



A computational prediction of SARS-CoV-2 structural protein inhibitors from *Azadirachta indica* (Neem)

Subhomoi Borkotoky  and Manidipa Banerjee

Kusuma School of Biological Sciences, Indian Institute of Technology Delhi, New Delhi, India

Communicated by Ramaswamy H. Sarma

ABSTRACT

The rapid global spread of the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) has created an unprecedented healthcare crisis. The treatment for the severe respiratory illness caused by this virus is primarily symptomatic at this point, although the usage of a broad antiviral drug Remdesivir has been allowed on emergency basis by the Food and Drug Administration (FDA). The ever-increasing death toll highlights an urgent need for development of specific antivirals. In this work, we have utilized docking and simulation methods to identify small molecule inhibitors of SARS-CoV-2 Membrane (M) and Envelope (E) proteins, which are essential for virus assembly and budding. A total of 70 compounds from an Indian medicinal plant source (*Azadirachta indica* or Neem) were virtually screened against these two proteins and further analyzed with molecular dynamics simulations, which resulted in the identification of a few common compounds with strong binding to both structural proteins. The compounds bind to biologically critical regions of M and E, indicating their potential to inhibit the functionality of these components. We hope that our computational approach may result in the identification of effective inhibitors of SARS-CoV-2 assembly.

ARTICLE HISTORY

Received 10 May 2020
Accepted 20 May 2020

KEYWORDS

SARS-CoV-2; natural compound; docking; molecular dynamics simulation; MM-PBSA

Introduction

Natural compounds like plant products constitute a rich resource for drug discovery (Patridge et al., 2016; Thomford et al., 2018; Wani et al., 1971). Anticancer drugs like taxol, and antimalarial compounds like quinine, were derived from plant sources like *Taxus brevifolia* and *Cinchona spp.* (Patridge et al., 2016; Thomford et al., 2018; Wani et al., 1971). Approximately a quarter of FDA or European Medical Agency (EMA) approved drugs are plant based (Thomford et al., 2018), which highlights the importance of plant-inspired compounds in the biomedical arena. Medicinal plants are thought to be a good source for antiviral compounds against multiple viruses including SARS-CoV-2 (Aanouz et al., 2020; Abdelli et al., 2020; Akram et al., 2018; Enmozhi et al., 2020). However, there are major obstacles in generating synthetic drug candidates by mimicking naturally existing compounds. Some common problems include difficulties in extraction of components from original sources and characterization, identifying potential targets, and setting up effective assays for measuring drug efficacy, safety and pharmacokinetics. Recent advances in computational methods can be effective in screening potential targets for newly identified molecules or repurposing licensed drugs.

Neem (*Azadirachta indica*), a member of the *Meliaceae* family, is a well-known medicinal plant, especially in the Indian subcontinent. The chemical constituents of Neem include azadirachtin, 7-desacetyl-7-benzoylazadiradione, 17-

hydroxyazadiradione, 7-desacetyl-7-benzoylgedunin, nimbin, nimbiol, polyphenolic flavinoids, etc (Alzohairy, 2016). Ethanol extracts of Neem leaves have been shown to exhibit anti-microbial properties, and Neem components have demonstrated free radical scavenging and anti-inflammatory activities (Alzohairy, 2016). Neem bark (NBE), when used at concentrations ranging from 50 to 100 µg/mL, has inhibitory effect on Herpes Simplex Virus (HSV) type-1 propagation (Tiwari et al., 2010). Cells treated with NBE inhibited HSV-1 glycoprotein mediated cell-to-cell fusion and polykaryocyte formation, suggesting a potential role of NBE at the viral fusion step. Leaf extract of Neem (*Azadirachta indica* A. Juss.) (NCL-11) has virucidal activity against Coxsackievirus B-4, and is thought to interfere at an early stage of the virus replication cycle (Badam et al., 1999). The proven antimicrobial property, and low toxicity of Neem extracts, makes this plant an excellent choice for harvesting and designing of potential antiviral components. A recent study has shown the beneficial effect of Neem bark on neuroinflammation caused by Mouse Hepatitis Virus (MHV), a murine coronavirus, by preventing cell-to-cell spread (Sarkar et al., 2020).

The novel coronavirus/SARS-CoV-2, like other members of the β -coronavirus family, utilizes several structural and non-structural proteins for receptor binding and cellular entry, replication, assembly and cell-to-cell spread. The Spike (S) glycoprotein, RNA-dependent RNA polymerase (RdRp) and 3C-like Main Protease (3CL^{Pro}) are utilized for cellular entry

and membrane fusion, viral RNA replication and viral poly-protein processing respectively. The 3D structures of these proteins have been resolved by X-ray crystallography or cryo-electron microscopy (Gao et al., 2020; Walls et al., 2020; Zhang et al., 2020). So far, various *in vitro* and computational studies (Bhardwaj et al., 2020; Caly et al., 2020; Elfiky, 2020; Wu et al., 2020; Zhang et al., 2020) have utilized these components as the main targets for drug screening and repurposing from various sources. Several other less explored protein components, for which 3D structures have not yet been resolved, also have essential roles to play in the viral life cycle and pathophysiology, and can be important drug targets. The membrane (M) protein is involved in virus assembly through M-M, M-S, and M-nucleocapsid (N) protein interactions (Arndt et al., 2010; Kuo et al., 2016). The Envelope (E) protein is a homopentameric, short, integral membrane protein of 76–109 amino acids (8.4 to 12 kDa) (Pervushin et al., 2009; Torres et al., 2006). E is a viroporin that appears to assist virus budding through an unknown mechanism (DeDiego et al., 2007). It also interacts with cellular adapter proteins through its C-terminal (PDZ)-binding motif (PBM), which contributes to cell-cell spread and viral pathophysiology (Jimenez-Guardeno et al., 2014; Schoeman & Fielding, 2019; Teoh et al., 2010; Yang et al., 2005). Although assembly of the viral envelope is coordinated by M, both M and E are required for the production and release of particles (Mortola & Roy, 2004). The removal of E protein from SARS-CoV leads to formation of immature or infection-incompetent progeny, indicating the critical role of E in Coronavirus biology (DeDiego et al., 2007).

We attempted docking of Neem compounds with modeled 3D structures of SARS-CoV-2 proteins M and E, which are essential for formation of virus particles. The top scoring compounds with the highest binding affinity were subjected to 100 ns Molecular Dynamics (MD) simulations to detect stability of binding. The results from these *in silico* binding studies are reported. It is possible that these compounds may prevent assembly of SARS-CoV-2 particles, thus reducing viral propagation. A combination of viral replication and assembly inhibitors may be a more effective regimen for therapeutic intervention.

Materials & methods

Target identification and homology modeling

The protein sequences for SARS-CoV-2 E and M proteins were obtained from the NCBI deposits of Wuhan-Hu-1 isolate (YP_009724392.1 and QHD43419.1 respectively). A 3D model for SARS-CoV-2 E was generated from the ITASSER server (Roy et al., 2010) by using the NMR structure of the SARS-CoV E protein as template (PDB ID: 5 × 29) (Surya et al., 2018), which has 88.71% sequence identity with its SARS-CoV-2 analogue. A pentameric form of the model was generated using MODELLER (Eswar et al., 2006). A 3D model for SARS-CoV-2 M was also generated from the ITASSER server. The models were further refined and energy minimized using the 3D refine server (Bhattacharya et al., 2016). 3D models were validated by Ramachandran plot analysis using the

RAMPAGE server (<http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>).

