of the hACE-2 receptor needs to be targeted by the small molecules or peptides to show any therapeutic effect. Therefore, the molecular docking studies were carried out by choosing the grid box wisely to include this region. As we see from the results presented in Table 2, the compounds are having relatively lower binding affinity for this target when compared to other targets discussed above. High affinity compounds for this target are found to be Nafamostat, Piperlonguminine, Piperine and Acalabrutinib. The binding sites for the top two high affinity compounds namely Nafamostat and Piperlonguminine are shown in (Fig. 7A & B) which show that the ligands target the expected peptidase domain of the receptor. The protein-ligand diagram for Nafamostat in hACE-2 protein is shown in (Fig. 7C). The hACE-2 peptidase domain (PD) is made of the residues 19 to 615 and C-terminal collectrin-like domain (CLD) is made of residues 616 to 768[56]. The protein-ligand interaction diagram shows that the ligand, Nafamostat interacted with the residues in the PD (such as Asn33, His34, Glu35, Glu37, Asp38, Leu39, Lys353) which is the requirement to intervene the protein-protein interaction to show any therapeutic effect[57]. The chemical structure of the top inhibitors of viral targets and hACE-2 receptor are shown in (Fig. 8A-G) and it

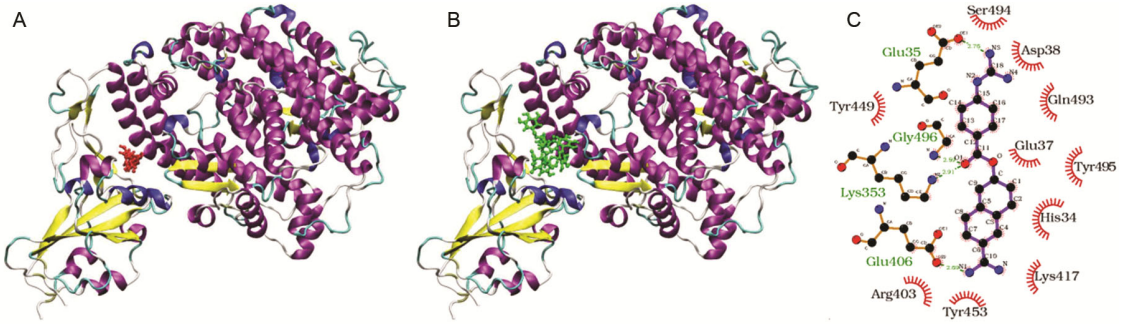

Fig. 6 — (A) Binding sites for Nafamostat within spike protein: hACE2 complex; (B) Binding sites for topmost phytochemical inhibitor, Repandusinic acid within spike protein: hACE2 complex; and (C) Protein-ligand interaction diagram computed for spike protein: hACE2 complex and Nafamostat ligand

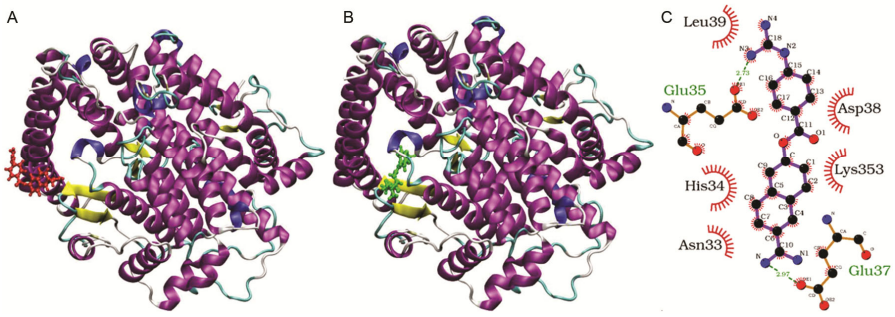

Fig. 7 — (A) Binding sites for Nafamostat within hACE2 receptor; (B) Binding sites for top-most phytochemical inhibitor within hACE2 receptor; and (C) Protein-ligand interaction diagram computed for hACE2 complex and Piperlonguminine ligand

