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

Journal of Molecular Graphics and Modelling 114 (2022) 108192

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/jmgm

Identification of phytocompounds as newer antiviral drugs against COVID-19 through molecular docking and simulation based study

Bipasa Kar [a], Budheswar Dehury [a], Mahender Kumar Singh [b], Sanghamitra Pati [a], Debdutta Bhattacharya [a,*]

[a] ICMR-Regional Medical Research Centre (Dept. of Health Research, Ministry of Health & Family Welfare, Govt. of India), Chandrasekharpur, Bhubaneswar, 751023, Odisha, India

[b] Data Science Laboratory, National Brain Research Centre, Gurgaon, Haryana, 122052, India

ARTICLE INFO

Keywords:
Molecular docking
ADMET analysis
Molecular dynamic simulation
Phytocompounds
COVID-19

ABSTRACT

COVID-19 pandemic has emerged as a global threat with its highly contagious and mutating nature. Several existing antiviral drugs has been worked on, without proper results and meanwhile the virus is mutating rapidly to create more infectious variant. In order to find some alternatives, phytocompounds can be opted as good one. In this study, three hundred phytocompounds were screened virtually against two viral proteins namely main protease and spike protein. Molecular docking and dynamic simulation study was used to find binding affinity, structural stability and flexibility of the complex. Pharmacokinetic properties were studied through ADMET analysis. To understand energy variation of the complex structure free energy landscape analysis was performed. Among three hundred phytocompounds virtual screening, three phytocompounds were selected for detailed molecular interaction analysis. Oleanderolide, Proceragenin A and Balsaminone A, showed strong binding affinity against both the target proteins and reflected conformational stability throughout the MD run. Oleanderolide, proceragenin A and balsaminone A has docking score −9.4 kcal/mol, −8.6 kcal/mol, and −8.1 kcal/mol respectively against main protease and same −8.3 kcal/mol docking score against spike protein. These three phytocompounds has high gastrointestinal absorption capacity. They were unexplored till now for their antiviral activity. Their promising in silico results suggests that they can be promoted in the long run for development of new antiviral drugs.

1. Introduction

The coronavirus disease (COVID-19) is a respiratory illness disease caused by SARS-CoV2, was first reported in Wuhan, China in December 2019 [20]. Since then more than 4 million people were died because of the disease and as of 3rd September 2021 the number of active cases around the world is more than 2 million (https://www.worldometers.info/coronavirus/).

This worldwide spreading novel virus is named SARS-CoV-2 belongs to the family coronaviridae. It has a genomic similarity of more than 80% with previously reported SARS-CoV virus. The coronavirus family (CoV) has single-stranded positive sense RNA as genetic material which is further surrounded by crown like spike protein (in Latin corona) [23]. The spike protein is responsible for viral entry into the host cell.

It is believed that bats are the most likely natural host for CoVs. Historically it was believed, that coronavirus was causing respiratory and enteric illness only in farm animals, but the emerging coronaviruses (beta-coronavirus) are proved to be pathogenic to humans as well, causing severe respiratory and gastrointestinal symptoms [3]. Unlike the previous pathogenic CoVs, the global pandemic causing SARS-CoV2 has a range of 0.2–9% with an average of 2% case fatality ratio, which also varied in different countries, but the highly contagious nature of the virus had made it the most consequential global health crisis. The pandemic has exposed the bottlenecks in global healthcare system to handle sudden outbreaks and the global economic structure is also shattered because of the prolonged shutdown [13].

The most common symptoms of COVID-19 are shortness of breath, cough, fever, fatigue, body ache, and in extreme cases it may result in lungs infection, kidney and multi-organ failure. Cohort studies have demonstrated that patients with co-morbidity have increased risk of

* Corresponding author. Scientist-D, Dept. of Microbiology, ICMR-Regional Medical Research Centre, (Dept. of Health Research, Ministry of Health & Family Welfare, Govt. of India), Chandrasekharpur, Bhubaneswar, 751023, India.

E-mail addresses: drdebdutta.bhattacharya@yahoo.co.in, drdebduttab.rmrc-od@gov.in (D. Bhattacharya).

https://doi.org/10.1016/j.jmgm.2022.108192
Received 19 November 2021; Received in revised form 7 April 2022; Accepted 8 April 2022
Available online 14 April 2022

Table 1
List of various interaction and interacting residues in the active site of Mpro and RBD with potential natural compounds, were listed from docked complexes.

Sl No	Compound Name	Main protease (6lu7)			RBD (6M0J)		
		Docking score (kcal/mol)	H bond	Hydrophobic interactions	Docking Score (kcal/mol)	H bond	Hydrophobic interactions
1.	Flavanomarein	−8	T^{29},T^{190},Q^{192}, Q^{189},R^{188}	L^{27},L^{141},G^{143}, E^{166},M^{165}, N^{142}	−7.5	C^{336},G^{339},F^{436}	N^{343},F^{338},F^{342},L^{368}, S^{371},L^{335},V^{367}
2.	Pindrolactone	−8.3	R^{131},D^{289},T^{199}	M^{270},G^{275},Y^{239}, L^{271},L^{287}, L^{280}	−8.1	W^{436}	F^{374},S^{371},S^{373}
3.	Cyanidin-3-glucoside	−8.3	T^{26}, Y^{54}, F^{140}, L^{141}, S^{144},C^{145}, H^{163}	N^{142},E^{166},M^{165},D^{187},Q^{189}, R^{188},G^{143},M^{49},H^{41},L^{27}	−7.3	R^{355},D^{428},T^{430}, P^{463},F^{525}	P^{426},F^{429},S^{514},F^{464}
4.	Diosgenin	−8.3	S^{144},L^{141}	A^{191},P^{168},D^{166},N^{142},H^{163}, C^{145},M^{165},Q^{189},T^{190}	−8.1	–	V^{362},D^{364},L^{335},V^{367}, C^{336},A^{363}, L^{368},F^{342}, S^{371},N^{343},S^{373}
5.	β-Amyrin	−7.9	K^{137}	D^{197},T^{198},N^{238},T^{199},Y^{237}, L^{272},L^{271},L^{287}, Y^{239}	−8.2	–	V^{367},G^{339},F^{342},N^{343},W^{436},L^{441}, S^{371},S^{373},F^{374},L^{368}
6.	Fumarprotocetraric acid	−8.3	G^{143},C^{145},E^{166},D^{187}	L^{141},H^{164},N^{142},Q^{189},M^{165}, H^{41},P^{52},M^{49},Y^{54},C^{44},R^{188}	−7.6	F^{347},S^{349},Y^{351}, N^{354},N^{450}	A^{352},Y^{451},A^{348},R^{346}, DV^{341},S^{399}, A^{344}
7.	α - Amyrin	−8	K^{137}, D^{197}	L^{271},L^{286},L^{287},L^{272},Y^{237}, T^{198},T^{199},Y^{239}	−8.5	–	V^{367},L^{368},G^{339},S^{371},N^{343},W^{436}, S^{373},N^{440},L^{441},F^{374},F^{342}
8.	Bryonolic acid	−7.9	K^{137},D^{289},M^{276}	R^{131},T^{199},Y^{239},L^{272},L^{287}, L^{286},G^{275}	−8.2	T^{430},F^{515}	R^{355},P^{463},F^{464},P^{426},F^{429},S^{514},D^{516}
9.	Citrullonol	−9.1	R^{105}	V^{104},$D1^{53}$,I^{152},F^{294},N^{151}, Q^{110},F^{8},S^{158}	−8.1	F^{515}	D^{428},P^{426},T^{430},S^{524},F^{464},G^{516},R^{355}
10.	Quassinoid analog	−8.4	T^{111},N^{151},S^{158}	I^{152},D^{153},I^{106},Q^{110},V^{104}, F^{294}	−7.3	–	D^{428}, P^{426}, T^{430},S^{514},F^{464},F^{515},E^{516}, R^{355},Y^{396}
11.	Astragalin	−8.4	Y^{54},L^{141},S^{144},H^{163}, D^{187},R^{188},T^{190}, Q^{192}	P^{168},E^{166},E^{89},H^{41},C^{145}, M^{165},F^{140}	−6.7	E^{516},L^{517},F^{515}, P^{463}	L^{518},F^{464},P^{426},D^{428},F^{429},T^{430},S^{514}
12.	Cynaroside	−8.1	T^{24},T^{25},T^{26},G^{143}, S^{144},R^{188}	M^{49},T^{45},Q^{189},E^{166},D^{187}, M^{165},H^{164},N^{142},C^{145},H^{41}	−7.4	R^{355},T^{430},R^{466}, F^{515}	D^{516},F^{464},P^{463},F^{429},S^{514},F^{429},W^{353}
13.	Epicatechin gallate	−8.2	F^{464},S^{514},F^{429},P^{426}	R^{355},F^{515},T^{430},D^{428},P^{463}	−7.0	E^{166},F^{140}	H^{172},H^{163},N^{142},S^{144},H^{164},M^{165}, R^{188},G^{143},L^{141},T^{26},H^{41},L^{27},D^{187}, Q^{189},T^{25},M^{29}
14.	Astilbin	−8.4	N^{142},G^{143},Y^{54},E^{166}, D^{187}	M^{49},H^{41},R^{188},Q^{189},C^{145}, M^{165},L^{141}	−6.9	T^{430},P^{463}	F^{464},D^{428},P^{426},S^{514},F^{525},F^{429}
15.	Baicalin	−8.3	T^{24},T^{25},T^{26},C^{44}, S^{144},C^{145}	S^{46},T^{45},G^{143},N^{142},M^{165}, Q^{189},M^{49}	−7.7	E^{340},A^{348},N^{354}	A^{352},R^{346},F^{347},S^{399},A^{344},V^{341}
16.	Guajavrin	−8.3	L^{141},S^{144},C^{145}, H^{163},E^{166}	H^{41},D^{187},N^{142},M^{165},Q^{189}, R^{188},G^{143},M^{49}	−7.1	N^{343},A^{344},S^{373}, R^{509}	S^{371},W^{436},N^{440},F^{342},F^{374},L^{441}
17.	**Balsaminone A**	**−8.1**	**L^{287}**	**L^{272},T^{199},Y^{237},Y^{239},R^{131}, L^{286},K^{137},D^{287}**	**−8.3**	**-**	**S^{373},F^{374},W^{436},S^{371},F^{338},N^{343}, F^{342},V^{367},L^{368},G^{339}**
18.	Canophyllic acid	−8	R^{131},M^{276}	A^{285},$L2^{71}$,T^{199},L^{287},K^{137}, D^{286},D^{289},G^{275}	−8.2	T^{430},D^{428}	R^{355},F^{464},P^{426},S^{514},F^{515},F^{429},P^{463}
19.	**Oleanderolide**	**−9.4**	**L^{271},L^{272},L^{287}**	**D^{289},L^{286},Y^{237},Y^{239},T^{199}, G^{275}**	**−8.3**	**K^{378}**	**G^{504},V^{503},D^{405},R^{408},V^{407},T^{376}, G^{404}**
20.	Oleanolic acid	−8.5	R^{131},D^{289}	K^{137},L^{287},L^{286},Y^{237},Y^{239}, T^{199},L^{272}	−8.2	S^{373}	V^{367},S^{371},L^{368},F^{342},N^{343},F^{374},W^{436}, L^{441}
21.	**Proceragenin A**	**−8.6**	**L^{271}**	**Y^{237},Y^{239},L^{272},T^{199},L^{287}, L^{286},K^{137}**	**−8.3**	**R^{355},T^{430}**	**D^{428},F^{429},D^{516},L^{517},Y^{396},F^{464}**
22.	Vitexin	−8	Y^{54},D^{187}	H^{41},M^{49},H^{164},M^{165},D^{166}, R^{188},Q^{189},T^{190}	−7.2	R^{355},F^{515}	T^{430},P^{426},P^{463},F^{429},F^{464},Y^{396}
23.	Betulic acid	−7.7	T^{199},N^{238}	L^{271},L^{272},L^{286},L^{287},D^{197}, R^{131},T^{198},Y^{237}	−7.4	Y^{508}	A^{411},Y^{380},G^{404},T^{376},^{407},S^{375},V^{503}, D^{405},R^{408}
24.	Avicularin	−8.8	Y^{54},L^{141},N^{142},E^{166}, D^{187}	G^{143},Q^{189},H^{41},F^{140},C^{145}, R^{188},M^{49},H^{163},L^{27},H^{172}, M^{165}	−7.2	D^{428},T^{430},P^{463}, F^{515},E^{516}	R^{355},Y^{396},F^{464},F^{429},S^{514},P^{426}
25.	Ursolic acid	−7.7	T^{199},M^{276}	G^{275},A^{285},N^{238},K^{137},Y^{239}, L^{287}	−8.3	S^{373}	N^{343},L^{441},N^{440},L^{368},G^{339},F^{342},F^{338}, W^{436},S^{371},F^{374},V^{367}
26.	Subulin	−7.9	T^{24},T^{26},S^{46},G^{143}, E^{166},T^{190}	T^{25},M^{49},N^{142},Q^{189},P^{168}, R^{188},Q^{192},H^{163},M^{165}	−7.8	L^{335},C^{336},G^{339}, N^{343},S^{371},S^{373}	F^{374},L^{368},V^{367},F^{338},P^{337},W^{436},L^{441}, F^{342}
27.	Silybin	−7.9	G^{143},H^{164},T^{190}	M^{165},E^{166},S^{144},C^{145},L^{141}, N^{142},P^{168},Q^{189},R^{188},H^{41}	−7.6	D^{428},T^{430},R^{466}, F^{515}	E^{465},F^{464},R^{355},E^{516},S^{514},P^{426}
28.	Hexanorcurcubitacin	−7.7	T^{111}	Q^{110},N^{151},F^{294},S^{158},I^{106}, V^{104}	−8.8	R^{355},T^{430},S^{514}	F^{429},Y^{396},E^{516},D^{428},P^{426},F^{464}
29.	Aromadendrin	−7.6	Y^{54},L^{141},G^{143},S^{144}, E^{166},D^{187}	N^{142},C^{145},H^{164},H^{41},R^{188}, M^{49},Q^{189},M^{165}	−6.5	R^{355},T^{430},S^{517}	F^{464},P^{426},E^{516},F^{515}
30.	Clausenin	−6.8	Q^{107}, S^{158}	F^{8},I^{106},V^{104},Q^{110},N^{151}, D^{153},F^{294}	−6.9	–	^{F342},N^{343},S^{371},F^{374},L^{368},W^{436},L^{441}
31.	Ritonavir	−6.4	R^{131},Y^{237}	N^{238},D^{289},T^{199},T^{198},D^{197}, L^{272},Y^{239},L^{286},L^{287}	−6.1	S^{373},N^{343},G^{339}	E^{340},P^{337},L^{335},V^{367},D^{364},C^{336},L^{368}, F^{342},F^{338}
32.	Favipiravir	−5.7	Q^{110},N^{151}	F^{294},T^{111},T^{292}	−5.3	A^{348},S^{399}	V^{341},R^{346},E^{340},A^{344},N^{354},F^{347}

