In silico analysis of selected alkaloids against main protease (M^{pro}) of SARS-CoV-2

Saksham Garg [b], Arpita Roy [a,b,*]

[a] Department of Biotechnology, School of Engineering & Technology, Sharda University, Greater Noida, India
[b] Department of Biotechnology, Delhi Technological University, Delhi, India

ARTICLE INFO

ABSTRACT

In the present situation, COVID-19 has become the global health concern due to its high contagious nature. It initially appeared in December 2019 in Wuhan, China and now affected more than 190 countries. As of now preventive measures are the sole solution to stop this disease for further transmission from person to person transmissions as there is no effective treatment or vaccine available to date. Research and development of new molecule is a laborious process; therefore, drug repurposing can be an alternative solution that involves the identification of potential compounds from the already available data. Alkaloids are potential source of therapeutic agents which might be able to treat novel COVID-19. Therefore, in the present study, twenty potential alkaloid molecules that possess antiviral activity against different viral diseases have taken into consideration and scrutinized using Lipinski's rule. Then out of twenty compounds seventeen were further selected for docking study. Docking study was performed using Autodock software and the best four molecule which provides maximum negative binding energy was selected for further analysis. Two alkaloids namely thalimonine and sophaline D showed potential activity to inhibit the M^{pro} but to confirm the claim further in-vitro studies are required.

1. Introduction

In late December of 2019, few patients were admitted to the hospitals of Wuhan city of China. The patients claimed to have acute respiratory distress and initially were diagnosed to have pneumonia like symptoms [1,2]. The tests revealed presence of an unknown infectious agent which was later identified to be a member of coronavirus family [3]. Coronaviruses (CoVs) are reported to cause major disorders in respiratory and digestive tracts. Infection is seen in both animals and humans. Some reports suggest that only certain species are infected by CoVs which include mammals, avian species and reptiles [4]. Human civilization had also seen earlier the outbreaks of CoVs such as severe acute respiratory syndrome (SARS)-CoV and middle east respiratory syndrome (MERS)-CoV. Now, along with previously known CoVs, the newly identified coronavirus is also considered to be of immense threat. Earlier known as 2019 novel-CoV, now WHO has named it as COVID-19. As per WHO current report on September 16, 2020, there are 29,724,918 confirmed cases all over the globe with 939,185 confirmed deaths (https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200615-covid-19). There are in total of 213 active territories

affected by the outbreak of virus. The virus is spreading on a logarithmic scale on a daily basis (https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technica). The outbreak now is referred to as a pandemic owing to its spread and transmission rate among the population [5]. Certain groups including children below 6 years of age, patients with underlying conditions like heart disease, lung disease, diabetes and elderly citizens are considered to be at a higher risk level. The virus majorly infects the upper respiratory tract causing a range of symptoms from mild to severe. In most cases mild symptoms like fever, dry cough and tiredness are seen while 1 out 6 patients of COVID develop severe symptoms like shortness of breath and become severely ill (WHO news). An average estimate of its incubation period is 1–14 days. The transmission is taking place from person to person via air droplets and currently, no definite therapy is available for this infection. COVID 19 (SARS-CoV-2) belongs to *Coronaviridae* family which contains single-stranded ribonucleic acid structure similar to SARS-CoV. Replication of viral particles depends on various proteins and targeting one of them can lead to inhibition of this infection. One of the potential characterized drug targets among the coronaviruses is main protease (M^{pro}) which is also known as 3CL^{pro}. Structural and amino acid analysis

* Corresponding author. Department of Biotechnology, School of Engineering & Technology, Sharda University, Greater Noida, India.
E-mail address: arbt2014@gmail.com (A. Roy).

showed that there is a high similarity (96%) in case of main protease (M^pro) of SARS-CoV-2 with other members of SARS-CoV family and majority of the residues are conserved in M^pro. This enzyme plays a key role in processing the polyproteins that are translated from viral RNA and it is essential for viral replication [6]. M^pro contains eleven cleavage sites on the large polyprotein and the recognition sequence at most sites is Leu-Gln (cleavage site) (Ser, Ala, Gly) [6]. Inhibition of this enzyme's activity would block the replication of virus because there are no known human proteases which show similar cleavage specificity [7]. A lot of investigations are going on to tackle and curb the virus infection. The secondary metabolites of plants are considered to be highly potential and able to prevent various diseases such as cancer, diabetics, viral infections, etc. Alkaloids are secondary metabolites which possess various pharmacological activities [8]. Over the past decade, due to shifted interest towards the natural compounds, various alkaloids are being tested for different purposes including the potential treatment against viral infections [9]. Therefore, in this study, twenty alkaloid molecules (ligand) were selected based on their antiviral potential reported in the literature. Lycorine is found in the members of *Amaryllidaceae* family and shows various medical importance [10]. Various studies reported its potential antiviral activity against Human Immunodeficiency Virus (HIV) [11], Herpes Simplex Virus type 1 (HSV1) [12], etc. Hemanthamine is another phytochemical present in *Lycoris radiate* and showed potential antiviral activity against influenza virus (H5N1) [13]. Berberine, an isoquinoline alkaloid found in a significant number of plants such as *Berberis vulgaris* (barberry), *Berberis aristata* (tree turmeric), *Mahonia aquifolium* (Oregon grape) and *Hydrastis canadensis* (goldenseal). It possesses antiviral activity against influenza virus (H1N1) [14]. Thalimonine another alkaloid presents in *Thalictrum simplex* showed potential to inhibit influenza virus [15]. Hippeastrine is derived from the member of *Amaryllidaceae* family and found in *Lycoris radiate*. It has been found to inhibit influenza virus (H5N1) [13]. Hirsutine is an indole alkaloid which is extracted from *Uncaria rhynchophylla*. It is a potential inhibitor of influenza virus (H3N2) [16]. 5alpha-Hydroxysophocarpine is isolated from *Sophora flavescens var. angustifolia* and possess anti-viral activity against hepatitis B virus [17]. Fangchinoline is found in *S. tetrandrae* and possess diverse biological activities. Its antiviral activity reported against HIV type 1 [18]. Tetrandrine is derived from *Stephania tetrandra* and act as a calcium channel blocker. Its antiviral activity reported against herpes simplex virus type-1, human coronavirus OC43 [19,20]. Cepharanthine (CEP) is extracted from *Stephania cepharantha* and possess potential antiviral activity against human coronavirus OC43 [20]. Skimmianine and methoxydihydronitidine both can be isolated from roots of *Zanthoxylum nitidum* and are being tested for potential activity against HBV [21]. Sophaline D again has shown to inhibit the pathogenicity of HBV virus and can be isolated from *Sophora alopecuriodes* [22]. A steroidal alkaloid; tomatidine derived from unripe stem and leaves green tomatoes showed reasonable inhibition of Chikungunya virus in various cell lines [23]. Emetine can be isolated from *Pyschotria ipecacuanha*. It has been considered a potential compound against the HIV-1 virus by interfering with the replication machinery of the virus [24]. Dendrobine was collected from *Dendrobium nobile* and has been investigated for anti-influenza A virus treatment (H1N1, H3N2) [25]. Homonojirimycin was first isolated from *Omphalea diandra*. Recently homonojirimycin extracted from *Commelina communis* was studied for antiviral activity against H1N1 virus [26]. 11-hydroxy Vittatine has been tested for potential activity against H5N1 [27]. Aloperine, a lupine alkaloid exhibited potential activity for anti-influenza A virus treatment. The alkaloid was isolated from *Sophora* sp. plants belonging to Fabaceae family [28]. Atropine was found competent against both DNA HSV-1 and RNA parainfluenza type-3 (PI-3) virus [29]. The present study aims to find a potential natural compound using computational approach i.e. repurposing the alkaloids against active site of M^pro of SARS-CoV-2 and which can further utilize for effectively treating COVID-19. Selected alkaloid molecules were subjected to molecular modelling techniques including

ADME analysis, molecular docking and molecular dynamic simulation to find out their interactions and stability with M^pro of SARS-CoV-2.

2. Material and methods

2.1. Protein/macromolecule

3-dimensional (3D) structure of SARS-CoV-2's main protease (M^pro) (pdb id 6LU7) was obtained from PDB website ("Protein Data Bank"). The structure of 6LU7 is formed of two A chains hence, making the molecule a homodimer. Along with the primary chains, a hetatom named N3 is found as an in-activator of the protein.

2.2. Ligands

PubChem repository ("PubChem") was used to obtain the structure of alkaloids required for the analysis in.sdf format. Later, Biovia Discovery Studio Visualizer v19.1.0.18287 was used to convert the structures into.pdb format. 3D structures of following compounds were used in this study: Lycorine (CID: 72378), Tetrandrine (CID: 73078), Fangchinoline (CID: 73481), Cepharanthine (CID: 10206), Hemanthamine (CID: 441593), Berberine (CID: 2353), Thalimonine (CID: 10893946), 5-alpha-Hydroxysophocarpine (CID: 15385686), Hippeastrine (CID: 441594), Hirsutine (CID: 3037884), Skimmianine (CID: 6760), 13-Methoxydihydronitidine (CID: 38845), Sophaline D (CID: 132991317), Tomatidine (CID: 65576), Emetine (CID: 10219), 11-Hydroxy Vittatine (CID: 602595), Homonojirimycin (CID: 159496), Aloperine (CID: 162147), Dendrobine (CID: 442523) and Atropine (CID: 174174).