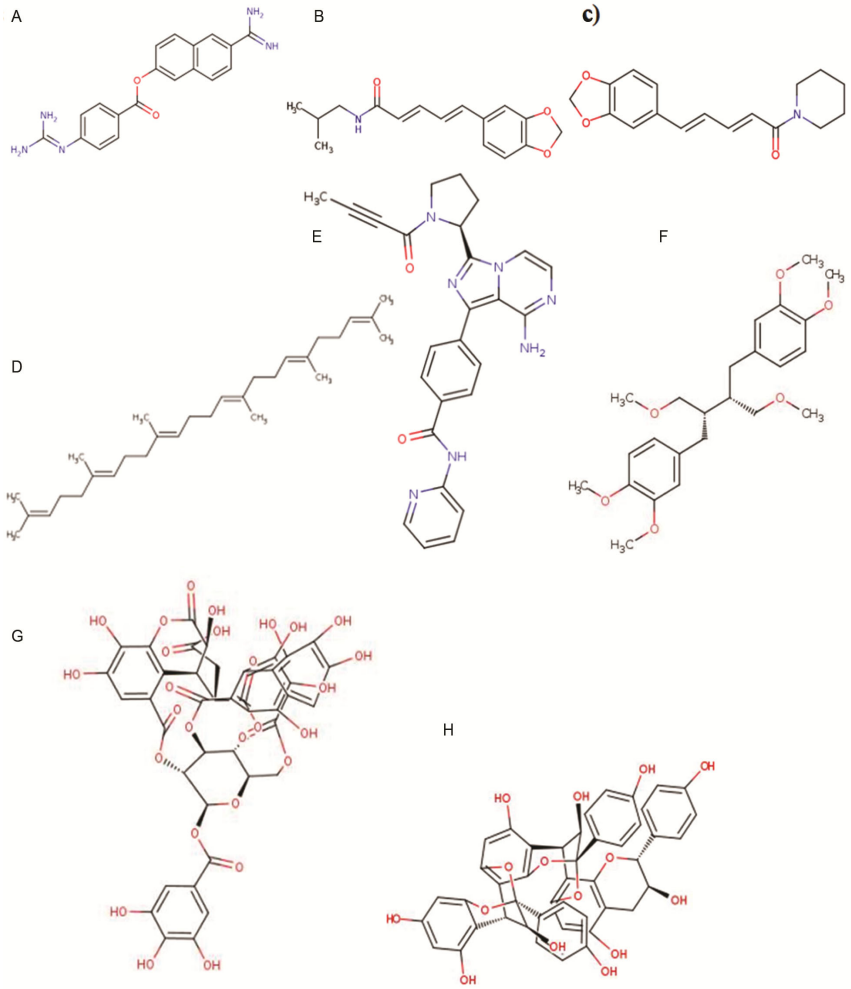

Fig. 8 — Chemical structures of selected high affinity inhibitors (A) Nafamostat; (B) Piperlonguminine; (C) Piperine; (D) Squalene; (E) Acalabrutinib; (F) Phyllanthin; (G) Chebulagic acid; and (H) Geranin-A

is observed that the compounds do not have any structural similarity except the two compounds namely Piperlonguminine and Piperine.

We have previously discussed about the issues related to accuracy of the force-field methods in estimating the binding free energies of organics. In order to rank the complexes, the methods should be able to estimate the binding free energies within a few kcal/mol. Force-field methods had seen certain critically negative remarks when used to rank the complexes of certain targets. Moreover, in our recent study we have shown that the force-field methods

such as molecular docking and MM-GBSA were not ranking reliably the monoamine oxidase-B and inhibitor complexes[58]. Therefore, we aimed to validate the binding free energy results using more reliable approach based on electronic structure theory. However, this theory cannot be straight away adopted for a large length scale systems like protein-ligand complexes and therefore we have employed QM fragmentation scheme for this. The residue-wise contributions as computed from MM-GBSA approach and QM fragmentation scheme for the top two inhibitors of 3CLpro are shown in (Fig. 9A-D). The subplots 9A and 9C correspond to Nafamostat while 9B and 9D correspond to Squalene. It is evidenced from (Fig. 8A & D) where their chemical structures are displayed, the former compound is having a number of polar groups that can form multiple hydrogen bonding with many residues in the catalytic site, refer to (Fig. 4A) while the latter is a hydrophobic molecule. It can be observed from (Fig. 9A & C) (similarly 9B & D) that there is a reasonable agreement between the MM-GBSA and QM fragmentation results on the residues contributing to the stabilization of the protein-ligand complex. In the case of Nafamostat, MET165 was the most dominantly contributing residue for the binding free energies in both cases *i.e.*, MM-GBSA and QM fragmentation results. Also, clusters of residues made of Arg188, Gln189, Thr190 and Phe140, Leu141, Asn142, Gly143, Ser144 and Cys145 were