Table 2
ADMET analysis of the phyto-compounds employed in this study.

Sl No	Compound name	Chemical formula	nVio	Hepatotoxicity	LD50 (mol/kg)	GI absorption (%)	BBB	SA	Bioavailability Score
1	Flavanomarein	C21H22O11	2 vio	no	2.73	40.3	no	5.06	0.17
2	Pindrolactone	C30H42O3	1 vio	yes	1.84	95.8	no	6.33	0.55
3	Cyanidin-3glucoside	C21H21ClO11	2 vio	no	2.54	29.9	no	5.3	0.17
4	Diosgenin	C27H42O3	1 vio	no	1.92	96.5	yes	6.94	0.55
5	Aromadendrin	C15H12O6	No vio	no	2.16	59.07	no	3.42	0.55
6	Beta-amyrin	C30H50O	1 vio	no	2.47	93.7	no	6.04	0.55
7	Fumarprotocetraric acid	C22H16O12	1 vio	no	2.43	28.7	no	3.87	0.11
8	Alpha.-Amyrin	C30H50O	1 vio	no	2.46	94.07	no	6.17	0.55
9	Bryonolic acid	C30H48O3	1 vio	yes	2.58	98.1	no	5.91	0.85
10	Citrullonol	C30H42O2	1 vio	yes	1.94	96.0	no	6.62	0.55
11	Quassinoid analog	C27H38O7	No vio	yes	4.04	80.0	no	6.77	0.55
12	Clausenin	C14H12O5	No vio	yes	2.01	95.3	yes	2.89	0.55
13	Astragalin	C21H20O11	2 vio	no	2.54	48.05	no	5.29	0.17
14	Cynaroside	C21H20O11	2 vio	no	2.54	37.5	no	5.42	0.55
15	(−)-Epicatechingallate	C22H18O10	2 vio	no	2.55	62.09	no	4.16	0.55
16	Astilbin	C21H22O11	2 vio	no	2.58	49.0	no	5.27	0.17
17	Baicalin	C21H18O11	2 vio	no	2.63	26.2	no	5.09	0.11
18	Guajavarin	C20H18O11	No vio	no	2.58	51.8	no	5.05	0.17
19	Balsaminone A	C21H12O5	No vio	no	2.43	95.60	no	3.44	0.55
20	Canophyllic acid	C30H50O3	1 vio	yes	2.26	100	no	5.27	0.85
21	Oleanderolide	C30H48O4	1 vio	no	2.25	92.9	no	6.64	0.55
22	Oleanolic acid	C30H48O3	1 vio	yes	2.34	99.9	no	6.08	0.85
23	ProcerageninA	C30H46O4	1 vio	no	2.15	91.9	no	6.92	0.55
24	Vitexin	C21H20O10	2 vio	no	2.59	46.69	no	5.12	0.55
25	Avicularin	C20H18O11	1 vio	no	2.54	57.2	no	5.63	0.85
26	Betulic acid	C30H48O3	1 vio	yes	2.25	99.76	no	5.04	0.17
27	Ursolic acid	C30H48O3	No vio	yes	2.34	100	no	6.21	0.85
28	Hexanorcucurbitacin I	C24H32O5	3 vio	no	2.33	90.7	no	5.74	0.55
29	Subulin	C28H32O16	No vio	no	2.52	32.6	no	6.49	0.17
30	Silybin	C25H22O10	2 vio	no	2.55	61.8	no	4.92	0.55

death [38]. In this extended period of the pandemic, COVID-19 has also developed several variants with increased transmissibility or virulence which has increased concern for the medical world. Although several studies were undertaken to find inhibitors against COVID-19, but none of them demonstrated much effectiveness.

In search for inhibitors against COVID-19, computer aided drug discovery (CADD) has emerged as a powerful tool in the drug discovery process, which reduces the time of screening of potential hit compounds in the task of finding lead drug candidates. In this study, we have selected the main protease (Mpro) protein because it is highly conserved in coronaviruses and plays important role in viral replication. Besides that, no proteases have been reported in humans with the same cleavage sites, so finding an inhibitor against Mpro is safe for human [37]. Spike protein has been chosen for another drug target because of its role in both receptor recognition (through S1 subunit) and membrane fusion (through S2 subunit). Spike protein binds with the acetyl-choline esterase (ACE2) receptor in the host cell and their increased affinity is the reason for the contagion nature of COVID-19 [18].

In this study, we have targeted two major viral proteins (Mpro and spike) and performed virtual screening; molecular docking and all-atoms molecular dynamics simulations with phytocompounds reported from twenty-five Indian medicinal plants and performed a detailed interaction study, which may lead us to select the potent compounds for further the clinical trial.

2. Materials and methods

To start with the virtual screening process, the crystallographic structures of the target proteins (6LU7 and 6M0J) were obtained from RCSB Protein Data Bank (PDB). COVID-19 main protease (Mpro) with PDB ID- 6LU7 and spike receptor binding domain (RBD) with PDB ID-6M0J were downloaded for further experimental analysis.

2.1. Receptor protein preparation

The X-ray crystallographic structures of the target proteins were selected based on their Resolution (6LU7:2.16A°, 6M0J: 2.45A°), R-value (6LU7:0.235A°, 6M0J: 0.227A°), and no mutation in the structure [19]. The protein structures were first checked through Swiss PDB viewer for any missing atoms, then further optimized using AutoDock tool by deleting all water molecules, heteroatom's and inhibitors followed by addition of polar H atom, Kollman charge and computation of Gasteiger charge. Finally, the 3D coordinates of the optimized protein structures bearing partial charges were saved in PDBQT format for the next step of virtual screening using AutoDock Vina [35].

2.2. Preparation of ligands

In consideration of the significant importance of Indian medicinal plants in controlling diseases, we have selected twenty-five commonly found Indian medicinal plants for screening their activity against COVID-19. We have searched phytocompounds from the respective plants through the IMMPAT database which is the largest repository of Indian medicinal plants [25]. The three-dimensional structures were saved as SDF files, which were further processed through PyMOL to convert into PDB files and optimized using AutoDock tools. For the compounds with only two-dimensional structures (available in PubChem), the three-dimensional structures were created through VegaZZ software and followed the previous process of optimization [27].

2.3. Docking of phytocompounds against target proteins

Molecular docking was performed to assess the binding energy of the protein-ligand interaction and to provide initial coordinates for the molecular dynamic simulation. The saved PDBQT proteins files were separately incorporated into AutoDock tools 1.5.6. The information regarding active site residues was collected from a previous literature study [31]. The active site residues were marked to draw the grid box/search area with the spacing of 0.375 Å. For main protease (Mpro), a grid box with dimension 66X66X66 $Å^3$ was centred at the coordinates X = −13.099, Y = 15.875, Z = 70.369 and for the spike protein RBD (6M0J) a grid box with dimension 50X50X50 $Å^3$was centred at the

