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

Computers in Biology and Medicine 127 (2020) 104063

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: http://www.elsevier.com/locate/compbiomed

In silico identification of Tretinoin as a SARS-CoV-2 envelope (E) protein ion channel inhibitor

Debajit Dey [a], Subhomoi Borkotoky [b], Manidipa Banerjee [b,*]

[a] School of Medicine, University of Maryland Baltimore, United States
[b] Kusuma School of Biological Sciences, Indian Institute of Technology Delhi, India

ARTICLE INFO

Keywords:
SARS-2
Viroporin
Virus-host interactions
Therapeutics
Drug

ABSTRACT

Viroporins are oligomeric, pore forming, viral proteins that play critical roles in the life cycle of pathogenic viruses. Viroporins like HIV-1 Vpu, Alphavirus 6 K, Influenza M2, HCV p7, and Picornavirus 2B, form discrete aqueous passageways which mediate ion and small molecule transport in infected cells. The alterations in host membrane structures induced by viroporins is essential for key steps in the virus life cycle like entry, replication and egress. Any disruption in viroporin functionality severely compromises viral pathogenesis. The envelope (E) protein encoded by coronaviruses is a viroporin with ion channel activity and has been shown to be crucial for the assembly and pathophysiology of coronaviruses. We used a combination of virtual database screening, molecular docking, all-atom molecular dynamics simulation and MM-PBSA analysis to test four FDA approved drugs - Tretinoin, Mefenamic Acid, Ondansetron and Artemether - as potential inhibitors of ion channels formed by SARS-CoV-2 E protein. Interaction and binding energy analysis showed that electrostatic interactions and polar solvation energy were the major driving forces for binding of the drugs, with Tretinoin being the most promising inhibitor. Tretinoin bound within the lumen of the channel formed by E protein, which is lined by hydrophobic residues like Phe, Val and Ala, indicating its potential for blocking the channel and inhibiting the viroporin functionality of E. In control simulations, tretinoin demonstrated a lower binding energy with a known target as compared to SARS-CoV-2 E protein. This work thus highlights the possibility of exploring Tretinoin as a potential SARS-CoV-2 E protein ion channel blocker and virus assembly inhibitor, which could be an important therapeutic strategy in the treatment for coronaviruses.

1. Introduction

Although viruses from different families have distinct life cycles; there exist striking mechanistic similarities in host interaction pathways. A few such commonalities include host membrane fusion by enveloped viruses, membrane disruption by non-enveloped viruses orchestrated by amphipathic viral peptides, and viroporin mediated membrane alteration/remodelling for facilitating virus replication and egress. Both enveloped and non-enveloped viruses utilize virally encoded membrane proteins or "viroporins" at different stages of their life cycles; and membrane remodelling by viroporins appears to be a crucial step during enveloped virus assembly and budding [1]. Examples of well-studied viroporins include M2 protein of Influenza A Virus, 6 K protein of Chikungunya Virus (CHIKV), Vpu of HIV-1 and p7 of HCV [1,2]. Small molecule inhibitors block multiple viroporins from different virus families; in spite of these components lacking in sequential or structural similarities [2]. This shows a degree of mechanistic equivalency between viroporins, indicating that they may have similar roles to execute in the viral life cycle [2,3].

The "E" or "Envelope" protein of coronaviruses is a viroporin produced during viral infection, although only a small percentage is incorporated into assembled virions [4]. The majority of the protein remains associated with ER, golgi and the ERGIC pathway, which are the sites for coronavirus assembly and release [5]. E is a small membrane protein of 76–109 amino acids, with a short N-terminal region of 7–12 amino acids, a central transmembrane domain (TMD) and a cytoplasmic tail, which is largely hydrophilic [4,6]. The major function of E in coronavirus biology appears to be to support assembly and release of particles [4,7]. It has been shown by several groups that association between the cytoplasmic tails of E and M (Membrane protein) of coronaviruses is essential for particle formation [8]. While M interacts with multiple structural proteins, the involvement of E is crucial for

* Corresponding author.
E-mail address: mbanerjee@bioschool.iitd.ac.in (M. Banerjee).

https://doi.org/10.1016/j.compbiomed.2020.104063
Received 3 August 2020; Received in revised form 25 September 2020; Accepted 14 October 2020
Available online 20 October 2020

membrane bending and scission during virus budding, leading to the "pinching off" of the virus from ERGIC [4,7,8]. Removal of E from the structural cassette of SARS-CoV has a devastating effect on virus assembly and replication, resulting in the production of immature or incompetent progeny [9].

The transmembrane domain of E is essential for its viroporin activity [3,6]. A synthetic version of the TMD assembles into pentamers that form cation-selective membrane pores. The ion channels generated by E of SARS-CoV in ERGIC membranes transports Ca^{2+}, which was shown to activate the NLRP3 inflammasome activation, and IL-1β overproduction [10]. This phenomenon was linked to the worsening of lung damage observed in SARS-CoV infected mice, and thus to the immunopathology associated with viral infection [10]. The ion-channel activity of E is crucial for the virus life cycle. Point mutations N15A and V25F introduced in SARS-CoV E causes disruption of its ion-channel activity [4,6,7], which, in turn, results in the accumulation of compensatory mutations in the protein that restore its membrane activity. However, the mechanistic link between the ion channel activity of E, and its membrane bending functionality that facilitates virus budding, is not clearly understood.

The E protein of coronaviruses contains several motifs, apart from the TMD, that are necessary for its functionality. A β-coil-β-motif, containing a highly conserved proline residue, is necessary for the targeting of the protein to the golgi complex. Mutation of this residue retargets E to the plasma membrane, which inhibits the assembly and release of CoV particles [11]. A PDZ-binding motif (PBM) located in the terminal four amino acids of E is involved in several, crucial host factor interactions that appear to contribute to the pathogenicity of the virus [12,13]. Mutations in the PBM also rapidly revert in cell culture. Some of the reported host partners of the E-protein, which interact through the PBM, include the B-cell lymphoma extra-large protein (Bcl-xL), which has anti-apoptotic properties; the protein associated with *Caenorhabditis elegans* lin-7 protein 1 (PALS1); syntenin, sodium/potassium (Na+/K+) ATPase α-1 subunit, and stomatin [4,13] Some of these interactions have been shown to contribute significantly to the pathogenicity of CoVs. For example, the interaction of E protein with PALSI has been shown to disrupt tight junctions in the lungs, allowing virus particles to cross the alveolar barrier [14]. Likewise, interaction of E with syntenin results in overexpression of inflammatory cytokines that is thought to contribute to the tissue damage caused by coronaviruses [4,13].

Given the importance of E in virus propagation and new virus assembly and budding, this protein can be considered as a crucial drug target for antiviral generation. Previous studies with viroporins from different virus families have reported amantadine as a broadly functional inhibitor of ion channel activity [2,6,15,16]. Also, a recent study has reported Gliclazide and Memantine as potential inhibitors of E protein channel activity [17].

In this work, we targeted the E protein of the novel coronavirus SARS-CoV-2 (nCoV19) for drug repurposing studies using computational techniques. Utilizing a combination of *in silico* molecular docking, 200ns all-atom molecular dynamics simulations, H-bonding and binding energy analysis (MM-PBSA), we tested four specific FDA-approved drugs, Tretinoin, Mefenamic acid, Ondansetron and Artemether were evaluated for their ability to bind SARS-2 E protein. Out of the four drugs Tretinoin was selected as the best candidate due to its ability to form extensive H-bonding interactions and high binding energy value (−412.8 kJ/mol). It is hoped that blocking of the ion-channels formed by the viroporin E can have a substantial detrimental effect on SARS-CoV-2 assembly and propagation.

