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

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Identification of Food Compounds as inhibitors of SARS-CoV-2 main protease using molecular docking and molecular dynamics simulations

Vijay H. Masand a,*, Md Fulbabu Sk b, Parimal Kar b, Vesna Rastija c, Magdi E.A. Zaki d

Abstract

SARS-CoV-2 has rapidly emerged as a global pandemic with high infection rate. At present, there is no drug available for this deadly disease. Recently, Mpro (Main Protease) enzyme has been identified as essential proteins for the survival of this virus. In the present work, Lipinski’s rules and molecular docking have been performed to identify plausible inhibitors of Mpro using food compounds. For virtual screening, a database of food compounds was downloaded and then filtered using Lipinski’s rule of five. Then, molecular docking was accomplished to identify hits using Mpro protein as the target enzyme. This led to identification of a Spermidine derivative as a hit. In the next step, Spermidine derivatives were collected from PubMed and screened for their binding with Mpro protein. In addition, molecular dynamic simulations (200 ns) were executed to get additional information. Some of the compounds are found to have strong affinity for Mpro, therefore these hits could be used to develop a therapeutic agent for SARS-CoV-2.

1. Introduction

COVID-19, a deadly disease with a high infection rate, is caused by the novel corona virus SARS-CoV-2. It was first reported from Wuhan (China) and then rapidly reached many countries. The virus spreads through human to human contact. Unfortunately, it has survived in different climatic conditions. The high infection rate as well as mortality rate of the disease is reflected from the data released by World Health Organization (https://covid19.who.int/). The emergence of this disease has led to serious socio-economic and health issues for human race [1–5].

In search of developing such drug it was essential to understand the nature, functioning and bio-chemistry of virus. Researchers quickly found that this virus has more than 70% genome similarity with previously reported corona virus SARS-CoV [1,2,10–13]. The high similarity was further found to exist in essential proteins like main protease, spike S protein, etc. of these two viruses. These proteins are essential for survival and replication of SARS-CoV-2.

Coronavirus (CoV) main protease (Mpro) is a key enzyme that participates in cleavage process of H-CoV polyproteins. Mpro, also known as Nsp5 and 3CLpro, homodimer consists of three domains, domain I (residues 8–101), domain II (residues 102–184) and domain III (residues 201–303), and a long loop (residues 185–200), which connects domains II and III [1,2,10–14]. The active site of this protein is situated in the gap between domains I and II, and the catalytic dyad of Cys145 and His41 is its important feature. Domain I and II have motifs that are representing the chymotrypsin catalytic domain, while domain III participates in the dimerization of protein and active enzyme production [15]. This protein is necessary for the processing of polyproteins and operates at 11 cleavage sites on the large polyprotein 1 ab. Fortunately, the cleavage specificity of this protein is different from human proteases; therefore an inhibitor of Mpro could be safe for humans [1,2,5,10–14].

The amino acid sequence and 3D-structures of this protein have been successfully resolved by researcher [15]. Consequently, Mpro has emerged as a valid target for developing a drug for COVID-19 using molecular docking. Molecular docking is a contemporary and rational approach to identify the important structural features that govern the activity profile of a molecule. It can be used for virtual screening to identify novel hits, which could be further optimized to develop a drug candidate.
There is an urgent need to develop a drug to control COVID-19. Unfortunately, transformation of a compound to drug is a time-consuming process due to optimization of ADME (Absorption, Distribution, Metabolism and Excretion) and minimization of toxicity. This could delay development of drug. However, many food compounds have innate ability to have better ADME related properties, or immune stimulatory effects without toxicity [16,17]. Therefore, transforming a food compound or its close derivative is a novel and plausible strategy to speed-up the process of developing a drug for COVID-19. In the present work, we have performed virtual screening of food compounds using Lipinski’s rule of five and molecular docking to screen potent inhibitors of this protein from food. In addition, thorough and systematic molecular dynamic simulations (200 ns) were executed to get additional information. The results could served as a tool to develop a safer drug for COVID-19.

