Computational studies towards identification of lead herbal compounds of medicinal importance for development of nutraceutical against COVID-19

Amit Kumar Srivastav[a], Sanjeev K. Gupta[b] and Umesh Kumar[a,*]

[a]School of Nano Sciences, Central University of Gujarat, Gandhinagar 382030, India

[b]Computational Materials and Nanoscience Group, Department of Physics, St. Xavier's College, Ahmedabad 380009, India

***Corresponding authors**: Dr Umesh Kumar (umesh.kumar@cug.ac.in)

Abstract

SARS-CoV-2 (COVID-19) disease is a new strain of coronavirus (2019-nCoV) which has spread all over the globe in a very short time becoming pandemic. Till date, there is no particular drug or therapy that can be used for the treatment of COVID-19 infection. This novel 2019-nCoV is a potential drug target in our recent study. In the present study, we have performed the in-silico study of SARS-CoV-2 structure with different herbal compounds of medicinal importance. We selected four viral key proteins of SARS-CoV-2 structure i.e ACE-2 Receptor, Main Protease (Mpro), APO Form, Cryo- electron microscopy structure for the Molecular docking followed by the molecular dynamic simulation. Using this simple in silico approach based on the molecular docking and MD simulation of protein and phytochemicals, we have identified potential lead candidates for the development of low cost nutraceuticals, which can be used against SARS-CoV-2 virus. Our analysis suggested that phytochemicals obtained from *Phyllanthus emblica* and *Azadirachta indica* have the highest potential to bind with ACE2 receptor or main protease of SARS-CoV-2, inhibiting the protease enzymatic activity. The lead compounds of herbal origin were docked and simulated on viral key proteins of SARS-CoV-2 structure to evaluate the binding affinity of these phytochemicals along with the type of interaction and its stability in terms of RMSD and Ramachandran plot. Further, these results were also verified by drug likeness properties by using SwissADME software. Overall, our results suggest that out of 14 herbal compounds, Nimbolide and Withaferin-A has great potential to be developed as low-cost nutraceuticals against SARS-CoV-2 virus, which is the need of hour. Though our results seem very promising, it needs to be validated by doing invitro and in-vivo studies accompanied by clinical trials in proper set up.

1. Introduction

Coronavirus Disease (SARS-Cov-2) is a highly infectious disease caused by newly discovered severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), that is responsible for respiratory illness effecting different parts of the respiratory system especially the lungs.[1, 2] This novel Corona virus was first identified and reported in December 2019, with the very first infection case for human in Wuhan City located at Hubei, China.[3] The novel coronaviruses are family of viruses comprises of enveloped RNA viruses.[4, 5] It is a zoonotic disease, which can be transmitted from animal to people causing the respiratory illness.[5] The World Health Organization (WHO) declared this disease as a pandemic on 11th March 2020 and SARS-CoV-2 as the deadliest virus. According the WHO report on 24th June, 2020, it infected 91,29,146 people globally and death toll rate reached 4,73,797. In such extreme health crisis, this virus is spreading at a much faster rate and scale than any previous coronavirus epidemics, which causes 5.18 % death rate of infected individual at present.[6] As the antigen is novel for a human host, public health is being seriously challenged. The spike protein on the surface of the novel corona virus particle plays an important role in the binding of the drug or inhibiting the protease enzymatic activity for targeting the active sites of 2019-nCoV.[7] Therefore, we can say that the world is in dire need for the fast development of drug, which is not yet available.[7] In the recent times, computational aided drug design methods have shown great importance as these techniques are expected to be faster and cost efficient.[8] In early stages of drug discovery in-silico methods aid in minimizing the risk of time and cost.[9] Therefore, computational aided drug design is one of the most popular and reliable technique for the development and assessment of the theoretical findings for drug discovery. Now days, phytochemicals are considered to be one of the key source of antiviral drug molecules.[10] Therefore, it is of global interest to identify and exploit naturally occurring phytochemicals or herbal compounds with lesser side effects, for the development of low cost nutraceuticals to control the SARS-CoV-2

viral infections.[11] In the present work, active target (Mpro, ACE2, spike protein etc.) of novel corona virus protein were screened with different Asian and European originated phytochemicals isolated from Amla (Phyllanthus emblica), Elderberry (Sambucus), Moringa (Moringa oleifera), Neem (Azadirachta indica), Tulsi (Ocimum tenuiflorum), Liquorice (Glycyrrhiza glabra), Ashwagandha (Withania somnifera) using molecular docking tools and molecular dynamic simulation. Molecular docking studies are based on the scoring algorithm, which calculates the binding energy (ΔG) between the ligands and the protein. The Binding energy predicts the strength of association or *binding affinity* between two molecules. Using these docking results, we performed Molecular dynamic simulation study for understanding the motions and structural behaviour of corona virus with high binding affinity phytochemicals at an atomic scale via theoretical principles. Our study includes 15 specific phytochemicals of herbal origin and four different structural motifs of SARS-CoV-2 virus for qualitative analysis along with the validation and comparison study. The list of various phytochemicals and different structural motifs of SARS-CoV-2 is given in Table 1&2. 3D structure of original SARS-CoV-2 (2019-nCoV) A: PDB (6LZG, ACE2 Receptor); B: PDB (5RH4, Mpro Structure). C: PDB (6M2Q, APO Structure); D: PDB (6Z43, Cryo EM Structure) is shown in Figure 1.

Table1: List of Phytochemicals

Herb	Phytochemicals/active ingredient
Amla	Chebulagic Acid
	Pedunculagin
Elderberry	Flavylium
	Cyanidin-3-glucoside
	Cyanidin-3-sambubioside
Moringa	Chlorogenic acid
Neem	Azadirachtin
	Nimbolide
Tulsi	Apigenin
	Oleanolic Acid
Liquorice	Glabridin
	Glycyrrhizin
	Liquiritigenin
Fennel	Trans-anethole
Ashwagandha	Withaferin-A

Table2: List of SARS-CoV-2 Structures

| Type of structure	Resolution	Description
Mpro Structure (5RH4)	1.34 Å	PanDDA analysis group deposition SARS-CoV-2 main protease fragment screen Crystal Structure of SARS-CoV-2 main protease in complex with Z1530425063 (Mpro-x2659).
ACE2 Receptor (6LZG)	2.50 Å	Structure of novel coronavirus spike receptor-binding domain complexed with its receptor ACE2.
APO Structure (6M2Q)	1.70 Å	SARS-CoV-2 3CL protease (3CL pro) APO structure (space group C21).
Cryo EM Structure (6Z43)	3.30 Å	Cryo-EM Structure of SARS-CoV-2 Spike: H11-D4 Nanobody Complex.
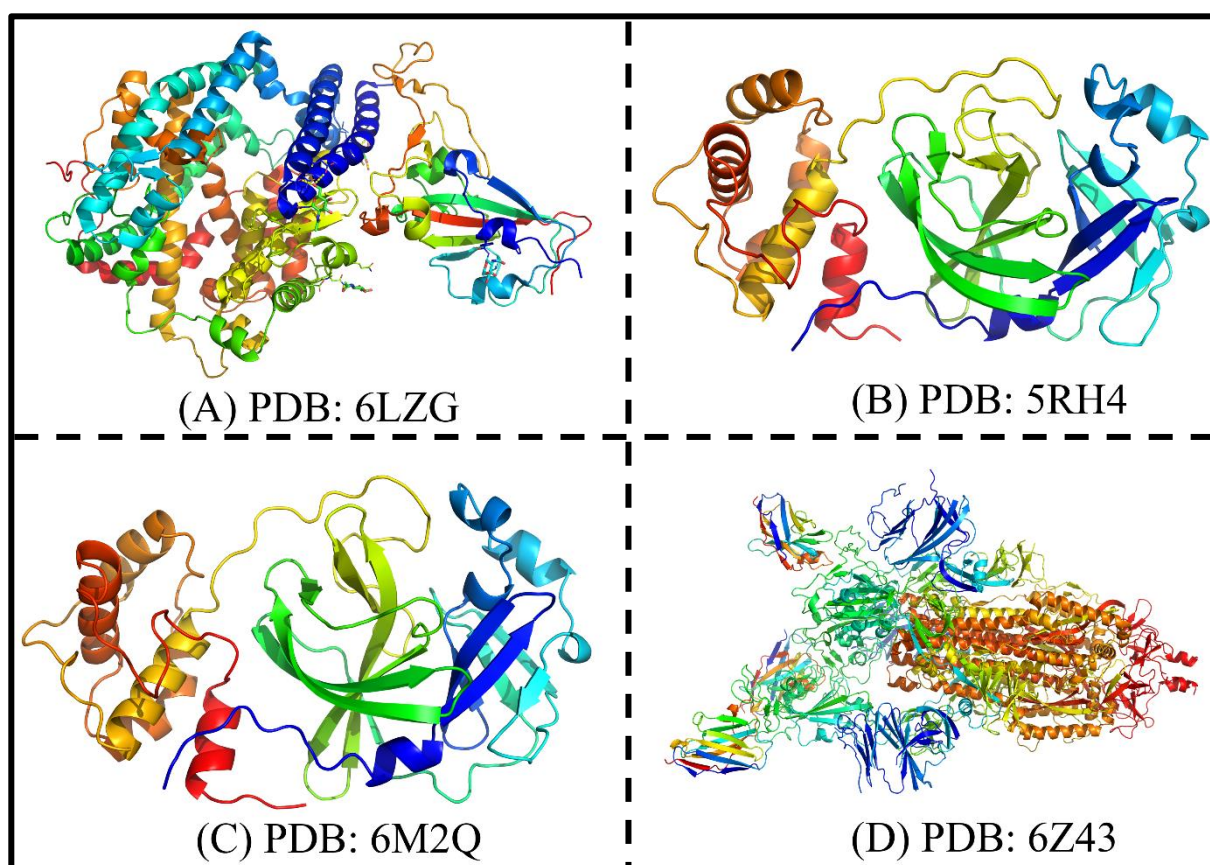
Figure 1: 3D structure of original SARS-CoV-2 (2019-nCoV) A: PDB (6LZG, ACE2 Receptor); B: PDB (5RH4, Mpro Structure). C: PDB (6M2Q, APO Structure); D: PDB(6Z43, Cryo EM Structure).Each color represents specific amino acid residues. Like bright red indicates ASP,GLU; yellow indicates CYS,MET; blue indicates LYS,ARG; orange represents SER,THR; mid blue indicates PHE,TYR; cyan indicates ASN,GLN ; light grey represents

GLY; green indicates LEU,VAL,ILE ;dark grey indicates ALA ; pink indicates TRP; pale blue represents HIS; flesh color indicates PRO amino acids.