2. Materials and methods

2.1. In silico 3D structure generation and homo-oligomerization of SARS-CoV-2 E protein

The structure for SARS-CoV-2 E protein monomer was generated using the I-TASSER (https://zhanglab.ccmb.med.umich.edu/I-TASSER/) web server [18,19]. The E protein sequence utilized for structure prediction was from the Wuhan-1 isolate (GenBank ID: QHD43418.1). The template used for structure prediction was the NMR structure of the E protein from SARS CoV (PDB ID: 5X29)[20]. The model with the best C score (−0.75) was chosen for further processing. The predicted model was then refined using GalaxyRefine server [21]. The final selected 3D structure had a clash score of 1.6 with 94.5% of residues in the favored region. The secondary structure content of the predicted E protein monomer had a 66.7% alpha-helical content. Since the E protein from SARS-CoV has been shown to form pentamers [6], the predicted 3D structure of the SARS-CoV-2 E protein monomer was subjected to homo-oligomerization using the GalaxyHomomer server [22] with the SARS CoV E protein pentamer as the template. The resultant pentameric arrangement was energy minimized prior to being used for docking and simulation studies.

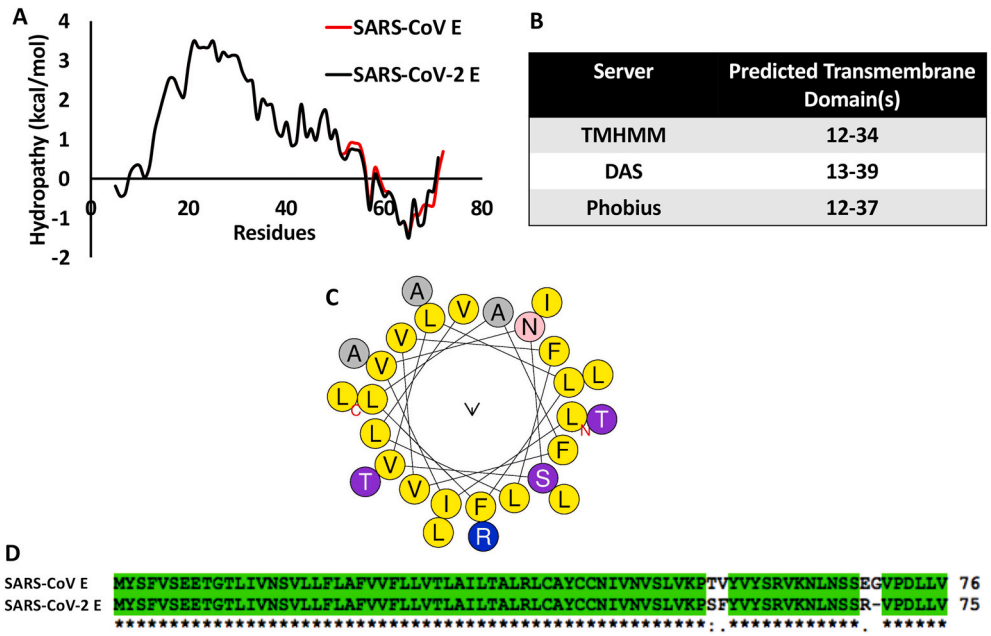
2.2. Virtual screening and molecular docking

All ligand structures were obtained from the DrugBank database (https://www.drugbank.ca/) [23]. For ligand screening, the DrugScreen server (http://cao.labshare.cn/drugscreen/) was utilized. Ligand search was conducted on a database of 1806 FDA-approved drug molecules, which generated a list of 50 top ranking molecules as output. Further screening was carried out based on the chemical nature, therapeutic target, side-effects and the status of FDA approval of molecules, which gave a total of 15 candidates [Supplementary Table T1]. Out of these four compounds were further tested.

The pentameric arrangement of SARS-CoV-2 E was docked with preselected compounds using the docking server CB-Dock (http://cao.labshare.cn/cb-dock/)[24]. CB-Dock performs blind docking using automated cavity-detection[24], followed by generation of docked poses using AutoDock Vina [25]. Selected ligand candidates were subjected to geometry optimization using the Universal Force Field [26,27]; and energy minimization with the steepest descent algorithm [28,29] using the Auto Optimize tool of Avogadro [30] prior to docking. Hydrogens and charges were added using the DockPrep tool of UCSF Chimera [31]. Binding affinity of docked ligands were evaluated using the PRODIGY web server [32]. Protein-ligand interactions of docked complexes were analyzed using Protein Ligand Interaction Profiler (PLIP) server (https://projects.biotec.tu-dresden.de/plip-web/plip/index) [33]. Visualization of protein-ligand interactions was carried out using UCSF Chimera [31] and PyMOL [34,35].

2.3. MD simulation

In order to understand the binding dynamics and stability of the protein-ligand complexes, multiple all-atom molecular dynamics simulations were carried out using GROMACS (v5.1.1) and the Gromos force field (43a1). Topology and parameters for each ligand was calculated using the PRODRG server [36]. A system consisting of E protein assembly, either with or without docked ligands, was embedded within a pre-equilibrated 392 POPC (1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine) membrane. System solvation was carried out using the Simple Point Charge (SPC) water model [37] followed by ionization with chloride (Cl^-) ions for system charge neutralization. 50,000 steps of energy minimization was performed using steepest descent algorithm [28,29]with a tolerance of 1000 kJ mol^{-1} nm^{-1}. This was followed by 500 ps of NVT (isothermal-isochoric ensemble) and 5ns of NPT (isothermal-isobaric ensemble) with position-restraints. The system temperature was maintained at 300 K using Noose-Hoover thermostat [38] and pressure was kept at 1 bar with Parrinello-Rahman pressure-coupling [39] with a compressibility of 4.5×10^{-5} bar^{-1} and 2ps time constant. Simulation timestep was kept at 2fs.Bond lengths were constrained using the LINCS (Linear Constraint Solver) algorithm [40]. Electrostatics were treated with PME (Particle mesh Ewald) [41]. A 1.2



Server	Predicted Transmembrane Domain(s)
TMHMM	12-34
DAS	13-39
Phobius	12-37

Fig. 1. (A) Kyte-Doolittle hydrophobicity plot of SARS CoV and SARS-CoV-2 E protein sequences. (B) Table showing the predicted transmembrane domain of SARS-CoV-2 E protein by three different transmembrane domain prediction servers [49–51] (C) Helical wheel plot of the transmembrane domain of SARS-CoV-2 E protein showing the amino acid composition, (D) Clustal Omega alignment of the full-length SARS CoV SARS-CoV-2 E protein sequences.

