Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Membrane-Mediated Sars-cov-2 Host Cell Entry: Potential Inhibitory Roles of Terpenoids in Silico

Gideon Ampoma Gyebi (gideonagyebi@gmail.com)
Bingham University https://orcid.org/0000-0002-1945-1739
Oludare Ogunyemi
Salem University
Ibrahim M. Ibrahim
Cairo University
Olalekan B. Ogunro
KolaDaisi University
Adegbenro P. Adegunloye
University of Ilorin
Saheed Afolabi
University of Ilorin

Research Article

Keywords: SARS-CoV-2, ACE2, TMPRSS2, spike protein, terpenoids, molecular docking

Posted Date: March 4th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-259624/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Abstract

Targeting viral cell entry proteins is an emerging therapeutic strategy for inhibiting the first stage of SARS-CoV-2 infection. In this study, 106 bioactive terpenoids from African medicinal plants were screened through molecular docking analysis against *human* angiotensin-converting enzyme 2 (*h*ACE2), *human* transmembrane protease serine 2 (TMPRSS2) and the S proteins of SARS-CoV-2, SARS-CoV and MERS-CoV. *In silico* ADMET and drug-likeness prediction, molecular dynamics simulation (MDS), binding free energy calculations and clustering analysis of MDS trajectories were performed on the top docked compounds to respective targets. The results revealed eight terpenoids with high binding tendencies to the catalytic residues of different targets. Pentacyclic terpenoids: 24-methylene cycloartenol and isoiguesterin interacted with the *h*ACE2 binding hotspots for the SARS-CoV-2 Spike protein. 11-hydroxy-2 -(3,4-dihydroxybenzoyloxy) abieta -5,7,9 (11),13-tetraene-12-one, 11-hydroxy-2 -(4-hydroxybenzoyloxy)-abieta- 5,7,9(11),13-tetraene-12-one and other abietane diterpenes interacted strongly with the S1-specificy pocket of TMPRSS2. 3-benzoylhosloppone and cucurbitacin interacted with the RBD and S2 subunit of SARS-CoV-2 spike protein respectively. The predicted druggable and ADMET favourable terpenoids formed structurally stable complexes in the simulated dynamics environment. These terpenoids provides core structure that can be exploited for further lead optimization to design drugs against SARS-CoV-2 cell mediated entry, subject to further *in vitro* and *in vivo* studies.

Introduction

The coronavirus disease-19 (COVID-19) caused by the new severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has been declared a public health emergency by the World Health Organization (WHO) [1-4]. The death toll from this virus has by far surpassed that of 2003 severe acute respiratory syndrome-coronavirus (SARS-CoV) and the 2012 Middle East respiratory syndrome coronavirus (MERS-CoV) outbreaks combined [5, 6]. The SARS-CoV-2 earlier known as 2019 novel corona virus (2019-nCoV) is evolutionarily related (80% identity) to SARS-CoV [7]. It causes multiple organ failure, which may present as fever, cough, shortness of breath, dyspnea, pneumonia, severe acute respiratory syndrome, kidney failure, and even death [8, 9].

Cell entry of coronaviruses depends on a fine interplay between the viral membrane spike (S) proteins and the host cell membrane proteins more importantly are the angiotensin-converting enzyme 2 (ACE2) and serine protease transmembrane protease serine 2 (TMPRSS2). [10]. The S-protein comprises two subunits, S1 as the receptor-binding domain (RBD) and S2 subunit for the fusion of viral membrane and the host cellular membrane. The SARS-CoV-2 relies on the host ACE2, for entry and the TMPRSS2 for S-protein priming. Upon binding of the S-protein to host receptor through the receptor-binding domain (RBD) in the S1 subunit, the S2 subunit mediates fusion of the viral envelope with the host membranes[11]. Although, the overall sequence similarities between S-protein of SARS-COV-2 and SARS-CoV are approximately ~76%, affinity between S-RBD of SARS-COV-2 and ACE2 is found to be approximately ten times higher when compared with SARS-CoV RBD [11-13]. This molecular interaction is responsible for regulating both the cross-species and higher human-to-human transmissions of SARS-CoV-2 [14, 15].

Therefore, these protein effectors of viral attachment, membrane fusion and cell entry are known as emerging targets for development of entry inhibitors, antibodies, and vaccines [14].

The use of phytomedicines as cheap alternatives to combat viral diseases and other infections, forms an integral component of African cultural practices, and hence a prominent feature in Africa [16-20]. Terpenoids are a well known class of phytochemicals of tremendous pharmaceutical value over time because of their relevant broad-spectrum utility in medicine [21, 22]. Screening a database of phytochemicals from indigenous African medicinal plants may help identify terpenoids with therapeutic potentials against the novel COVID-19 pandemic. Therefore, this study explores computational based screening of terpenoids from indigenous African medicinal plants as potential inhibitors of the emerging protein targets responsible for coronavirus cell entry and subsequent infection.

Methods

2.1 Protein preparation

The crystal structures of proteins for the docking studies were retrieved from the Protein Databank (http://www.rcsb.org) with their various PDB identification codes [1r42: Angiotensin-Converting Enzyme 2 (ACE2) [23]; 2OQ5: type II transmembrane serine proteinases (TMPRSS2) [24]; 6vw1: 2019-nCoV chimeric receptor-binding domain complexed with its receptor human ACE2 (ACE2-RBD) [25] and coronoviruses spike protein (6VSB: SARS-CoV-2) [3]; (5X5B: SARS-CoV) [26] and (5x5c : MERS-CoV) [26]. All the crystal structures were prepared by removing existing ligands and water molecules while missing hydrogen atoms were added using Autodock version 4.2 program (Scripps Research Institute, La Jolla, CA). Subsequently, non-polar hydrogens were merged while polar hydrogen was added to each protein. The well-ordered scheme was repeated for each protein and thereafter saved into dockable pdbqt format for molecular docking.

2.2 Ligand preparation

One hundred and six (106) bioactive terpenoids from African medicinal plants were complied base on literature search. Structure Data Format (SDF) of the reference inhibitors (S1: MLN-4760; S2: Camostat and S3: Nelfinavir mesylates) and 106 bioactive terpenoids derived from African plants were retrieved from the PubChem database (www.pubchem.ncbi.nlm.nih.gov) and converted to mol2 chemical format using Open babel [27]. Other compounds that were not available on the database were drawn with Chemdraw version 19 and converted to mol2 chemical format. Polar hydrogen charges of the gasteiger-type were assigned and the nonpolar hydrogen molecules were merged with the carbons and the internal degrees of freedom and torsions were set to zero. The protein and ligand molecules were further converted to the dockable pdbqt format using Autodock tools.

2.3 Molecular docking

Molecular docking was performed to evaluate the binding energy and to provide initial coordinates and topology parameters for the MD simulations. Virtual screening of human enzymes and active regions of the coronaviruses spike protein and determination of binding affinities were carried out using AutoDock Vina [28] and the binding scores from vina analysis were further validated by BINDSURF [29]. Docking of bioactive terpenoids and reference compounds against human ACE2, human TMPRSS2 and : SARS-CoV-2 spike protein was performed by AutoDock Vina to locate alternate binding sites enclosing the whole macromolecules with a very extended grid (60 Å × 60 Å × 60Å) to reveal all the possible interaction sites applying exhaustiveness values of 8. Pdbqt form of each protein and terpenoid were uploaded into their respective columns of Autodock Vina and BINDSURF (the online tool) to run the interaction. The compounds were then ranked by their binding affinity scores. The molecular interactions between proteins and selected compounds with higher binding affinity to the proteins were viewed with Discovery Studio Visualizer version 16.

2.4 Molecular Dynamics Simulation

Molecular Dynamics simulations were carried out on selected compounds, to evaluate their binding interactions with. The protein–ligand for the top docked terpenoids to SARS-CoV-2 spike (S) protein, human angiotensin-converting enzyme 2 (ACE2) and transmembrane protease serine 2 (TMPRSS2) from AutoDock Vina docking step were used in Molecular Dynamics Simulation (MDS) using NAMD software [30]. Necessary files for MDS were generated using CHARMM-GUI webserver [31, 32]. For each complex, the system was minimized for 10000 steps then a production run for 100 ns was performed. Temperature was set to be 310 K and salt concentration was set to be the physiological concentration 0 .154 M NaCl. Afterwards, calculations of Backbone-Root Mean Square Deviation (RMSD), Per residue Root Mean Square Fluctuations (RMSF), Radius of Gyration (RoG), Surface Accessible Surface Area (SASA) were performed using VMD TK console scripts [33].

2.5 Binding Free Energy calculation and Clustering analysis

Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) [34] calculation was performed using AmberTools 20 [35] on the results from MDS. TTClust version 4.7.2 were used to cluster the trajectory automatically according to the elbow method and produce a representative structure for each cluster [36]. These representative conformations were analyzed using Protein Ligand Interaction Profiler (PLIP) to know the interacting amino acids and the types of interactions [37].

2.6 Drug-likeness and ADMET studies

The top terpenoids that demonstrated highest binding affinity for ACE2, TMPRSS2 and active regions of SARS-CoV-2 spike protein were subjected to several drug-likeness predictive descriptors which an orally bio-active drug should comply to criteria for drug likeness [38, 39]. The predicted Absorption, Distribution, Metabolism, Excretion and Toxicity (ADMET) study was analysed using the admetSAR webserver [40].

The SDF file and SMILES of the compounds were downloaded from PubChem database to calculate ADMET properties using default parameters.

Results

3.1. Molecular docking

Figure 1 provides a flow chart showing the stepwise screening of African derived terpenoids for potential inhibitors of membrane-mediated SARS-CoV-2 cell entry proteins.

The result from the docking analysis of the reference inhibitors and bioactive terpenoids with the human ACE2, TMPRSS2 and SARS-CoV-2 spike protein is shown in table S3 (supplementary material). The top 20 terpenoids with the highest binding affinity for the ACE2 were further analyzed for binding interactions with SARS-CoV-2 chimeric Receptor-Binding Domain complexed with its human receptor ACE2 (ACE2-RBD) and the S protein of SARS-CoV and MERS-CoV as shown in table 2.

The docking analysis revealed that the reference inhibitor (MLN-4760) to the human ACE2 protein had binding energy of **-7.7** Kcal/mol respectively, while Camostat an inhibitor of TMPRSS2 had a binding energy of -7.6 Kcal/mol as represented in figure 3. It was further observed that the topmost docked terpenoids to the ACE2 had higher binding affinity for the S protein of SARS-CoV and MERS-CoV than SARS-CoV-2. More than 10 terpenoids had higher binding affinity than the 3 inhibitors used in this study table S1 (supplementary material). The top 20 docked compounds to SARS-CoV-2 S-proteins had higher binding affinity than nelfinavir mesylates S2 (supplementary material).

From the binding scores generated by the interacting terpenoids with the ACE2 and TMPRSS2 proteins, the 2 best docked terpenoids with the highest binding affinity are: 24-methylene cycloarteno and isoiguesterin with the corresponding binding energy of -9.7, and -9.5 Kcal/mol respectively. The 2 best docked terpenoids to SARS-CoV-2 S protein are 3-benzoylhosloppone and cucurbitacin with binding energies of -9.4 and -9.3 Kcal/mol respectively. 3-benzoylhosloppone had the highest binding affinity for SARS-CoV-2 S protein and the second top binding affinity to MERS-CoV S protein (Figure 3).

3.1.1 Amino acid interaction of selected terpenoids with target proteins.

The amino acid interactions of the human target proteins (ACE2 and TMPRSS2) with reference inhibitors and plant derived terpenoids that demonstrated the highest binding tendencies are represented in table 1. In the same way, the amino acid residues of the coronaviruses S protein that interacted with reference inhibitors and terpenoids with the highest binding affinity are shown in table 2. The interacting residues of ACE2 and TMPRSS2 with respective ligand groups were majorly through hydrophobic interactions and H-bond. Few H-bonding below 3.40 Å were observed with coronaviruses S protein (table 4).

The binding of MLN-4760 to ACE2 showed that it was docked into the N terminus and zinc-containing subdomain I of ACE2 (figure 4a). MLN-4760 exhibited several types of hydrophobic interactions (Pi-

Sigma, Pi-Pi T-Shaped, Pi-Alkyl and Alkyl) with TYR^{510} PHE^{504} MET^{360}, LYS^{363} and CYS^{344}, a Salt and attractive charges to ARG^{514}, ARG^{518} and ARG^{278} and hydrogen bond to TYR^{515}, THR^{371}, PRO^{346} and ARG^{273} (figure 4a).

24-methylene cycloartenol the best docked terpenoid was docked into the C terminus-containing subdomain II of ACE2 but interacted with different residue as with the case of N-acetyl-D-glucosamine (figures 4b). 24-methylene cycloartenol interacted via H-bond to TRP^{163}, SER^{170} and TYR^{497}.A Pi-Alkyl interaction was also observed with TYR^{613}, PRO^{492} and VAL^{491}. Isoiguesterin interacted via H-bond to ASP^{350}, TYR^{385} and ASN^{394}. A Pi-Alkyl and Alkyl interactions was observed with the ALA^{99}, PHE^{40}, PHE^{390} and LEU^{73}, TRP^{69} residues respectively in a similar binding pattern with MLN-4760 (figure 4c).

Camostat was docked into the S1-specificity pocket of TMPRSS2 (figure 5a). It interacted via conventional H-bond to five amino residues (ARG^{41}, SER^{195}, TRP^{215}, ALA^{190} and ASP^{189}) and via carbon hydrogen bond to GLN^{192} of TMPRSS2. The conventional H-bond was formed in the direction of the guanidine group in this order: first ester bond, second ester bond, while the last three residues interacted with amidino nitrogen of guanidine group respectively. The phenyl ring was responsible for the carbon-hydrogen bond with GLN^{192} (figure 5a). T3 and T4 were docked into S1-specificity pocket of TMPRSS2 in a similar binding pattern as in the case of camostat (figure 5b & 5c). The only difference observed between the binding pattern of T3 and T4 was an additional H-bond between T3 with ARG^{41} (figure 5b).

Nelfinavir mesylates an inhibitor of SARS-CoV and MERS-CoV S protein interacted in its best docked conformation to the S protein of SARS-CoV-2 in a different manner. Nelfinavir mesylates was docked into the S2 Subunit of SARS-CoV S protein (figure 7a). The same inhibitor was docked into to the N-terminal domain (NTD) region of the S1 subunit of SARS-CoV-2 and MERS-CoV S protein (figure 6a & 8a).

