Theaflavin 3-gallate inhibits the main protease (Mpro) of SARS-CoV-2 and reduces its count in vitro

Mahima Chauhan
CSIR-Institute of Himalayan Bioresource Technology
Vijay Kumar Bhardwaj
CSIR-Institute of Himalayan Bioresource Technology
Asheesh Kumar
CSIR-Institute of Himalayan Bioresource Technology
Vinod Kumar
CSIR-Institute of Himalayan Bioresource Technology
Pawan Kumar
CSIR-Institute of Himalayan Bioresource Technology
M. Ghalib Enayathullah
Centre for Cellular and Molecular Biology
Jessie Thomas
Centre for Cellular and Molecular Biology
Joel George
Centre for Cellular and Molecular Biology
Bokara Kiran Kumar
Centre for Cellular and Molecular Biology
Rituraj Purohit
CSIR-Institute of Himalayan Bioresource Technology
Arun Kumar (arunkumar@ihbt.res.in)
CSIR-Institute of Himalayan Bioresource Technology
Sanjay Kumar
CSIR-Institute of Himalayan Bioresource Technology

Article

Keywords: SARS-CoV-2, Theaflavin 3-gallate, Mpro, Steered MD-simulations

Posted Date: March 15th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1444367/v1

Abstract

Main protease (M^{pro}) of SARS-CoV-2 is crucial for its replication/infection and has been recognized as an attractive drug target. In this study, we identified theaflavin 3-gallate as an inhibitor of M^{pro} protein of SARS-CoV-2 with IC_{50} value of 18.48 ± 1.29 μM. Compared to theaflavin, theaflavin 3-gallate exhibited superior antiviral activity and at a concentration of 200 μM reduced the viral count by 75% (viral particles reduced from $10^{6.7}$ to $10^{6.1}$). Time-dependent analyses of conventional and steered MD-simulations revealed stronger interactions of theaflavin 3-gallate with the active site residues of M^{pro} than the standard molecule GC373 and theaflavin. Taken together, our findings suggest that theaflavin 3-gallate can be developed into a potential lead molecule against SARS-CoV-2.

1. Introduction

The ongoing COVID-19 pandemic due to SARS-CoV-2 has paralyzed the whole world, which motivated the scientific community around the world to find possible remedies[1]. The COVID-19 outbreak originated in December 2019 has developed into a global pandemic in just a few months, spreading to about more than 222 countries, areas, or territories[2–4]. The SARS-CoV-2 belongs to the family of enveloped, single-stranded, positive-sense, and very diverse RNA viruses[1]. The genome of SARS-CoV-2 is composed of about 30,000 nucleotides: its replicase gene encodes for two overlapping polyproteins namely, pp1a and pp2ab, required for the replication and transcription of virus[4–6]. The polyproteins are proteolytically processed, mainly by the 33.8-kDa main protease (M^{pro}) also known as 3C-like protease[7]. M^{pro} undergoes autolytic cleavage from polyproteins pp1a and pp1ab following which it cleaves the polyprotein at 11 conserved sites[7]. Such functional importance of M^{pro} in the life cycle of the virus along with the absence of its closely related homologs in human beings recognize M^{pro} as an attractive target for the anti-viral drug designs[7]. Though several vaccines against SARS-CoV-2 have been developed and granted emergency use authorization by the Food and Drug Administration, there is still uncertainty about their long-term side effects and no strong scientific data is available regarding the safety of these vaccines for pregnant or breastfeeding women[8]. Similar is the case for children under 16–18 years of age as clinical trials in adolescents have only just begun and it will take time to reach certain decisions. Therefore, potential molecules that can inhibit the viral replication and disrupt the interaction between the viral protein and host receptor can add significant value in managing COVID-19[9]. To facilitate rapid drug discoveries or drug repurposing, several antiviral compounds were screened *in-silico* in various laboratories all around the world but the studies still lack sufficient wet lab experimentations following clinical trials to support their claims.

Many natural molecules, and their derivatives have entered different stages of drug design, including clinical trials against diseases like cardiovascular diseases, malaria, Human immunodeficiency virus-AIDS, etc.[10–12], and more than 50% of approved drugs are reported to be based on natural products[13]. Therefore, natural molecules have caught the attention of researchers to hunt for anti-SARS-CoV-2 drug

candidates and several natural molecules belonging to different classes and sources of origin have been explored through *in silico* and experimental investigations[14,15]. Natural molecules like theaflavin, rutin, curcumin, salvionic acid, and many flavones were reported to have anti-SARS-CoV-2 activity[14–18]. Among natural molecules, tea polyphenols are well known to have anti-HIV effect anti-cancerous, anti-oxidative, anti-mutagenic, and anti-diabetic, and hypocholesterolemic activities[19,20]. The beneficial effects of green tea, oolong tea, and black tea are well-known since long ago[21]. Recent studies have also shown anti-SARS-CoV-2 potential of several bioactive tea molecules. For example, epigallocatechin gallate was shown to disrupt the interaction of spike protein with ACE2 receptors[22]; epicatechin-3,5-di-O-gallate, epigallocatechin-3,5-di-O-gallate, and epigallocatechin-3,4-di-O-gallate as inhibitors of RNA-Dependent RNA Polymerase[23]; epigallocatechin gallate, epicatechin gallate and gallocatechin-3-gallate as potential inhibitors of M^{pro} [24]. Through *in-silico* studies, our group have previously recognized three lead molecules, oolonghomobisflavan-A, theasinensin-D, and theaflavin-3-O-gallate as a potential inhibitors of M^{pro} protein of SARS-CoV-2 with docking scores higher than repurposed drugs Atazanavir, Darunavir, and Lopinavir[23]. Since the average content of oolonghomobisflavan-A (1.1 mg/100g) and theasinensin-D (1.8 mg/100g) as compared to theaflavin 3-gallate (148.6 mg/100g) in different types of tea is lower, we selected the later for detailed studies[25]. Additionally, oolonghomobisflavan-A and theasinensin-D are poorly soluble in water[26] while theaflavin and its derivatives are highly soluble[27]. Considering the above stated reasons and the huge therapeutic potential of theaflavin 3-gallate, we performed the lacking wet-lab studies to test the inhibition potential of this molecule along with control molecules GC376 and theaflavin (known inhibitors of M^{pro}) against M^{pro} protein of SARS-CoV-2 and tested it on the live virus under *in vitro* conditions. Further, *in silico* studies were conducted to shed insights into the mechanism of action of theaflavin 3-gallate on M^{pro}.