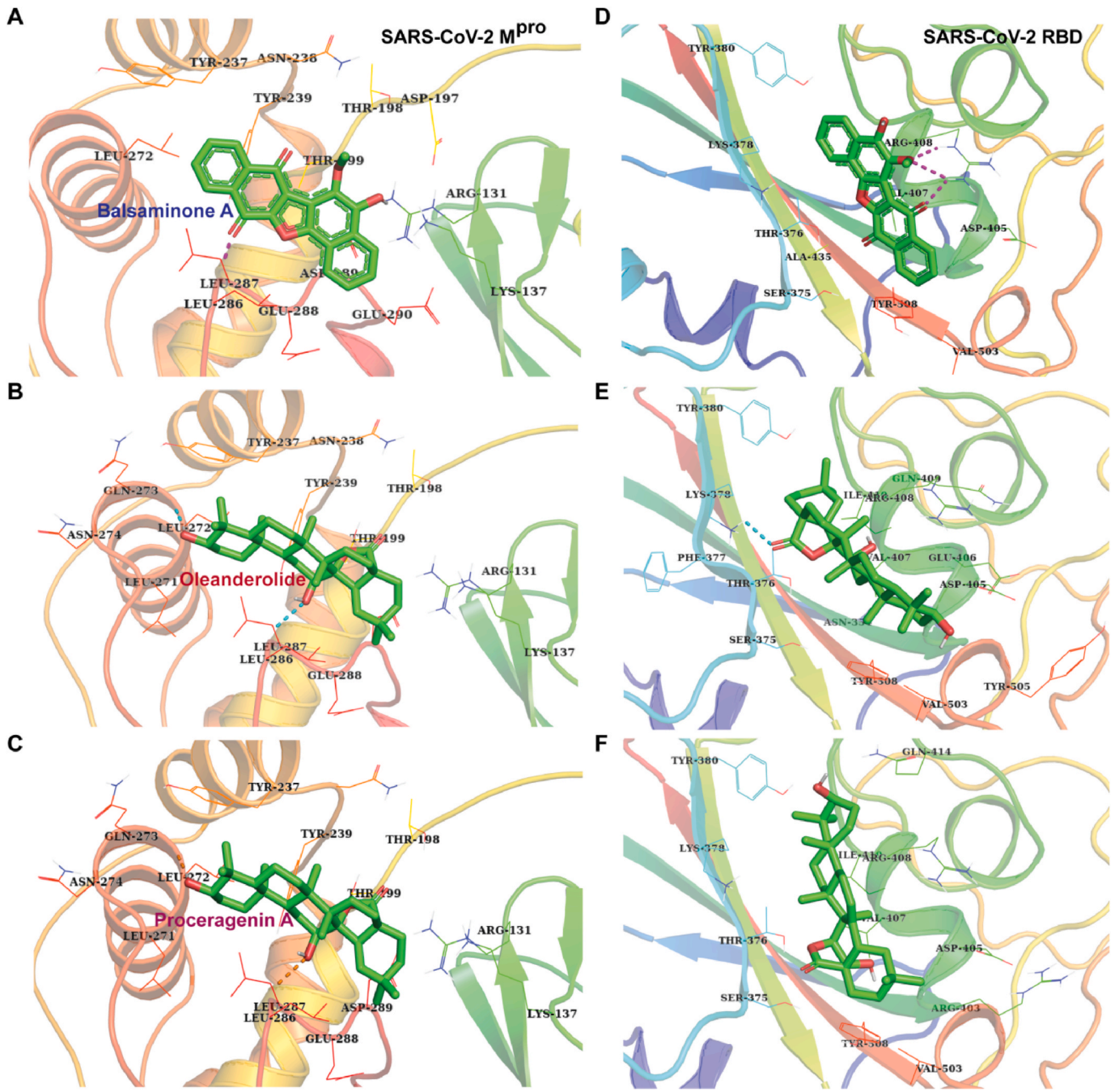

Fig. 1. Non-bonded interaction of phytocompounds with SARS-CoV-2 M^{Pro} and RBD obtained after docking with Autodock Vina [**A**: Mpro-Balsaminone A; **B**: Mpro-Oleanderolide; **C**: Mpro-Proceragenin A; **D**: RBD-Balsaminone A; **E**: RBD-Oleanderolide; **F**: RBD-Proceragenin A]. The protein is shown in cartoon representation and the ligands are shown in stick format while interacting amino acids (shown in stick format) are labelled. The dotted lines indicate the hydrogen bonds.

coordinates X = −32.231, Y = 15.038, Z = 28.403. For both the proteins, the selected ligands were incorporated individually and with the in-house scripts that were incorporated in the Autodock Vina program, a virtual screening process was performed. The 300 phytocompounds from twenty-five plants were screened and ranked according to the binding energies. The docking results were confirmed by re-docking the ligands with corresponding proteins. After docking, the structures were visualized using PyMOL (www.pymol.org). Antiviral drugs Flavipiravir, Ritonavir were used as reference standards for selecting phytocompounds [2]. For best 1% (30) compounds, docking scores and interactive amino acid residues were listed in Table 1.

2.4. ADMET calculation

The previously selected thirty phytocompounds were also evaluated for their ADMET (Absorption, Distribution, Metabolism, Excretion, Toxicity) properties through SwissADME and pkCSM server [10,28]. Lipinski Rule of five (R05) violation, GI absorption, BBB permeability, hepatotoxicity, oral rat acute toxicity (LD50), synthetic and bioavailability scores were into consideration for selection of the compounds for MD simulation. The pharmacokinetic parameters help to understand the reaction after drug entry into the body. ADMET analysis helps to narrow down the search process for selecting candidate drugs.

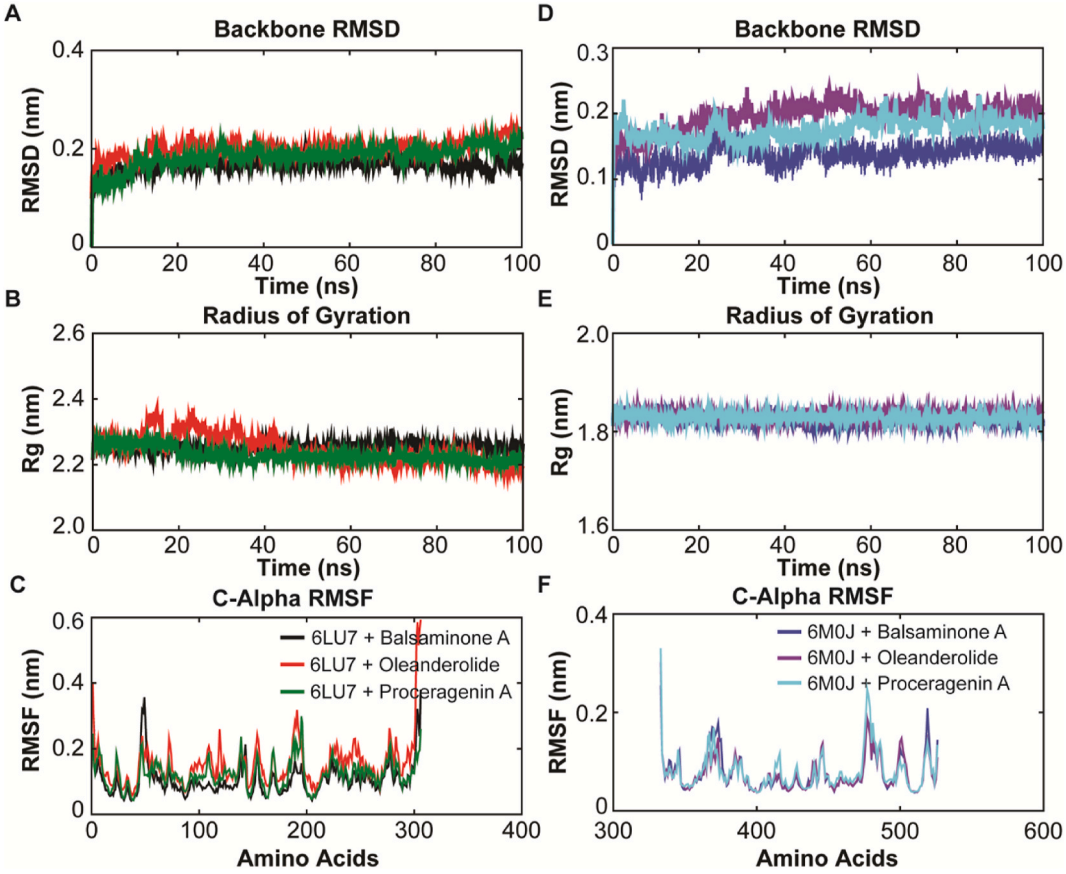

Fig. 2. Intrinsic dynamics stabilities of the top ranked phytocompounds (Balsaminone A, Oleanderolide and Proceragenin A complexed with SARS-CoV-2 M^{Pro} and RBD during all-atoms MD simulation of 100 ns. Colours mentioned on the bottom panel are followed in the upper images as well. (**A**) Dynamics stabilities of the SARS-CoV-2 M^{Pro}-phytocompound complex systems measured by the plotting the RMSD of the backbone atoms with respect to the initial structure used for production MD over the time scale of 100 ns. (**B**) Radius of gyration of SARS-CoV-2 M^{Pro} over the time scale of 100 ns. (**C**) The average Cα-RMSF profile of the each amino acid during the last 50 ns of MD in SARS-CoV-2 M^{Pro}-phytocompound complexes. (**D**) Backbone RMSD displaying stability of SARS-CoV-2 RBD-phytocompound complex systems with respect to the initial structure used for production MD over the time scale of 100 ns. (**E**) Radius of gyration of SARS-CoV-2 RBD complexes over the time scale of 100 ns. (**F**) The average Cα-RMSF profile of the each amino acid during the last 50 ns of MD in SARS-CoV-2 RBD-phytocompound complexes. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

2.5. Molecular interaction analysis

The selected phytocompounds were studied their protein-ligand interacting atoms using LigPlot^{+} v1.4.5 [36] and PLIP [1]. The structures were visualized and illustrated through PyMOL and Adobe Illustrator. Non-bonded interactions like hydrophobic interactions, hydrogen bonds, salt bridges were marked for each docked complex. These results help us to identify important interacting amino acid residues and the mode of interaction.

2.6. Molecular Dynamic (MD) simulation

From the ADMET analysis, we have selected three compounds were effective against both Mpro (6LU7) and spike (6M0J) proteins. The three selected phytocompounds along with the plants' name, from where they have been collected are – balsaminone A (*Impatients balsamina*), oleanderolide (*Nerium oleander*),proceragenin A(*Calotropis procera*).

The Molecular dynamic simulation was performed with those six complexes to study the dynamic behavior, stability, and conformational flexibility using Gromacs 2018.3 version [4,29]. Charmm36 force field was applied to generate topology for proteins and ligands [22]. A cubic water box with TIP3P water model was used for structure solvation. The complex structure was electro-neutralized by adding 0.15 M NaCl. The structure was energy minimized using the steepest descent algorithm in 5000 steps. After energy minimization, the structure was equilibrated through NVT and NPT ensembles for 5ns and 10ns respectively. Finally, MD simulation was performed for 100ns at 300K with 2fs time step using Leap-Frog integrator [8].

2.7. MD trajectory analysis

The quality of the trajectories was analyzed using GROMACS utility toolkit through determining conformational flexibility, stability of the trajectories. Quality assurance parameters like root mean square deviation (RMSD), root mean square fluctuation (RMSF), radius of gyration (Rg), solvent accessible surface area (SASA), and intermolecular H Bond, were calculated using *gmx rms*, *gmx rmsf*, *gmx gyrate*, *gmx sasa*, *gmx hbond* built-in commands in Gromacs respectively. XmGrace, PyMOL, BIOVIA Discovery Studio Visualizer were used for visualization of 2D graphs and structures respectively and the plots were created in Adobe Illustrator.

