

Anti-COVID-19 terpenoid from marine sources: a docking, ADMET and molecular dynamics study

Nayim Sepay^{a,*}, Aishwarya Sekar^b, Umesh C Halder^{a,*}, Abdullah Alarifi^c, Mohd Afzal^c

^a*Department of Chemistry, Jadavpur University, Kolkata-70032, India.*

^b*Department of Bioinformatics, Stella Maris College (Autonomous), Chennai, Tamilnadu-600 086, India.*

^c*Catalytic Chemistry Research Chair, Department of Chemistry, College of Science, King Saud University, Riyadh 11451, Saudi Arabia.*

Methods

The study focused on the SARC-CoV-2 M^{Pro}, i.e. PDB ID: 5r7y, 5r7z, 5r80, 5r81, 5r82, 5r83, 5r84, 6lu7 and 6y7m for *in silico* studies to find some SARC-CoV-2 M^{Pro} inhibitors from a library of natural products. The three dimensional structure of SARC-CoV-2 M^{Pro} enzymes were collected from the RCSB website.¹

Density Functional Theory

The energy minimized structures of all the NPs were obtained with the help of the density functional theory (DFT). The optimization of ground state geometries of the NPs with B3LYP functional at 6-311G level of theory in gaseous state. The job have been performed with the Gaussian 09W Revision D.01 program² on the Windows platform.

Molecular docking

Energy minimized structure of all the investigated NPs which was obtained from DFT optimization were used for docking simulations with the SARC-CoV-2 M^{Pro} (PDB ID: 6lu7) protein structure. The docking studies was performed using AutoDock 4.2.0 applications through Autodock tools at the Windows platform.³ The MGL Tools was utilized in the preparation of the structure of the NPs and the proteins in appropriate formats which were required for the calculations. In the case of SARC-CoV-2 M^{Pro} enzyme with the NPs, the partial atomic charges (Gasteiger charges) have been allocate after putting hydrogens to all the atoms of the protein as well as the NPs, separately. Here, the NPs structures allowed as flexible moiety and the SARC-CoV-2 M^{Pro} enzyme structure kept as rigid during the docking studies. The ten conformers of NPs inside the active site of the SARC-CoV-2 M^{Pro} enzyme having minimum potential energy were obtained through subsequent 20 000 precise docking step with 1000 exhaustiveness parameter inside the 60 x 60 x 60 Å³ grid box using a Lamarckian genetic algorithm.

Protein structure modelling.

The protein crystal structure data of all the SARC-CoV-2 M^{Pro} enzymes were downloaded from Protein Data Bank (PDB ID: 5r7y, 5r7z, 5r80, 5r81, 5r82, 5r83, 5r84, 6lu7 and 6y7m). The protein structures utilized in all the further studies were prepared by Discovery studio 2017 R2 client. The pictures of the protein used here was made with MolSoft-ICM browser, Meastro 11.1, Samson core and Discovery studio 2017 R2 client. The SARC-CoV-2 M^{Pro} protein-NPs docked complex with lowest potential energy structures also analyzed by the aforesaid software.

Molecular Electrostatic Potential (MEP) Analysis

The energy minimized structures of the NPs gained from DFT were further utilized for MEP calculations. The same functional used for DFT has also been employed to generate the

electrostatic potential map throughout the atomic framework of the NPs molecules. Here, the basis set was 6-31g with the 0.03 iso values. All these calculations were performed on the Windows version of Gaussian 09W software with D1 revision.²

ADME study

The SwissADME web server was utilized for all the ADME calculations⁴ of the NPs showed top most binding affinity toward SARS-CoV-2 M^{pro}. The server have a strong data base to predict physicochemical properties like lipophilicity, water solubility, drug likeness, pharmacokinetics and medicinal properties with high accuracy.

Toxicity

The probability of Cardiac toxicity for all the NPs having high binding score in docking studies were calculated by Pred-hERG which is the only web-accessible computational server for this toxicity.⁵ All the other type of toxicity of these NPs and FDA approved anti-viral drugs have been predicted using PROTOX-II.⁶ In this case, we have considered the acute toxicity, organ toxicity, toxicological endpoint, nuclear receptor signaling pathways and stress response pathway.

Molecular Dynamics simulation

Understanding the stability of protein upon ligand binding is significantly improved by molecular dynamics simulation studies. Molecular dynamics simulation of the SARS CoV2 Main Protease and the ligand **T3** (from marine sponges) was performed with Groningen Machine for Chemical Simulation (GROMACS) version 2020.2. Topology parameters for protein and ligand were generated With GROMOS96 54a7 force field and Ligand topology was obtained from the PRODRG2 server. The protein -ligand system was embedded in a cubic box of approximate size with periodic boundary conditions using a simple point charge water solvation model (38012 water molecules)⁷. The overall system was neutralized by adding 4 Na⁺ ions in solution and the SHAKE algorithm was used to constrain all bond lengths involving hydrogen atoms. Particle Mesh Ewald method with a cutoff of 12 Å was applied to treat the long range electrostatic interactions. The processed system was suitably minimized followed by the NPT and NVT ensemble equilibration steps at a uniform temperature and pressure of 300 K and 1 bar, respectively maintained for each system with Parrinello–Rahman barostat. The trajectories were saved at every 2 fs time step and the production MD simulation of the protein-ligand complex was performed for 95 nanoseconds⁸.

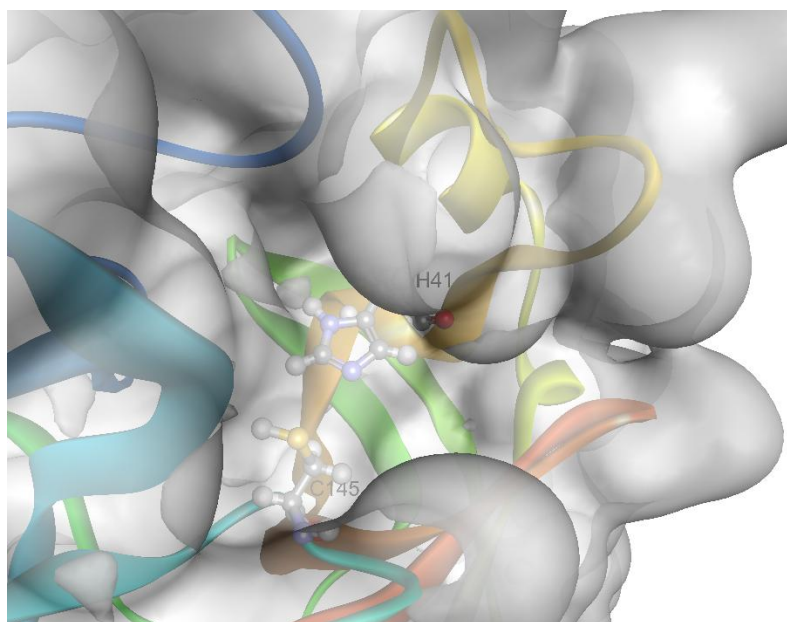


Figure S1: The H41 and C145 amino acid residue inside the active site of SARS-CoV-2 M^{pro}.

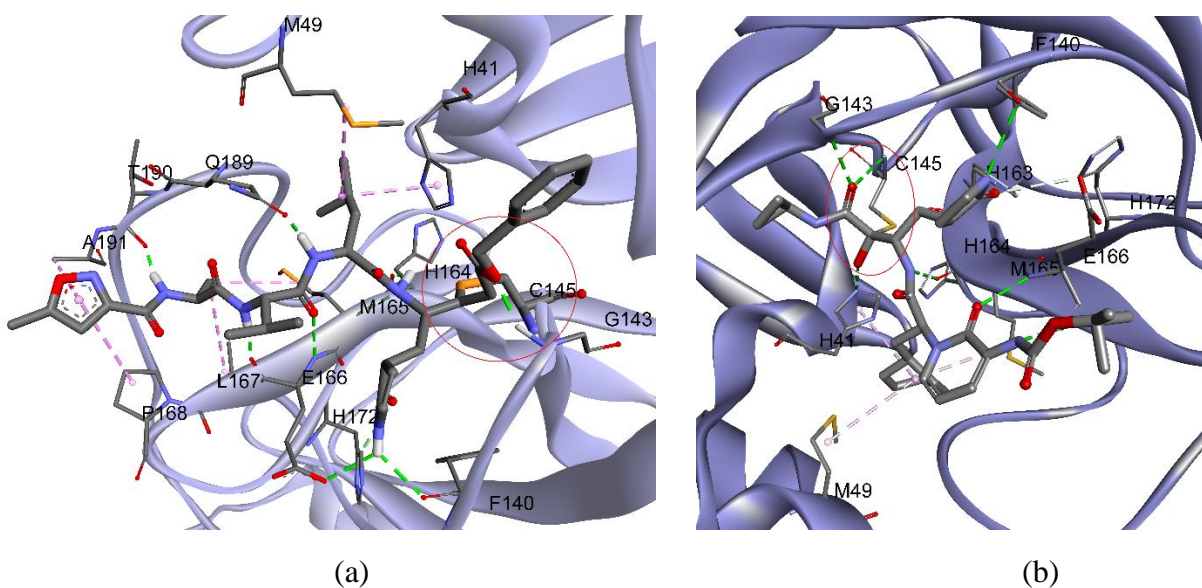


Figure S2: The peptide based ligand attached inside the active site of SARS-CoV-2 M^{pro} with a covalent bond by C145 through (a) conjugated addition to α,β -unsaturated ester and (b) nucleophilic addition to active ketone.

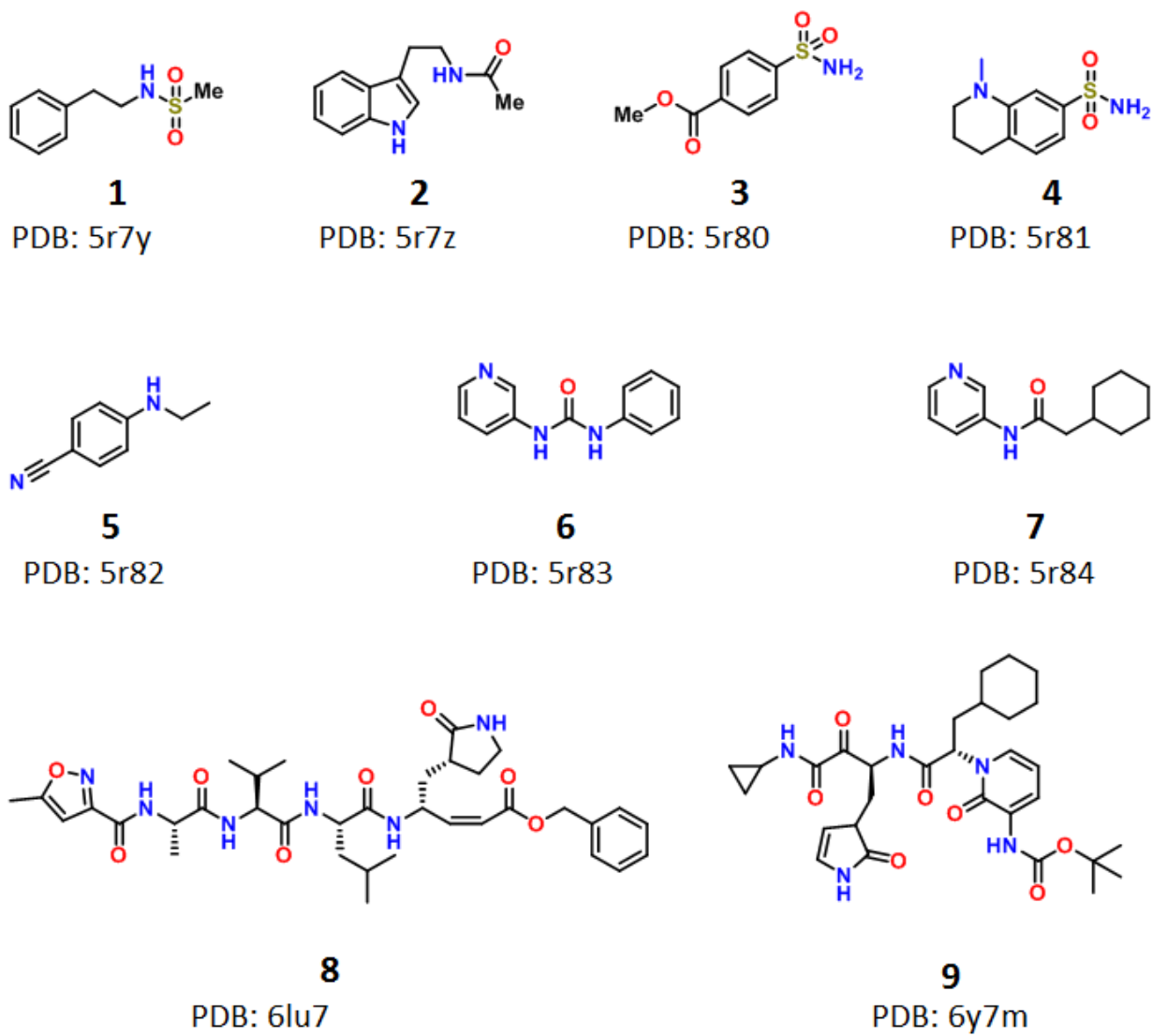


Figure S3: The small ligands found inside the crystal structure of SARS-CoV-2 M^{pro}.