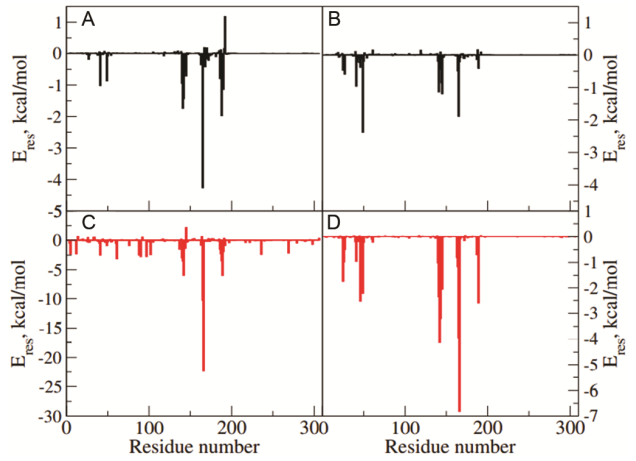

Fig. 9 — (A) Residue-wise contributions from different residues of 3CLpro with ligand Nafamostat from MM-GBSA approach; (B) Residue-wise contributions from different residues of 3CLpro with ligand, squalene from MM-GBSA approach; (C) Residue-wise contributions from different residues of 3CLpro with ligand Nafamostat from QM fragmentation scheme; and (D) Residue-wise contributions from different residues of 3CLpro with ligand, squalene from QM fragmentation scheme

contributing to the stabilization as we see from MM-GBSA (Fig. 9A) which is in concurrence with the QM fragmentation based results (Fig. 9C). It is also worth noting that the QM predicted interaction energies are quantitatively larger in magnitude when compared to MM-GBSA based energies. For example, the contribution from the residue 165 to the binding free energy was –4.3 kcal/mol from MM-GBSA approach while the same corresponds to <–20 kcal/mol QM fragmentation approach and this has to be attributed to the neglect of solvation part in the later approach (QM fragmentation). In the former approach, the free energies were computed for the residue-ligand complex in water solvent where the solvation free energies were accounted by using an implicit solvent model. In the latter approach, this was not the case but the interaction energies were computed for the residues in the vacuum like environment. In the case of the Squalene, comparably similar set of residues are dominantly contributing to the binding free energies as predicted from MM-GBSA (Fig. 9B). The key residues were Hie41 (–1.0 kcal/mol), Ser46 (–0.4 kcal/mol), Met49 (–2.4 kcal/mol), Leu141 (–1.1 kcal/mol), Ser144 (–0.9 kcal/mol), Cys145 (–1.2 kcal/mol), His164 (–0.5 kcal/mol), and Met165 (–1.9 kcal/mol). The residues 41, 46 and 49 contributed with –1.0, –2.5 and –2.2 kcal/mol as predicted from the QM fragmentation-based approach. Similarly, the residues 140, 141, 142, 143 and 145 contributed with –1.0, –2.4, –4.1, –3.2 and –1.2 and –2.1 kcal/mol respectively which are quantitatively larger than the MM-GBSA predicted values. But it is worth reporting that many of the MM-GBSA predicted residues which were dominantly contributing to the stabilization were also the same as predicted from QM fragmentation scheme except that the latermethod underestimates the energetics and as we mentioned which has to be attributed to the neglect of solvent environment. Therefore, such solvent effect needs to be incorporated for the QM fragmentation scheme to predict the energetics with a better accuracy and so this method can serve as a golden standard for validating other binding free energy estimation approaches. The schematic representation of the potential antiviral effects of the selected phytochemicals against SARS-CoV-2 is shown in (Fig. 10).

