ChemistrySelect

Research Article
doi.org/10.1002/slct.202200055

Chemistry Europe
European Chemical Societies Publishing

www.chemistryselect.org

Medicinal Chemistry & Drug Discovery

Identification of Alkaloids from *Terminalia chebula* as Potent SARS- CoV-2 Main Protease Inhibitors: An *In Silico* Perspective

Rajesh Ghosh[+],[a] Vishnu Nayak Badavath[+],[b] Snehasis Chowdhuri,*[a] and Anik Sen*[c]

Natural compounds in medicinal plants are best remedies for different diseases and are important to develop new drugs. This work was dedicated to understand the role of different natural compounds of *Terminalia Chebula*, a well-known herbal plant, in the treating of Covid 19. In this article, we have investigated interactions of such natural compounds from *Terminalia Chebula* with the main protease (M^{pro}) of the SARS-CoV-2, which is a key component for cleavage of viral polyprotein, and an important target for the development of drugs towards COVID-19. We have performed molecular docking study on 22 different molecules of *Terminalia Chebula* and proposed that 7 of the natural compounds (triterpenoids and sterols) interacts with a comparable or stronger interactions than the inhibitor N3. Molecular dynamics simulations (100 ns) revealed that 7 M^{pro}-*Terminalia Chebula* complexes are stable, conformationally less fluctuated, slightly less compact, and marginally expanded than ligand-free conformation of M^{pro}. The intermolecular H-bonding and detailed MM/PBSA and MM-GBSA analysis showed Daucosterol interaction to be the most strong, whereas comparable interactions were observed for Arjunetin, Maslinic acid, and Bellericoside. Our study suggested that these natural compounds can act as potent M^{pro} inhibitors for SARS-CoV-2, and may evolve as promising anti-COVID-19 drugs in the near future.

Introduction

Since December 2019, the novel coronavirus (SARS-CoV-2) has created an outbreak of respiratory disease worldwide and is a successor of the two previous massive coronavirus outcomes, severe acute respiratory syndrome coronavirus (SARS-CoV) in 2002–03, and Middle-East respiratory syndrome coronavirus (MERS-CoV) in 2012.[1] The virus was named SARS-CoV-2[2–4] due to a great similarity of the RNA genome with the SARS-CoV. The acceleration of the spread of the virus was due to human-to-human cough, sneeze, and touch from an infected person i.e. through the common droplet infection.[4–6] The present studies on such disease have been categorized into three different groups (a) Vaccine development, (b) Designing new drugs, and (c) Repurposing of drugs and here we have focused on the third. The SARS CoV-2 M^{pro} is a member of homologous cysteine proteases that are needed for viral replication.[7,8] It is a two protomer homodimer in its active form and hRg as one active site in each homodimer chain,[7,9] The structure of the protomer of M^{pro} consists of three major domains. Domain I and II contains amino acid residues 8–101 and 102–184, respectively, and mostly consist of β-barrels. Domain III contains amino acid residues 201–303 and mainly consists of α-helices.[7] The catalytic site comprising of cysteine (Cys145) and histidine (His41) amino acid moiety, is located at the cleft of domain I and domain II.[7] The catalytic dyad His 41 and Cys145 present in the center of the site is responsible for forming strong interactions with inhibitors.[10] The complete active sites of MPro are the amino acid residues surrounding this catalytic dyad. The active site consists of two subsites S1 and S2, where the S1-subsite comprises the residues Phe 140, Gly 143, Cys 145, His 163, Glu 166, and His 172, and the residues Thr 25, His 41, Met 49, Met 165 and Gln 189 are for S2 -subsite, respectively.[8]

The M^{pro} is one of the most important proteins for the virus as it plays a key role in mediating viral replication and transcription, and several studies have been performed to target such proteins for the 'b' and 'c' groups.[7,11–20] As there are no human proteases with a homolog of M^{pro}, it is an ideal target for drug design as inhibitors are less likely to be toxic for humans.[2] The details regarding the M^{pro} architecture and brief mechanism are also discussed in earlier studies.[21–24] Natural compounds obtained from medicinal plants are used in the treatment of various diseases and also help in the development of different drugs. It has shown important roles in disease prevention and treatment in various diseases caused by bacteria or viruses.[25] It has also been largely accepted that numerous pharmacologically active drugs are derived from natural resources, including medicinal plants, which show lower side effects and are available in affordable prices too.[26,27]

[a] *R. Ghosh,[+] Dr. S. Chowdhuri*
School of Basic Sciences, Indian Institute of Technology Bhubaneswar, Bhubaneswar, India
E-mail: snehasis@iitbbs.ac.in
[b] *Dr. V. N. Badavath[+]*
Chitkara College of Pharmacy, Chitkara University, Punjab, 140401, India
[c] *Dr. A. Sen*
Department of Chemistry, Institute of Science, GITAM (Deemed to be University), Visakhapatnam 530045, Andhra Pradesh, India
E-mail: anikchem@gmail.com
[+] *Equal Authorship*

Supporting information for this article is available on the WWW under https://doi.org/10.1002/slct.202200055

In recent times, the role of natural products has also been very important for the development of drugs for the SARS-CoV 2.[28–31] A total of 51 different medicinal plants with very high properties of anti-viral, antioxidant, and anti-inflammatory were studied to check the activities on the SARS-CoV-2 by Upadhyay et. al.[32] Their results showed the M^{pro} inhibitory activity of a plethora of compounds from three different medicinal plants: green and black tea (Camellia sinensis) and Haritaki (*Terminalia Chebula*).[32]

In this article, we have focussed on the natural compounds extracted from Haritaki or *Terminalia Chebula*, an important traditional medicinal plant found in different parts of Asia, especially in India. Haritaki belongs to the genus Terminalia, which contains approximately 250 species that are distributed throughout the globe. It possesses effective anti-bacterial and anti-viral activity against various bacterial strains[33,34] and is well known for treating various other diseases like gastrointestinal and urinary tract diseases, fever, cough, diarrhoea, wound infections, skin diseases, urinary tract infection, and candidiasis. *Terminalia Chebula* consists of 131 natural compounds comprising of 31 different tannins, 6 different phenolic carboxylic compounds, 3 phenols, 19 different terpenoids, and triterpene saponins, 7 flavonoids, 2 sterols, and other hydrocarbon-based compounds.[35] We have performed a screening study on the 13 different terpenoids and triterpene saponins (TTS), 7 flavonoids, and 2 sterols of the *Terminalia Chebula* to identify the inhibitory effect on the M^{pro} of the SARS-CoV-2 through the molecular docking analysis using Autodock software.[36] It is found that seven compounds (five TTS and two sterols) of *Terminalia Chebula* show high interaction with the protease while compared to the inhibitor N3,[7] and other repurposed drugs. Further Molecular Dynamics simulations were performed on these 7 natural compounds to understand their affinity to be a potential inhibitor for SARS-CoV-2 main protease M^{pro}. Our results states, that the best interaction was observed for Daucosterol, Arjunetin, and Bellericoside. A previous study by Chowdhuri and co-workers, stated that N3 docked complex is stabilized by multiple hydrophobic and hydrogen bond interactions, and the pharmacokinetics analysis was performed for the complex, which revealed that N3 showed hepatotoxicity, which makes it carcinogenic for humans.[37] Our selected natural compounds were extracted from the traditional *Terminalia Chebula* medicinal plant and are well known for their various remedies from ages, and such compounds have been proven to be non-toxic for the human body.[38–40]

Result and Discussion:

Molecular docking studies

The 22 natural compounds from *Terminalia Chebula* (Figure 1) were downloaded from PubChem database server and were optimized with B3LYP/6-31G* basis set by using Gaussian 09 software.[41] Positive frequency confirmed that each optimized structure corresponds to the global minimum. These optimized structures were subjected to docking studies to identify the