2. Experimental section

2.1. Database collection, curation and filtering

In the present work, FooDB (https://foodb.ca/accessed on 11th May 2020) was selected as it comprises a rich collection of food constituents (26,467 compounds). The food database used in the present work comprises a variety of molecules thus covering a broad chemical space. As a part of data curation, all charged molecules, duplicate entries, organometallic compounds, etc. were removed. Then, Lipinski’s rule of five was used to filter this database, except that the lower limit for molar mass was set to 150 due to the large size of active site of Mpro (main protease) of SARS-CoV-2 [1,2,5,10–14]. This reduced the pool to a dataset of 7,486 molecules only. In order to narrow down the search, only polyphenols, coumarins, polyamines only were selected. Thereafter, Spermidine de-rivatives were collected from PubMed database (23 molecules). Molecular docking of these 23 spermidine derivatives was accomplished for Mpro protein. The protocol followed in the present work has been summarized in Fig. 1.

2.2. Molecular docking

The three-dimensional structure of COVID-19 main protease (COV19-MPpro) in complex with peptidomimetic inhibitor, N3 (pdb: 6lu7), was downloaded from the Protein Data Bank (PDB, https://www.rcsb.org/).

The three-dimensional coordinates of water molecules were removed from protein structure using BIOVIA Discovery Studio 4.5 (Dassault Systems, USA). Program iGEMDOCK [18] (BioXGEM, Taiwan) was used for removing co-crystallized ligands and performing molecular docking. Molecular docking was performed on optimized structures (force field: MMFF94) of compounds. Genetic parameters for molecular docking were set on: population size 200; generations 70; number of solution or poses: 3. Active site of MPpro according the bounded synthetic peptidomimetic inhibitor, N3.

After the docking procedure, protein-compound interaction profiles of electrostatic (Elec), hydrogen-bonding (Hbond), and van der Waals (vdW) interactions were generated. Docking poses were ranked by combining the pharmacological interactions and energy-based scoring function \( E/kcal mol^{-1} \): \( E = vdw + Hbond + Elec \) [18]. Results were viewed and analyzed with BIOVIA Discovery Studio 4.5.

2.3. Molecular dynamics simulations

After docking studies of MPpro and compared the energy-based scoring functions of each food compound, we selected the top three food compound complexes for studying the docked structure’s thermodynamics stability. These three best-docked complexes are complex1 (MPpro/7), complex2 (MPpro/17), and complex3 (MPpro/85). We also rename the food compound 7, 17, and 85 to ligand1, ligand2, and ligand3. The molecular dynamics simulations were performed with the AMBER18 [19] package using ff14SB [20] force field for MPpro and updated generalized Amber force field (GAF2) [21] for food compounds. The protonation states of the charged residues were determined using the Propka 3.1 module [22]. The inhibitors were assigned AM1-BCC [23] charge, which was calculated by utilizing the Antechamber module [24]. The complexes were solvated in a truncated octahedron periodic box with an explicit TIP3P [25] water model and set a buffer distance cut-off at 10 Å from any edge to any protein atom. To make the whole system charge-neutral, we add a suitable number of Na⁺ ions to each system. The protein, complex, and food compounds topology and coordinates were prepared with the tLeap [26] module of AMBER suite.

Energy minimization of each system was performed by using the very famous steepest descent and conjugate gradient algorithms. All bond lengths involving hydrogen atoms were constrained by the SHAKE algorithm [27]. The particle mesh Ewald summation (PME) [28] approach was employed to treat long-range electrostatic interactions between the MPpro and food compounds. For all cases, the nonbonded Coulomb cut-off was fixed at 10 Å. An overall pressure and a temperature equal to 1 atm and 300 K were used with a time-frequency of 2 fs. The temperature was...
kept constant inside the box with the Langevin thermostat [29] temperature coupling method and Berendsen Barostat [30] to monitor the system pressure. All the steps of MD run and parameters, we adopted from our previous studies [31–37]. Finally, each system was subjected to 200 ns production MD run with a simulation time step of 2 fs at the NPT ensemble. Overall, we accumulated 20000 conformations for each simulation, and we used 2000 snapshots from the last 100 ns trajectories for binding affinity calculations.

2.4. Trajectory and binding free energy analysis

Trajectory analysis was performed using the AmberTools19 Cpptraj module [38], and all the plots are generated by matplotlib [39]. The binding affinity of ligand1, ligand2, and ligand3 toward SARS-CoV-2 3CL Mpro, were calculated by the molecular mechanics generalized Born (MM-GBSA) method [40] for gas phase, the desolvation free energy, and the conformational entropy decomposition scheme. All the parameters used in this calculation energy contributions at the residual level using by same MM-GBSA software used.