We also further provide the binding free energies computed using finite temperature MD trajectories and the results are only provided for the selected high affinity compounds (Table 3) for each of the viral

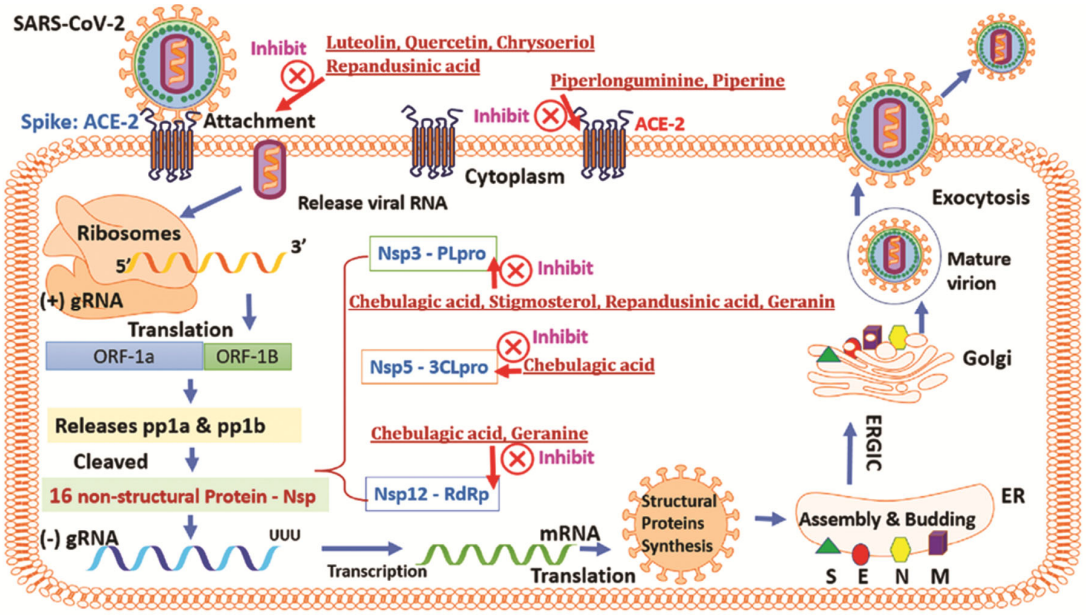

Fig. 10 — Mechanism of multi-targeting potential of selected phytochemicals

Table 3 — The binding free energies computed using MM-GBSA approach for high affinity phytochemicals and trial compounds with four vital targets of Covid-19 virus and hACE2 receptor. The results correspond to 300 K and 1 atm pressure.

Compound	ΔE_{vdw}	ΔE_{elec}	ΔG_{GB}	ΔG_{SA}	$\Delta G_{binding}$
		3CLpro			
Piperine	–35.24	–9.64	18.60	–4.50	–30.80
Squalene	–49.69	–4.26	21.47	–6.47	–38.95
Scutellarein	–33.48	–33.24	40.36	–4.39	–30.74
Phyllanthin	–45.38	–8.80	26.56	–5.90	–33.52
Nafamostat	–34.31	–38.54	49.73	–4.46	–27.58
		PLpro			
Chebulagic acid	–43.77	–43.47	68.23	–5.67	–24.68
Luteolin	–27.14	–37.74	42.12	–3.82	–26.59
Phyllanthin	–32.56	–0.84	16.82	–4.35	–20.93
Acalabrutinib	–42.18	–40.70	49.82	–5.50	–38.56
Nafamostat	–34.61	–82.93	77.49	–5.21	–45.26
		RdRp			
Chebulagic acid	–58.65	–207.36	228.32	–8.30	–45.99
Geranin	–49.24	–95.90	145.20	–7.30	–7.25
Acalabrutinib	–46.35	–32.72	62.63	–5.89	–22.33
Nafamostat	–23.96	–158.26	161.53	–4.39	–25.09
		Spike:hACE2			
Chrysoeriol	–41.82	–29.44	40.35	–5.14	–36.06
Luteolin	–32.58	–45.40	48.70	–4.57	–33.84
Quercetin	–36.06	–46.45	54.15	–5.09	–33.46
Repandusinic acid	–56.82	–212.13	218.14	–9.35	–60.15
Nafamostat	–41.59	–66.61	76.65	–5.40	–36.94
		hACE2			
Piperine	–33.69	–18.44	29.23	–4.51	–27.40
Piperlonguminine	–30.14	–15.61	30.44	–4.24	–19.56
Squalene	–29.42	–2.20	17.75	–3.83	–17.71
Acalabrutinib	–33.24	–20.38	40.32	–3.94	–17.25
Nafamostat	–22.21	–21.51	32.54	–3.19	–14.38