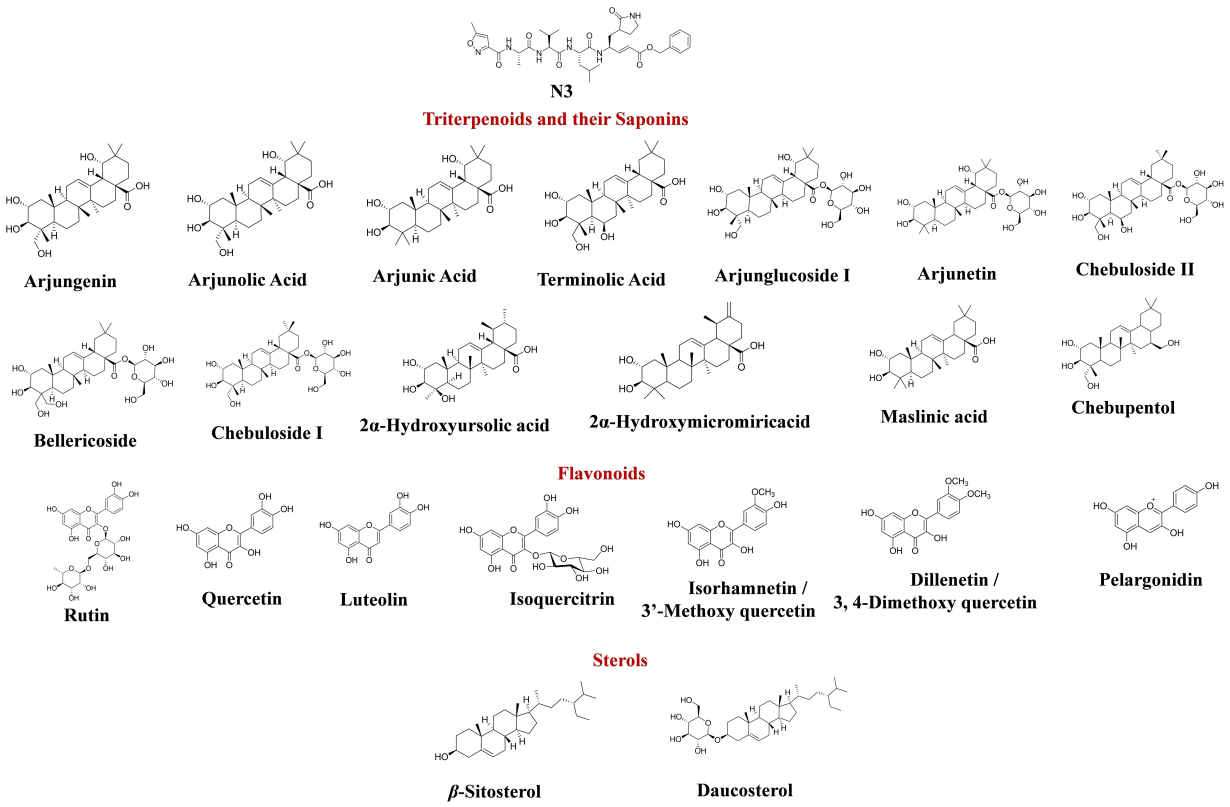

Figure 1. The schematic structures of isolated natural compounds from *Terminalia Chebula* along with native N3.

potent compounds, with the co-crystallized X-ray crystal structure of main protease (M^{pro}) of the SARS-CoV-2, (PDB id: 6LU7) with native N3. The SARS-CoV-2 main protease (M^{pro}) is very important for viral replication, and as it is dissimilar to human proteases, can act as a perfect target for inhibitor research.[7,8] Some known drugs under investigation for their potential use of treatment of COVID-19, were also docked with the protease for a comparison. The structures of the Remdesivir, Favipiravir, Chloroquine, Hydroxychloroquine, and Niclosamide were obtained from the PubChem Database.[42] The natural compounds with higher free binding energy (kcal/mol) and with good inhibitory constant (Ki) than N3-M^{pro} complex were considered for further Molecular dynamics simulations along with the native N3.

All the natural compounds studied in this work were well accommodated in the active site of M^{pro} with good free binding energy (kcal/mol) and inhibitory constant (Ki) when compared with N3-M^{pro} complex and the repurposed drugs, given in Table 1. The interaction of the best docked natural compounds with the M^{pro} is given in figure 2. The number of hydrogen bonding interactions for the seven different M^{pro}-natural compounds of *Terminalia Chebula* complexes, including the respective amino acids of M^{pro} involved in these hydrogen bond, are also shown in table 2. The details of the hydrogen bonding of the inhibitors for the specific atoms responsible for the different hydrogen bonding interactions associated with the amino acid residues of the M^{pro} are given in the supporting information. It showed that the all the hydrogen bondings for the inhibitors with the M^{pro} except for a single one for the Arjunetin proved to be hydrogen bond acceptor.

Molecular Dynamics (MD) Simulation Studies

We have carried out a 100 ns MD simulation using GROMOS96 53a6[43] force field with GROMACS software[44] to investigate the stability as well as the interactions of the 7 different M^{pro}-natural compounds of *Terminalia Chebula* chosen from the 22 natural compounds which were docked against the M^{pro} due to their high binding affinity (higher or comparable to the inhibitor N3), as mentioned in Table-1. Several structural

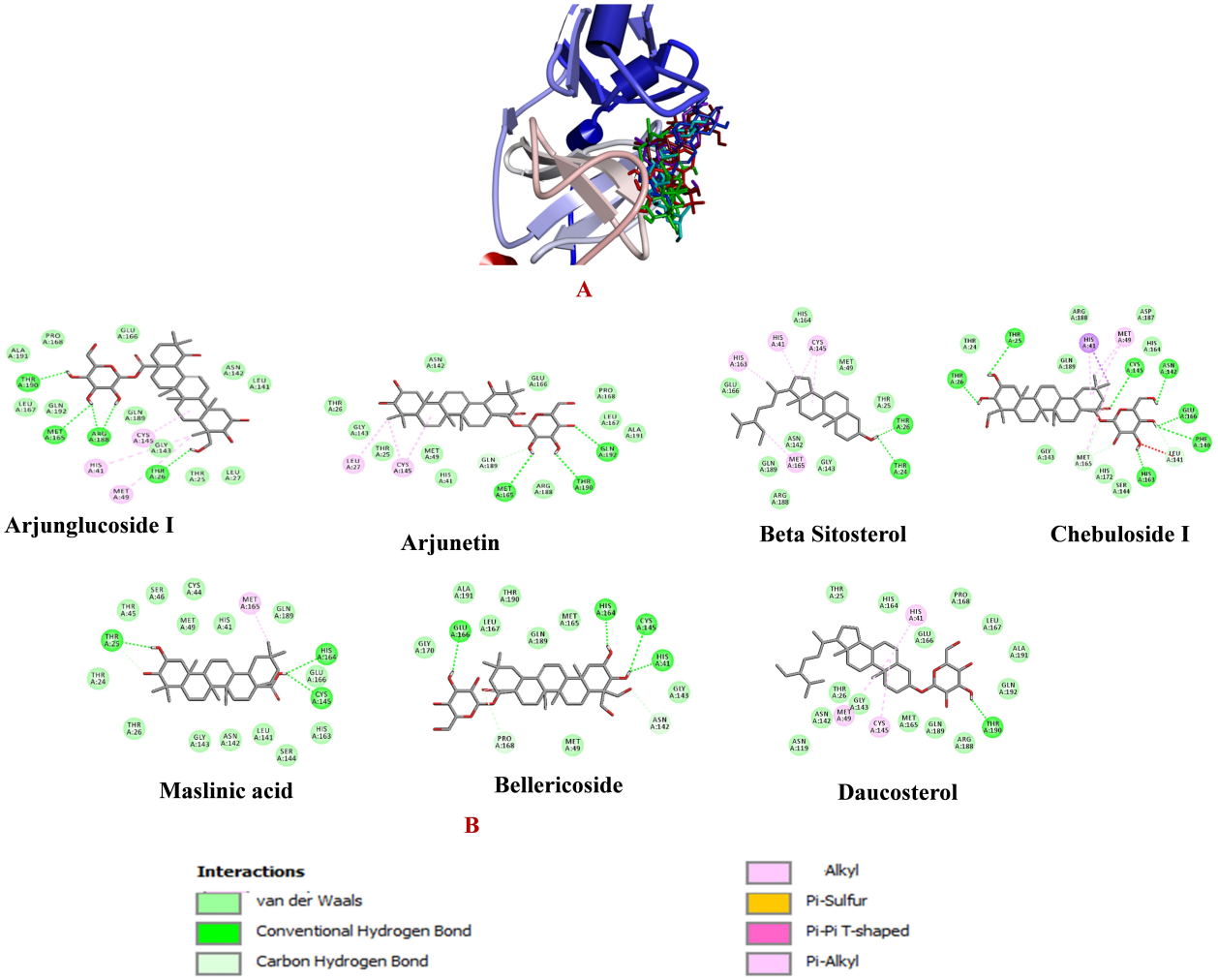

Figure 2. Orientation and binding interaction of potent seven compounds (**Arjunglucoside**I, **Arjunetin**, **Beta sitosterol**, **Chebulosidel**, **Maslinic acid**, **Bellericoside**, and **Daucosterol**) in the active site of M^{pro} (PDB: 6LU7). **A**. Superimposed of all the natural compounds in the active site of M^{Pro} ; **B**. The schematic diagrams of the interactions of the seven different ligands with the amino acids in the protein.