2. Results

2.1 Theaflavin 3-gallate inhibited M^{pro} protein of SARS-CoV-2

M^{pro} was inhibited by more than 80% after incubation with 100 μM concentration of theaflavin and theaflavin 3-gallate for 30 min. Accordingly, it was tested for inhibition by these molecules at lower concentrations. The IC_{50} of theaflavin and theaflavin 3-gallate against M^{pro} protein was calculated to be 22.22 ± 1.4 μM and 18.48 ± 1.29 μM, respectively (Fig. 1). A known inhibitor of M^{pro} protein named GC376 was kept as a positive control. The IC_{50} for GC376 was calculated to be 0.24 ± 0.04 μM (Fig. 1) which is slightly lower than the reported value of 0.42 μM with the kit.

2.2. Incubation of SARS-CoV-2 with theaflavin 3-gallate reduced viral count *in vitro* in a dose-dependent manner

Incubation of theaflavin and theaflavin 3-gallate with SARS-CoV-2 resulted in the reduction of viral count to different extents (Fig. 2). At lower concentrations of 12.50, 25 μM, theaflavin and theaflavin 3-gallate could reduce the viral count to different extents below 24%. However, the viral count was reduced by 26% with theaflavin and 42% with theaflavin 3-gallate at 100 μM. At 200 μM, theaflavin and theaflavin 3-gallate showed a reduction of 40% and 75%, respectively. Treatment with theaflavin 3-gallate at 200 μM reduced the viral particles from $10^{6.7}$ to $10^{6.1}$. As anticipated, treatment with remdesivir (positive control) resulted in inhibition of virus at a concentration of 0.5, 0.75 and 1.0 μM (Fig. 2).

2.3. Theaflavin 3-gallate interacts with the active site residues of M^{pro}

The availability of crystal structures of protein allowed us to generate and visualize the interactions between protein and their respective ligands. The optimal poses of GC373 (the active form of the prodrug GC376), theaflavin, and theaflavin 3-gallate with the M^{pro} of SARS-CoV-2 were generated by employing molecular docking methodology. The best poses with the highest docking score were selected, as shown in Fig. 3. The docking scores in terms of interaction energy for GC373, theaflavin, and theaflavin 3-gallate were 50.54 kcal/mol, 57.41 kcal/ml, and 74.35 kcal/mol respectively. The standard inhibitor (GC373) formed three conventional hydrogen bonds (H-bonds), two carbon hydrogen bonds, and pi-pi interactions with the active site of M^{pro}. Theaflavin showed higher number of H-bonds than the standard molecule. A total of eight H-bonds were observed between theaflavin and the binding pocket residues of M^{pro}. The residue Glu166 interacted via three H-bonds, while the residues His41, His164, Pro168, Arg188, and Gln189 formed H-bond each. Among the docked molecules, theaflavin 3-gallate showed the highest number of H-bonds. Theaflavin 3-gallate established eleven H-bonds with residues of the active site of SARS-CoV-2 M^{pro}. The residues involved in H-bonds with theaflavin 3-gallate were Ser46, Ser144, Asn142, Gly143, Met165, Glu166, Pro168, Asp187, and Arg188. The pi-pi interactions were shown by residues Met49 and Cys145 of the binding pocket. All the three molecules interacted with the key residues of the binding site. Many natural molecules were shown to target the same binding site to inhibit the M^{pro} of SARS-CoV-2. The stability of binding poses of theaflavin 3-gallate with M^{pro} were further accessed by performing MD simulations.

2.4. MD simulation studies suggested stable protein topology and global stability of the protein-ligand complexes

MD simulations present various time-dependent analyses to analyze protein-ligand dynamics at the molecular level. We calculated the root mean square deviation (RMSD) of backbone C-α atoms to analyze the effect of ligand binding on protein topology (Fig. 4a). We observed initial deviations from the starting structures during the simulation in RMSD values for all the three trajectories. The RMSD values for protein structure with theaflavin stabilized at ~ 10 ns and followed the same trajectory till the end of simulation. On the other hand, the RMSD values proteins with GC373 and theaflavin 3-gallate stabilized

towards the end of the simulation. However, the average RMSD values never exceeded over 0.4 nm during the simulation. The average RMSD values for the entire simulation run for GC373, theaflavin, and theaflavin 3-gallate were 0.27 nm, 0.22 nm, and 0.36 nm, respectively. The low RMSD values and stable trajectories validated the structural stability and indicated that the ligand binding had no impact on the overall protein topology.

2.5. Theaflavin 3-gallate formed higher number of H-bonds than GC373 and theaflavin

We analyzed the total number of H-bonds formed between M^{pro} and selected molecules during the entire simulation (Fig. 4b). Among the three molecules, theaflavin 3-gallate formed highest number of H-bonds throughout the entire simulation run. The average number of H-bonds for GC373, theaflavin, and theaflavin 3-gallate were 2.15, 3.26, and 6.85, respectively. We observed no H-bonds between GC373 and M^{pro} after 80 ns of simulation time. A few conformations of M^{pro}-theaflavin 3-gallate complex showed upto 13 H-bonds during the simulation. These results suggest high potential of theaflavin 3-gallate over theaflavin and GC373 to interact with the active site of M^{pro}.

2.6. Theaflavin 3-gallate showed exclusive interactions with binding site of M^{pro}

The MD trajectories could be used to visualize the interactions between protein and ligands at different time intervals during the simulation. We also took snapshots of protein-ligand complexes at different time intervals to analyze the interaction pattern of GC373 (Supplementary Fig. S2), theaflavin (Supplementary Fig. S3), and theaflavin 3-gallate (Fig. 5) with the binding sites of M^{pro}. The standard molecule GC373 initially interacted by showing H-bonds, pi-pi, and hydrophobic interactions, while towards the end of simulation, it only interacted with the active site residues via hydrophobic interactions. The binding site of M^{pro} formed the greatest number of H-bonds with theaflavin 3-gallate in all the conformations when compared to GC373 and theaflavin. We also observed formation of some exclusive interactions (Val186, His164, His41) at different time intervals between M^{pro} and theaflavin 3-gallate during the simulation. These interactions were not observed for GC373 and theaflavin. Also, theaflavin formed weaker hydrophobic interactions with residue Cys145, while theaflavin 3-gallate interacted with cys145 by stronger H-bonds, pi-sulfur, and pi-alkyl interactions. These results confirmed that theaflavin 3-gallate adhered tightly to the binding site by interacting with different residues through the entire course of simulation.