targets studied. In this case along with the total binding free energies, various contributions such as van der Waals, electrostatic, polar and non-polar solvation energies are also presented. As can be seen, even though the magnitude of binding free energies is lowered (when compared to the results presented in Table 2), still these compounds exhibit considerable binding free energies towards their relevant targets. Exception has been observed in the case of Geranin, which has become a compound with lower binding affinity and this has to be attributed to the fact that it tends to interact with larger number of solvent molecules than to the target protein due to many hydroxyl groups present in the molecule. This is also revealed by the highly positive polar solvation free energies which is much larger in magnitude than the electrostatic interaction with protein itself. Overall, the ranking of compounds can be based on the binding free energies computed using the low temperature MD trajectories. However, care should be taken when the ligands are highly flexible and are of polar in nature.

Conclusion

The potency of phytochemicals was studied in comparison with compounds under clinical trials for treating COVID-19 infection. Inhibitors for SARS-COV-2 targets (3CLpro, PLpro, RdRp and spike protein) and hACE-2 were studied using scoring (molecular docking) and implicit solvent binding free energy calculation (MM-GBSA) approaches. Key residues in the catalytic binding sites of 3CLpro, PLpro and RdRp were found to be targeted by the top inhibitors identified in this study. The therapeutic values of selected phytochemicals were also evidenced by its binding to the interfacial region making more contacts with receptor binding domain of spike protein. Moreover, the selected high affinity compounds were found to target the peptidase domain of hACE-2 that is responsible for spike protein recognition. Overall, Chebulagic acid, Geranine and Repandusinic acid act as multitargeting drug-cocktail by effectively inhibiting 3CLpro, PLpro and RdRp targets and also weakening protein-protein interaction between spike protein and hACE-2 as schematically represented in Fig. 10. The present study provides a scientific support validating multi-targeting potency of selected phytochemicals that can make COVID-19 therapy effective since even when one specific viral target mutates, the other protein targets can be inhibited.

Acknowledgement

This work was supported by the grants from the Swedish Infrastructure Committee (SNIC) for the projects “In-silico Diagnostic Probes Design” (snic2020-5-2). JJ gratefully acknowledges MHRD-RUSA 2.0 [F. 24/51/2014-U, Policy (TNMulti-Gen), Dept. of Edn. Govt. of India], DST INDO-TAIWAN (GITA/DST/TWN/P-86/2019 dated: 04/03/2020) and Department of Biotechnology-Bioinformatics Centre (DBT-BIC)-No.BT/PR40154/BTIS/137/34/2021 and TANSCHE (RGP/2019-20/ALU/HECP-0049 dated: 27/04/2021) for funding and infrastructure facilities provided to the Department of Bioinformatics, Alagappa University.

Conflict of interest

All authors declare no conflict of interest.

References

1 Deepak P, Dharmendra T, Atul B, Sagar D, Shanti PC & Manohar P, Global research trends of interleukin-6 in SARS-CoV-2 infection. *Indian J Biochem Biophys*, 59, (2022) 528.

2 Ahmad M, Dwivedy A, Mariadasse R, Tiwari S, Kar D, Jeyakanthan J & Biswal BK, Prediction of Small Molecule Inhibitors Targeting the Severe Acute Respiratory Syndrome Coronavirus-2 RNA-dependent RNA Polymerase. *ACS Omega*, 5 (2020) 18356.

3 Chitra JP, Jeyakanthan J & Rajendran SM, Post-acute sequelae of SARS-CoV-2 Delta variant infection: A report of three cases in a single family. *Indian J Biochem Biophys*, 59 (2022) 777.

4 Murugan NA, Pandian CJ & Jeyakanthan J, Computational investigation on *Andrographis paniculata* phytochemicals to evaluate their potency against SARS-CoV-2 in comparison to known antiviral compounds in drug trials. *J Biomol Struct Dyn*, 39 (2021) 4415.