2.8. Principal Component (PC) analysis

Principal component analysis (PCA) is a multivariate statistical analysis tool that is used to reduce the dimensionality of a dataset while preserving the variability as much as possible. This method is used to understand the correlated motion of the residues to a set of linearly uncorrelated variables which are called (PC) principal components [12]. To analyze principal component, *gmx covar* and *gmx anaeig* built-in

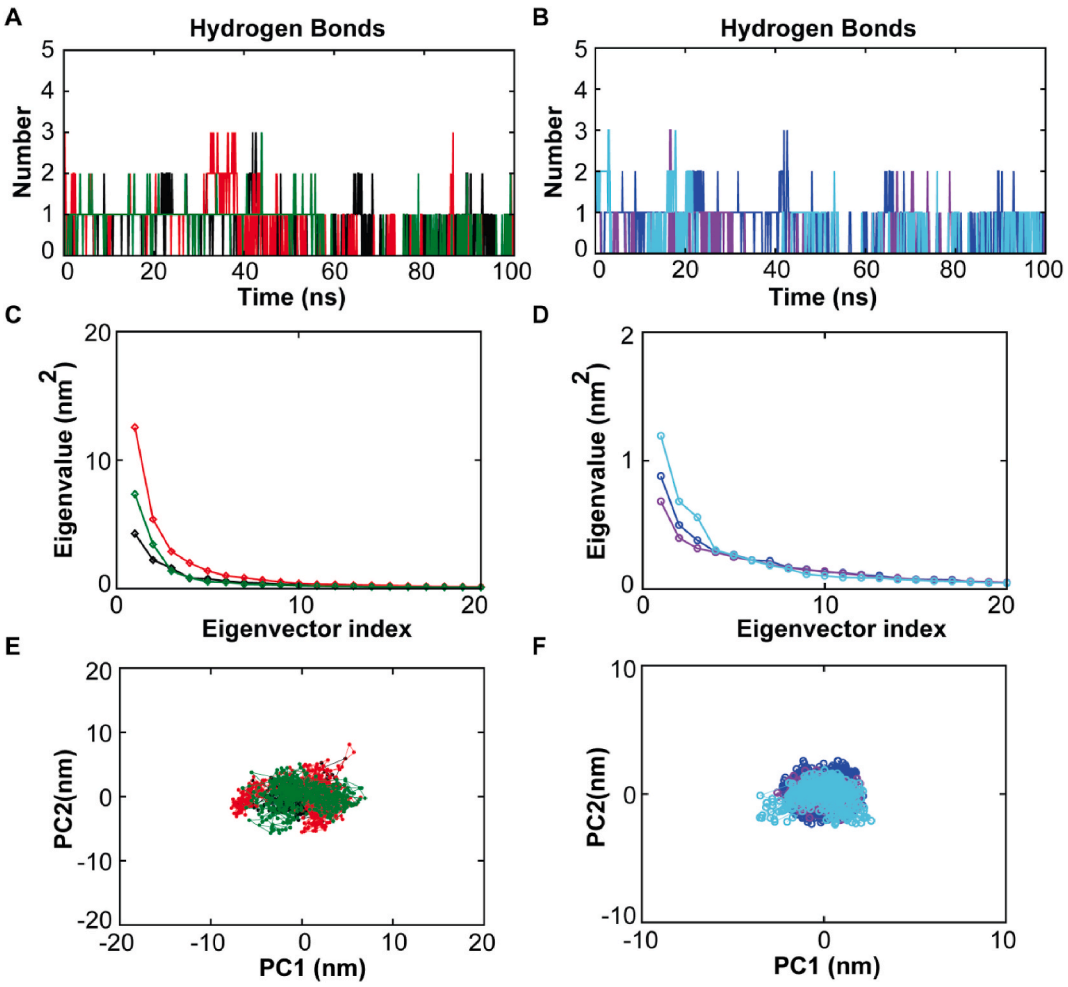

Fig. 3. Dynamics of intermolecular hydrogen bonds and principal component analysis of the SARS-CoV-2 M^{Pro} and RBD complexes with Balsaminone A, Oleanderolide and Proceragenin A during all-atoms MD simulation of 100 ns. In the **left panel black** represents **6LU7+BalsaminoneA**, **red** represents **6LU7+Oleanderolide**, **green** represents **6LU7+ProcerageninA** and in the **right panel blue** represents **6M0J + BalsaminoneA**, **violet** represents **6M0J + Oleanderolide**, and **cyan blue** represents **6M0J + ProcerageninA**] **(A)** Intermolecular H-bond dynamics in SARS-CoV-2 M^{Pro}-phytocompound complexes during 100 ns MD simulation. **(B)** Intermolecular H-bond dynamics in SARS-CoV-2 RBD-phytocompound complexes during 100 ns MD simulation. **(C)** Plot displaying the eigenvalues vs. eigenvectors obtained from the main-chain atoms covariance matrix constructed from the last 50 ns MD trajectory of SARS-CoV-2 M^{Pro}-phytocompound complex systems (only the first 20 eigenvectors are considered for representation). (**D**) Plot displaying the eigenvalues vs. eigenvectors obtained from the main-chain atoms covariance matrix constructed from the last 50 ns MD trajectory of SARS-CoV-2 RBD-phytocompound complex systems. (**E**) Projection of the motion of the SARS-CoV-2 M^{Pro}-phytocompound complexes in phase space along the PC1 and PC2. (**F**) Projection of the motion of the SARS-CoV-2 RBD-phytocompound complexes in phase space along the PC1 and PC2. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

modules of GROMACS were used. In GROMACS, *gmx covar* tool is used to construct the covariance matrix. The matrix generates diagonal eigenvectors (PCs) which represent the correlated motion of the protein and eigenvalues indicate atomic contribution of the motion of protein-ligand complex system. PC Analysis was done by calculating eigenvectors along with eigenvalues and associated projections of the first two principal components where *gmx anaeig* tool was used [34]. Here covariance matrix was constructed from the last 50ns of MD trajectory of the protein-ligand complex system and the first 20 eigenvectors were used for representation. The porcupine plot was generated using modevectors.py script using PC1 and PC2 as parameters.

2.9. Free energy landscape analysis

Free energy landscape (FEL) is a popular tool in bioinformatics to analyze the folding or aggregation of the complex after MD simulation, which in terms gives an idea about the stability of the protein [15]. Principal component values were used as order parameters for the calculation. The *gmx sham* module was used to obtain the FEL of each complex using the whole MD trajectories. In this study, the free energy landscape of protein-ligand complexes was analyzed using Gibbs free energy calculation equation

$$Gi = - K_B T \ln (Ni/Nmax)$$

Where K_B stands for Boltzmann's constant, T for temperature (300K), Ni for population bin I, and Nmax depicts the most inhabited population [11]. A color model was used for representation of the energy level in plot, where red and blue showed the highest and lowest energy level respectively.

2.10. Cluster analysis

Geometric clustering (*gmx cluster*) was performed to explore heterogeneity in the protein structure after performing MD simulation. Gromos clustering algorithm by Daura et al., 1999 was employed. During analysis, C^{α} RMSD cut off 0.2 nm was set for identifying

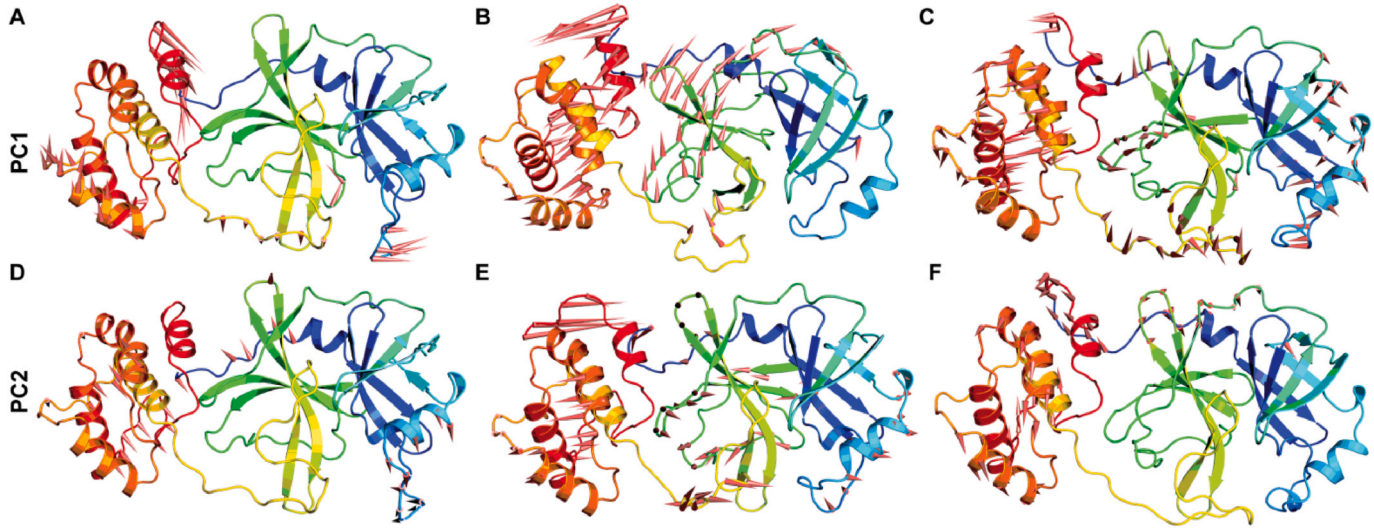

Fig. 4. PCA displaying global motions occupied by the top two principal components in SARS-CoV-2 MPro-phytocompound complexes during the last 50 ns MD. **(A)** Porcupine plot displaying the motion of SARS-CoV-2 MPro-Balsaminone A complex obtained from PC1. **(B)** Porcupine plot displaying the motion of SARS-CoV-2 MPro-Oleanderolide complex obtained from PC1. **(C)** Porcupine plot displaying the motion of SARS-CoV-2 MPro-Proceragenin A complex obtained from PC1. **(D)** Porcupine plot displaying the motion of SARS-CoV-2 MPro-Balsaminone A complex obtained from PC2. **(E)** Porcupine plot displaying the motion of SARS-CoV-2 MPro-Oleanderolide complex obtained from PC2. **(F)** Porcupine plot displaying the motion of SARS-CoV-2 MPro-Proceragenin A complex obtained from PC2. obtained from PC2.

structurally similar clusters and the top 2 ranked clusters with maximum screenshots were analyzed in detail and compared with docked conformations [30].

3. Results and discussion

The COVID-19 pandemic has become a global threat that impacted human health care facilities and the socio-economic structure of the world. Studies have shown that, during the first wave of COVID-19, there was increased use of azithromycin for treatment, but later it was declared by WHO and CDC that, antibiotics have no direct impact on COVID-19. Although some vaccines have arrived in the market, but the effectiveness of the vaccines and fighting the pandemic through only vaccination is not recommended. Besides in largely populated countries complete vaccination will take a lot of time and the virus is evolving rapidly by developing variants, which are proved to be more deadly in nature [16]. Besides that, with COVID-19 prevalence the intake of antibiotics has also increased tremendously [24].

So it is high time to find some drugs for treatment of the infected persons with confidence, so that the spread of the virus can be blocked. In this present study, we have selected Mpro and spike protein (RBD) as drug targets, because both of them are essential for viral entry and replication. Case studies reported effectiveness of traditional Chinese medicines in the containment and prevention of COVID-19 infection in China [9]. Indian medicinal plants are also well known for their medicinal properties. Dual targeting of proteins can be a good way to select phytocompounds for next step of clinical trial. In virtual screening, molecular docking followed by MD simulation has become a classic way to screen the inhibitors. In MD simulation, various utility tool kit has been developed, which help to understand detailed intermolecular interaction, conformational flexibility, energy variation of the system, that gives a very clear prospect of the suppose mechanism of action of the compound against target protein.

3.1. Screening of phytocompounds against SARS CoV-2 viral proteins

Molecular docking between protein and ligand reflects binding affinity between them and the lowest docking score is considered to have the highest binding affinity. In our study, 300 phytocompounds were categorized into three categories for each protein – category I (−3 to -5 kcal/mol), category II (−5.1 to -7 kcal/mol), and category III (−7.1 kcal/mol to above) (**S1**).

Among 300 phytocompounds, 161 phytocompounds for Mpro protein and 150 phytocompounds for spike protein (RBD) belong to category II. The docking scores of ritonavir against Mpro and spike protein (RBD) were −6.4 kcal/mol and −6.1 kcal/mol respectively. Another antiviral drug favipiravir has a lower binding affinity than ritonavir. The docking scores of favipiravir against Mpro and spike protein were −5.7 kcal/mol and −5.3 kcal/mol respectively. Among 300 compounds, we have selected the best 30 compounds for further analysis. The docking scores and ligand efficiency values were listed in Table 1.