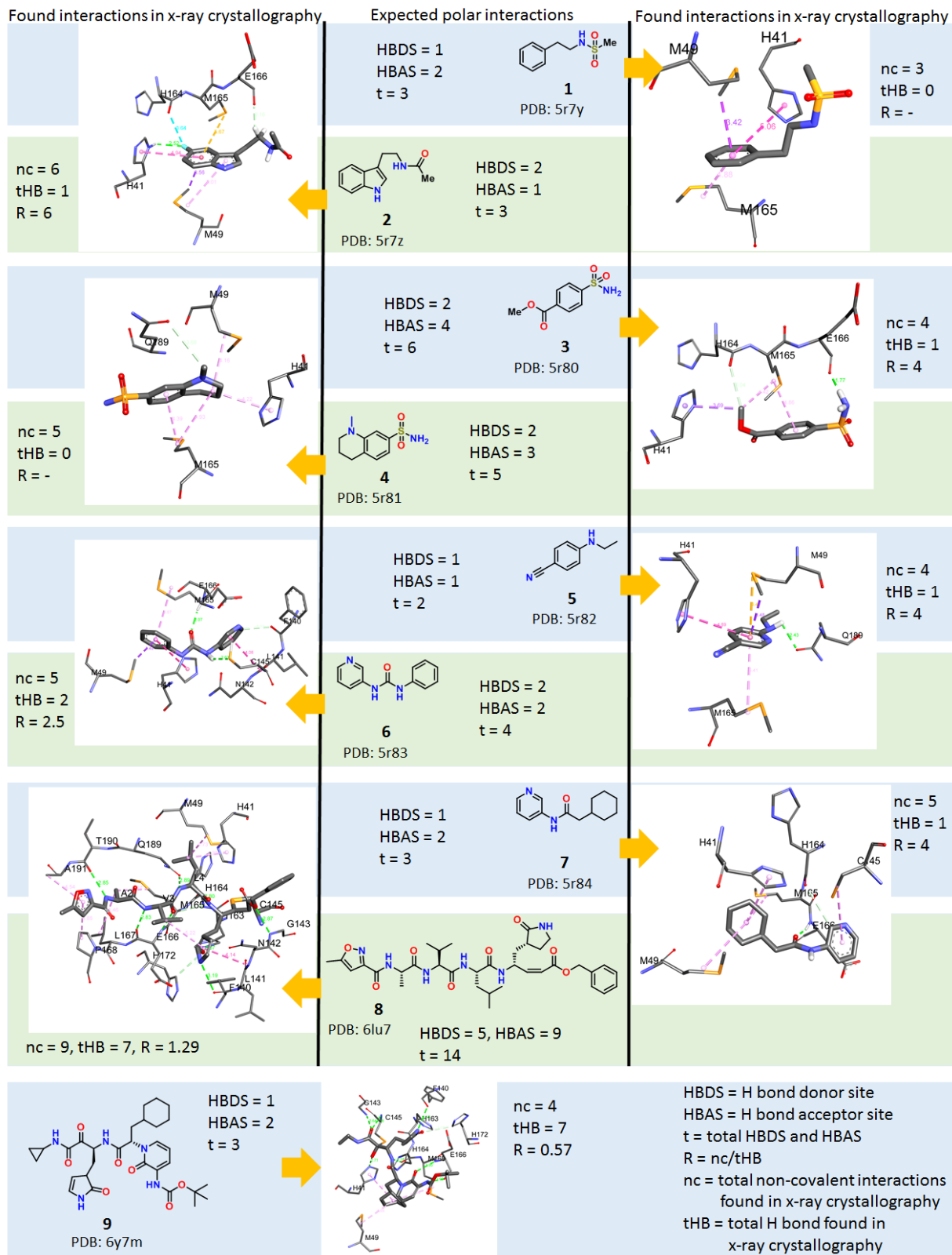


Figure S4: The small ligands-SARC-CoV-2 M^{Pro} enzyme interactions found inside the active site of the protein shows hydrophobic preference over hydrophilicity.

Alkaloids

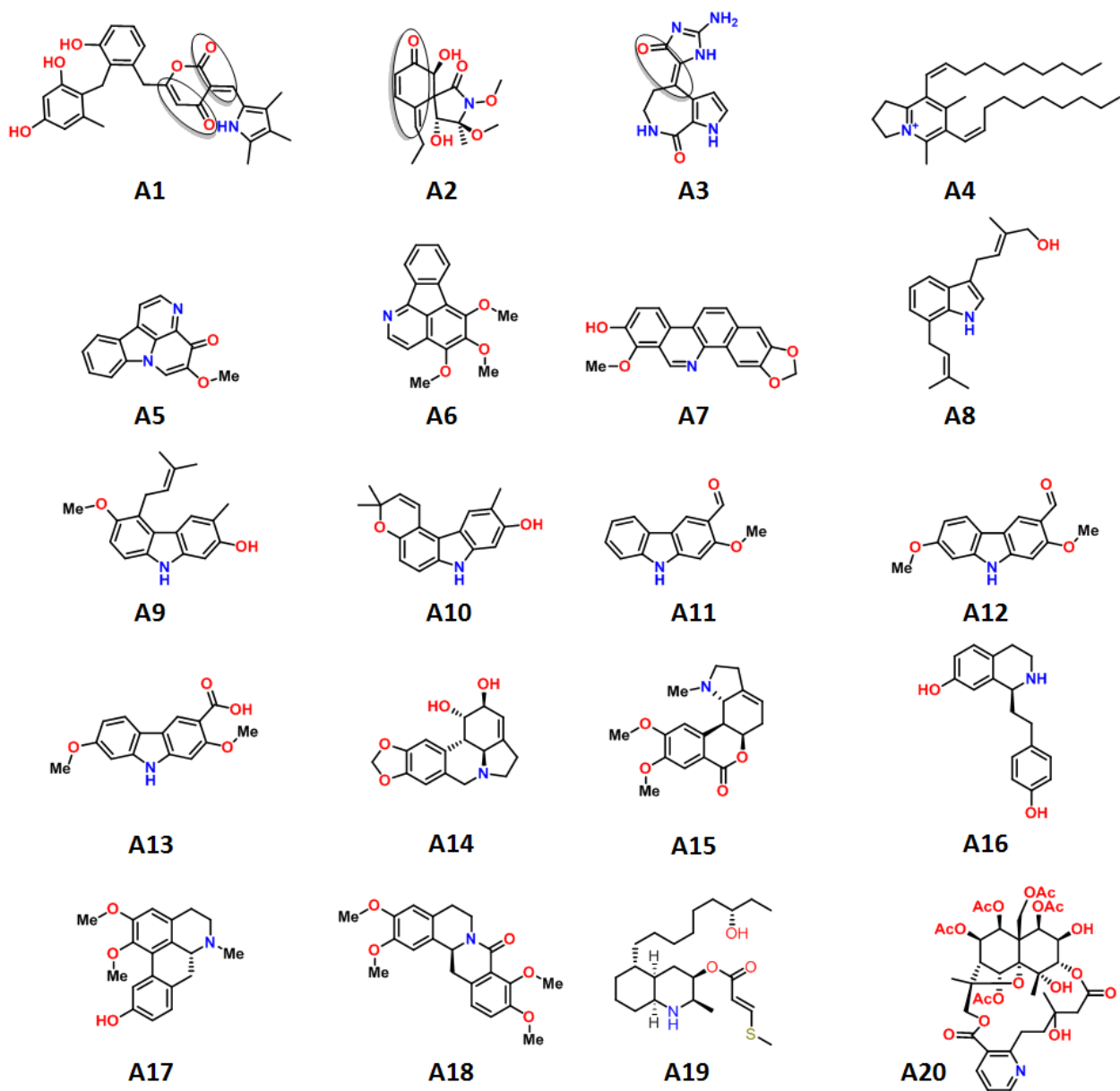


Figure S5: The structures of natural alkaloids **A1-A20**.

Terpenoids

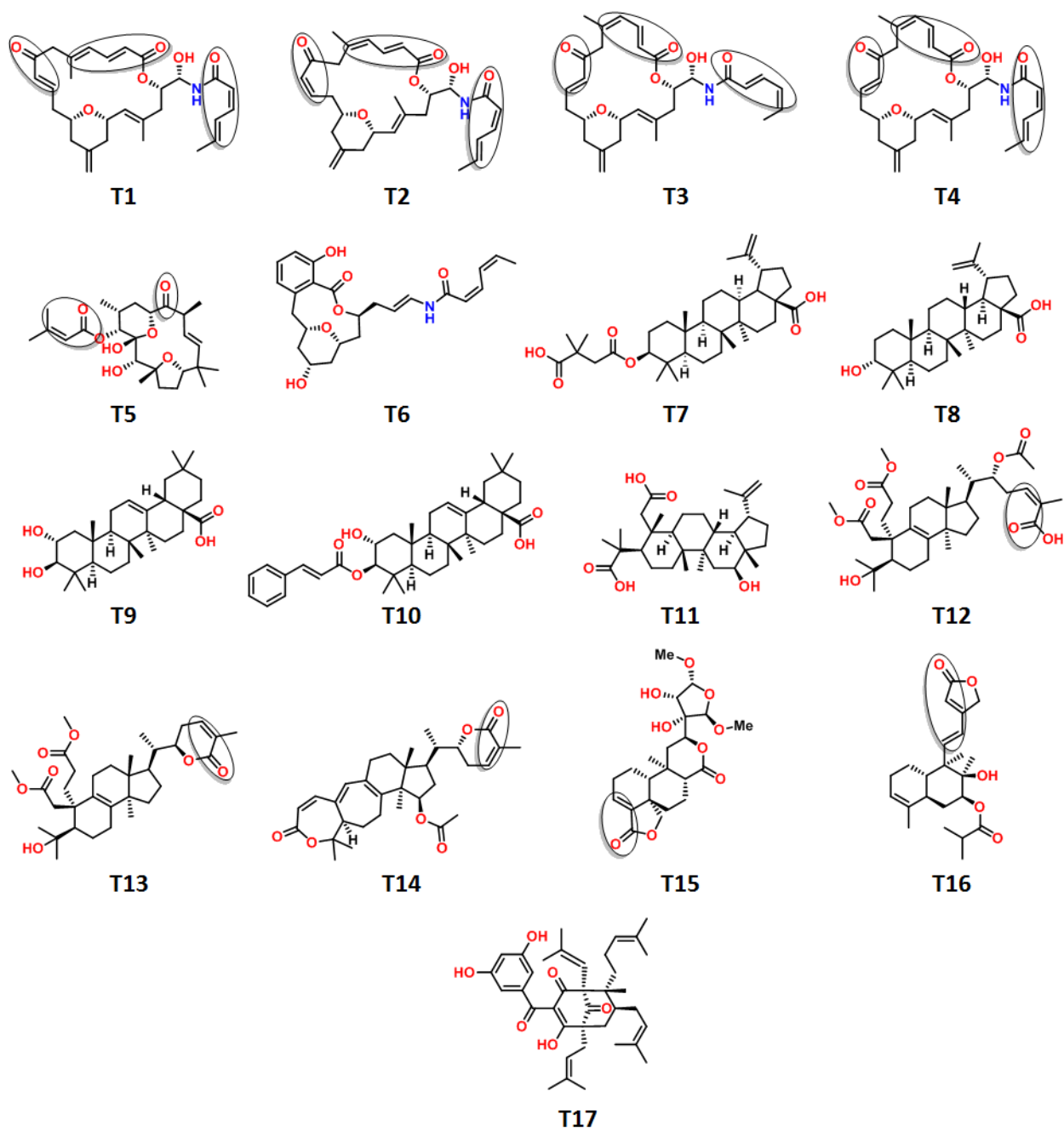
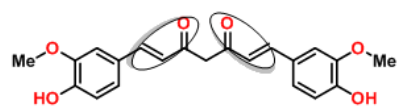
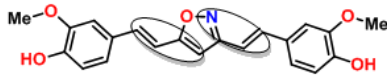


Figure S6: The structures of natural terpenoids **T1-T17**.

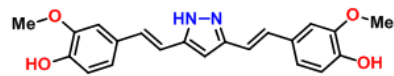
Polyphenols



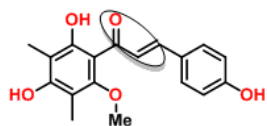
P1



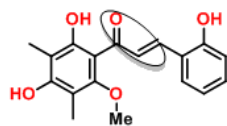
P2



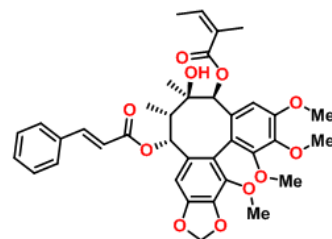
P3



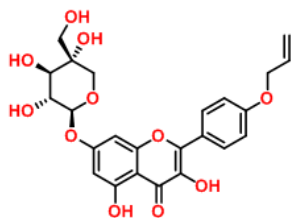
P4



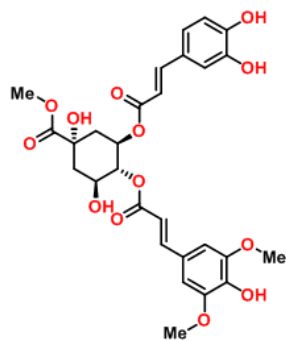
P5



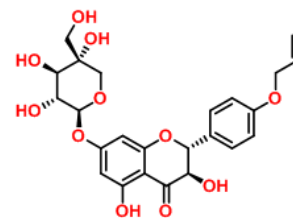
P6



P7



P8



P9

Figure S7: The structures of polyphenolic NPs **P1-P9**.