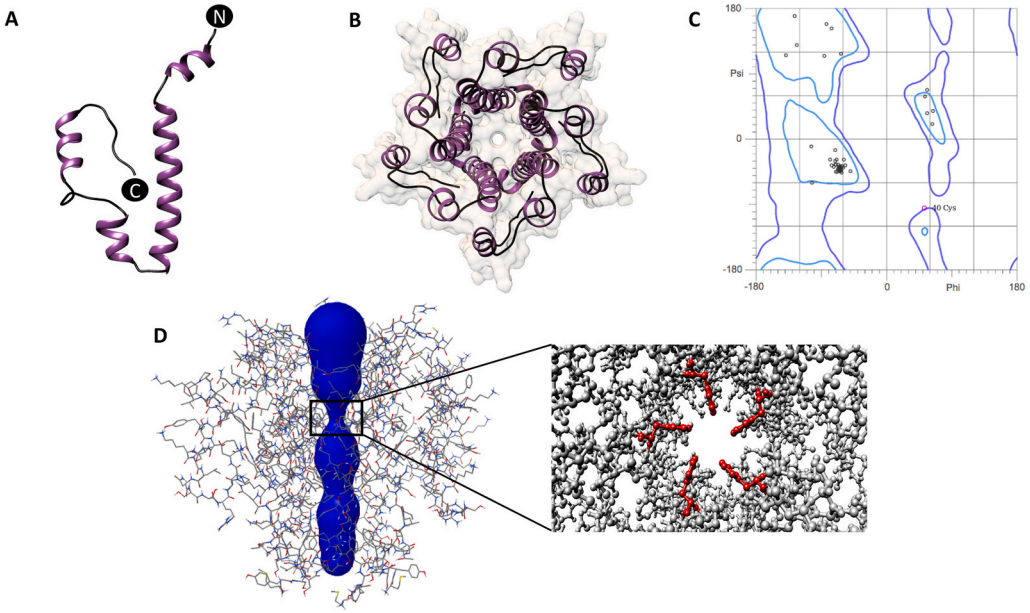


Fig. 2. (A) Predicted structure of the SARS-CoV-2 E protein monomer in ribbon representation (purple: alpha-helix; black: coil), (B) Top view of the pentameric arrangement of SARS-CoV-2 E protein, (C) Ramachandran plot of the predicted SARS-CoV-2 E protein monomer, (D) Line representation of the pentameric SARS-CoV-2 E protein showing the predicted ion channel. Inset shows the positioning of the Phe^{26} residues (red) at the narrowest point of the ion channel.

nm cut-off was used for both coulombic and van der Waals's interactions. Post equilibration, unrestrained production run was carried out for 200 ns. Trajectory analysis was carried out using Gromacs analysis tools and VMD [42].

Additionally, two control simulations were also carried out with the known targets of Tretinoin (DB00755) and Mefenamic acid (DB00784) which are Human Retinoic Acid Receptor Gamma and Cyclooxygenase-2 respectively. Briefly, the crystal structures of the complexes (2LBD and 5IKR) were subjected to 200 ns all-atom MD simulation under explicit solvent conditions at 300 K and 1 bar atmospheric pressure. Triplicate MD runs with random initial distributions of the atomic velocities were performed amounting to total simulation time of 600 ns for each system.

2.4. MM-PBSA analysis

MM-PBSA stands for Molecular Mechanics Poisson-Boltzmann Surface Area. It is a method of choice for calculating the Gibb's free energy of binding in protein-ligand and protein-protein simulation studies [43–45]. It is considered an intermediate in terms of accuracy and computational cost between empirical scoring and alchemical perturbation techniques [46]. g_mmpbsa was utilized for quantifying the Gibb's free energy of binding of the protein-ligand complexes from the

Table 1
Details of selected drugs for the study from DrugBank.

S No.	DrugBank ID	Approval Status	MW (Average)	Approved Usage
1	DB00755	Approved, Investigational, Nutraceutical	300.4	Acne treatment, photodamaged skin, keratinization disorders and Acute Promyelocytic Leukemia.
2	DB00784	Approved	241.3	Anti-inflammatory agent
3	DB00904	Approved	293.4	Prevention of nausea and vomiting associated with emetogenic cancer chemotherapy
4	DB06697	Approved	298.4	Anti-malarial agent

generated trajectories. It uses the following equations for its calculations.

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

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

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

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

where.

$\Delta G_{complex}$ = total free energy of protein-ligand complex
$\Delta G_{protein}$, ΔG_{ligand} = free energies of protein and ligand in solvent respectively
ΔG_X = free energy for each entity
ΔE_{MM} = vacuum potential energy (total energy of bonded and nonbonded interactions)
ΔG_{solv} = solvation free energy including polar (ΔG_{polar}) and nonpolar ($\Delta G_{nonpolar}$) energies
$T\Delta S$ = entropic contribution of free energy in vacuum. T and ΔS is the temperature and entropy respectively.

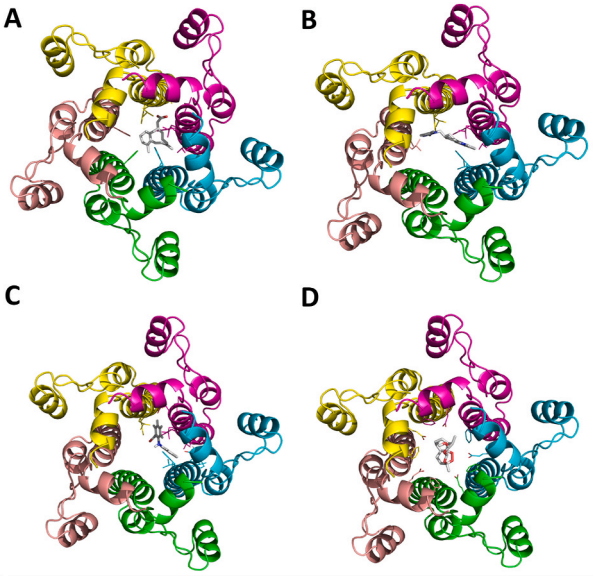
Since the g_mmpbsa tool does not calculate the entropic term (S), the term TΔS MM was not included in this study. Thus, the ΔG bind was calculated as follows, where ΔG_{SASA} is equivalent to $\Delta G_{nonpolar}$ since the calculation of nonpolar solvation energy was based on the SASA model in this study,

$\Delta G_{bind} = (\Delta E_{elec} + \Delta E_{vdw}) + (\Delta G_{polar} + \Delta G_{SASA})$

Comprehensive information regarding the concept and the protocol for use in Gromacs is described [46]. In addition to determining overall free energy of binding between the protein and ligand, residue level contribution towards the overall binding energy was also determined for in-depth insight on the mechanism of binding. The following formula was used,

$$\Delta RxBE = \sum_{i=0}^{n} \left(Ai^{bound} - Ai^{free}\right)$$

where Ai^{bound} and Ai^{free} represent energies of *i*th atom from xth residue with and without ligand whereas n is total number of atoms which makes up the residue.



E

Compound ID	Binding Affinity, ΔG (Kcal/mol)	Hydrophobic interactions
DB00755	-8.6	Leu^{18}, Leu^{21}, Ala^{22}, Val^{25}, Phe^{26}
DB00904	-8.5	Ala^{22}, Val^{25}, Phe^{26}
DB00784	-7.9	Leu^{18}, Leu^{21}, Ala^{22}
DB06697	-7.0	Phe^{4}

Fig. 3. (A–D) Relative positioning of ligands (A: Tretinoin-DB00755; B: Mefenamic acid-DB00784; C: Ondansetron-DB00904 and D: Aremether-DB06697) with respect to SARS-CoV-2 E. Protein is represented as ribbons (Monomer 1–5: yellow, magenta, cyan, green and beige) and ligands are represented as ball and stick. (E) PRODIGY server predicted binding affinity values in ΔG (Kcal/mol), for each tested compound and their interacting partner.