2.7. Van der Wall and electrostatic energies majorly contributed to tight binding of theaflavin 3-gallate to M^{pro} protein of SARS-CoV-2

The Molecular Mechanics Poisson-Boltzmann Surface Area (MMPBSA) is a well-established end-state post-processing approach to calculate the free energy of binding of the ligands to their respective receptors. The total binding energy obtained by the MMPBSA approach is a cumulative sum of individual energies including the polar solvation energy, van der Waal energy, SASA energy, and Electrostatic energy. The MMPBSA results are summarized in Table 1. Among the selected molecules, theaflavin 3-gallate showed the least binding energy (highest affinity) of -47.62 kcal/mol, followed by theaflavin and GC373. The van der Waal energy contributed most significantly to the binding between ligands and M^{pro} of SARS-CoV-2. The non-favorable contributions by the polar solvation energy were highest for theaflavin 3-gallate. However, the large difference in positive contributions by van der Waal and electrostatic energy contributed to the least binding energy of theaflavin 3-gallate in comparison to GC373 and theaflavin. The final binding energy of theaflavin 3-gallate provided a strong rationale for its high inhibition potential against both the proteins. The per-frame trajectories of binding energies for all the three complexes throughout the simulation are provided in Supplementary Fig. S4.

Table 1
The components of binding free energy calculated by the MMPBSA method.

Complex	van der Waal energy (kcal/mol)	Electrostatic energy (kcal/mol)	Polar solvation energy (kcal/mol)	SASA energy (kcal/mol)	Binding energy (kcal/mol)
GC373	-21.96	-5.57	21.29	-2.89	-9.13
Theaflavin	-35.50	-11.44	39.93	-4.19	-11.21
Theaflavin 3-gallate	-66.44	-18.35	42.33	-5.15	-47.62

2.6. A strong external pulling force is required to unbind theaflavin 3-gallate from its respective binding sites on M^{pro}

SMD simulations can offer the qualitative perspectives of interactions and conformational perturbations between ligand and its neighboring amino acids by inducing ligand unbinding along the MD simulation pathway. In this study, the well equilibrated protein-ligand complexes were taken as starting points for carrying out SMD simulations. A pulling simulation was performed on all the complexes with a constant pulling velocity of 0.01 nm/ps and a spring constant of 250 kJ/mol/nm^2. The representative force profiles of GC373, theaflavin, theaflavin 3-gallate unbinding from binding sites of the M^{pro} are shown in Fig. 6. A linear increase in the time-dependent external pulling force was observed for the initial phase of SMD simulation for all the complexes. The standard molecule GC373 required the lowest amount of

external pulling force to completely unbind from the binding pocket. The peak values of external pulling force experienced by theaflavin and theaflavin 3-gallate were ~ 350.02 kJ/mol/nm and ~ 368.47 kJ/mol/nm respectively. Afterward, the external pulling force gradually decreased and the ligand was completely out of the binding site. The peak of pull force trajectory for theaflavin 3-gallate lied towards the right of theaflavin and GC373, suggesting that theaflavin was bound to the active site for a longer duration than both the compounds. These results show that theaflavin 3-gallate required a strong external pulling force to separate it from the binding site of M^{pro} of SARS-CoV-2.

4. Discussion And Conclusions

Within the past two decades, three extremely pathogenic and deadly human coronaviruses namely, SARS-CoV, MERS, SARS-CoV-2 have emerged[28]. They belong to the group β-coronavirus and draw more attention because of their ability to cross animal-human barriers, ultimately becoming major human pathogens[29]. The timely development of antiviral substances is of utmost importance, which in such a short span is extremely challenging. Moreover, the conventional drug takes years to develop and to get into the market. Drug repurposing is a choice in which the molecules/compounds already known for their therapeutic effects can be screened and tested for inhibition of SARS-CoV-2.

Since time immemorial, natural compounds have been used as therapeutic agents for the treatment of several ailments[30]. Plant-based natural compounds offer a rich reservoir for novel antiviral drug development. Some natural medicines have been reported to possess antiviral activities against many virus strains including herpes simplex virus, coronavirus, influenza virus, hepatitis B and C viruses, and HIV virus[31–33]. Our previous *in-silico* studies showed that tea molecule theaflavin-3-O-gallate has the affinity to bind to the substrate-binding pocket of M^{pro} of SARS-CoV-2[34]. To quantify the inhibition on a scale, we tested theaflavin 3-gallate against M^{pro} protein of SARS-CoV-2, and calculated its IC_{50}. The IC_{50} is a quantitative measure that represents how much of a particular inhibitory substance is needed to inhibit a given biological process or biological component by 50% under *in vitro* conditions. With our aforesaid method, the IC_{50} values of theaflavin 3-gallate against M^{pro} was calculated to be 18.48 ± 1.29 μM. Also, incubation of theaflavin 3-gallate with SARS-CoV-2 led to inhibition of the virus as evident by reduction of viral count by 75% at a concentration of 200 μM.

To understand the mechanism of inhibitory action of theaflavin 3-gallate we performed molecular docking studies which provide a platform to predict the estimated binding affinity and optimal binding pose between receptor and ligand. The binding poses were compared with a standard drug GC373 and theaflavin. This method has been implemented for more than three decades and a significant number of experimental drugs have been identified and developed accordingly[14,15,35]. Our molecular docking analysis showed that theaflavin 3-gallate interacts strongly with the binding sites on M^{pro} with a higher docking score than GC373 and theaflavin. Our results showed that Theaflavin 3-gallate acquired the active site of M^{pro} by interacting with many residues (Asn142, Ser144, His145 His163, and Glu166) crucial for dimerization and biological activity. Many experimental and computational studies have