All of these compounds have shown higher binding affinity than the standard reference antiviral drugs. Among those thirty compounds, Oleanderolide (−9.4 kcal/mol) and α -Amyrin (−8.5 kcal/mol) has shown the highest binding affinity against Mpro and RBD protein respectively. In comparison with other phytocompounds, Clausenin (−6.8 kcal/mol) has low binding affinity against Mpro and Aromadendrin (−6.5 kcal/mol) has the lowest binding affinity against RBD. But among the above listed compounds, 8 phytocompounds were found to have docking score ≤ -8.0 kcal/mol.

3.2. Evaluations of ADMET properties

The results of ADMET analysis for the best 30 phytocompounds were listed in Table 2. It was stated that any drug which violates two or more rules of Lipinski Rule of Five (R05) will not be orally active [17]. Six compounds have been found to fulfil Lipinski's Rule of five. Quassinoid analog, Pindrolactone, Bryonolic acid, Citrullonol, Canophyllic acid, Oleanolic acid, Ursolic acid, Betulic acid and Clausenin are hepatotoxic and rests three Aromadendrin, Guajaverin, Subulin has very low intestinal absorption capacity. Based on the high binding affinity against the target proteins and non-hepatotoxicity, high gastrointestinal absorption (>90%), the three phytocompounds were chosen for further analysis.

As of now, several studies have employed virtual screening of phytocompounds against SARS-CoV-2 proteins [6,26] to discover potent compounds against Covid-19. Although Oleanderolide and Proceragenin A has slight violation for one Rule of Lipinski (mLogP> 4.15) but rest of the parameters justify the reasons for this selection.

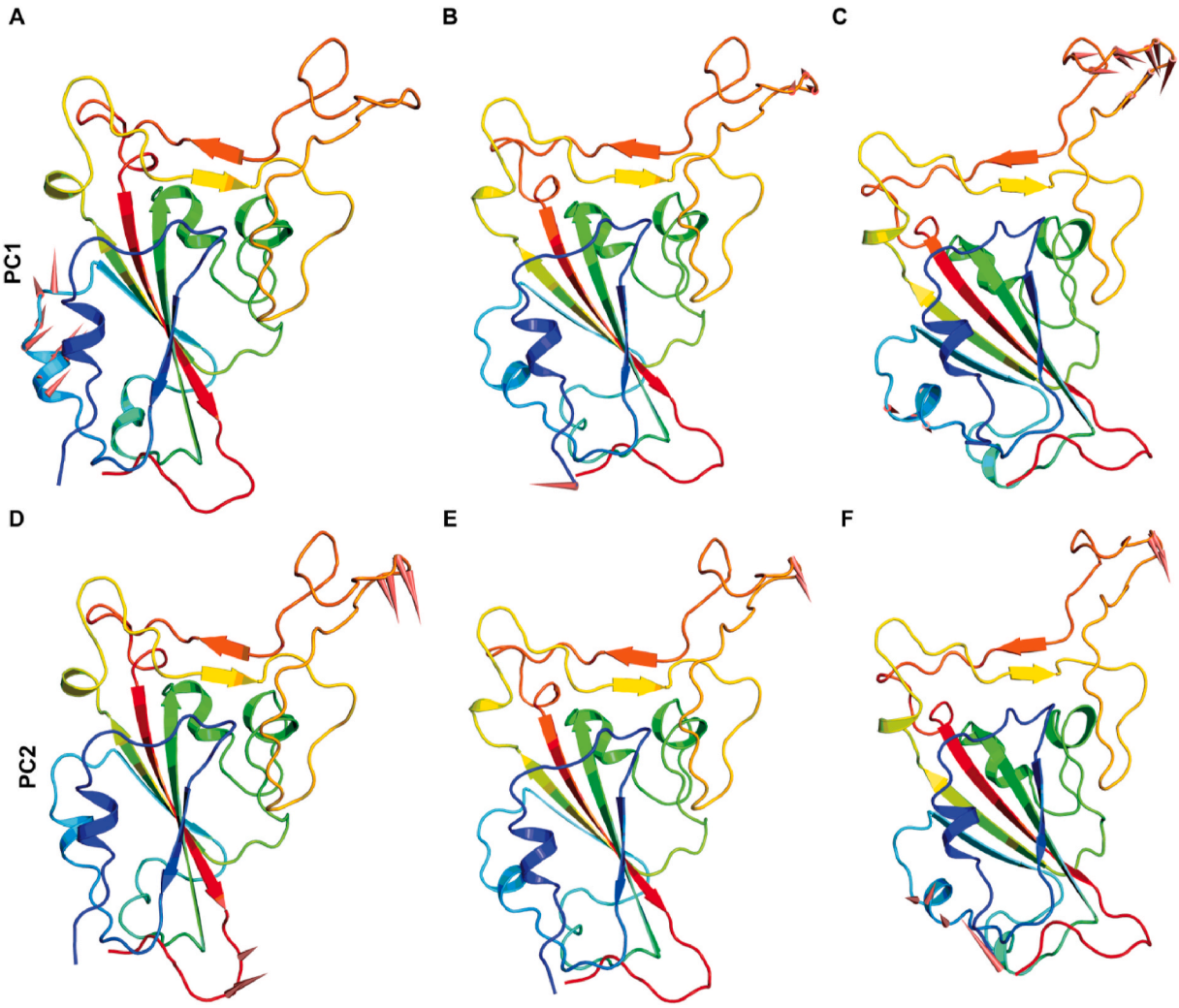

Fig. 5. PCA displaying global motions occupied by the top two principal components in SARS-CoV-2 RBD-phytocompound complexes during the last 50 ns MD. (A) Porcupine plot displaying the motion of SARS-CoV-2 RBD-Balsaminone A complex obtained from PC1. **(B)** Porcupine plot displaying the motion of SARS-CoV-2 RBD-Oleanderolide complex obtained from PC1. **(C)** Porcupine plot displaying the motion of SARS-CoV-2 RBD-Proceragenin A complex obtained from PC1. **(D)** Porcupine plot displaying the motion of SARS-CoV-2 RBD-Balsaminone A complex obtained from PC2. **(E)** Porcupine plot displaying the motion of SARS-CoV-2 RBD-Oleanderolide complex obtained from PC2. **(F)** Porcupine plot displaying the motion of SARS-CoV-2 RBD-Proceragenin A complex.

3.3. Molecular interaction analysis

Studies have shown that the most commonly found bonds between protein-ligand complex are – hydrophobic interaction and Hydrogen bonds [14]. In the present study, six docked complexes were analyzed in detail for molecular interaction (Fig. 1). Oleanderolide has −9.4 kcal/-mol docking score against (6LU7) Mpro protein and it forms two hydrogen bonds with Thr199 (3.0 Å) and Leu287 (3.07 Å) residues of the protein. There are four hydrophobic interactions with Tyr237, Tyr239, Leu286 and Leu287 residues.

Proceragenin A with a docking score of −8.6 kcal/mol with Mpro and has developed three H bonds with Arg131 (3.60 Å), Thr199 (2.89 Å) and Leu271 (3.02 Å) and created six hydrophobic interactions with Tyr237, Tyr239, Leu272, Leu286 and Leu287 residues.

Another studied complex (6lu7-Balsaminone A) has a docking score −8.1 kcal/mol which has formed two H bonds with Asp197 (3.54 Å), and Leu287 (2.87 Å), and three hydrophobic interactions with Tyr237, Leu272 and Leu286 residues of the protein. Balsaminone A also has created two salt bridges with Arg131 (4.51 Å), and Lys137 (4.34 Å). Salt bridge is proved to be the most strong interaction in comparison with H bonds and hydrophobic interactions [14].

All the three above-mentioned phytocompounds have shown good binding affinity with the same docking score −8.3 kcal/mol against spike-RBD protein (6M0J). So molecular interactions were also studied for those complexes.Oleanderolide-6M0J has formed one H bond with Ser375 (4.03 Å) one hydrophobic interaction with Val503 and one salt bridge interaction with Lys 378. Proceragenin A has created three hydrogen bonds with Arg355 (2.88 Å), Asp428 (4.01 Å) and Thr430 (3.07 Å) residues and three hydrophobic interactions with Thr430, Phe464 residues of the protein. Another phytocompound BalsaminoneA formed two H bonds with Ser371 (2.50 Å) and Ser373 (3.04 Å) residues and six hydrophobic interactions with Phe338, Phe342, Val367, Leu368, Phe374 and Trp436 residues of the protein.

Hydrogen bonds are suggested to increase the binding affinity of the protein-ligand complex by removing protein bound water into a bulk solvent [7]. Hydrogen bond interactions of selected phytocompounds with target proteins were shown in **S2.** The above mentioned six docked complexes were further subjected to molecular dynamic simulation to understand the stability of the docked complex.

3.4. Analysis of MD trajectories

After performing molecular docking, the binding affinity between protein and ligand has become clear in a rigid environment, but the

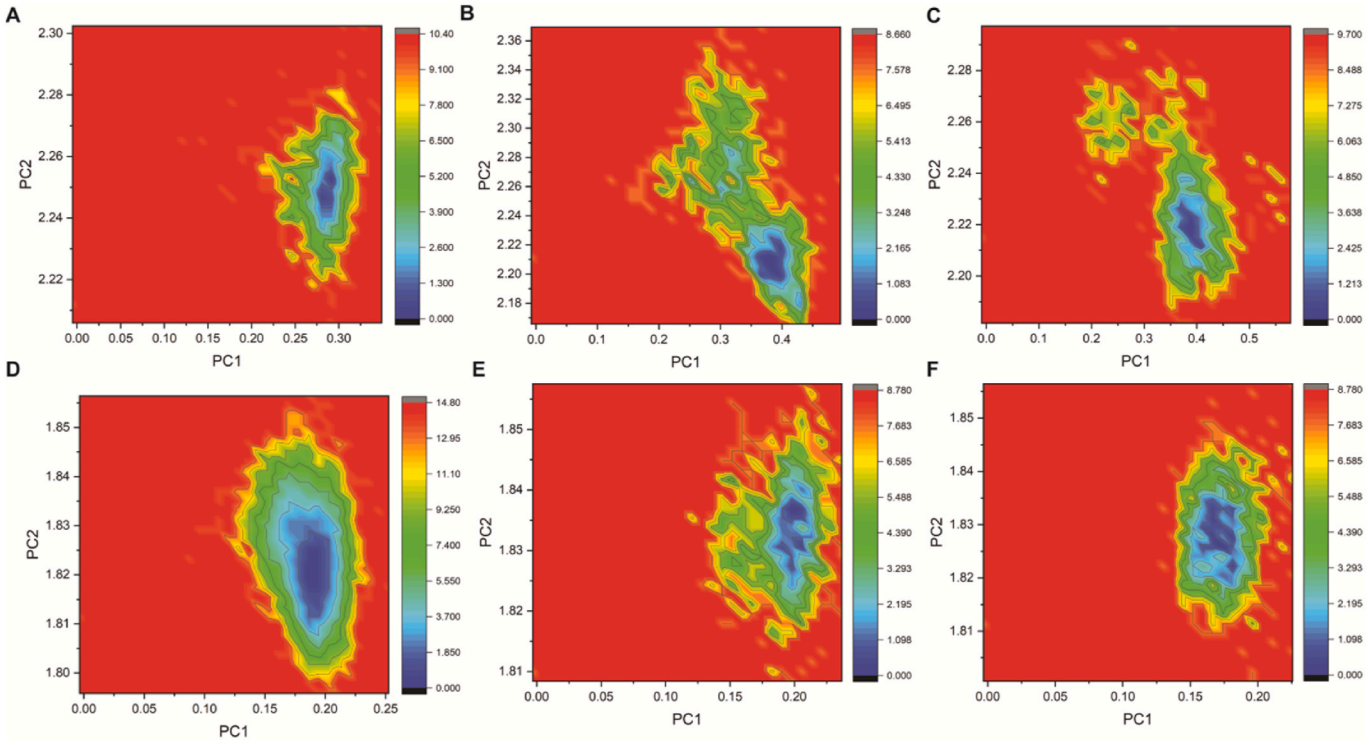

Fig. 6. Free energy landscape (FEL) profiles of the SARS-CoV-2 M^{Pro} and RBD-phytocompound complexes as a function of PC1 and PC2. The colored scale plot shows the free energy profile (kcal/mol) and the dark blue colored profiles are deep minima. [**(A)** M^{Pro}-Balsaminone A; **(B)** M^{Pro}-Oleanderolide; **(C)** M^{Pro}-Proceragenin A; **(D)** RBD-Balsaminone A complex; **(E)** RBD-Oleanderolide complex**; (F)** RBD-Proceragenin A. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

binding properties in the solvated state can fully be analyzed only through MD simulation. A total 100 ns MD run was performed to understand structural stability of each complex.