Virtual screening and molecular docking

A natural compound ligand library from the Indian Medicinal Plants, Phytochemistry and Therapeutics (IMPPAT) database (Mohanraj et al., 2018) was used for docking. A set of 70 compounds from Neem (*Azadirachta indica*) was downloaded from the database (Supplementary material Table S1). Virtual screening of the compounds with the 3D models of M and E was carried out using AutoDock Vina (Trott & Olson, 2010) with default configuration parameters. The size of the grid box was chosen to encompass all possible binding sites of E and M. The top-ranked five poses of *Azadirachta indica* compounds bound to each structural protein were used for further analysis.

Molecular dynamics simulation

The top scoring complexes were subjected to MD simulations for further analysis. Simulations were performed using GROMACS 5 with Gromos force field (Oostenbrink et al., 2004). The parameters and topologies of the ligands were calculated by PRODRG server prior to MD simulation (Schüttelkopf & van Aalten, 2004). Each complex was immersed in a cubic box of 1.2 nm containing SPC water molecules. Na^+ and Cl^- ions were added to neutralize the charge of the systems. Energy minimization was performed using the steepest descent method for 50,000 steps for all systems with a tolerance of $1000 \text{ kJ mol}^{-1} \text{ nm}^{-1}$. For long-range interactions, the PME method was used with a 1.2 nm cut-off and a Fourier spacing of 0.16 nm. Equilibrations were carried out for 1 ns for each system with constant number of particles, volume, and temperature (NVT; with modified Berendsen thermostat with velocity rescaling (Bussi et al., 2007) at 310 K and a 0.1 ps time step, Particle Mesh Ewald coulomb type (Kawata & Nagashima, 2001) for long-range electrostatics with Fourier spacing 0.16); and constant number of particles, pressure, and temperature (NPT; Parrinello–Rahman pressure coupling (Martoňák et al., 2003) at 1 bar with a compressibility of $4.5 \times 10^{-5} \text{ bar}^{-1}$ and a 2 ps time constant). Finally, the equilibrated systems were subjected to 100 ns MD simulation with time-steps of 2 fs. Bond-lengths were constrained using the Linear Constraint Solver (LINCS) algorithm (Hess et al., 1997). Quality of receptor-ligand complexes were analyzed using in-built Gromacs utilities.

MM-PBSA binding free energy calculations

The *g_mmpbsa* tool (Baker et al., 2001; Kumari et al., 2014) was employed for calculating the binding free energy of protein–ligand complexes from MD trajectories. Post simulation MM-PBSA (Molecular Mechanics Poisson – Boltzmann Surface Area) binding free energy calculations have been used extensively for screening inhibitors (Khan et al., 2020) and was found to correlate reasonably well with experimental results

(Ekhteiri Salmas et al., 2017; Rajkumari et al., 2018). This module estimates Gibb's free energy of binding using the MM-PBSA method as described by the equations below:

$$\Delta G_{\text{bind}} = \Delta G_{\text{complex}} - (\Delta G_{\text{protein}} + \Delta G_{\text{ligand}})$$

Where, $\Delta G_{\text{complex}}$, $\Delta G_{\text{protein}}$ and ΔG_{ligand} signify the total free energy of the protein–ligand complex and total free energies of the isolated protein and ligand in solvent, respectively.

$$\Delta G_x = \Delta E_{\text{MM}} + \Delta G_{\text{solv}} - T\Delta S_{\text{MM}}$$

$$\Delta G_{\text{solv}} = \Delta G_{\text{polar}} + \Delta G_{\text{nonpolar}}$$

$$\Delta E_{\text{MM}} = E_{\text{bonded}} + E_{\text{nonbonded}} = E_{\text{bonded}} + (E_{\text{elec}} + E_{\text{vdw}})$$

ΔG_x is the free energy for each individual entity. $T\Delta S$ refers to the entropic contribution to the free energy in vacuum where T and S denote the temperature and entropy, respectively. ΔG_{solv} is the free energy of solvation where G_{polar} and G_{nonpolar} are the electrostatic and non-electrostatic contributions to the solvation free energy. The vacuum potential energy, ΔE_{MM} , includes the energy of both bonded as well as non-bonded interactions. The *g_mmpbsa* tool also allows decomposing the total binding free energy into the contribution made by each residue.

Pharmacokinetic parameters prediction

Pharmacokinetic properties assist in the early stages of drug discovery by identifying safety and effectiveness of the compounds. Predicted values of toxicity and some pharmacokinetic parameters were obtained from IMPPAT database which were calculated with *admetSAR* server (Cheng et al., 2012).

Results

Modeling of structural proteins of SARS-CoV-2

The 3D model for SARS-CoV-2 E-protein, generated from I-TASSER server with a C-score of -0.75 , was further refined and energy minimized using 3D refine. The best model was chosen based on the lowest $3D^{\text{refine}}$ (4938.24) and *RWplus* (-10970.54) scores, which indicated a better and more physically realistic model. The model, which contained 65.9% helix, 0% sheet and 34.1% other secondary structure elements, was found to have 92.9% residues in favored region, 5.4% in the allowed region and 1.8% in the outlier region of Ramachandran plot (Supplementary material Figure S2). The 3D model for M-protein obtained from I-TASSER with a C-score $= -3.35$ and was similarly refined and energy minimized, and the best model chosen based on the lowest $3D^{\text{refine}}$ (14559.9) and *RWplus* (-43870.59) scores. The 3D model for SARS-CoV-2 M protein (Figure 1B) contained 43.7% helix, 9% sheet and 47.3% other secondary structure elements, and was found to have 86.4% residues in favored, 10.9% in allowed and 2.7% residues in the outlier regions of Ramachandran plot respectively (Supplementary material Figure S2).

Molecular docking

For both E and M proteins, the top five best scoring compounds in docking studies were selected for further analysis. Details of the docking scores and interactions of each compound with the respective proteins are discussed below.

Interaction with SARS-CoV-2 E

All five best scoring compounds in terms of binding energy (Table 1) displayed strong interaction with the transmembrane and C-terminal domains of the E-protein. Almost all compounds showed hydrophobic interaction with residues from all five monomeric chains in the E-protein pentamer. Nimbolin A, which had the highest binding energy of -11.2 kcal/mol, formed 16 hydrophobic interactions. The other compounds: Nimocin, 7-Deacetyl-7-benzoylgedunin, 24-Methylenecycloartanol and Cycloeucalenone -interacted with the E-protein pentamer with binding energies of -11.0 kcal/mol, -10.9 kcal/mol, -10.8 kcal/mol and -10.5 kcal/mol respectively, with 10 or more hydrophobic interactions in each case (Table 1, Figures 2A and 3A).

Interaction with SARS-CoV-2 M

The best scoring compounds in terms of binding energy (Table 2) formed primarily hydrophobic interactions with the transmembrane and C-terminal domains. Nimocin had the highest binding energy of -10.2 kcal/mol, with 12 hydrophobic interactions. 24-Methylenecycloartan-3-one, Phytosterol, Beta-Amyrin and Nimbolin A had binding energies of -10 kcal/mol, -9.5 kcal/mol, -9.4 kcal/mol and -9.4 kcal/mol respectively with 10, 9, 8 and 17 hydrophobic interactions. Three of the compounds - 24-Methylenecycloartan-3-one, Phytosterol and Beta-Amyrin - formed hydrophobic interactions with Ser 111 and Phe 112 from the conserved region in the C-terminal domain (Table 2, Figures 2B and 3B).