Peptides

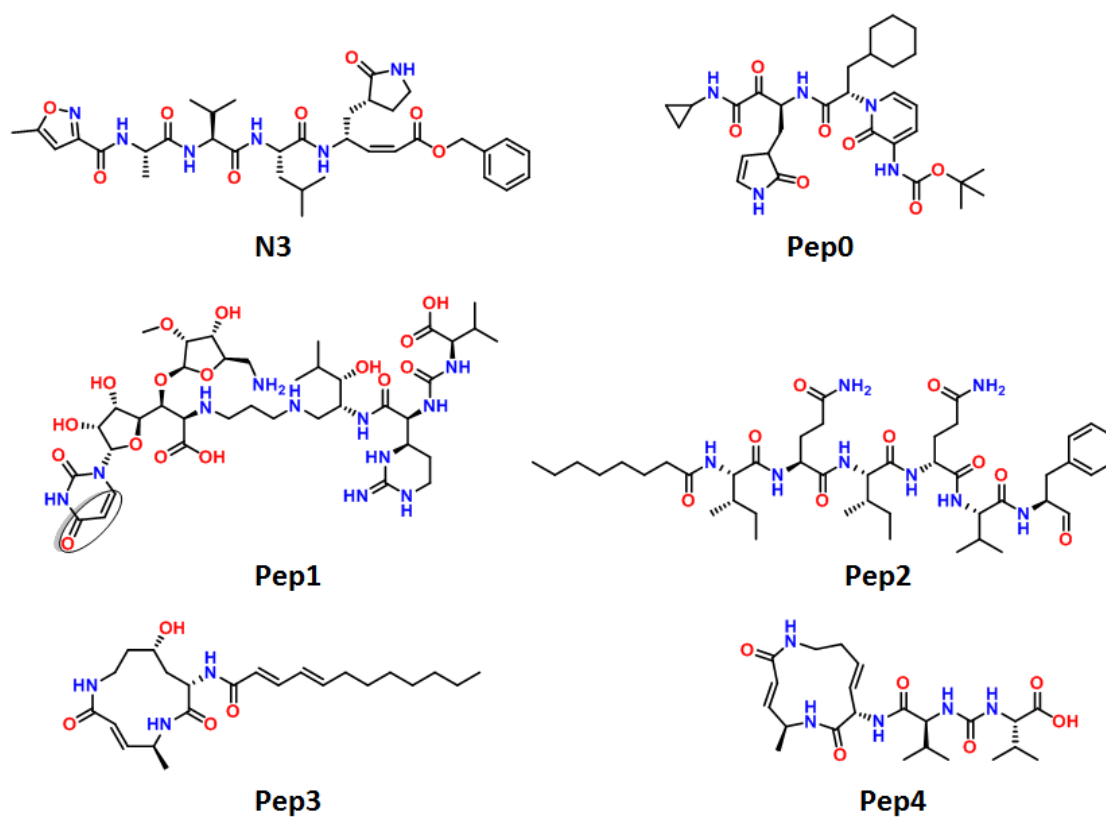


Figure S8: The structures of peptide based NPs (**Pep1-Pep4**) and small ligands crystallized with the SARC-CoV-2 M^{pro} enzyme.

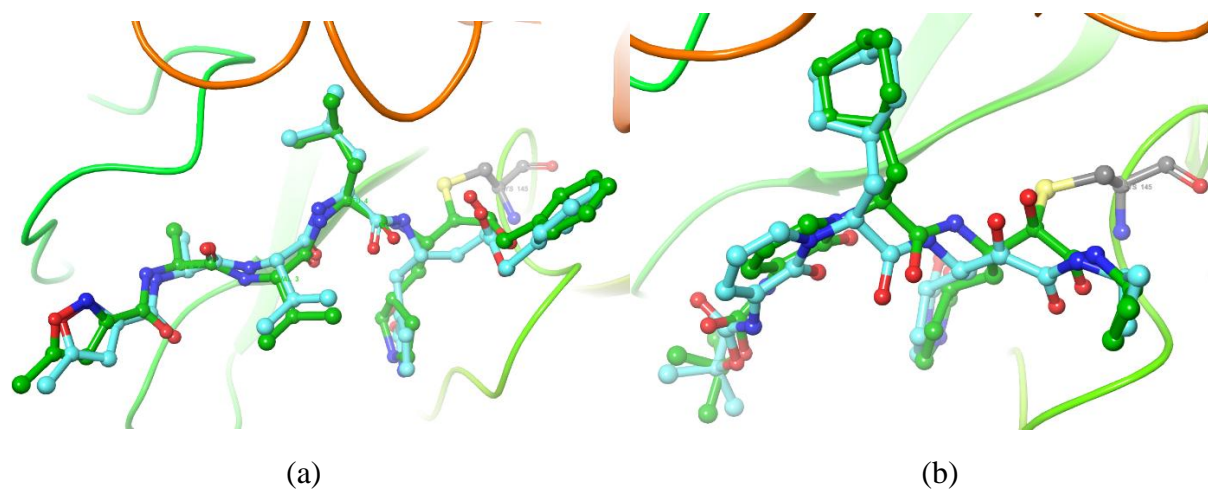
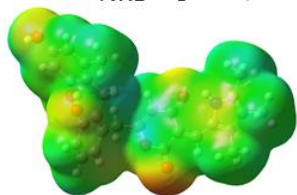
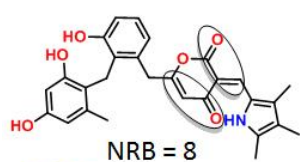


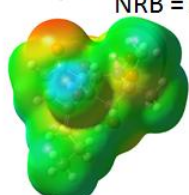
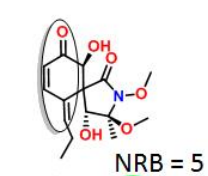
Figure S9: Overly of crystal structure and docked structure of SARC-CoV-2 M^{pro} enzyme with (a) **N3** and (b) **Pep0** ligand

(a) Alkaloids

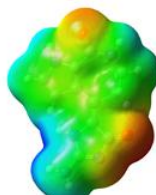
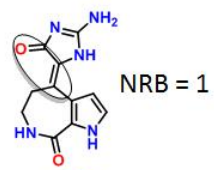
Number of rotatable bond = NRB



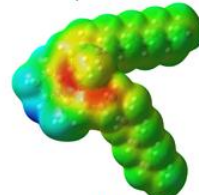
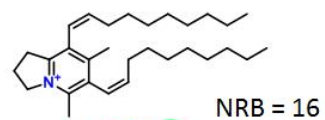
A1



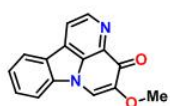
A2



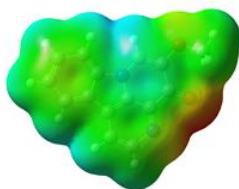
A3



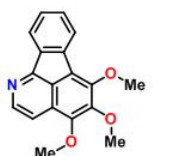
A4



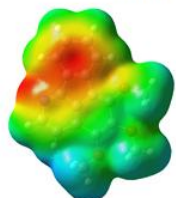
NRB = 1



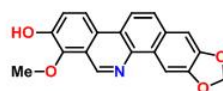
A5



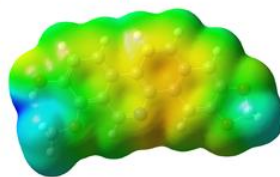
NRB = 3



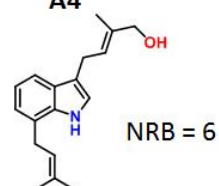
A6



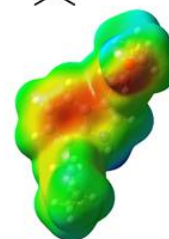
NRB = 1



A7

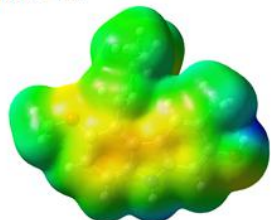
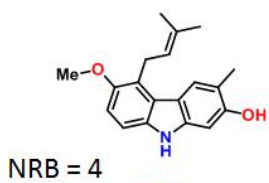


NRB = 6

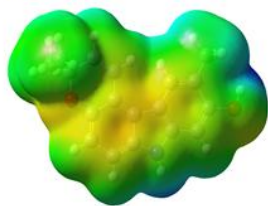
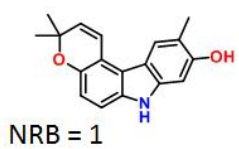


A8

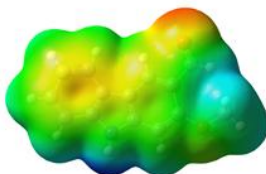
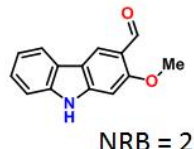
Continued...



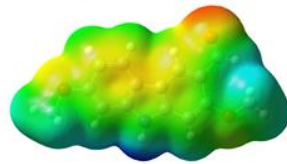
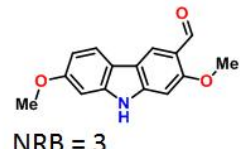
A9



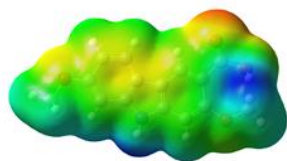
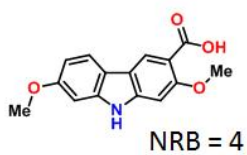
A10



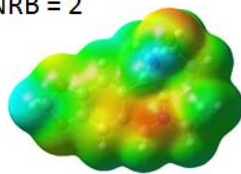
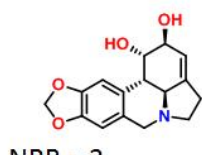
A11



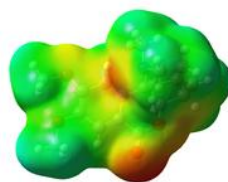
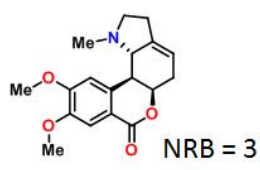
A12



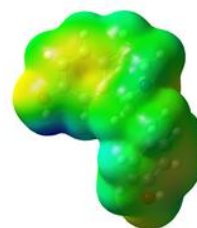
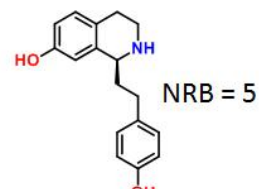
A13



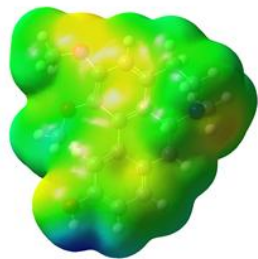
A14



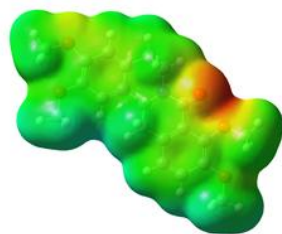
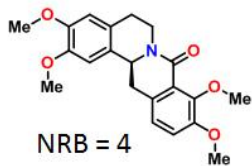
A15



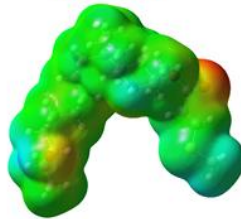
A16



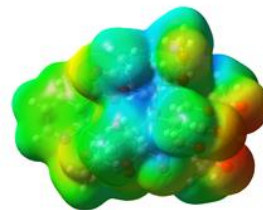
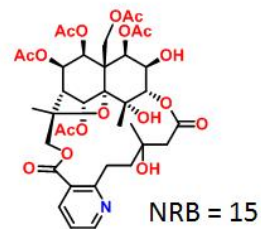
A17