The angiotensin- converting enzyme-2 (ACE2) receptor is a prominent receptor on the surface of many cells.[12] Using the spike-like protein on its surface, the SARS-CoV-2 virus binds to ACE2 – like a key being inserted into a lock – prior to entry and infection of cells.[13, 14] Hence, ACE2 acts as a cellular doorway – a receptor for the virus that causes transmission of COVID-19.[15, 16] While the virulence of this novel corona virus is due to the presence of main protease, which is primarily responsible for virus replication.[17, 18] Therefore, we have focused on the main protease and ACE2 receptor as the target protein to identify the best inhibitory drug molecule of herbal origin by using in silico computational methods like molecular docking and simulations. Along with these two structures, we also focused on APO and Cryo EM Structure of SARS-Cov-2 for the comparative study between all the motifs. Further, these results were also validated by investigating their pharmacokinetics and bioavailability using SwissADME to create BOILED-Egg model based on gastrointestinal absorption and brain access.[19] This model can be applied in a variety of settings, starting from the filtering of lead compounds at the early steps of drug discovery, to the evaluation of drug candidates for development. The aim of present study is to screen various herbal compounds based on their inhibitory activity against receptor and main protease of SARS-Cov-2 virus by means of computational studies towards development of low-cost nutraceutical with less side effects. The rationale behind selecting these herbal plants was that they are already known for their immune boosting or antiviral property.[10, 20, 21] Though our results seem very promising, it needs to be validated by doing invitro and in-vivo studies accompanied by clinical trials in proper set up.

2. Materials and methods

2.1 Molecular docking process

Molecular docking studies were carried out using the Auto Dock Vina software for the investigation and analysis of the molecular interaction between different phytocompounds and SARS-Cov-2 virus.[22] The crystal structures of SARS-CoV-2 (PDB code: 5RH4, 6LZG, 6M2Q, 6Z43) were obtained from the Protein Data Bank (www.pdb.org). For docking analysis, the studied ligands were sketched using pymol interface. The water molecules and metals were stripped out from the proteins pdb files .The polar hydrogen atoms were added to the amino acid residues. Missing side chains and residues were fixed and protein energies were minimized using the Amber10 force field. In addition, gasteiger charges were assigned to all atoms of the

protein. The resulted structures were used for docking analysis in Autodock programs. The phytochemicals were extracted from the PubChem database in the SDF format and were converted to PDB format using PyMOL. Then the protein in PDBQT format was used as an input for the AUTOGRID program. Docking calculations were carried out using the Lamarckian genetic algorithm (LGA), and all parameters were the same for each docking.[23] The grid box conformation is given in Table 3. The exhaustiveness was set 24 for each docking calculation.[24]

Table 3: Molecular Docking grid box specification.

	Centre			Size		
	X	Y	Z	X	Y	Z
Mpro Structure (5RH4)	11.822	0.702	4.799	74	89	106
ACE2 Receptor (6LZG)	-26.283	18.290	-13.899	80	74	92
APO Structure (6M2Q)	37.109	-12.751	49.702	98	78	94
Cryo EM Structure (6Z43)	218.126	218.609	210.753	112	110	124

These output files were stored in PDBQT files, each having nine poses. Different poses were analyzed in the AutoDock tool and pymol. The pose with the highest binding energy or affinity was selected for further molecular dynamics simulation analysis by using the GROMACS software.

2.2 Pharmacokinetics and bioavailability investigation by SwissADME web tool

The Lipinski rule of five signifies the drug ability of the compound. This was calculated using SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules.[25] The molecular weight, number of hydrogen bond donor and acceptor, number of Lipinski violations, number of rotatable bonds all were calculated and represented as BOILED EGG construction.

2.3 MD SIMULATION

The Ligand-ACE2-RBD complex and Ligand_Mpro complex were obtained after the molecular docking studies. The simulation systems for ACE2-RBD complex and Ligand_Mpro

complex were prepared using the GROMACS software.[26] MD simulation was performed with CHARMM36 force field using the NAMD package.[27] The protein complex with selected phytochemicals was solvated with TIP3P water molecules in 10*10*10 Å box. The systems were neutralized by adding counter ions and periodic boundary conditions were applied. The Particle Mesh Ewald method[28] was used to calculate the long-range electrostatic interactions while the SHAKE algorithm was used to constrain the hydrogen bonds.[29] NPT and NVT ensemble was used with periodic boundary conditions. Pressure was fixed at 1 atm, while the temperature was set at 300 K. The particle-mesh Ewald method was used to evaluate the Coulomb interactions. 2 fs of time step was used in all MD simulations. Initially, water was equilibrated for 200 ps at 300 K after fixing the protein and energy minimization of 1000 steps. 1000 steps of energy minimization of the whole system were performed, and further equilibration for 400 ps at 310 K after releasing the protein was done. Simulation run was performed of 25 ns. The trajectory data were saved at every 1 ps to analyze the change in the dynamics of ACE2-RBD, Mpro structural binding interface using the VMD. The root mean square deviation, Ramachandran plot and snapshots at every 5 ns were calculated for all the sixteen simulated systems (1. ACE2-Chebulagic Acid, 2. ACE2-Pedunculagin, 3. ACE2-Azadirachtin, 4. ACE2- Nimbolide, 5. Mpro-Chebulagic Acid, 6. Mpro-Pedunculagin, 7. Mpro-Azadirachtin, 8. Mpro- Nimbolide, 9. APO-Chebulagic Acid, 10. APO-Pedunculagin, 11. APO-Azadirachtin, 12. APO- Nimbolide, 13. CryoEM-Chebulagic Acid, 14. CryoEM-Pedunculagin, 15. CryoEM-Azadirachtin, 16. CryoEM- Nimbolide) to predict the reasons attributable to their stability. The graphs were plotted using the origin and VMD software.

3.   Result and Discussion

3.1 Molecular Docking

The molecular docking studies with various phytochemicals highlighting the different binding and pocket sites for the 2019-nCoV are depicted in Table 4,5 & 6. Each phytochemical show 9 conformation of binding pockets for targeting sites. Each conformation has different binding affinity with respect to the rmsd value of lower base and upper base. Docking study results for the binding of these phytochemicals are shown in Figure 2. The SARS-CoV-2 structure has different active binding or targeting sites for each different phytochemical. These active target sites can be used for targeted drug delivery to inhibit the protease enzymatic activity. The detailed description of active binding or targeting sites for each phytochemical are shown in Table 4,5 & 6.

Table 4: Highest Docking scores of different 15 phytochemicals with active site residues of ACE2 6LZG structure

Herb	Phytochemicals/ active ingredient	ACE2	Amino Acid Residues	Binding Domain	Types of Interaction
Amla	Chebulagic Acid	-18.4	TYR-196(3.1), GLN-102(3.1,3.4), GLN-98(2.8,2.9,3.4), LSY-562(2.8), ASN-210(2.6,3.2,3.4), ASP-206(2.4), ALA-396(2.2)	ACE2	H- bond, π-π, π- cation
	Pedunculagin	-17.5	GLN-102(3.2), GLN-98(2.8,2.9,3.1), ASN-210(2.7,3.0,3.4), TYR-196 (2.5)	ACE2	H- bond, π-π, π- cation
Elderberry	Flavylium	-9.5	MET-366 (2.8,2.9), PHE-438(2.2), GLU-430 (2.7,2.9), LYS-541(3.3), THR-434(3.0), ASN-290 (2.8,2.9)	ACE2	H- bond
	Cyanidin-3-glucoside	-10.2	GLU-564 (2.4,2.7), TRP-566(3.1), ASP-206(2.8), GLN-98 (3.2,3.2), ASN-210(3.0)	ACE2	π- cation, H- bond
	Cyanidin-3-sambubioside	-8.2	ARG-403 (2.8,3.1), TYR-505(2.8, 2.9), ARG-393 (3.0), LYS-417(3.2), GLN-409(2.9,3.1,3.2), ARG-408 (3.3,3.0,3.1,3.2)	ACE2, RBD (β6, β5)	H- bond, π-π, π- cation
Moringa	Chlorogenic acid	-8.1	GLN-102(2.9), TYR-196(2.8), ARG-514(3.22.8,3.0), TRP-203(3.0), ASP-509(2.3,2.1), SER-511(3.3,3.2,2.8,3.0)	ACE2	H- bond, π-π, π- cation

Neem	Azadirachtin	-12.4	GLU-37(2.7), GLY-496(3.4), TYR-505(3.3), PHE-497(3.6)	ACE2, RBD (β6, β5)	H- bond, π-π
	Nimbolide	-16.7	THR434, LYS-541, GLN-429, THR-414, ILE-291	ACE2	π- cation
Tulsi	Apigenin	-9.3	ILE-291(2.9,3.1), THR-414(2.1)	ACE2	H- bond
	Oleanolic Acid	-9.6	ASN-394(2.8), ARG-393(2.1)	ACE2	H- bond
Liquorice	Glabridin	-9.7	ASP-367(2.2,2.7), ILE-291(3.4)	ACE2	H- bond, π-π
	Glycyrrhizin	-15.3	ASP-367(2.4), LYS-441(3.3), MET-366(2.8), ASP-292(2.1)	ACE2	H- bond
	Liquiritigenin	-8.9	ILE-291(3.1), MET-366(2.9), GLU-430(2.8,3.0)	ACE2	H- bond
Fennel	trans-anethole	-6.9	HIS-540(2.3,2.2), GLU-430(3.0), ASN-290(2.7,2.9)	ACE2	H- bond
Ashwagandha	Withaferin-A	-9.3	THR-371(3.3), ILE-291(3.0)	ACE2	H- bond

Table 5: Highest Docking scores of different 15 phytochemicals with active site residues of Mpro Structure (5RH4)

Herb	Nutraceuticals active ingredient	Mpro	Amino Acid Residues	Binding Domain	Types of Interaction
Amla	Chebulagic Acid	-14.6	GLU290(2.9), LYS-5(2.7,2.5,2.5,2.6), GLN-127(2.9,3.0), PHE-8(3.2,3.7), CYS-128(2.0), ASN-151(2.9,3.2), PHE-112(2.1)	Domain-II	H- bond, π-π, π- cation
	Pedunculagin	-14.2	ASN-238(2.6), ASP-197(2.8), LYS-137(2.8,2.9), THR-199(3.3), LEU-287(2.6), ASP-289(3.2)	Domain-III	H- bond, π-cation
Elderberry	Flavylium	-6.8	GLN-192(3.0,3.2), ASN-142(3.2), MET-165(2.9),	Domain-I,	H- bond