Molecular dynamics simulation analysis

Molecular dynamics simulations upto 100 ns were performed to analyze the stability of the protein-ligand complexes as described before. The stability of the complexes was assessed by Root Mean Square Deviation (RMSD) of protein backbone and ligand. In case of E-protein complexes, the RMSD of the protein backbone was consistent after 50 ns for all complexes. The overall ligand RMSD was also stable in the final 50 ns of simulations for all ligands except Cycloeucalenone, which showed comparatively higher fluctuations (Figure 4A). In case of M-protein complexes, the protein backbone RMSD stabilized after 80 ns. The overall ligand RMSD was stable throughout simulation time in case of all ligands except Phytosterol, where comparatively higher fluctuations were observed initially, which was stabilized after 80 ns (Figure 4B).

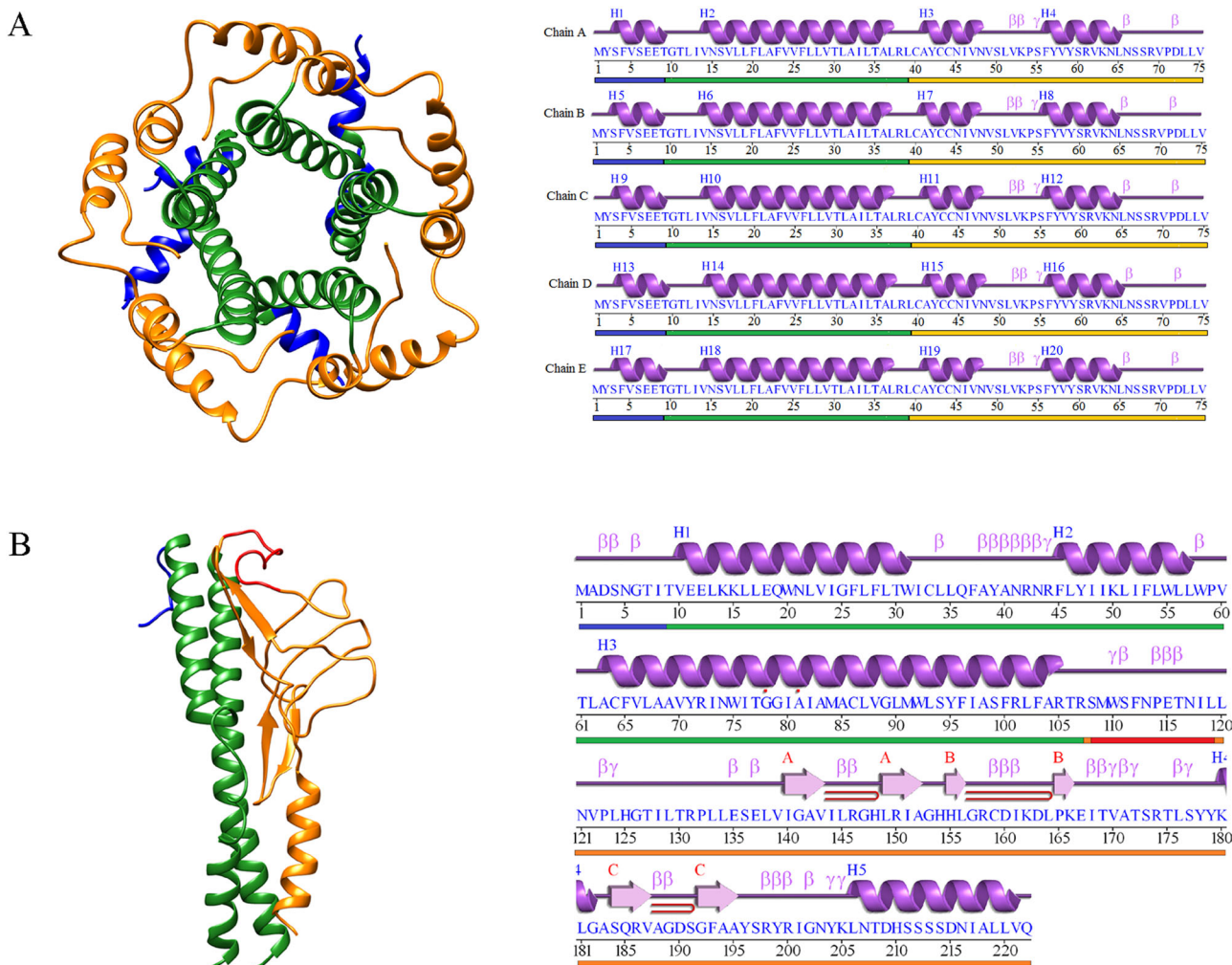


Figure 1. Modeling of SARS-CoV-2 E and M proteins: A) 3D representation of a pentameric form of E-protein, with the N-terminal, transmembrane, and C-terminal domains of each monomer highlighted in blue, green and orange respectively and secondary structure representation is shown on right panel. B) 3D representation of the M-protein, with the N-terminal, transmembrane and C-terminal domains highlighted in blue, green and orange respectively. The C-terminal conserved region is highlighted in red and secondary structure representation is shown on right panel.

MM/PBSA

In order to further quantify the binding affinity between the structural proteins and Neem compounds, the binding free energy and energy contribution of residues were calculated. The energy components E_{MM} , G_{polar} , and $G_{nonpolar}$ of each complex were calculated from 201 snapshots that were extracted at every 0.1 ns from the production trajectories from the final 20 ns. The binding free energy between the protein and inhibitors were decomposed into the contribution of each residue using MM-PBSA approach. Only the residues with contributions greater than -1.5 kJ/mol were shown.

E-protein

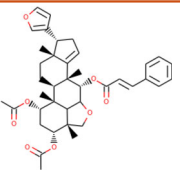
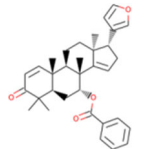
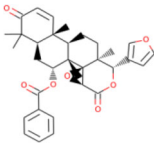
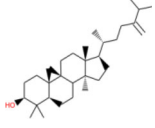
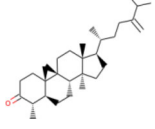
Among the 5 complexes, 7-Deacetyl-7-Benzoylgedunin and Nimbolin A showed the lowest binding free energy values of -297.87 kJ/mol and -297.49 kJ/mol respectively. Nimocin, 24-Methylenecycloartanol and Cycloeucaenone displayed binding free energies of -277.20 kJ/mol, -199.84 kJ/mol and