Root mean square deviation (RMSD) of the protein backbone denotes the stability of the structure during MD run. Three protein-ligand complexes (with 6LU7) has found little fluctuations in RMSD till first 20ns, then the system became almost stable with an average RMSD-0.2 nm (Fig. 2A). For spike protein-ligand complexes (RBD), the structures were seemed to fluctuate till the first 40ns, and after 80ns, the systems appeared to be stable. Among three protein-ligand complexes, Balsaminone A-6M0J has shown the least RMSD value for protein backbone, which indicates more structural stability (Fig. 2D). Similar to protein RMSD, ligand RMSD values were calculated (**S3 C,D**). The probability distribution of RMSD for each complex was shown in **S3 A,B**.

Radius of gyration (Rg) indicates the compactness of the system. Stability in the graph indicates conformational stability of the complex structure during MD simulation. Among three protein-ligand complexes with 6LU7 protein, ProcerageninA showed the most structural stability with an average compact of gyrate-2.25 nm and Oleanderolide was found to be conformationally stable after 40ns (Fig. 2B). But all the protein-ligand complexes with 6M0J protein showed structural stability during the whole MD simulation with an average compact radius of −1.85 nm (Fig. 2E). The steady value of Rg indicates stable folded complex structure, but variations in Rg over time indicate misfolding of the complex structure. A higher value of Rg indicates more flexibility of the conformation. Additionally, solvent-accessible surface area (SASA) indicates the contracted nature of the protein (Figure. S3 E,F). In trajectory analysis, the structural fluctuation and degree of flexibility of the protein were analyzed through (RMSF) of Cα atoms of protein-ligand complexes. A close examination of Fig. 2C,F indicates all the six complexes for both the proteins showed a similar pattern in RMSF. A close inspection of the residues in putty representation, revealed that, loops showing more fluctuations and more structural mobility (**S4**). Ligand binding residues with higher peaks indicate their mobility and participation in ligand binding.

3.5. Intermolecular H-bonds and PCA analysis

Intermolecular stability providing H bond dynamics for each complex were depicted in Fig. 3 (A,B). During the 100ns MD run, all of the complexes showed different intermolecular H bonding patterns, but among them, 6LU7-Oleanderolide, 6LU7-Balsaminone A, 6M0J-Oleanderolide, and 6M0J-Proceragenin A yielded a higher number of hydrogen bonds in comparison with other complexes. During the first 40 ns, an increase in the number of H bonds was observed, but later on a constant of one or more H bonds was observed for each system.

Principle component analysis (PCA) is used to study the collective motion of each complex, based on eigenvectors and eigenvalues. To get a better knowledge about the conformational flexibility of each atom, MD trajectories were thrown to a phase space to yield a spectrum of EVs, which represents a single component of the motion, symbolic of the direction of the motion [5]. A close inspection of each docked complex showed a sharp decrease in the eigenvalues after the first two among the twenty eigenvalues. This indicates considerable flexibility of the docked protein-ligand complex structure during the initial stages of simulation, which is further reduced with simulation interval. Oleanderolide in complex with 6LU7 and Proceragenin A in complex with RBD showed a higher trace value as compared to the rest (Fig. 3C and D). High trace value indicates greater flexibility, which perfectly correlates with the higher occupancy of the phase space occupied in the scatter plot (Fig. 3E and F). This result perfectly supports RMSF analysis data.

To get a better understanding, porcupine plots were generated, by capturing eigenvectors on extreme projections on both PC1 and PC2 where the direction and length of the arrow reflects motion direction and strength respectively. The large arrows indicate increased motion of the complex structures. Here in Fig. 4, porcupine plots were generated

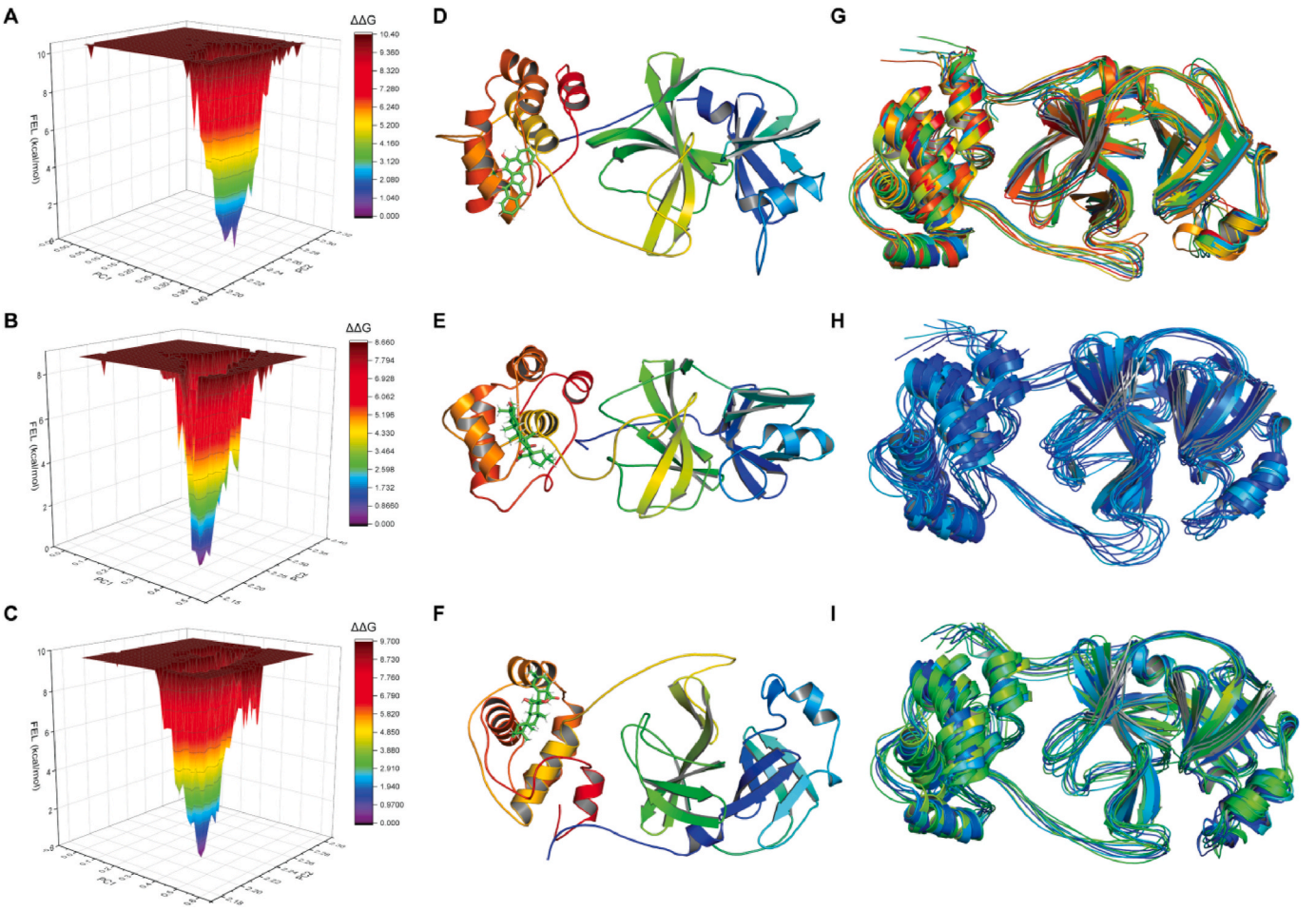

Fig. 7. Three-dimensional FEL profile, snapshot from low energy conformational state and clustering analysis of SARS-CoV-2 M^{Pro}-phytocompound complexes. (A–C) FEL profiles of SARS-CoV-2 M^{Pro}-Balsaminone A, Oleanderolide complex and Proceragenin A complexes. **(D–F)** Snapshots extracted from the minima (low energy conformational states) of SARS-CoV-2 M^{Pro}-Balsaminone A, Oleanderolide complex and Proceragenin A complexes. **(G–I)** Structurally superimposed view of the clustered snapshots of SARS-CoV-2 M^{Pro}-Balsaminone A, Oleanderolide complex and Proceragenin A complexes obtained from RMSD based clustering employed in GROMACS with a cut-off of 0.2 nm.

using both PC1 and PC2 order parameters for the protein (Mpro) in combination with phytocompounds. The Mpro-Oleanderolide complex showed more outward arrows in the middle domain for both PC1 and PC2 and the C-terminal (red) domain showed both inward and outward arrows. In comparison to Mpro protein, RBD showed less motion in porcupine plot (Fig. 5). However, similar to the scatter plot results, both 6LU7-Oleanderolide, and RBD-Proceragenin A showed a high degree of movement (Figs. 4 and 5). A great movement was observed in the beta-strand connecting loop region in both the complexes.

3.6. Free energy landscape (FEL) and clustering

The free energy landscape helps to understand conformational variability along with energy minimization. In Fig. 6, the dark blue color represents energy minima which mean the lowest energy.The degree of dispersion in the plot indicates conformational flexibility. A larger area was observed for 6LU7-Oleanderolide which also supports the previous RMSF, PCA result in terms of flexibility. There are so no significant conformational changes been detected in the complex structures.

In Figs. 7 and 8, we have extracted the least energy conformational state from FEL analysis and compared them with the top ranked cluster which represents no significant structural variation. Thus, it indicates the structural uniformity is maintained throughout the MD run. To identify similar structures throughout the MD simulation trajectory, clustering analysis was performed using Gromacs clustering algorithm with Cα-RMSD cut off 0.2 nm. For 6LU7-Balsaminone A, 11 clusters were found with an average RMSD - 0.22 and for 6LU7-Oleanderolide, 44 clusters were found with an average RMSD -0.33. For 6LU7 - Proceragenin A complex structure, 21 clusters with an average RMSD - 0.25was found. With the same RMSD cut off for RBD protein, Balsaminone - RBD had 2 clusters with an average RMSD -0.17, Oleanderolide-6M0J had 4 clusters with an average RMSD - 0.18 and 3 clusters for ProceragininA with an average RMSD - 0.17. In the free energy profile (Figs. 7 and 8) violet color represents energy minima, so large energy minima depicts more stability, but separate peaks in violet indicate transition in the conformation of protein towards more thermodynamically stable state. In Fig. 7A-C, 6LU7-Oleanderolide and in Fig. 8A-C, RBD-Balsaminone A depicts more stability which perfectly correlates the previous FEL profile. RBD-Proceragenin A denotes transition towards thermodynamic stability.

Structurally superimposed view of clustered snapshots were represented in Fig. 7G-I and Fig. 8G-I. This free energy profile results clearly correlates with the previous RMSD and radius of gyrate (Rg) data for understanding stability of the system.

In computer aided drug design (CADD), understanding electrostatic surface potential of protein and binding interaction of ligands is an important tool which defines binding affinity of intermolecular complex. In Fig. 9, it is clearly depicted that, ligands which bind on charged patches of the protein, indicates more stability in behavior throughout MD run and shows high binding affinity.