5 Christian G, Krishnasamy R, Sekar P, Murugan R, Ponnappan S & Ramasamy M, Phytochemical Screening of Adathodai Kudineer a Siddha Concoction and Evaluation of Binding Affinity of Its Constituents with SARS-CoV-2 Spike Protein and ACE2 Receptor Spike Protein Complex Through Molecular Docking *In Silico* Approach. *Int J Ayurvedic Med*, 12(2) (2020) 366.

6 Kiran G, Karthik L, Devi S, Sathiyarajeswaran P, Kanakavalli K, Kumar K & Kumar DR, *In Silico* Computational Screening of Kabasura Kudineer-Official Siddha Formulation and JACOM against SARS-CoV-2 Spike protein. *Int J Ayurvedic Med*, 13(1) (2020) 100324.

7 Pitchiah Kumar M, Sundaram KM & Ramasamy M, Coronavirus Spike (S) Glyco-protein (2019-Ncov) Targeted Siddha Medicines Kabasura Kudineer and Thonthasura Kudineer–*In silico* Evidence for Corona Viral Drug. *Asian J Pharm Res Health Care*, 11 (2019) 1.

8 Saravanan J, Devasia N, Gopalasatheeskumar K, Devan SV, Kokila TK & Sanjay M, Anti-Inflammatory, Antipyretic and Antibacterial Study of Kabasura kudineer choornam. *Int J Curr Adv Res*, 7 (2018) 9992.

9 Venkateswaran P, Millman I & Blumberg BS, Effects of an extract from *Phyllanthus niruri* on hepatitis B and woodchuck hepatitis viruses: *In vitro* and *In vivo* studies. *Proc Natl Acad Sci,* 84 (1987) 274

10 Wahyuni T, Azmi D, Permanasari A, Adianti M, Tumewu L, Widiandani T, Utsubo CA. Widyawaruyanti A, Fuad A & Hak H, Anti-viral activity of *Phyllanthus niruri* against Hepatitis C virus. *Malays Appl Biol*, 48 (2019) 105.

11 Lee MR & Sun Y, Improving docking accuracy through molecular mechanics generalized born optimization and scoring. *J Chem Theory Comput*, 3 (2007) 1106.

12 He X, Zhu T, Wang X, Liu J & Zhang JZ, Fragment Quantum Mechanical Calculation of Proteins and Its Applications. *Acc Chem Res*, 47 (2014) 2748.

13 Ryde U & Soderhjelm P. Ligand-Binding Affinity Estimates Supported by Quantum-Mechanical Methods. *Chem Rev*, 116 (2016) 5520.

14 Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, Han L, He J, He S & Shoemaker BA, PubChem substance and compound databases. *Nucl acids Res*, 44 (2016) D1202.

15 Wishart DS, Knox C, Guo AC, Cheng D, Shrivastava S, Tzur D, Gautam B &Hassanali M, DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nucl acids Res*, 36 (2008) D901.

16 O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T & Hutchison GR. Open Babel: An open chemical toolbox. *J Cheminformatics*, 3 (2011) 33.

17 Promi R, Amy G, Jaffarguriqbal S, Nanette M, Wachter, Jonathan R & Richard S, Solvent Effects on the [3+2] Cycloaddition of 2-Furfural Oxime and Ethyl Propiolate: Unexpected Change in Regioselectivity. *World J Org Chem*, 5 (2017) 6.

18 Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS & Olson AJ, Autodock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J Comput Chem*, 16 (2009) 2785.

19 Berman H, Henrick K, Nakamura H & Markley JL, The worldwide Protein Data Bank (wwPDB): ensuring a single, uniform archive of PDB data. *Nucl acids Res*, 35 (2007) D301.

20 Yin W, Mao C, Luan X, Shen D, Shen Q, Su H, Wang X, Zhou F, Zhao W & Gao M, Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. *Science*, 363 (2020) 1499.

21 Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT & Veesler D, Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell*, 181 (2) (2020) 281.

22 Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh CL, Abiona O, Graham BS & McLellan JS, Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science,* 367 (2020) 1263.

23 Yan R, Zhang Y, Li Y, Xia L, Guo Y & Zhou Q, Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science,* 367 (2020) 1444.