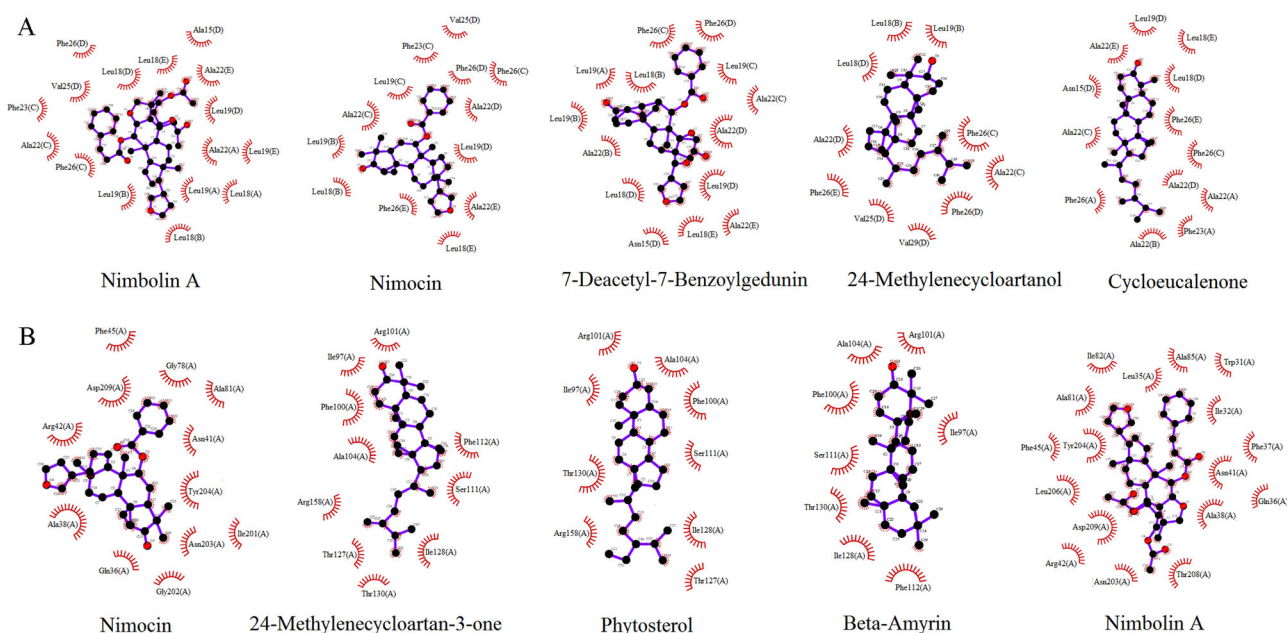
-193.81 kJ/mol respectively (Table 3). Among all observed interactions, the contribution of van der Waals and polar solvation interactions to the binding free energies were more than that of others. Leu 18, Leu 19, Leu 21, Ala 22, Val 25 and Phe 26, were found to be common among all interacting residues in the complexes, and contributed strongly to the binding energy values (Figure 5).

M-Protein

The Nimbolin A- M protein complex showed the lowest binding free energy value (-188.99 kJ/mol), followed by 24-Methylenecycloartan-3-one (-173.33 kJ/mol), Beta-Amyrin (-161.55 kJ/mol), Nimocin (-155.77 kJ/mol), and Phytosterol (-123.70 kJ/mol) (Table 4). In this case also, the contribution of van der Waals and polar solvation energies were the highest. Common residues contributing strongly to the binding energy in the complexes of M with Nimbolin A and Nimocin included Met 84, Ala 85, Arg 200 and Tyr 204. Interaction of 24-Methylenecycloartan-3-one, Beta-Amyrin and Phytosterol was with an alternate site of M protein. Per residue

Table 1. Docking results of top five ligands with the E-protein of SARS-CoV-2.

Compound	Chemical Structure	Binding Energy (kcal/mol)	H-bonds	Hydrophobic Interactions (Chain ID)
Nimbolin A		-11.2	NA	Leu 18 (A), Leu 19 (A), Ala 22 (A), Leu 19 (B), Leu 18 (B), Ala 22 (C), Phe 23 (C), Phe 26 (C), Asn 15 (D), Leu 18 (D), Leu 19 (D), Val 25 (D), Phe 26 (D), Leu 18 (E), Leu 19 (E), Ala 22 (E)
Nimocin		-11.0	NA	Leu 18 (B), Leu 19 (B), Leu 19 (C), Ala 22 (C), Phe 23 (C), Phe 26 (C), Leu 19 (D), Ala 22 (D), Val 25 (D), Phe 26 (D), Ala 22 (E), Leu 18 (E), Phe 26 (E)
7-Deacetyl-7-Benzoylgedunin		-10.9	NA	Leu 19 (A), Leu 18 (B), Leu 19 (B), Ala 22 (B), Leu 19 (C), Ala 22 (C), Phe 26 (C), Ala 15 (D), Leu 18 (D), Leu 19 (D), Ala 22 (D), Phe 26 (D), Leu 18 (E), Ala 22 (E)
24-Methylenecycloartanol		-10.8	NA	Leu 18 (B), Leu 19 (B), Ala 22 (C), Phe 26 (C), Leu 18 (D), Ala 22 (D), Val 25 (D), Phe 26 (D), Val 29 (D), Phe 26 (E)
Cycloeucalenone		-10.5	NA	Ala 22 (A), Phe 23 (A), Phe 26 (A), Ala 22 (B), Ala 22 (C), Phe 26 (C), Ala 15 (D), Leu 18 (D), Leu 19 (D), Ala 22 (D), Leu 18 (E), Ala 22 (E), Phe 26 (E)

**Figure 2.** 2D interaction diagram of five compounds with the highest binding energies against SARS-CoV-2 E and M proteins. The diagrams were generated with LigPlot⁺⁺.

contribution (≥ 1.5 kJ/mol) to the binding energy is shown in Figure 6. Residues in the conserved C-terminal region (SMWSFNPETNIL) of M protein (Arndt et al., 2010), like Met 109, Trp 110, Phe 112, Pro 114 and Ile 118 were some of the highest contributors to the binding energy values.

Pharmacokinetic properties of the top compounds

The pharmacokinetic properties from filtering analyses (Supplementary material Table S3) suggested that all top compounds have high probability of absorption and blood

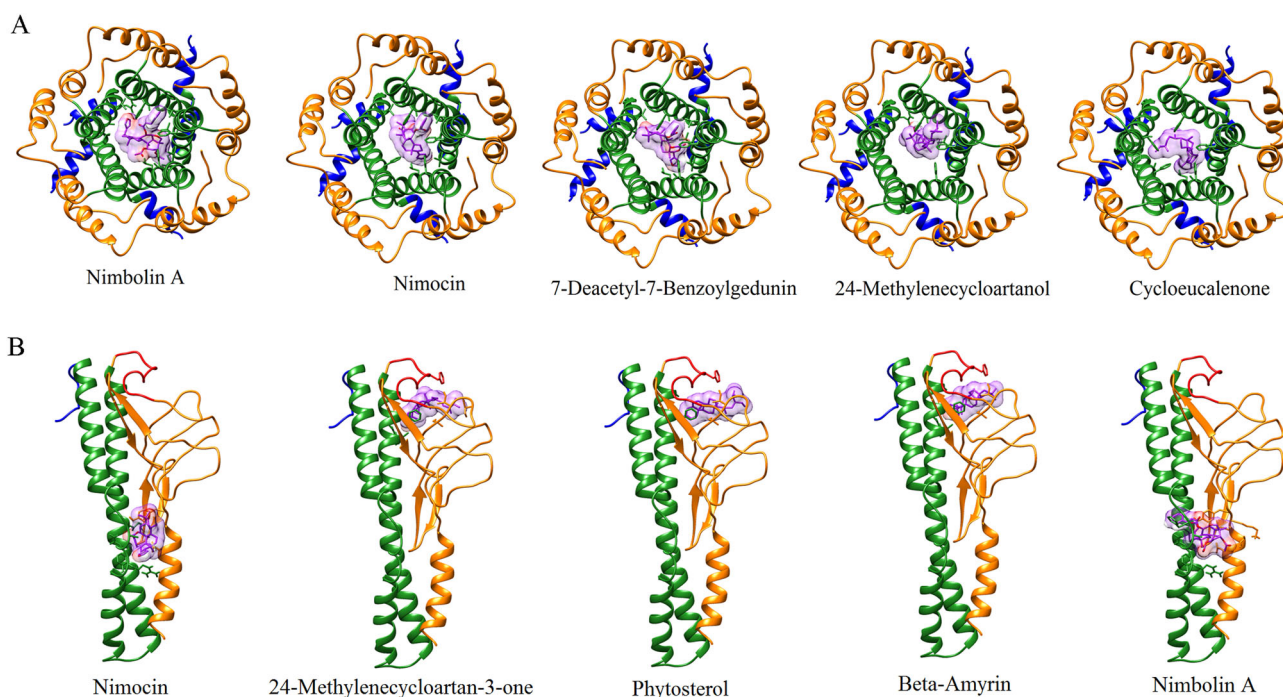
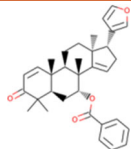
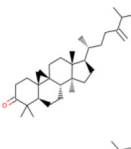
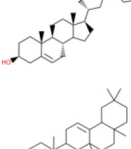
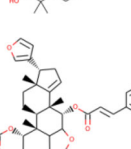
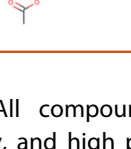


Figure 3. 3D interaction diagram of five compounds with the highest binding energies against SARS-CoV-2 E and M proteins. The ligands are shown in violet surface representation.