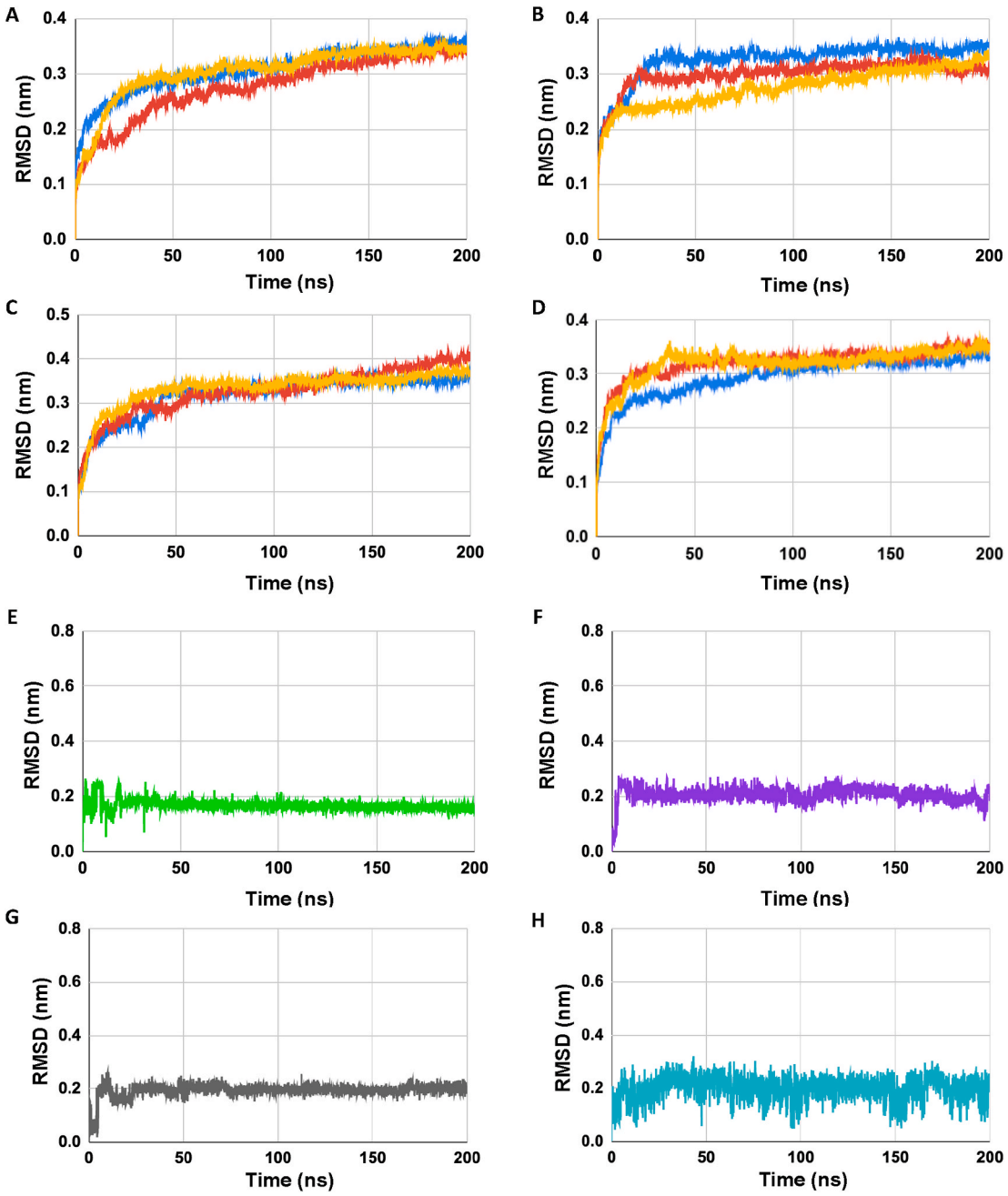


Fig. 4. (A–D) Shows the RMSD plots of the C-alpha backbone of SARS CoV-2 E protein in presence of complexed ligands. (A) SARS CoV-2 E-DB00755, (B) SARS CoV-2 E-DB00784, (C) SARS CoV-2 E-DB00904, and (D) SARS CoV-2 E-DB06697 respectively (triplicate runs). (E–H) shows the RMSD trace of ligands in complex with SARS CoV-2 E protein where, (E) DB00755 (Tretinoin, green), (F) DB00784 (Mefenamic acid, purple), (G) DB00904 (Ondansetron, grey) and (H) DB06697 (Artemether, blue) respectively (). For (A–D) color coding is run1 (blue), run2 (red), run3 (yellow).

3. Results

3.1. Novel coronavirus envelope protein is a membrane spanning viroporin

The envelope or E protein of Coronaviruses is a single transmembrane domain containing protein with hydrophilic termini. Sequence based analysis of SARS-CoV-2 E protein using hydrophobicity plots [Fig. 1A] and multiple transmembrane domain prediction servers [Fig. 1B] suggested the presence of a conserved, prominent, central transmembrane domain between residues 12–39. Sequence composition of the predicted TMD (12–39 aa) of the SARS-CoV-2 E protein indicated the presence of multiple hydrophobic residues like Alanine, Valine, Phenylalanine and Leucine [Fig. 1C]. The N and C termini were hydrophilic, as expected [Fig. 1A]. Both the hydrophobicity profile and the sequence of E were found to be highly conserved between SARS-CoV and SARS-CoV-2 [Fig. 1A and D]. The SARS CoV E protein structure was earlier solved by solution NMR [6]. In the absence of available 3D structures for the SARS-CoV-2 E protein, *in silico* structural prediction was carried out using I-TASSER [Fig. 2A], with the NMR structure of SARS CoV E as the template. The predicted 3D structure of SARS-CoV-2 E protein was dominated by the central helical transmembrane domain with N and C-terminus oriented opposite to each other [Fig. 2A]. Ramachandran plot analysis revealed that 94.5% of the residues are in the favored region with a total of two outliers [Fig. 2C]. The pentameric assembly [Fig. 2B and D] showed the most likely transmembrane

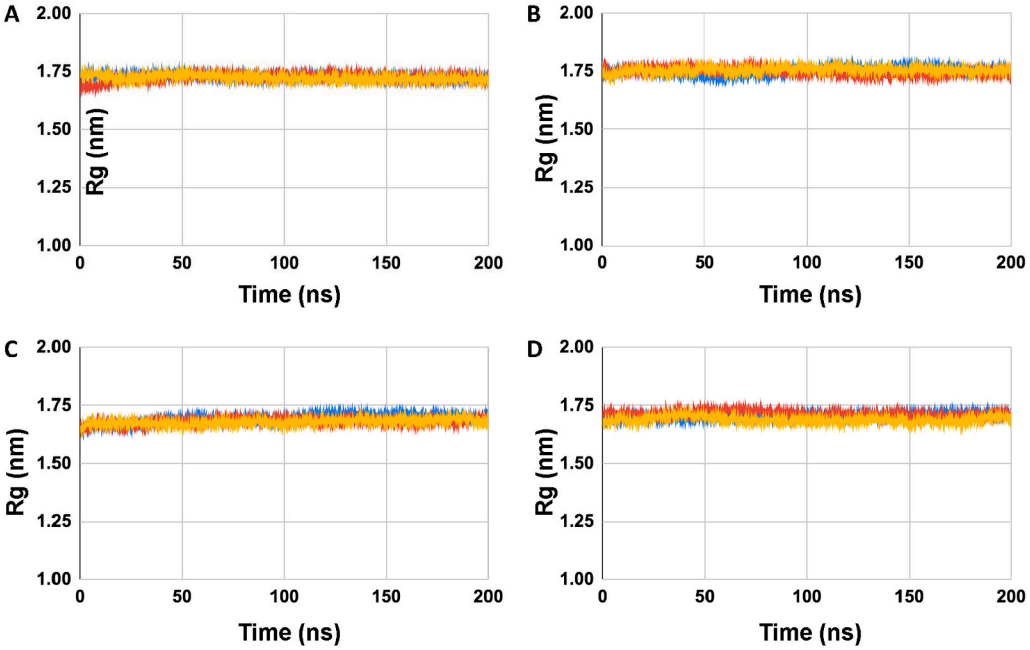


Fig. 5. (A–D) Shows the radius of gyration plots of SARS CoV-2 E protein in presence of complexed ligands. (A) SARS CoV-2 E-DB00755, (B) SARS CoV-2 E-DB00784, (C) SARS CoV-2 E-DB00904, and (D) SARS CoV-2 E-DB06697 respectively (triplicate runs). Color coding is run1 (blue), run2 (red), run3 (yellow).