Table 1. The binding free energy and inhibitory constant of native N3, triterpenoids and saponins of *Terminalia Chebula* with M^{pro}.

S. No.	Compound	Inhibitory Constant (Ki)	Free binding energy (kcal/mol)	H-bonds (bond length in Å)
	N3	307.96 nM	-8.88	Cys145 (1.91 Å)
Repurposed drugs				
A	Ebselen	18.66 μM	−6.45	Glu166 (2.18 Å), Cys145
B	Favipiravir	378.76 μM	−4.67	–
C	Remdesivir	676.24 nM	−8.42	Glu166-NH (2.09 Å), Glu166-OH (1.92 Å)
D	Chloroquine	7.27 μM	−7.01	Glu166 (2.04 Å), Arg188 (1.78 Å)
E	Hydroxychloroquine	3.80 μM	−7.39	His164 (1.88 Å), Ser144 (2.21 Å), Leu141 (1.69 Å)
F	Niclosamide	2.66 μM	−7.60	Gly143(1.76 Å), Cys145(2.16 Å), Glu166 (2.068 Å)
	Triterpenoids and their Saponins			
1	Arjungenin	1.73 uM	−7.86	Gln189 (1.92), Thr26 (1.75)
2	Arjunolic acid	1.94 μM	−7.79	Gln192 (2.09), Thr190 (1.72)
3	Arjunic acid	644.15 nM	−8.45	Thr26 (1.96), Gln189 (1.81), Glu166 (2.12)
4	Terminolic acid	2.08 μM	−7.75	Cys145 (2.15), Glu166 (1.95), Gly143 (1.76), Met165 (1.68)
5	Arjunglucoside I	137.93 nM	−9.36	Thr26 (1.84), Asn142 (2.218), Gln189 (1.97)
6	Arjunetin	41.18 nM	−10.08	Thr26 (2.09), Thr190 (1.73), Gln192 (2.24), Arg188 (2.23), Asn142 (2.20)
7	Chebuloside II	1.44 μM	−7.97	Glu166 (2.06), Asn142 (1.98), Thr26 (2.11)
8	Bellericoside	588.17 nM	−8.50	Thr26 (2.20), Glu166 (1.86), Arg188 (1.81)
9	Chebuloside I	271.75 nM	−8.96	Asn142 (2.13), His163 (2.22), Glu166 (1.95)
10	2α-Hydroxyursolic acid	795.73 nM	−8.32	Gly143 (2.05), Thr24 (1.97), Thr25 (1.86)
11	2α-Hydroxymicromiricacid	705.45 nM	−8.39	Thr26 (1.77), Glu166 (2.20)
12	Maslinic acid	494.77 nM	−8.60	Gln192 (2.09), Thr190 (1.86)
13	Chebupentol	852.24 nM	−8.28	His164 (1.80) Thr26 (1.92), Gln189 (1.96),
Flavonoids isolated from *Terminalia Chebula*.				
1	Rutin	1.11 μM	−8.12	Asn142(2.10 Å), His164 (2.15 Å), Glu166 (1.93 Å), Thr190 (2.06 Å), Gln189 (1.79 Å)
2	Quercetin	1.50 μM	−7.94	His164 (2.23 Å), Glu166 (2.12 Å), Asp187 (2.02 Å)
3	Luteolin	2.44 μM	−7.66	Glu166 (2.03), Gln192 (2.16), Thr190 (2.06) and π-π interaction with Tyr54
4	Isoquercitrin	729.80 nM	−8.37	Glu166 (2.11), His164 (2.17), Asn142 (2.20), Asp187 (1.96)
5	Isorhamnetin/ 3′-Methoxy quercetin	708.59 nM	−8.39	Glu166 (2.18), Gln192 (2.03), Asp187 (2.14), and one π-π interaction with Tyr54
6	Dillenetin/ 3, 4-Dimethoxy quercetin	2.04 μM	−7.76	Thr190 (2.03) and one π-π interaction with His41
7	Pelargonidin	644.34 nM	−8.45	Glu166 (1.95), His164 (2.16), Gln192 (2.18), Asp187 (2.14), Thr190 (1.93)
Sterols isolated from *Terminalia Chebula*				
1	β-Sitosterol	241.37 nM	−9.03	Thr26 (1.75 Å)

Table 1. continued

S. No.	Compound	Inhibitory Constant (Ki)	Free binding energy (kcal/mol)	H-bonds (bond length in Å)
2	Daucosterol	402.16 nM	−8.73	Thr26 (2.07)

Table 2. Hydrogen bond interactions of identified potent natural compounds with the active site of SARS CoV-2 M^{pro}.

Natural compounds interaction with M^{pro}	Number of H-bonds	Amino acids of M^{pro} involved in H-bonding	Hydrogen bond distance (Å)
M^{pro}-ArjunglucosideI	5	Thr190	2.3
		Met165	2.4
		Arg188	2.3,2.35
		Thr26	2.35
M^{pro}-Arjunetin	3	Gln192	2.4
		Thr190	2.45
		Met165	2.45
M^{pro}-Beta sitosterol	2	Thr26	2.2
		Thr24	2.5
M^{pro}-ChebulosideI	7	Thr26	2.2
		Thr25	2.5
		Cys145	2.5
		Asn142	2.3
		Glu166	2.2
		Phe140	2.4
		His163	2.1
M^{pro}-Maslinic acid	3	Thr25	2.2
		His164	2.5
		Cys145	2.5
M^{pro}-Bellericoside	4	Glu166	2.3
		His164	2.3
		Cys145	2.4
		His41	2.2
M^{pro}-Daucosterol	1	Thr190	2.3

properties like the Root Mean Square Deviation (RMSD) for the complex stability, Root Mean Square Fluctuation (RMSF) for the conformational fluctuations, the radius of gyration (Rg) for understanding the structural compactness, and Solvent Accessible Surface Area (SASA) were investigated using the MD simulations. The RMSD of α-carbon atoms of the ligand-free conformation of M^{pro}, and complexes of M^{pro}-N3, and seven M^{pro}-natural compounds (Figure 3) were estimated in this study. It is found that for the ligand-free conformation of M^{pro} the RMSD maintained a constant value of ~0.21–0.22 nm in the initial 2 ns to 17 ns, which gradually increased to a value of ~0.35 nm till 25 ns, and after that, a decrease of the value to ~0.31 nm was observed till 65 ns and remained stationary towards the end of the molecular simulation. The average RMSD values of the seven different M^{pro}-natural complexes along with M^{pro}-N3, and ligand-free conformation of M^{pro} are given in Table 3. It has been observed that except M^{pro}-Arjunglucoside I, all other M^{pro}-natural compound complexes have comparable results with M^{pro}-N3. The RMSD pattern for M^{pro}, M^{pro}-N3, and all other M^{pro}-natural compound complexes are shown in figure 3. The RMSD shows slight fluctuations until 20 ns, but no such fluctuations were observed for the later part (Figure 3). The total number of intermolecular hydrogen bonds formed during the entire MD simulation was calculated in this

Table 3. Average values of the RMSD, RMSF, Rg, SASA, and the total number of intermolecular hydrogen bonds formed during MD simulation for different systems.

System	RMSD (nm)	RMSF (nm)	Rg (nm)	SASA (nm^2)
M^{pro} (ligand free)	0.3090 ± 0.0528	0.1937 ± 0.0218	2.195 ± 0.0219	151.4483 ± 0.0234
M^{pro}-N3	0.2593 ± 0.0425	0.1575 ± 0.0239	2.1911 ± 0.0215	155.3979 ± 0.0321
M^{pro}-Arjunglucoside I	0.3680 ± 0.0472	0.1532 ± 0.0285	2.1962 ± 0.0218	152.1000 ± 0.0218
M^{pro}-Arjunetin	0.2541 ± 0.0398	0.1552 ± 0.0283	2.1968 ± 0.0217	153.7193 ± 0.0221
M^{pro}-Beta sitosterol	0.2463 ± 0.0426	0.1451 ± 0.0412	2.1979 ± 0.0321	152.7391 ± 0.0328
M^{pro}-Chebuloside_I	0.2741 ± 0.0314	0.1565 ± 0.0232	2.1966 ± 0.0298	152.6999 ± 0.0224
M^{pro}-Maslinic acid	0.2379 ± 0.0384	0.1236 ± 0.0216	2.1979 ± 0.0291	152.2632 ± 0.0219
M^{pro}-Bellericoside	0.2401 ± 0.0411	0.1315 ± 0.0223	2.1984 ± 0.0215	152.8833 ± 0.0232
M^{pro}-Daucosterol	0.2539 ± 0.0374	0.1558 ± 0.0243	2.1978 ± 0.0214	152.9458 ± 0.0227