Table 2. Docking results of top five ligands with the M-protein of SARS-CoV-2.

Compound	Chemical Structure	Binding Energy (kcal/mol)	H-bonds	Hydrophobic Interactions
Nimocin		-10.2	NA	Gln 36, Ala 38, Asn 41, Arg 42, Phe 45, Gly 78, Ala 81, Ile 201, Gly 202, Asn 203, Tyr 204, Asp 209
24-Methylenecycloartan-3-one		-10	NA	Ile 97, Phe 100, Arg 101, Ala 104, Ser 111, Phe 112, Thr 127, Ile 128, Thr 130, Arg 158
Phytosterol		-9.5	NA	Ile 97, Phe 100, Arg 101, Ala 104, Ser 111, Thr 127, Ile 128, Thr 130, Arg 158.
Beta-Amyrin		-9.4	NA	Ile 97, Phe 100, Arg 101, Ala 104, Ser 111, Phe 112, Ile 128, Thr 130
Nimbolin A		-9.4	NA	Trp 31, Ile 32, Leu 35, Gln 36, Phe 37, Ala 38, Asn 41, Arg 42, Phe 45, Ala 81, Ile 82, Ala 85, Asn 203, Tyr 204, Leu 206, Thr 208, Asp 209.

brain barrier permeability. All compounds displayed low CYP450 inhibitory promiscuity, and high probability of being non-carcinogenic and non-toxic.

Discussion

The novel coronavirus SARS-CoV-2 constitutes an unprecedented physical, physiological and financial threat to the

global population. Till now, this virus has caused ~3.94 million infections and ~275,000 deaths all over the world. All movement and activities around the world have come to a grinding halt in an effort to contain the virus. As of now, the treatment is primarily symptomatic; but clinical trials with an investigational antiviral drug Remdesivir has shown reduced recovery time in COVID-19 patients (Hendaus, 2020).

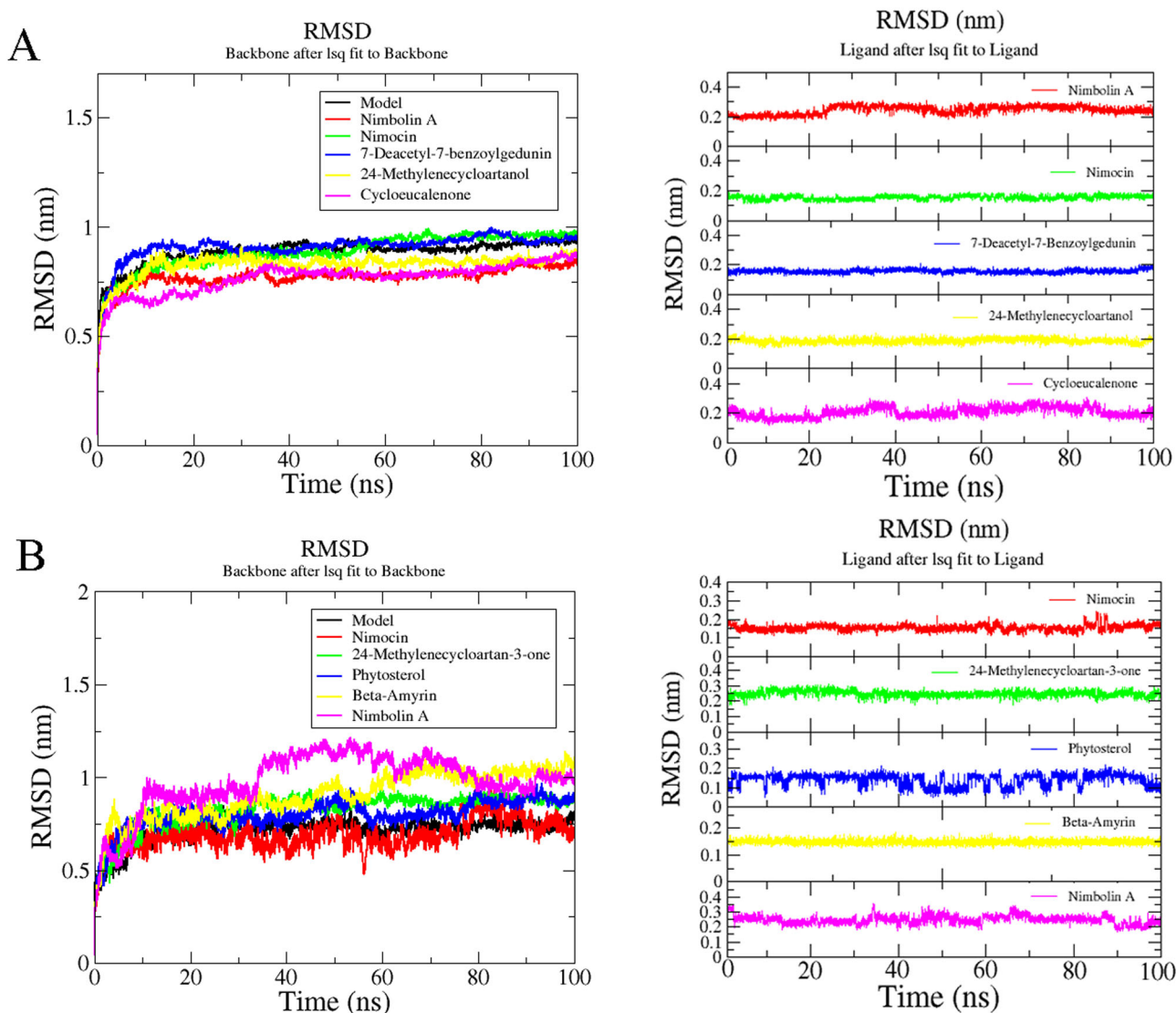


Figure 4. Molecular Dynamics (MD) simulation results of Protein-ligand complexes: RMSD of protein backbone over time and RMSD of ligands over time for A) E-protein complexes and B) M-protein complexes.

Table 3. Binding free energy (MM/PBSA) and their components (kJ/mol) of E-protein complexes.