Table 2
Hydrogen bond occupancy values for SARS-CoV-2 E – drug complexes.

DrugBank ID	Protein-ligand Hydrogen bond	Percentage Occupancy (≥10%)
DB00755	Phe^{23}	58.9
	Phe^{20}	35.6
	Leu^{21}	25.0
	Ala^{22}	13.9
DB00784	Val^{24}	44.9
	Val^{25}	39.5
	Ala^{22}	10.6

channel [47]lined with residues $Val^{14, 25, 29}$, $Leu^{18, 21, 34}$, Ala^{22}, Phe^{26}, Thr^{30},Ile^{33} and Leu^{37} [Fig. 2B and D]contributed by each monomer. Channel lining residues were identified using PoreWalker server [48]. The narrowest region of the pentameric channel, which had a diameter of 6.5 Å, was lined by Phe^{26}contributed by each monomer. This region of the channel formed a constricted hydrophobic pocket [Fig. 2D].

3.2. *Docking studies highlight important residues involved in ligand binding*

Blind molecular docking was carried out with the 15 selected drugs from the DrugBank database using the CB-DOCK web server. Based on the number of binding modes within the channel lumen- Tretinoin

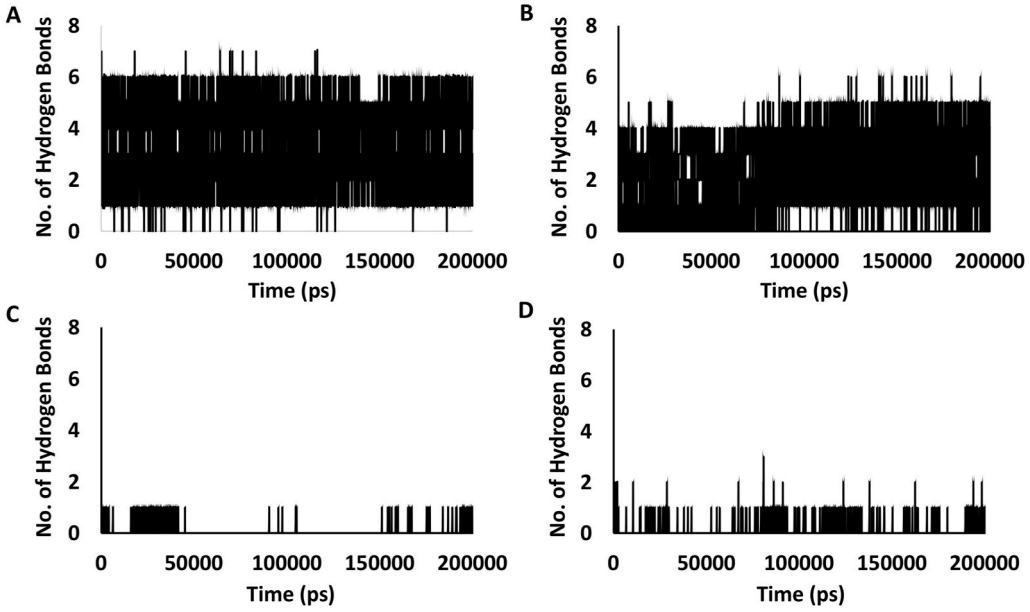


Fig. 6. Graphs showing number of hydrogen bonds formed during the simulation period for the ligands - (A) Tretinoin (DB00755), (B) Mefenamic acid (DB00784), (C) Ondansetron (DB00904) and (D) Artemether (DB06697).

Table 3
Binding free energy (MM-PBSA) analysis of SARS-CoV-2 E-drug complexes.

DrugBank ID	Binding Energy (kJ/mol)	Van der waal's Energy (kJ/mol)	Electrostatic Energy (kJ/mol)	Polar Solvation Energy (kJ/mol)	SASA Energy (kJ/mol)
DB00755	−412.8 ± 12.8	−185.2 ± 13.4	−502.7 ± 14.6	293.0 ± 7.3	−17.9 ± 0.9
DB00784	−381.4 ± 11.9	−191.7 ± 11.1	−424.9 ± 13.2	249.7 ± 6.8	−14.4 ± 0.6
DB00904	−227.1 ± 9.9	−232.5 ± 9.5	−6.8 ± 2.5	28.6 ± 2.2	−16.4 ± 0.7
DB06697	−173.5 ± 9.5	−187.2 ± 8.7	−2.4 ± 0.8	31.5 ± 3.0	−15.4 ± 0.8

(DB00755), Mefenamic acid (DB00784), Ondansetron (DB00904) and Artemether (DB06697) were selected [Table 1, Supplementary Fig. S1] for further analysis. Binding affinity calculation of docked complexes [Fig. 3E] showed that Tretinoin possessed the highest binding affinity of −v8.6 kcal/mol followed by Ondansetron (−8.5 kcal/mol), Mefenamic Acid (−7.9 kcal/mol) and Artemether (−7.0 kcal/mol). Three of the four ligands -Tretinoin, Mefenamic acid and Ondansetron -bound to the narrowest region of the channel, lined by Phe^{26} residues [Fig. 2D]. Artemether, on the other hand, bound towards the N terminal region [Fig. 2D]. In all cases, the ligands interacted via hydrophobic interactions with the protein residues from all chains lining the channel [Fig. 3A–E]. Tretinoin formed a total of 11 interactions involving the residues Leu^{18} (Chain C, D), Leu^{21} (Chain C), Ala^{22} (Chain C), Val^{25} (Chain C) and Phe^{26} (Chains A-E). Ondansetron was observed to form 7 hydrophobic interactions with Ala^{22} (Chain B, C, D), Val^{25} (Chain B) and Phe^{26} (Chain B, C). Mefenamic acid exhibited 6 interactions with residues Leu^{18} (Chain B, C), Leu^{21} (Chain D) and Ala^{22} (Chain B, C) [Fig. 3A–E]. Lastly, Artemether showed interactions with Phe^{4} (Chains A-E) only [Fig. 3A–E].

3.3. Molecular dynamics studies

All-atom molecular dynamics simulation was carried out with the SARS-CoV-2 E protein pentamer and E protein-ligand complexes embedded within a pre-equilibrated POPC membrane for a total of 200 ns. RMSD (Root Mean Square Deviation) analysis of the C-alpha backbones of the E protein monomers showed that the systems were stable during the simulation time period [Fig. 4A-D], Supplementary Fig. S2]. All the four ligands also showed a relatively stable RMSD trace except Artemether which showed relatively large fluctuations during simulation [Fig. 4E-H]. Radius of gyration values of SARS CoV-2 E showed a steady state in presence of all four ligands which indicates that the protein maintained a relatively compact state during the simulation time period [Fig. 5A-D].