3-benzoylhosloppone with the highest binding affinity for SARS-CoV-2 S protein interacted via H-bond to THR^{547}; Alkyl interaction to PHE^{541} and Pi-Alkyl interaction to PRO^{589} and LEU^{546}. The region of interaction was between the CTD and SD1 region of S1 subunit of SARS-CoV-2 S protein. Cucurbitacin B was docked to the S2 subunit of SARS-CoV-2 S protein but interacted with different amino acid residue. The interaction of cucurbitacin B to the protein was via H-bond to ARG^{1091}, ASN^{914}, THR^{912} and GLN^{1113}; Pi-Sigma bond to PHE^{1121} and Alkyl interaction to ILE^{1114} and GLY^{1124} (figure 6c).

The same pattern of interaction was observed in both 7-Deacetoxy-7-oxogedunin and 3-friedelanone to the S2 subunit of SARS-CoV S protein. Both terpenoids interacted via a H-bond to ARG^{982} and GLY^{726} of the S2 subunit. While 7-deacetoxy-7-oxogedunin interacted with the upstream helix and central helix, 3-friedelanone interacted with the connecting region of the S2 subunit. A hydrophobic interaction via Pi-Alkyl and alkyl bonds was observed with the remaining amino acid residue (table 2; figure 7b & 7c)

7-Deacetoxy-7-oxogedunin interacted via H-bond to the SER^{51} residue of N-terminal domain of the S1 subunit of MERS-CoV S protein. A Pi-Pi T-shaped interaction was formed between 7-deacetoxy-7-

oxogedunin and PHE^{354}; HIS^{670} of MERS-CoV S protein. Other hydrophobic interactions via Pi-Alkyl and Pi-Sigma bonds were observed to with the remaining amino acid residues (table 4; figure 8a & 8b). 3-benzoylhosloppone interacted via: Pi-Sigma interaction to (PHE^{341}) of NTD; Pi-Pi Stacking to (MET^{698}) of SD2; Pi-Alkyl interaction to (LYS^{689}) of SD2; and an Alkyl interaction to (LEU^{344} and ILE^{337}) of NTD with the S1 subunit (figure 8c).

In summary, the binding of ligands to various proteins revealed eight terpenoid with remarkable binding affinities. Those with very good interactions with ACE2 and TMPRSS2 are 24-methylene cycloartenol; isoiguesterin; 11-Hydroxy-2 - (3,4-dihydroxybenzoyloxy) abieta-5,7,9(11),13-tetraene-12-one; and 11-Hydroxy-2 -(4-hydroxybenzoyloxy)-abieta-5,7,9(11),13-tetraene-12-one. Similarly, 3-benzoylhosloppone, and cucurbitacin B interacted well with SARS-CoV-2 spike protein, while 7-deacetoxy-7-oxogedunin, and 3-friedelanone interacted well with SARS-CoV and MERS-CoV spike protein.

3.1.2 Energy profile of best docked terpenoids to respective proteins

The overall energy profiles of terpenoid-protein complexes in the selected clusters with the best docked poses are shown in figures 9-11. Figure 9a-11a shows the breakdown of the binding energy of the selected cluster into different contributions. Gauss 1 (blue) and 2 (leaf green) bars: represent the non-bonding interactions, red bar: repulsion, light blue bar: Hydrophobic, Purple bar: Hydrogen bonds, light green bar: rotational forces, while the black bar represents total binding affinity which is a representative contribution of all bonding and non-bonding interactions between the terpenoids and the protein residues. The contributions of the various type of interaction as presented in graph (figure 9a-11a) shows that of the total binding energy of **-9.7** Kcal/mol exhibited by the binding of 24-methylene cycloartenol to the ACE2, -2.1 and 1.8 Kcal/mol of hydrophobic and H-bond energies respectively was contributed, while the rest was contributed by non-bonding interaction mainly van der Waals, repulsive and rotational forces. A H-bond, Hydrophobic interaction and repulsive energy of -2.8 -0.8 and +2.3 Kcal/mol respectively was contributed to the total binding energy of -**10.0** Kcal/mol between T3 and TMPRSS2. A hydrophobic interaction energies of -2.1, -0.6 and -1.5, a H-bond energies of 0.3, -0.6, -0.3 Kcal/mol was contributed to the total binding energy of the spike protein of SARS-CoV-2, SARS-CoV and MERS-CoV with respective terpenoids. The rest of the energy was contributed by non-binding interactions.

Figures 9b-11b shows the overall energy profile of the ligand-receptor complex of the selected cluster, showing the individual energetic contributions for each atom in the ligand. The colour indication is similar to (a) above.

3.2 Molecular Dynamics Simulation

Four compounds which are, camostat, 11-hydroxy-2 - (3,4-dihydroxybenzoyloxy) abieta-5,7,9(11),13-tetraene-12-one, 24-methylene cycloartenol, and 3-benzoylhosloppone, were analysed for their interactions with SARS-CoV-2 Spike glycoprotein (S protein), Angiotensin-converting enzyme 2 (ACE2), and Transmembrane protease serine 2 (TMPRSS2) proteins. Molecular dynamics simulation was done on each of the protein-ligand complexes and the trajectories were analyzed. The Radius of Gyraion (RoG),

Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF), and Surface Accessible Surface Area (SASA) results were calculated for each trajectory and are shown in figure 11. There was no observed difference between the RoG of TMPRSS2_camostat and TMPRSS2_(11-hydroxy-2 - (3,4-dihydroxybenzoyloxy) abieta-5,7,9(11),13-tetraene-12-one) complexes, while that of ACE2 is larger and S protein has the largest values because it is a monomer of the SARS-CoV-2 spike trimer. All of them are fluctuating about certain values. The RMSD values of TMPRSS2_(T3), ACE2-24_methylene cycloartenol), TMPRSS2_camostat, and S protein_(3-benzoylhosloppone) complexes are around 2.13 Å, 3.6 Å, 2.14 Å, and 16.78 Å, respectively. While the RMSF values for TMPRSS2_(11-hydroxy-2 - (3,4-dihydroxybenzoyloxy) abieta-5,7,9(11),13-tetraene-12-one), ACE2_24-methylenecycloartenol), TMPRSS2_camostat, and S protein_(3- benzoylhosloppone) complex are fluctuating around 0.68 Å, 1.29 Å, 0.73 Å, and 7.36 Å, respectively.

TMPRSS2_(11-hydroxy-2 - (3,4-dihydroxybenzoyloxy) abieta-5,7,9(11),13-tetraene-12-one), ACE2_(24-methylene cycloartenol), and TMPRSS2_camostat complexes have a spike in the end of their RMSF results indicating the motion of the terminals. The spikes in the middle and the start of the RMSF of ACE2_(24-methylene cycloartenol) complex between amino acid 265 and amino acid 443 and spikes in S protein_(3-benzoylhosloppone) complex corresponds to the loops in the two protein respectively (Figure 12). The values of SASA can be found to be nearly stable for each complex but differ from each other.

Molecular Mechanics/Generalized Born Surface Area (MMGBSA) algorithm in AmberTools 20 was utilized to calculate the ligand binding free energy. All frames (~1000 frame) were used in this calculation for each protein-drug complex. Figure 14 shows the binding affinity in Kcal/mol from MMGBSA analysis with Standard Deviation as error bars for each protein-drug complex. The best binding affinity (more negative) is for TMPRSS2_camostat (-53.5059 Kcal/mol) which indicates the strong binding between them.

Table 4 shows the number of clusters, representative frame produced for each trajectory, and the interaction types using PLIP webserver. Hydrophobic, H-bond, salt-bridges, pi-cation and pi-stacking are the types of interactions found by PLIP webserver. Most of complexes have H-bond and hydrophobic interactions, with TMPRSS2_camostat having the largest number of bonds in each cluster compared to other complexes. Figure 15 shows the protein-drug cluster representatives for the protein-ligand complexes and the mode of interaction in the enlarged part of the image. Images were generated using PyMol software V 2.2.2.

3.3 Drug likeness and Pharmacokinetic properties of selected terpenoids.

The result generated from the Lipinski and ADMET filtering analyses are represented in table 4 and figure S1 (supplementary file).

Four terpenoids T1, T3, T5 and T6 fulfilled the requirement for Lipinski analysis of the rule of-five with corresponding favourble predicted ADMET parameters.

The *in silico* predictive pharmacokinetic and ADMET properties from the filtering analyses suggested T1, T3, T5 and T6 with a high probability of absorption, subcellular distribution, low toxicity. Though pharmacokinetic analysis indicated T1 (Table 4) to be less soluble while the ADME/tox analysis indicated high aqueous solubility, ability to pass the high human intestinal absorption, low acute oral toxicity with a good bioavailability score as exhibited by T3, T5 and T6 (Table 4).

Discussion

Interference with several proteins that mediate viral attachment, membrane fusion, and cell entry of coronaviruses is an emerging therapeutic strategy for preventing COVID-19 infection [10, 41]. This principle was earlier demonstrated with HIV [42, 43] and SARS-CoV [44]. Earlier screening and prospecting of therapeutic phytocompound have been reported for both SARS-CoV and MERS-CoV [45-48]. Cell-based assays have shown the antiviral potentials of specific plant terpenoids against Severe Acute Respiratory Syndrome Coronavirus (SARS- CoV) [45, 49]. This study was therefore undertaken to identify potential inhibitors of membrane-mediated SARS-CoV-2 entry proteins form the class of the plant derived terpenoids. Specifically, two triterpenes namely 24-methylene cycloartenol and isoiguesterin were reported to target ACE2 as well as the host-virus interface (S-protein-ACE2 Receptor Complex). These compounds interacted with adjacent residues in the conserved domain, apparently portraying its ability to bind and block interactions of hotspot 31 residues. The residues near lysine 31, and tyrosine 41, 82–84, and 353–357 in human ACE2 are important for the binding of S-protein of coronavirus [50]. The hotspots: 31 and 353, makes salt bridge between Lys31 and Glu35, and the hotspot 353, comprising a salt bridge between Lys353 and Asp38, and are both buried in hydrophobic environment, therefore interaction within this region is suggested to affect the binding of its substrate [51]. Abietane diterpenes, namely 11-hydroxy-2 -(3,4-dihydroxybenzoyloxy) abieta-5,7,9(11),13-tetraene-12-one (T3), and 11-hydroxy-2-(4-hydroxybenzoyloxy)-abieta- 5,7,9(11),13-tetraene-12-one (T4) showed the strongest interaction with with TMPRSS2. In a similar binding pattern to camostat, these compounds were fitted into the S1-specificity pocket. They interacted with residue ALA^{190}, ASP^{189} and GLN^{192} that are known to be part of the amino acid found at the basement of the pocket. ASP^{189} at the bottom of the pocket is known to determine the specificity of the S1 pocket for basic residues Arg and Lys at position P1 of the substrate [24]. The result showed that the hydroxybenzoyloxyl moiety of the terpenoids (T3 and T4) was responsible for atleast 75% of the H-Bond with the protein. It was further observed that just as in the case of benzamidine (the native ligand) and camostat, the hydroxybenzoyloxyl moiety of the two terpenoids points with its hydroxyl group towards the carboxylate group of ASP^{189} forming strong H-bonds with ASP^{189} and other residue in the pocket. For camostat, the phenylquanidine moiety pointed into the hydrophobic pocket with the negatively charged ASP^{189} at its bottom. Unlike the H-bond formed between the amidino nitrogen of the phenylquanidine and benzamidine, in T3 and T4 the H-Bonds were formed mainly with the hydroxyl, and carboxylate group. A striking similarity observed was that the ester bond that linked both the phenylquanidine moiety of camostat and the hydroxybenzoyloxyl moiety of T3 and T4 to the remaining structural unit of the compounds formed strong H-Bonds to the same residue SER^{195}. The phenyl group of the hydroxybenzoyloxy moiety of T3 and T4 further interacted with hydrophobic interactions to CYS^{119}

and CYS^{219} just as the peptide planes of the bonds between Trp215–Gly216 and Cys191–Gln192 sandwich the phenyl ring of benzamidine [24, 52] The additional hydrophobic interaction by T3 and T4 may have been responsible for the exhibited higher binding affinities than camostat and benzamidine. Furthermore, while the hydroxybenzoyloxy moiety was directed toward the hydrophobic cleft created by ASP^{189}, the abietane agylcon interacted with the imidazol ring of HIS^{57} of the S2 pocket that is found next to the S1 pocket and ARG^{41} (in the case of T4) which are outside the hydrophobic cleft. A similar interaction as the later was observed with camostat. The strong similarity in the binding pattern and even a far strong binding affinity than the camostat and benzamidine indicates that T3, T4 and other abietane diterpenes especially those with hydroxybenzoyloxyl moiety attached to the abietane aglycon are potential inhibitors of TMPRSS2 thus preventing some coronaviruses from entering host [24]. It is known that, like SARS-CoV, SARS-CoV-2 S protein recognizes and binds to host-cell receptor angiotensin-converting enzyme 2 (ACE2) using a transmembrane protease serine 2 (TMPRSS2) which activates the S protein to facilitate viral fusion and entry into cells [9]. It is important to note that serine protease inhibitors like camostat mesylate, which blocks the activity of TMPRSS2 [53], has been approved in Japan for human use. Related compounds with antiviral activity potentiates as an anti-SARS-CoV-2 agent [54]. Also some abietane terpenoids have been identified to exhibit *in vitro* anti-SARS-CoV activity [45]. This is corroborates the result of our study that shows that abietane diterpenes exhibits a wide spectrum and multiplicity of protein binding; and may thereby specifically execute a complete blockage of viral entry.

With regards to coronavirus S-proteins two compounds, 3- Benzoylhosloppone and Cucurbitacin B, were of utmost interest. While 3-benzoylhosloppone interacted with amino acid residue of the RBD and SD1 region of the S1 subunit, Cucurbitacin B was docked into the S2 subunit of SARS-CoV-2 S protein. The former subunit is responsible for receptor recognition while the later mediates the fusion of viral membrane and the host cellular membrane [55]. These terpenoids may prevent interaction of spike protein with its host cell receptor, thereby preventing entry of the virus into the host cell. 3-benzoylhosloppone has been reported for its antimalarial property while Cucurbitacin B is an anticancer agent [56, 57]. The MDS analysis of the top docked with their complexed proteins were stable and could be therefore subjected to experimental processes in further studies. From the Lipinski, pharmacokinetic and ADMET filtering analyses, we identified four druggable and non-toxic, natural terpenoids that exhibited strong binding tendency to the various protein targets that mediates coronavirus-host cell entry. The result from the predicted filtering analyses of the four compounds showed parameters that suggest a favourable *in silico* ADMET and pharmacokinetic properties The terpenoids expressed high probability of human intestinal absorption. They were also non-substrate to the permeability-glycoprotein (P-gp) [58], expressed capability to cross the blood brain barrier (BBB). SARS-CoV-2 has been reported to infect the brain, thus indicating its ability to cross the blood brain barrier (BBB) [59], compounds that can cross the BBB will be beneficail in the overal all viral clearance, The four terpenoids did not show indication of mutagenicity *in silico*, thereby they may not cause genetic mutations. The compounds did not display inhibitory potential for the various cytochrome P450, thus may not adversely affect phase I drug metabolism in the liver. These terpenoids are therefore considered as potential drug candidates.