Ligands	Binding Energy kJ/mol	Van der Waal Energy	Electrostatic Energy	Polar Solvation Energy	SASA Energy
Nimbolin A	-297.49 ± 14.05	-329.32 ± 12.37	-5.24 ± 3.53	67.80 ± 8.19	-30.73 ± 1.42
Nimocin	-277.20 ± 9.90	-294.12 ± 10.33	-1.25 ± 1.27	43.28 ± 3.64	-25.10 ± 1.10
7-Deacetyl-7-Benzoylgedunin	-297.87 ± 9.72	-323.11 ± 9.94	-4.18 ± 1.89	54.94 ± 3.76	-25.51 ± 1.10
24-Methylenecycloartanol	-199.84 ± 9.70	-215.18 ± 11.18	-15.12 ± 3.90	52.86 ± 5.11	-22.40 ± 1.38
Cycloeucaenone	-193.81 ± 9.16	-212.18 ± 10.06	0.35 ± 3.03	39.60 ± 8.44	-21.57 ± 1.28

Although FDA has issued an emergency use authorization for Remdesivir (Barrett, 2020; Wang et al., 2020), the requirement of more specific and effective antiviral drugs against SARS-CoV-2 is evident. Multiple vaccine candidates are being tested, however the licensing and mass production of an effective vaccine is expected to require a minimum period of 1–1.5 years. Prior to the emergence of SARS-CoV-2, the world has witnessed the emergence of SARS-CoV and MERS-CoV. This indicates that the chance of future emergence of another novel human pathogen from the coronavirus family, with the capability to cause a pandemic, cannot be dismissed (Menachery et al., 2015). Thus, identification of crucial steps in the virus life cycle, and developing specific inhibitors

targeted towards these processes, appears to be a viable strategy.

Computational methods tender a fast and cost-efficient approach to design potential inhibitors. We have utilized virtual screening and molecular dynamics simulation to identify small molecules that bind to the structural proteins M and E of SARS-CoV-2 with. M and E are necessary and sufficient to form particles of other β -coronaviruses (Mortola & Roy, 2004), therefore they are thought to have a central role in virus assembly. Interaction of the M-protein with all other major structural proteins like N, S and E is central to virus assembly. The contribution of the ion-channel protein E in membrane interaction and bending is deemed necessary for

contributed highly to the binding free energy. Post simulation MM/PBSA analysis showed that residues within the pentameric channel such as Leu 18, Leu 19, Leu 21, Ala 22, Val 25 and Phe 26 were also strong contributors to binding free energy, indicating robust interaction with the screened compounds. Our studies indicate that these compounds may have the ability to inhibit ion channel activity of SARS-CoV-2 Envelope protein, thus hampering virus assembly.

In case of M-protein, docking analysis with Neem compounds revealed two possible binding sites. Nimocin and Nimbolin A interacted primarily with the transmembrane domain, with contributions from the last 20 residues at the C-terminus. The C-terminal tail, including residues Lys 199, Gly 201, Tyr 203 and Arg 204, are required for localization of the M-protein of MERS-CoV in the Trans-Golgi Network (TGN) (Perrier et al., 2019). Localization of M to this site is crucial for virion assembly. Our post simulation MM-PBSA analysis showed that Gly 202 and Tyr 204 are among the strongest contributors to the binding free energy from the C-terminal end of M-protein. The residues GNY (202–204) are conserved in members of the β -coronavirus family (Supplementary material Figure S5).

24-Methylenecycloartan-3-one, Phytosterol and Beta-Amyrin were found to interact with the C-terminal region, primarily with the conserved domain CD (SMWSFNPETNIL), which is essential for virus assembly (Arndt et al., 2010). Thus, the compounds identified in this study may hamper virus assembly by interacting with crucial regions of the M-protein involved in TGN localization and protein-protein interaction.

In conclusion, we have identified a few possible inhibitors of the E and M proteins of SARS-CoV-2 using molecular docking, MD simulation and binding free energy calculations. These compounds, derived from Neem, displayed stable binding and interactions with crucial regions of E and M required for assembly; and were predicted to have good pharmacokinetic properties. Nimbolin A showed the strongest binding free energy with both E and M proteins. Other compounds: Nimocin and Cycloartanols (24-Methylenecycloartanol and 24-Methylenecycloartan-3-one) were also common ligands, binding strongly to both proteins. A recent study has shown ameliorating effect of Neem extract on propagation and pathophysiology of another member of the coronavirus family (Sarkar et al., 2020), reinforcing the potential of these compounds as potential therapeutic options. Experimental validation and optimization of these natural compounds might add value to the development of specific therapeutics against SARS-CoV-2.

Acknowledgements

Authors would like to thank the Central Hybrid Supercomputing Cluster, Indian Institute of Technology Delhi for providing computational facilities.

Disclosure statement

The authors of this manuscript declare no conflict of interest.

ORCID

Subhomai Borkotoky  <http://orcid.org/0000-0001-6752-4296>

References

- Aanouz, I., Belhassan, A., El-Khatibi, K., Lakhliifi, T., El-Ldrissi, M., & Bouachrine, M. (2020). Moroccan Medicinal plants as inhibitors against SARS-CoV-2 main protease: Computational investigations. *Journal of Biomolecular Structure and Dynamics*, 1–9. <https://doi.org/10.1080/07391102.2020.1758790>
- Abdelli, I., Hassani, F., Bekkel Brikci, S., & Ghalem, S. (2020). In silico study the inhibition of angiotensin converting enzyme 2 receptor of COVID-19 by *Ammoides verticillata* components harvested from Western Algeria. *Journal of Biomolecular Structure and Dynamics*, 1–14. <https://doi.org/10.1080/07391102.2020.1763199>
- Akram, M., Tahir, I. M., Shah, S. M. A., Mahmood, Z., Altaf, A., Ahmad, K., Munir, N., Daniyal, M., Nasir, S., & Mehboob, H. (2018). Antiviral potential of medicinal plants against HIV, HSV, influenza, hepatitis, and coxsackievirus: A systematic review. *Phytotherapy Research: PTR*, 32(5), 811–822. <https://doi.org/10.1002/ptr.6024>
- Alzohairy, M. A. (2016). Therapeutics role of *Azadirachta indica* (Neem) and their active constituents in diseases prevention and treatment. *Evidence-Based Complementary and Alternative Medicine*, 2016, 7382506. <https://doi.org/10.1155/2016/7382506>
- Arndt, A. L., Larson, B. J., & Hogue, B. G. (2010). A conserved domain in the coronavirus membrane protein tail is important for virus assembly. *Journal of Virology*, 84(21), 11418–11428. <https://doi.org/10.1128/JVI.01131-10>
- Badam, L., Joshi, S. P., & Bedekar, S. S. (1999). 'In vitro' antiviral activity of neem (*Azadirachta indica*. A. Juss) leaf extract against group B coxsackieviruses. *Journal of Communicable Diseases*, 31(2), 79–90.
- Baker, N. A., Sept, D., Joseph, S., Holst, M. J., & McCammon, J. A. (2001). Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences of the United States of America*, 98(18), 10037–10041. <https://doi.org/10.1073/pnas.181342398>
- Barrett, J. (2020). FDA issues emergency use authorization for remdesivir in COVID-19. Retrieved May 1, 2020, from <https://www.drugtopics.com/covid-19/fda-issues-emergency-use-authorization-remdesivir-covid-19>
- Bhardwaj, V. K., Singh, R., Sharma, J., Rajendran, V., Purohit, R., & Kumar, S. (2020). Identification of bioactive molecules from Tea plant as SARS-CoV-2 main protease inhibitors. *Journal of Biomolecular Structure and Dynamics*, 1–13. <https://doi.org/10.1080/07391102.2020.1766572>
- Bhattacharya, D., Nowotny, J., Cao, R., & Cheng, J. (2016). 3Drefine: An interactive web server for efficient protein structure refinement. *Nucleic Acids Research*, 44(W1), W406–W409. <https://doi.org/10.1093/nar/gkw336>
- Boopathi, S., Poma, A. B., & Kolandaivel, P. (2020). Novel 2019 coronavirus structure, mechanism of action, antiviral drug promises and rule out against its treatment. *Journal of Biomolecular Structure and Dynamics*, 1–10. <https://doi.org/10.1080/07391102.2020.1758788>
- Bussi, G., Donadio, D., & Parrinello, M. (2007). Canonical sampling through velocity rescaling. *The Journal of Chemical Physics*, 126(1), 014101. <https://doi.org/10.1063/1.2408420>
- Caly, L., Druce, J. D., Catton, M. G., Jans, D. A., & Wagstaff, K. M. (2020). The FDA-approved drug ivermectin inhibits the replication of SARS-CoV-2 in vitro. *Antiviral Research*, 178, 104787. <https://doi.org/10.1016/j.antiviral.2020.104787>
- Cheng, F., Li, W., Zhou, Y., Shen, J., Wu, Z., Liu, G., Lee, P. W., & Tang, Y. (2012). admetSAR: A comprehensive source and free tool for assessment of chemical ADMET properties. *Journal of Chemical Information and Modeling*, 52(11), 3099–3105. <https://doi.org/10.1021/ci300367a>
- DeDiego, M. L., Alvarez, E., Almazán, F., Rojas, M. T., Lamirande, E., Roberts, A., Shieh, W.-J., Zaki, S. R., Subbarao, K., & Enjuanes, L. (2007). A severe acute respiratory syndrome coronavirus that lacks the E gene is attenuated in vitro and in vivo. *Journal of Virology*, 81(4), 1701–1713. <https://doi.org/10.1128/JVI.01467-06>