2.3. ADME analysis

Lipinski's Rule of Five was applied to 20 compounds that were selected for this investigation. Filters like Molecular weight of the ligand (<500Da), high lipophilicity (LogP<5), Number of hydrogen bonds donors (<5), Number of hydrogen bond acceptors (<10) and Molar refractivity (40–130) (Ghose Rule) were used to carry out the further selection of ligand molecules. Violation of more than 2 of the above stated parameters debarred further analysis of particular molecule [30]. Parameter details were taken from PubChem portal ("PubChem") except molar refractivity. Molar refractivity was calculated with the use of an online tool SwissADME ("SwissADME").
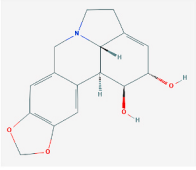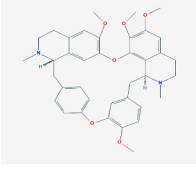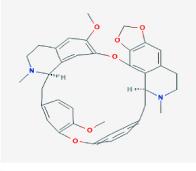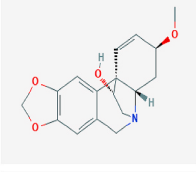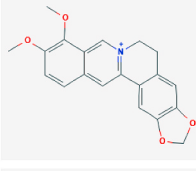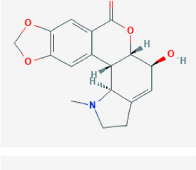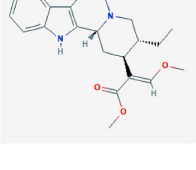
2.4. Molecular docking

Autodock v4.2.6 was used to carry out the molecular docking of each ligand with 6LU7. Both protein and ligand optimization were carried out by removing water, adding polar hydrogens and charges including Kollman and Gasteiger. Finally, the removal of N3 i.e. hetatom from macromolecule concluded the preparation of ligand and protein. A grid box of 60 × 60 × 60 with a spacing of 0.375 Å was used for active site residues. Lamarckian GA output was obtained from docking studies. For each ligand, docking procedure was repeated thrice and out of 10 conformations obtained for each docking run, the final best conformation with minimum binding energy was taken into consideration and converted into a 2-dimensional diagram showing interaction of ligand with active site residues using Biovia Discovery Studio Visualizer v19.1.0.18287.

2.5. Bioavailability radar
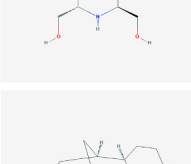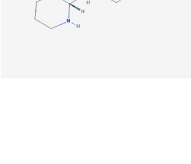
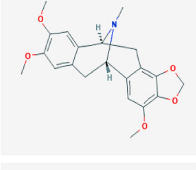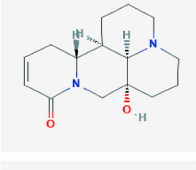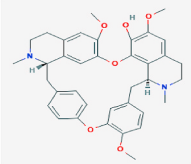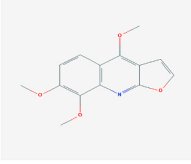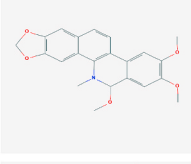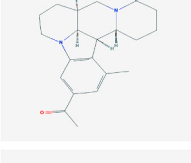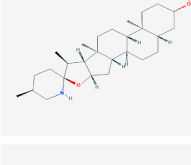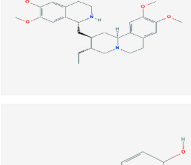A more comprehensible analysis of physiochemical properties was used to continue the filtration of potent ligand molecule. Bioavailability radars of ligands having better results than control compound were obtained using SwissADME web based tool. A total of 6 parameters were used to scrutinize the compounds: solubility, size, polarity, lipophilicity, flexibility and saturation. Ligands deviating from the standardized

Table 1
ADME analysis of Compounds.

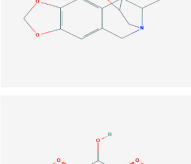| S.No.	Compound Name	Compound Structure	Analysis	
1.	Lycorine		Molecular Weight (<500 Da)	287.31
			Lipophilicity (LogP <5)	0
			H bond donor (<5)	2
			H bond acceptor (<10)	5
			Molar Refractivity (40–130)	78.4
			Violation	0
2.	Tetrandrine		Molecular Weight (<500 Da)	622.7
			Lipophilicity (LogP <5)	6.4
			H bond donor (<5)	0
			H bond acceptor (<10)	8
			Molar Refractivity (40–130)	186.07
			Violation	3
3.	Cepharanthine		Molecular Weight (<500 Da)	606.7
			Lipophilicity (LogP <5)	6.5
			H bond donor (<5)	0
			H bond acceptor (<10)	8
			Molar Refractivity (40–130)	179.15
			Violation	3
4.	Hemanthamine		Molecular Weight (<500 Da)	301.34
			Lipophilicity (LogP <5)	1.3
			H bond donor (<5)	1
			H bond acceptor (<10)	5
			Molar Refractivity (40–130)	83.02
			Violation	0
5.	Berberine		Molecular Weight (<500 Da)	336.4
			Lipophilicity (LogP <5)	3.6
			H bond donor (<5)	0
			H bond acceptor (<10)	4
			Molar Refractivity (40–130)	94.87
			Violation	0
6.	Thalimonine		Molecular Weight (<500 Da)	369.4
			Lipophilicity (LogP <5)	3.1
			H bond donor (<5)	0
			H bond acceptor (<10)	6
			Molar Refractivity (40–130)	103.56
			Violation	0
7.	5-alpha-Hydroxysophocarpine		Molecular Weight (<500 Da)	262.35
			Lipophilicity (LogP <5)	0.7
			H bond donor (<5)	1
			H bond acceptor (<10)	3
			Molar Refractivity (40–130)	103.56
			Violation	0
8.	Hippeastrine		Molecular Weight (<500 Da)	315.32
			Lipophilicity (LogP <5)	0.7
			H bond donor (<5)	1
			H bond acceptor (<10)	6
			Molar Refractivity (40–130)	83.56
			Violation	0
9.	Hirsutine		Molecular Weight (<500 Da)	368.5
			Lipophilicity (LogP <5)	3.4
			H bond donor (<5)	1
			H bond acceptor (<10)	4
			Molar Refractivity (40–130)	110.39
			Violation	0
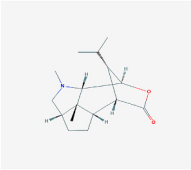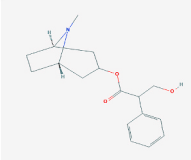
Table 1 (*continued*)

S.No.	Compound Name	Compound Structure	Analysis	
10.	Fangchinoline		Molecular Weight (<500 Da)	608.7
			Lipophilicity (LogP <5)	6.1
			H bond donor (<5)	1
			H bond acceptor (<10)	8
			Molar Refractivity (40–130)	181.6
			Violation	3
11.	Skimmianine		Molecular Weight (<500 Da)	259.26
			Lipophilicity (LogP <5)	2.84
			H bond donor (<5)	0
			H bond acceptor (<10)	5
			Molar Refractivity (40–130)	70.99
			Violation	0
12.	13-Methoxydihydronitidine		Molecular Weight (<500 Da)	379.41
			Lipophilicity (LogP <5)	4.23
			H bond donor (<5)	0
			H bond acceptor (<10)	5
			Molar Refractivity (40–130)	49.39
			Violation	0
13.	Sophaline D		Molecular Weight (<500 Da)	338.44
			Lipophilicity (LogP <5)	2.49
			H bond donor (<5)	0
			H bond acceptor (<10)	2
			Molar Refractivity (40–130)	40.62
			Violation	0
14.	Tomatidine		Molecular Weight (<500 Da)	415.65
			Lipophilicity (LogP <5)	6.21
			H bond donor (<5)	2
			H bond acceptor (<10)	3
			Molar Refractivity (40–130)	127.7
			Violation	1
15.	Emetine		Molecular Weight (<500 Da)	480.64
			Lipophilicity (LogP <5)	4.74
			H bond donor (<5)	1
			H bond acceptor (<10)	6
			Molar Refractivity (40–130)	147.05
			Violation	1
16.	11-Hydroxy Vittatine		Molecular Weight (<500 Da)	287.31
			Lipophilicity (LogP <5)	0.73
			H bond donor (<5)	2
			H bond acceptor (<10)	5
			Molar Refractivity (40–130)	78.29
			Violation	0
17.	Homonojirimycin		Molecular Weight (<500 Da)	193.2
			Lipophilicity (LogP <5)	−2.42
			H bond donor (<5)	6
			H bond acceptor (<10)	6
			Molar Refractivity (40–130)	46.17
			Violation	1
18.	Aloperine		Molecular Weight (<500 Da)	232.36
			Lipophilicity (LogP <5)	1.59
			H bond donor (<5)	1
			H bond acceptor (<10)	2
			Molar Refractivity (40–130)	78.82
			Violation	0

Table 1 (*continued*)

S.No.	Compound Name	Compound Structure	Analysis	
19.	Dendrobine		Molecular Weight (<500 Da)	263.38
			Lipophilicity (LogP <5)	3
			H bond donor (<5)	0
			H bond acceptor (<10)	3
			Molar Refractivity (40–130)	78.41
			Violation	0
20.	Atropine		Molecular Weight (<500 Da)	289.37
			Lipophilicity (LogP <5)	1.83
			H bond donor (<5)	1
			H bond acceptor (<10)	4
			Molar Refractivity (40–130)	84.51
			Violation	0

values suggested non oral bioavailability and hence were debarred from further testing [31].