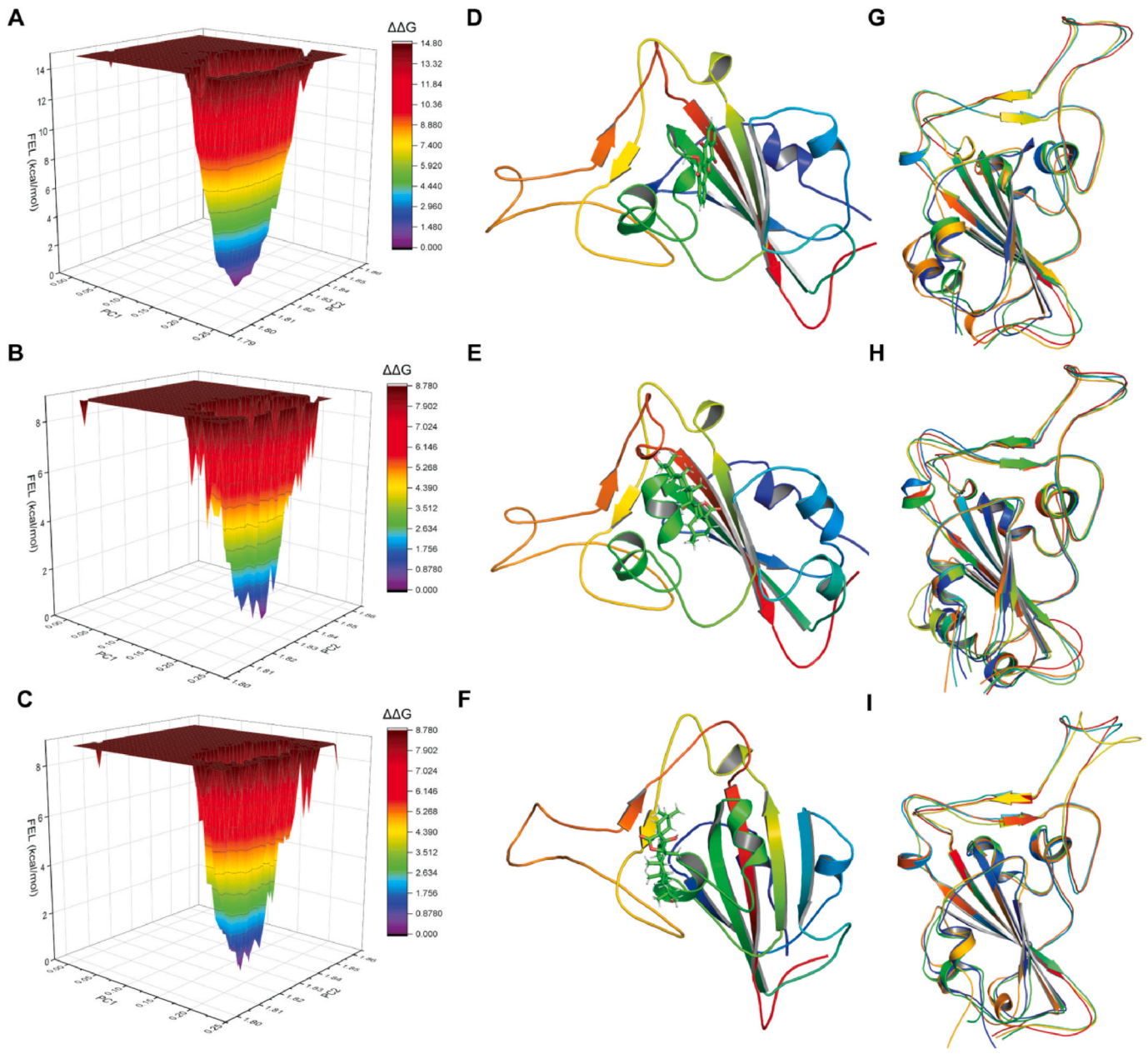

Fig. 8. Three-dimensional FEL profile, snapshot from low energy conformational state and clustering analysis of SARS-CoV-2-RBD-phytocompound complexes. (A–C) FEL profiles of SARS-CoV-2 RBD-Balsaminone A, Oleanderolide complex and Proceragenin A complexes. **(D–F)** Snapshots extracted from the minima (low energy conformational states) of SARS-CoV-2 RBD-Balsaminone A, Oleanderolide complex and Proceragenin A complexes. **(G–I)** Structurally superimposed view of the clustered snapshots of SARS-CoV-2 RBD-Balsaminone A, Oleanderolide complex and Proceragenin A complexes obtained from RMSD based clustering employed in GROMACS with a cut-off of 0.2 nm.

Phytocompounds selected in this study were Oleanderolide from *Nerium oleander,* proceragenin A from *Calotropis procera* and balsaminone A was derived from *Impatiens balsamina* plants, which all are common Indian medicinal plants. Similar to our study, Sen et al. [31], also targeted the same two proteins and screened 10 Indian spices plants for their antiviral activity. Naltrexone, a FDA approved drug was targeted against spike RBD and has shown to bind in the cavity of RBD-ACE2 receptor [8]. Another study targets both Mpro and RDRP protein and screened phytocompounds against them. They selected remdesivir as reference drug. They found Mulberrosides and Punigluconin showed high binding affinity against both the proteins and Emblicanin A, Nimbolide against RdRp while Andrographolides, Anolignans, Chebulic acid, Mimusopic acid showed better binding affinity against Mpro as compared with the reference drug. (A [32]. Similarly, in our study we have used ritonavir and favipiravir as reference drugs. The three selected compounds has shown high binding affinity than those reference drugs. Although there were previous studies [21]; S [33]. reported that different phytocompounds can be found effective against COVID-19 proteins, but our study for the first time reporting these three phytocompounds with strong binding affinity against both proteins. The phytocompounds reported in this study also showed positive results in ADMET property analysis, which suggests that there is a high probability that they will perform well in clinical trial. Diversity of phytocompounds tells that still phytocompounds are very much unexplored and they can be opted as better alternative for treatment.

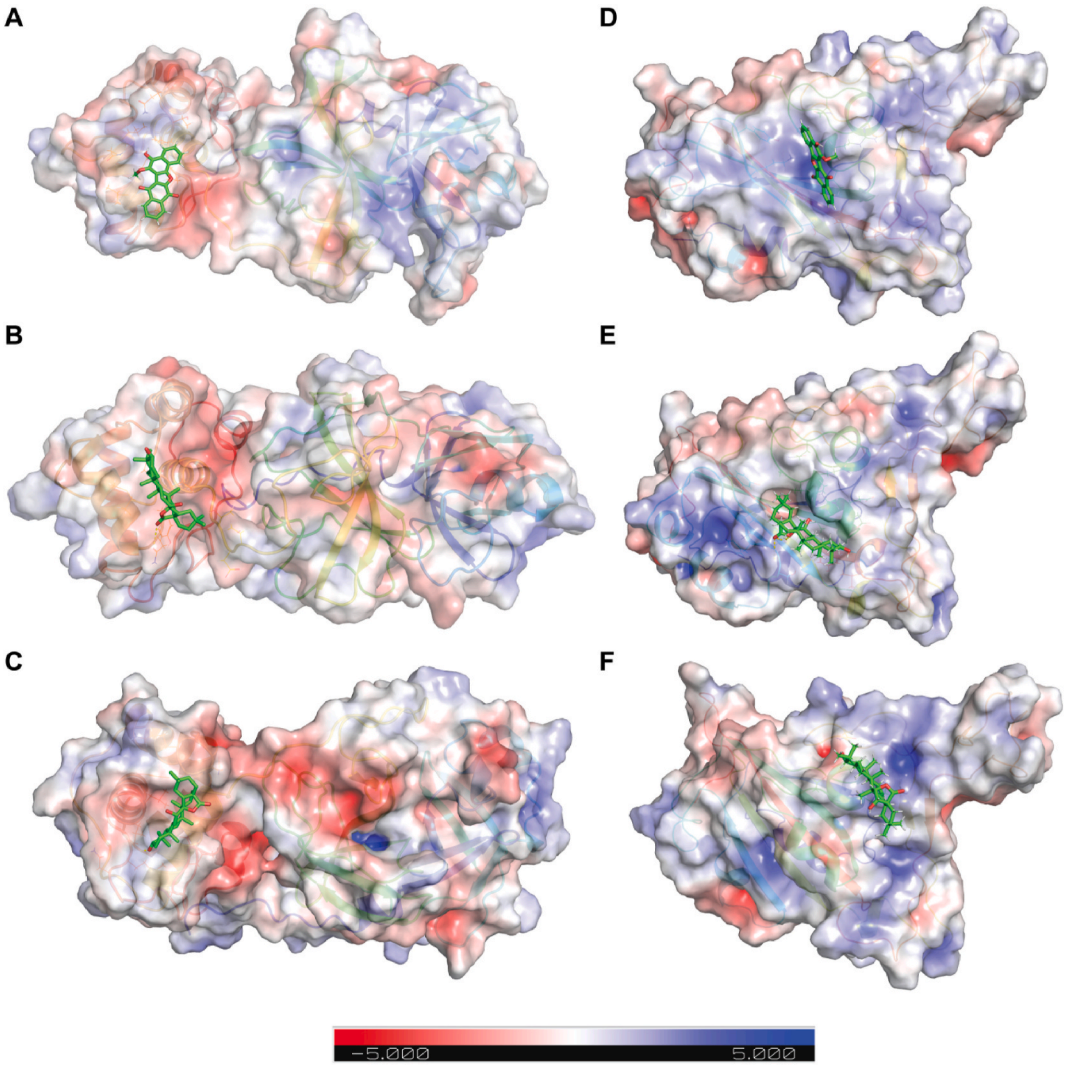

Fig. 9. Electrostatic surface potential maps of the representative conformations extracted from low-energy structural ensembles of each FEL from SARS-CoV-2 M^{Pro}(6LU7) and RBD(6M0J)-phytocompound complexes. [(A–C): 6LU7-Balsaminone A, 6LU7-Oleanderolide, 6LU7-Proceragenin A; **(D–F)**: 6M0J-Balsaminone A, 6M0J-Oleanderolide complex, 6M0J-Proceragenin A. The image was generated using APBS-Plugin in PyMOL, where Blue, Red and White color indicates positive, negative and neutral charge surfaces in SARS-CoV-2 M^{Pro} and RBD respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

4. Conclusion

Till date accurate treatment strategy for ongoing COVID-19 pandemic is limited and controlling COVID-19 pandemic through vaccination is challenging due to the coverage and rapid emergence of variants. To cope up with viral variant development speed, in the present scenario discovery of new antiviral drugs has become inevitable. Computational technique with detailed analysis of the complex structure can pave a path for drug discovery. Natural products are non-toxic to human and have immense antimicrobial potential. In the present study, three compounds named – Oleanderolide, Proceragenin A, Balsaminone A appeared as good antiviral drug candidate. Throughout the MD simulation trajectory analysis and principal component analysis, these three compounds displayed a consistency in stability in the complex structure. But among the three phytocompounds, Oleanderolide for Mpro and Proceragenin A for spike-RBD have demonstrated encouraging results. From the above good results found in *in silico* studies, it can be concluded that, these three novel phytocompounds (Oleanderolide, Proceragenin A, Balsaminone A) has promising future as possible drug candidates against COVID-19 and should be further evaluated *in vitro* and *in vivo* clinical trial for their activity. So to find new antiviral drugs, a good combination of computational and medicinal techniques can be effective and which can offer the mankind new antiviral drugs in near future.

Ethics approval

The study was ethically approved by the institutional human ethical committee of ICMR – Regional Medical Research Centre, Bhubaneswar.