A18



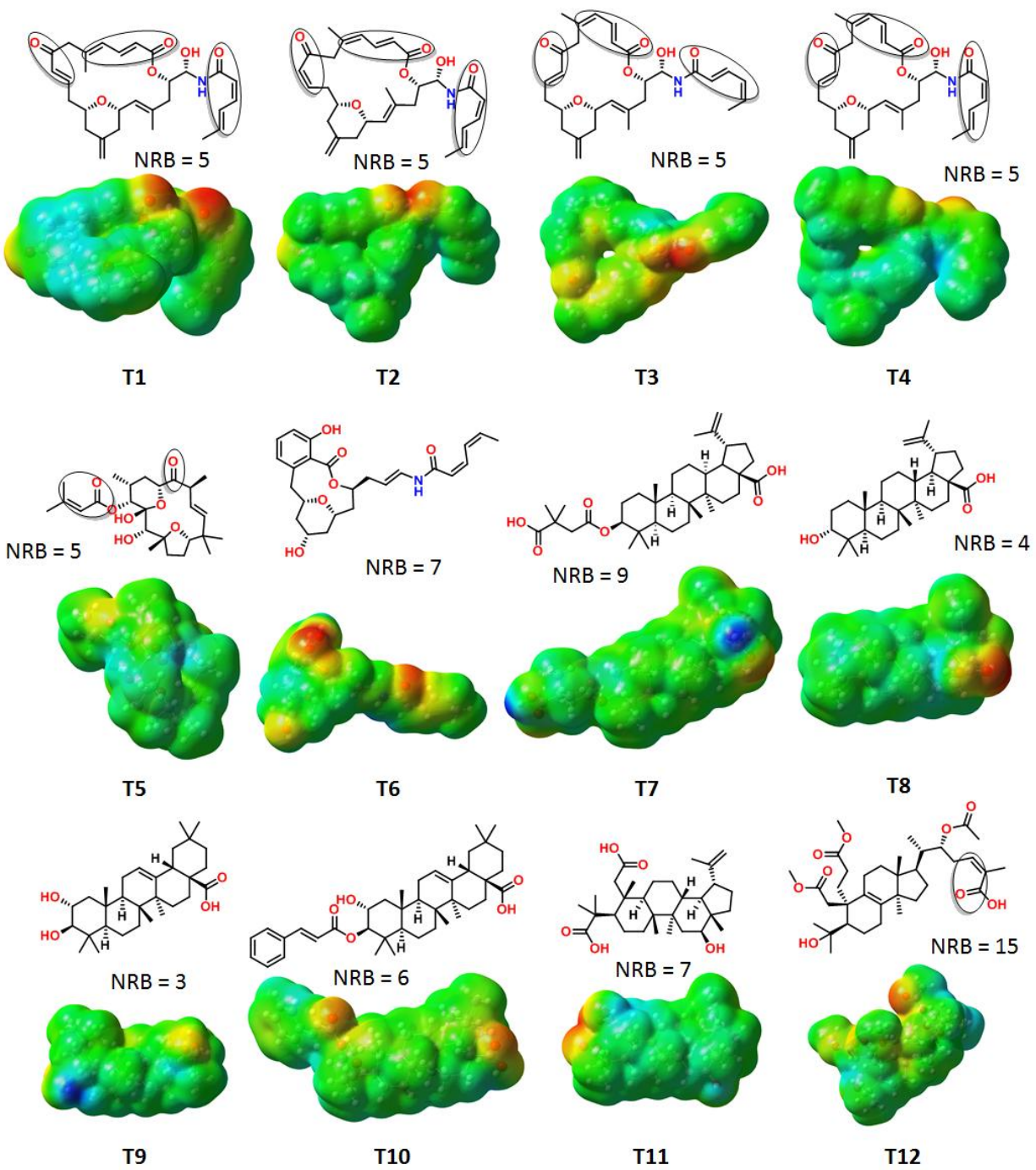
A19



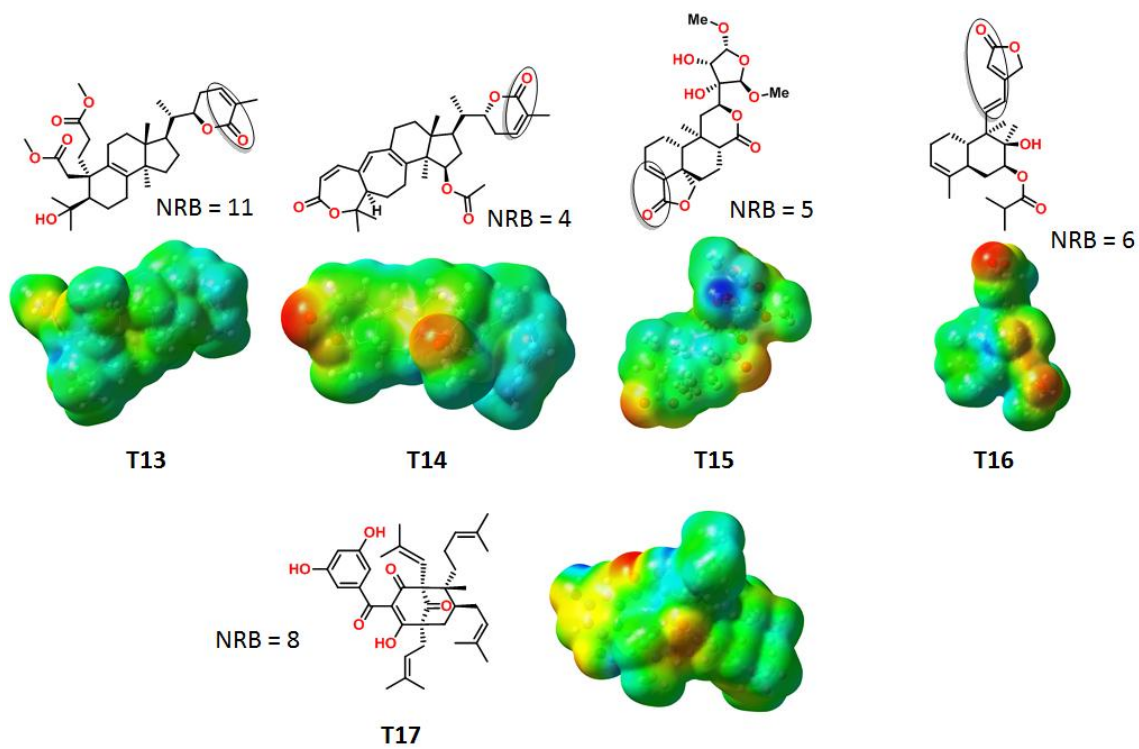
A20

Continued...

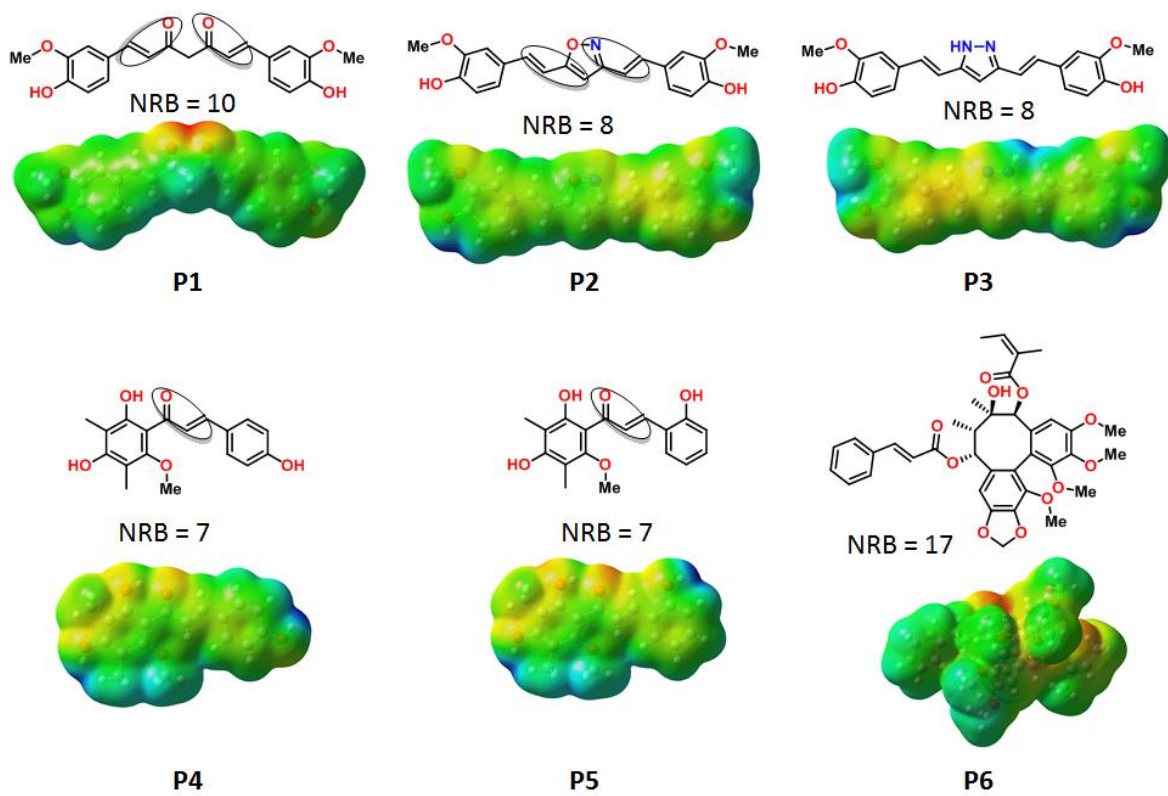
(b) Terpenoids



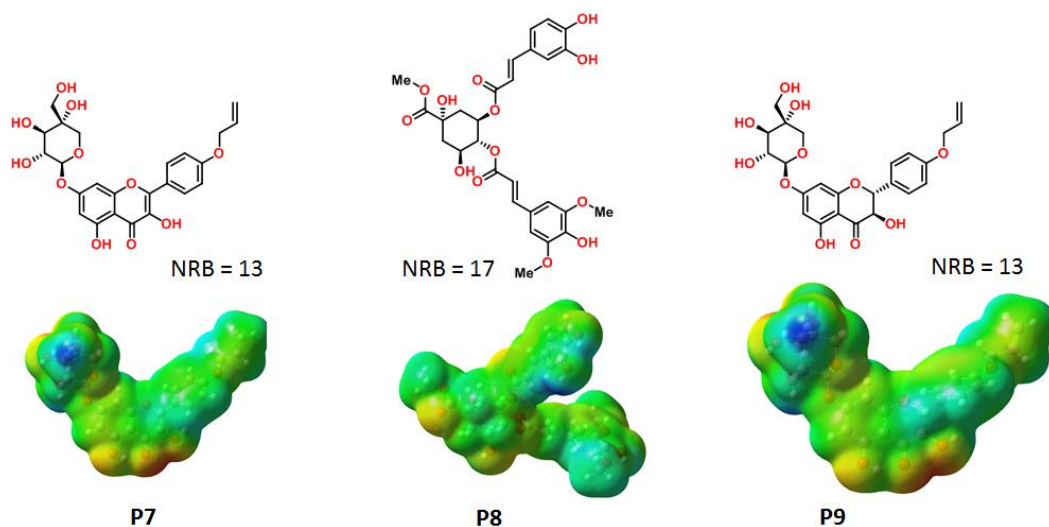
Continued...



(c) Polyphenolic



Continued...



(d) Peptides

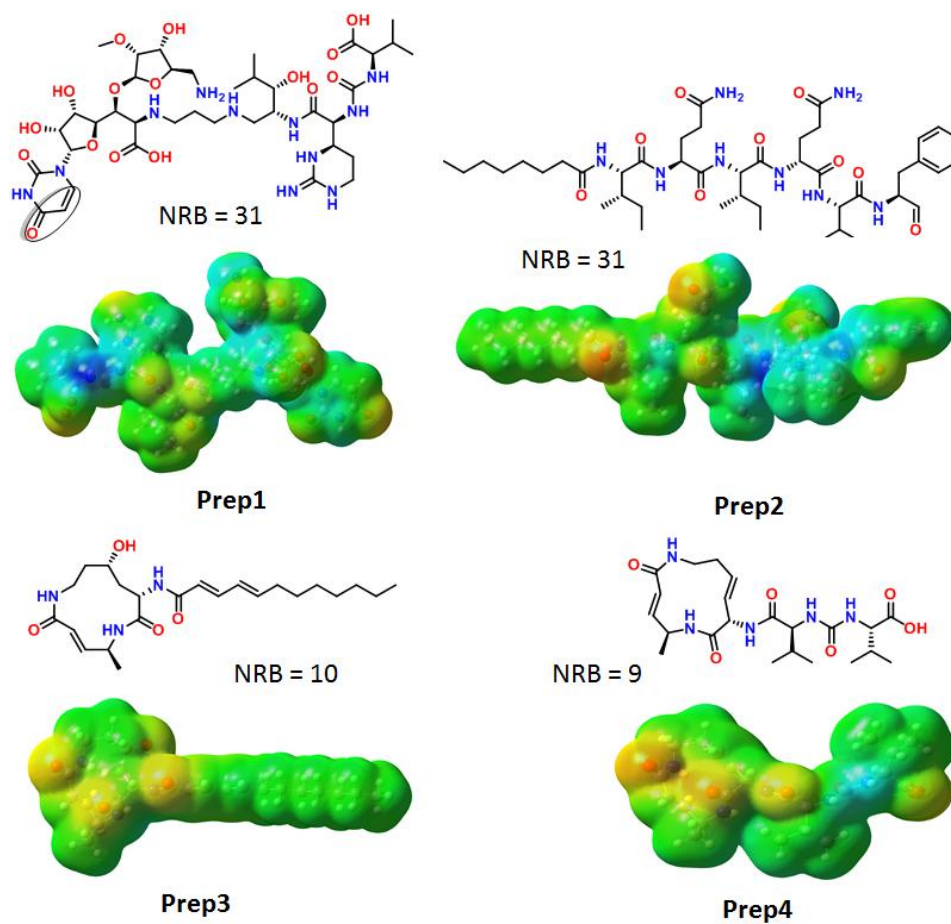


Figure S10: Molecular electrostatic potential surface and number of rotatable bonds of (a) Alkaloids, (b) Terpenoids, (c) Polyphenols and (d) Peptides.