2.6. Molecular dynamics

Simulation of receptor-ligand docked complex was done using Desmond-Maestro 2020 module. By default, Desmond uses algorithms which are entitled to perform high speed and precision-based MD simulations. For conformational stability of protein-ligand complex, a system setup based on default algorithms was used. The protein-ligand complex was solvated in orthorhombic TIP3P water model. System was neutralized with three counter ions of Na^+ at 0.15 M salt concentration. Optimized potentials for liquid simulations-AA (OPLS-AA) 2005 force field was assigned for the system. Employed SHAKE/RATTLE algorithm limited the movement of all the covalent bonds with default tolerance value of 10^{-7} Å. NVT was utilized as an ensemble class to relax the system before the simulation begins. The simulation was set at 300 K and 1 bar pressure for 50 ns. Thermostat method was set on Nose-Hoover chain with relaxation time of 1ps while Martyna−Tuckerman−Tobias−Klein (MTTK) barostat method was employed with 2ps relaxation time to maintain temperature and pressure. The approaches taken to set up the system was combined by RESPA integrator. A trajectory of 50ns was setup with 1000 frames to show the interaction between protein and ligand based on the literature reports ("User Manual, Impact 6.7" 2015; [32]).

2.7. Biological activity prediction

Prediction of biological activity was done using the PASS webserver ("PASS webserver"). PASS webserver allows us to predict the potential biological effect of the compound using MNA (multilevel neighbors of atoms) descriptors. This particular method prediction is entirely based on the chemical structure of the compound [33,34].

3. Results and discussion

3.1. ADME analysis

The selected 20 compounds were subjected to a preliminary comparison between few properties of ligand molecule (Table 1). This comparison helped to rule out a few compounds which were less likely to be a drug candidate. Along with all the properties as mentioned in Lipinski's Rule of 5, one of the Ghose Rule was also taken into consideration as filter for comparison among the different alkaloids. Qualifying majority of the parameters doesn't guarantee a particular compound to be a drug; it only tells about drug-likeness and helps in eliminating the weak compounds in the preclinical phase only. The result of analysis showed that among the selected 20 alkaloids, tetrandrine, cepharanthine and fangchinoline showed violation of more than two

parameters, hence they are eliminated from further studies. The remaining 17 compounds were considered to be as a potential drug candidate and molecular docking studies were performed to predict the interaction of protein-ligand complex.
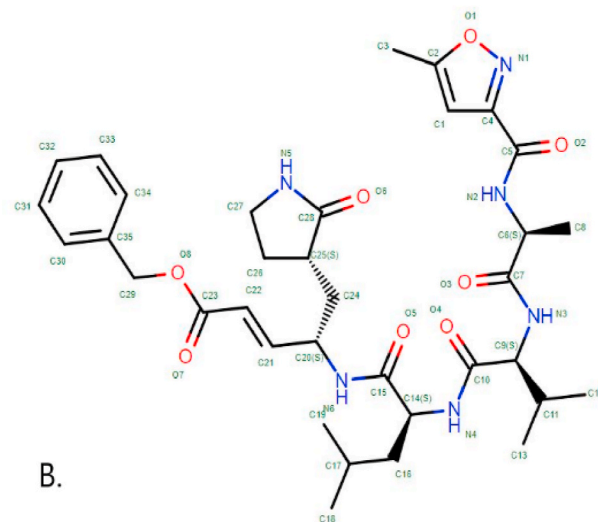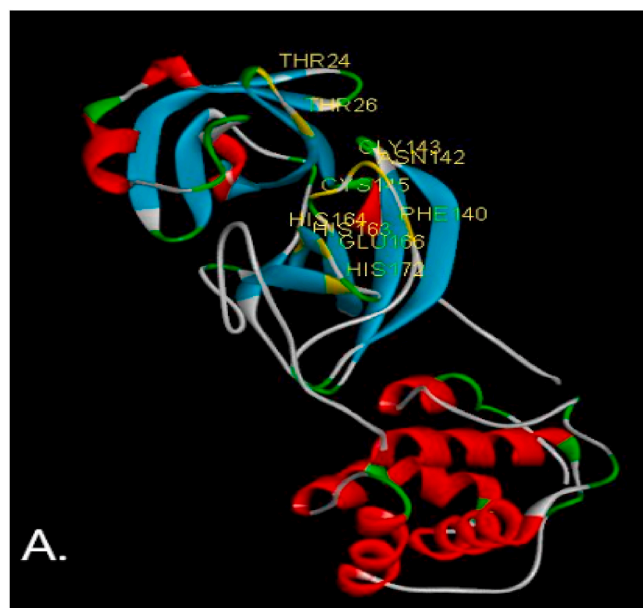

Fig. 1. (A) Structure of M^pro (6LU7), Yellow markings indicate the active site residues (B) Native ligand of the M^pro protein.

Table 2
Comparative results of molecular docking analysis with different ligands.

S. No.	Alkaloid	Binding Energy (ΔG) (Kcal/mol)	Ligand Efficiency	Inhibition Constant (μM)	Intermolecular Energy	Vdw H-Bond Desolvation
1.	Lycorine	−7.41	−0.35	3.73	−8.0	−7.67
2.	Hemanthamine	−7.21	−0.33	5.18	−7.81	−7.19
3.	Berberine	−7.82	−0.31	1.85	−8.42	−8.38
4.	Thalimonine	−8.39	−0.31	0.706	−9.29	−8.84
5.	5-alpha-Hydeoxysophocarpine	−7.04	−0.31	6.96	−7.33	−6.88
6.	Hippeastrine	−7.74	−0.34	2.13	−8.04	−7.52
7.	Hirsutine	−7.7	−0.29	2.28	−9.19	−0.29
8.	Skimmianine	−6.05	−0.32	36.73	−6.95	−6.95
9.	13-Methoxydihydronitidine	−7.55	−0.27	2.91	−8.45	−8.39
10.	Sophaline D	−8.79	−0.35	0.36266	−9.08	−8.96
11.	Tomatidine	−9.58	−0.32	0.09544	−9.88	−9.36
12.	Emetine	−10.17	−0.29	0.03535	−12.25	−10.47
13.	11-Hydroxy Vittatine	−7.14	−0.34	5.85	−7.74	−7.41
14.	Homonojirimycin	−4.05	−0.31	1008	−6.41	−5.61
15.	Aloperine	−6.81	−0.4	10.15	−6.81	−6.35
16.	Dendrobine	−6.6	−0.35	14.44	−7.5	−7.01
17.	Atropine	−7.44	−0.35	3.54	−9.23	−9.01
18.	N3 (Native Inhibitor)	−8.15	−0.17	1.06	−13.52	−13.35

3.2. Molecular docking

Molecular docking is a computational method used in the field of drug designing. Docking results predict the residues of protein that are interacting with ligand's atoms and generates a lowest energy conformation complex. All the seventeen selected alkaloids were docked at the active site of M^{pro} (PDB ID: 6LU7) protein of SARS-CoV-2. THR24, THR26, PHE140, ASN142, GLY143, CYS145, HIS163, HIS164, GLU166, HIS172 (Fig. 1A) were found at the active site of the M^{pro} protein. The binding energy was taken as the comparison parameter and N3 (Native inhibitor) was considered as control molecule. Binding energy of N3 with 6LU7 was −8.15 kcal/mol (Table 2 and Fig. 2E). Thalimonine, Emetine, Sophaline D and Tomatidine were showing better binding energy than N3.