shown potent inhibitor molecules interacting with these residues[7,16,17,23,36,37]. The stability of binding poses was validated by analysis of different MD-driven time-dependent analyses. The low deviations in RMSD values for all the three structures suggested that the binding of ligands on protein had no impact on stability of the binding pocket. We also compared the number of H-bonds formed by GC373, theaflavin, and theaflavin 3-gallate with the binding site of M^{pro} throughout the simulation. The analysis of the H-bond profiles of the three molecules showed that theaflavin 3-gallate formed highest number of bonds during the simulation. To get an in-depth insight of the molecular interactions during the simulations, protein-ligand conformations at different time intervals were extracted from the MD trajectories. In comparison to GC373 and theaflavin, theaflavin 3-gallate formed exclusive H-bonds with residues His41, Val186, and His164 at different time intervals during the simulation. These interactions were not observed for GC373 and theaflavin. Also, theaflavin 3-gallate interacted with residue Cys145 by much stronger interactions than theaflavin. The residue Cys145 is involved in the formation of catalytic dyad of the active site of M^{pro} [7]. The strong interactions shown by theaflavin 3-gallate with M^{pro} were further validated and compared with GC373 and theaflavin by caculating the binding free energy by the MM-PBSA method. The MM-PBSA is an efficient, reliable method for evaluating protein-ligand interactions[38,39]. The binding free energy results confirmed that theaflavin 3-gallate showed the highest binding affinity for M^{pro} than GC373 and theaflavin. Our results showed that the van der Waal energy contributed most favorably to the binding of theaflavin 3-gallate with M^{pro}. Moreover, we also performed SMD simulations to analyze the amount of external force required for unbinding GC373, theaflavin, and theaflavin 3-gallate from the binding pocket. Our SMD results demonstrated that theaflavin 3-gallate required highest amount of external force to unbind it from the binding pocket of the M^{pro} of SARS-CoV-2. Also, theaflavin 3-gallate remained at the active site for a longer duration than both the GC373 and theaflavin during the pulling simulations, suggesting strong interactions with the residues of the binding site.

In conclusion, *in-silico* and experimental results suggested theaflavin 3-gallate as a potential candidate molecule that could be rapidly developed as a therapeutic agent to fight COVID-19. Theaflavin 3-gallate performed better than the standard molecule GC373 and theaflavin in both the *in-silico* and experimental analyses. Theaflavin 3-gallate is a major component of black which is already known for its anti-oxidant properties and is the most consumed beverage in the world. Since theaflavin 3-gallate is already consumed by humans through tea for ages and being edible, crossing cell-cell barriers in the body[40] makes it a good potential inhibitor to be used against SARS-CoV-2. Either the molecule alone or in formulations with other such anti-viral compounds as a cocktail can provide an effective first line of defense against diseases associated with coronaviruses.

5. Material And Methods

Theaflavin (molecular weight = 564.49 g/mol) and theaflavin 3-gallate (Molecular weight = 716.61g/mol) were procured from (PhytoLab GmbH & Co. KG, Germany). The authenticity and purity of the molecules were validated by Mass spectrometry analysis using UHPLC-IM-QTOF 6560 instrument (Agilent, USA)

equipped with PDA detector and hyphenated to the Q-TOF MS/MS (Supplementary Fig. S1). Stock solutions (5 mM) of theaflavin and theaflavin 3-gallate were prepared in methanol. Appropriate dilutions were made in assay buffer to test the molecules at various concentrations keeping methanol concentration less than 1% in the final reaction volume.

All the assays were carried out at room temperature (RT: 25 ± 2°C) in triplicates. As an initial screen theaflavin and theaflavin 3-gallate were tested for inhibition of M^{pro} protein at 100 µM concentration and later tested at lower concentrations to calculate the IC_{50} value.

5.1. *In vitro* inhibition assay of M^{pro}

MBP-tagged (SARS-CoV-2) assay kit was used (Cat. No. Catalog #79955-1; BPS Bioscience, USA) to study the inhibition of M^{pro} by theaflavin 3-gallate. GC376 and theaflavin (known inhibitors of M^{pro}) was kept as positive controls. In brief, reactions were performed in a final assay volume of 50 µl in a 96 well plate wherein each reaction consisted of 30 µl of M^{pro} (150 ng) prepared in an assay buffer containing 1 mM DTT, 10 µl of fluorogenic protease substrate (DABCYL-KTSAVLQSGFRKME-EDANS) to a final concentration of 50 µM, and 10 µl of either theaflavin or theaflavin 3-gallate. Appropriate assay control was also kept by replacing inhibitor with inhibitor buffer. Sample controls were included to account for any interference caused by color of the molecules. The plate was sealed with the plate sealer and incubated overnight at RT. The fluorescence intensity was measured in a microtiter plate-reading fluorimeter (Biotek Synergy™ H1, USA) with an excitation wavelength of 360 nm and detection of emission at a wavelength of 460 nm. All the values were corrected by subtracting the blank values. The percent inhibition was calculated using the following formula: % inhibition = [(positive control – test inhibitor)/positive control] × 100.

5.3. Calculation of IC_{50}

To calculate the inhibitory concentration (IC_{50}) value, theaflavin and theaflavin 3-gallate were tested at concentrations of 0.1, 0.6, 1.0, 5.0, 10, 25, 50, 75, and 100 µM against M^{pro} protein of SARS-CoV-2. GC376 was tested at 0.1, 0.2, 0.5, 1, and 100 µM. The inhibitor concentration was plotted against the percent inhibition and the IC_{50} value was calculated using the non-linear regression equation of the resulting graph using GraphPad Prism software version 8.0.2.

5.4. Drug treatment and detection of SARS-CoV-2 using an RT-qPCR assay

The individual antiviral effect of theaflavin and theaflavin 3-gallate were tested against the SARS-CoV-2 (Indian/a3i clade/2020 isolate) virus *in-vitro*, using Vero cells (Green African Monkey). The vero cells, prior to infection, were maintained in Dulbecco Minimum Essential Medium (DMEM, Gibco) with 10% Fetal Bovine Serum (Gibco) at 37□C, 5% CO2. 24 hours prior to infection, the cells were trypsinized and seeded in a 96 well plate. 80–90% of cell confluency was considered for infection. Briefly, 50, 100, 150 &

200 (μM) concentrations of theaflavin and theaflavin 3-gallate was used to enumerate the antiviral effect. Initially the cells were primed with different concentrations of theaflavin and theaflavin 3-gallate for 2 hours. Later, the medium containing the drug candidates was replaced with the virus inoculum (0.1 MOI) along with drug dilutions made in DMEM without FBS for 3 hours. Post-infection, the media was replaced with DMEM media along with 10% FBS containing drugs and maintained in an incubator at 37□C, 5% CO2 until 72 hours. Post 72 hours, the cell supernatants were collected for enumerating the cell released viral RNA particles using quantitative real time PCR.

Viral RNA extraction was performed via the Vira/Pathogen Extraction Kit (Applied Biosystems, Thermo Fisher Scientific) according to the protocol described by manufacturer. Viral supernatants (200μL) after centrifugation from the experimental groups were aliquoted into deep well plates and were subjected with, the lysis buffer that contained 260μL, Binding Solution; 10μL, Binding Beads; 5μL Proteinase K (w.r.t to mentioned sample volume) from the Extraction Kit. The RNA extraction step was performed using the Kingfisher Flex (version 1.01, Thermo Fisher Scientific) according to the manufacturer's protocol. The obtained RNA was stored at -80 until further use. The COVID-19 RT-PCR Detection Kit (Fosun 2019-nCoV qPCR, Shanghai Fosun Long March Medical Science Co. Ltd.) was used. The primers provided along with the kit amplify the Envelope gene (E; ROX labeled), Nucleocapsid gene (N-JOE labeled), and Open Reading Frame1ab (ORF1ab, FAM-labeled) specific to SARS-CoV-2. SARS-CoV-2 cDNA showing a Ct ~ 28 was used as an assay control. A linear regression equation, obtained from the known log viral particle count and Ct using RT-qPCR, was established for N- and ORF1ab (via the COVID-19 RT-PCR Detection Kit) genes specific to SARS-CoV-2[37].