24 Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S, Zhang Q, Shi X, Wang Q, Zhang L & Wang X, Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature*, 581 (2020) 215.

25 Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TA, Rempfer C, Bordoli L, Rosalba L & Torsten S. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res*, 46 (2018) W296.

26 Khaerunnisa S, Kurniawan H, Awaluddin R, Suhartati S & Soetjipto S, Potential Inhibitor of COVID-19 Main Protease (Mpro) From Several Medicinal Plant Compounds by Molecular Docking Study. *Prepr,* (2020) 1.

27 Baez-Santos YM, John SES & Mesecar AD, The SARS-coronavirus papain-like protease: structure, function and inhibition by designed antiviral compounds. *Antivir Res*, 115 (2015) 21.

28 Narayanan N & Nair DT, Vitamin B12 May Inhibit RNA-Dependent-RNA Polymerase Activity of nsp12 from the COVID-19. *Virus*, 72(10) (2020) 2112.

29 Elfiky AA, Anti-HCV, nucleotide inhibitors, repurposing against COVID-19. *Life Sci*, 248 (2020) 117477.

30 Gordon CJ, Tchesnokov EP, Woolner E, Perry JK, Feng JY, Porter DP & Gootte M, Remdesivir is a direct-acting antiviral that inhibits RNA-dependent RNA polymerase from severe acute respiratory syndrome coronavirus 2 with high potency. *J Biol Chem*, 295 (2020) 6785.

31 Breneman C & Wiberg K, Determining atom-centered monopoles from molecular electrostatic potentials: The need for high sampling density in formamide conformational analysis. *J Comput Chem*, 11 (1990) 361.

32 Wang J, Wolf R, Caldwell J, Kollman P & Case, D. Development and Testing of a General AMBER Force Field. *J Comput Chem,* 34 (2004) 1157.

33 Pastor RW, Brooks BR & Szabo A, An analysis of the accuracy of Langevin and molecular dynamics algorithms. *Mol Phys*, 65 (1988) 1409.

34 Berendsen HJ, Postma J, Van Gunsteren, WF, DiNola A & Haak J, Molecular dynamics with coupling to an external bath. *J Chem Phys*, 81 (1984) 3684.

35 Case DA, Betz RM, Cerutti DS, Cheatham TE, Darden TA, Duke RE, Giese TJ, Gohlke H, Goetz AW, Homeyer N, Izadi S, Janowski P, Kaus J, Kovalenko A, Lee TS, LeGrand S, Li P, Lin C, Luchko T, Luo R, Madej B, Mermelstein D, Merz KM, Monard G, Nguyen HT, Nguyen I, Omelyan A, Onufriev DR, Roe A, Roitberg C, Sagui CL, Simmerling WM, Botello-Smith J, Swails RC, Walker J, Wang RM, Wolf X, Wu L & Kollman PA, *Amber 16 Reference Manual* (Covers Amber16 and AmberTools 16) University of California, San Francisco, (2016)

36 Rastelli G, Rio AD, Degliesposti G & Sgobba M, Fast and Accurate Predictions of Binding Free Energies Using MM-PBSA and MM-GBSA. *J Comput Chem*, 31 (2010) 797.

37 Genheden S & Ryde U, How to obtain statistically converged MM/GBSA results. *J Comput Chem*, 31(2010) 837.

38 Su PC, Tsai CC, Mehboob S, Hevener KE & Johnson ME, Comparison of radii sets, entropy, QM methods, and sampling on MM-PBSA, MM-GBSA, and QM/MM-GBSA ligand binding energies of F. tularensisenoyl-ACP reductase (FabI). *J Comput Chem*, 36 (2015) 1859.

39 Genheden S & Ryde U, Will molecular dynamics simulations of proteins ever reach equilibrium? *Phys Chem Chem Phys*, 14 (2012) 8662.

40 Sun H, Li Y, Shen M, Tian S, Xu L, Pan P, Guan Y & Hou T, Assessing the performance of MM/PBSA and MM/GBSA methods, 5. Improved docking performance using high solute dielectric constant MM/GBSA and MM/PBSA rescoring. *Phys Chem Chem Phys*, 16 (2014) 22035.