				Domain-II	
	Cyanidin-3-glucoside	-9.1	SER-144(3.3), GLY-138 (2.4), GLU-166(3.1), HIS-41(3.3), ASN-142(2.8,3.0)	Domain-I, Domain-II	H- bond
	Cyanidin-3-sambubioside	-7.6	GLN-192(2.8), THR-190(3.3,2.5), HIS-41(3.2), LEU-141(2.5) PHE-140(1.9),	Domain-I, Domain-II	H- bond, π-π
Moringa	Chlorogenic acid	-7.4	LYS-137(2.9), ARG-131(3.5,2.9), TYR-239(2.8), THR-199(3.3,3.1,3.0), LEU-271(2.7),	Domain-II, Domain-III	H- bond, π-π, π- cation
Neem	Azadirachtin	-11.5	GLY-143(3.0,3.3,3.4) THR-26(3.1)	Domain-I, Domain-II	H- bond, π-cation
	Nimbolide	-12.3	GLN-110(2.5), SER158(3.0), THR-111(2.8)	Domain-I	H- bond
Tulsi	Apigenin	-8.2	SER-144(3.0), HIS-163(3.1), GLU-166(3.1), THR-190(3.2)	Domain-I, Domain-II	H- bond, π-π
	Oleanolic Acid	-9.5	GLU-166(2.1,3.0), THR-24(3.3), THR-25 (3.0,2.8)	Domain-I, Domain-II	H- bond, π-π, π- cation
Liquorice	Glabridin	-8.9	ALA-191(2.1), THR-190(3.4), ASN-142(3.2)	Domain-I, Domain-II	H- bond, π-π
	Glycyrrhizin	-13.3	GLN-192(2.8), THR-190(2.4), THR-45(2.5), GLY-143(2.7)	Domain-I, Domain-II	H- bond
	Liquiritigenin	-7.5	THR-190(3.3), GLN-192(2.8), HIS-41(3.1)	Domain-I, Domain-II	H- bond
Fennel	trans-anethole	-6.1	GLN-127(3.0)	Domain-II	H- bond

| Ashwagandha | Withaferin-A | -8.3 | THR-24(1.9,2.6) THR-26(2.9,3.2), GLN-189(2.5) | Domain-I, Domain-II | H- bond |

Table 6: Highest Docking scores of different 15 phytochemicals with active site residues of APO Structure (6M2Q) Cryo EM Structure (6Z43)

Herb	Phytochemicals / active ingredient	APO form	Amino Acid Residues	Cryo-EM Structure	Amino Acid Residues
Amla	Chebulagic Acid	-15.1	LEU-287(2.9,3.1), THR-199(3.3), ASP-17(2.4), ARG-131(3.3,3.3)	-20.8	THR-376(2.4,2.8,2.9), TYR-489(2.3,2.4,2.5), THR-385(3.2,3.5), TYR-369(2.4), SLR-383 (2.8)
	Pedunculagin	-14.3	LEU-167(2.8), MET-49(2.9), SER-46(2.8,3.2), THR-26(3.0,3.3), HIS-41(3.2,3.3), CYS-145(2.6)	-17.7	PHE-377 (2.4), SER-383(3.2), TYR-369(2.8,2.3) (2.8), ALA-475(2.2,2.8), TYR-421(2.8)
Elderberry	Flavylium	-6.9	LEU-287(3.0), ASP-289(2.9)	-10.4	GLU-661, PRO-853, SER-698
	Cyanidin-3-glucoside	-9.5	PHE-140(2.5), GLY-143(3.2), ASN-142(2.9), LEU-167(2.3), HIS-41(3.2), GLU-166(2.7), HIS-163(2.9)	-11.5	LYS-529, ASN-544, THR-523, ASN-388
	Cyanidin-3-sambubioside	-8.0	LYS-5(3.1) GLU-290(2.1), ARG-131(3.0), ASP-289(2.6), TRP-207(3.1), LEU-282(2.0,2.2), SER-284(3.0)	-9.1	ALA-348(2.5,2.7) ARG-509(2.8,2.5) PRO-426(2.9,2.5), SER-469(2.0), CYS-432(2.5), THR-430(2.3,2.5)
Moringa	Chlorogenic acid	-7.4	THR-199(2.9,2.9,3.3), ARG-131(2.9)	-9.5	THR-333(2.7), PRO-337(2.5), THE-338(2.0,2.8), SER-349(2.5)
Neem	Azadirachtin	-11.4	GLY-143(3.0,3.3,3.4) THR-26(3.1)	-13.3	THR-1027 (3.2,3.4,3.1), ARG-

					1039 (2.9,3.5), GLU-725 (2.5)
	Nimbolide	-12.5	HIS-41(2.4), GLN-110(3.0)	-14.7	LYS-1038 (2.9), TRP-886 (3.1)
Tulsi	Apigenin	-8.2	HIS-164(2.2,2.0), LEU-167(2.5), ALA-129(2.3), LEU-141(3.0)	-9.9	ASN-481(2.1), ILE-468(2.5), PHE456(2.1), LYS-528(1.9,2.3), VAL-102(2.8,3.0)
	Oleanolic Acid	-9.3	GLY-109(2.9), GLN-110(3.1,3.3), THR-292(2.2), ASN-203(2.5)	-10.1	PHE-168(2.7), ALA-520(2.9), SER-514 (3.2), LEU-229 (2.2,2.3), SER-514(2.4,2.6)
Liquorice	Glabridin	-7.7	ASN-214(2.4), ASN-203(2.8), GLY-283(2.9,3.2)	-11.2	LEU-84(2.4), ASP-88(2.2), PRO-521(2.7), PHE-329(3.1)
	Glycyrrhizin	-12.7	THR-198(2.7), PRO-132(3.0,3.2), GLN-107(2.0), LEU-287(1.9)	-15.9	GLU-324(2.2,2.4), LYS-41(1.9), VAL-227(2.7), ASP-198(2.6), PHE-541(3.1,3.3), ASN-388(2.1)
	Liquiritigenin	-7.4	ARG-105(2.7), LEU-177(2.0,2.1), PHE-181(2.4), TYR-161(3.3)	-10.9	GLY-447(3.1), LEU-44(2.4), TYR-453(2.1)
Fennel	trans-anethole	-6.1	ASN-203(2.5), PHE-291(2.8), GLU-290(2.3,2.2)	-7.1	LYS-528(2.9), ASN-540(3.2), VAL-327(3.0)
Ashwagandha	Withaferin-A	-8.4	THR-199(2.6), GLY-109(2.4,2.2), ASP-289(2.9)	-10.9	GLY-55(2.6), ASP-467(2.8), GLY-232(2.3)

The phytochemicals compounds Chebulagic Acid, Pedunculagin of Amla (Phyllanthus emblica) and Azadirachtin, Nimbolide of Neem exhibited highest binding energies for the SARS-CoV-2 virus structure. These phytochemicals like Chebulagic Acid (-20.8 to -14.6 kcal/mol), Pedunculagin (-17.7 to -14.2 kcal/mol), Azadirachtin (-13.3 to -11.3 kcal/mol) and Nimbolide (-16.7 to -12.2 kcal/mol) (see Fig 2,3) could become potential candidates for targeting or inhibiting the protease enzyme of the SARS-CoV-2 viruses. The binding energy values against Mpro and ACE2 receptor for these compounds are also good in comparison to other phytochemicals.
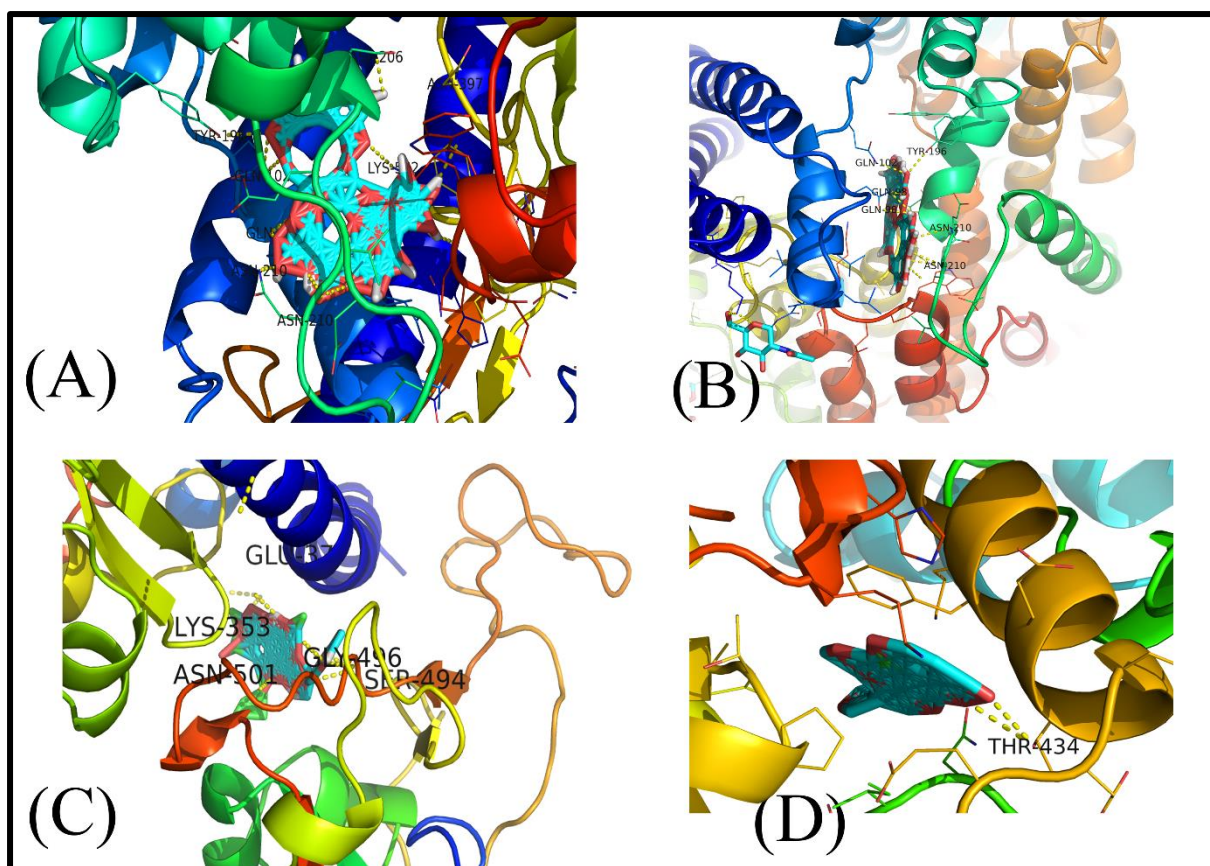
Fig:2 Docked structure of (A) Chebulagic Acid (B) Pedunculagin (C) Azadirachtin (D) Nimbolide with Structure of novel coronavirus spike receptor-binding domain complexed with its receptor ACE2.