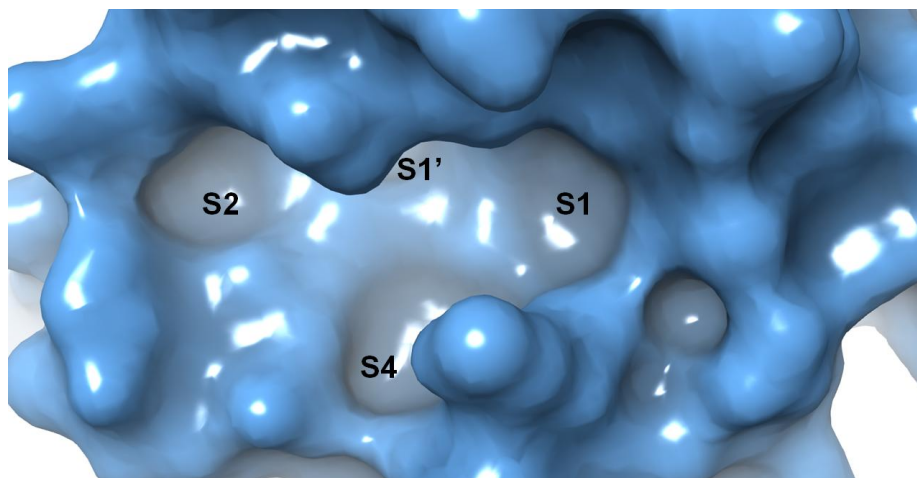
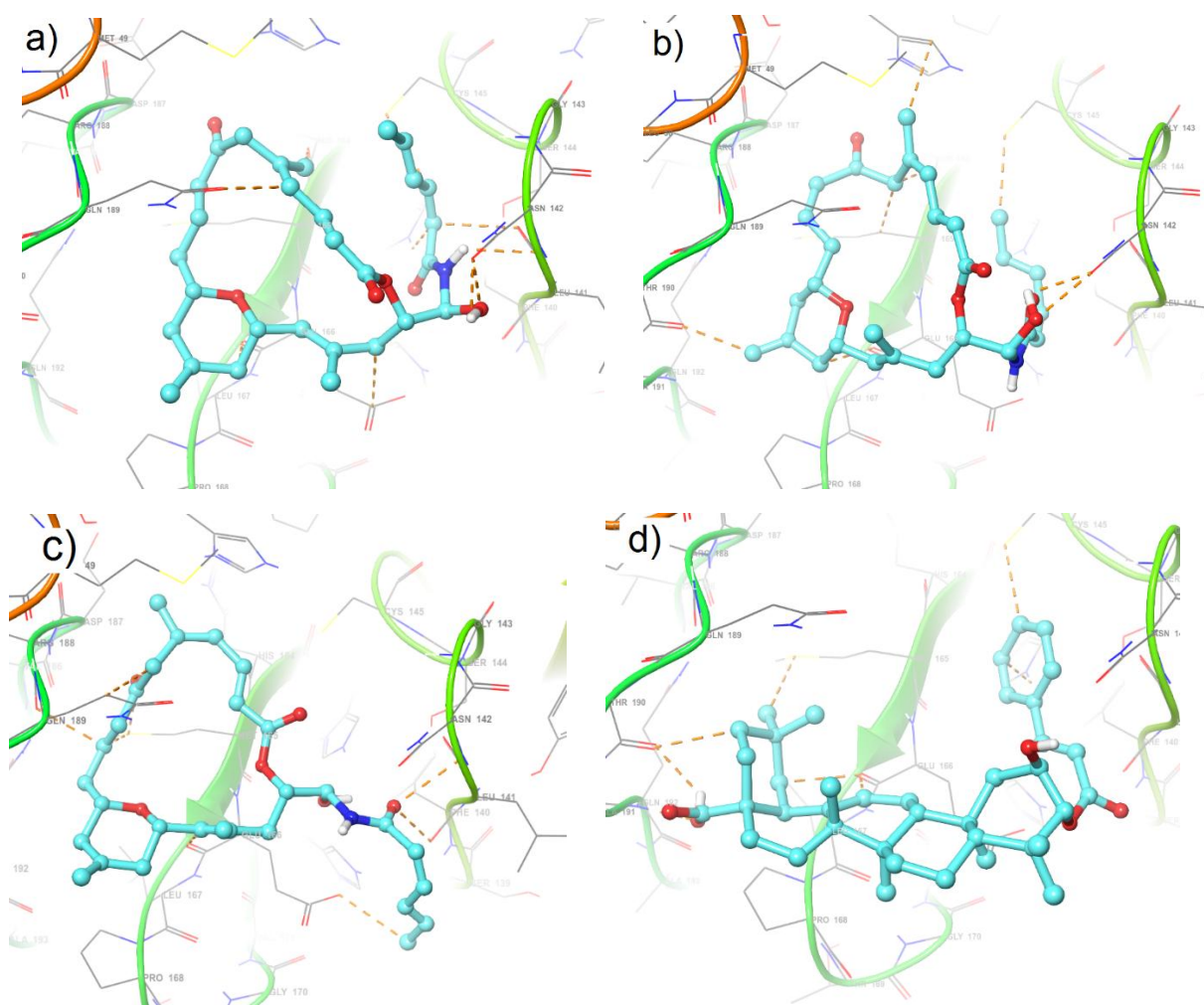


Figure S11: The four sites inside the active site of SARS-CoV-2 M^{Pro} enzyme



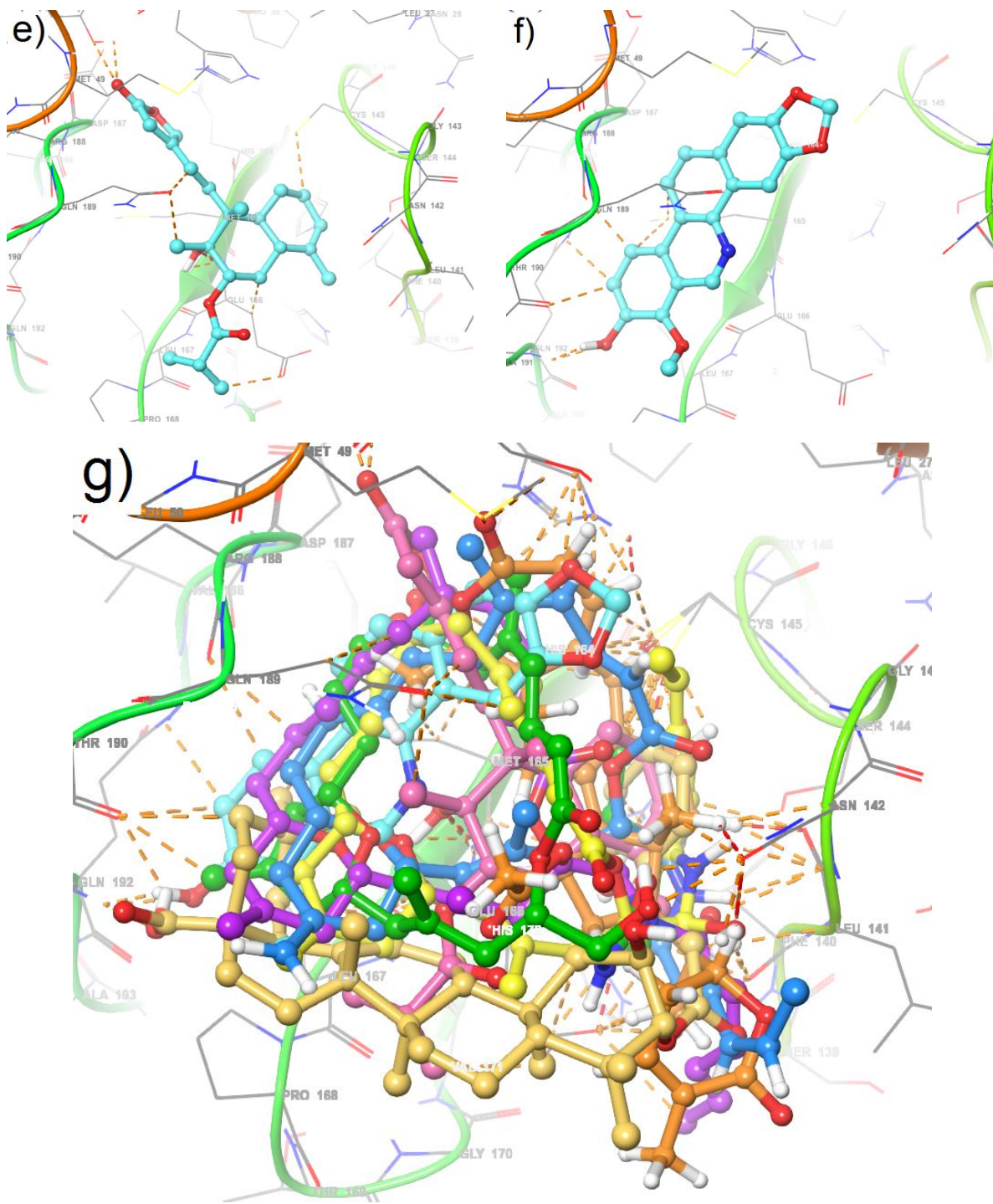
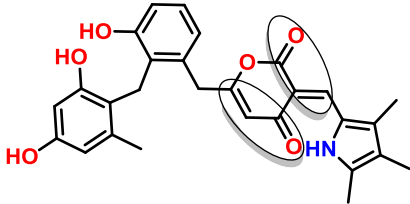
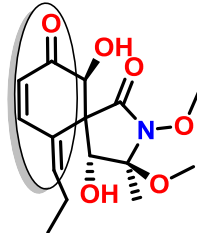
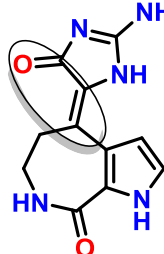
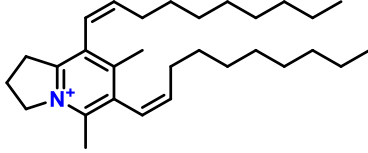
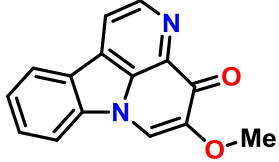
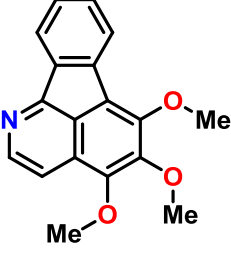
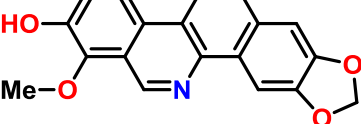
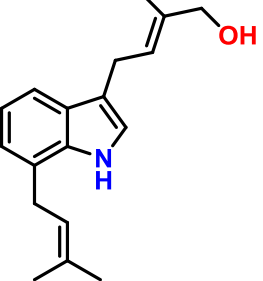
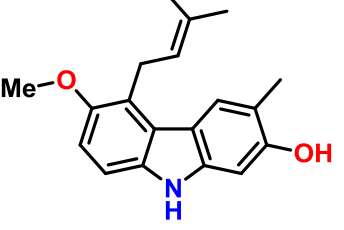
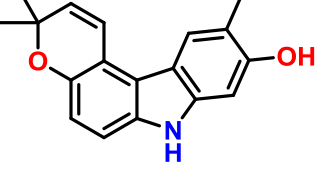
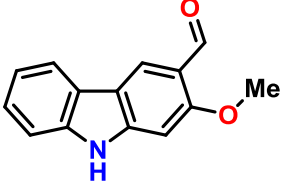
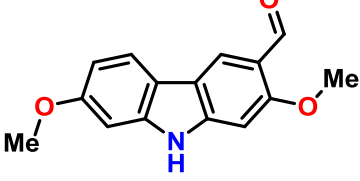
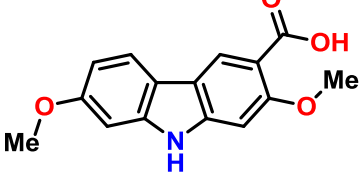
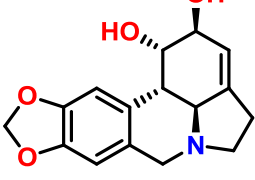
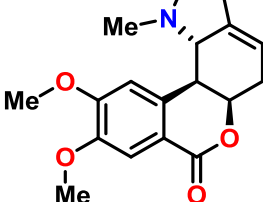


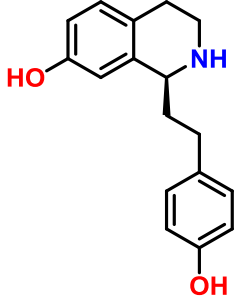
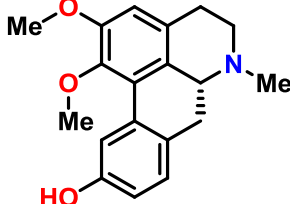
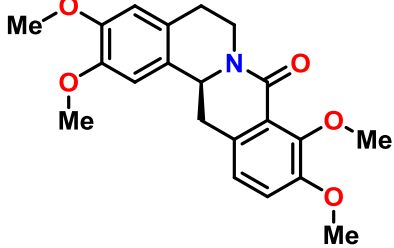
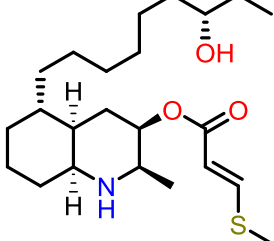
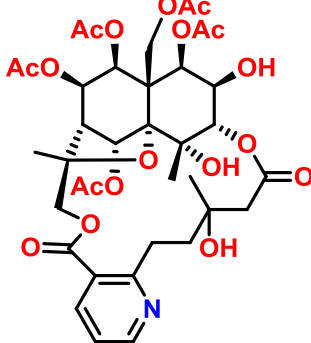
Figure S12: Docking poses and the binding interactions at the active site of SARC-CoV-2 M^{PRO} enzyme of (a) T1, (b)T2, (c)T4, (d)T10, (e)T14 as well as (f)T16 and overlying docked structures of (g) T3, T14, T1, T2, T4, T10 and T16 at the active site of protein. The pictures also show that His41 and C145 interact the NPs.