Thalimonine-M^{pro} complex showed the least binding energy of −8.39 kcal/mol. SER144 and CYS145 were forming conventional H bonds with CYS145 forming 2 H bonds along with Pi-donor and Alkyl interactions. Electrostatic bond formation was seen in case of HIS41 with both benzene rings present in the compound. Other interactions were also observed with HIS41 residue like Pi-Pi cation (Fig. 2A) and Pi-alkyl. Two sulfur bond formation was taking place with sulfur containing residues i.e. MET165 and MET49. One more alkyl bond formation takes place with MET49. HIS164 and ASP187 were forming a Pi-donor H-bond with the ligand.

Sophaline D-M^{pro} docked complex gave −8.79 kcal/mol as the minimum binding energy. Interaction can be classified into 6 categories namely: Conventional H-bond, Alkyl, pi-alkyl, carbon hydrogen and van der Waal bonds. Among the various residues, HIS163 was seen bonding via conventional hydrogen bond. HIS41 formed pi-alkyl bond and CYS145 and MET165 can were bound to ligand by alkyl bond formation. HIS164 residue was found interacting via carbon hydrogen bond. 12 residues stabilizing the complex by van der Waals attractive force can be observed too (Fig. 2B).

Tomatidine-M^{pro} complex was found to have minimum binding energy of −9.58 kcal/mol. Only 1 conventional hydrogen bond was observed with HIS164 residue. GLN189 bonded using carbon hydrogen bond. LEU141, MET49 and CYS145 interacted with the ligand by forming alkyl bonds while due to presence of ring structure in the residues HIS41, HIS163 and HIS172 formed pi-alkyl type of bond with tomatidine. Along with these major interaction 8 more van der Waal attractive interaction were seen in the bound protein-ligand complex (Fig. 2C).

Emetine-M^{pro} complex showed minimum binding energy of −10.17 kcal/mol. A total of 6 types of bond formation was observed in the docked complex. HIS163 and GLU166 were forming conventional H-bond with emetine. GLU166 was also interacting using a pi-anion interaction. MET165 alone was forming alkyl as well as pi alkyl bonds. LEU141 and HIS164 formed carbon hydrogen bond with the ligand. In the surrounding 11 other residues can be seen interacting through weak van der Waals interaction thereby, adding to the stability of the overall conformation (Fig. 2D).

3.3. Bioavailability radar

Comparison with control (N3) left us with 4 ligands viz., thalimonine, emetine, tomatidine and sophaline D. Observing bioavailability radars of four compounds revealed that all the ligands except tomatidine were found orally bioavailable. Tomatidine failed to follow the standard values of lipophilicity and insolubility. Other three ligands can be seen to lying in the shaded pink region of standard values (Fig. 6).

3.4. MD simulations

Three compounds were shortlisted, thalimonine, sophaline D and emetine. 6LU7 in complex with mentioned three ligands was subjected to a 50ns MD simulation individually. The average potential energy of the thalimonine-M^{pro}, sophaline D-M^{pro} and emetine-M^{pro} complexes was found to be −115,805 kcal/mol, −117,514 kcal/mol and −117,360 kcal/mol respectively. Along with potential energy total energy parameter was also calculated during simulation time, −93,656 kcal/mol, −95,415 kcal/mol and −95,270 kcal/mol respectively were the total energy values obtained.

3.5. Structural deviation and compactness

Upon binding of a small molecule, the stability of the conformation of the protein is a fundamental property to explore and thus, gain insight about changes and dynamics of the protein structure. Root mean square deviation (RMSD), Root mean square fluctuation (RMSF), Radius of Gyration (rGyr) and Solvent accessible surface area (SASA) plots can be utilized to estimate the properties like compactness of the protein.

RMSD plot (Fig. 3A) of thalimonine-M^{pro} complex can be observed getting stabilized at the end of the simulation suggesting protein stability after binding of the ligand. Though, a lot of steep peaks can be seen during the simulation which are a result of internal vibration of molecule. A maximum deviation of 1 Å was seen which is considered acceptable for the protein stability. RMSD graph (Fig. 4A) of 6LU7 in complex with sophaline D ligand can be observed ending with a lower
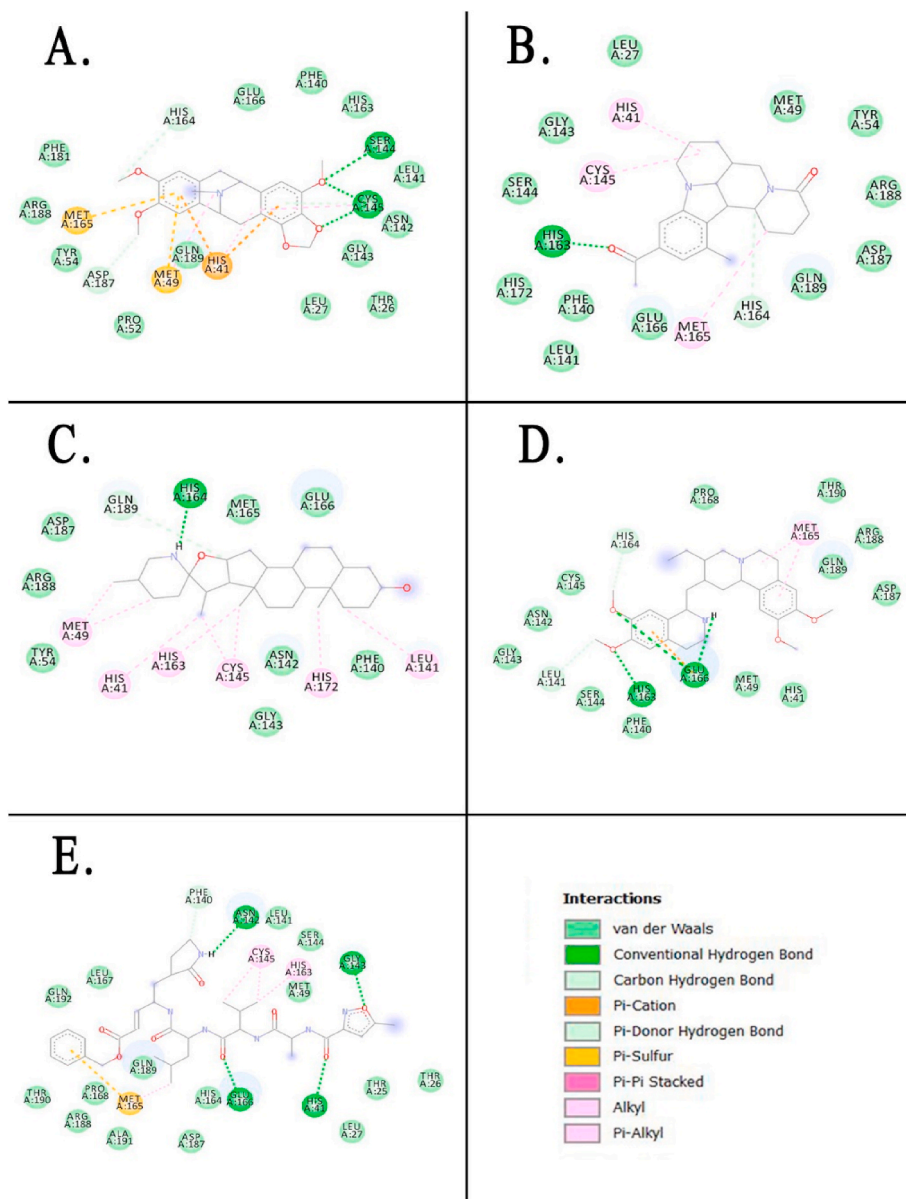
Fig. 2. (A) Interaction of thalimonine with M^{pro} (B) interaction of sophaline D with M^{pro} (C) interaction of tomatidine with M^{pro} (D) interaction of emetine with M^{pro} (E) interaction of N3 (control) with M^{pro}.

RMSD indicating stabilized protein. A sudden peak with deviation of nearly 1 Å can be seen at about 45ns which might be the result of sudden internal vibration in the protein but henceforth from 45 to 50ns RMSD plot can be seen going towards lower numbers in the graph. RMSD plot (Fig. 5A) of M^{pro} with emetine was obtained and during the most time of the simulation the values can be seen fluctuating around a fixed number. It was only after 40ns of the simulation of RMSD plot went a little higher than before. A structural deviation was seen in the trajectory and the deviation is nearly 1 Å only hence, the elevated graph can be attributed to internal vibrations.

RMSF plot helped in understanding the residual vibrations in three different complexes. Since, alpha helixes and beta strands are rigid hence, less fluctuation should be seen with the residues involved in such secondary structures. RMSF plot (Fig. 3B) of 6LU7 with thalimonine can be observed to follow the same principle with high peaks at the unstructured part of the protein. The residues of the binding pocket can be seen obtaining lower values of the graph indicating less conformational change at the binding pocket. Similarly, RMSF plot (Fig. 4B) of M^{pro} with sophaline D can be observed to follow the same lines, residues making

alpha helixes and beta sheets with the lower values. Relatively, the RMSF graph obtains less values with less peaks. RMSF graph of 6LU7-emetine (Fig. 5B) complex can be seen obtaining relatively higher values and peaks. Though the residues at binding site and secondary structure obtained smaller values. Also, in all three graphs the B-factor graph can be running mostly parallel to the RMSF plot indicating towards protein stability.