5.5. Molecular Docking

We used the resolved crystal structures of M^{pro} (PDB ID: 6M0K)[30] protein of SARS-CoV-2 for carrying out molecular docking study through the CDOCKER protocol of Discovery Studio[41]. The structure of theaflavin and theaflavin 3-gallate were downloaded in SDF format from PubChem (CID: 169167)[42]. The Gaussian protocol with DFT was used for ligand geometry and energy minimization[43]. The binding site for M^{pro} was defined by taking the reference of the co-crystallized inhibitor (11a). The binding site with the best docking score was reported. The sphere coordinates for the M^{pro} binding pocket were 11.62, 11.88, and 68.70 with a radius of 12.00 Å. The other parameters were kept as default. Of all the binding conformations predicted, the top five binding poses with the highest CDOCKER energy were reported.

5.6. Molecular dynamics (MD) simulations and thermodynamic free energy calculations

The best binding poses of theaflavin 3-gallate with M^{pro} were subjected to MD simulations by GROMACS package[44–46]. We used the CHARMM36 force field to generate ligand topologies by CGenFF server[47,48]. Similarly, the protein topologies were prepared by CHARMM36 force field by employing the "gmx pdb2gmx" script of GROMACS. The protein-ligand complexes for simulations were prepared by appending ligand topologies to the protein topologies. To maintain the electroneutrality of the system,

sodium, and chloride ions were added by employing the “gmx genion” script. Further steps of MD simulations were carried out by following the protocol defined in our previous studies[49,50,34]. The simulations were carried out for 100 ns. The simulation trajectories of both the complexes were used to determine the root mean square deviations (RMSD) of backbone C-α atoms, extracting binding poses at different time intervals, and calculating thermodynamic binding free energies. We utilized the Molecular Mechanics Poisson-Boltzmann Surface Area (MM-PBSA) method to calculate free energies of protein-ligand binding. MM-PBSA is a classical and validated method for modeling protein-ligand interactions[39].

5.7. Steered molecular dynamics (SMD) simulations

SMD allows to analyze the amount of external force required to unbind ligand from its binding site in protein[51]. One of the most extensively used GROMACS package was employed to perform SMD simulations. The end terminals of proteins were processed for preparing protein topologies, while exactly the same ligand topologies were used as used in conventional MD simulations. Subsequently, protein-ligand complexes were solvated in a rectangular box of dimensions (8.5 X 8.3X 25) Å. Steepest descent method was employed for energy minimization of protein-ligand complexes. Further, NPT run was carried out at 100 ps to equilibrate the energy minimized complexes. For the execution of the pulling code a spring constant of 250 kJ/mol/nm^2 and a constant velocity of 0.01 nm/ps was maintained. For the execution of the pulling code a spring constant of 250 kJ/mol/nm^2 and a constant velocity of 0.01 nm/ps was maintained. During pulling simulations, the following equation was utilized to calculate the external force:

$F = -k [X_{pull} (t) – X_{pull} (0) -vt]$

where F = external pulling force; k = spring constant; v = constant velocity; X_{pull} (t) = position of atom at time t and X_{pull} (0) = initial position.

Declarations

Acknowledgement

Authors thank the Council of Scientific and Industrial Research (CSIR) for funding the research through “Biotechnological interventions for sustainable bio-economy generation through characterization, conservation, prospection, and utilization of Himalayan bioresources (MLP-201)” and the Department of Biotechnology through “Testing for diagnosis and molecular and digital Surveillance of COVID-19 (OLP-0043)”. MC, VKB, and AK are grateful to CSIR for support through Junior Research Fellowship. BKK duly acknowledges CSIR, GOI for funding his research through in house grant. AK is highly thankful to the Science and Engineering Research Board, Department of Science & Technology, Govt. of India for the Ramanujan Fellowship (File no. SB/S2/RJN-022/2017; CSIR-IHBT/GAP-0266). The manuscript represents IHBT publication no. xxx.

Conflict of interest

The authors declare no conflict of interest

References

1. Weiss, S. R. Forty years with coronaviruses. *J. Exp. Med.* **217**, (2020).
2. Mamedov, T. *et al.* Plant-produced glycosylated and *in vivo* deglycosylated receptor binding domain proteins of SARS-CoV-2 induce potent neutralizing responses in mice. *Viruses* **13**, 1595 (2021).
3. Zhou, P. *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273 (2020).
4. Naqvi, A. A. T. *et al.* Insights into SARS-CoV-2 genome, structure, evolution, pathogenesis and therapies: Structural genomics approach. *Biochim. Biophys. Acta. Mol. Basis Dis.* **1866**, 165878 (2020).
5. Xu, X. *et al.* Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission. *Sci. China. Life Sci.* **63**, 457–460 (2020).
6. Lu, R. *et al.* Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* **395**, 565–574 (2020).
7. Jin, Z. *et al.* Structure of M^{pro} from SARS-CoV-2 and discovery of its inhibitors. *Nature* **582**, 289–293 (2020).
8. Favilli, A. *et al.* Effectiveness and safety of available treatments for COVID-19 during pregnancy: a critical review. *J. Matern.Fetal Neonatal Med.* 1–14 (2020).
9. Gordon, D. E. *et al.* A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **583**, 459–468 (2020).
10. Kurapati, K. R. V., Atluri, V. S., Samikkannu, T., Garcia, G. & Nair, M. P. N. Natural products as Anti-HIV agents and role in HIV-associated neurocognitive disorders (HAND): A brief overview. *Front. Microbiol.* **6**, 1444 (2016).
11. Wells, T. N. C. Natural products as starting points for future anti-malarial therapies: Going back to our roots? *Malar. J.* **10**, 1–12 (2011).
12. Mashour, N. H., Lin, G. I. & Frishman, W. H. Herbal medicine for the treatment of cardiovascular disease: clinical considerations. *Arch. Intern. Med.* **158**, 2225–2234 (1998).
13. Harvey, A. L. Natural products in drug discovery. *Drug Discov. Today* **13**, 894–901 (2008).
14. Ibrahim, M. A. A., Abdeljawaad, K. A. A., Abdelrahman, A. H. M. & Hegazy, M. E. F. Natural-like products as potential SARS-CoV-2 M^{pro} inhibitors: *in-silico* drug discovery. *J. Biomol. Struct. Dyn.* **39**, 1–13 (2021).
15. Ibrahim, M. A. A. *et al.* Rutin and flavone analogs as prospective SARS-CoV-2 main protease inhibitors: *In silico* drug discovery study. *J. Mol. Graph. Model.* **105**, 107904 (2021).
16. Ibrahim, M. A. A. *et al. In silico* drug discovery of major metabolites from spices as SARS-CoV-2 main protease inhibitors. *Comput. Biol. Med.* <bvertical-align:super;>126</bvertical-align:super;>, (2020).