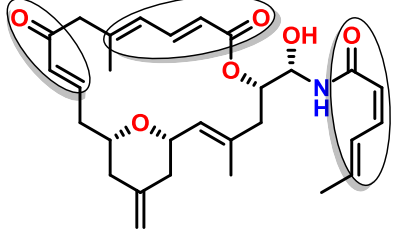
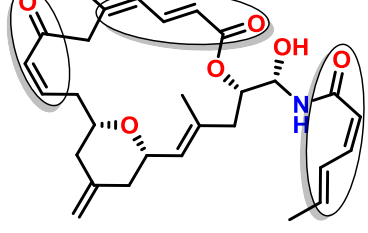
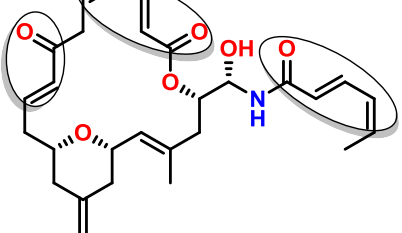
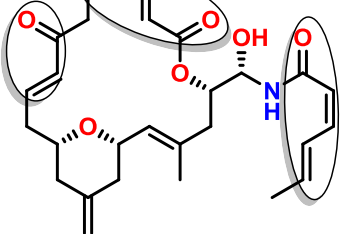
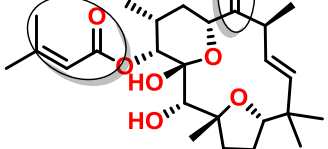
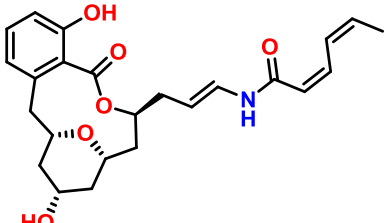
Table S1: Chemical structures, source and reported activities of the NPs

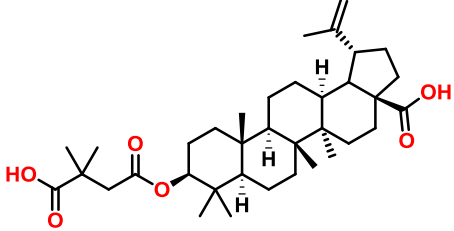
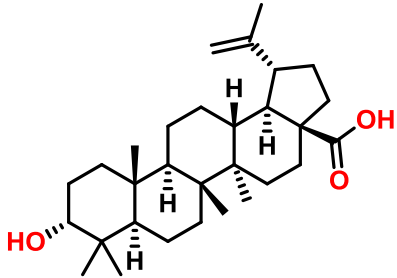
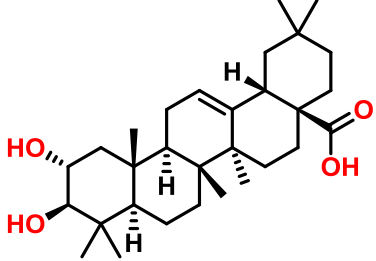
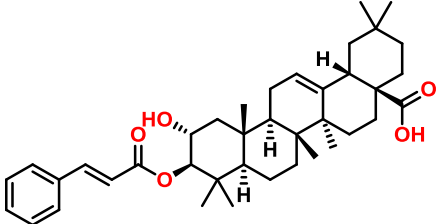
Compound	Source and Activity	Reference
<p>Alkaloid</p> <p>A1</p> 	<p>Marin microb: mangrove-derived <i>Streptomyces albus</i>.</p>	<p><i>Nat. Prod. Rep.</i> 2020,37, 175-223</p>
<p>A2</p> 	<p>Fungi: <i>Cochliobolus lunatus</i> anti-IAV activity.</p>	<p><i>Nat. Prod. Rep.</i>, 2020,37, 175-223</p>
<p>A3</p> 	<p>Sponge: <i>Stylissa carteri</i> anti HIV1 activity.</p>	<p><i>Nat. Prod. Rep.</i>, 2020,37, 175-223</p>
<p>A4</p> 	<p>Anibamine, a novel pyridine quaternary alkaloid, Anibamine competed for the binding of I-gp120 to human CCR5 with an IC50 of 1 μM.</p>	<p><i>Nat. Prod. Rep.</i>, 2010, 27, 1781–1800</p>
<p>A5</p> 	<p>Anewcanthin-4-one type alkaloid, drymaritin isolated from <i>Drymaria diandra</i> (also known as <i>Drymaria cordata</i> subsp. <i>diandra</i>) exhibited anti-HIV activity in H9 lymphocytes with an EC50 value of 0.699 μg/mL and a TI of 20.6</p>	<p><i>Nat. Prod. Rep.</i>, 2010, 27, 1781–1800</p>

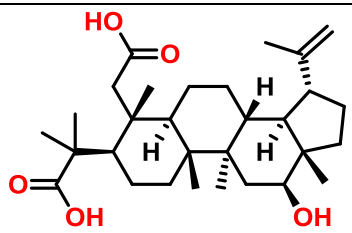
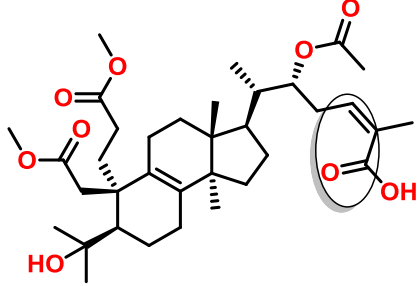
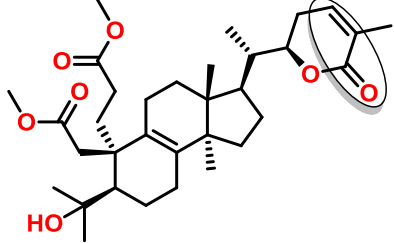
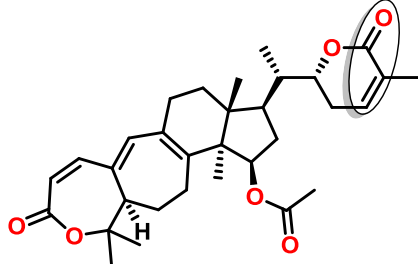
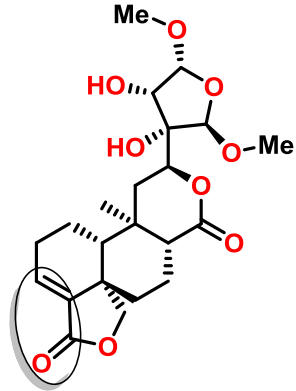
A6		<p>Norruffscine Plant: <i>Pericampylus glaucus</i>, a climbing shrub widely distributed in the southwest of China. Compound showed EC50 values of 10.9 and SI values of 45.7 against HIV-1 in C8166 cells.</p>	<p><i>Nat. Prod. Rep.</i>, 2010, 27, 1781–1800</p>
A7		<p>Decarine Plant: root bark of <i>Zanthoxylum ailanthoides</i> potent anti-HIV, EC50 values of <0.1 mg/mL and TI values of >226.</p>	<p><i>Nat. Prod. Rep.</i>, 2010, 27, 1781–1800</p>
A8		<p>(E)-3-(3-hydroxymethyl-2-butenyl)-7-(3-methyl-2-butenyl)-1H-indole was isolated from the twigs and leaves of <i>Glycosmis montana</i> native to China. Potent anti-HIV activity with an IC50 value of 1.17 mg/mL and an SI of 11.68</p>	<p><i>Nat. Prod. Rep.</i>, 2010, 27, 1781–1800</p>
A9		<p>Glybomine was isolated from the twigs and leaves of <i>Glycosmis montana</i>. Earlier these alkaloids were reported from <i>G. arborea</i>. <i>in vitro</i> inhibitory activity against HIV replication in C8166 cells. Anti-HIV activity with an IC50 of 9.73 μ/mL.</p>	<p><i>Nat. Prod. Rep.</i>, 2010, 27, 1781–1800</p>
A10		<p>Glycoborinine (7), was isolated from the twigs and leaves of <i>Glycosmis montana</i>. Earlier these alkaloids were reported from <i>G. arborea</i>. <i>in vitro</i> inhibitory activity against HIV replication in C8166 cells. Anti-HIV activity with an IC50 of 4.47 μg/mL.</p>	<p><i>Nat. Prod. Rep.</i>, 2010, 27, 1781–1800</p>

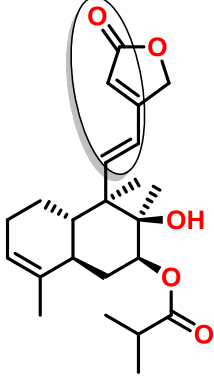
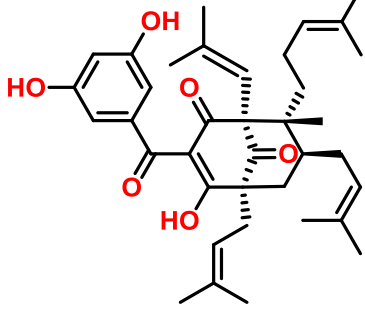
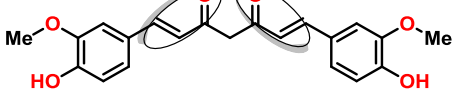
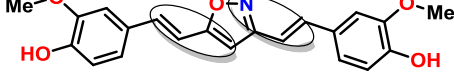
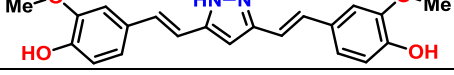
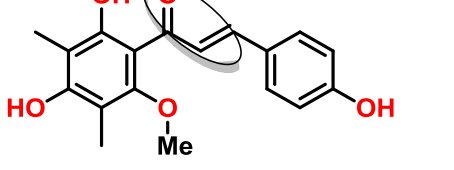
A11		<p>O-methylmukonal isolated from the rhizomes and roots of the plant: <i>Clausena excavata</i>. Compound showed anti-HIV-1 activity in a syncytial assay, with EC₅₀ values of 12.0 μM and TI of 56.7</p>	<p><i>Nat. Prod. Rep.</i>, 2010, 27, 1781–1800</p>
A12		<p>3-formyl-2,7-dimethoxycarbazole isolated from the rhizomes and roots of the plant: <i>Clausena excavata</i>. Compound showed anti-HIV-1 activity in a syncytial assay, with EC₅₀ values of 29.1 μM and TI 8.0</p>	<p><i>Nat. Prod. Rep.</i>, 2010, 27, 1781–1800</p>
A13		<p>clauszolineJ isolated from the rhizomes and roots of plant: <i>Clausena excavata</i>. Compounds showed anti-HIV-1 activity in a syncytial assay, with EC₅₀ values of 34.2 μM and TI of 1.6</p>	<p><i>Nat. Prod. Rep.</i>, 2010, 27, 1781–1800</p>
A14		<p>Amaryllidaceae alkaloids (lycorine) isolated from the bulbs of the plant: <i>Leucojum vernum</i>, possess anti-HIV-1 activity in MT4 cells with ID₅₀ values of 0.4 μg/mL.</p>	<p><i>Nat. Prod. Rep.</i>, 2010, 27, 1781–1800</p>
A15		<p>Amaryllidaceae alkaloids homolycorine isolated from the bulbs of the plant: <i>Leucojum vernum</i> possess anti-HIV-1 activity in MT4 cells with ID₅₀ values 7.3 μg/mL.</p>	<p><i>Nat. Prod. Rep.</i>, 2010, 27, 1781–1800</p>