Radius of gyration (rGyr) is yet another parameter which is linked with the tertiary structure and general conformational state defining our understanding of compactness and folding of the protein. All three complexes were allowed to yield the rGyr plot. A fluctuation of 0.1 Å, 0.1 Å and 0.4 Å around the average values were observed for M^{pro}-thalimonine (Fig. 3C), M^{pro}-sophaline D (Fig. 4C) and M^{pro}-emetine (Fig. 5C) respectively. Though, in case of M^{pro}-emetine complex a sudden drop in rGyr was observed between 15ns and 25ns but immediately after 25ns the graph was restored to its average value. This sudden deviation might be due to protein's packing. All three plots attained the equilibrium around their average value thus, suggesting complex stoutness through the simulation course.
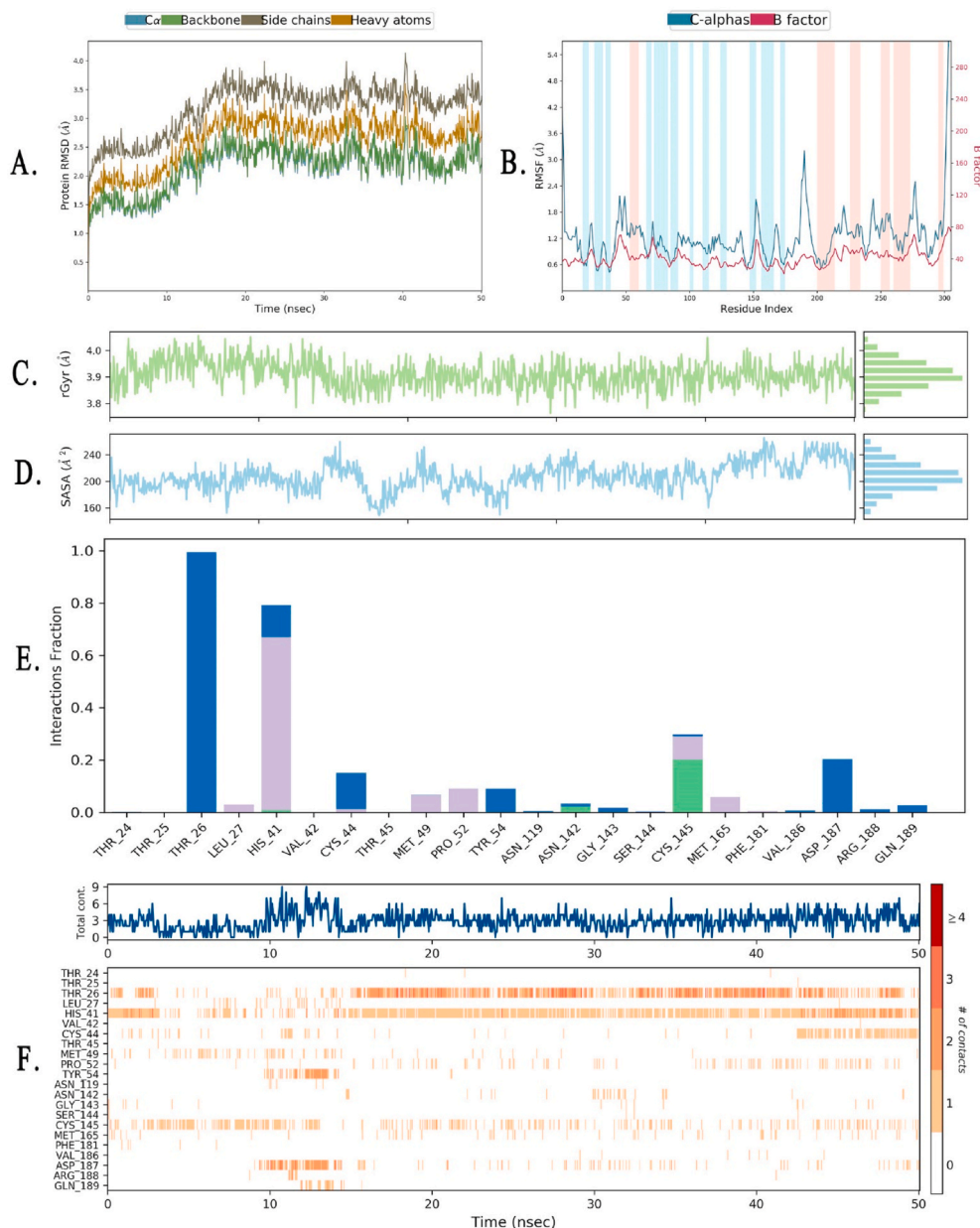
Fig. 3. Thalimonine dynamic results (A) RMSD plot (B) RMSF (Background pink strips represent alpha-helix and blue strip beta sheets) (C) Radius of Gyration (rGyr) (D) Solvent Accessible Surface Area (SASA) (E) Protein-Ligand Contacts (Green = H-bond, Purple = Hydrophobic, Blue = Water bridges) (F) Timeline of ligand contacts.

SASA is the surface area accessible to the water molecule and study of SASA can be used to investigate the conformational dynamics. Average value of SASA observed were: 200 Å^2, 120 Å^2 and 450 Å^2 respectively for 6LU7-thalimonine (Fig. 3D), 6LU7-sophaline D (Fig. 4D) and 6LU7-emetine (Fig. 5D) complex. In each case result, the equilibrium is reached around the average values. Emetine showed highest SASA value suggesting larger surface are in contact with water molecules i.e., more number of inner residues coming in contact with the solvent while thalimonine and sophaline D complexes are more stable with less inner residues interacting with surroundings. Though the attained equilibrium in each case suggests that the complexes are structurally stable.

3.6. Interaction dynamics and secondary structural analysis

Computing the 6LU7-thalimonine complex, a new residue, THY25

interacted with ligand almost 100% of the simulation time via water bridge formation. Out of all residues, four residues i.e. CYS145, ASN142, HIS41 and TYR54 formed H-bond with ligand. HIS41 showed pi-pi stacking majority of the simulation time and forming water bridges for the rest of the remaining simulation time. Majority of the amino acids including THR24, CYS44, TYR54, ASN119, ASN142, GLY143, SER144, CYS 145, VAL 186, ASP187, ARG188 and GLN189 were involved in water bridge bond formation which were slightly less weak than conventional H-bonds. Other residues like LEU27, MET49, PRO52, MET165 were involved in hydrophobic interaction which include pi-pi stacking, pi-cation and other non-specific interactions (Fig. 3E). Fig. 3F represents a timeline representation of number of contacts made by residues with ligand. The upper panel showed total number of contacts at any instant throughout the trajectory course. On average at any given time 3–4 contacts were made by the 6LU7 residues. The lower panel showed which residue makes contact with ligand molecule at any instantaneous
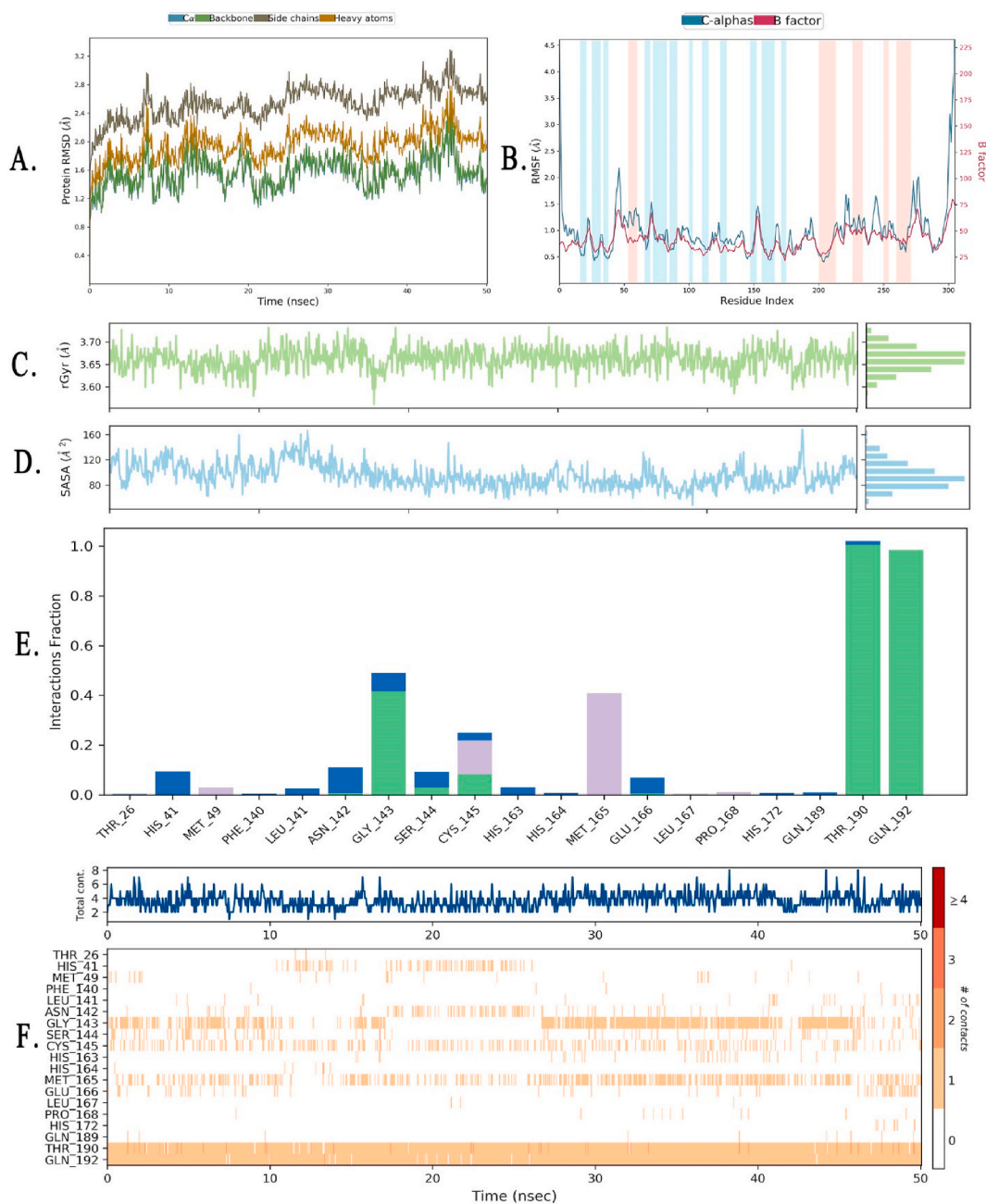
Fig. 4. Sophaline D dynamic results (A) RMSD plot (B) RMSF (Background pink strips represent alpha-helix and blue strip beta sheets) (C) Radius of Gyration (rGyr) (D) Solvent Accessible Surface Area (SASA) (E) Protein-Ligand Contacts (Green = H-bond, Purple = Hydrophobic, Blue = Water bridges) (F) Timeline of ligand contacts.