41 Hou T, Wang J, Li Y & Wang W, Assessing the performance of the MM/PBSA and MM/GBSA methods: II. The accuracy of ranking poses generated from docking. *J Comput Chem*, 32 (2011) 866.

42 Zhang X, Perez-Sanchez HC & Lightstone F, A Comprehensive Docking and MM/GBSA Rescoring Study of Ligand Recognition upon Binding Antithrombin. *Curr Top Med Chem,* 17 (2017) 1631.

43 Murugan NA, Nordberg A & Agren H, Different positron emission tomography tau tracers bind to multiple binding sites on the tau fibril: insight from computational modelling. ACS *Chem Neurosci*, 9 (2018) 1757.

44 Ehrlich S, Moellmann J & Grimme S, Dispersion-corrected density functional theory for aromatic interactions in complex systems. Acc Chem Res, 46 (2013) 916.

45 Jin Z, Du X, Xu Y, Deng Y, Liu M, Zhao Y, Zhang B, Li X, Zhang L & Peng C, Structure of M pro from SARS-CoV-2 and discovery of its inhibitors. *Nature*, 582 (2020) 289.

46 Khan SA, Zia K, Ashraf S, Uddin R &Ul-Haq Z, Identification of chymotrypsin-like protease inhibitors of SARS-CoV-2 *via* integrated computational approach. *J Biomol Struct Dyn*, 39 (2021) 260.

47 Kneller DW, Phillips G, O'Neill HM, Jedrzejczak R, Stols L, Langan P, Joachimiak A, Coates L & Kovalevsky A, Structural Plasticity of the SARS-CoV-2 3CL Mpro Active Site Cavity Revealed by Room Temperature X-ray Crystallography. *Nat Commun*, 11 (2020) 3202.

48 Aftab SO, Ghouri MZ, Masood MU, Haider Z, Khan Z, Ahmad A & Munawar N, Analysis of SARS-CoV-2 RNA-dependent RNA polymerase as a potential therapeutic drug target using a computational approach. *J Transla Med*, 18 (2020) 1.

49 Xia S, Zhu Y, Liu M, Lan Q, Xu W, Wu Y, Ying T, Liu S, Shi Z & Jiang S, Fusion mechanism of 2019-nCoV and fusion inhibitors targeting HR1 domain in spike protein. *Cell Mol Immunol*, 17 (2020) 765.

50 Nguyen HL, Lan PD, Thai NQ, Nissley DA, O'Brien EP & Li MS, Does SARS-CoV-2 Bind to Human ACE2 More Strongly Than Does SARS-CoV? *J Phys Chem B*, 124 (2020) 7336.

51 Meng XY, Zhang HX, Mezei M & Cui M, Molecular docking: a powerful approach for structure-based drug discovery. *Curr Comput Aided Drug Des*, 7(2) (2011)146.

52 Villoutreix BO, Eudes & Miteva MA, Structure-based virtual ligand screening: recent success stories. *Comb Chem High Throughput Screening*, 12 (2009) 1000–1016.

53 Yin C, Genotyping coronavirus SARS-CoV-2: methods and implications. *Genomics*, 112 (2020) 3588.

54 Hillen HS, Kokic G, Farnung L, Dienemann C, Tegunov D & Cramer P, Structure of replicating SARS-CoV-2 polymerase. *Nature*, 584 (2020) 154.

55 Huang Y, Yang C, Xu X, Xu W & Liu SW, Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19. *Acta Pharmacol Sin* 41(2020) 1141.

56 Samavati L & Uhal BD, ACE2, Much More Than Just a Receptor for SARS-COV-2. *Front Cell Infect Microbiol*, 10 (2020) 317.

57 Greenidge PA, Kramer C, Mozziconacci JC & Wolf RM, MM/GBSA binding energy prediction on the PDBbind data set: successes, failures, and directions for further improvement. *J Chem Inf Model*, 53 (2013) 201.

58 Murugan NA, Muvva C, Jeyarajpandian C, Jeyakanthan J & Subramanian V, Performance of Force-Field-and Machine Learning-Based Scoring Functions in Ranking MAO-B Protein–Inhibitor Complexes in Relevance to Developing Parkinson's Therapeutics. *Inter J Mol Sci*, 21 (2020) 7648.