Declaration of competing interest

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

Acknowledgments

The authors gratefully acknowledge all the healthcare workers for their tireless dedication at each level to fight COVID-19. The authors are thankful to the Indian Council of Medical Research, New Delhi for providing financial support for the study through intramural funding

and Council of Scientific and Industrial Research for providing Senior Fellowship to the 1st author.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmgm.2022.108192.

References

[1] M.F. Adasme, K.L. Linnemann, S.N. Bolz, F. Kaiser, S. Salentin, V.J. Haupt, M. Schroeder, Plip 2021: expanding the scope of the protein-ligand interaction profiler to DNA and RNA, Nucleic Acids Res. 49 (W1) (2021) W530–W534, https://doi.org/10.1093/nar/gkab294.

[2] C. Anderson, Reference 2000, Ref. User Serv. Q. 38 (3) (1999) 241.

[3] A. Banerjee, K. Kulcsar, V. Misra, M. Frieman, K. Mossman, Bats and coronaviruses, Viruses 11 (1) (2019) 7–9, https://doi.org/10.3390/v11010041.

[4] H.J.C. Berendsen, D. van der Spoel, R. van Drunen, GROMACS: a message-passing parallel molecular dynamics implementation, Comput. Phys. Commun. 91 (1–3) (1995) 43–56, https://doi.org/10.1016/0010-4655(95)00042-E.

[5] S. Bharadwaj, A. Dubey, U. Yadava, S.K. Mishra, Exploration of natural compounds with anti-SARS-CoV-2 activity via inhibition of SARS-CoV-2 M pro 22 (September 2020) (2021) 1361–1377, https://doi.org/10.1093/bib/bbaa382.

[6] S. Borkotoky, M. Banerjee, A computational prediction of SARS-CoV-2 structural protein inhibitors from Azadirachta indica (Neem), J. Biomol. Struct. Dyn. 39 (11) (2021) 4111–4121, https://doi.org/10.1080/07391102.2020.1774419.

[7] D. Chen, N. Oezguen, P. Urvil, C. Ferguson, S.M. Dann, T.C. Savidge, Regulation of protein-ligand binding affinity by hydrogen bond pairing, Sci. Adv. 2 (3) (2016), https://doi.org/10.1126/sciadv.1501240.

[8] A. Choubey, B. Dehury, S. Kumar, B. Medhi, P. Mondal, Naltrexone a potential therapeutic candidate for COVID-19, J. Biomol. Struct. Dyn. (2020) 1–8, https://doi.org/10.1080/07391102.2020.1820379, 0(0).

[9] L. Chu, F. Huang, M. Zhang, B. Huang, Y. Wang, Current status of traditional Chinese medicine for the treatment of COVID-19 in China, Chin. Med. 16 (1) (2021) 1–15, https://doi.org/10.1186/s13020-021-00461-y.

[10] A. Daina, O. Michielin, V. Zoete, SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules, Sci. Rep. 7 (October 2016) (2017) 1–13, https://doi.org/10.1038/srep42717.

[11] R. Dash, M.C. Ali, N. Dash, M.A.K. Azad, S.M. Zahid Hosen, M.A. Hannan, I. S. Moon, Structural and dynamic characterizations highlight the deleterious role of SULT1A1 R213H polymorphism in substrate binding, Int. J. Mol. Sci. 20 (24) (2019), https://doi.org/10.3390/ijms20246256.

[12] B. Dehury, M.C. Patra, J. Maharana, J. Sahu, P. Sen, M.K. Modi, M.D. Choudhury, M. Barooah, Structure-based computational study of two disease resistance gene homologues (Hm1 and Hm2) in maize (Zea mays L.) with implications in plant-pathogen, PLoS One 9 (5) (2014), https://doi.org/10.1371/journal.pone.0097852.

[13] M. El-Dairi, R.J. House, Optic nerve hypoplasia, in: Handbook of Pediatric Retinal OCT and the Eye-Brain Connection, 2019, pp. 285–287, https://doi.org/10.1016/B978-0-323-60984-5.00062-7.

[14] R. Ferreira De Freitas, M. Schapira, A systematic analysis of atomic protein-ligand interactions in the PDB, MedChemComm 8 (10) (2017) 1970–1981, https://doi.org/10.1039/c7md00381a.

[15] Frauenfelder, H., Sligar, S. G., & Wolynes, P. G. (n.d.). Proteins.

[16] J.L. Goodman, J.D. Grabenstein, M.M. Braun, Answering key questions about COVID-19 vaccines, JAMA - J. Am. Med. Ass. 324 (20) (2020) 2027–2028, https://doi.org/10.1001/jama.2020.20590.

[17] L. Guan, H. Yang, Y. Cai, L. Sun, P. Di, W. Li, G. Liu, Y. Tang, ADMET-score-a comprehensive scoring function for evaluation of chemical drug-likeness, MedChemComm 10 (1) (2019) 148–157, https://doi.org/10.1039/C8MD00472B.

[18] Y. Huang, C. Yang, X. feng Xu, W. Xu, S. wen Liu, Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19, Acta Pharmacol. Sin. 41 (9) (2020) 1141–1149, https://doi.org/10.1038/s41401-020-0485-4.

[19] Z. Jin, X. Du, Y. Xu, Y. Deng, M. Liu, Y. Zhao, B. Zhang, X. Li, L. Zhang, C. Peng, Y. Duan, J. Yu, L. Wang, K. Yang, F. Liu, R. Jiang, X. Yang, T. You, X. Liu, H. Yang, Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors, Nature 582 (7811) (2020) 289–293, https://doi.org/10.1038/s41586-020-2223-y.

[20] S. Keretsu, S.P. Bhujbal, S.J. Cho, Rational approach toward COVID-19 main protease inhibitors via molecular docking, molecular dynamics simulation and free energy calculation, Sci. Rep. 10 (1) (2020) 1–14, https://doi.org/10.1038/s41598-020-74468-0.

[21] A.U. Khan, A. Ali, G. Srivastava, A. Sharma, Potential Inhibitors Designed against NDM-1 Type Metallo- β - Lactamases : an Attempt to Enhance Efficacies of Antibiotics against Multi-Drug-Resistant Bacteria, 2017, https://doi.org/10.1038/s41598-017-09588-1. *April*, 1–14.

[22] J.S. Knutson, 基因的改变NIH public access, Bone 23 (1) (2014) 1–7, https://doi.org/10.1002/jcc.23354.CHARMM36.

[23] S. Kumar, P.P. Sharma, U. Shankar, D. Kumar, S.K. Joshi, L. Pena, R. Durvasula, A. Kumar, P. Kempaiah, Poonam, B. Rathi, Discovery of new hydroxyethylamine analogs against 3CLproProtein target of SARS-CoV-2: molecular docking, molecular dynamics simulation, and structure-activity relationship studies, J. Chem. Inf. Model. 60 (12) (2020) 5754–5770, https://doi.org/10.1021/acs.jcim.0c00326.

[24] M.A.B. Lucien, M.F. Canarie, P.E. Kilgore, G. Jean-Denis, N. Fénélon, M. Pierre, M. Cerpa, G.A. Joseph, G. Maki, M.J. Zervos, P. Dely, J. Boncy, H. Sati, A. del Rio, P. Ramon-Pardo, Antibiotics and antimicrobial resistance in the COVID-19 era: perspective from resource-limited settings, Int. J. Infect. Dis. 104 (52) (2021) 250–254, https://doi.org/10.1016/j.ijid.2020.12.087.

[25] K. Mohanraj, B.S. Karthikeyan, R.P. Vivek-Ananth, R.P.B. Chand, S.R. Aparna, P. Mangalapandi, A. Samal, IMPPAT: a curated database of Indian medicinal plants, phytochemistry and therapeutics, Sci. Rep. 8 (1) (2018) 1–17, https://doi.org/10.1038/s41598-018-22631-z.

[26] B. Nouadi, A. Ezaouine, M. El Messal, M. Blaghen, F. Bennis, F. Chegdani, Prediction of anti-COVID 19 therapeutic power of medicinal Moroccan plants using molecular docking, Bioinf. Biol. Insights 15 (2021), https://doi.org/10.1177/11779322211009199.

[27] A. Pedretti, L. Villa, G. Vistoli, VEGA: a versatile program to convert, handle and visualize molecular structure on Windows-based PCs, J. Mol. Graph. Model. 21 (1) (2002) 47–49, https://doi.org/10.1016/S1093-3263(02)00123-7.

[28] D.E.V. Pires, T.L. Blundell, D.B. Ascher, pkCSM: predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures, J. Med. Chem. 58 (9) (2015) 4066–4072, https://doi.org/10.1021/acs.jmedchem.5b00104.

[29] S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M.R. Shirts, J. C. Smith, P.M. Kasson, D. Van Der Spoel, B. Hess, E. Lindahl, Gromacs 4.5: a high-throughput and highly parallel open source molecular simulation toolkit, Bioinformatics 29 (7) (2013) 845–854, https://doi.org/10.1093/bioinformatics/btt055.

[30] C.R. Sahoo, S.K. Paidesetty, B. Dehury, R.N. Padhy, Molecular dynamics and computational study of Mannich-based coumarin derivatives: potent tyrosine kinase inhibitor, J. Biomol. Struct. Dyn. 38 (18) (2020) 5419–5428, https://doi.org/10.1080/07391102.2019.1701554.

[31] D. Sen, P. Debnath, B. Debnath, S. Bhaumik, S. Debnath, Identification of potential inhibitors of SARS-CoV-2 main protease and spike receptor from 10 important spices through structure-based virtual screening and molecular dynamic study, J. Biomol. Struct. Dyn. (2020) 1–22, https://doi.org/10.1080/07391102.2020.1819883, 0(0).

[32] A. Sharma, J. Vora, D. Patel, S. Sinha, P.C. Jha, N. Shrivastava, Identification of natural inhibitors against prime targets of SARS-CoV-2 using molecular docking, molecular dynamics simulation and MM-PBSA approaches, J. Biomol. Struct. Dyn. (2020) 1–16, https://doi.org/10.1080/07391102.2020.1846624, 0(0).

[33] S. Sharma, S. Sharma, P.P. Singh, I.A. Khan, *Potential Inhibitors against NDM-1. 00* (00), 2020, https://doi.org/10.1089/mdr.2019.0315.

[34] S.S. Swain, S.K. Paidesetty, B. Dehury, J. Sahoo, S.C. Vedithi, N. Mahapatra, T. Hussain, R.N. Padhy, Molecular docking and simulation study for synthesis of alternative dapsone derivative as a newer antileprosy drug in multidrug therapy, J. Cell. Biochem. 119 (12) (2018) 9838–9852, https://doi.org/10.1002/jcb.27304.

[35] O. Trott, A.J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, J. Comput. Chem. 31 (2) (2009), https://doi.org/10.1002/jcc.21334. NA-NA.

[36] A.C. Wallace, R.A. Laskowski, J.M. Thornton, LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions the LIGPLOT program automatically generates schematic 2-D representations of protein-ligand complexes from standard Protein Data Bank file input, Protein Eng. 8 (2) (1995) 127–134.

[37] L. Zhang, D. Lin, X. Sun, U. Curth, C. Drosten, L. Sauerhering, S. Becker, K. Rox, R. Hilgenfeld, Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved a-ketoamide inhibitors, Science 368 (6489) (2020) 409–412, https://doi.org/10.1126/science.abb3405.

[38] F. Zhou, T. Yu, R. Du, G. Fan, Y. Liu, Z. Liu, J. Xiang, Y. Wang, B. Song, X. Gu, L. Guan, Y. Wei, H. Li, X. Wu, J. Xu, S. Tu, Y. Zhang, H. Chen, B. Cao, Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study, Lancet 395 (10229) (2020) 1054–1062, https://doi.org/10.1016/S0140-6736(20)30566-3.