Conclusion

A virtual screening approach was successfully applied to identify plant derived terpenoids as potential inhibitor of coronavirus cells entry proteins. The pentacyclic terpenoids, 4-methylene cycloartenol and isoiguesterin interacted strongly with binding sites residues that are known to interfere with the activity of ACE2. Compounds: 11-Hydroxy-2 - (3,4-dihydroxybenzoyloxy) abieta -5,7,9 (11),13-tetraene-12-one (T3), and 11-Hydroxy-2 -(4-hydroxybenzoyloxy)-abieta- 5,7,9(11),13-tetraene-12-one (T4) and other abietane diterpenes exhibited a similar binding pattern to the S1-specificity pocket of TMPRSS2 as camostat (reference inhibitor). They also showed wide spectrum and multiplicity of entry protein binding. The terpenoids binding conformations in the complexes were stable in a simulated dynamic environment. Since the identified lead compounds showed drug-likeness and low toxicity as indicated by the *in silico* pharmacokinetically relevant molecular descriptors, they are postulated as potential inhibitors that can be considered for further *in vitro* and *in vivo* studies towards developing entry inhibitors against the ongoing coronavirus pandemic.

Declarations

Acknowledgments

The authors appreciate the members of the ***BioNet-AP***: Bioinformatics Network on African Phytomedicine for COVID-19 research.

-**Ethical** Approval: Not applicable
-**Consent to Participate:** Not applicable

-**Consent to Publish:** Not applicable
-**Authors Contributions**

1. A. Gyebi Conceived and designed the analysis
2. M. Ogunyemi Performed molecular docking analysis
3. M. Ibrahim Performed molecular simulations
4. B. Ogunro Wrote manuscript
5. P. Adegunloye Interprets results and wrote manuscript
6. O. Afolabi Editing and review of manuscript

-Funding: The research did not receive any funding or grants

-Conflicts of interest: The authors declare no conflicting interest

-Availability of data and materials

The authors confirm that the data supporting the findings of this study are available within the article [and/or] its supplementary materials.

References

1. Taniguchi, Y., Nishikawa, H., Maeda, N., & Terada, Y. (2020). Breathlessness, pleural effusions, fibromas, and Meigs syndrome: look beyond the chest and don't delay! *The Lancet, 395*, e32.
2. Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y., Qiu, Y., Wang, J., Liu, Y., Wei, Y., Xia, J., Yu, T., Zhang, X., & Zhang, L. (2020). Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet, 395*, 507–513.
3. Wrapp, D., Nianshuang, W., Corbett, K., Goldsmith, J., Hsieh, C.-L., Abiona, O., Graham, B., & McLellan, J. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation, Science (New York, N.Y.), 367 (2020) eabb2507.
4. WHO, Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19), Geneva, (2020).
5. Hui, D. S., E, I. A., Madani, T. A., Ntoumi, F., Kock, R., Dar, O., Ippolito, G., McHugh, T. D., Memish, Z. A., Drosten, C., Zumla, A., Petersen, E., The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health - The latest 2019 novel coronavirus outbreak in Wuhan, China, International journal of infectious diseases: IJID : official publication of the International Society for Infectious Diseases, 91 (2020) 264–266.
6. Srinivasan, S., Cui, H., Gao, Z., Liu, M., Lu, S., Mkandawire, W., Narykov, O., Sun, M., Korkin, D., Structural Genomics of SARS-CoV-2 Indicates Evolutionary Conserved Functional Regions of Viral Proteins, Viruses, 12 (2020).
7. Chen, Y. W., Yiu, C.-P., & Wong, K.-Y. (2020). Prediction of the SARS-CoV-2 (2019-nCoV) 3C-like protease (3CLpro) structure: virtual screening reveals velpatasvir, ledipasvir, and other drug repurposing candidates. *F1000Research, 9*, 129.
8. Liu, C., Zhou, Q., Li, Y., Garner, L. V., Watkins, S. P., Carter, L. J., Smoot, J., Gregg, A. C., Daniels, A. D., Jervey, S., Albaiu, D., Research and Development on Therapeutic Agents and Vaccines for COVID-19 and Related Human Coronavirus Diseases, ACS Central Science, (2020).
9. 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. Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods, Acta Pharmaceutica Sinica B, (2020).
10. Chen, H., & Du, Q. Potential natural compounds for preventing SARS-CoV-2 (2019-nCoV) infection, 2020.
11. Coutard, B., Valle, C., de Lamballerie, X., Canard, B., Seidah, N. G., & Decroly, E. (2020). The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade. *Antiviral research, 176*, 104742.

12. Chan, J. F., & Kok, K. H. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan, 9 (2020) 221–236.
13. Choudhary, S., Malik, Y., Tomar, S., Identification of SARS-CoV-2 Cell Entry Inhibitors by Drug Repurposing Using in Silico Structure-Based Virtual Screening Approach, 2020.
14. Zhang, H., Penninger, J. M., Li, Y., Zhong, N., & Slutsky, A. S. Angiotensin-converting enzyme 2 (ACE2) as a SARS-CoV-2 receptor: molecular mechanisms and potential therapeutic target, Intensive Care Medicine, (2020).
15. Wan, Y., Shang, J., Graham, R., Baric, R. S., & Li, F. (2020). Receptor Recognition by the Novel Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS Coronavirus. *Journal of virology, 94*, e00127–e00120.
16. Beuscher, N., Bodinet, C., Neumann-Haefelin, D., Marston, A., & Hostettmann, K. (1994). Antiviral activity of African medicinal plants. *Journal of Ethnopharmacology, 42*, 101–109.
17. Ojo, O., Oluyege, J., & Famurewa, O. Antiviral properties of two Nigerian plants, Afr. J. Plant Sci, 3 (2009).
18. Tung, N., Kwon, H.-J., Kim, J.-H., Ra, J., Ding, Y., Kim, J., & Kim, Y. H. (2010). Antiinfluenza Diarylheptanoids from the Bark of Alnus japonica. *Bioorganic & medicinal chemistry letters, 20*, 1000–1003.
19. Bagla, V., McGaw, L., & Eloff, J. (2012). The antiviral activity of six South African plants traditionally used against infections in ethnoveterinary medicine. *Veterinary microbiology, 155*, 198–206.
20. Ndhlala, A., Amoo, S., Ncube, B., Moyo, M., Nair, J., & van Staden, J. Antibacterial, Antifungal, and Antiviral Activities of African Medicinal Plants, 2013, pp. 621–651.
21. Jaeger, R., & Cuny, E. (2016). Terpenoids with Special Pharmacological Significance: A Review. *Natural Product Communications, 11*, 1934578X1601100946.
22. Gyebi, G. A., Ogunro, O. B., Adegunloye, A. P., Ogunyemi, O. M., & Afolabi, S. O. Potential inhibitors of coronavirus 3-chymotrypsin-like protease (3CLpro): An in silico screening of alkaloids and terpenoids from African medicinal plants, Journal of Biomolecular Structure and Dynamics, (2020) 1–19.
23. Towler, P., Staker, B., Prasad, S., Menon, S., Tang, J., Parsons, T., Ryan, D., Fisher, M., Williams, D., Dales, N., Patane, M., & Pantoliano, M. (2004). ACE2 X-Ray Structures Reveal a Large Hinge-bending Motion Important for Inhibitor Binding and Catalysis. *The Journal of biological chemistry, 279*, 17996–18007.
24. Kyrieleis, O. J., Huber, R., Ong, E., Oehler, R., Hunter, M., Madison, E. L., & Jacob, U. (2007). Crystal structure of the catalytic domain of DESC1, a new member of the type II transmembrane serine proteinase family. *The FEBS journal, 274*, 2148–2160.
25. Shang, J., Ye, G., Shi, K., Wan, Y., Luo, C., Aihara, H., Geng, Q., Auerbach, A., & Li, F. Structural basis for receptor recognition by the novel coronavirus from Wuhan, DOI: 10.21203/rs. 2.24749/v1, (2020).
26. Yuan, Y., Cao, D., Zhang, Y., Ma, J., Qi, J., Wang, Q., Lu, G., Wu, Y., Yan, J., Shi, Y., Zhang, X., & Gao, G. F. (2017). Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains. *Nature Communications, 8*, 15092.

27. O'Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., & Hutchison, G. R. (2011). Open Babel: An open chemical toolbox. *Journal of cheminformatics, 3*, 33.
28. Trott, O., & Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem, 31*, 455–461.
29. Sanchez-Linares, I., Perez-Sanchez, H., Cecilia, J. M., & Garcia, J. M. High-Throughput parallel blind Virtual Screening using BINDSURF, BMC bioinformatics, 13 Suppl 14 (2012) S13.
30. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., & Kale, L. K.J.J.o.c.c. Schulten, Scalable molecular dynamics with NAMD, 26 (2005) 1781–1802.
31. Brooks, B. R., Brooks, C. L. III, Mackerell Jr, A. D., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., & Bartels, C. S.J.J.o.c.c. Boresch, CHARMM: the biomolecular simulation program, 30 (2009) 1545–1614.
32. Lee, J., Cheng, X., Swails, J. M., Yeom, M. S., Eastman, P. K., Lemkul, J. A., Wei, S., Buckner, J., & Jeong, J. C. Y.J.J.o.c.t. Qi, computation, CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field, 12 (2016) 405–413.
33. Humphrey, W., & Dalke, A. (1996). K.J.J.o.m.g. Schulten. *VMD: visual molecular dynamics, 14*, 33–38.
34. Miller, B. R. III, McGee, T. D., Jr, J. M., Swails, N., Homeyer, H., & Gohlke, A.E.J.J.o.c.t. Roitberg, computation, MMPBSA. py: an efficient program for end-state free energy calculations, 8 (2012) 3314–3321.
35. Case, D., Belfon, K., Ben-Shalom, I., Brozell, S., Cerutti, D., Cheatham, T. III, Cruzeiro, V., Darden, T., Duke, R., Giambasu, G., AMBER 2020, (2020).
36. Tubiana, T., Carvaillo, J.-C., & Boulard, Y. S.p.J.J.o.c.i. Bressanelli, modeling, TTClust: a versatile molecular simulation trajectory clustering program with graphical summaries, 58 (2018) 2178–2182.
37. Salentin, S., Schreiber, S., Haupt, V. J., & Adasme, M. F. M.J.N.a.r. Schroeder, PLIP: fully automated protein–ligand interaction profiler, 43 (2015) W443-W447.
38. Lipinski, C. A. (2000). Drug-like properties and the causes of poor solubility and poor permeability. *Journal of pharmacological and toxicological methods, 44*, 235–249.
39. Nickel, J., Gohlke, B. O., Erehman, J., Banerjee, P., Rong, W. W., Goede, A., Dunkel, M., & Preissner, R. (2014). SuperPred: update on drug classification and target prediction. *Nucleic acids research, 42*, W26–W31.
40. 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*, 3099–3105.
41. Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., Schiergens, T., Herrler, G., Wu, N.-H., Nitsche, A., Müller, M., Drosten, C., Pöhlmann, S., SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor, Cell, (2020).

42. Derdeyn, C. A., Decker, J. M., Sfakianos, J. N., Wu, X., O'Brien, W. A., Ratner, L., Kappes, J. C., Shaw, G. M., & Hunter, E. (2000). Sensitivity of human immunodeficiency virus type 1 to the fusion inhibitor T-20 is modulated by coreceptor specificity defined by the V3 loop of gp120. *Journal of virology, 74*, 8358–8367.
43. Hartt, J. K., Liang, T., Sahagun-Ruiz, A., Wang, J. M., Gao, J. L., & Murphy, P. M. The HIV-1 cell entry inhibitor T-20 potently chemoattracts neutrophils by specifically activating the N-formylpeptide receptor, Biochemical and biophysical research communications, 272 (2000) 699–704.
44. Adedeji, A., Severson, W., Jonsson, C., Singh, K., Weiss, S., & Sarafianos, S. Novel Inhibitors of SARS-CoV Entry acting by Three Distinct Mechanisms, Journal of virology, 87 (2013).
45. Wen, C. C., Kuo, Y. H., Jan, J. T., Liang, P. H., Wang, S. Y., Liu, H. G., Lee, C. K., Chang, S. T., Kuo, C. J., Lee, S. S., Hou, C. C., Hsiao, P. W., Chien, S. C., Shyur, L. F., & Yang, N. S. (2007). Specific plant terpenoids and lignoids possess potent antiviral activities against severe acute respiratory syndrome coronavirus. *Journal of medicinal chemistry, 50*, 4087–4095.
46. Ryu, Y. B., Park, S. J., Kim, Y. M., Lee, J. Y., Seo, W. D., Chang, J. S., Park, K. H., Rho, M. C., & Lee, W. S. (2010). SARS-CoV 3CLpro inhibitory effects of quinone-methide triterpenes from Tripterygium regelii. *Bioorg Med Chem Lett, 20*, 1873–1876.
47. Park, J. Y., Kim, J. H., Kim, Y. M., Jeong, H. J., Kim, D. W., Park, K. H., Kwon, H. J., Park, S. J., Lee, W. S., & Ryu, Y. B. (2012). Tanshinones as selective and slow-binding inhibitors for SARS-CoV cysteine proteases. *Bioorganic & medicinal chemistry, 20*, 5928–5935.
48. Shen, L., Niu, J., Wang, C., Huang, B., Wang, W., Zhu, N., Deng, Y., Wang, H., Ye, F., Cen, S., & Tan, W. High-Throughput Screening and Identification of Potent Broad-Spectrum Inhibitors of Coronaviruses, Journal of virology, 93 (2019).
49. Xiao, S., Tian, Z., Wang, Y., Si, L., Zhang, L., & Zhou, D. (2018). Recent progress in the antiviral activity and mechanism study of pentacyclic triterpenoids and their derivatives. *Medicinal research reviews, 38*, 951–976.
50. Li, W., Zhang, C., Sui, J., Kuhn, J. H., Moore, M. J., Luo, S., Wong, S. K., Huang, I. C., Xu, K., & Vasilieva, N. (2005). Receptor and viral determinants of SARS-coronavirus adaptation to human ACE2. *The EMBO journal, 24*, 1634–1643.
51. Wu, K., Li, W., Peng, G., Li, F., Crystal structure of NL63 respiratory coronavirus receptor-binding domain complexed with its human receptor, Proceedings of the National Academy of Sciences, 106 (2009) 19970–19974.
52. Gyebi, G. A., Adegunloye, A. P., Ibrahim, I. M., Ogunyemi, O. M., Afolabi, S. O., & Ogunro, O. B. Prevention of SARS-CoV-2 cell entry: insight from in silico interaction of drug-like alkaloids with spike glycoprotein, human ACE2, and TMPRSS2, J Biomol Struct Dyn, (2020) 1–25.
53. Zhou, Y., Vedantham, P., Lu, K., Agudelo, J., Carrion, R. Jr., Nunneley, J. W., Barnard, D., Pöhlmann, S., McKerrow, J. H., & Renslo, A. R. (2015). Protease inhibitors targeting coronavirus and filovirus entry. *Antiviral research, 116*, 76–84.