17. Ibrahim, M. A. A. *et al. In Silico* mining of terpenes from red-sea invertebrates for SARS-CoV-2 Main Protease (M^{pro}) inhibitors. *Mol.* <bvertical-align:super;>26</bvertical-align:super;>, 2082 (2021).
18. Jang, M. *et al.* Tea polyphenols EGCG and theaflavin inhibit the activity of SARS-CoV-2 3CL-protease *in vitro*. *Evidence-Based Complement. Altern. Med.* <background-color:#66FF66;bvertical-align:super;>2020</background-color:#66FF66;bvertical-align:super;>, (2020).
19. Yang, C. S., Lambert, J. D. & Sang, S. Antioxidative and anti-carcinogenic activities of tea polyphenols. *Arch. Toxicol.* **83**, 11 (2009).
20. Ponmurugan, P., Kavitha, S., Suganya, M. & Gnanamangai, B. M. Tea polyphenols chemistry for pharmaceutical applications. *Tea - Chem. Pharmacol.* **1**, 1 (2019).
21. Nakai, M. *et al.* Inhibitory effects of oolong tea polyphenols on pancreatic lipase *in vitro*. *J. Agric. Food Chem.* **53**<bvertical-align:super;>,</bvertical-align:super;> 4593–4598 (2005).
22. Liu, J. *et al.* Epigallocatechin gallate from green tea effectively blocks infection of SARS-CoV-2 and new variants by inhibiting spike binding to ACE2 receptor. *Cell Biosci.* **11**, 1–15 (2021).
23. Bhardwaj, V. K. *et al.* Bioactive molecules of tea as potential inhibitors for RNA-Dependent RNA Polymerase of SARS-CoV-2. *Front. Med.* **8**, (2021).
24. Ghosh, R., Chakraborty, A., Biswas, A. & Chowdhuri, S. Evaluation of green tea polyphenols as novel corona virus (SARS CoV-2) main protease (M^{pro}) inhibitors–an *in silico* docking and molecular dynamics simulation study. *J. Biomol. Struct. Dyn.* **39**, 4362–4374 (2021).
25. Jesús Naveja, J., Rico-Hidalgo, M. P., Medina-Franco, J. L. & Minkiewicz, P. Analysis of a large food chemical database: chemical space, diversity, and complexity. *F1000Research* <bvertical-align:super;>7</bvertical-align:super;>, (2018).
26. Wishart, D. S. *et al.* HMDB: the Human Metabolome Database. *Nucleic Acids Res.* **35**, D521 (2007).
27. Koch, W. Theaflavins, Thearubigins, and Theasinensins. *Handb. Diet. Phytochem.* 975–1003 (2020).
28. Zhu, Z. *et al.* From SARS and MERS to COVID-19: a brief summary and comparison of severe acute respiratory infections caused by three highly pathogenic human coronaviruses. *Respir. Res.* **21**, (2020).
29. Coleman, C. M. & Frieman, M. B. Coronaviruses: important emerging human pathogens. *J. Virol.* **88**, 5209–5212 (2014).
30. Kar, P., Sharma, N. R., Singh, B., Sen, A. & Roy, A. Natural compounds from Clerodendrum spp. as possible therapeutic candidates against SARS-CoV-2: An *in silico* investigation. *J. Biomol. Struct. Dyn.* **39**, 4774–4785 (2020).
31. Xu, H. X. *et al.* Screening of traditional medicines for their inhibitory activity against HIV-1 protease. *Phyther. Res.* **10**, 207–210 (1996).
32. Kurokawa, M. *et al.* Antiviral traditional medicines against herpes simplex virus (HSV-1), poliovirus, and measles virus in vitro and their therapeutic efficacies for HSV-1 infection in mice. *Antiviral Res.* **22**, 175–188 (1993).

33. Du, J. *et al.* Antiviral flavonoids from the root bark of Morus alba L. *Phytochemistry* **62**, 1235–1238 (2003).
34. Sharma, J. *et al.* An *in-silico* evaluation of different bioactive molecules of tea for their inhibition potency against non structural protein-15 of SARS-CoV-2. *Food Chem.* **346**, 128933 (2021).
35. Pinzi, L. & Rastelli, G. Molecular Docking: Shifting Paradigms in Drug Discovery. *Int. J. Mol. Sci.* **20**, (2019).
36. Bhardwaj, V. K., Singh, R., Das, P. & Purohit, R. Evaluation of acridinedione analogs as potential SARS-CoV-2 main protease inhibitors and their comparison with repurposed anti-viral drugs. *Comput. Biol. Med.* **128**, 104117 (2021).
37. Dai, W. *et al.* Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. *Science* **368**, 1331–1335 (2020).
38. Wang, C. *et al.* Calculating protein-ligand binding affinities with MMPBSA: Method and error analysis. *J. Comput. Chem.* **37**, 2436–2446 (2016).
39. Kumari, R., Kumar, R. & Lynn, A. G-mmpbsa -A GROMACS tool for high-throughput MM-PBSA calculations. *J. Chem. Inf. Model.* **54**, 1951–1962 (2014).
40. Li, B. *et al.* Theaflavins inhibit glucose transport across Caco-2 cells through the downregulation of the Ca2+/AMP-activated protein kinase-mediated glucose transporter SGLT1. *J. Funct. Foods* **75**, 104273 (2020).
41. BIOVIA Discovery Studio - BIOVIA - Dassault Systèmes®.
42. Kim, S. *et al.* PubChem Substance and Compound databases. *Nucleic Acids Res.* **44**, D1202–D1213 (2016).
43. Wilson, P. B., & Grootveld, M. (Eds.). Computational Techniques for Analytical Chemistry and Bioanalysis. Vol. 20. *Royal Society of Chemistry*. (2020)
44. Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).
45. Van Der Spoel, D. *et al.* GROMACS: fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718 (2005).
46. Hess, B., Kutzner, C., Van Der Spoel, D. & Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **4**, 435–447 (2008).
47. Vanommeslaeghe, K. *et al.* CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **31**, 671–690 (2010).
48. Yu, W., He, X., Vanommeslaeghe, K. & MacKerell, A. D. Extension of the CHARMM General Force Field to sulfonyl-containing compounds and its utility in biomolecular simulations. *J. Comput. Chem.* **33**, 2451–2468 (2012).
49. Bhardwaj, V. K., Singh, R., Sharma, J., Das, P. & Purohit, R. Structural based study to identify new potential inhibitors for dual specificity tyrosine-phosphorylation- regulated kinase. *Comput. Methods Programs Biomed.* **194**, (2020).