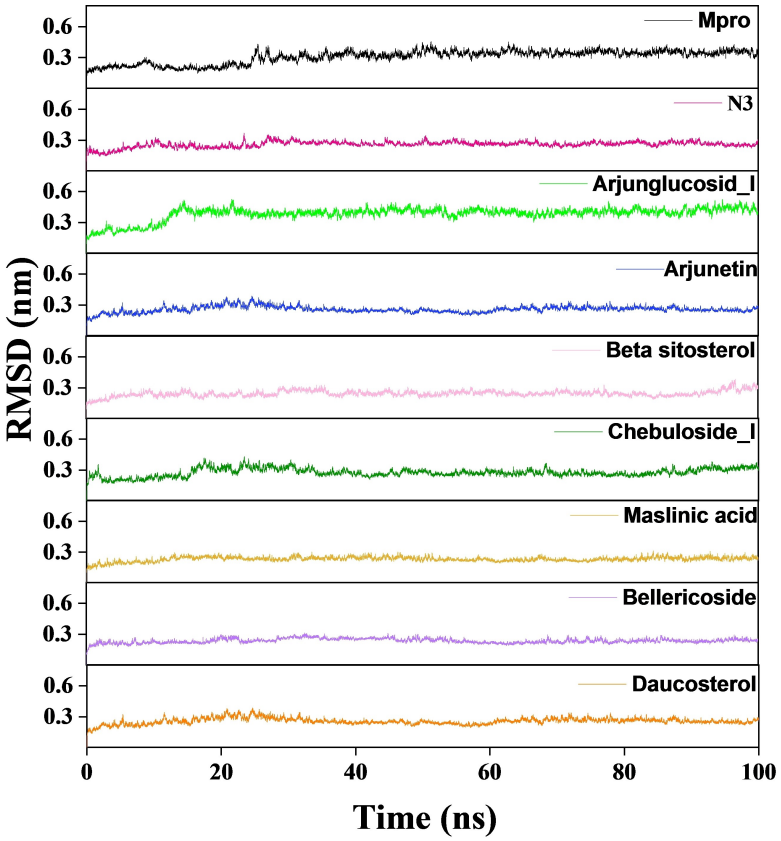

Figure 3. The RMSD plots of identified potent compounds and native drug N3, along with ligand-free conformation of M^{pro} of SARS-CoV-2.

study as it can help to estimate the conformational stability of the M^{pro}-natural compounds of *Terminalia Chebula* complexes (Figure 4). It was observed that on average, 547 intermolecular hydrogen bonds were present in the ligand-free conformation of M^{pro}, whereas all other M^{pro}-complexes show a higher number of hydrogen bonds which suggested the formation of a stable complex. Furthermore, hydrogen bonds for the N3 as well as the natural compounds with the M^{pro} have been plotted with the 100-ns MD simulation as given in figure 5. The plots indicated that for the N3, there were approximately 6 hydrogen bonds till 60 ns but the lasr 40 ns only 3/4 hydrogen bonds were observed. For the natural compounds, 5–8 hydrogen bonds were observed for the entire 100 ns time frame except for the Arjunetin, where it showed an increase in the average

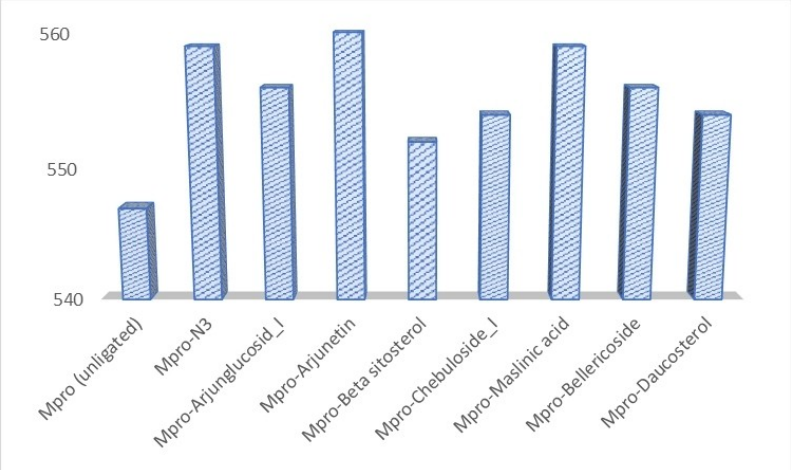

Figure 4. Calculated total number of hydrogen bonds for the ligand-free conformation of M^{pro} and complexed M^{pro}.

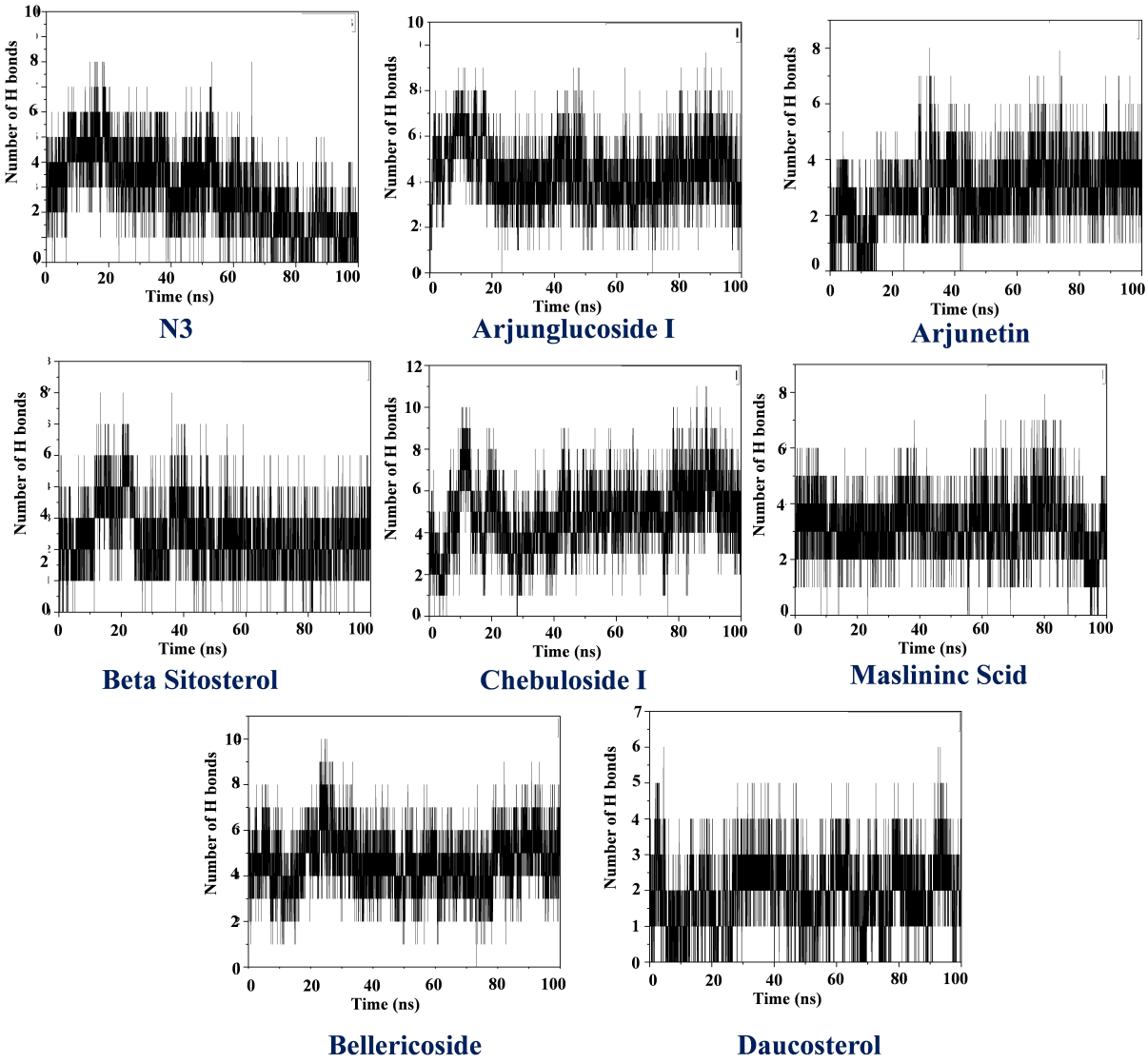

Figure 5. The number of Hydrogen bonds of the N3 and the natural compounds with M^{pro} plotted along with the 100-ns MD simulation.

number of hydrogen bonds from 4 to 5/6 after the first 30 ns run. Arjunglucoside I, Bellericoside, and Chebuloside I have shown an average number of Hydrogen bonds 6–8 during the 100 ns run whereas, on average 4–6 number of hydrogen bonds were found in β-silostrerol, and Maslinic acid, and 3–5 number of hydrogen bonds are observed in Daucosterol.