- Ekhteiari Salmas, R., Unlu, A., Bektaş, M., Yurtsever, M., Mestanoglu, M., & Durdagi, S. (2017). Virtual screening of small molecules databases for discovery of novel PARP-1 inhibitors: Combination of in silico and in vitro studies. *Journal of Biomolecular Structure & Dynamics*, 35(9), 1899–1915. <https://doi.org/10.1080/07391102.2016.1199328>
- Elfiky, A. A. (2020). SARS-CoV-2 RNA dependent RNA polymerase (RdRp) targeting: An in silico perspective. *Journal of Biomolecular Structure and Dynamics*, 1–9. <https://doi.org/10.1080/07391102.2020.1761882>
- Enmozhi, S. K., Raja, K., Sebastine, I., & Joseph, J. (2020). Andrographolide as a potential inhibitor of SARS-CoV-2 main protease: An in silico approach. *Journal of Biomolecular Structure and Dynamics*, 1–7. <https://doi.org/10.1080/07391102.2020.1760136>
- Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M. S., Eramian, D., Shen, M.-Y., Pieper, U., & Sali, A. (2006). Comparative protein structure modeling using Modeller. *Current Protocols in Bioinformatics*, 15, 5.6.1–5.6.30. <https://doi.org/10.1002/0471250953.bi0506s15>
- Gao, Y., Yan, L., Huang, Y., Liu, F., Zhao, Y., Cao, L., Wang, T., Sun, Q., Ming, Z., Zhang, L., Ge, J., Zheng, L., Zhang, Y., Wang, H., Zhu, Y., Zhu, C., Hu, T., Hua, T., Zhang, B., ... Rao, Z. (2020). Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Science*, 368(6492), 779–782. <https://doi.org/10.1126/science.abb7498>
- Hendaus, M. A. (2020). Remdesivir in the treatment of Coronavirus Disease 2019 (COVID-19): A simplified summary. *Journal of Biomolecular Structure and Dynamics*, 1–10. <https://doi.org/10.1080/07391102.2020.1767691>
- Hess, B., Bekker, H., Berendsen, H. J., & Fraaije, J. G. (1997). LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, 18(12), 1463–1472. [https://doi.org/10.1002/\(SICI\)1096-987X\(199709\)18:12 < 1463::AID-JCC4 > 3.0.CO;2-H](https://doi.org/10.1002/(SICI)1096-987X(199709)18:12 < 1463::AID-JCC4 > 3.0.CO;2-H)
- Jimenez-Guardeno, J. M., Nieto-Torres, J. L., DeDiego, M. L., Regla-Nava, J. A., Fernandez-Delgado, R., Castano-Rodriguez, C., & Enjuanes, L. (2014). The PDZ-binding motif of severe acute respiratory syndrome coronavirus envelope protein is a determinant of viral pathogenesis. *PLoS Pathogens*, 10(8), e1004320. <https://doi.org/10.1371/journal.ppat.1004320>
- Kawata, M., & Nagashima, U. (2001). Particle mesh Ewald method for three-dimensional systems with two-dimensional periodicity. *Chemical Physics Letters*, 340(1-2), 165–172. [https://doi.org/10.1016/S0009-2614\(01\)00393-1](https://doi.org/10.1016/S0009-2614(01)00393-1)
- Khan, S. A., Zia, K., Ashraf, S., Uddin, R., & Ul-Haq, Z. (2020). Identification of chymotrypsin-like protease inhibitors of SARS-CoV-2 via integrated computational approach. *Journal of Biomolecular Structure and Dynamics*, 1–10. <https://doi.org/10.1080/07391102.2020.1751298>
- Kumari, R., Kumar, R., & Lynn, A. (2014). g_mmpbsa-a GROMACS tool for high-throughput MM-PBSA calculations. *Journal of Chemical Information and Modeling*, 54(7), 1951–1962. <https://doi.org/10.1021/ci500020m>
- Kuo, L., Hurst-Hess, K. R., Koetzner, C. A., & Masters, P. S. (2016). Analyses of coronavirus assembly interactions with interspecies membrane and nucleocapsid protein chimeras. *Journal of Virology*, 90(9), 4357–4368. <https://doi.org/10.1128/JVI.03212-15>
- Martónák, R., Laio, A., & Parrinello, M. (2003). Predicting crystal structures: The Parrinello-Rahman method revisited. *Physical Review Letters*, 90(7), 075503. <https://doi.org/10.1103/PhysRevLett.90.075503>
- Menachery, V. D., Yount, B. L., Debbink, K., Agnihotram, S., Gralinski, L. E., Plante, J. A., Graham, R. L., Scobey, T., Ge, X.-Y., Donaldson, E. F., Randell, S. H., Lanzavecchia, A., Marasco, W. A., Shi, Z.-L., & Baric, R. S. (2015). A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence. *Nature Medicine*, 21(12), 1508–1513. <https://doi.org/10.1038/nm.3985>
- Mohanraj, K., Karthikeyan, B. S., Vivek-Ananth, R. P., Chand, R. P. B., Aparna, S. R., Mangalampandi, P., & Samal, A. (2018). IMPPAT: A curated database of Indian Medicinal Plants, Phytochemistry and Therapeutics. *Scientific Reports*, 8(1), 4329. <https://doi.org/10.1038/s41598-018-22631-z>
- Mortola, E., & Roy, P. (2004). Efficient assembly and release of SARS coronavirus-like particles by a heterologous expression system. *FEBS Letters*, 576(1-2), 174–178. <https://doi.org/10.1016/j.febslet.2004.09.009>
- Nieto-Torres, J. L., DeDiego, M. L., Verdía-Báguena, C., Jimenez-Guardeno, J. M., Regla-Nava, J. A., Fernandez-Delgado, R., Castaño-Rodriguez, C., Alcaraz, A., Torres, J., Aguilera, V. M., & Enjuanes, L. (2014). Severe acute respiratory syndrome coronavirus envelope protein ion channel activity promotes virus fitness and pathogenesis. *PLoS Pathogens*, 10(5), e1004077. <https://doi.org/10.1371/journal.ppat.1004077>
- Oostenbrink, C., Villa, A., Mark, A. E., & Van Gunsteren, W. F. (2004). A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *Journal of Computational Chemistry*, 25(13), 1656–1676. <https://doi.org/10.1002/jcc.20090>
- Patridge, E., Gareiss, P., Kinch, M. S., & Hoyer, D. (2016). An analysis of FDA-approved drugs: Natural products and their derivatives. *Drug Discovery Today*, 21(2), 204–207. <https://doi.org/10.1016/j.drudis.2015.01.009>
- Perrier, A., Bonnin, A., Desmaret, L., Danneels, A., Goffard, A., Rouillé, Y., Dubuisson, J., & Belouzard, S. (2019). The C-terminal domain of the MERS coronavirus M protein contains a trans-Golgi network localization signal. *The Journal of Biological Chemistry*, 294(39), 14406–14421. <https://doi.org/10.1074/jbc.RA119.008964>
- Pervushin, K., Tan, E., Parthasarathy, K., Lin, X., Jiang, F. L., Yu, D., Vararattanavech, A., Soong, T. W., Liu, D. X., & Torres, J. (2009). Structure and inhibition of the SARS coronavirus envelope protein ion channel. *PLoS Pathogens*, 5(7), e1000511. <https://doi.org/10.1371/journal.ppat.1000511>
- Rajkumari, J., Borkotoky, S., Murali, A., Suchiang, K., Mohanty, S. K., & Busi, S. (2018). Attenuation of quorum sensing controlled virulence factors and biofilm formation in *Pseudomonas aeruginosa* by pentacyclic triterpenes, betulin and betulinic acid. *Microbial Pathogenesis*, 118, 48–60. <https://doi.org/10.1016/j.micpath.2018.03.012>
- Roy, A., Kucukural, A., & Zhang, Y. (2010). I-TASSER: A unified platform for automated protein structure and function prediction. *Nature Protocols*, 5(4), 725–738. <https://doi.org/10.1038/nprot.2010.5>
- Sarkar, L., Puthala, R. K., Safiriyu, A. A., & Sarma, J. D. (2020). *Azadirachta indica* A. Juss ameliorates mouse hepatitis virus-induced neuroinflammatory demyelination by modulating cell-to-cell fusion in an experimental animal model of multiple sclerosis. *Frontiers in Cellular Neuroscience*, 14, 116. <https://doi.org/10.3389/fncel.2020.00116>
- Schoeman, D., & Fielding, B. C. (2019). Coronavirus envelope protein: Current knowledge. *Virology Journal*, 16(1), 69. <https://doi.org/10.1186/s12985-019-1182-0>
- Schüttelkopf, A. W., & van Aalten, D. M. F. (2004). PRODRG: A tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallographica Section D Biological Crystallography*, 60(Pt 8), 1355–1363. <https://doi.org/10.1107/S0907444904011679>
- Surya, W., Li, Y., & Torres, J. (2018). Structural model of the SARS coronavirus E channel in LMPG micelles. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1860(6), 1309–1317. <https://doi.org/10.1016/j.bbame.2018.02.017>
- Teoh, K.-T., Siu, Y.-L., Chan, W.-L., Schlüter, M. A., Liu, C.-J., Peiris, J. S. M., Bruzzone, R., Margolis, B., & Nal, B. (2010). The SARS coronavirus E protein interacts with PALS1 and alters tight junction formation and epithelial morphogenesis. *Molecular Biology of the Cell*, 21(22), 3838–3852. <https://doi.org/10.1091/mbc.E10-04-0338>
- Thomford, N. E., Senthelane, D. A., Rowe, A., Munro, D., Seele, P., Maroyi, A., & Dzobo, K. (2018). Natural products for drug discovery in the 21st century: Innovations for novel drug discovery. *International Journal of Molecular Sciences*, 19(6), 1578. <https://doi.org/10.3390/ijms19061578>
- Tiwari, V., Darmani, N. A., Yue, B. Y., & Shukla, D. (2010). In vitro antiviral activity of neem (*Azadirachta indica* L.) bark extract against herpes simplex virus type-1 infection. *Phytotherapy Research: PTR*, 24(8), 1132–1140. <https://doi.org/10.1002/ptr.3085>
- Torres, J., Maheswari, U., Parthasarathy, K., Ng, L., Liu, D. X., & Gong, X. (2007). Conductance and amantadine binding of a pore formed by a lysine-flanked transmembrane domain of SARS coronavirus envelope protein. *Protein Science*, 16(9), 2065–2071. <https://doi.org/10.1110/ps.062730007>
- Torres, J., Parthasarathy, K., Lin, X., Saravanan, R., Kukol, A., & Liu, D. X. (2006). Model of a putative pore: The pentameric alpha-helical bundle of SARS coronavirus E protein in lipid bilayers. *Biophysical Journal*, 91(3), 938–947. <https://doi.org/10.1529/biophysj.105.080119>