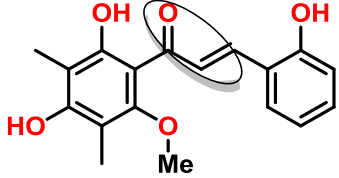
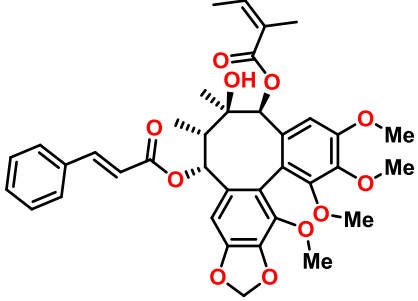
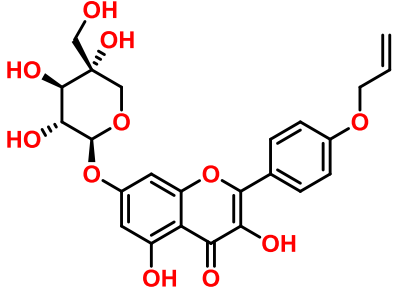
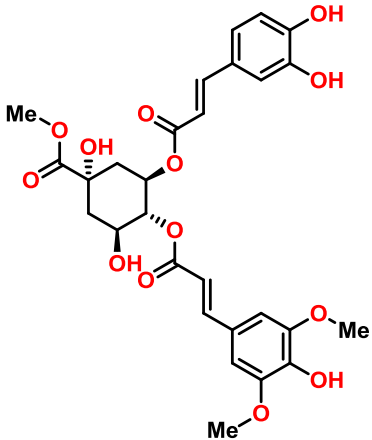
A16		<p>Alkaloids, (+)-1-(R) coclaurine Plant: leaves of <i>Nelumbo nucifera</i>, a perennial aquatic crop grown throughout Asia, showed anti-HIV activity with EC50 values of 0.8 µg/mL and TI values of >125.</p>	<p><i>Nat. Prod. Rep.</i>, 2010, 27, 1781– 1800</p>
A17		<p>Nuciferine, isolated from the plant: <i>Nymphaea caerulea</i> anti-HIV activity EC50 of 0.8 µg/mL and TI of 36.</p>	<p><i>Nat. Prod. Rep.</i>, 2010, 27, 1781– 1800</p>
A18		<p>(-)-8-oxotetrahydropalmatine were isolated from plant : <i>Pericampylus glaucus</i>, a climbing shrub widely distributed in the south west of China. Compound showed EC50 values of 14.1 µM and SI values of 18.8 against HIV-1 in C8166 cells.</p>	<p><i>Nat. Prod. Rep.</i>, 2010, 27, 1781– 1800</p>
A19		<p>Isolated from a marine invertebrate animal <i>Didemnum</i> <i>sp.</i> (a tunicate).</p>	<p><i>Nat. Prod. Rep.</i> 2020, 37, 175-223</p>
A20		<p>tripfordine A Plant: roots of poisonous liana <i>Tripterygium wilfordii</i>. The highly toxic roots are known as ‘lei gong teng’ in Chinese folklore and used as an Insecticide. anti-HIV activity, with an EC50 of 13.4 µg/mL and TI of 1.9</p>	<p><i>Nat. Prod. Rep.</i>, 2010, 27, 1781– 1800</p>
<p>Terpenes</p>			

T1		<p>Isolated from the marine Sponge <i>Cacospongia mycofijiensis</i> Exhibit nanomolar cytotoxicity toward the HL-60 cell line, are antimitotic, and induce in vitro tubulin polymerization.</p>	<p><i>J. Nat. Prod.</i> 2018, <i>81</i>, 2539-2544</p>
T2		<p>Isolated from the marine Sponge <i>Cacospongia mycofijiensis</i> Exhibit nanomolar cytotoxicity toward the HL-60 cell line, are antimitotic, and induce in vitro tubulin polymerization.</p>	<p><i>J. Nat. Prod.</i> 2018, <i>81</i>, 2539-2544</p>
T3		<p>Isolated from the marine Sponge <i>Cacospongia mycofijiensis</i> Exhibit nanomolar cytotoxicity toward the HL-60 cell line, are antimitotic, and induce in vitro tubulin polymerization.</p>	<p><i>J. Nat. Prod.</i> 2018, <i>81</i>, 2539-2544</p>
T4		<p>Isolated from the marine Sponge <i>Cacospongia mycofijiensis</i>. Exhibit nanomolar cytotoxicity toward the HL-60 cell line, are antimitotic, and induce in vitro tubulin polymerization.</p>	<p><i>J. Nat. Prod.</i> 2018, <i>81</i>, 2539-2544</p>
T5		<p>Microb: bacteria <i>Excoecaria venenta</i> The anti-viral activity of these compounds against the chikungunya virus.</p>	<p><i>Nat. Prod. Rep.</i>, 2016, <i>33</i>, 1227</p>
T6		<p>Apicularen (22) anti-HIV and anti-HCV hits, inhibitor of HPV replication (SI= 3–6 in Hela cells).</p>	<p><i>Nat. Prod. Rep.</i>, 2015, <i>32</i>, 29-48</p>

T7		<p>Plant: <i>Platanus acerifolia</i> (London plane tree) Usable Part: bark Induces apoptosis through the mitochondrial pathway. A typical decrease in bcl-2 and cyclin D1 gene expression and increase in bax gene expression was observed in several cancer cell lines; inhibits NF-kB expression in androgen refractory, human prostate cancer cells (PC-3).</p>	<p><i>Nat. Prod. Rep.</i>, 2019, 36, 1654-1686</p>
T8		<p>Plant: <i>Platanus acerifolia</i> (London plane tree) Usable Part: bark Activity induces apoptosis through the mitochondrial pathway. A typical decrease in bcl-2 and cyclin D1 gene expression and increase in bax gene expression was observed in several cancer cell lines; inhibits NF-kB expression in androgen refractory, human prostate cancer cells (PC-3).</p>	<p><i>Nat. Prod. Rep.</i>, 2016, 33, 1227</p>
T9		<p>Maslinic acid an oleanane triterpene identified from <i>Crataegus oxyacantha</i>, apple peel and was also isolated quantitatively from olive fruit and show effect only on the Caco-2 cell line (EC50: 15.4 mM), anti-proliferative effect against the HT-29 cell line (EC50: 101.2 mM) without any necrotic effects.</p>	<p><i>Nat. Prod. Rep.</i>, 2009, 26, 1321–1344</p>
T10		<p>Isolated from <i>Crataegus oxyacantha</i>, showed minor antiproliferative effects against HepG2 (EC50: 17.9–20.6 mM), MCF-7 (EC50:</p>	<p><i>Nat. Prod. Rep.</i>, 2009, 26, 1321–1344</p>

		20.9–29.2 mM), and Caco-2 (EC50: 8.9–14.2 mM),	
T11		<i>Stauntonia obovatifoliola</i> anti-HIV-1 protease activity (EC50: 17.8 mM).	<i>Nat. Prod. Rep.</i> , 2019 , 36, 1654-1686
T12		Fungi: Mushroom <i>Ganoderma colossum</i> . anti-protease activity (tested against HIV-1) with EC50 values of 14.6.	<i>Nat. Prod. Rep.</i> , 2019 , 36, 1654-1686
T13		Fungi: Mushroom <i>Ganoderma colossum</i> . anti-protease activity (tested against HIV-1) with EC50 values of 24.7	<i>Nat. Prod. Rep.</i> , 2019 , 36, 1654-1686
T14		Fungi: Mushroom <i>Ganoderma colossum</i> . anti-HIV-1 protease inhibition activity. with EC50 values of 15.3	<i>Nat. Prod. Rep.</i> , 2019 , 36, 1654-1686
T15		Isolated from the plant <i>Dichrocephala benthamii</i> (Asteraceae) possess anti-HIV integrase activity.	<i>Nat. Prod. Rep.</i> , 2019 , 36, 1654-1686

T16		<p>Known as micafungin or FK463,41 is now marketed as the intravenous antifungal drug Mycamine essential component of fungal cell walls.</p>	<p><i>Nat. Prod. Rep.</i>, 2019,36, 1654-1686</p>
T17		<p>Plant: <i>Symphonia globulifera</i> purified from other species, belonging to the Clusiaceae family, such as <i>Garcinia livingstonei</i>, and <i>G. macrophylla</i> Usable Part: Roots antibacterial, anti-HIV, trypanocidal, anti-malarial, anti-cancer, and leishmanicidal activities, anti-parasitic activity. IC50 value 2.95 μM.</p>	<p><i>Nat. Prod. Rep.</i>, 2016,33, 372-381</p>
Phenolic Compounds			
P1		<p>Plant: <i>Curcuma longa</i>. HIV1 protease.</p>	<p><i>Nat. Prod. Rep.</i>, 2011, 28, 1937</p>
P2		<p>Synthesised derivatives of P1</p>	
P3		<p>Synthesised derivatives of P1</p>	
P4		<p>Plant: <i>Cleistocalyx operculatus</i> have neuraminidase (NA) inhibiting properties, with (E)-4,2,4- trihydroxy-6-methoxy-3,5-dimethylchalcone (IC50 3.31–20.45 μM) Strongest effects against NA from the novel influenza H1N1 .</p>	<p><i>Nat. Prod. Rep.</i>, 2012 29, 11</p>

P5		<p>Plant: <i>Cleistocalyx operculatus</i> with 2,2,4-trihydroxy-6-methoxy-3,5-dimethylchalcone (IC₅₀ 2.55–28.12 μM) strongest effects against NA from the oseltamivir resistant H1N1 (H274Y mutant).</p>	<p><i>Nat. Prod. Rep.</i>, 2012 29, 11</p>
P6		<p>Tiegusanin G (29) was one of 14 new lignans isolated from a 70% aqueous acetone extract of the aerial parts of plant <i>Schisandra propinqua</i> var. <i>sinensis</i>, popularly known as 'tie-gusan' in China. anti-HIV-1 activity in the C8166 cell line, with an EC₅₀ value of 7.9 mM and TI of > 25.</p>	<p><i>Nat. Prod. Rep.</i>, 2010, 27, 1781–1800</p>
P7		<p>6-(g,g-dimethylallyl)dihydroquercetin-7-O-beta-D-glucoside) Plant: <i>Ochna integerrima</i> (Methanolic extract) anti-HIV-1 activity with an EC₅₀ of 14.0 μg/mL.</p>	<p><i>Nat. Prod. Rep.</i>, 2010, 27, 1781–1800</p>
P8		<p>Methyl-5-O-caffeoyl-3-O-sinapoylquininate Plant: <i>Gardenia jasminoides</i> (isolated by bioactivity-guided fractionation of the ethyl acetate extract of this plant) HIV-1IN inhibitory activity with an IC₅₀ of 20 μg/mL.</p>	<p><i>Nat. Prod. Rep.</i>, 2010, 27, 1781–1800</p>

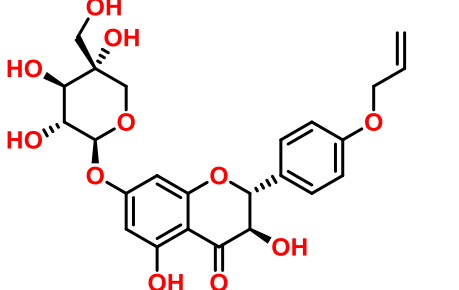
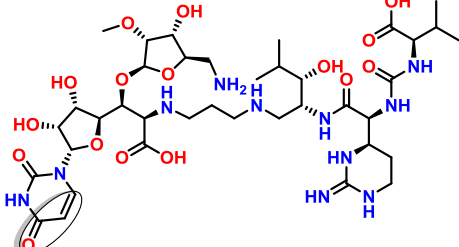
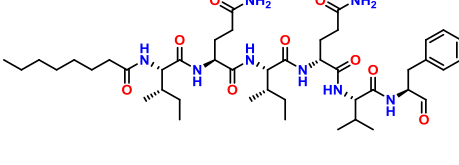
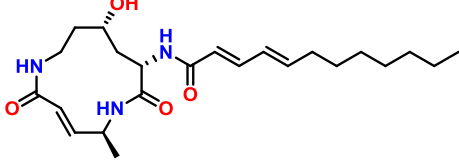
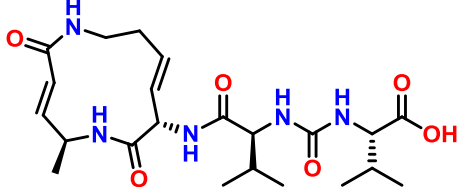
<p>P9</p>		<p>Flavonoid glycosides 6-(g,g-dimethylallyl)dihydrokaempferol-7-O-beta-D-glucoside Plant: <i>Ochna integerrima</i> (Methanolic extract) anti-HIV-1 activity with an EC50 of 29.1 µg/mL.</p>	<p><i>Nat. Prod. Rep.</i>, 2010, 27, 1781–1800</p>
<p>Peptides</p>			
<p>Pep1</p>		<p>Muraymycin C1 found to have good enzyme inhibitory activity and no toxicity.</p>	<p><i>Nat. Prod. Rep.</i>, 2016, 33, 372-381</p>
<p>Pep2</p>		<p>antipain a microbial alkaline protease inhibitor Bacterium: <i>Streptomyces albulus</i> NRRL B-3066.</p>	<p><i>Nat. Prod. Rep.</i>, 2019, 36, 1654-1686</p>
<p>Pep3</p>		<p>glidobactins Syrbactins are potent and irreversible inhibitors of the eukaryotic proteasome by oxa-Michael-type addition of a catalytic threonine at the a,b-unsaturated amide.</p>	<p><i>Nat. Prod. Rep.</i>, 2019, 36, 1654-1686</p>
<p>Pep4</p>		<p>syringolins Syrbactins are potent and irreversible inhibitors of the eukaryotic proteasome by oxa-Michael-type addition of a catalytic threonine at the a,b-unsaturated amide.</p>	<p><i>Nat. Prod. Rep.</i>, 2019, 36, 1654-1686</p>