Additionally, We also analyzed the type of interaction between the phytochemicals and amino acid residues within active site of SARS-CoV-2 virus structure. In ACE2 motif, Chebulagic acid is forming π- cation and H-bond interaction with ASN-210. The benzene ring shows π- π interaction with GLN-102, GLN-98., while TYR-196 and LYS-562 formed the H bond interaction with the ACE 2 receptor of SARS-CoV-2 virus. This vast number of interactions gives the complex its stability with the −18.4 kcal/mol binding energy. Pedunculagin, the active ingredient of Amla (Phyllanthus emblica) formed H-bond, π-π, and π- cation interaction with GLN-102, TYR-196; GLN-98; and ASN-210 in ACE2 domain respectively. Azadirachtin and Nimbolide the other phytochemical compound of Neem (Azadirachta indica) also showed higher binding affinity towards the SARS-CoV-2 structure with H- bond and π- cation interaction with corona virus structure.
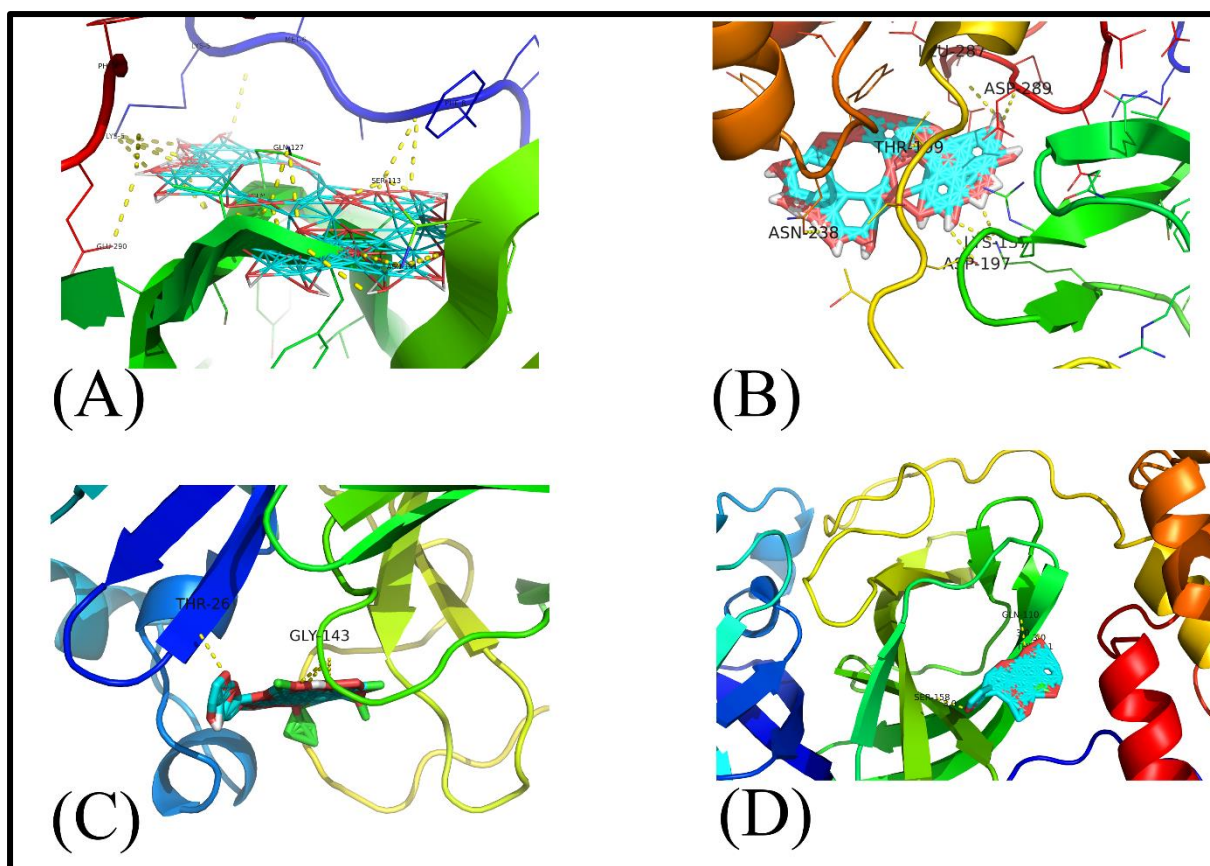
Fig:3 Docked structure of (A) Chebulagic Acid (B) Pedunculagin (C) Azadirachtin (D) Nimbolide with SARS-CoV-2 main protease Mpro Structure of novel coronavirus.

The details of types of interactions are given in Table 4, 5 for Mpro and ACE2 receptor. Many phytochemical compounds show hydrophobic contacts towards the SARS-CoV-2 structure with π-π, and π- cation interactions. Docking studies with Mpro protease also exhibit similar types of interaction with different phytochemicals e.g Chebulagic acid formed H-bond with the O-H group to the GLU-290, PHE-8,PHE-112, CYS-128, while due to the presence of phenol ring in the structure it also formed the π-π, and π- cation interactions with the LYS-5, GLN-127 and ASN-151 respectively. The details of the bond length and interaction type with the binding domain is described in the Table 4, 5 of docking results of ACE2 and Mpro structure.

3.2 MD SIMULATION

Molecular dynamics simulation studies were performed using the GROMACS software. Molecular Dynamics (MD) simulation is an attractive approach to explore the real-time dynamics and conformational stability of a protein upon binding of a ligand. Simulation studies of 25 ns for the highest docking affinity of phytochemical based docked structure were carried out. The Root Mean Square Deviation (RMSD) from the initial crystal structure, protein-ligand binding contacts in terms of hydrogen bonding and electrostatic interactions were also estimated to get an insight into the overall stability of the system.

Table 7: RMSD values of simulated ligand- protein docked complex

Herb	Phytochemicals/ active ingredient		RMSD (Å)				
			5 ns	10 ns	15 ns	20 ns	25 ns
Amla	Chebulagic Acid	ACE2	2.11	3.47	5.11	6.33	6.55
		M pro	1.59	2.60	4.80	6.40	7.84
		Cryo-EM Structure	2.95	5.88	7.52	7.94	8.06
		APO form	1.80	4.66	5.98	6.24	6.69
	Pedunculagin	ACE2	0.32	2.16	3.27	4.52	5.00
		M pro	1.99	3.70	4.74	5.57	6.64
		Cryo-EM Structure	3.30	6.05	7.29	8.42	9.10
		APO form	2.05	3.26	4.55	5.37	5.68
Neem	Azadirachtin	ACE2	0.94	3.56	5.70	6.99	7.28
		M pro	2.4	5.69	6.10	7.27	7.70
		Cryo-EM Structure	1.55	5.76	7.27	9.62	13.51
		APO form	2.03	4.96	5.92	6.77	7.04
	Nimbolide	ACE2	1.60	2.81	4.42	5.67	6.12
		M pro	1.23	3.00	5.06	7.10	7.48
		Cryo-EM Structure	3.51	5.54	6.60	8.18	11.03
		APO form	0.22	2.74	4.06	5.27	5.96

Our results from MD simulation shows that the ACE2-complex and APO form complex have least RMSD value indicating higher stability in terms of electrostatic interactions and overall stability of the simulated structure. The Root Mean Square Deviation (RMSD) conveys the information of overall stability of the protein-ligand complex in terms of deviation from the initial structure. As depicted in Table 7 all the fifteen systems were significantly stable with variable deviation. All the systems have projected RMSD around 6-9 Å except Azadirachtin

and Nimbolide for Cryo-Em Structure, which shows the RMSD up to 11-13 Å. Upon calculating the average deviation, Chebulagic Acid presented an average RMSD of 7.28 Å with slight increase fluctuation in Cα backbone around 20000ps and was stabilized in the remaining simulation. Similarly, in the case of Pedunculagin the analysis of RMSD of Cα backbone has prompted an average RMSD of 6.60 Å. For Azadirachtin bound complex, the average RMSD was found to be 8.88 Å with the overall RMSD of around 9-10 Å suggesting its stability. An increasing fluctuation was observed in the starting and upto 15000ps and stable thereafter throughout simulation. In case of Nimbolide bound complex to target protein, the average RMSD of around 7.64 Å was observed. The Nimbolide bound complex shows the least fluctuation throughout the simulation process among the fifteen selected systems. However, RMSD was significantly stable throughout the simulation with negligible fluctuation as compared to other systems inferring the formation of most stable ligand bounded protein complex of simulated system followed by a stable RMSD values. The data indicates that all the systems showed stable internal motion. SEE Fig.4,5
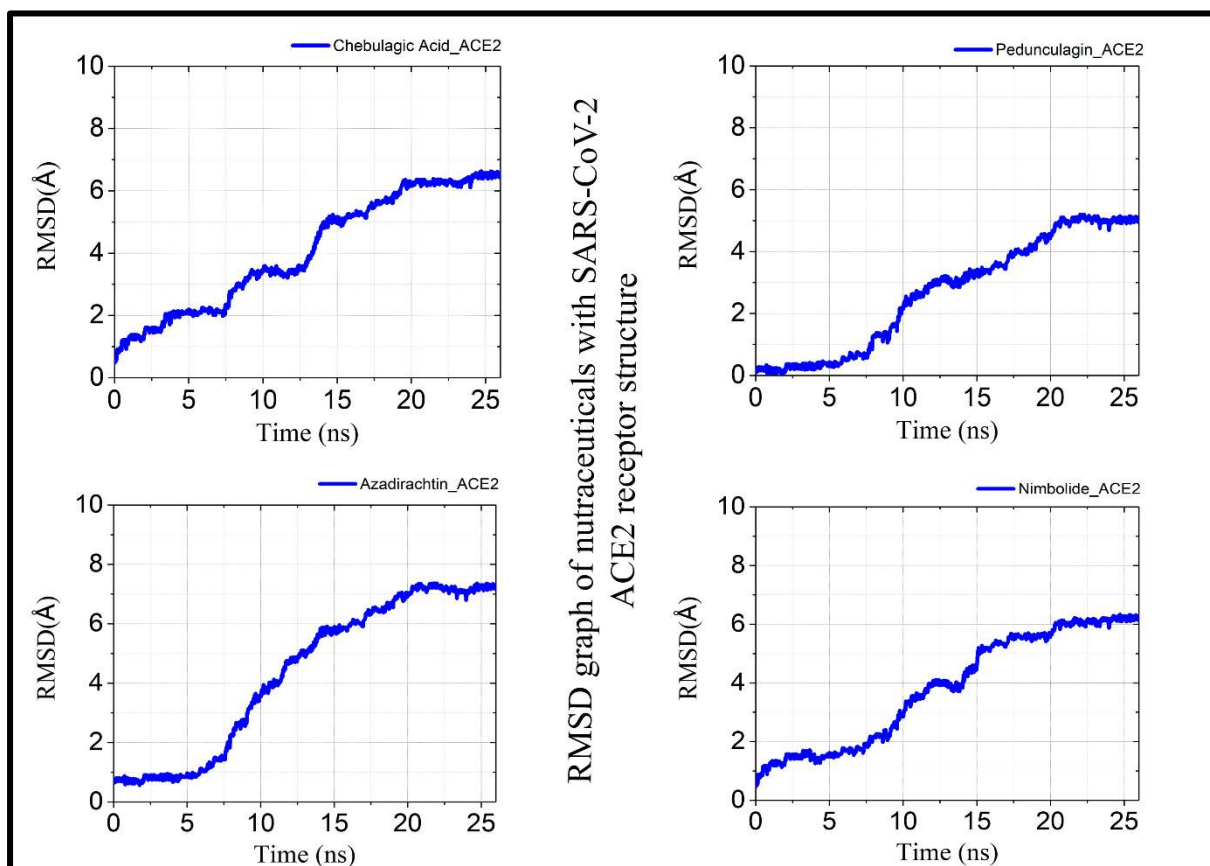