- Trott, O., & Olson, A. J. (2010). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, 31(2), 455–461. <https://doi.org/10.1002/jcc.21334>
- Verdia-Baguena, C., Nieto-Torres, J. L., Alcaraz, A., DeDiego, M. L., Torres, J., Aguilera, V. M., & Enjuanes, L. (2012). Coronavirus E protein forms ion channels with functionally and structurally-involved membrane lipids. *Virology*, 432(2), 485–494. <https://doi.org/10.1016/j.virol.2012.07.005>
- Walls, A. C., Park, Y. J., Tortorici, M. A., Wall, A., McGuire, A. T., & Velesler, D. (2020). Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell*, 181(2), 281–292.e6. <https://doi.org/10.1016/j.cell.2020.02.058>
- Wang, Y., Zhang, D., Du, G., Du, R., Zhao, J., Jin, Y., Fu, S., Gao, L., Cheng, Z., Lu, Q., Hu, Y., Luo, G., Wang, K., Lu, Y., Li, H., Wang, S., Ruan, S., Yang, C., Mei, C., ... Wang, C. (2020). Remdesivir in adults with severe COVID-19: A randomised, double-blind, placebo-controlled, multi-centre trial. *The Lancet*, 395(10236), 1569–1578. [https://doi.org/10.1016/S0140-6736\(20\)31022-9](https://doi.org/10.1016/S0140-6736(20)31022-9)
- Wani, M. C., Taylor, H. L., Wall, M. E., Coggon, P., & McPhail, A. T. (1971). Plant antitumor agents. VI. The isolation and structure of taxol, a novel antileukemic and antitumor agent from *Taxus brevifolia*. *Journal of the American Chemical Society*, 93(9), 2325–2327. <https://doi.org/10.1021/ja00738a045>
- Wu, C., Liu, Y., Yang, Y., Zhang, P., Zhong, W., Wang, Y., Wang, Q., Xu, Y., Li, M., Li, X., Zheng, M., Chen, L., & Li, H. (2020). Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. *Acta Pharmaceutica Sinica B*. <https://doi.org/10.1016/j.apsb.2020.02.008>
- Yang, Y., Xiong, Z., Zhang, S., Yan, Y., Nguyen, J., Ng, B., Lu, H., Brendese, J., Yang, F., Wang, H., & Yang, X.-F. (2005). Bcl-xL inhibits T-cell apoptosis induced by expression of SARS coronavirus E protein in the absence of growth factors. *The Biochemical Journal*, 392(Pt 1), 135–143. <https://doi.org/10.1042/BJ20050698>
- Zhang, L., Lin, D., Sun, X., Curth, U., Drosten, C., Sauerhering, L., Becker, S., Rox, K., & Hilgenfeld, R. (2020). Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α -ketoamide inhibitors. *Science*, 368(6489), 409–412. <https://doi.org/10.1126/science.abb3405>