Table 1

| Compound (pose) | Total energy (kcal mol⁻¹) with COV19-MPro |
|-----------------|------------------------------------------|
| 85 (0)          | –141.27                                  |
| 78 (1)          | –122.71                                  |
| 68 (0)          | –120.38                                  |
| 30 (0)          | –119.89                                  |
| 77 (1)          | –119.68                                  |
| 4 (1)           | –118.55                                  |
| 76 (1)          | –115.57                                  |
| 79 (1)          | –115.30                                  |
| 3 (0)           | –114.57                                  |
| 33 (0)          | –114.46                                  |
| 65 (2)          | –113.05                                  |
| Remdesivir (4)  | –116.05                                  |

Table 2

| Compound (pose) | Total energy (kcal mol⁻¹) with COV19-MPro |
|-----------------|------------------------------------------|
| 85 (0)          | –141.27                                  |
| 7 (2)           | –130.79                                  |
| 17 (0)          | –129.08                                  |
| 6 (1)           | –122.15                                  |
| 18 (0)          | –121.54                                  |
| 12 (2)          | –121.45                                  |
| 8 (1)           | –120.71                                  |
| 13 (0)          | –116.97                                  |
| 19 (1)          | –114.85                                  |
| 20 (0)          | –112.92                                  |
| 22 (0)          | –102.14                                  |
| Remdesivir (4)  | –116.05                                  |

Table 3

Energies of the main interactions of N1, N10-dicoumaroylspermidine (compound 85) with amino acid residuals of binding site of COVID-19 Mpro. (M = main chain; S = side chain).

| H bonds | Residual     | Energy | vdW interactions |
|---------|--------------|--------|-----------------|
| M-Ser 144 | –3.50     | M-Thr 25 | –1.87           |
| S-Ser 144 | –2.50     | S-Thr 25 | –3.03           |
| M-Cys 145 | –3.50     | S-His 163 | –3.05          |
| S-Cys 145 | –3.85     | S-His 41 | –1.68           |
| M-Phe 140 | –3.50     |        |                 |
| S-Glu 166 | –2.44     |        |                 |
| S-Asn 142 | –2.57     |        |                 |
| M-Gly 143 | –0.48     |        |                 |

(faba bean) and Pyrus communis (pear). Considering that spermidine obtained the best docking results for both receptors related to the SARS-CoV-2, further molecular docking was performed on a new set of 24 spermidine derivatives, including compound 85 from the last set.

The data for 24 spermidine derivatives are given in Supplementary Files 2 (Table SF2). The best eleven results of molecular docking performed on COVID-19 main protease are presented in Table 2.

From the set of spermidine derivatives, compound 85 again showed the highest affinity for the binding to the COV19-Mpro. Comparing the total energies of interaction for the set of 106 food compounds with energies of spermidine derivatives, three spermidine derivatives have shown better results for COV19-Mpro. Thus, spermidines have proven to be the leading food compounds for the treatment of COVID-19. Observing the mode of interactions of the best ranked compounds, we will try to define structural characteristics important for the inhibition.

Compound 85 is an amine, a polyphenol, a secondary amino compound and a secondary carboxamide in which each of the primary amino groups has been mono-acetylated by formal condensation with trans-coumaric acid. In Table 3 are given the energies of the main interactions with amino acid residuals. Fig. 2 shows the docking pose and interactions of molecule 85 at the N3-binding site of COVID-19 Mpro (pdb: 6lu7).

Compound 85 forms interactions with residuals in catalytic dyad composed of Cys144 and His41. It forms several strong hydrogen bonds: oxygen atom from one amide group forms hydrogen bonds with Cys145 (2.69 Å) and Ser144 (3.17 Å); nitrogen atom from amine group forms hydrogen bonds with Phe140 (3.01 Å) and Gln166 (3.25 Å), while nitrogen atom from second amide group forms H-bond with Asn142 (3.23 Å). Phenolic hydroxyl group generates carbon hydrogen bond with Asn142 (3.59 Å). Phenol ring from one coumaric acid generates π-π stacked interactions with Tyr118 (5.66 Å). Fig. 3 presents the surface of COVID-19 Mpro coloured by hydrogen bond type in complex with compound 85. Figure shows how compound 85 is situated in the gap between domains I and II. Recently, the crystal structure of COVID-19
virus Mpro in complex with N3 elucidated specific interactions of N3 with Mpro [55]. N3 creates hydrogen bond with His 163 at S1 subsite and makes van der Waals contacts with Pro 168 and the backbone of residues 190–191. The bulky benzyl group forms van der Waals interactions with Thr 24 and Thr 25, similar as compound 85. Also, N3 forms multiple hydrogen bonds with the main chain of the residues in the substrate-binding pocket helping to lock the inhibitor inside the substrate-binding pocket. As stated, N3 and compound 85 bind to Mpro's in a similar mode.