moment of the trajectory. During the middle and later stages of the simulation THR26 was interacting with ligand in a significant amount, often making more than 4 contacts. HIS41 was also stabilizing the ligand over 50ns time. At the later stages of the trajectory CYS44 also plays role in stabilizing the overall interaction with TYR54, CYS145 and ASP187.

Considering the 6LU7-sophaline D complex, in total of 5 H-bonds can be observed. THR190 and GLN192 both residues interacted with the residue for 100% of the simulation time. Along with these residues, GLY143, CYS145 and SER144 were also seen to form H-bond with the ligand for a brief period of time. Few residues were also seen to change the nature of bond formation. GLY143 for majority of interaction time formed H-bond while for a brief time period also interacted through water bridge formation while SER144 interacting in a vice-versa manner. CYS145 was seen to deviate along all three types of

interaction, for almost equal amount of fraction, H-bond and hydrophobic interactions can be observed while water bridge relatively for a lower fraction. On the other and MET145, MET49, PRO168 interacted with hydrophobic interactions. Other residues like HIS41, LEU141, ASN142, HIS163, GLU166, GLN189 all contributed towards the stability of the complex via water bridge bond formation (Fig. 4E). Fig. 4F shows the timeline representation of the residues interacting with the ligand at any specific time. An average of 4–6 bonds were observed with the ligand at any given point of time during the 50ns simulation. GLN192 and THR190 can be seen to hold the ligand in the binding pocket during the full 50ns time. While other residues like HIS41, CYS145, GLY143 and MET165 binding with the ligand in a very scattered manner. Throughout the trajectory at no given period of time the ligand is devoid of any interaction.
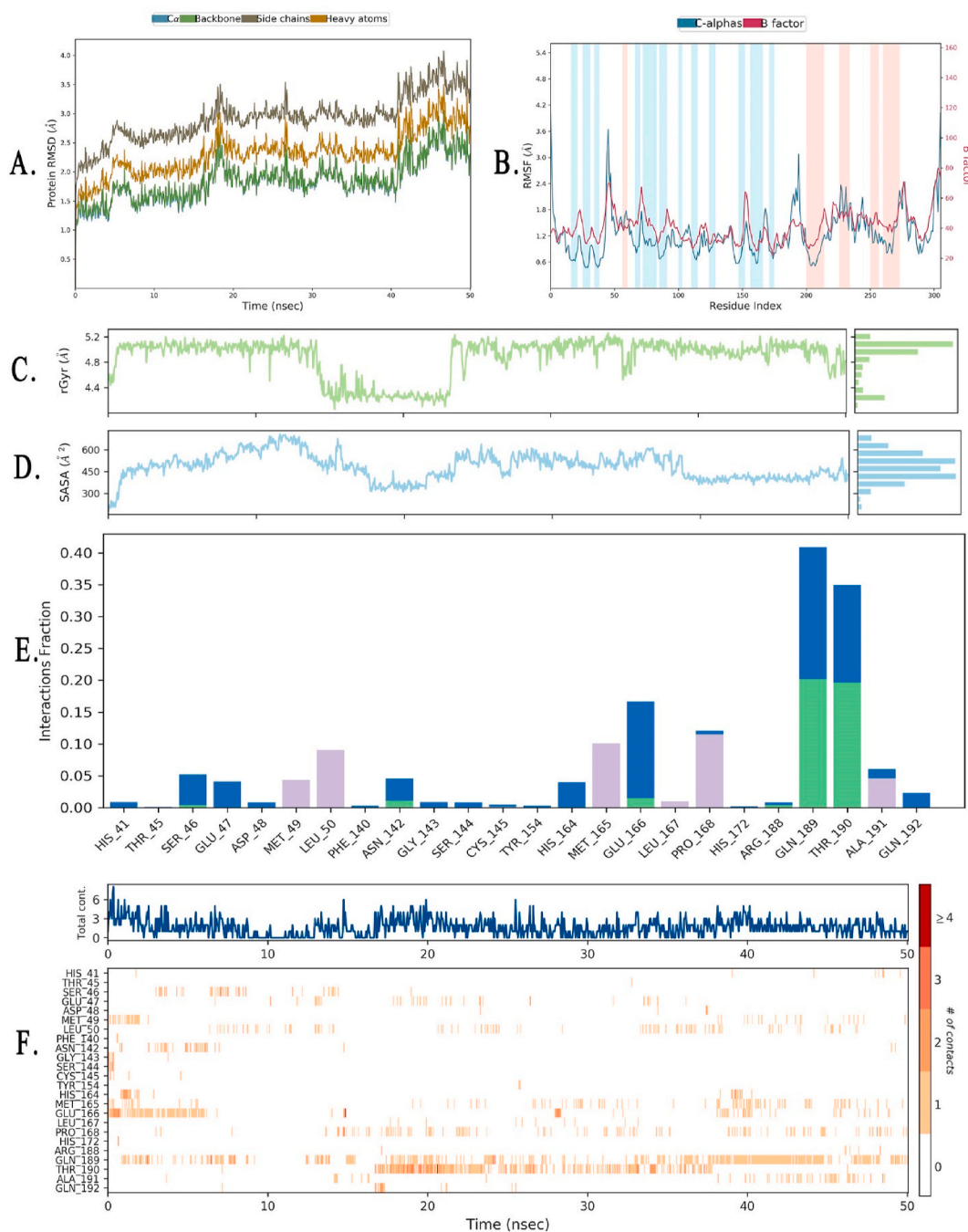
Fig. 5. Emetine dynamic results (A) RMSD plot (B) RMSF (Background pink strips represent alpha-helix and blue strip beta sheets) (C) Radius of Gyration (rGyr) (D) Solvent Accessible Surface Area (SASA) (E) Protein-Ligand Contacts (Green = H-bond, Purple = Hydrophobic, Blue = Water bridges) (F) Timeline of ligand contacts.

In the case of 6LU7-emetine complex, the highest interaction fraction can be observed to be only 40%. Many residues formed 2 types of bonds. GLN189 and THR190 formed H-bond and water bridge interaction for almost equal time. ASN142, GLU166 and SER46, all interacted via water bridge formation for majority of their respective times and very briefly by forming H-bond with the ligand. PRO168 and ALA199 both formed hydrophobic interactions along with water bridges. 12 other residues individually formed water bridges and hydrophobic interactions for very short fraction of time contributing to the complex stability (Fig. 5E). The timeline of the residues in contact with ligand can be seen in Fig. 5F. Observations from the upper panel suggested 2 contacts to an average number of contacts. We can also see significant number of intervals were no contact with the ligand was made. The lower panel

confirms the same. THR190 and GLN189 were seen interacting most with emetine though, contacts are scattered. Other residues also make contact with ligand but overall, the contacts are distributed unevenly with intermediate complete absence of any interactions.