The Root Mean Square Fluctuations (RMSF) of the α-carbon atoms of ligand-free conformation of M^{pro}, M^{pro}-N3, and M^{pro}-natural compounds are given in Figure 6. A slight fluctuation in the range of 0.3 nm was observed for the domain, I and II. It was also observed from the RMSF profiles that all the systems experience more conformational fluctuations in domain III. For the residues 45–60 in domain I, M^{pro}-natural compound complexes and M^{pro}-N3 show lower fluctuations than the ligand-free conformation of M^{pro} . The average RMSF value for ligand-free conformation of M^{pro} was ~0.194 nm, and that of M^{pro}-N3 complex was ~0.157 nm (Table 3). The M^{pro}-natural compound complexes observed lower fluctuations, especially in domains II, and III compared to the ligand-free conformation of M^{pro} and M^{pro}-N3 complexes, as shown in figure 6.

The stability of the protein can also be understood by the calculation of the radius of gyration (R_g). The calculated Rg values for the ligand-free conformation of M^{pro}, and the M^{pro}-natural compound complexes are given in figure 7, and table 3, to assess the compactness of the complexes. The average Rg value for M^{pro} and M^{pro}-N3 were observed to be identical with a slight increase in case of M^{pro}-Natural compounds (Table 3). The M^{pro}-natural compound complexes are thus slightly less compact than that of ligand-free conformation of M^{pro} and M^{pro}-N3 complex. The solvent-accessible surface area (SASA) values were also calculated to assess the extent of expansion of protein volume in each system (Figure 7 and Table 3). A little expansion of the M^{pro} occurs upon interaction with the natural compounds, as observed from the SASA values given in figure 7 and Table 3.

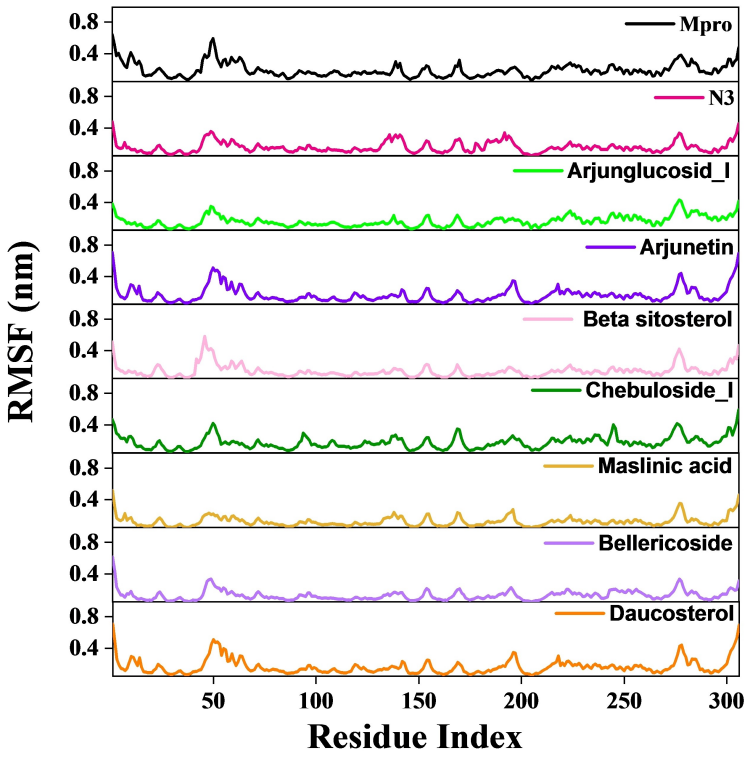

Figure 6. RMSF profiles of identified potent natural compounds along with ligand-free conformation of M^{pro} and M^{pro}-N3 complexes against the amino acid residues of M^{pro}.

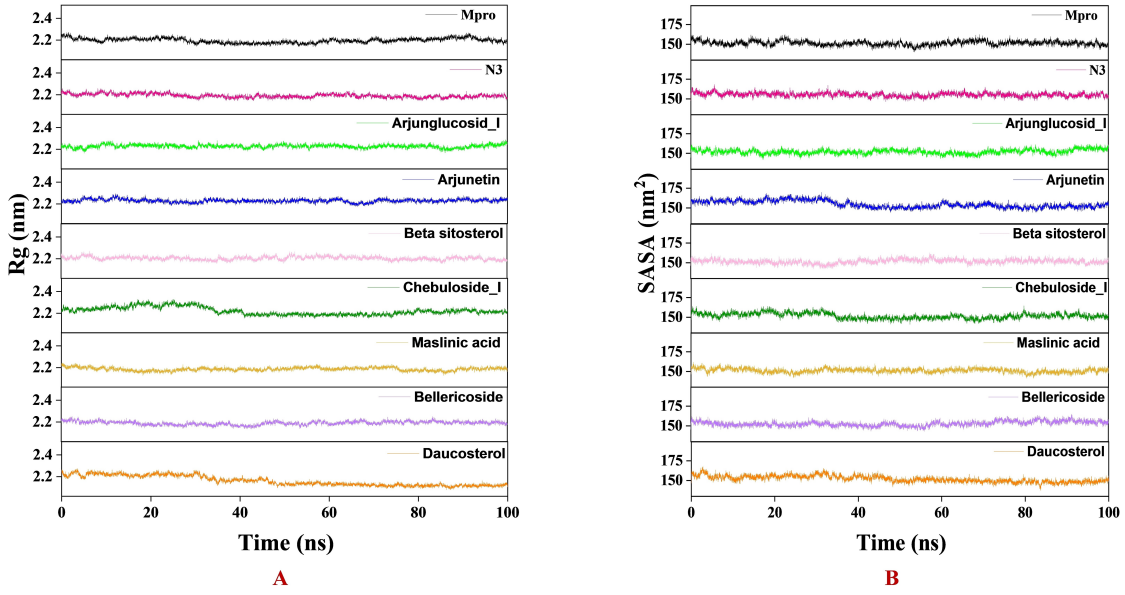

Figure 7. (A) The Rg plots of identified potent natural compounds along with ligand-free conformation of M^{pro} and M^{pro}-N3 complexes.(B) The SASA plots of identified potent natural compounds along with ligand-free conformation of M^{pro} and M^{pro}-N3 complex.

MM-GBSA

The binding free energy of the complexes of the M^{pro}- natural compounds along with the M^{pro}-N3 was calculated using the MM-GBSA method. The MM-GBSA was calculated on the coordinates of each complex after each 5 ns simulation run, and 20 binding energies (100 ns) are given in figure 8. The average binding energy values for each complex are given in table 4. It is observed that M^{pro}-Daucosterol complex shows higher binding free energy (−57.81 kcal/mol) when compared

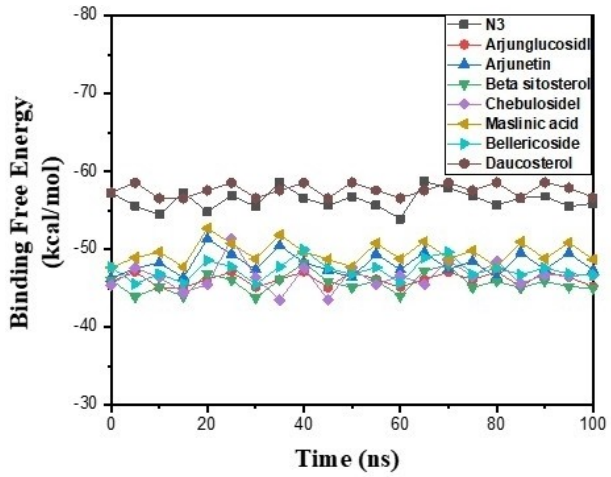

Figure 8. MM-GBSA binding free energy profiles of two Mpro-standard complexes and seven Mpro-*Terminalia Chebula* complexes. The binding free energy values of Mpro-N3 complex, M^{pro}-Arjunglucoside I, M^{pro}-Arjunetin, M^{pro}-Beta sitosterol, M^{pro}-Chebuloside I, M^{pro}-Maslinic acid, M^{pro}-Bellericoside, and M^{pro}-Daucosterol complex were represented throughout the entire 100 ns simulation trajectory.

Table 4. MM-GBSA values of the natural compounds along with M^{pro}-N3 and complexes. The standard deviation was reported as an error (±) associated with free energy differences. Free energies are given in kcal/mol.