An obvious advantage of using food compounds is that most of the food compounds are less toxic to human body therefore they could serve as excellent starting points for developing a drug for a disease.

3.1. Structural analysis of food complexes from molecular dynamics simulations

Firstly, we computed the root-mean-square deviations (RMSD) of backbone atoms of protein for all the complexes relative to their initial structures. The time evolution of the RMSD of three complexes is shown in Fig. 4(A). It is evident that RMSD values for 3CL Mpro complexes with ligand1, ligand2, and ligand3 remain stable after 75 ns depicting the convergence of simulations. Overall, within the last 125 ns, all systems got converged.

The average values of RMSDs for three complexes are listed in Table 4. The average values vary between 1.61 ± 0.02 Å and 2.92 ± 0.06 Å. The highest deviation was observed for the complex1 (2.92 ± 0.06 Å), while the lowest was obtained for complex2 (1.62 ± 0.02 Å). It suggests that complex2 and complex3 are more stable than complex1 in the last 125 ns simulations based on average RMSD deviations. Higher RMSDs of complex1 is mainly attributed to extended loop rearrangement (residue 185–200) and domain III (residue 200–306) fluctuations.

Besides, to get more insight into the extent to which the binding of food compounds affects the Mpro structural fluctuation and flexibility in a single amino acid level during the MD simulation, the root mean square fluctuation (RMSF) was analyzed, see Fig. 4(B). From the RMSF plot, we observed that complex1 is more flexible in the domain III region. Mainly, domain III is involved in forming a homodimer to be a reason for higher fluctuations in food compound, and domain I and II involved in ligand binding. We observed that those residues involved in food compound binding interactions show lower fluctuation than non-binding residues.
from the domain I and domain II. Residues Thr25, Thr26, Thr45, Cys44, Ser46, His41 are shows low fluctuations, and Leu50, Pro52, Asn51, Asn53, Glu55 shows high fluctuations from the domain I. Similarly, Ser144, Cys145, His164, Ser144, Gln189, Thr190, Asn142, Pro168, Met165, Gln192, and Ala191 residues from domain II shows low fluctuations. This suggests that food compound binding pocket residues are less flexible. The lower degree of fluctuation gives us a better compactness and rigidity complex structure.

As we know that the radius of gyration (Rg) helps us to understand the compactness of receptor protein, keeping mind, we monitored the Rg of Mpro through the entire 200 ns simulation and the time evolution of Rg shown in Fig. 4(C). The average values of Rg of three complexes are listed in Table 4. As evident from Table 4, all three complexes show more or less similar compactness. It could sometimes significantly affect the protein structure after binding a new ligand and side by side the solvent-accessible surface area (SASA) of changes. We know that SASA is very important for non-polar solvation energy measurement, directly affecting the ligand binding. Therefore, we also explored the SASA for three complexes and plotted with respect to simulation time, shown in Fig. 4(D). The average value of SASA varies from 14067.56 ± 32.02 Å² to 14498.94 ± 32.46 Å². The highest being reported for complex1 and low for the complex3.

To further explain the conformational stability, we also measured structural variations, RMSD of protein in its ligand-binding pocket, including all amino acids that fall within a radius of 5 Å from the ligand, see in Fig. 5(A). Complex1 binding site RMSD fluctuates around 1 Å up to 100 ns after that deviation is increased by two-fold and fluctuating at

| System   | RMSD (Å) | Rg (Å)          | SASA (Å²)          |
|----------|----------|-----------------|--------------------|
| Complex1 | 2.92 ± 0.06 | 21.86 ± 0.02    | 14498.94 ± 32.46   |
| Complex2 | 1.61 ± 0.02 | 21.96 ± 0.02    | 14213.10 ± 56.30   |
| Complex3 | 1.78 ± 0.03 | 21.98 ± 0.01    | 14067.56 ± 32.02   |

Fig. 3. Surface of COVID-19 Mpro (pdb: 6lu7) coloured by hydrogen bond type, with receptor donors coloured in green and receptor acceptors in cyan in complex with compound N1,N10-dicoumaroyl spermidine compound (85). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Fig. 4. (A) Time evolution of root-mean-square deviations (RMSDs) of backbone atoms, (B) the root-mean-square fluctuations (RMSFs) of Cα atoms, (C) radius of gyration, Rg of Cα atoms, and (D) solvent accessible surface area (SASA) of Mpro of three complexes relative to their respective energy minimized structure.