50. Singh, R., Bhardwaj, V., Das, P. & Purohit, R. Natural analogues inhibiting selective cyclin-dependent kinase protein isoforms: a computational perspective. *J. Biomol. Struct. Dyn.* **38**, 5126–5135 (2020).
51. Boubeta, F. M. *et al.* Lessons learned about steered molecular dynamics simulations and free energy calculations. *Chem. Biol. Drug Des.* **93**, 1129–1138 (2019).

Figures

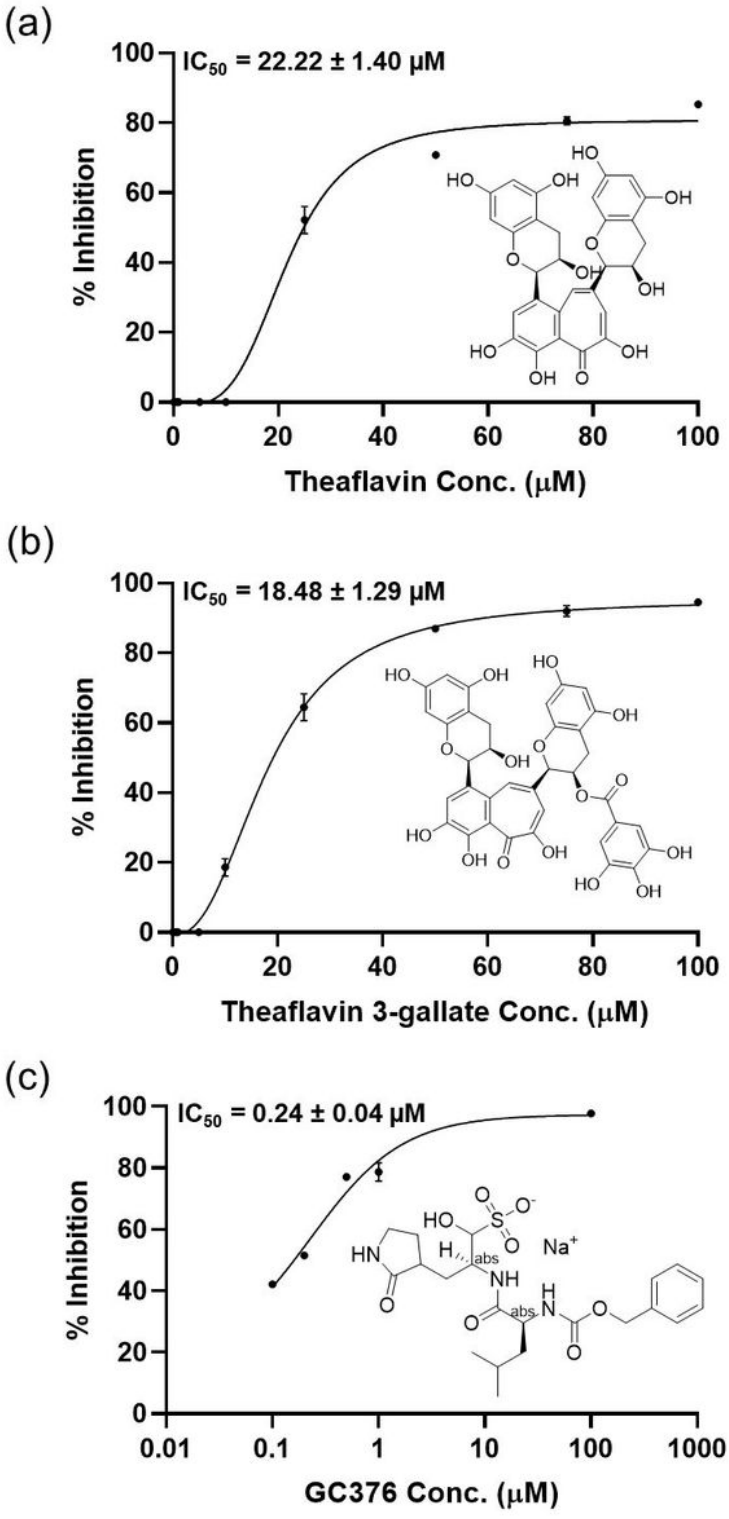

Figure 1

Inhibition of M^{pro} protein of SARS-CoV-2 by theaflavin 3-gallate. The inhibition of M^{pro} protein by a) theaflavin (positive control) b) theaflavin 3-gallate and c) GC376 (positive control) was measured in the presence of increasing concentrations of these molecules. The structures of molecules are shown along with the IC_{50} curves. Dose-response curves for IC_{50} values were determined by nonlinear regression. Data represent mean ± SE., n = 3 independent replicates.