54. Yamamoto, M., Matsuyama, S., Li, X., Takeda, M., Kawaguchi, Y., Inoue, J., & Matsuda, Z. (2016). Identification of nafamostat as a potent inhibitor of Middle East respiratory syndrome coronavirus S protein-mediated membrane fusion using the split-protein-based cell-cell fusion assay. *Antimicrobial agents and chemotherapy, 60*, 6532–6539.
55. Zhou, P., Yang, X.-L., Wang, X.-G., Hu, B., Zhang, L., Zhang, W., Si, H.-R., Zhu, Y., Li, B., & Huang, C.-L. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature, 579*, 270–273.
56. Garg, S., Kaul, S. C., & Wadhwa, R. (2018). Cucurbitacin B and cancer intervention: Chemistry, biology and mechanisms. *International journal of oncology, 52*, 19–37.
57. Achenbach, H., Waibel, R., Nkunya, M. H., & Weenen, H. (1992). Antimalarial compounds from Hoslundia opposita. *Phytochemistry, 31*, 3781–3784.
58. Lin, J. H., & Yamazaki, M. (2003). Role of P-glycoprotein in pharmacokinetics. *Clinical pharmacokinetics, 42*, 59–98.
59. Zanin, L., Saraceno, G., Panciani, P. P., Renisi, G., Signorini, L., Migliorati, K., & Fontanella, M. M. SARS-CoV-2 can induce brain and spine demyelinating lesions, Acta Neurochirurgica, (2020) 1–4.

Tables

Table 1: Interacting amino acid residue of human ACE2 and TMPRSS2 with the top binding terpenoids from African phytochemicals

Bioactive compound	Human Protein targets	Interacted residues	Protein atom involved in H-Bonding (BOND DISTANCE)
S1 (MLN-4760)	ACE2	ARG^{514} ARG^{518} ARG^{278} TYR^{510} PHE^{504} MET^{360} LYS^{363} CYS^{344}	TYR^{515}(3.44) THR^{371} (3.03) PRO^{346} (3.08) ARG^{273} (2.93)
24-Methylene cycloartenol **(T1)**		TRP^{163} SER^{170} TYR^{497} TYR^{613} PRO^{492} VAL^{491} SER^{167}	TRP^{163} (3.22) SER^{170} (2.81) TYR^{497} (3.27)
Isoiguesterin (**T2)**		ASP^{350} $TYR^{385}$$ASN^{394}$ ALA^{99} PHE^{40} PHE^{390} LEU^{73} TRP^{69}	ASP^{350} (3.27) TYR^{385} (3.27) ASN^{394} (3.27)
S2 (Camostat)	TMPRSS2	ARG^{41} SER^{195} ALA^{190} ASP^{189} TRP^{215} GLN^{192}	ARG^{41} SER^{195} ALA^{190} ASP^{189} TRP^{215}
11-Hydroxy-2 - (3,4-dihydroxybenzoyloxy)abieta -5,7,9(11),13-tetraene-12-one **(T3)**		ARG^{41} GLN^{192} SER^{195} ALA^{190} ASP^{189} CYS^{191} HIS^{57} CYS^{191}	ARG^{41} (2.41)GLN^{192} (2.89)SER^{195} (2.89)ALA^{190} (2.65)ASP^{189}(2.39)
11-Hydroxy-2 -(4-hydroxybenzoyloxy)-abieta-5,7,9(11),13-tetraene-12-one **(T4)**		GLN^{192} ASP^{189} ALA^{190} SER^{195} HIS^{57} SER^{214} TRP^{192} CYS^{219}	GLN^{192} (2.32) ASP^{189} (2.62) ALA^{190} (2.27) SER^{195} (2.32)

Table 2: Interacting amino acid residue of Spike protein of coronaviruses with the top binding terpenoids from selected African phytochemicals

Bioactive compound	Coronavirus spike proteins	Interacted residues	Protein atom involved in H-Bonding(BOND DISTANCE)
(S3) Nelfinavir mesylates	SARS-Cov-2	THR^{886} ASP^{867} PRO^{869} PRO^{862} VAL^{860} SER^{730} HIS^{1058}	THR^{886} **(3.48)** ASP^{867} **(2.13)** SER^{730} **(2.57)** HIS^{1058} **(2.03)**
3-Benzoylhosloppone (**T5**)		THR^{547} PHE^{541} LEU^{546} PRO^{589}	THR^{547}**(3.03)**
Cucurbitacin B (**T6**)		ARG^{1091} ASN^{914} THR^{912} GLN^{1113} PHE^{1121} ILE^{1114} GLY^{1124}	ARG^{1091} **(2.93)** ASN^{914} **(3.32)** THR^{912} **(2.95)** GLN^{1113} **(2.89)**
(S3) Nelfinavir mesylates	SARS-CoV	SER^{556} THR^{535} THR^{559} PHE^{558} PRO^{575} PHE^{527}	SER^{556} **(2.14)** THR^{535} **(2.38, 2.59)** THR^{559} **(3.30)**
7-Deacetoxy-7-oxogedunin (**T7**)		ARG^{982} GLY^{726} VAL^{958} PHE^{837}	ARG^{982} **(2.73, 2.16)** GLY^{726} **(2.52)**
3-Friedelanone (**T8**)		ARG^{982} GLY^{726} VAL^{958} PHE^{837} VAL^{945} LYS^{836} LEU^{948} ASN^{838}	ARG^{982} (3.23) GLY^{726} (3.03) ASN^{838} (3.12)
(S3) Nelfinavir mesylates	MERS-CoV	SER^{51} ARG^{335} HIS^{348} HIS^{670} LEU^{344} ILE^{337} PHE^{354} LYS^{668}	SER^{51} (2.90) ARG^{335} (2.89)
7-Deacetoxy-7-oxogedunin (**T7**)		SER^{51} HIS^{348} HIS^{670} ILE^{337} PHE^{354} LEU^{344} ARG^{335}	SER^{51} (2.74)
3-Benzoylhosloppone		LYS^{689} PHE^{341} MET^{698} VAL^{958} LEU^{344} ILE^{337}	

Table 3: Shows the number of clusters produced from TTClust, its representative frame for each of the protein-ligand complexes, and the interactions between the ligand and the protein from PLIP webserver for that frame.

TMPRSS2_(11-Hydroxy-2 - (3,4-dihydroxybenzoyloxy)abieta-5,7,9(11),13-tetraene-12-one) complex				
CLUSTER NUMBER (REPRESENTATIVE FRAME)	**Hydrophobic**	**H-bond**	**Salt-bridges**	**Pi-cation**
Cluster 1 (frame 140)	W215	A190 (2)	None	None
Cluster 2 (frame 853)	**R41** - T62	S39 – H40 - **R41**	**R41**	**R41**
Cluster 3 (frame 977)	T61	None	None	**R41**
ACE2 _(24-Methylene cycloartenol) complex				
CLUSTER NUMBER (REPRESENTATIVE FRAME)	**Hydrophobic**			
CLUSTER 1 (FRAME 172)	Y255 (2)- P612			
CLUSTER 2 (FRAME 721)	L142 (2) – I151 – L162 (2)			
TMPRSS2_ camostat complex				
cluster number (representative frame)	**Hydrophobic**	**H-bond**	**Salt-bridges**	**Pi-stacking**
Cluster 1 (frame 92)	None	**A190** (2) – Q192 – **D217** – E218 – A220	**D189**	H57
Cluster 2 (frame 618)	Q192 - V213	R41 – **A190** – S195 – **D217** – E218	H57 - **D189**	None
Cluster 3 (frame 284)	Q192	**A190** (2) – S195 – **D217** (2)	**D189**	None
Cluster 4 (frame 728)	None	**A190** (2) – **D217** (2)	**D189**	None
Cluster 5 (frame 915)	Q192	**A190** (2) – **D217** (2) – A220	**D189**	None
S protein_(3- Benzoylhosloppone) complex				
CLUSTER NUMBER (REPRESENTATIVE FRAME)	**Hydrophobic**	**H-bond**	**Pi-stacking**	
CLUSTER 1 (FRAME 184)	L546 – **V576** – I587 (2) - P589	None	None	
CLUSTER 2 (FRAME 631)	F541 – **F543** – L546 – T549 - P589	T573	**F543**	
CLUSTER 3 (FRAME 935)	F541 – **F543** – L546 – F565 – 573 - **V576**	None	None	

Amino acid residues were represented by single letter code. Bold amino acids are common in each protein-drug complex

Table 4: Physicochemical properties of the top binding terpenoids from African plants to ACE2, TMPRSS2 and S protein of SARS-Cov-2

a) Lipinski filter analysis				
Lipinski filters	**T1**	**T3**	**T5**	**T6**
Molecular weight (g/mol)	454.77	450.52	402.48	558.70
Num. heavy atoms	33	33	30	40
Num. rotatable bonds	5	4	4	6
Num. H-bond acceptors	1	6	4	8
Hydrogen bond donor	1	3	0	3
MLogP	7.30	2.96	3.79	1.76
Molar Refractivity	144.50	126.11	116.15	150.94
Lipinski violation	1	0	0	1
(b) admet SAR				
Absorption (Probability)				
Blood-Brain Barrier	BBB+ (0.96)	BBB+ (0.60)	BBB+ (0.61)	BBB+ (0.81)
Human Intestinal Absorption	HIA+ (0.99)	HIA+ (0.92)	HIA+ (0.92)	HIA+ (0.97)
Bioavailability Score	0.55	0.55	0.55	0.55
Caco-2 Permeability	Caco2+ (0.79)	Caco2+ (0.59)	Caco2+ (0.59)	Caco2+ (0.61)
P-glycoprotein Substrate	Substrate (0.73)	Substrate (0.78)	Non-inhibitor (0.58)	Substrate (0.79)
P-glycoprotein Inhibitor	Non-inhibitor (0.65)	Non-inhibitor (0.74)	Non-inhibitor (0.74)	Non-inhibitor (0.61)
Renal Organic Cation Transporter	Inhibitor (0.75)	Inhibitor (0.90)	Non-inhibitor (0.90)	Non-inhibitor (0.87)
Distribution (Probability)				
Subcellular localization	Lysosome (0.55)	Mitochondria (0.86)	Mitochondria (0.86)	Mitochondria (0.77)
Metabolism				
CYP450 Substrate	Substrate (0.77) Non-inhibitor (0.78)	Substrate Non-inhibitor (0.83)	Non-substrate (0.65) inhibitor (0.80)	Inhibitor (0.79) Non-substrate (0.83)

Toxicity				
AMES Toxicity	Non AMES toxic (0.71)	AMES toxic (0.87)	Non AMES toxic (0.87)	Non AMES toxic (0.84)
Carcinogens	Non-carcinogens (0.92)	Non-carcinogens (0.90)	Non-carcinogens (0.90)	Non-carcinogens (0.92)
Acute Oral Toxicity	III (0.77)	III (0.59)	III (0.57)	I (0.78)
Rat Acute Toxicity LD_{50}, mol/kg	3.2804	2.5370	2.5370	3.8742
Aqueous solubility (LogS)	-4.76258	-4.5550	-4.7201	-4.5035
Pharmacokinetics				
Lower GI absorption	Low	High	High	low
Log K_p (skin permeation) cm/s	-1.48	-5.58	-5.33	-7.83

4-Methylene cycloartenol **(T1);** 11-Hydroxy-2 - (3,4-dihydroxybenzoyloxy)abieta -5,7,9(11),13-tetraene-12-one**(T3);** 3- Benzoylhosloppone (**T5) and** Cucurbitacin B (**T6)**

Figures

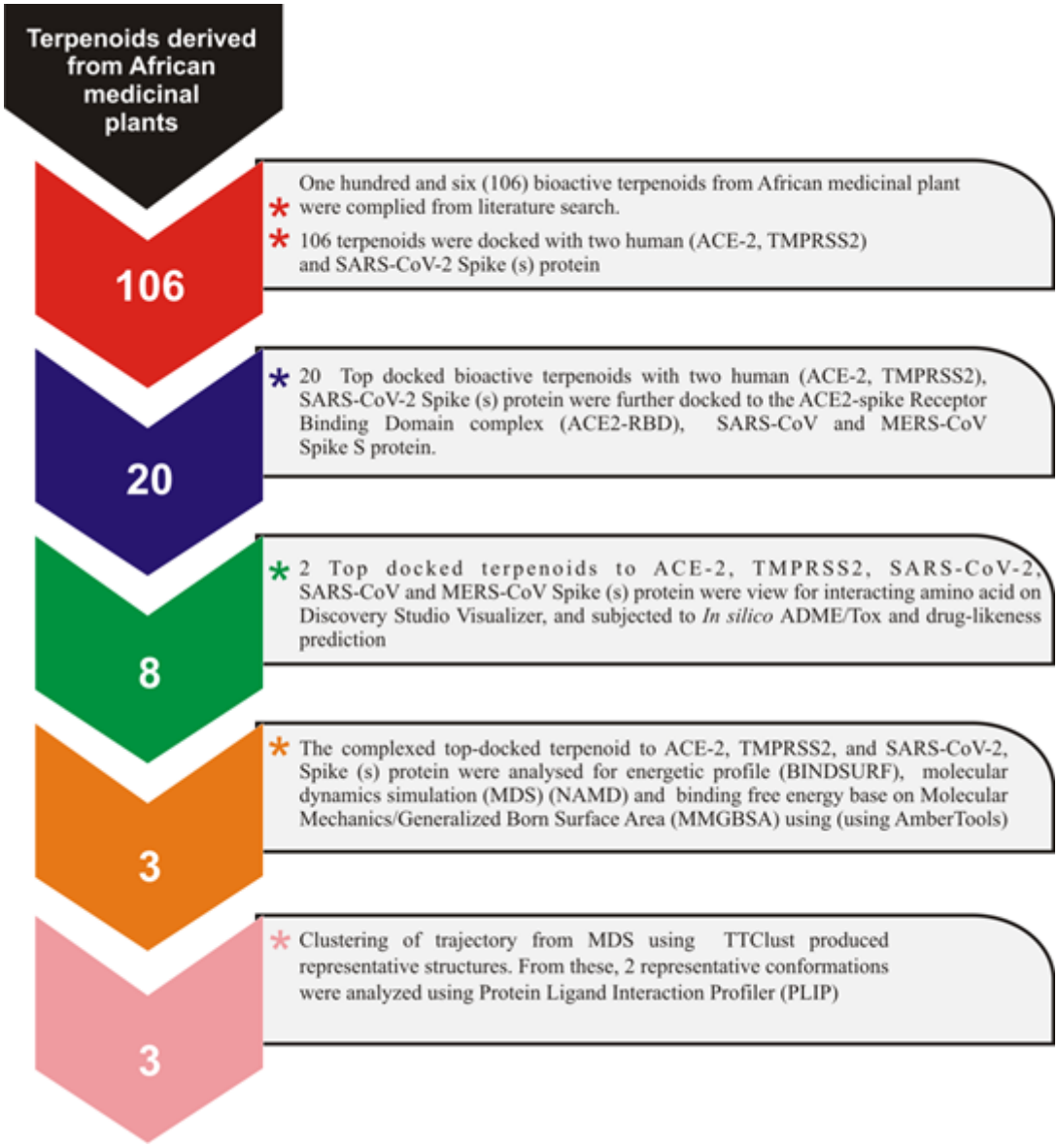

Figure 1

Flow chart showing the stepwise screening of African derived terpenoids for potential inhibitors of membrane-mediated SARS-CoV-2 cell entry