System	Binding Free Energy (kcal/mol)
M^{pro}-N3	−56.29 ± 0.19
M^{pro}-Arjunglucosidel	−46.12 ± 0.55
M^{pro}-Arjunetin	−48.18 ± 0.24
M^{pro}-Beta sitosterol	−45.59 ± 0.14
M^{pro}-Chebulosidel	−46.48 ± 0.32
M^{pro}-Maslinic acid	−49.54 ± 0.35
M^{pro}-Bellericoside	−47.32 ± 0.19
M^{pro}-Daucosterol	−57.58 ± 0.32

Table 5. The ΔG_{bind} for the natural compounds complexes along with M^{pro}-N3 were calculated using the MM/PBSA method. The standard deviation was reported as an error (±) associated with free energy differences. Free energies are given in kcal/mol.

Name	Total Binding Energy (ΔG_{bind}) in kcal/mol
M^{pro}-N3	−66.65 ± 7.31
M^{pro}-Arjunglucoside I	−41.51 ± 6.48
M^{pro}-Arjunetin	−47.96 ± 6.73
M^{pro}-β-sitosterol	−41.05 ± 6.04
M^{pro}-Chebuloside I	−42.01 ± 5.63
M^{pro}-Maslinic acid	−49.46 ± 5.76
M^{pro}-Bellericoside	−47.32 ± 4.94
M^{pro}-Daucosterol	−67.32 ± 6.73

with M^{pro}-N3 complex (−56.02 kcal/mol). For all other M^{pro}-natural compound complexes, the average binding free energy was a slightly lower than that of M^{pro}-N3 complex (Table 4 and Figure 8). The higher MM-GBSA value (ΔG_{bind}) for the M^{pro}-natural compound complexes of *Terminalia Chebula* is mostly contributed by the SASA and hydrophobic interactions. Our calculations suggested that these seven natural compounds obtained from the *Terminalia Chebula* can be allocated for further studies for the development of natural compound-based drugs for the treatment of SARS-CoV2.

MM/PBSA

The MM-GBSA relies on certain snapshots during the MD simulation to calculate the absolute ligand binding interaction of the natural compounds with the M^{pro}. Herein, we have also calculated the binding free energies of the interactions using the MM/PBSA method. It has been observed that all the seven natural compounds show strong binding affinity to the M^{pro} (Table 5), similar to the MM-GBSA calculations. Daucosterol shows the highest binding free energy (−67.32 kcal/mol). A

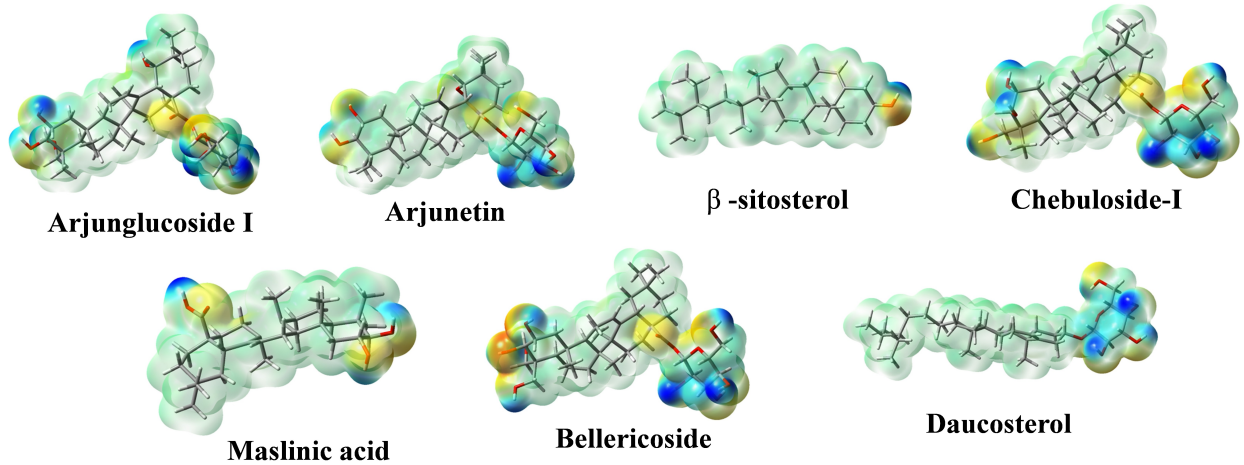

Figure 9. The calculated Molecular Electrostatic Potential surface for the identified potent 7 natural compounds. The red parts indicate hydrogen bonding donor, the blue parts indicate hydrogen bond acceptors sites, and the π-interaction sites are given as green color, and yellow patches showed higher electron density than the green sites. The surfaces shown correspond to an isosurface value of 0.03 electrons/a.u.^{3}.

Table 6. The overall percentage of secondary structure elements of identified potent natural compounds along with ligand-free conformation of M^{pro} and M^{pro}-N3 complexes.

Systems	Structure (α-helix + β-sheet + β-bridge + Turn) (%)	Coil (%)	β-sheet (%)	β-bridge (%)	Bend (%)	Turn (%)	α-helix (%)	5-helix (%)	3-helix (%)
M^{pro} (ligand free)	58.0	28.0	25.0	2.0	13.0	10.0	21.0	0.00	1.0
M^{pro}-N3	58.0	28.0	26.0	1.0	12.0	9.0	22.0	0.00	2.0
M^{pro}-Arjunglucosi-del	59.0	28.0	26.0	2.0	12.0	9.0	22.0	0.00	1.0
M^{pro}-Arjunetin	59.0	29.0	26.0	2.0	11.0	9.0	22.0	0.00	1.0
M^{pro}- β- sitosterol	59.0	27.0	26.0	2.0	11.0	9.0	22.0	0.00	2.0
M^{pro}-Chebulosidel	61.0	27.0	26.0	1.0	10.0	10.0	22.0	0.00	2.0
M^{pro}-Maslinic acid	59.0	27.0	27.0	2.0	13.0	9.0	21.0	0.00	1.0
M^{pro}-Bellericoside	60.0	28.0	27.0	1.0	11.0	10.0	22.0	0.00	1.0
M^{pro}-Daucosterol	61.0	27.0	27.0	1.0	0.10	11.0	22.0	0.00	2.0

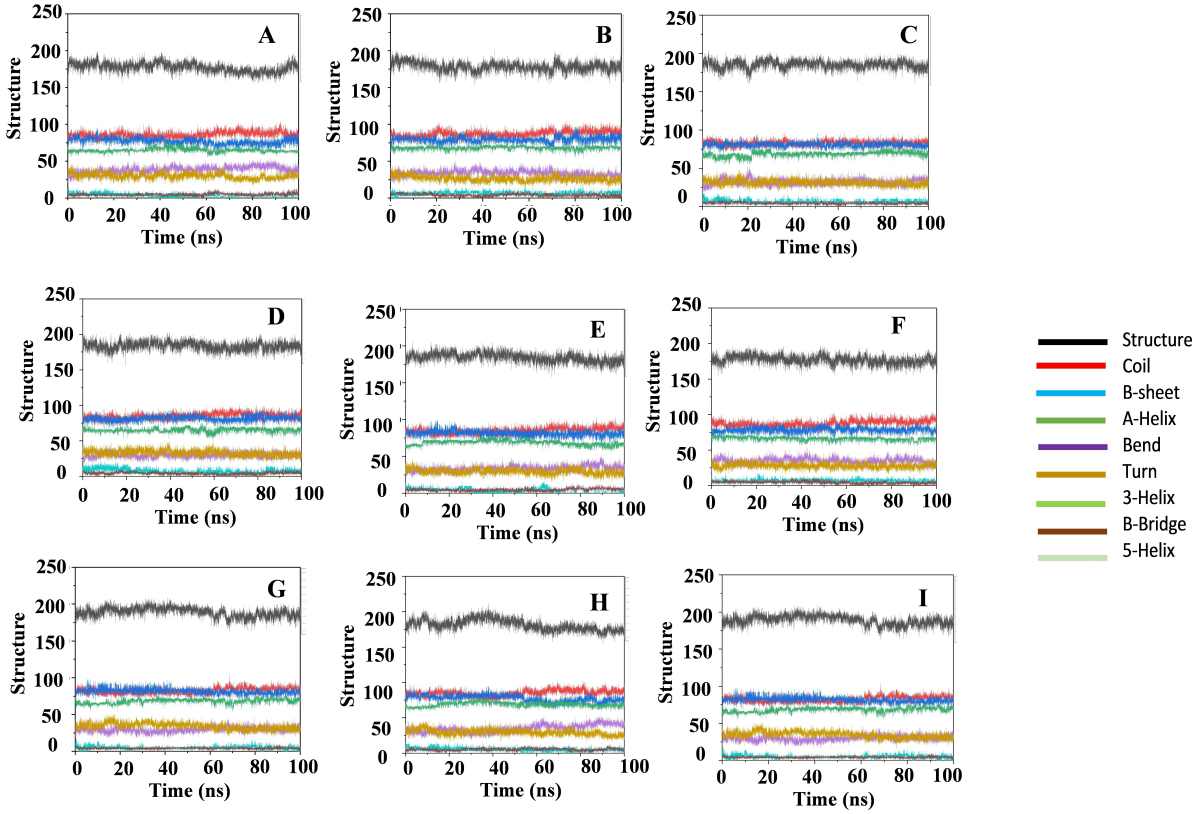

Figure 10. Analysis of M^{pro} secondary structure in different systems during 100 ns MD simulation. The secondary structure content of M^{pro} in (A) ligand free conformation of M^{pro}, (B) M^{pro}-N3 system, (C) M^{pro}-Arjunglucoside I, (D) M^{pro}-Arjunetin, (E) M^{pro}- β-Sitosterol and (F) M^{pro}-Chebuloside I, (G), M^{pro}-Maslinic acid, (H) M^{pro}-Bellericoside, and (I) M^{pro}-Daucosterol.