Secondary structural dynamics can be used to understand the conformational changes in the protein and folding behavior of the protein. It is a result of changes in the residue upon binding of different ligands. Table 4 represents the percentage composition of secondary structure of the whole protein. Analysis showed that with thalimonine as the ligand, 41.87% of all the residues were involved in secondary structure formation. This number was observed slightly decreased with sophaline D as the ligand (41.56%) and with the emetine as the ligand molecule further drop was observed (40.19%). Both alpha-helixes and
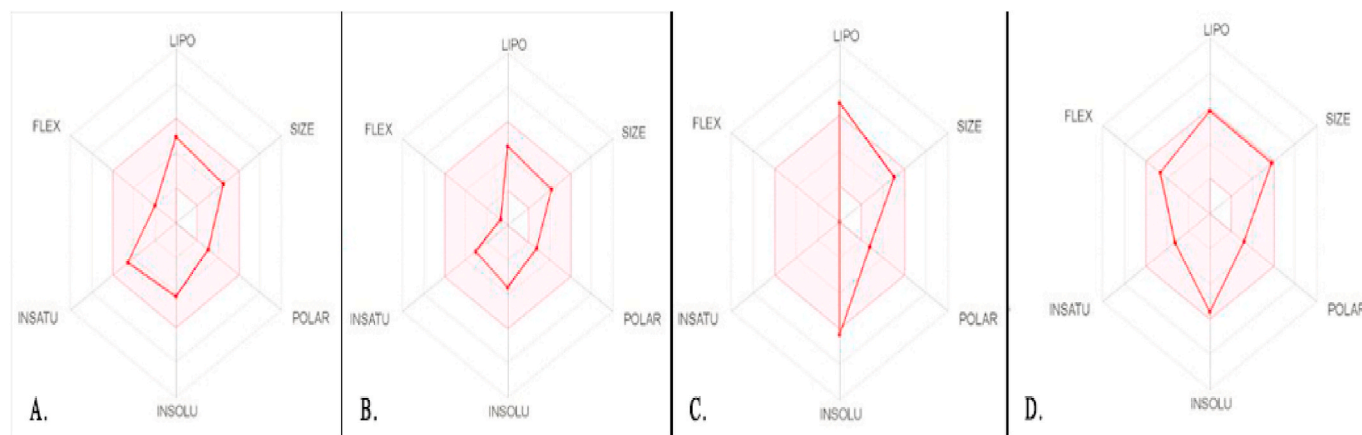
Fig. 6. Bioavailability radar (A) thalimonine (B) sophaline D (C) tomatidine (D) emetine.

Table 3
Biological activity prediction of the selected three ligands.

Compound Name	Pa	Pi	Biological Activity
Thalimonine	0.827	0.003	Antitussive
Sophaline D	0.325	0.204	Antiviral (Rhinovirus)
Emetine	0.429	0.016	Antitussive

Pa = probability to be active; Pi = probability to be inactive.

Table 4
Protein Secondary Structure Element (SSE) analysis.

Compound Name	% Total SSE	% Alpha Helix	% Beta Strand
Thalimonine	41.87	17.83	24.04
Sophaline D	41.56	17.56	24.00
Emetine	40.19	16.56	23.63

beta-strands were reduced.

3.7. Biological activity prediction

PASS server was being used to make the effective biological activity predictions. The analysis resulted in similar biological activity for the selected three compounds. Thalimonine, sophaline D and emetine have shown antitussive, antiviral (Rhinovirus) and antitussive predictions respectively. All the biological activities predicted were related to respiratory ailments and Pa ranged from 0.325 to 0.827 when Pi < Pa. Table 3 shows the predicted activities of the three compounds.

M^{pro} of COVID19 is predicted to contain 306 amino acid residues and structural analysis showed that it has total of 3 domains. The binding site is located in-between domain I and domain II. Both domains contribute a single residue for the formation of a catalytic site which is essential for substrate binding. Domain I contain His41 residue and domain II contains Cys145 residue which are conserved in SARS-CoV family [6]. From literature survey it was found that His41 acts as a base and –SH group of Cys145 mediates the electrostatic trigger to start the chemical reaction and in collaboration it is known as Cys-His dyad. Functional M^{pro} is an essential requirement for maturation of the virus and the catalytic dyad is essential to carry out substrate conversion by M^{pro}. Even a slight mutation in either of the residues leads to enzyme inactivation [35]. Hence, considering the genomic and structural similarities the dyad becomes a potential target for M^{pro} of Covid 19. Comparing all the filters, parameters and mechanisms, thalimonine and sophaline D can be considered for the further analysis to inhibit the M^{pro} of COVID-19. Both were predicted to have respiratory track related biological activity and can be seen interacting with the catalytic dyad of the 6LU7. Emetine showed the maximum negative binding energy of all docked complexes

but the lack of interaction with the catalytic dyad residues and frequent absence of any contact with ligand during simulation makes emetine a hard choice.

4. Conclusion

For combating the pandemic caused by the COVID19 many studies are aiming to find antiviral therapies. Most of the studies as of now are concentrated towards finding a cure in chemically synthesized or earlier used drugs (drug repurposing) for other diseases. In this study, the aim was to find a natural solution which can effectively solve the problem. Considering that as a fact, a five-step filter study was conducted consisting of ADME analysis, followed by molecular docking, bioavailability radar analysis, molecular dynamic simulation and prediction of biological activity on different alkaloids which are known to show some antiviral activity. In this study, all the twenty alkaloids which were chosen filtered through ADME parameters. Further, the compounds with binding energy lower than N3, thalimonine, sophaline D, tomatidine and emetine were used for bioavailability testing which ruled out tomatidine and further the remaining three were subjected to molecular dynamic simulation study. After the MD simulation biological activity was also predicted. Overall, the study concludes two alkaloids namely thalimonine and sophaline D having potential activity to inhibit the M^{pro} but to confirm the claim further in-vitro studies need to be conducted.

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.

References

[1] RS Houlston, J Cheadle, SE Dobbins, A Tenesa, AM Jones, K Howarth, SL Spain, P Broderick, E Domingo, S Farrington, JGD Prendergast, AM Pittman, E Theodoratou, CG Smith, B Olver, A Walther, RA Barnetson, M Churchman, EEM Jaeger, S Penegar, E Barclay, L Martin, M Gorman, R Mager, E Johnstone, R Midgley, I Niittymäki, S Tuupanen, J Colley, S Idziaszczyk, COGENT Consortium, HJW Thomas, AM Lucassen, DGR Evans, ER Maher, CORGI Consortium, COIN Collaborative Group, Collaborative Group, T Maughan, A Dimas, E Dermitzakis, JB Cazier, LA Aaltonen, P Pharoah, DJ Kerr, LG Carvajal-Carmona, H Campbell, MG Dunlop, IPM Tomlinson, Meta-analysis of three genome-wide association studies identifies susceptibility loci for colorectal cancer at 1q41, 3q26.2, 12q13.13 and 20q13.33, Nat Genet 42 (11) (2010) 973–977, https://doi.org/10.1038/ng.670.
[2] Bogoch II, A. Watts, A. Thomas-Bachli, C. Huber, M.U.G. Kraemer, K. Khan, Pneumonia of unknown aetiology in Wuhan, China: potential for international spread via commercial air travel, J. Trav. Med. 27 (2) (2020), https://doi.org/10.1093/jtm/taaa008.
[3] S. Zhao, Q. Lin, J. Ran, S.S. Musa, G. Yang, W. Wang, Y. Lou, D. Gao, L. Yang, D. He, M.H. Wang, Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: a data-driven

analysis in the early phase of the outbreak, Int. J. Infect. Dis. 92 (2020) 214–217, https://doi.org/10.1016/j.ijid.2020.01.050.

[4] Y.S. Malik, S. Sircar, S. Bhat, K. Sharun, K. Dhama, M. Dadar, R. Tiwari, W. Chaicumpa, Emerging novel coronavirus (2019-nCoV)—current scenario, evolutionary perspective based on genome analysis and recent developments, Vet. Q. 40 (1) (2020) 68–76, https://doi.org/10.1080/01652176.2020.1727993.

[5] W.C. Culp, Coronavirus disease 2019, In Pract. 14 (6) (2020), https://doi.org/10.1213/xaa.0000000000001218 e01218.

[6] M.U. Mirza, M. Froeyen, Structural elucidation of SARS-CoV-2 vital proteins: computational methods reveal potential drug candidates against main protease, Nsp12 polymerase and Nsp13 helicase, J. Pharm. Anal. (2020), https://doi.org/10.1016/j.jpha.2020.04.008.

[7] 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 α-ketoamide inhibitors, Science (2020), https://doi.org/10.1126/science.abb3405 eabb3405.

[8] A. Roy, A review on the alkaloids an important therapeutic compound from plants, 10.231/JIM.0b013e3181948b37, Int. J. Plant Biotechnol. 3 (2) (2017) 1–9.