Figure 4: Root mean square deviation graph for ACE2 receptor docked phytochemical.
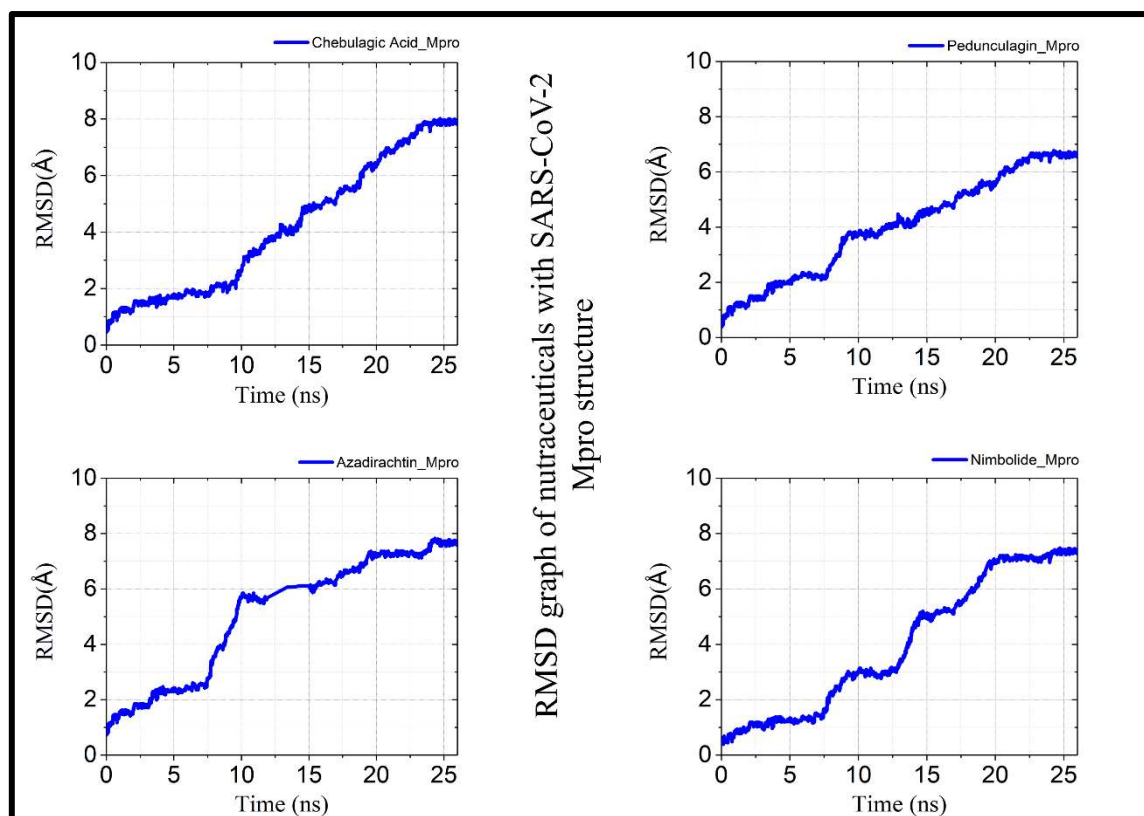
Figure 5: Root mean square deviation graph for Mpro (main protease) docked with phytochemical.

Further, Ramachandran plot was plotted in 2D to analyze the changes in the coordinates of the simulated system. We analyzed and confirmed the changes in the conformational structure of all ligand-protein complexes for the energetically activated coordinates region for the backbone dihedral angles ψ against φ of the amino acid residues. (Fig6, 7) Validation of simulated system using Ramachandran plot were performed with the PROCHECK server[30].(Fig:8) It revealed that simulated Chebulagic Acid with ACE2 receptor has 89.9% (594) residues of COVID-19 protease were in the most favoured regions[A,B,L], followed by 8.9% (59) in additional allowed regions [a,b,l,p], 0.5% (3) in generously allowed region[~a,~b,~l,~p] and 0.8%(5) in the disallowed regions. Overall, G factor for the predicted structure was - 0.05 (Table-8). The G-factor provides a measure of the normalcy of stereo-chemical property of a protein model. Values below - 0.5 shows unusual stereo-chemical property while values below - 1.0 show a highly unusual property. Since G value obtained for the predicted model in the present study is not less than - 0.5, it is suggestive of satisfactory quality check for the presence of any closely related homologue of SARS-CoV-2.[30] This background data further prompted us to continue our hunt for potential phytochemical candidate for its inhibitory activity against receptor and

main protease of the SARS-CoV-2 virus as the target of foremost importance. Complete result of Ramachandran plot is tabulated in Table no.8
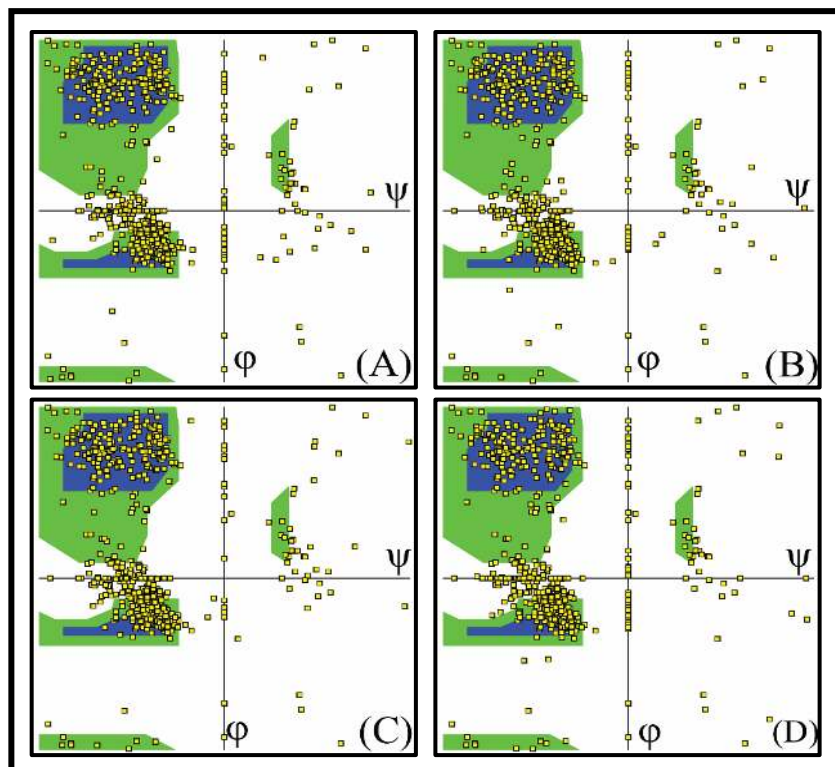

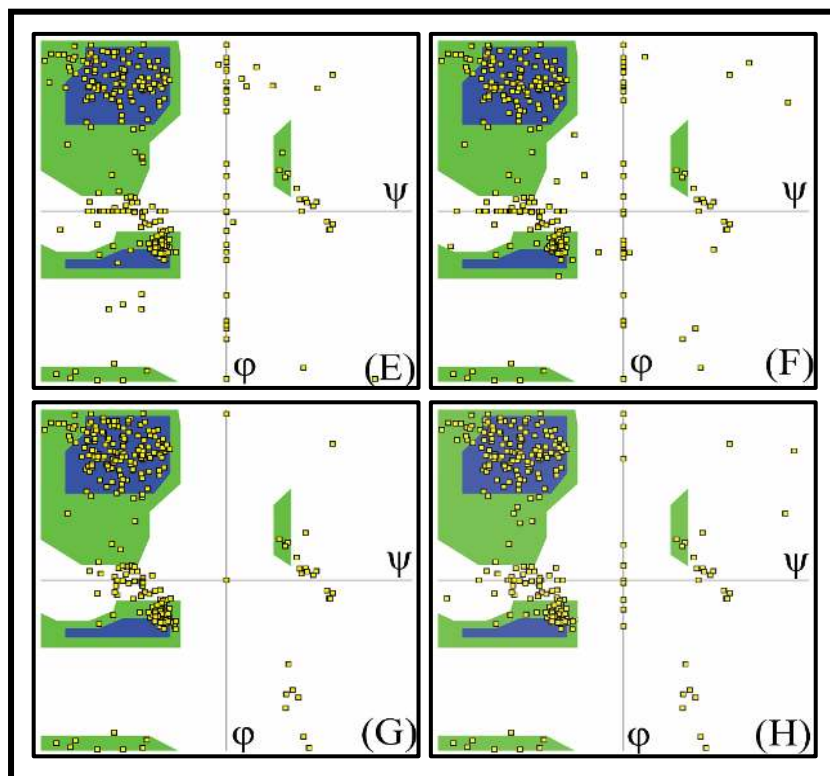Figure 6: Ramachandran Plot for ACE2 (main protease) docked with phytochemical.