Table 4

The average backbone RMSD, the radius of gyration (Rg), and solvent accessible surface area (SASA) for the best three complexes. The data are reported as average ± standard error of the mean (SEM).
In the case of complex2, initial 30 ns, we see some drifting in the RMSD values, and after that up to 150 ns binding pocket residues fluctuating around 1 Å. Finally, it reached an equilibrium stage at a slightly higher RMSD value, i.e., 1.8 Å in the last 50 ns. For complex3, we see the increasing pattern of RMSD value in the first 50 ns after that reached stable equilibrium and fluctuating around 2 Å up to 200 ns. Overall, it suggests that the binding of a food compound in the binding site of Mpro stabilized the pocket conformation.

Furthermore, to determine the dynamics of the food compound throughout the simulations, the potential of mean force (PMF) was plotted w.r.t to RMSD of the food compound and shown in Fig. 5 (B). Fig. 5(B) showed that the ligand1 in complex1 showed a single global minimum at ~ 1.7 Å and exhibiting a very narrow peak and room temperature accessible secondary minimum at ~2.1 Å, suggesting the stability of the inhibitor in the binding site of the complex1. Ligand2 in complex2 also showed a single global minimum at ~2.2 Å and exhibiting a very narrow peak. The other secondary minimum structure was obtained at ~1.7 Å, but the energy barrier between these adjacent structures was high and found to be ~1.5 kcal/mol. Complex3 PMF profile shows a broad peak at 4 Å, and the conformational sampling space is wider relative to the other two complexes. Overall, the PMF profile of ligand molecules suggests that the ligand2 in the binding site is more flexible, reflecting on the binding affinity.

Finally, we monitored the CoM distance of food compounds and two vital domains involved in ligand binding, domain I and domain II, shown in Fig. 5 (C, D). In both cases, complex1 and complex2 maintain a stable distance. Fig. 5(C) indicates that ligand2 is positively shifted toward domain I, and ligand1 is 5 Å farther away from the domain I relative to ligand2. Initially, ligand3 is much far away from the domain I, but within 50 ns–125 ns, distance decreases, and the entire last 75 ns, it is stable around 18 Å. Similarly, ligand1 is shuffled toward domain II, and ligand 2 is 5 Å farther away from domain II than ligand1. Ligand3 shows a very flexible distance from domain II compared to ligand1 and 2. This time evolution of distances gives an important message to understand food compound behavior inside the binding pocket of SARS-CoV-2 main protease, Mpro.

### Table 5

| COMPONENTS | COMPLEX1 | COMPLEX2 | COMPLEX3 |
|------------|----------|----------|----------|
| ΔE(elec)   | -48.7 ± 0.2 | -22.6 ± 0.2 | -18.6 ± 0.2 |
| ΔE(int)    | 67.6 ± 0.2 | 38.7 ± 0.1 | 23.8 ± 0.1 |
| ΔG(np)     | -6.7 ± 0 | -5.1 ± 0 | -5.6 ± 0 |
| ΔG(solv)   | 60.9 ± 0.2 | 33.6 ± 0.1 | 28.2 ± 0.1 |
| ΔG(pot)    | 18.9 ± 0.4 | 16.1 ± 0.2 | 15.2 ± 0.2 |
| ΔG(sep)    | -94.6 ± 0.3 | -99.7 ± 0.2 | -95.5 ± 0.2 |
| ΔG(int)    | -33.7 ± 0.4 | -26.1 ± 0.2 | -31.3 ± 0.2 |

*ΔG(solv) = ΔG(np) + ΔG(pot),
ΔG(pot) = ΔE(elec) + ΔG(pot),
ΔG(int) = ΔE(int) + ΔG(int)”