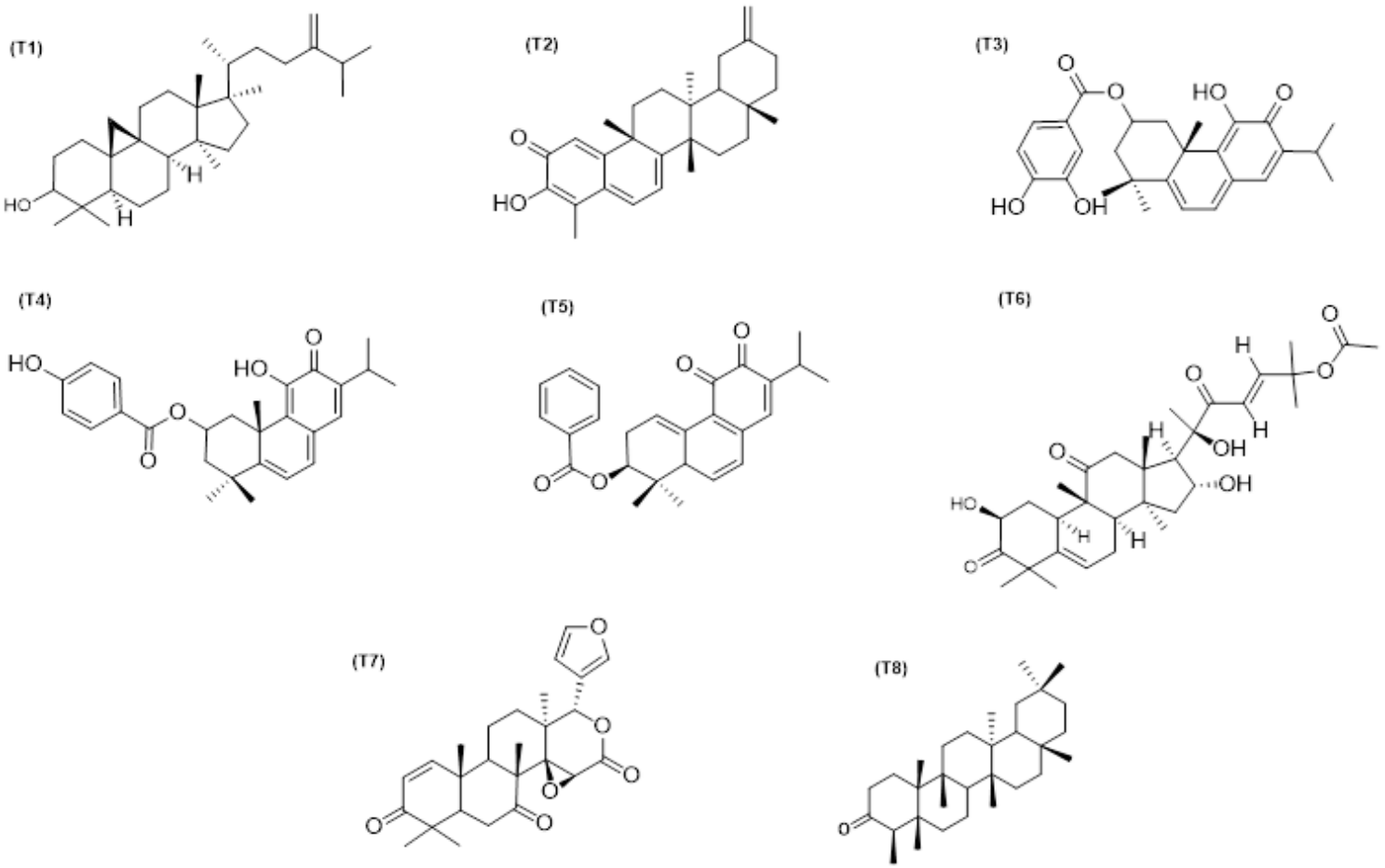

Figure 2

Chemical structure of terpenoid with remarkable binding energyto human ACE2, TMPRSS2 and SARS-coronoviruses S protein (T1) 24-Methylene cycloartenol; (T2) Isoiguesterin;(T3)11-Hydroxy-2 - (3,4-dihydroxybenzoyloxy)abieta-5,7,9(11),13-tetraene-12-one; (T4)11-Hydroxy-2 -(4-hydroxybenzoyloxy)-abieta-5,7,9(11),13-tetraene-12-one; (T5) 3-Benzoylhosloppone; (T6) Cucurbitacin B; (T7) 7-Deacetoxy-7-oxogedunin and (T8) 3-Friedelanone.

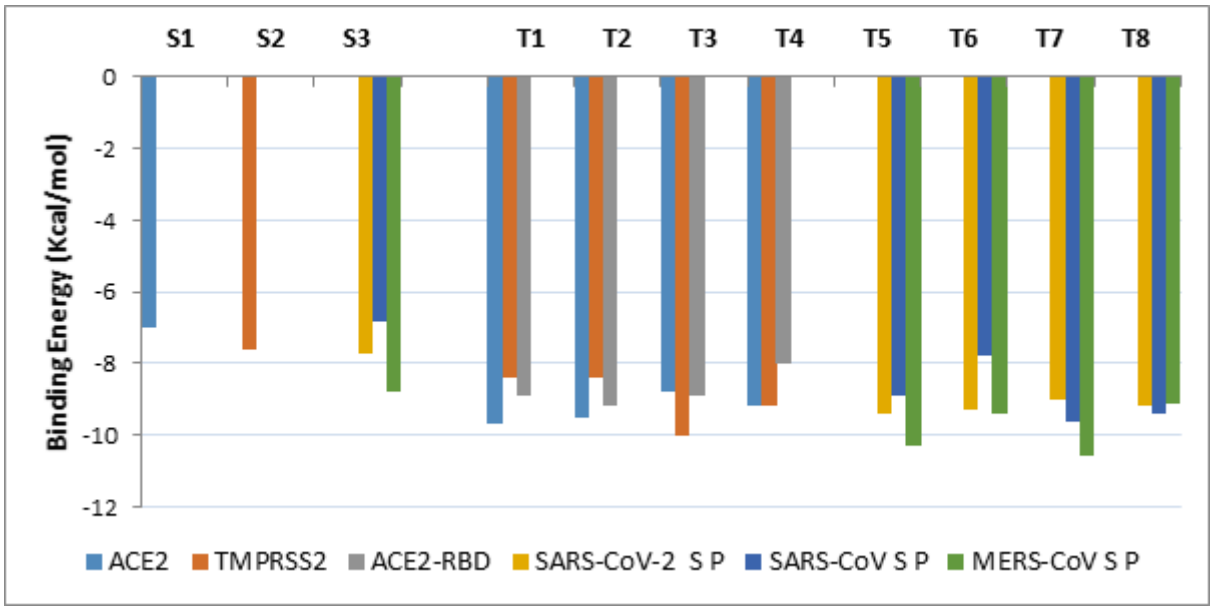

Figure 3

AutoDock scores (binding energies) of standard drugs and top bioactive terpenoids with human Angiotensin-Converting Enzyme 2 (ACE2), Transmembrane Protease Serine 2 (TMPRSS2), and ACE2-Spike Receptor Binding Domain complex (ACE2-RBD) and (*S P) spike protein of coronaviruses. S1: MLN-4760; S2: Camostat S3: Nelfinavir mesylates T1: 24-Methylene cycloartenol; T2: Isoiguesterin; T3:11-Hydroxy-2 - (3,4-dihydroxybenzoyloxy)abieta-5,7,9(11),13-tetraene-12-one; T4: 11-Hydroxy-2 -(4-hydroxybenzoyloxy)-abieta-5,7,9(11),13-tetraene-12-one; T5: 3-Benzoylhosloppone; T6: Cucurbitacin B; T7: 7-Deacetoxy-7-oxogedunin and T8: 3-Friedelanone.

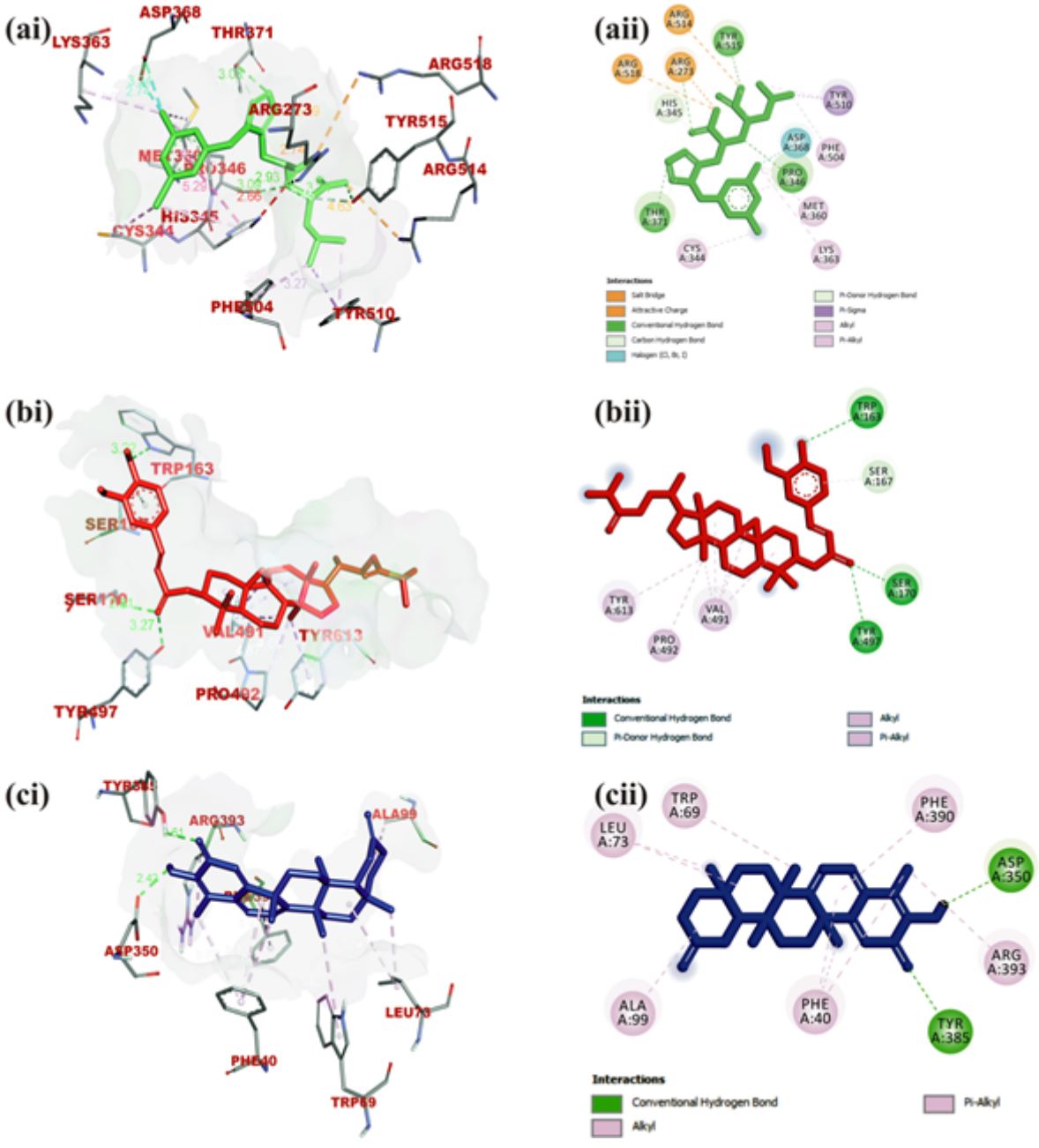

Figure 4

Visualization of interacting amino acid residues of human ACE2 with ligands in 3D (i) and 2D (ii) representation: Ligands in stick representation are presented in different colours: (a) Green: S1 (MLN-4760) (b) Read: 24-methylene cycloartenol (c)Blue: Isoiguesterin. Types of interactions are represented by Green- dotted lines: H-bond interactions, light purple-dotted line: hydrophobic interactions (Pi-Alkyl, Alkyl & pi-stacking) purple-dotted line: Pi-Pi T Shaped, yellow-dotted lines: Pi-sulphur interactions, pi-stacking interactions. Three-letter amino acids are in red colour