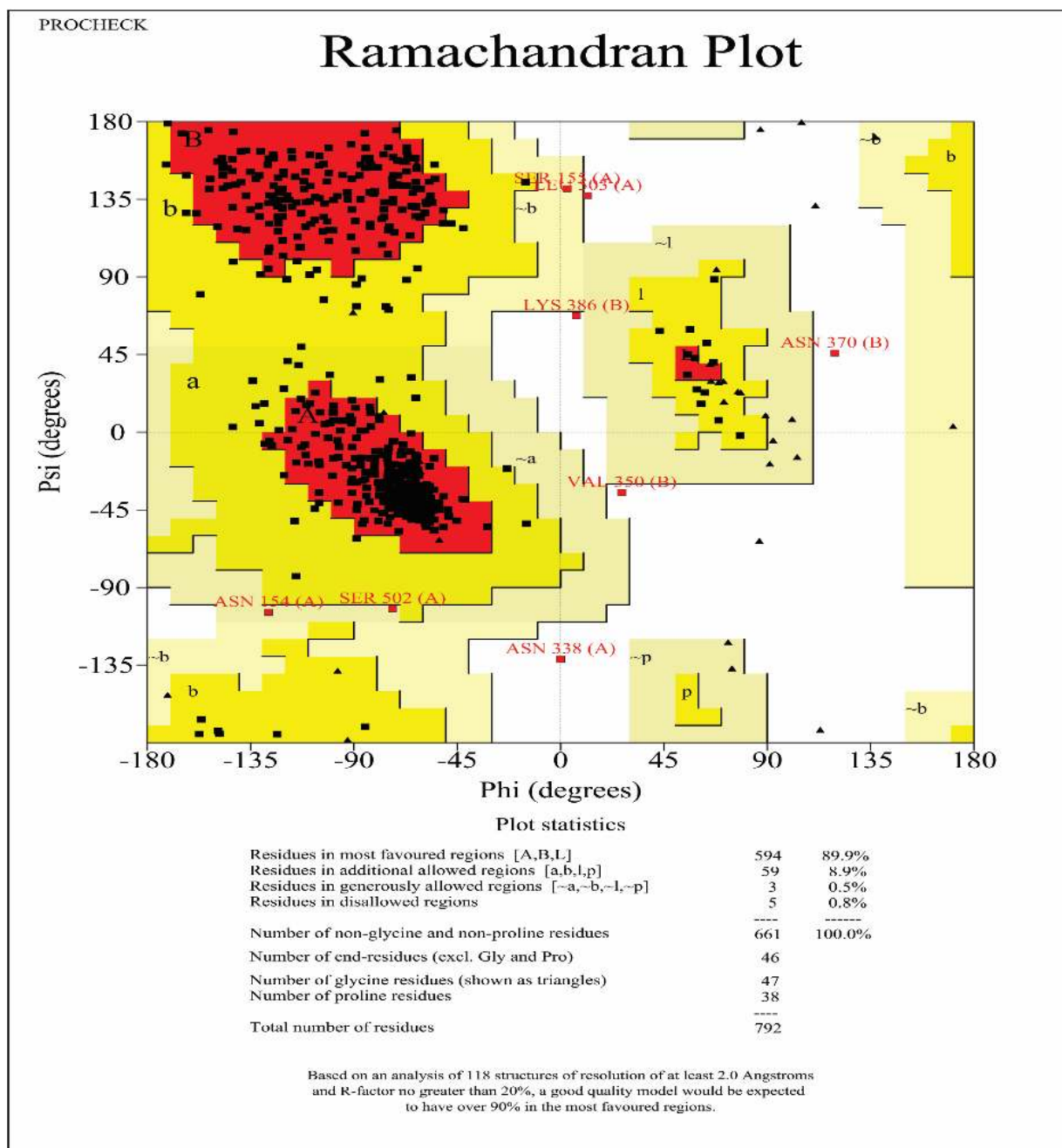Figure 7: Ramachandran Plot for Mpro (main protease) docked with phytochemical.

Figure 8: Validation of simulated system using Ramachandran plot performed with the PROCHECK server

Phytochemicals/ active ingredient		Most Favoured Regions[A,B,L]	Additional Allowed Regions[a,b,l,p],	Generously Allowed Region	Disallowed Regions	G-Factor
Chebulagic Acid	ACE2	89.9%	8.9	0.5	0.8	-0.05
	M pro	88.6	10.1	0.7	0.6	-0.11
	Cryo-EM Structure	83.6	15.2	0.5	0.7	-0.29
	APO form	88.3	10.8	0.4	0.5	-0.18
Pedunculagin	ACE2	87.6	11.3	0.6	0.5	-0.09
	M pro	87.9	11.2	0.4	0.5	-0.17
	Cryo-EM Structure	85.6	13.1	0.7	0.6	-0.20
	APO form	88.3	10.5	0.4	0.8	-0.19
Azadirachtin	ACE2	89.4	9.6	0.4	0.6	-0.14
	M pro	88.9	10.1	0.7	0.3	-0.07
	Cryo-EM Structure	84.2	14.5	0.9	0.4	-0.25
	APO form	88.6	9.9	0.8	0.7	-0.15
Nimbolide	ACE2	89.9	9.4	0.5	0.3	-0.16
	M pro	89.9	9.0	0.5	0.4	-0.08
	Cryo-EM Structure	87.6	11.4	0.3	0.7	-0.22
	APO form	89.7	9.1	0.8	0.4	-0.14

Table 8: Validation report of Ramachandran Pot based on PROCHECK server for Most Favoured Regions[A,B,L], Additional Allowed Regions[a,b,l,p], Generously Allowed Region Disallowed Regions and G-Factors.
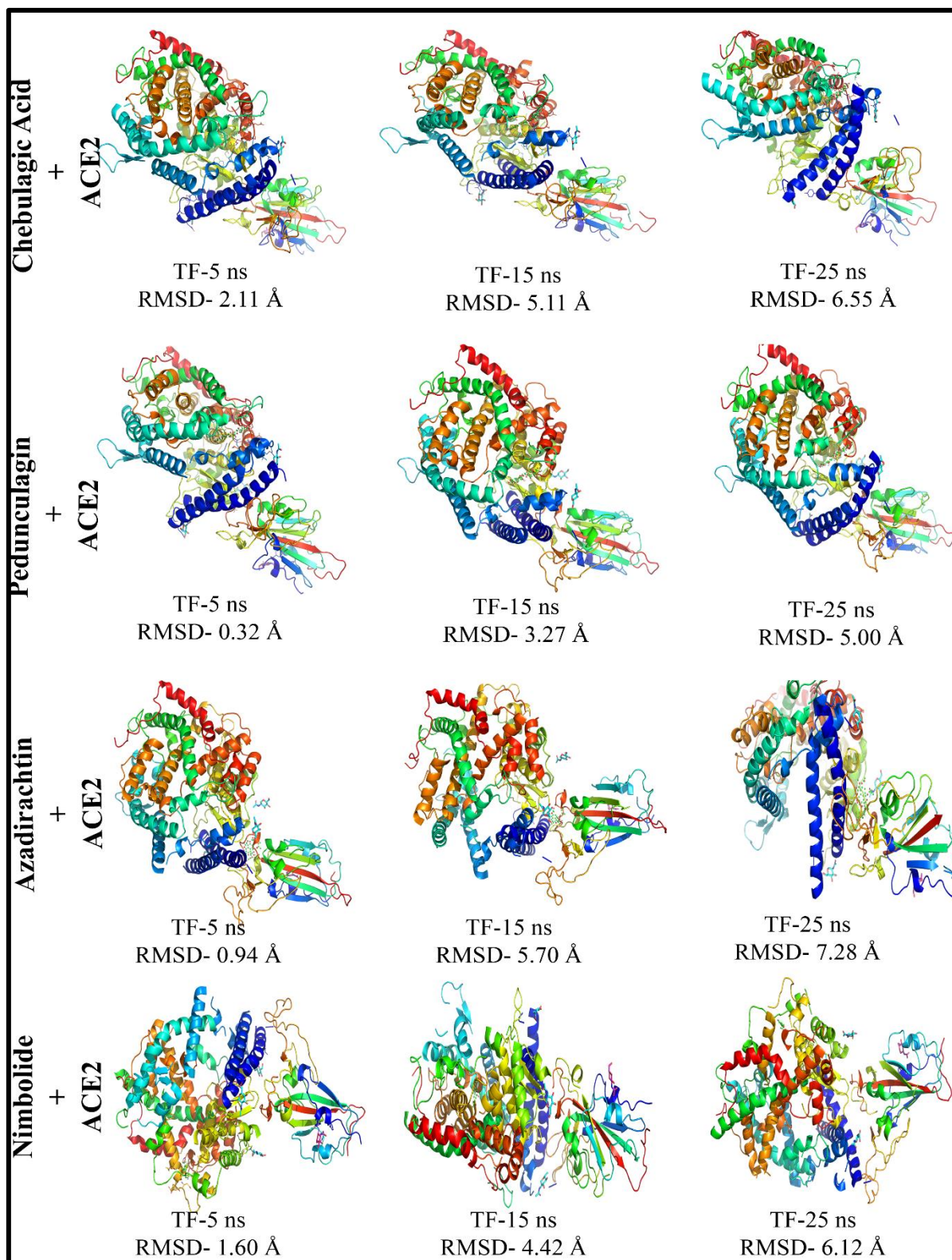
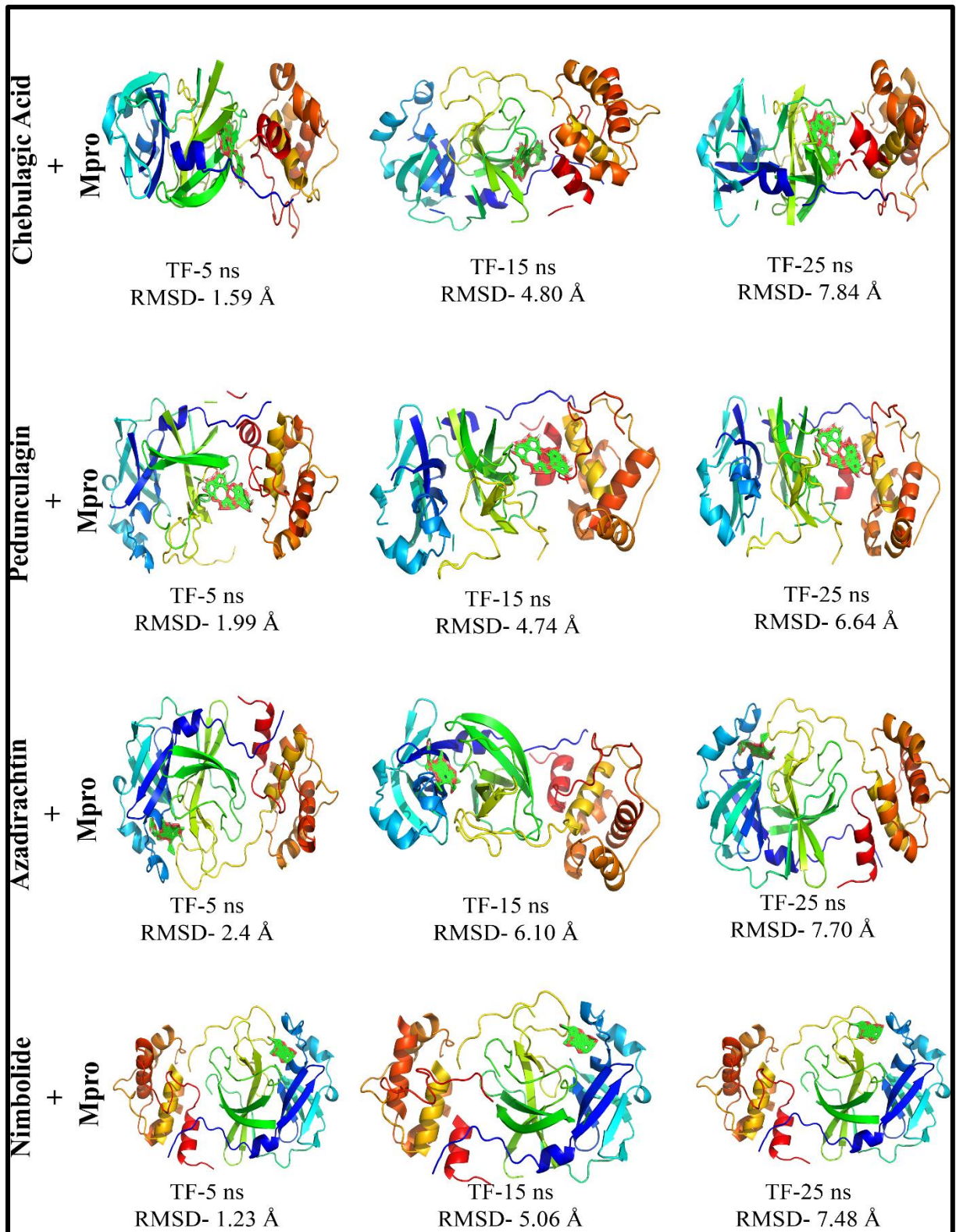Figure 9: Snapshot of SARS-CoV-2 of ACE2 receptor simulated system with phytochemicals