very comparable binding energy was observed for the Arjunetin, Maslinic acid, and Bellericoside with the N3 inhibitor. While, Arjunglucoside I, β-sitosterol, and Chebuloside I show lower binding affinity and thus may not act as effective inhibitors. The reason behind such binding affinities is displayed in table 5 and the supplementary information (Table S2). It has been observed that the ligand-binding affinity is mostly dominated by two parameters, (i) Van Der Waals interaction (ΔG_{VDW}) and the (ii) polar solvation energy (ΔG_{PB}), whereas the first one is favorable for the ligand interaction. From the calculations we find that Van Der Waals energy were observed in the sequence of Daucosterol > Maslinic acid > N3 > Arjunetin > Bellericoside > Chebulosidel > Arjunglucosidel > β-Sitosterol, whereas the unfavourable polar solvation energy (ΔG_{PB}) is observed as Chebuloside I > Maslinic acid > Daucosterol > Arjunetin > β-Sitosterol > Arjunglucosidel > N3 > Bellericoside. These two dominated parameters, along with the electrostatic interactions (ΔG_{elec}), and the non-polar solvation energy (ΔG_{SA}), revealed the highest binding energy for the Daucosterol, and comparable binding energies for the N3, Arjunetin, Bellericoside, and Maslinic acid (in the order of Daucosterol > N3 > Maslinic acid > Arjunetin > Bellericoside). The components

of the MM/PBSA, ΔG_{VDW} (Van Der Waals), ΔG_{elec} (Electrostatic), ΔG_{PB} (polar solvation, Poisson-Boltzmann term), and ΔG_{SA} (non-polar solvation) for the natural compounds along with M^{pro}-N3 and complexes are given in the supporting information.

Molecular Electrostatic Potential Analysis

Molecular electrostatic potential calculations (MESP) were performed for all the seven natural compounds using Density Functional Theory, as electronic distribution is a very important factor to predict the binding site to a protein. In the MESP surface in figure 9, the red-colored regions denote regions for the H-bond acceptor and are the places where the high electronegative atoms are available. The electron-poor atoms are designated in blue, which acts as H-bond donors. The neutral areas are shown as green regions, where π- and other types of π-staking interactions are more important (Figure 9). The MESP analysis clearly proved the number of hydrogen bonding possible for each of the natural compounds, which was observed in the MD results (Table 2 and Figure 9). The enhancement of the electron density is favorable for the π-staking interaction. The reason behind the binding energy of the Daucosterol, Arjunetin, Maslinic acid, and Bellericoside natural compounds can be high electron density surface as observed in our calculation with a large green region indicating that it is suitable for multiple π-staking interactions along with the hydrophilic and hydrophobic properties. The red, yellow, and blue spaces on the large green surface of compounds balance the hydrophilic and hydrophobic properties of the ligand, which are essential for good binding to the protein for all the different compounds.

Secondary structural analysis during MD simulation

The secondary structural analysis for all the complexes was calculated to check the structural changes in the entire trajectory (Figure 10, Table 6). No major fluctuations were observed in the 100 ns run for the secondary structural elements of M^{pro} (α-helix and β-sheet) as well as for M^{pro}-N3, M^{pro}-natural compound complexes. Thus, due to minor or no changes in the secondary structure proved very minimal protein fluctuation and high stability for the complexes, which leads to good interaction.

The content of α-helix, β-sheet, β-bridge and turn in M^{pro} was 21%, 25%, 2% and 10%, respectively (Table 6). Altogether, these findings suggested that overall structural conformation, including secondary conformation of M^{pro} is unaltered even when complex formation with M^{pro}-natural compounds (Arjunglucoside I, Arjunetin, β-sitosterol, Chebuloside I, Maslinic acid, Bellericoside, and Daucosterol).

Conclusion

In this study, we have investigated the inhibition potency of the terpenoids and triterpene saponins, flavonoids, and sterols present in the *Terminalia Chebula* against M^{pro} SARS-CoV-2, using an in silico approach. Among the twenty–two natural compounds studied here, only seven showed the highest interaction energies in the molecular docking calculations compared to the N3 and were further investigated with Molecular Dynamics calculations. All seven natural compounds showed interactions with the key catalytic residues (His41 and Cys145) of M^{pro}. The RMSD and RMSF profiles of the natural compound complexes with the M^{pro} show high stability and less conformational fluctuations. Furthermore, the Rg and SASA analysis revealed less compactness of the natural compound complexes than the ligand-free conformation of M^{pro} and M^{pro}-N3 complex. During 100 ns MD simulation, the natural compounds show a stable number of hydrogen bonds with the protein, whereas for N3 there was a drastic change after 60 ns run, though there is no change in the total number of intermolecular hydrogen bonds. The MM-GBSA analysis further reconfirmed that M^{pro}-natural compound complexes of *Terminalia Chebula* are quite stable along with the M^{pro}-N3 complex and the Daucosterol show the best interaction, and similar interaction energies were calculated with the MM/PBSA calculations. Both methods also revealed comparable interaction energies of Arjunetin, Maslinic acid, and Bellericoside with the N3. The MESP analysis showed an agreement for the number of hydrogen bonding possible for each of the natural compounds as observed in the MD results. It also showed that the compounds Daucosterol, Arjunetin, Maslinic acid and Bellericoside have high electron density surface which makes them suitable for multiple π-staking interaction along with the hydrophilic and hydrophobic properties. Our overall findings revealed that out of the twenty-two natural compounds of *Terminalia Chebula* studied here, the most potent inhibitor is Daucosterol. The other potential inhibitors revealed from our calculations are Arjunetin, Maslinic acid, and Bellericoside, which have the potency to inhibit the proteolytic activity of M^{pro} and may be used for the treatment of COVID-19. This study is unique as such a detailed analysis on the potency of the extracts of the *Terminalia Chebula* on the SARS-CoV-2 is performed for the first time. Further in-vitro and in-vivo experimental studies are necessary before repurposing of such natural compounds to develop a safe and effective non-toxic drug against COVID-19.

Supplementary Information

A detailed computational method section is given in the supplementary information along with the MM/PBSA and MM/GBSA components and binding energies of several other Natural compounds.

Acknowledgements

RG acknowledges IIT Bhubaneswar for providing fellowship. The authors thank IIT Bhubaneswar, India for providing facility for computational resources. AS thanks GITAM (Deemed to be) University and the UGC Start UP Grant for providing support in this work.

Conflict of Interest

The authors declare that they have no conflicts of interest with the contents of this article.

Data Availability Statement

Research data are not shared.