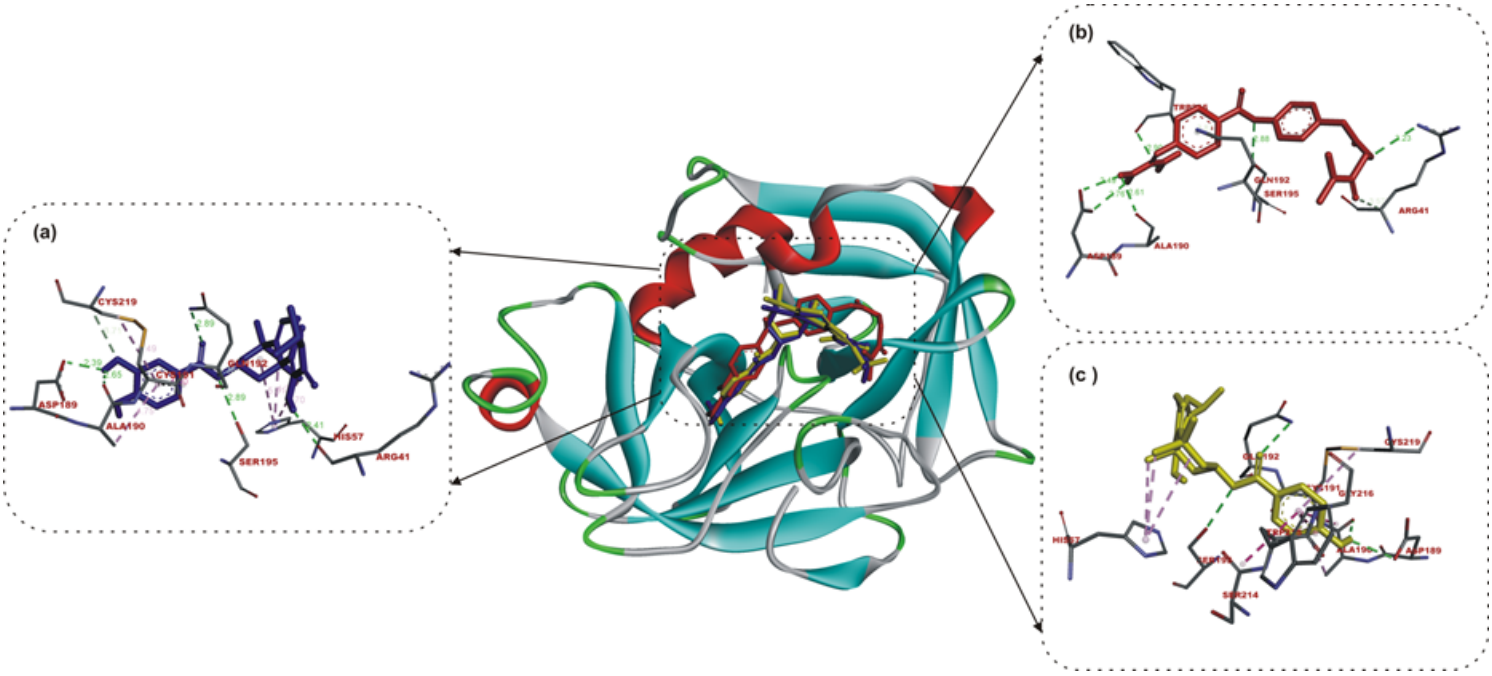

Figure 5

The interactive view of ligands in binding cavity of human TMPRSS2 in cartoon representation. Ligands in sticks representation are represented by colours (a) Green:Camostat (b) Red:11-Hydroxy-2 - (3,4-dihydroxybenzoyloxy)abieta -5,7,9(11),13-tetraene-12-one (c) Yellow: 11-Hydroxy-2 -(4-hydroxybenzoyloxy)-abieta-5,7,9(11),13-tetraene-12-one. Types of interactions are represented by Green-dotted lines: H-bond interactions, light purple-dotted line: hydrophobic interactions (Pi-Alkyl, Alkyl & pi-stacking) purple-dotted line: Pi-Pi T Shaped, yellow-dotted lines: Pi-sulphur interactions, pi-stacking interactions. Three-letter amino acids are in red colour

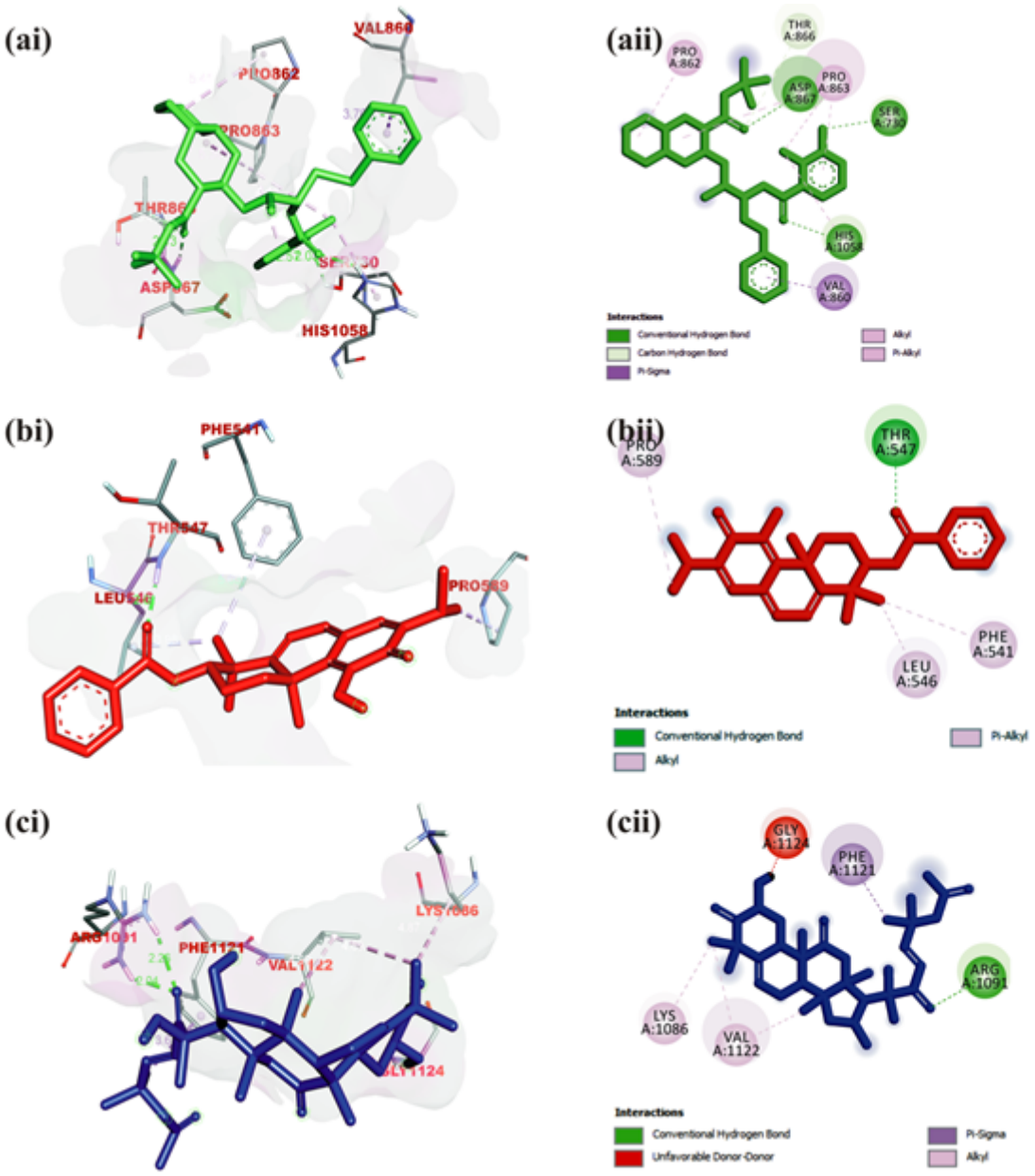

Figure 6

Visualization of interacting amino acid residues of SARS-Cov-2 Spiike proten with ligands in 3D (i) and 2D (ii) representation: Ligands in stick representation are presented in different colours: (a) Green: Nelfinavir mesylates (S3) (b) Red: 3- benzoylhosloppone (c) Blue: Cucurbitacin B. Types of interactions are represented by Green- dotted lines: H-bond interactions, light purple-dotted line: hydrophobic interactions (Pi-Alkyl, Alkyl & pi-stacking) purple-dotted line: Pi-Pi T Shaped, yellow-dotted lines: Pi-sulphur interactions, pi-stacking interactions. Three-letter amino acids are in red colour

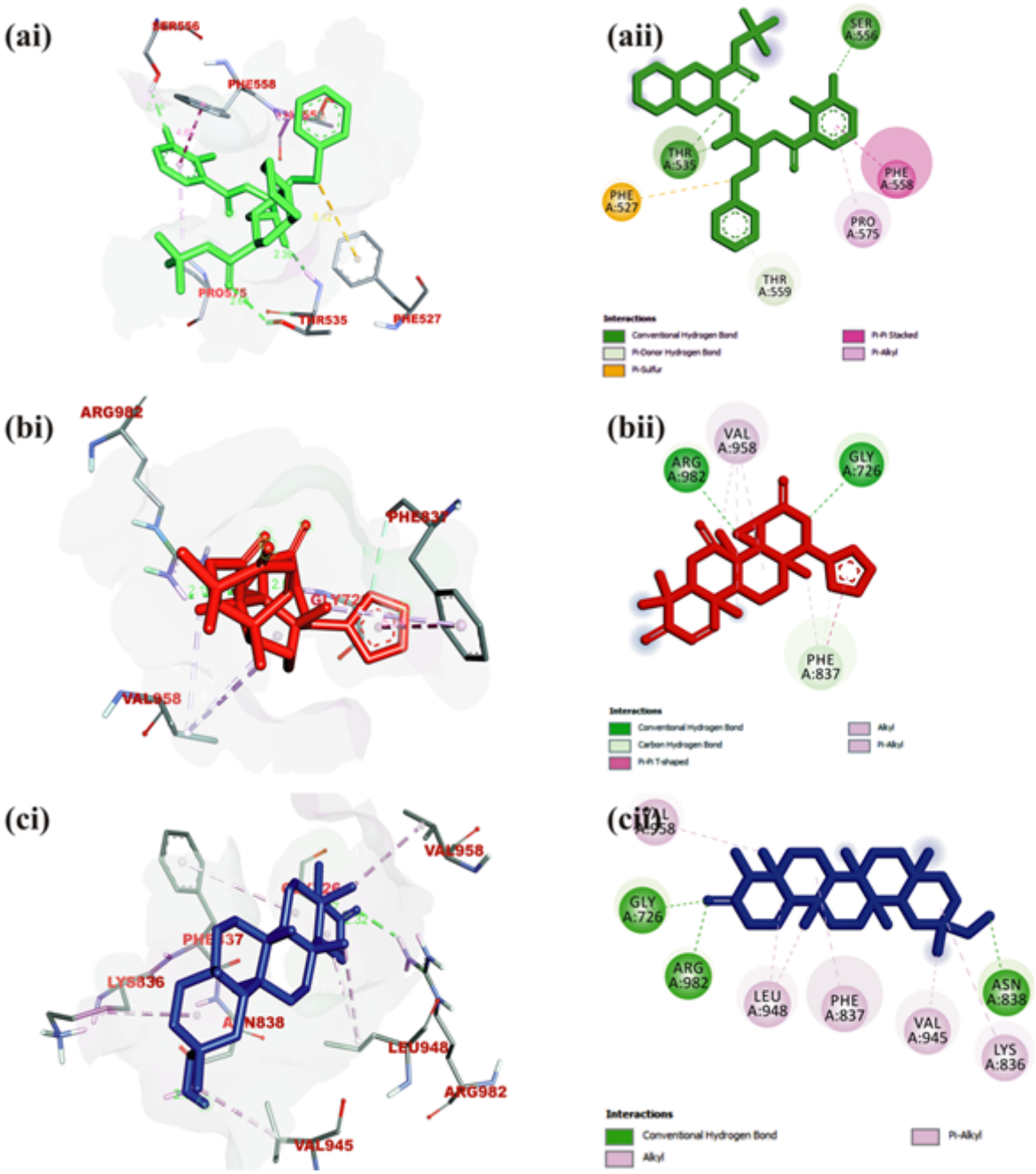

Figure 7

Visualization of interacting amino acid residues of SARS-Cov-2 Spiike proten with ligands in 3D (i) and 2D (ii) representation: Ligands in stick representation are presented in different colours: (a) Green: Nelfinavir mesylates (S3) (b) Red: 7-deacetoxy-7-oxogedunin (c) Blue: 3-friedelanone.. Types of interactions are represented by Green- dotted lines: H-bond interactions, light purple-dotted line: hydrophobic interactions (Pi-Alkyl, Alkyl & pi-stacking) purple-dotted line: Pi-Pi T Shaped, yellow-dotted lines: Pi-sulphur interactions, pi-stacking interactions. Three-letter amino acids are in red colour