Table S2: Predicted data of lipophilicity and water solubility of the NPs

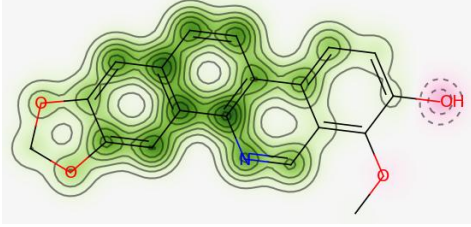
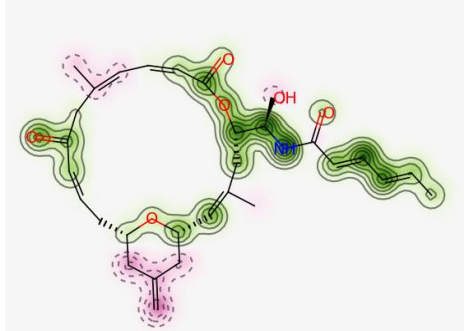
	A7	T3	T14	T1	T2	T4	T10	T16
iLOGP	3.35	4.31	4.05	4.31	4.31	4.31	5.2	3.89
XLOGP3	4.12	3.51	5.57	3.51	3.51	3.51	9.17	3.37
WLOGP	3.98	4.33	6.17	4.33	4.33	4.33	8.36	3.73
MLOGP	2.08	2.49	4.9	2.49	2.49	2.49	6.15	3.27
Silicos-IT Log P	3.86	3.54	5.54	3.54	3.54	3.54	7.19	3.98
Consensus Log P	3.48	3.63	5.25	3.63	3.63	3.63	7.21	3.65
ESOL Log S	-4.9	-4.79	-6.33	-4.79	-4.79	-4.79	-9.13	-4.04
ESOL Solubility (mg/ml)	3.98E-03	7.96E-03	2.47E-04	7.96E-03	7.96E-03	7.96E-03	4.51E-07	3.53E-02
ESOL Solubility (mol/l)	1.25E-05	1.61E-05	4.72E-07	1.61E-05	1.61E-05	1.61E-05	7.49E-10	9.08E-05
ESOL Class	M.S.	M.S.	P.S.	M.S.	M.S.	M.S.	P.S.	M.S.
Ali Log S	-5.1	-5.33	-6.99	-5.33	-5.33	-5.33	-10.83	-4.58
Ali Solubility (mg/ml)	2.52E-03	2.30E-03	5.37E-05	2.30E-03	2.30E-03	2.30E-03	8.97E-09	1.03E-02
Ali Solubility (mol/l)	7.88E-06	4.63E-06	1.03E-07	4.63E-06	4.63E-06	4.63E-06	1.49E-11	2.64E-05
Ali Class	M.S.	M.S.	P.S.	M.S.	M.S.	M.S.	Insolu.	M.S.
Silicos-IT LogSw	-6.34	-3.33	-5.81	-3.33	-3.33	-3.33	-8.01	-3.3
Silicos-IT Solubility (mg/ml)	1.47E-04	2.32E-01	8.06E-04	2.32E-01	2.32E-01	2.32E-01	5.85E-06	1.93E-01
Silicos-IT Solubility (mol/l)	4.60E-07	4.69E-04	1.54E-06	4.69E-04	4.69E-04	4.69E-04	9.71E-09	4.98E-04
Silicos-IT class	P.S.	Solu.	M.S.	Solu.	Solu.	Solu.	P.S.	Solu.

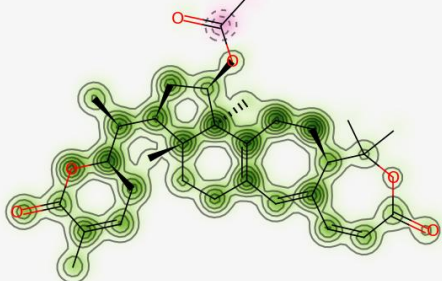
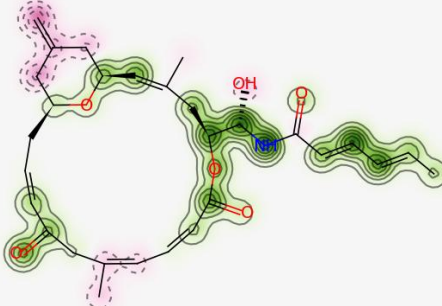
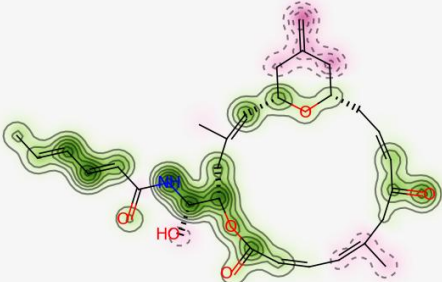
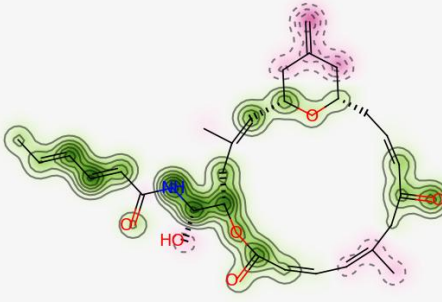
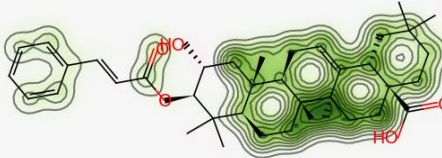
Solu. = Soluble, M.S. = Moderately soluble, P.S. = Poorly soluble, Insolu. = Insoluble.

Table S3: Predicted data of pharmacokinetics, drug likeness and medicinal chemistry of the NPs

	A7	T3	T14	T1	T2	T4	T10	T16
GI absorption	High	High	High	High	High	High	Low	High
BBB permeant	Yes	No	No	No	No	No	No	Yes
CYP1A2 inhibitor	Yes	Yes	Yes	Yes	Yes	Yes	No	No
CYP2C19 inhibitor	Yes	No	No	No	No	No	No	No
CYP2C9 inhibitor	Yes	No	No	No	No	No	No	No
CYP2D6 inhibitor	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
CYP3A4 inhibitor	Yes	Yes	No	Yes	Yes	Yes	No	Yes

Table S4: Predicted data of cardiotoxicity of NPs compounds⁵

Compound	Prediction / Potency	Confidence	Applicability domain (AD)	Probability Map
A7	Non-cardiotoxic (-)	50%	No (Value= 0.23 and limit = 0.26)	
T3	Non-cardiotoxic (-)	50%	No (Value= 0.19 and limit = 0.26)	

T14	Weak or Moderate Potential cardiotoxic (+)	60% 50%	No (Value= 0.22 and limit = 0.60)	
T1	Non-cardiotoxic (-)	50%	No (Value= 0.19 and limit = 0.26)	
T2	Non-cardiotoxic (-)	50%	No (Value= 0.19 and limit = 0.26)	
T4	Non-cardiotoxic (-)	50%	No (Value= 0.19 and limit = 0.26)	
T10	Weak or Moderate Potential cardiotoxic (+)	60% 50%	No (Value= 0.25 and limit = 0.26)	

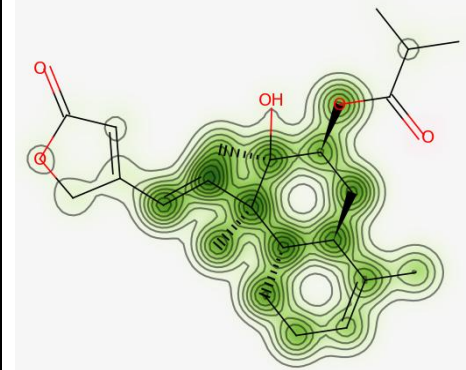
T16	Weak or Moderate Potential cardiotoxic (+)	60% 50%	No (Value= 0.21 and limit = 0.26)	
------------	---	----------------	---------------------------------------	--

Table S5: Predicted data of toxicity of NPs compounds⁶

	A7	T3	T14	T1	T2	T4	T10	T16
Predicted LD50: mg/kg	778	10000	55	10000	10000	10000	6000	55
Predicted Toxicity Class:	4	6	3	6	6	6	6	3
Average similarity	80.07	64.92	59.66	64.92	64.92	64.92	65.49	67.09
Prediction accuracy: %	70.97	68.07	63.38	68.07	68.07	68.07	68.07	68.07
Organ toxicity								
Hepatotoxicity (Probability)	0.60	0.68	0.80	0.68	0.68	0.68	0.79	0.84
Toxicity end points								
Carcinogenicity	0.60	0.54	0.50	0.54	0.54	0.54	0.54	0.57
Immunotoxicity	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.98
Mutagenicity	0.75	0.65	0.92	0.65	0.65	0.65	0.91	0.82
Cytotoxicity	0.60	0.57	0.67	0.57	0.57	0.57	0.86	0.71
Tox21-Nuclear receptor signalling pathways								
Aryl hydrocarbon Receptor (AhR)	0.68	0.96	0.98	0.96	0.96	0.96	0.98	0.98
Androgen Receptor (AR)	0.96	0.96	0.58	0.96	0.96	0.96	0.67	0.60
Androgen Receptor Ligand Binding Domain (AR-LBD)	0.91	0.97	0.64	0.97	0.97	0.97	0.72	0.61
Aromatase	0.89	0.87	0.63	0.87	0.87	0.87	0.89	0.69
Estrogen Receptor Alpha (ER)	0.91	0.88	0.57	0.88	0.88	0.88	0.74	0.58
Estrogen Receptor Ligand Binding Domain (ER-LBD)	0.90	0.93	0.70	0.93	0.93	0.93	0.90	0.97
Peroxisome Proliferator Activated Receptor Gamma (PPAR-Gamma)	0.96	0.94	0.97	0.94	0.94	0.94	0.96	0.92

Tox21-Stress response pathways								
Nuclear factor (erythroid-derived 2)-like 2/antioxidant responsive element (nrf2/ARE)	0.59	0.90	0.95	0.90	0.90	0.90	0.56	0.90
Heat shock factor response element (HSE)	0.59	0.90	0.95	0.90	0.90	0.90	0.56	0.90
Mitochondrial Membrane Potential (MMP)	0.55	0.84	0.69	0.84	0.84	0.84	0.64	0.59
Phosphoprotein (Tumor Suppressor) p53	0.81	0.81	0.93	0.81	0.81	0.81	0.92	0.81
ATPase family AAA domain containing protein 5 (ATAD5)	0.87	0.94	0.94	0.94	0.94	0.94	0.94	0.94
Toxicity Target	--	--	--	--	--	--	--	--
Weak Active, Strong Active, Weak Inactive, Strong Inactive								

Table S6: Predicted data of toxicity of some FDA approved anti-viral drugs⁶

	Rilpivirine	Dolutegravir	Glecaprevir	Grazoprevir	Paritaprevir
Predicted LD50: mg/kg	2450	1600	68	68	200
Prediction accuracy: %	67.38	54.26	23	54.26	54.26
Organ Toxicity					
Hepatotoxicity (Probability)	0.55	0.76	0.56	0.59	0.51
Toxicity end points					
Carcinogenicity	0.69	0.62	0.58	0.58	0.63
Immunotoxicity	0.94	0.98	0.96	0.99	0.99
Mutagenicity	0.60	0.56	0.64	0.62	0.63
Cytotoxicity	0.76	0.62	0.50	0.54	0.54
Nuclear receptor signalling pathways					
Aryl hydrocarbon Receptor (AhR)	0.66	0.89	0.90	0.87	0.90
Androgen Receptor (AR)	0.99	0.94	0.94	0.92	0.95
Androgen Receptor Ligand Binding Domain (AR-LBD)	0.98	0.97	0.91	0.93	0.93
Aromatase	0.82	0.88	0.83	0.92	0.94
Estrogen Receptor Alpha (ER)	0.61	0.84	0.87	0.88	0.89

Estrogen Receptor Ligand Binding Domain (ER-LBD)	0.95	0.95	0.95	0.97	0.97
Peroxisome Proliferator Activated Receptor Gamma (PPAR-Gamma)	0.98	0.97	0.92	0.85	0.88
Stress response pathways					
Nuclear factor (erythroid-derived 2)-like 2/antioxidant responsive element (nrf2/ARE)	0.96	0.95	0.88	0.88	0.92
Heat shock factor response element (HSE)	0.96	0.95	0.88	0.88	0.92
Mitochondrial Membrane Potential (MMP)	0.81	0.76	0.62	0.61	0.73
Phosphoprotein (Tumor Suppressor) p53	0.84	0.76	0.85	0.85	0.87
ATPase family AAA domain containing protein 5 (ATAD5)	0.83	0.94	0.92	0.95	0.95

Weak Active, Strong Active, Weak Inactive, Strong Inactive

1. <https://www.rcsb.org/>
2. Gaussian 09, Revision D.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2009.
3. Sousa; S. F.; Fernandes, P. A.; Ramos, M. J. Protein-ligand docking: current status and future challenges. *Proteins*. **2006**, *65*, 15-26.
4. Daina, A.; Michielin, O.; Zoete, V., SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules, *Scientific Reports*, **2017**, *7*, 42717
5. <http://predherg.labmol.com.br/>
6. http://tox.charite.de/protox_II/index.php?site=compound_input

7. Mishra, P., Günther, S. New insights into the structural dynamics of the kinase JNK3. *Sci Rep* 8, 9435 (2018). <https://doi.org/10.1038/s41598-018-27867-3>
8. Hospital, A., Goñi, J. R., Orozco, M., & Gelpí, J. L. (2015). Molecular dynamics simulations: advances and applications, *Advances and applications in bioinformatics and chemistry : AABC* , 8, 37–47. <https://doi.org/10.2147/AABC.S70333>