Figure 10: Snapshot of SARS-CoV-2 of Mpro main protease simulated system with phytochemicals

Moreover, in addition to the Ramachandran and rmsd values, the trajectories of amino acid residues in the designated domain sites of ACE2 receptor and Mpro such as (TYR-196, GLN-102, ASN-210, GLN-98,LYS-541, ASP-367, MET-366, GLU-430, ILE-291,HIS-41, ASN-142, THR-190, GLN-192) play a vital role in the binding of phytochemical/active ingredient. The visual structural analysis of Molecular dynamic simulation of 25 ns trajectories proposed that all the computational docked complexes engaged in significant binding interactions with the hotspot residues of the targeted SARS-CoV-2 protein structure (Figure 3A). While analysing the simulated condition, it was observed that Chebulagic Acid exhibits strong and stable significant interactions with the active site residues of all form of SARS-Cov-2 structure (ACE2, Mpro, APO, Cryo-Em). Similarly, Nimbolide showed less fluctuated root mean square deviation and highly strong interaction with main protease and ACE2 receptor in terms of H bond and π-cation interaction. These interactions are probably responsible for the strong affinity of screened phytochemicals causing the inhibition of the main protease and ACE2 receptor. All the interactions of ligand-receptor were highly conserved during the course of simulation with high occupancies and accommodating the ligand to be held at the binding pocket firmly. For detailed analysis, we have also taken the snapshot of each simulated system at every 5ns interval upto 25 ns. (Fig 9, 10).

Looking at these results it seems that there are multiple potential phytochemicals that can act as lead compounds for the development of low-cost nutraceutical against SARS-CoV-2 virus. Therefore, we further stringent our criteria for screening of lead compounds by subjecting them to pharmacokinetics and bioavailability investigation by using SwissADME tool. This tool allowed us to create BOILED-Egg model predicting gastrointestinal absorption and brain access of these phytochemicals, which is very important for drug discovery process. Interestingly after putting all these phytochemicals to pharmacokinetics and bioavailability investigation to create BOILED-Egg model, the number of eligible lead candidates reduced to 3 i.e Apigenin, Nimbolide and Withaferin A. Out of these three herbal compounds, Apigenin can be excluded because of its inability to get effluated from the central nervous system. Withaferin-A can also be excluded since its docking score was not very high compared to others. Finally, only Nimbolide was left with good binding affinity towards both ACE2 receptor and main protease of SARS-CoV-2 as well as excellent pharmacokinetics and bioavailability as per Lipinski rule of 5 (Fig-11) (See detailed Table ESI-Table1), fulfilling the criteria of lead compound that can easily be developed as low cost nutraceutical against SARS-CoV-2 which can be easily taken via oral route because of its high gastrointestinal absorption.
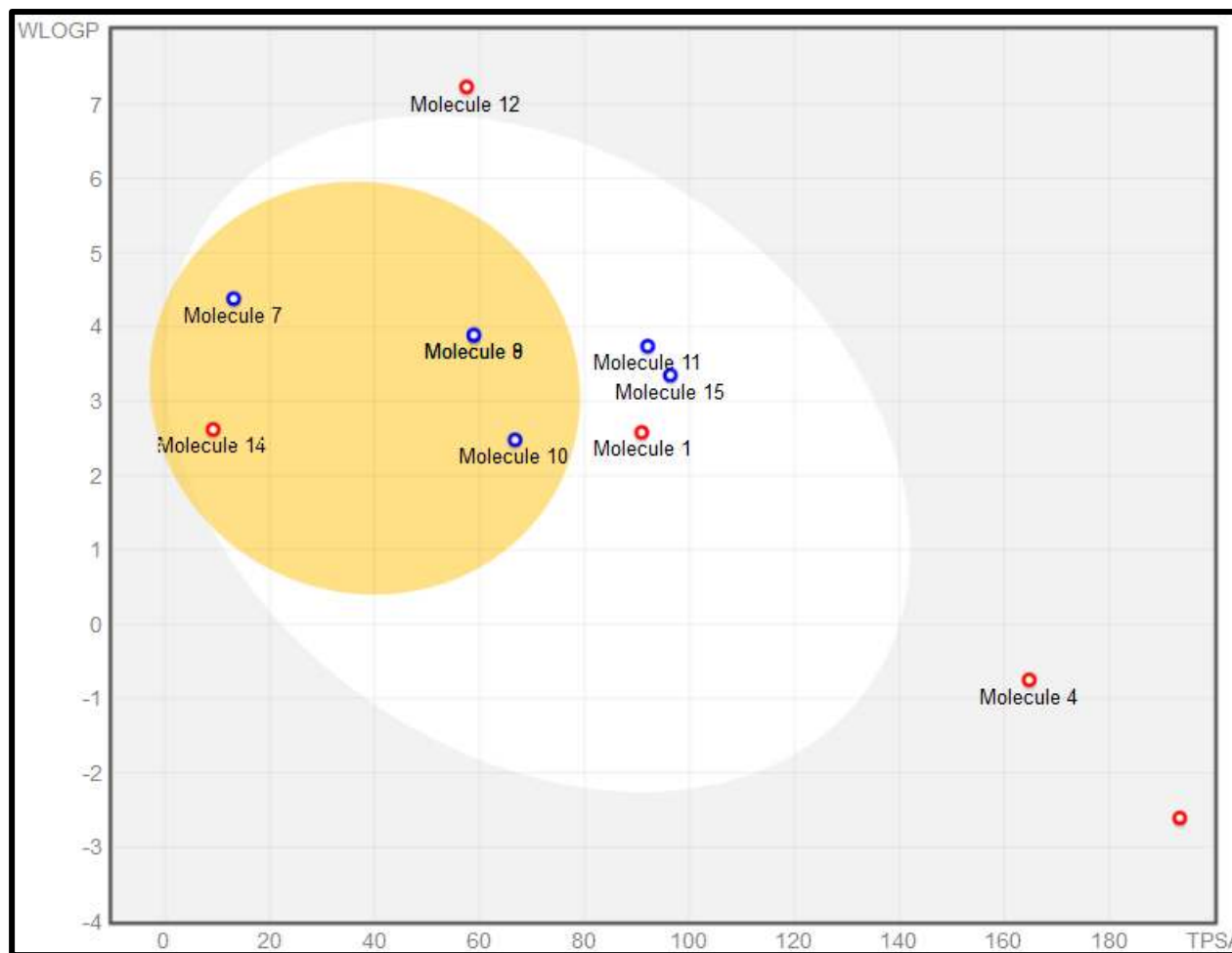
Molecule- 1: Apigenin

Molecule- 2: Azadirachtin

Molecule- 3: Chebulagic acid

Molecule-4: Chlorogenic acid

Molecule-5: Cyanidin 3 glucoside

Molecule-6: Cyanidin 3-sambubioside_Elderberry

Molecule-7: Flavylium

Molecule-8: Glabridin

Molecule-9: Glycirizzin

Molecule-10: Liquiritigenin

Molecule-11: Nimbolide

Molecule-12: Oleanolic acid

Molecule-13: Pedunculagin

Molecule-14: trans-anethole

Molecule-15: Withaferin A

Figure 11: Combined Boiled egg representation (created by using Swiss ADME) of fifteen phytochemicals of herbal origin depicting their pharmacokinetic behaviour and bioavailability based on their gastrointestinal absorption and brain access capability.