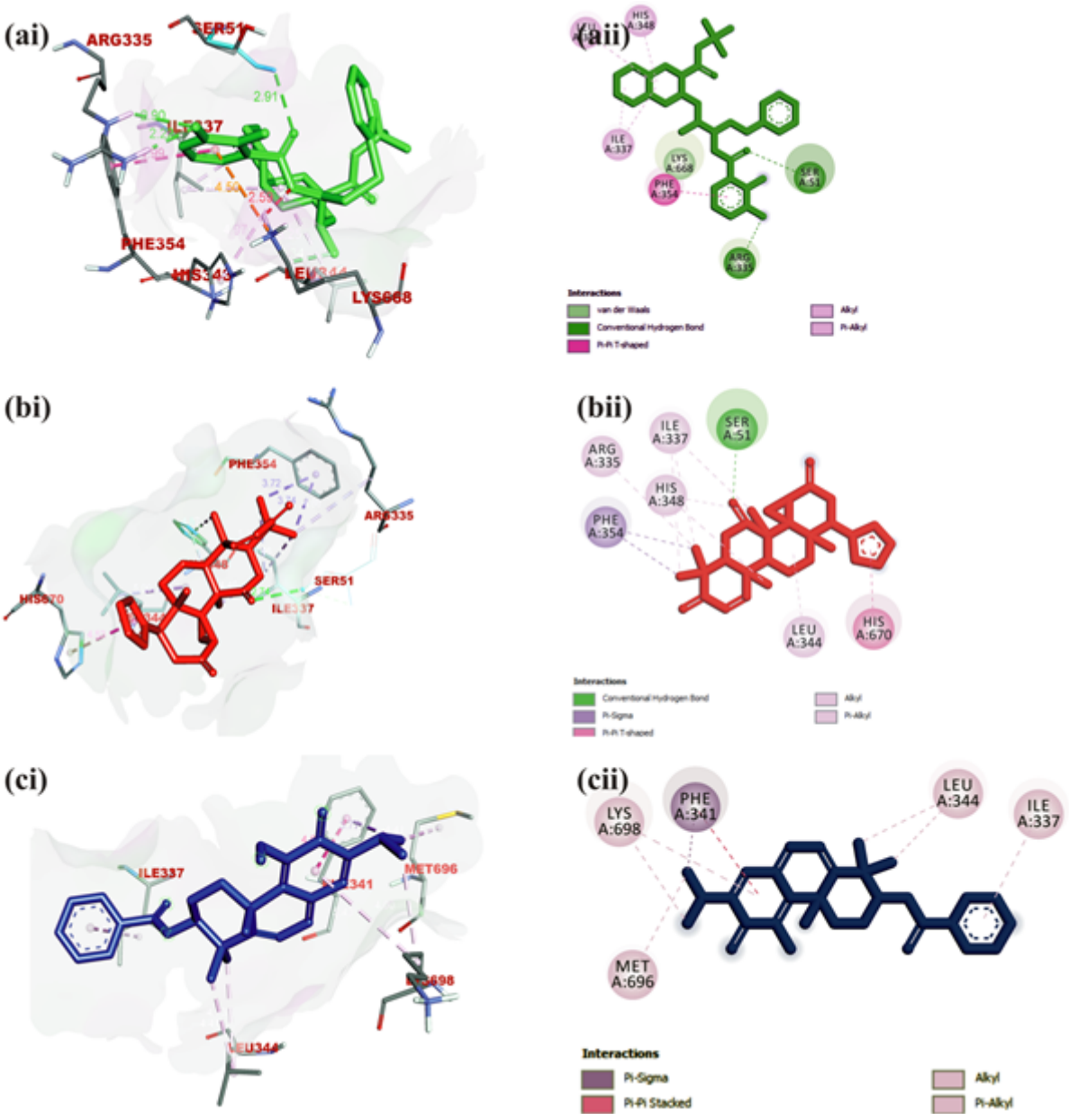

Figure 8

Visualization of interacting amino acid residues of SARS-Cov-2 Spiike proten with ligands in 3D (i) and 2D (ii) representation: Ligands in stick representation are presented in different colours: (a) Green: Nelfinavir mesylates (S3) (b) Red: 7-deacetoxy-7-oxogedunin (c) Blue: 3- benzoylhosloppone. Types of interactions are represented by Green- dotted lines: H-bond interactions, light purple-dotted line: hydrophobic interactions (Pi-Alkyl, Alkyl & pi-stacking) purple-dotted line: Pi-Pi T Shaped, yellow-dotted lines: Pi-sulphur interactions, pi-stacking interactions. Three-letter amino acids are in red colour

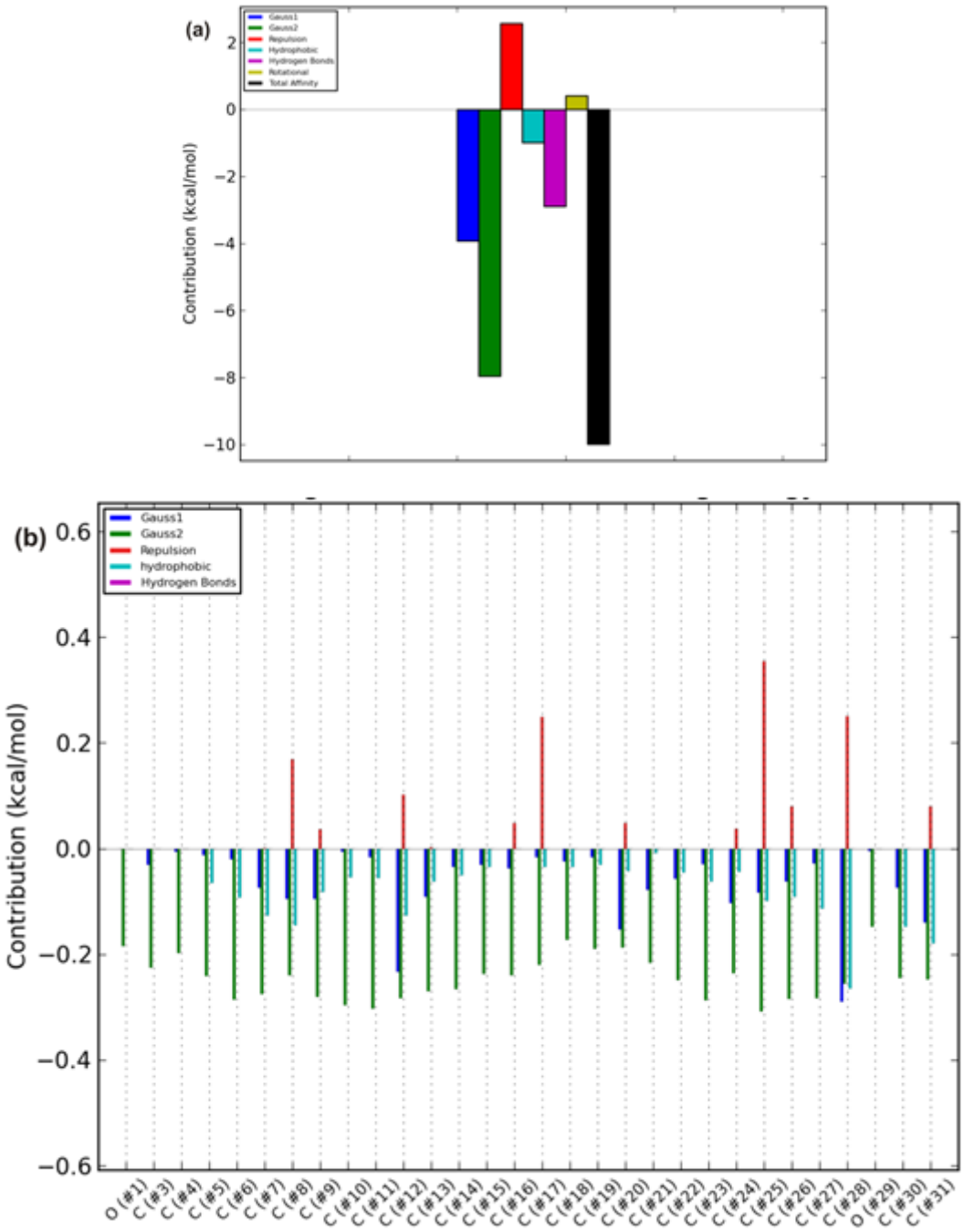

Figure 9

Energy profile of 24-Methylene cycloartenol binding groups in human ACE2: (a) Energetic contribution to the Binding energy (d) Energetic contributions for each atom in the ligand. Number of poses in selected cluster: 68, best pose: 116 and binding site coordinate: 39.14, 35.33, and 12.71

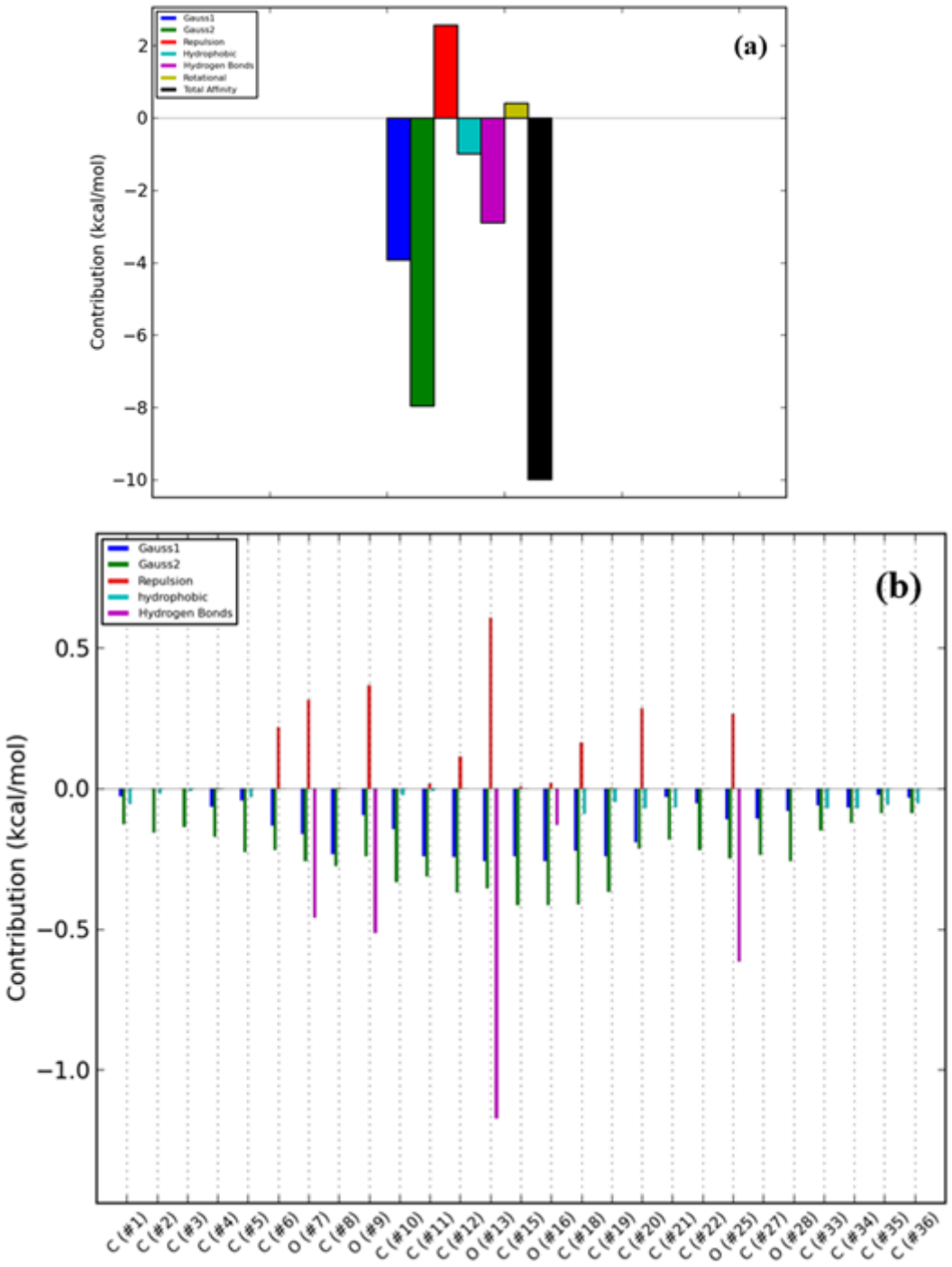

Figure 10

Energy profile of T3 binding groups in human TMPRSS2: (a) Energetic contribution to the Binding energy (d) Energetic contributions for each atom in the ligand. Number of poses in selected cluster: 87, best pose: 40 and binding site coordinate: -2.96, 26.97, and 23.55

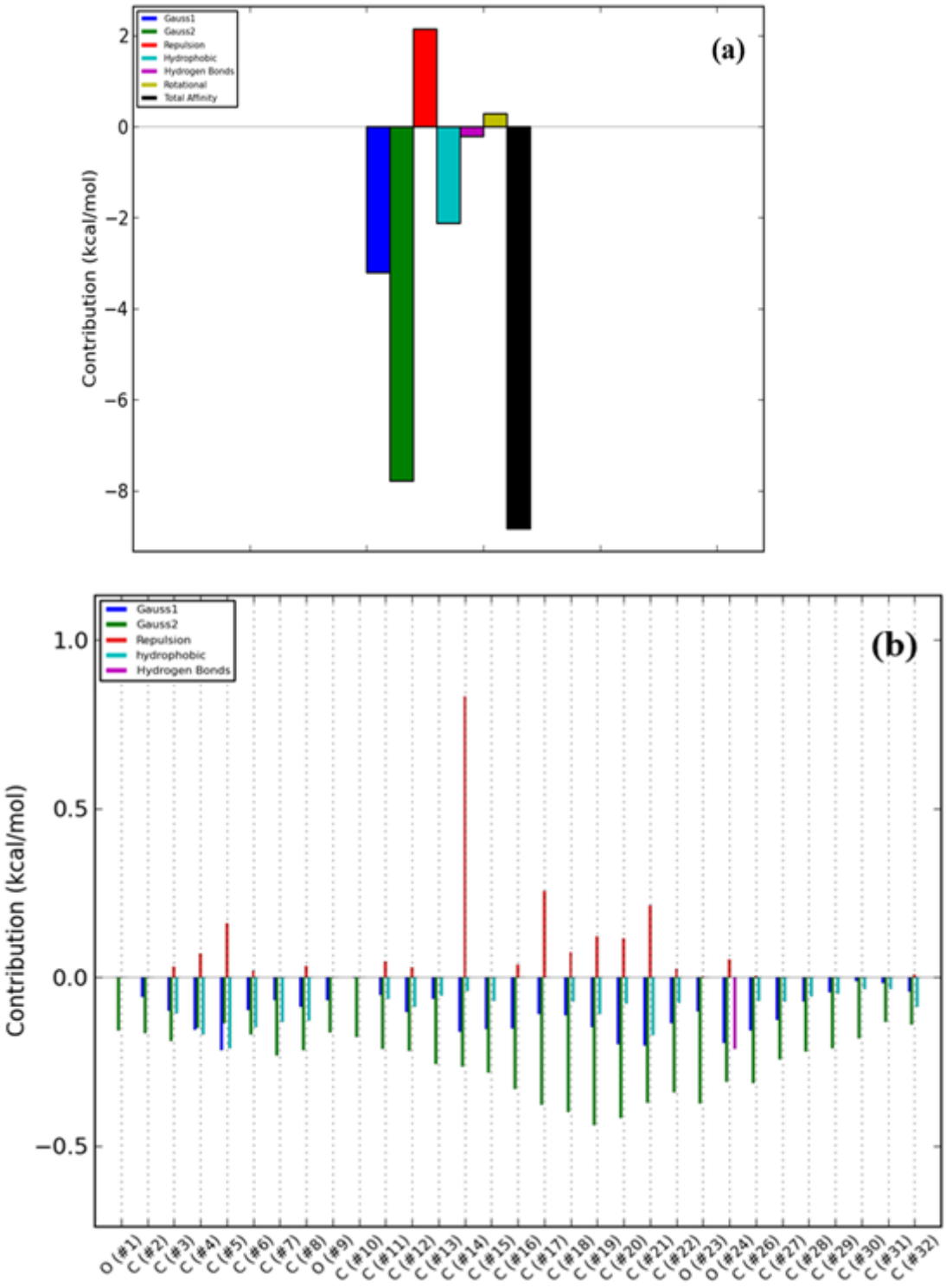

Figure 11

Energy profile of 3- benzoylhosloppone binding groups in SARS-Cov-2 S protein (a) Energetic contribution to the Binding (b) Energetic contributions for each atom in the ligand. Number of poses in selected cluster: 49, best pose: 571 and binding site coordinate: 214.85, 246.53, and 212.68