[9] Saksham Garg, Ashutosh Anand, Yash Lamba, Arpita Roy, Molecular docking analysis of selected phytochemicals against SARS-CoV-2 Mpro receptor, Vegetos (2020), https://doi.org/10.1007/s42535-020-00162-1. In press.

[10] D. Lamoral-Theys, A. Andolfi, G Van Goietsenoven, A. Cimmino, B Le Calvé, N. Wauthoz, V. Mégalizzi, T. Gras, C. Bruyère, J. Dubois, V. Mathieu, A. Kornienko, R. Kiss, A. Evidente, Lycorine, the main phenanthridine Amaryllidaceae alkaloid, exhibits significant antitumor activity in cancer cells that display resistance to proapoptotic stimuli: an investigation of Structure–Activity relationship and mechanistic insight, J. Med. Chem. 52 (20) (2009) 6244–6256, https://doi.org/10.1021/jm901031h.

[11] L. Szlávik, Minárovits J. Gyuris Á, P. Forgo, J. Molnár, J. Hohmann, Alkaloids from leucojum vernum and antiretroviral activity of Amaryllidaceae alkaloids, Planta Med. 70 (9) (2004) 871–873, https://doi.org/10.1055/s-2004-827239.

[12] M. Ieven, D.A.V.A. van den Berghe, Plant antiviral agents. IV. Influence of lycorine on growth pattern of three animal viruses, Planta Med. 49 (2) (1983) 109–114.

[13] J. He, W.B. Qi, L. Wang, J. Tian, P.R. Jiao, G.Q. Liu, W.C. Ye, M. Liao, Amaryllidaceae alkaloids inhibit nuclear-to-cytoplasmic export of ribonucleoprotein (RNP) complex of highly pathogenic avian influenza virus H5N1, Influ. Other Respi. Virus. 7 (6) (2013) 922–931, https://doi.org/10.1111/irv.12035.

[14] Y.Q. Yan, Y.J. Fu, S. Wu, H.Q. Qin, X. Zhen, B.M. Song, Y.S. Weng, P.C. Wang, X. Y. Chen, Z.Y. Jiang, Anti-influenza activity of berberine improves prognosis by reducing viral replication in mice, Phyther. Res. 32 (12) (2018) 2560–2567, https://doi.org/10.1002/ptr.6196.

[15] J. Serkedjieva, M. Velcheva, Vitro anti-influenza virus activity of the pavine alkaloid (-)-Thalimonine isolated from Thalictrum simplex L, Antivir. Chem. Chemother. 14 (2) (2003) 75–80, https://doi.org/10.1177/095632020301400202.

[16] H. Takayama, Y. Iimura, M. Kitajima, N. Aimi, K. Konno, H. Inoue, M. Fujiwara, T. Mizuta, T. Yokota, S. Shigeta, K. Tokuhisa, Y. Hanasaki, K. Katsuura, Discovery of anti-influenza A virus activity of a corynanthe-type indole alkaloid, hirsutine, in vitro and the structure-activity relationship of natural and synthetic analogs, Bioorg. Med. Chem. Lett 7 (24) (1997) 3145–3148, https://doi.org/10.1016/S0960-894X(97)10154-8.

[17] P.L. Ding, H. Huang, P. Zhou, D.F. Chen, Quinolizidine alkaloids with anti-HBV activity from Sophora tonkinensis, Planta Med. 72 (9) (2006) 854–856, https://doi.org/10.1055/s-2006-946639.

[18] Z. Wan, Y. Lu, Q. Liao, Y. Wu, X. Chen, Fangchinoline inhibits human immunodeficiency virus type 1 replication by interfering with gp160 proteolytic processing, PLoS One 7 (6) (2012) e39225,, https://doi.org/10.1371/journal.pone.0039225.

[19] S. Hu, J. Dutt, T. Zhao, C.S. Foster, Tetrandrine potently inhibits herpes simplex virus type-1-induced keratitis in BALB/c mice, Ocul. Immunol. Inflamm. 5 (3) (1997) 173–180, https://doi.org/10.3109/09273949709116892.

[20] Min Kim, Lee Jang, Song Shin, Kim Kim, Kwon Jin, Natural bis-benzylisoquinoline alkaloids-tetrandrine, fangchinoline, and cepharanthine, inhibit human coronavirus OC43 infection of MRC-5 human lung cells, Biomolecules 9 (11) (2019) 696,, https://doi.org/10.3390/biom9110696.

[21] G. Yang, D. Chen, Alkaloids from the roots of Zanthoxylum nitidum and their antiviral and antifungal effects, Chem. Biodivers. 5 (9) (2008) 1718–1722, https://doi.org/10.1002/cbdv.200890160.

[22] Y.B. Zhang, X.L. Zhang, N.H. Chen, Z.N. Wu, W.C. Ye, Y.L. Li, G.C. Wang, Four matrine-based alkaloids with antiviral activities against HBV from the seeds of Sophora alopecuroides, Org. Lett. 19 (2) (2017) 424–427, https://doi.org/10.1021/acs.orglett.6b03685.

[23] B. Troost, L.M. Mulder, M. Diosa-Toro, van de Pol D, I.A. Rodenhuis-Zybert, J. M. Smit, Tomatidine, a natural steroidal alkaloid shows antiviral activity towards chikungunya virus in vitro, Sci. Rep. 10 (1) (2020) 1–12, https://doi.org/10.1038/s41598-020-63397-7.

[24] A.L.C. Valadão, C.M. Abreu, J.Z. Dias, P. Arantes, H. Verli, A. Tanuri, RS De Aguiar, Natural plant alkaloid (Emetine) inhibits HIV-1 replication by interfering with reverse transcriptase activity, Molecules 20 (6) (2015) 11474–11489, https://doi.org/10.3390/molecules200611474.

[25] R. Li, T. Liu, M. Liu, F. Chen, S. Liu, J. Yang, Anti-influenza A virus activity of dendrobine and its mechanism of action, J. Agric. Food Chem. 65 (18) (2017) 3665–3674, https://doi.org/10.1021/acs.jafc.7b00276.

[26] G.B. Zhang, B. Zhang, X.X. Zhang, F.H. Bing, Homonojirimycin, an alkaloid from dayflower inhibits the growth of influenza A virus in vitro, Acta Virol. 57 (1) (2013) 85–86, https://doi.org/10.4149/av_2013_01_85.

[27] M.T. Moradi, A. Karimi, Z. Lorigooini, Alkaloids as the natural anti-influenza virus agents: a systematic review, Toxin Rev. 37 (1) (2018) 11–18, https://doi.org/10.1080/15569543.2017.1323338.

[28] Z. Dang, K. Jung, L. Zhu, W. Lai, H. Xie, K.H. Lee, L. Huang, C.H. Chen, Identification and synthesis of quinolizidines with anti-influenza A virus activity, ACS Med. Chem. Lett. 5 (8) (2014) 942–946, https://doi.org/10.1021/ml500236n.

[29] B. Özçelik, M. Kartal, I. Orhan, Cytotoxicity, antiviral and antimicrobial activities of alkaloids, flavonoids, and phenolic acids, Pharm. Biol. 49 (4) (2011) 396–402, https://doi.org/10.3109/13880209.2010.519390.

[30] C.A. Lipinski, Lead- and drug-like compounds: the rule-of-five revolution, Drug Discov. Today Technol. 1 (4) (2004) 337–341, https://doi.org/10.1016/j.ddtec.2004.11.007.

[31] 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 (1) (2017) 42717,, https://doi.org/10.1038/srep42717.

[32] V. Kumar, J.K. Dhanjal, S.C. Kaul, R. Wadhwa, D. Sundar, Withanone and caffeic acid phenethyl ester are predicted to interact with main protease (M pro) of SARS-CoV-2 and inhibit its activity, J. Biomol. Struct. Dyn. (2020) 1–13, https://doi.org/10.1080/07391102.2020.1772108.

[33] R.K. Goel, D. Singh, A. Lagunin, V. Poroikov, PASS-assisted exploration of new therapeutic potential of natural products, Med. Chem. Res. 20 (9) (2011) 1509–1514, https://doi.org/10.1007/s00044-010-9398-y.

[34] M. Gulzar, S.B. Syed, F.I. Khan, P. Khan, S. Ali, G.M. Hasan, P. Taneja, M.I. Hassan, Elucidation of interaction mechanism of ellagic acid to the integrin linked kinase, Int. J. Biol. Macromol. 122 (2019) 1297–1304, https://doi.org/10.1016/j.ijbiomac.2018.09.089.

[35] G.G. Chang, Quaternary Structure of the SARS Coronavirus Main Protease. Molecular Biology of the SARS-Coronavirus, Springer, Berlin Heidelberg, Berlin, Heidelberg, 2010, pp. 115–128, https://doi.org/10.1007/978-3-642-03683-5_8.