In the BOILED-Egg predictive model (Figure 11) white region indicates the physicochemical space of molecules with highest probability of being absorbed by the gastrointestinal route whereas the yellow region (yolk) depicts the physicochemical space of molecules with highest probability to permeate the brain. Red dots represent the molecule predicted not to be effluated from the central nervous system by p-glycoprotein and vice versa for blue dots.

4. Conclusion

We performed molecular docking studies followed by the molecular dynamic simulation to evaluate and compare their energy scores and binding affinity, structural mechanism with SARS-CoV-2 main protease, ACE2 receptor, APO, Cryo-EM structure. Our in-silico study provides fruitful insight about the possible role of structural flexibility and efficacy during interactions between different SARS-CoV-2 motifs and the phytochemicals. We also report the preliminary molecular docking in silico studies with several phytochemicals, which have shown promising anti-SARS-CoV-2 activity. For the treatment of any viral disease, the receptor and protease of virus are considered as the most primary and favourable target. Therefore, we focused our study towards four different SARS-CoV-2 motifs to cover a range of targets. Our in-silico experiment based on molecular docking and molecular dynamic simulation approach, we investigated 15 different types of phytochemicals derived from plants of Asian and European origin, towards their potential application for the development of low cost nutraceuticals against SARS-CoV-2 virus, with less side effects. Our studies give an insight into the possible binding sites, the pocket(s) of ligand- protein interaction, best conformer, its orientation and interactions with the respective targets. Based on their binding affinity, we found that some phytochemicals like Chebulagic Acid, Pedunculagin, Azadirachtin, and Nimbolide compounds may act as potential inhibitors of SARS-CoV-2 ACE2 receptor and main protease by showing promising interactions and most stable ligand bounded protein complex with the novel corona virus. Further we also screened these four phytochemicals for pharmacokinetics and bioavailability by using SwissADME. Taken together our results based on molecular docking, simulation, pharmacokinetics and bioavailability determination due to gastrointestinal absorption and brain access suggest that there is only one herbal compound that fits all the criteria i.e Nimbolide showing great potential for the development of low-cost nutraceuticals against corona virus. Some previous study reports that structural stability of the SARS-CoV-2 Mpro protein with respect to flexible loop mutations indicated that the virus' mutability will further pose a challenge for the development of of small-molecule inhibitors.[31] As per these results our lead compound i.e Nimbolide is not a very small molecule and is also known for its antiviral and immune boosting property. Further it showed a greater potential to bind ACE2 as well as inhibit the Mpro enzymatic activity of SARS-CoV-2 with high binding affinity, stability under simulated condition and excellent pharmacokinetics resulting in enhanced bioavailability in the gastrointestinal tract. This study strongly recommends the significance and importance of phytochemicals of herbal origin for controlling SARS-CoV-2 virus by inhibiting the receptor and protease enzymatic activity. These phytochemical compounds have higher potential to inhibit the SARS-CoV-2 spike protein and the main protease of the virus, which showed stronger binding affinity including the stable complexation with ACE2 receptor and Mpro than remdesivir drug, which has been previously reported a good inhibitor for SARS-CoV–2 in vitro[32]. Our computational study lays the groundwork for computer-aided development of nutraceutical against SARS-CoV-2 with

less side effects in quick time, which is the need of hour. Although our results seem very promising, but it needs to be further validated by both in vitro and in vivo studies accompanied by clinical trials in proper setup.

5. References

1.   Bauch, C. T.; Lloyd-Smith, J. O.; Coffee, M. P.; Galvani, A. P., Dynamically modeling SARS and other newly emerging respiratory illnesses: past, present, and future. *Epidemiology (Cambridge, Mass.)* **2005,** *16* (6), 791-801.

2.   Tian, S.; Hu, W.; Niu, L.; Liu, H.; Xu, H.; Xiao, S. Y., Pulmonary Pathology of Early-Phase 2019 Novel Coronavirus (COVID-19) Pneumonia in Two Patients With Lung Cancer. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer* **2020,** *15* (5), 700-704.

3.   Xu, X.; Chen, P.; Wang, J.; Feng, J.; Zhou, H.; Li, X.; Zhong, W.; Hao, P., Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission. *Science China Life Sciences* **2020,** *63* (3), 457-460.

4.   Li, F., Structure, Function, and Evolution of Coronavirus Spike Proteins. *Annual Review of Virology* **2016,** *3* (1), 237-261.

5.   Boopathi, S.; Poma, A. B.; Kolandaivel, P., Novel 2019 coronavirus structure, mechanism of action, antiviral drug promises and rule out against its treatment. *Journal of Biomolecular Structure and Dynamics* **2020**, 1-10.

6.   Adhikari, S. P.; Meng, S.; Wu, Y.-J.; Mao, Y.-P.; Ye, R.-X.; Wang, Q.-Z.; Sun, C.; Sylvia, S.; Rozelle, S.; Raat, H.; Zhou, H., Epidemiology, causes, clinical manifestation and diagnosis, prevention and control of coronavirus disease (COVID-19) during the early outbreak period: a scoping review. *Infectious Diseases of Poverty* **2020,** *9* (1), 29.

7.   Walls, A. C.; Park, Y.-J.; Tortorici, M. A.; Wall, A.; McGuire, A. T.; Veesler, D., Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **2020,** *181* (2), 281-292.e6.

8.   Osakwe, O., Chapter 5 - The Significance of Discovery Screening and Structure Optimization Studies. In *Social Aspects of Drug Discovery, Development and Commercialization*, Osakwe, O.; Rizvi, S. A. A., Eds. Academic Press: Boston, 2016; pp 109-128.

9.   Tomar, V.; Mazumder, M.; Chandra, R.; Yang, J.; Sakharkar, M. K., Small Molecule Drug Design. In *Encyclopedia of Bioinformatics and Computational Biology*, Ranganathan, S.; Gribskov, M.; Nakai, K.; Schönbach, C., Eds. Academic Press: Oxford, 2019; pp 741-760.

10.   Ben-Shabat, S.; Yarmolinsky, L.; Porat, D.; Dahan, A., Antiviral effect of phytochemicals from medicinal plants: Applications and drug delivery strategies. *Drug Delivery and Translational Research* **2020,** *10* (2), 354-367.

11.   Mani, J. S.; Johnson, J. B.; Steel, J. C.; Broszczak, D. A.; Neilsen, P. M.; Walsh, K. B.; Naiker, M., Natural product-derived phytochemicals as potential agents against coronaviruses: A review. *Virus Research* **2020,** *284*, 197989.

12.   Yan, R.; Zhang, Y.; Li, Y.; Xia, L.; Guo, Y.; Zhou, Q., Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* **2020,** *367* (6485), 1444.

13.   Wan, Y.; Shang, J.; Graham, R.; Baric, R. S.; Li, F., Receptor Recognition by the Novel Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS Coronavirus. *Journal of Virology* **2020,** *94* (7), e00127-20.

14.   Shang, J.; Ye, G.; Shi, K.; Wan, Y.; Luo, C.; Aihara, H.; Geng, Q.; Auerbach, A.; Li, F., Structural basis of receptor recognition by SARS-CoV-2. *Nature* **2020,** *581* (7807), 221-224.

15.   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,** *46* (4), 586-590.

16.   Lan, J.; Ge, J.; Yu, J.; Shan, S.; Zhou, H.; Fan, S.; Zhang, Q.; Shi, X.; Wang, Q.; Zhang, L.; Wang, X., Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **2020,** *581* (7807), 215-220.

17.   Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C.; Duan, Y.; Yu, J.; Wang, L.; Yang, K.; Liu, F.; Jiang, R.; Yang, X.; You, T.; Liu, X.; Yang, X.; Bai, F.; Liu, H.; Liu, X.; Guddat, L. W.; Xu, W.; Xiao, G.; Qin, C.; Shi, Z.; Jiang, H.; Rao, Z.; Yang, H., Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. *Nature* **2020,** *582* (7811), 289-293.

18.   Mengist, H. M.; Fan, X.; Jin, T., Designing of improved drugs for COVID-19: Crystal structure of SARS-CoV-2 main protease Mpro. *Signal Transduction and Targeted Therapy* **2020,** *5* (1), 67.

19.   Daina, A.; Zoete, V., A BOILED-Egg To Predict Gastrointestinal Absorption and Brain Penetration of Small Molecules. *ChemMedChem* **2016,** *11* (11), 1117-21.

20.   Xiao, J.; Bai, W., Bioactive phytochemicals. *Critical Reviews in Food Science and Nutrition* **2019,** *59* (6), 827-829.

21.   Schmitz, H.; Chevaux, K., Defining the Role of Dietary Phytochemicals in Modulating Human Immune Function. In *Nutrition and Immunology: Principles and Practice*, Gershwin, M. E.; German, J. B.; Keen, C. L., Eds. Humana Press: Totowa, NJ, 2000; pp 107-119.

22.   Trott, O.; Olson, A. J., AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* **2010,** *31* (2), 455-461.

23.   Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J., AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem* **2009,** *30* (16), 2785-2791.

24.   Forli, S.; Huey, R.; Pique, M. E.; Sanner, M. F.; Goodsell, D. S.; Olson, A. J., Computational protein–ligand docking and virtual drug screening with the AutoDock suite. *Nature Protocols* **2016,** *11* (5), 905-919.

25.   Daina, A.; Michielin, O.; Zoete, V., SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci Rep* **2017,** *7* (1), 42717.

26.   Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E., GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015,** *1-2*, 19-25.

27.   Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D., Jr., CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat Methods* **2017,** *14* (1), 71-73.

28.   Petersen, H. G., Accuracy and efficiency of the particle mesh Ewald method. *The Journal of Chemical Physics* **1995,** *103* (9), 3668-3679.

29.   Elber, R.; Ruymgaart, A. P.; Hess, B., SHAKE parallelization. *Eur Phys J Spec Top* **2011,** *200* (1), 211-223.

30.   Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M., PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography* **1993,** *26* (2), 283-291.

31.   Bzówka, M.; Mitusińska, K.; Raczyńska, A.; Samol, A.; Tuszyński, J. A.; Góra, A., Structural and Evolutionary Analysis Indicate that the SARS-CoV-2 Mpro is an Inconvenient Target for Small-Molecule Inhibitors Design. *bioRxiv* **2020,** 2020.02.27.968008.

32.   Zhou, Y.; Hou, Y.; Shen, J.; Huang, Y.; Martin, W.; Cheng, F., Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2. *Cell Discovery* **2020,** *6* (1), 14.