Out of four compounds, only Tretinoin and Mefenamic acid formed stable hydrogen bonds [Table 2] with E protein residues within the channel lumen, apart from hydrophobic interactions. Tretinoin formed multiple high occupancy H-bonds (≥10%) with Phe^{20}, Leu^{21} Ala^{22} and Phe^{23}, with the strongest interaction being with Phe^{23} that exhibited an occupancy value of 58.9%. Mefenamic acid on the other hand, formed H-bonds with Ala^{22}, Val^{24} and Val^{25}, with the strongest interaction with Val^{24}having an occupancy value of 44.9%. Quantification of the total number of H-bonds formed during the simulation time period also showed a steady decrease in the number and frequency of H-bonds in the following order (Tretinoin > Mefenamic Acid > Artemether > Ondansetron) [Fig. 6A–D]. In terms of hydrophobic interactions, Tretinoin and Mefenamic acid formed multiple hydrophobic interactions with Val^{17}, Leu^{18}, Leu^{21}, Ala^{22}, Phe^{23}, Val^{25} and Phe^{26}. However, unlike in case of H-bonds, multiple hydrophobic interactions were observed for Ondansetron and Artemether with residues like Phe^{4}, Glu^{8}, Ile^{13}, Leu^{21}, Ala^{22} and Phe^{26}.

3.4. MM-PBSA analysis

To further evaluate the mechanistics of binding of the ligands to the SARS-CoV-2 E protein pentamer during simulation, MM-PBSA analysis was carried out.

It was observed that for Tretinoin and Mefenamic acid, electrostatic energy and polar solvation energies were the major contributors towards the binding energy; whereas, for binding of Ondansetron and Artemether to the channel, van der Waals's energy played the major role [Table 3, Supplementary Table 3]. Out of the four compounds tested,

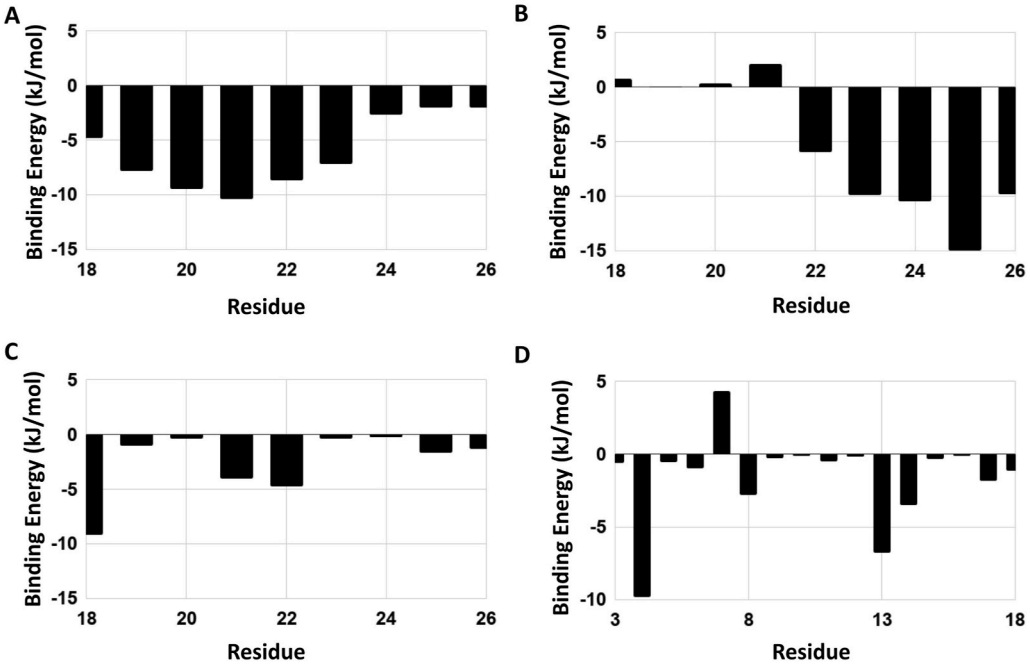


Fig. 7. Binding energy contribution per residue of the binding sites corresponding to the ligands - (A) Tretinoin (DB00755), (B) Mefenamic acid (DB00784), (C) Ondansetron (DB00904) and (D) Artemether (DB06697).

Tretinoin showed the maximum binding energy value of −412.8 kJ/mol followed by Mefenamic acid (−381.4 kJ/mol), Ondansetron (−227.1 kJ/mol) and lastly Artemether (−173.5 kJ/mol) respectively. At the residue level, the major contributors for binding to Tretinoin and Mefenamic acid were Met^{1}, Leu^{19}, Phe^{20}, Leu^{21}, Ala^{22}, Phe^{23}, Val^{24}, Val^{25}, Phe^{26}, Arg^{38}, Lys^{53}, Arg^{61}, Lys^{63} and Arg^{69}. For Ondansetron and Artemether, Phe^{4}, Ile^{13} and Leu^{18} were majorly involved in binding [Fig. 7A-D].

4. Discussion

The current pandemic caused by the novel coronavirus SARS-CoV-2 has resulted in 4.5 million infections and 308,000 deaths so far. To combat this unprecedented public health emergency, developing methods of prevention and cure is crucial. Besides designing a safe and efficacious vaccine, effective treatment options are absolutely essential for already infected patients or identified asymptomatic carriers. Thus, understanding common host-virus interactions, and developing inhibitors targeted towards these processes, appears to be a viable strategy. Structure-based drug design and drug repurposing efforts against protein components of SARS-CoV-2 have primarily focused on the main protease ($3CL^{pro}$), the Spike protein (S) and the viral replicase [52–59], because of their prominent roles in the life cycle of the virus. However, existing data on other coronaviruses indicate that the small E protein has an outsize role in the life cycle of these viruses [9]. The functional similarity of E with viroporins from other virus families, and the previous success in identifying inhibitors for these components [2,6,15,16], indicates that similar strategies may been utilized to identify or repurpose inhibitors against E of SARS-CoV-2.

Recent studies have highlighted the potential of targeting SARS-CoV-2 E protein as a therapeutic strategy. Compounds like Rutin, Doxycycline [60], Nimbolin A [61], Belachinal, Macaflavanone E, Vibsanol B [62] have been proposed as potential inhibitors of the SARS-2 E protein. However, their experimental validation remains. We initiated a thorough search for potential inhibitors of the SARS-CoV-2 E protein by screening 1806 FDA-approved drugs using the DrugScreen web server. The top 50 hits were further screened based on specific attributes of the drugs, resulting in a final list of 15 candidates [Supplementary Table T1]. Four of these candidates (Tretinoin, DB00755; Mefenamic Acid, DB00784; Ondansetron, DB00904 and Artemether, DB06697) were tested for their suitability as inhibitors of SARS-CoV-2 E protein by docking and simulation studies. Ligand binding analysis of docked complexes showed a predominance of hydrophobic interactions in all cases [Fig. 3A–E]. Except Artemether, all other ligands occupied a region in the pentameric channel formed by the E protein. This channel includes protein residues 18–26 and is characterized by the predominance of hydrophobic amino acids like Leucine, Phenylalanine and Valine. The region forms a constriction in the channel diameter, in the form of a hydrophobic cavity. Interestingly, in two studies [61,62] residues Val^{25} and Phe^{26} of the SARS-2 E protein were found to play an important role in binding to potential inhibitors. Our study validates the importance of these residues in forming hydrophobic interactions and intermolecular H-bonds with potential inhibitors. This indicates that the hydrophobic stretch encompassing residues 18–26 might be the major ligand-binding site in SARS-CoV-2 E protein for a variety of small molecules. The blocking of the SARS-CoV-2 E ion channels by small molecules has the potential to strongly inhibit the viroporin activity of E, and consequently nullify its contribution to virus assembly. A previous NMR study on SARS-CoV E protein had demonstrated binding of an inhibitor Hexamethylene Ameloride (HMA) mediated primarily by Asn^{15} [6]. Based on recent literature and our current work, the mode of inhibitor binding appears to be different for the SARS-CoV-2 E protein channel, in spite of strong sequence similarities between the E proteins. This variation might be attributed to the differences in the molecular architecture of the tested compounds.