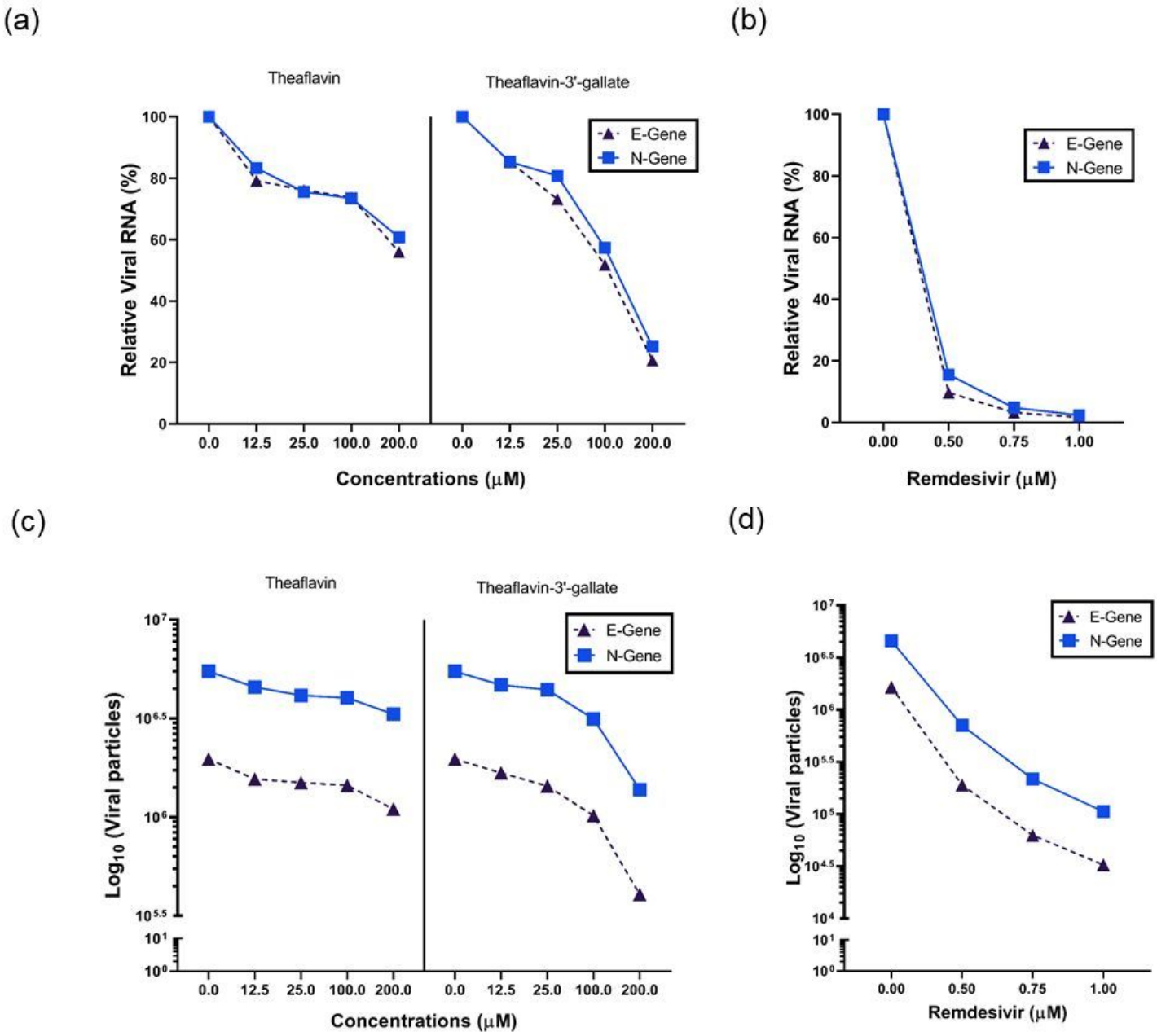

Figure 2

Effect of theaflavin 3-gallate on the inhibition of SARS-CoV-2. Response to theaflavin (positive control), theaflavin 3-gallate and remdesivir (positive control) in Vero cells at 50, 100, 150, and 200 μM was calculated using quantitative PCR of N and E viral genes. Graphs represent relative viral RNA % (a, b) and

log reduction in viral particles (c, d) after treatment with theaflavin, theaflavin 3-gallate and positive control remdesivir, respectively.

Figure 3

Analysis of docking results. The docking poses are shown in (a) 3D and (b) 2D representations for (i) GC373, (ii) theaflavin, and (iii) theaflavin 3-gallate.

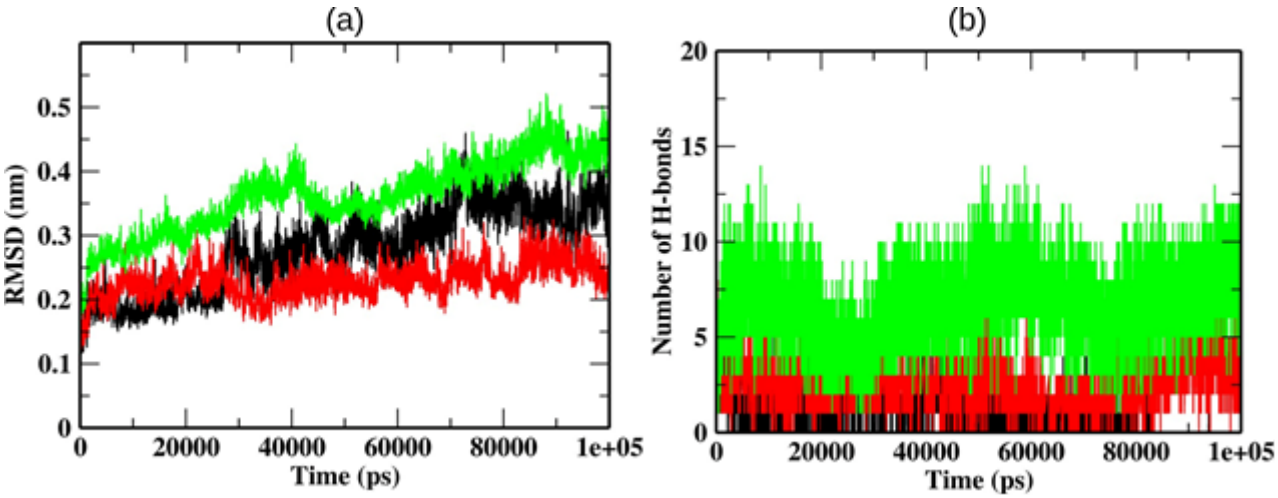

Figure 4

Analysis of MD trajectories showing the (a) RMSD of backbone Cα atoms for M^{pro} bound to different ligands and (b) number of H-bonds formed between M^{pro} and ligands during the entire simulation. The color-coding scheme is as follows: GC373 (black), theaflavin (red), and theaflavin 3-gallate (green).

Figure 5

Representation of the molecular interactions at the binding site of theaflavin 3-gallate-M^{pro} complex. The binding poses are shown for time intervals (a) 20 ns, (b) 40 ns, (c) 60 ns and (d) 80 ns.

Figure 6

Analysis of SMD results showing: (a) typical external force profiles of GC373 (black), theaflavin (red), and theaflavin 3-gallate (green). The position of (a) GC373, (b) theaflavin, and (c) theaflavin 3-gallate at different time intervals during SMD simulations. The color coding is as follows: 100 ns (blue), 194 ns (magenta), 242 ns (orange), and 400 ns (cyan).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryInformation.doc