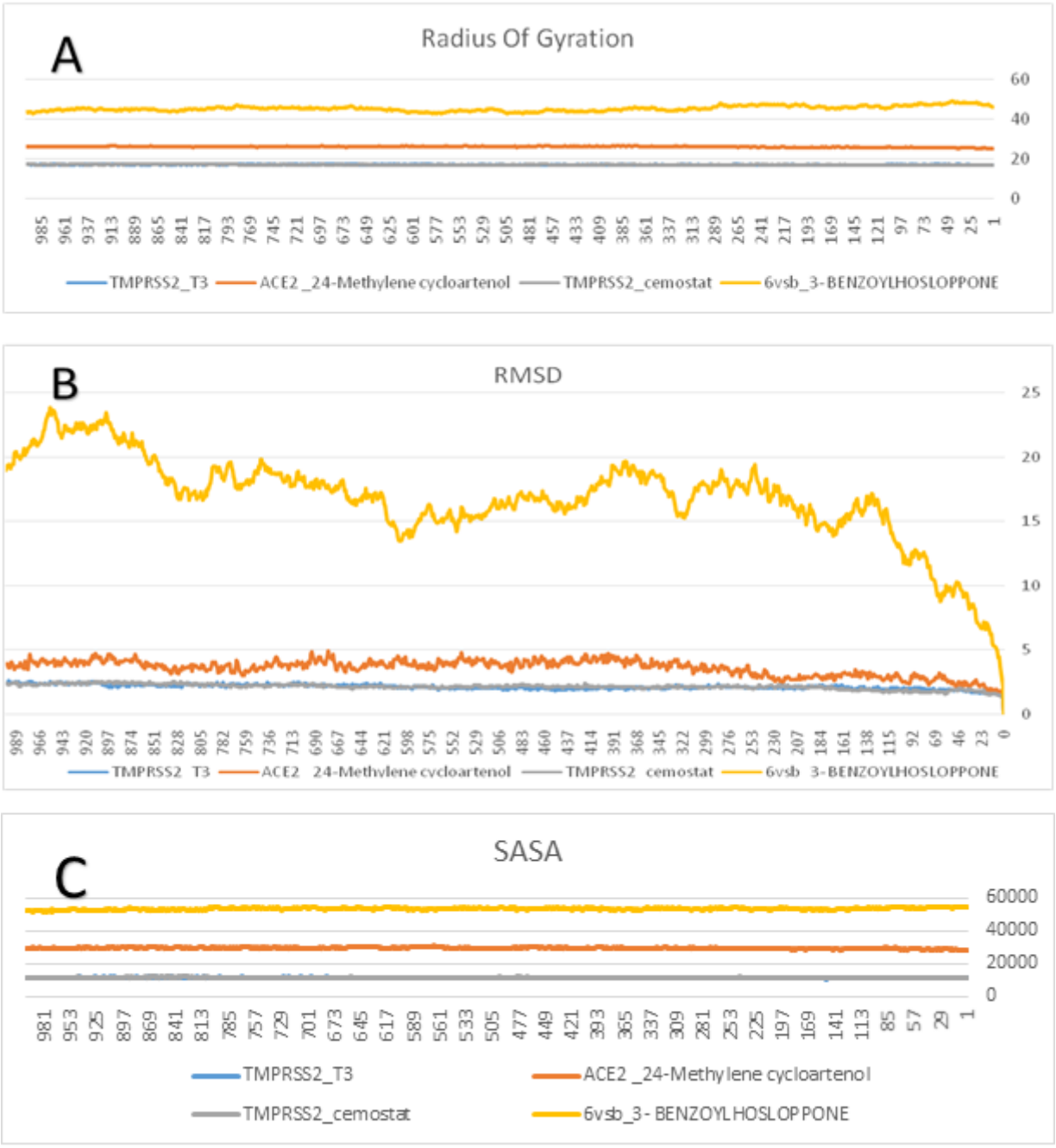

Figure 12

The Radius of Gyration, Root Mean Square Deviation (RMSD), and Surface Accessible Surface Area (SASA) for each of protein-ligand complexes

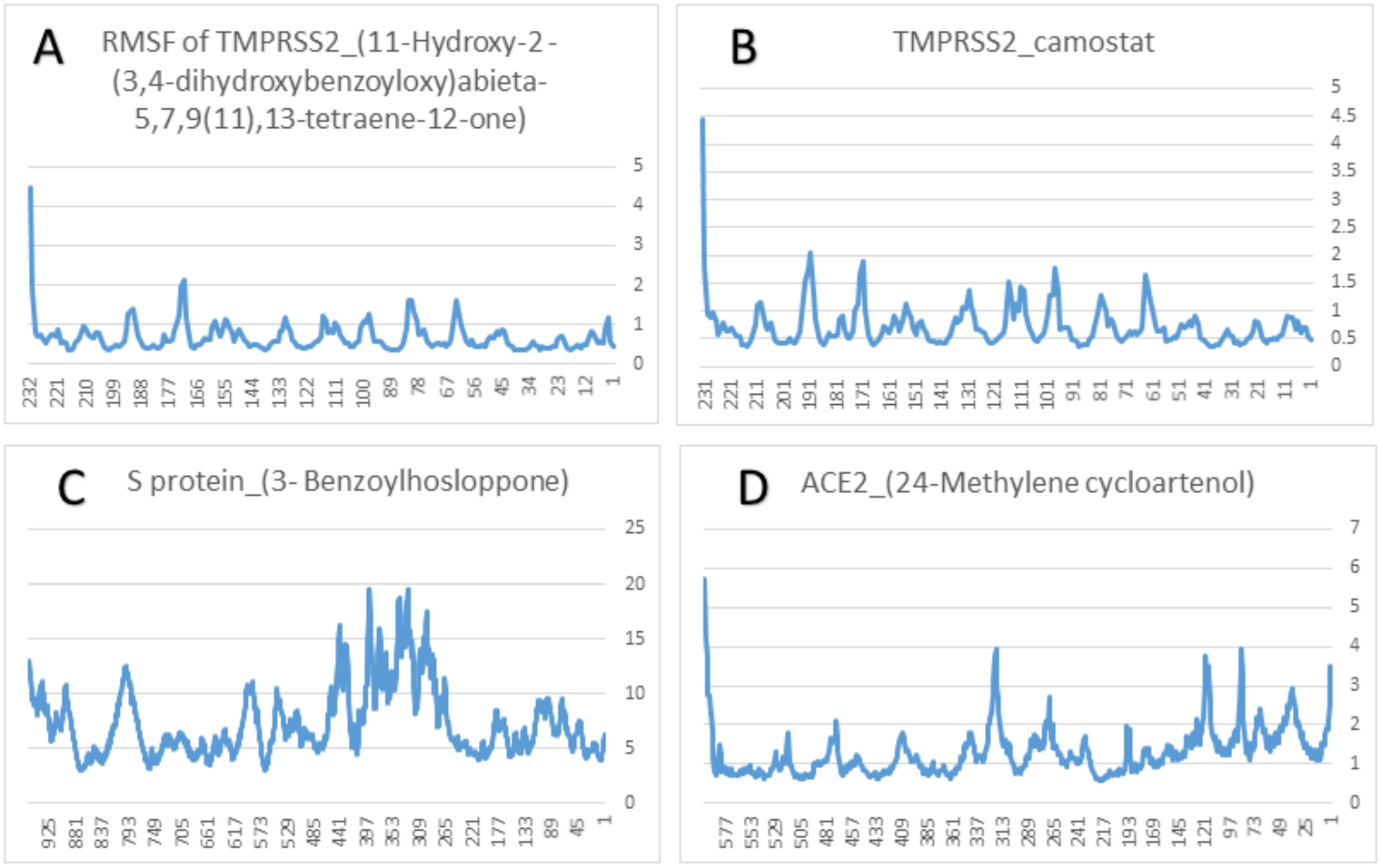

Figure 13

The Root Mean Square Fluctuation for (A) TMPRSS2_(11-Hydroxy-2 - (3,4-dihydroxybenzoyloxy)abieta-5,7,9(11),13-tetraene-12-one), (B) TMPRSS2_camostat, (C) S protein_(3- Benzoylhosloppone), and (D) ACE2_(24-Methylenecycloartenol), respectively.

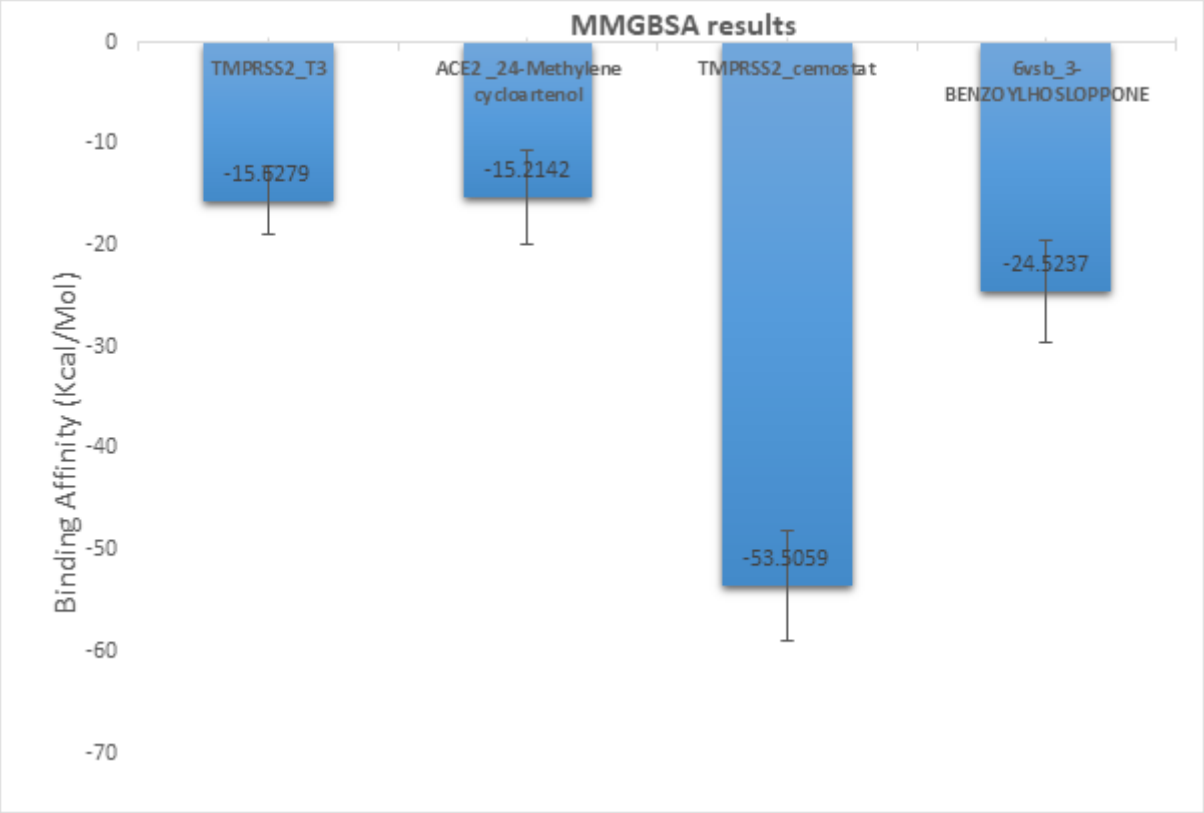

Figure 14

The binding free energy from MMGBSA analysis for each protein-drug complex. Error bars represent standard deviation. T3 = 11-Hydroxy-2 - (3,4-dihydroxybenzoyloxy) abieta-5,7,9(11),13-tetraene-12-on.

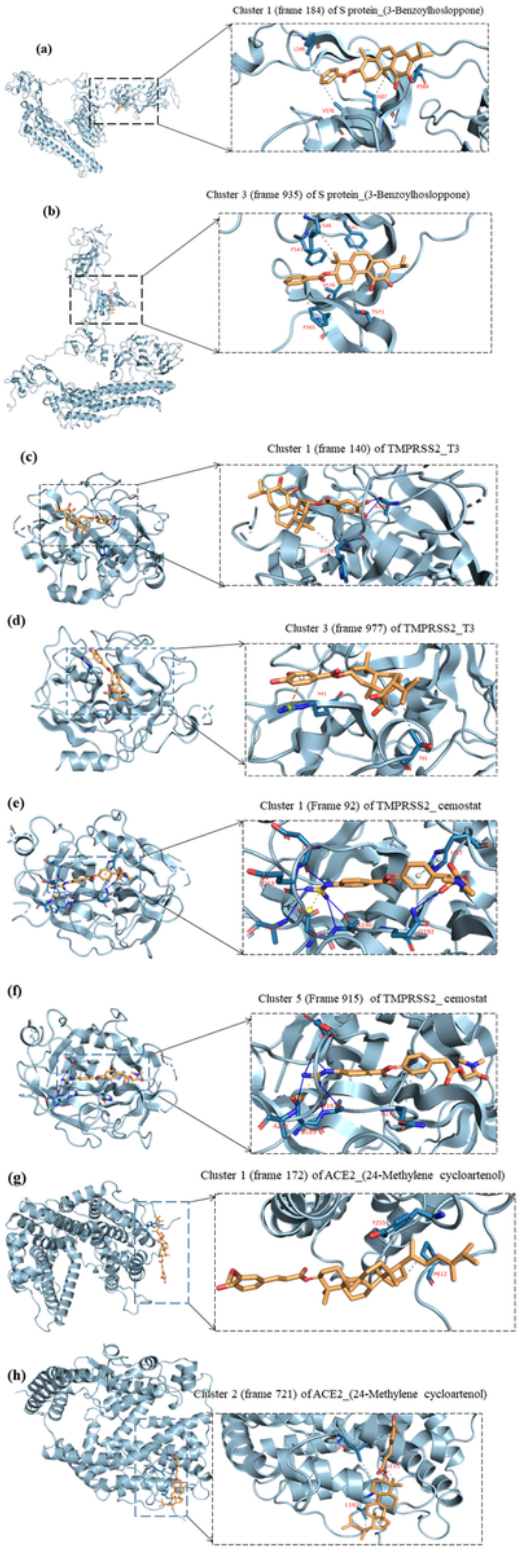

Figure 15

The representative structure for each cluster in cartoon representation, ligands in sticks representation and the types of interactions. Gray-dotted line: hydrophobic interactions, blue lines: H-bond interactions, yellow-dotted lines: salt-bridges interactions, and green-dotted lines: pi-stacking interactions. Single-letter amino acids are in red color.

Supplementary Files

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

- SupplementaryData.docx