Keywords: Molecular Docking · Molecular Dynamics · SARS CoV-2 Main protease · Natural Products · *Terminalia Chebula*

[1] R. Čivljak, A. Markotić, I. Kuzman, *Croat. Med. J.* **2020**, *61*, 1–4.
[2] F. Wu, S. Zhao, B. Yu, Y. M. Chen, W. Wang, Z. G. Song, Y. Hu, Z. W. Tao, J. H. Tian, Y. Y. Pei, M. L. Yuan, Y. L. Zhang, F. H. Dai, Y. Liu, Q. M. Wang, J. J. Zheng, L. Xu, E. C. Holmes, Y. Z. Zhang, *Nature* **2020**, *579*, 265–269.
[3] A. E. Gorbalenya, S. C. Baker, R. S. Baric, R. J. de Groot, C. Drosten, A. A. Gulyaeva, B. L. Haagmans, C. Lauber, A. M. Leontovich, B. W. Neuman, D. Penzar, S. Perlman, L. L. M. Poon, D. v. Samborskiy, I. A. Sidorov, I. Sola, J. Ziebuhr, *Nat. Microbiol.* **2020**, *5*, 536–544.
[4] P. Zhou, X. lou Yang, X. G. Wang, B. Hu, L. Zhang, W. Zhang, H. R. Si, Y. Zhu, B. Li, C. L. Huang, H. D. Chen, J. Chen, Y. Luo, H. Guo, R. di Jiang, M. Q. Liu, Y. Chen, X. R. Shen, X. Wang, X. S. Zheng, K. Zhao, Q. J. Chen, F. Deng, L. L. Liu, B. Yan, F. X. Zhan, Y. Y. Wang, G. F. Xiao, Z. L. Shi, *Nature* **2020**, *579*, 270–273.
[5] G. Bolcato, M. Bissaro, M. Pavan, M. Sturlese, S. Moro, *Sci. Rep.* **2020**, *10*, 20927.
[6] Z. M. Chen, J. F. Fu, Q. Shu, Y. H. Chen, C. Z. Hua, F. B. Li, R. Lin, L. F. Tang, T. L. Wang, W. Wang, Y. S. Wang, W. Z. Xu, Z. H. Yang, S. Ye, T. M. Yuan, C. M. Zhang, Y. Y. Zhang, *World J. Pediatr.* **2020**, *16*, 240–246.
[7] 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, X. Yang, F. Bai, H. Liu, X. Liu, L. W. Guddat, W. Xu, G. Xiao, C. Qin, Z. Shi, H. Jiang, Z. Rao, H. Yang, *Nature* **2020**, *582*, 289–293.
[8] L. Zhang, D. Lin, X. Sun, U. Curth, C. Drosten, L. Sauerhering, S. Becker, K. Rox, R. Hilgenfeld, *Science* **2020**, *368*, 409–412.
[9] T. Muramatsu, C. Takemoto, Y. T. Kim, H. Wang, W. Nishii, T. Terada, M. Shirouzu, S. Yokoyama, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 12997–13002.
[10] Q. W. Wang, Y. Su, J. T. Sheng, L. M. Gu, Y. Zhao, X. X. Chen, C. Chen, W. Z. Li, K. S. Li, J. P. Dai, *PLoS One* **2018**, *13*, 1–19.
[11] V. N. Badavath, A. Kumar, P. K. Samanta, S. Maji, A. Das, G. Blum, A. Jha, A. Sen, *J. Biomol. Struct. Dyn.* **2020**, *DOI:10.1080/07391102.2020.1845800*.
[12] X. Liu, X. J. Wang, *J. Genet. Genomics* **2020**, *47*, 119–121.
[13] V. N. Holanda, E. M. de A. Lima, W. V. da Silva, R. T. Maia, R. de L. Medeiros, A. Ghosh, V. L. de M. Lima, R. C. B. Q. de Figueiredo, *J. Biomol. Struct. Dyn.* **2021**, *18*, 1–19.
[14] E. Singh, R. J. Khan, R. K. Jha, G. M. Amera, M. Jain, R. P. Singh, J. Muthukumaran, A. K. Singh, *J. Genet. Eng. Biotechnol.* **2020**, *18*, 69.
[15] N. Sepay, A. Sekar, U. C. Halder, A. Alarifi, M. Afzal, *J. Mol. Struct.* **2021**, *1228*, 129433.
[16] R. Ghosh, A. Chakraborty, A. Biswas, S. Chowdhuri, *J. Mol. Struct.* **2021**, *1229*, 129489.
[17] A. Gupta, H. X. Zhou, *ACS Comb. Sci.* **2020**, *22*, 826–832.
[18] D. A. Abdelrheem, S. A. Ahmed, H. R. Abd El-Mageed, H. S. Mohamed, A. A. Rahman, K. N. M. Elsayed, S. A. Ahmed, *J. Environ. Sci. Health Part A* **2020**, *55*, 1373–1386.
[19] H. H. Fan, L. Q. Wang, W. L. Liu, X. P. An, Z. D. Liu, X. Q. He, L. H. Song, Y. G. Tong, *Chin. Med. J.* **2020**, *133*, 1051–1056.
[20] Y. W. Chen, C. P. B. Yiu, K. Y. Wong, *F1000Research* **2020**, *9*, 129.
[21] V. K. Bhardwaj, R. Singh, P. Das, R. Purohit, *Comput. Biol. Med.* **2021**, *128*, 104117.
[22] V. K. Bhardwaj, R. Singh, J. Sharma, V. Rajendran, R. Purohit, S. Kumar, *J. Biomol. Struct. Dyn.* **2020**, *39*, 3449–3458.
[23] R. Ghosh, A. Chakraborty, A. Biswas, S. Chowdhuri, *J. Biomol. Struct. Dyn.* **2021**, *39*, 6747–6760.
[24] R. Ghosh, A. Chakraborty, A. Biswas, S. Chowdhuri, *J. Biomol. Struct. Dyn.* **2022**, *40*, 2647–2662.
[25] D. J. Newman, G. M. Cragg, *J. Nat. Prod.* **2020**, *83*, 770–803.
[26] L. J. V. Piddock, *Lancet Infect. Dis.* **2012**, *12*, 249–253.
[27] C. T. Walsh, T. A. Wencewicz, *J. Antibiot.* **2014**, *67*, 7–22.
[28] T. Capell, R. M. Twyman, V. Armario-Najera, J. K. C. Ma, S. Schillberg, P. Christou, *Trends Plant Sci.* **2020**, *25*, 635.
[29] B. T. P. Thuy, T. T. A. My, N. T. T. Hai, L. T. Hieu, T. T. Hoa, H. Thi Phuong Loan, N. T. Triet, T. T. van Anh, P. T. Quy, P. van Tat, N. van Hue, D. T. Quang, N. T. Trung, V. T. Tung, L. K. Huynh, N. T. A. Nhung, *ACS Omega* **2020**, *5*, 8312–8320.
[30] T. Erdogan, *J. Mol. Struct.* **2021**, *1242*, 130733.
[31] S. kumar Arumugam, S. S. Dash, K. Mitra, M. Doble, S. N. Gummadi, *ChemRxiv.* **2020**, *DOI:10.26434/CHEMRXIV.12600587*.
[32] S. Upadhyay, P. K. Tripathi, M. Singh, S. Raghavendhar, M. Bhardwaj, A. K. Patel, *Phytother. Res.* **2020**, *34*, 3411–3419.
[33] D. Lee, K. H. Boo, J. K. Woo, F. Duan, K. H. Lee, T. K. Kwon, H. Y. Lee, K. Z. Riu, D. S. Lee, *J. Appl. Biol. Chem.* **2011**, *54*, 295–298.
[34] R. Rathinamoorthy, G. Thilagavathi, *Int. J. PharmTech Res.* **2014**, *6*, 97–114.
[35] R. P. Nishanth, T. Prasad, R. G. Jyotsna, P. K. Reddy, P. Reddanna, *J. Herbs, Spices Med. Plants* **2014**, *20*, 402–420.
[36] G. M. Morris, H. Ruth, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, A. J. Olson, *J. Comput. Chem.* **2009**, *30*, 2785.
[37] R. Ghosh, A. Chakraborty, A. Biswas, S. Chowdhuri, *J. Biomol. Struct. Dyn.* **2021**, *39*, 4362–4374.
[38] A. Bag, S. K. Bhattacharyya, R. R. Chattopadhyay, *Asian Pac. J. Trop. Med.* **2013**, *3*, 244.
[39] A. Varghese, J. Savai, N. Pandita, R. Gaud, *Toxicol. Rep.* **2015**, *2*, 806.
[40] R. Paniagua-Pérez, E. Madrigal-Bujaidar, S. Reyes-Cadena, D. Molina-Jasso, J. Pérez Gallaga, A. Silva-Miranda, O. Velazco, N. Hernández, G. Chamorro, *J. Biomed. Biotechnol.* **2005**, *2005*, 242–247.
[41] M. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.a. Petersson, *Gaussian 09, Revision D. 01*. Gaussian, Inc., Wallingford CT **2009**.
[42] S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, Q. Li, B. A. Shoemaker, P. A. Thiessen, B. Yu, L. Zaslavsky, J. Zhang, E. E. Bolton, *Nucleic Acids Res.* **2019**, *47*, D1102-D1109.
[43] C. Oostenbrink, A. Villa, A. E. Mark, W. F. van Gunsteren, *J. Comput. Chem.* **2004**, *25*, 1656–1676.
[44] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindah, *SoftwareX* **2015**, *1–2*, 19–25.

Submitted: January 5, 2022
Accepted: March 28, 2022