In our study, Tretinoin formed the strongest hydrophobic interactions within this cavity as evaluated by binding affinities [Fig. 3E]. Control simulations of Tretinoin and Mefenamic acid in complex with their respective known targets (Human Retinoic Acid Receptor Gamma (PDB ID: 2LBD) and Human Cyclooxygenase-2 (PDB ID: 5IKR) respectively) showed comparatively lower binding energy values [Supplementary Table 2], which indicates that there is a strong possibility for these compounds to bind to SARS-2 E protein preferentially, and with higher affinity. The complexes were then subjected to 200 ns of explicit all-atom MD simulations in presence of a lipid membrane. The presence of the membrane was necessary to provide a native-like environment for the pentameric ion channel. Evaluation of protein drug interactions during the simulation time period revealed formation of high occupancy hydrogen bonds in case of Tretinoin and Mefenamic acid apart from hydrophobic interactions [Table 2]. However, Ondansetron and Artemether established H-bonds sporadically [Fig. 5]. This indicated that Tretinoin and Mefenamic acid are able to interact more extensively with the channel. MM-PBSA analysis also showed that Tretinoin exhibited the strongest binding energy value among the four drugs tested. It is hoped that the role of proposed E-protein channel inhibitors on SARS-CoV-2 assembly can be biochemically validated in future.

5. Conclusion

In this work, we have utilized *in silico*-based techniques of molecular docking and MD simulations to identify potential inhibitors against the Envelope (E) protein of SARS-CoV-2, which is a relatively less explored but very critical component of the virus. Based on docking and simulation studies, Tretinoin emerged as the best candidate with the ability to inhibit the viroporin activity of SARS-CoV-2 E and affect virus assembly; and should be further tested *in vitro* and *in vivo*. The other three FDA approved drugs could also be evaluated against SARS-CoV-2 for their ability to inhibit the virus propagation by disrupting the ion channel functionality of E.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank Dr. Manish Agarwal and Mr. Puneet Singh of the HPC Supercomputing Facility PADUM, IIT Delhi for computational support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2020.104063.

References

[1] K. Wang, S. Xie, B. Sun, Viral proteins function as ion channels, Biochim. Biophys. Acta Biomembr. 1808 (2011) 510–515.
[2] D. Dey, et al., The effect of amantadine on an ion channel protein from Chikungunya virus, PLoS Neglected Trop. Dis. 13 (2019).
[3] J. Torres, W. Surya, Y. Li, D.X. Liu, Protein-protein interactions of viroporins in coronaviruses and paramyxoviruses: new targets for antivirals? Viruses 7 (2015) 2858–2883.
[4] D. Schoeman, B.C. Fielding, Coronavirus envelope protein: current knowledge, Virol. J. 16 (2019).
[5] J.L. Nieto-Torres, et al., Subcellular location and topology of severe acute respiratory syndrome coronavirus envelope protein, Virology 415 (2011) 69–82.
[6] K. Pervushin, et al., Structure and inhibition of the SARS coronavirus envelope protein ion channel, PLoS Pathog. 5 (2009).
[7] W. Surya, M. Samso, J. Torres, Structural and functional aspects of viroporins in human respiratory viruses: respiratory syncytial virus and coronaviruses, in:

Respiratory Disease and Infection - A New Insight, 2013, https://doi.org/10.5772/53957.
[8] Y.L. Siu, et al., The M, E, and N structural proteins of the severe acute respiratory syndrome coronavirus are required for efficient assembly, trafficking, and release of virus-like particles, J. Virol. 82 (2008) 11318–11330.
[9] M.L. DeDiego, et al., A severe acute respiratory syndrome coronavirus that lacks the E gene is attenuated in vitro and in vivo, J. Virol. 81 (2007) 1701–1713.
[10] J.L. Nieto-Torres, et al., Severe acute respiratory syndrome coronavirus E protein transports calcium ions and activates the NLRP3 inflammasome, Virology 485 (2015) 330–339.
[11] J.R. Cohen, L.D. Lin, C.E. Machamer, Identification of a golgi complex-targeting signal in the cytoplasmic tail of the severe acute respiratory syndrome coronavirus envelope protein, J. Virol. 85 (2011) 5794–5803.
[12] R.T. Javier, A.P. Rice, Emerging theme: cellular PDZ proteins as common targets of pathogenic viruses, J. Virol. 85 (2011) 11544–11556.
[13] J.M. Jimenez-Guardeño, et al., The PDZ-binding motif of severe acute respiratory syndrome coronavirus envelope protein is a determinant of viral pathogenesis, PLoS Pathog. 10 (2014).
[14] K.T. Teoh, et al., The SARS coronavirus E protein interacts with PALS1 and alters tight junction formation and epithelial morphogenesis, Mol. Biol. Cell **21** (2010) 3838–3852.
[15] S.D. Cady, et al., Structure of the amantadine binding site of influenza M2 proton channels in lipid bilayers, Nature 463 (2010) 689–692.
[16] D. Kozakov, G.Y. Chuang, D. Beglov, S. Vajda, Where does amantadine bind to the influenza virus M2 proton channel? Trends Biochem. Sci. 35 (2010) 471–475.
[17] Prabhat Pratap Singh Tomar, Isaiah T. Arkin, SARS-CoV-2 E protein is a potential ion channel that can be inhibited by Gliclazide and Memantine, Biochem. Biophys. Res. Commun. 530 (1) (2020) 10–14, https://doi.org/10.1016/j.bbrc.2020.05.206.
[18] J. Yang, Y. Zhang, Protein structure and function prediction using I-tasser, Curr. Protoc. Bioinforma. 52 (2015), 5.8.1-5.8.15.
[19] Y. Zhang, I-TASSER server for protein 3D structure prediction, BMC Bioinf. 9 (2008).
[20] W Surya, Y Li, J Torres, NMR structure of the SARS Coronavirus E protein pentameric ion channel, Biochim Biophys Acta 1860 (6) (2018), https://doi.org/10.1016/j.bbamem.2018.02.017, 1317-1317.
[21] L. Heo, H. Park, C. Seok, GalaxyRefine: protein structure refinement driven by side-chain repacking, Nucleic Acids Res. 41 (2013).
[22] M. Baek, T. Park, L. Heo, C. Park, C. Seok, GalaxyHomomer: a web server for protein homo-oligomer structure prediction from a monomer sequence or structure, Nucleic Acids Res. 45 (2017) W320–W324.
[23] D.S. Wishart, et al., DrugBank 5.0: a major update to the DrugBank database for 2018, Nucleic Acids Res. 46 (2018) D1074–D1082.
[24] Y. Liu, et al., CB-Dock: a web server for cavity detection-guided protein–ligand blind docking, Acta Pharmacol. Sin. 41 (2020) 138–144.
[25] S. Hashmi, S. Al-Salam, AutoDock Vina: improving the speed and, Int. J. Clin. Exp. Pathol. 8 (2015) 8786–8796.
[26] Á. Jász, Á. Rák, I. Ladjánszki, G. Cserey, Optimized GPU implementation of merck molecular force field and universal force field, J. Mol. Struct. 1188 (2019) 227–233.
[27] O. Molecules, A universal force field, Molecules (1992) 10035–10046.
[28] S.K. Mishra, B. Ram, S.K. Mishra, B. Ram, Steepest descent method, in: Introduction to Unconstrained Optimization with R, 2019, pp. 131–173, https://doi.org/10.1007/978-981-15-0894-3_6.
[29] S.S. Petrova, A.D. Solov'Ev, The origin of the method of steepest descent, Hist. Math. 24 (1997) 361–375.
[30] M.D. Hanwell, et al., Avogadro: an advanced semantic chemical editor, visualization, and analysis platform, J. Cheminf. 4 (2012).
[31] E.F. Pettersen, et al., UCSF Chimera, J. Comput. Chem. 25 (2004) 1605–1612.
[32] L.C. Xue, J.P. Rodrigues, P.L. Kastritis, A.M. Bonvin, A. Vangone, PRODIGY: a web server for predicting the binding affinity of protein-protein complexes, Bioinformatics 32 (2016) 3676–3678.
[33] S. Salentin, S. Schreiber, V.J. Haupt, M.F. Adasme, M. Schroeder, PLIP: fully automated protein-ligand interaction profiler, Nucleic Acids Res. 43 (2015) W443–W447.
[34] W.L. DeLano, Pymol: an open-source molecular graphics tool, CCP4 Newsl. Protein Crystallogr. 40 (2002) 82–92.
[35] PyMOL. Schrödinger, Molecular visualization system, Available at: https://pymol.org, 2018.
[36] A.W. Schüttelkopf, D.M.F. Van Aalten, PRODRG: a tool for high-throughput crystallography of protein-ligand complexes, Acta Crystallogr. Sect. D Biol. Crystallogr. 60 (2004) 1355–1363.
[37] S.M. Gopal, A.B. Kuhn, L.V. Schäfer, Systematic evaluation of bundled SPC water for biomolecular simulations, Phys. Chem. Chem. Phys. 17 (2015) 8393–8406.
[38] D.J. Evans, B.L. Holian, The Nose-Hoover thermostat, J. Chem. Phys. 83 (1985) 4069–4074.
[39] M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: a new molecular dynamics method, J. Appl. Phys. 52 (1981) 7182–7190.
[40] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaije, LINCS: a linear Constraint solver for molecular simulations, J. Comput. Chem. 18 (1997) 1463–1472.
[41] U. Essmann, et al., A smooth particle mesh Ewald method, J. Chem. Phys. 103 (1995) 8577–8593.
[42] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph. 14 (1996) 33–38.
[43] S. Borkotoky, C.K. Meena, A. Murali, Interaction analysis of T7 RNA polymerase with heparin and its low molecular weight derivatives - an In-silico approach, Bioinf. Biol. Insights 10 (2016) 155–166.
[44] H. Gohlke, D.A. Case, Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex ras-raf, J. Comput. Chem. 25 (2004) 238–250.
[45] N. Tran, T. Van, H. Nguyen, L. Le, Identification of novel compounds against an R294K substitution of influenza A (H7N9) virus using ensemble-based drug virtual screening, Int. J. Med. Sci. 12 (2015) 163–176.
[46] R. Kumari, R. Kumar, O.S.D.D. Consortium, A. g Lynn, _ mmpbsa - a GROMACS tool for MM-PBSA and its optimization for high-throughput binding energy calculations, J. Chem. Inf. Model. 54 (2014) 1951–1962.
[47] T. Bin Masood, S. Sandhya, N. Chandra, V. Natarajan, CHEXVIS: a tool for molecularchannel extraction and visualization, BMC Bioinf. 16 (2015).
[48] M. Pellegrini-Calace, T. Maiwald, J.M. Thornton, PoreWalker: a novel tool for the identification and characterization of channels in transmembrane proteins from their three-dimensional structure, PLoS Comput. Biol. 5 (2009).
[49] M. Cserzo, E. Wallin, I. Simon, G. von Heijne, A. Elofsson, Prediction of Transmembrane Alpha-Helices in Prokaryotic Membrane Proteins: the Dense Alignment Surface Method. Protein Engineering Design and Selection, 1997.
[50] A. Krogh, B. Larsson, G.V. Heijne, E.L. Sonnhammer, Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes, J. Mol. Biol. 305 (3) (2001) 567–580, https://doi.org/10.1006/jmbi.2000.4315.
[51] L. Käll, A. Krogh, E.L.L. Sonnhammer, Advantages of combined transmembrane topology and signal peptide prediction-the Phobius web server, Nucleic Acids Res. 35 (2007), https://doi.org/10.1093/nar/gkm256.
[52] I. Aanouz, et al., Moroccan Medicinal plants as inhibitors of COVID-19: computational investigations, J. Biomol. Struct. Dyn. (2020) 1–12, https://doi.org/10.1080/07391102.2020.1758790.
[53] S. Boopathi, A.B. Poma, P. Kolandaivel, Novel 2019 coronavirus structure, mechanism of action, antiviral drug promises and rule out against its treatment, J. Biomol. Struct. Dyn. (2020) 1–14, https://doi.org/10.1080/07391102.2020.1758788.
[54] A.D. Elmezayen, A. Al-Obaidi, A.T. Şahin, K. Yelekçi, Drug repurposing for coronavirus (COVID-19): in silico screening of known drugs against coronavirus 3CL hydrolase and protease enzymes, J. Biomol. Struct. Dyn. (2020) 1–12, https://doi.org/10.1080/07391102.2020.1758791.
[55] S.K. Enmozhi, K. Raja, I. Sebastine, J. Joseph, Andrographolide as a potential inhibitor of SARS-CoV-2 main protease: an in silico approach, J. Biomol. Struct. Dyn. (2020) 1–10, https://doi.org/10.1080/07391102.2020.1760136.
[56] R.S. Joshi, et al., Discovery of potential multi-target-directed ligands by targeting host-specific SARS-CoV-2 structurally conserved main protease, J. Biomol. Struct. Dyn. (2020) 1–16, https://doi.org/10.1080/07391102.2020.1760137.
[57] S.A. Khan, K. Zia, S. Ashraf, R. Uddin, Z. Ul-Haq, Identification of chymotrypsin-like protease inhibitors of SARS-CoV-2 via integrated computational approach, J. Biomol. Struct. Dyn. (2020), https://doi.org/10.1080/07391102.2020.1751298.
[58] O. Mitjà, B. Clotet, Use of antiviral drugs to reduce COVID-19 transmission, The Lancet Global Health 8 (2020) e639–e640.
[59] C. Liu, et al., Research and development on therapeutic agents and vaccines for COVID-19 and related human coronavirus diseases, ACS Cent. Sci. 6 (2020) 315–331.
[60] D. Bhowmik, R. Nandi, R. Jagadeesan, N. Kumar, A. Prakash, D. Kumar, Identification of potential inhibitors against SARS-CoV-2 by targeting proteins responsible for envelope formation and virion assembly using docking based virtual screening, and pharmacokinetics approaches, Infect. Genet. Evol. 84 (2020) 104451, https://doi.org/10.1016/j.meegid.2020.104451 [published online ahead of print, 2020 Jul 5].
[61] S. Borkotoky, M. Banerjee, A computational prediction of SARS-CoV-2 structural protein inhibitors from *Azadirachta indica* (Neem), J. Biomol. Struct. Dyn. (2020) 1–17, https://doi.org/10.1080/07391102.2020.1774419. Advance online publication.
[62] Manoj Gupta, Sarojamma Vemula, Ravindra Donde, Gayatri Gouda, Lambodar Behera, Ramakrishna Vadde, In-silico approaches to detect inhibitors of the human severe acute respiratory syndrome coronavirus envelope protein ion channel, Journal of Biomolecular Structure and Dynamics (2020) 1–11, https://doi.org/10.1080/07391102.2020.1751300.