To elucidate the binding mechanism of three food compounds to the Mpro using MM-GBSA analysis, we have computed the total binding energy and its various components contributing to the binding free energy (ΔGbind). The MM-GBSA based binding affinity calculations were performed on the production simulation trajectories. The various components include van der Waals interactions (ΔE(vdW)), electrostatic interactions (ΔE(elec)), polar solvation free energy (ΔG(solv)), and non-polar solvation free energy (ΔG(pot)), which are listed in Table 5 and shown in Fig. 6. It is evident that in all cases, food compound-Mpro complexation is
favored by the van der Waal interactions ($\Delta E_{vdW}$) and electrostatic interactions ($\Delta E_{ele}$). The effects of solvent around the Mpro can also be studied, and it depicts that non-polar solvation free energy ($\Delta G_{np}$) favors the complexation. In contrast, the polar solvation free energy ($\Delta G_{pol}$) disfavor the complex formation.

It is evident from Table 5, the predicted binding free energies ($\Delta G_{bind}$) are $-33.7 \pm 0.4$ kcal/mol, $-26.1 \pm 0.2$ kcal/mol, and $-31.3 \pm 0.2$ kcal/mol for complex1, complex2, and complex3, respectively. It suggests that the food compound, ligand1 (7), and ligand3 (85) binds strongly with the Mpro in comparison to ligand2 (17). As shown in Table 5, that for all
complexes, van der Waals (ΔE_{vdw}) varies between −37.1 ± 0.1 kcal/mol and −45.9 ± 0.1 kcal/mol, while in electrostatic interactions, ΔE_{elec} ranges from −18.6 ± 0.2 to −48.7 ± 0.3 kcal/mol. The van der Waal interactions favor the most in complex2 and complex3 compared to the electrostatic interactions but in complex1 electrostatic interaction is higher than van der Waals. Furthermore, for all the electrostatic interaction components, ΔE_{elec} are over-compensated by the polar desolvation energy, ΔG_{pol}, suggesting that the total polar (ΔG_{pol} + elec) components are unfavorable to the food compound binding toward M^{pro}.

It is evident from Table 5 that the energy of ΔE_{vdw} and ΔE_{elec} for the complex2 was less favorable than complex1 as found to be −22.6 ± 0.2 kcal/mol and −37.1 ± 0.1 kcal/mol, respectively. So, the affinity of the three food compounds increases in the following order: ligand2 (17) < ligand3 (85) < ligand1 (7). Our results revealed that ligand1 (7) binds most strongly to M^{pro} due to the higher value of total internal molecular mechanics energy, ΔE_{internal} is more favorable to the binding than the other two food compounds.

To further explore the critical residues involved in the food compounds’ binding mechanism to SARS-CoV-2 main protease, we computed the per-residue decomposition of free energy using the MM-GBSA [56] method. The approach of per-residue based contributions is useful to determine the binding mechanisms at a residual atomic level. The different energy contributions from each residue’s backbone and side-chain are shown in Fig. 7 and listed in Table 6. Here all the reported interacting residues energy contribution is −1.0 kcal/mol or higher.

It is evident from Table 6 that the number of highly favorable residues in the food compound binding is more or less the same. As shown in Fig. 7(A), we observed that residues involved in binding ligand1 (7) are Ser144, Gln189, Asn142, Pro168, Cys145, and Met165. All these residues are located in domain II and form intense contact with ligand1. In Fig. 6(B), we found that residues Thr25, Thr45, Met49, His41, Ser46, Gln189, Cys44, and Leu27 are the most energy contributing amino acids to the binding of ligand2 (17). Most of these residues are located in domain I and make close contact with ligand2. Similarly, Table 6 and

Table 6

| Residue     | E_{vdw} | E_{elec} | G_{pol} | G_{side-chain} | G_{backbone} | G_{total} |
|-------------|---------|----------|---------|----------------|--------------|-----------|
| Complex1    |         |          |         |                |              |           |
| Ser144      | −0.7    | −2.6     | 1.6     | −0.1           | −0.7         | −1.1      | −1.8      |
| Gln189      | −1.3    | −2.1     | 2.0     | −0.3           | −1.0         | −0.7      | −1.7      |
| Asn142      | −1.9    | −2.1     | 2.8     | −0.3           | −0.6         | −0.9      | −1.5      |
| Pro168      | −1.7    | 0.3      | 0.1     | −0.1           | −1.2         | −0.2      | −1.4      |
| Cys145      | −1.2    | −0.2     | 0.2     | −0.1           | −0.7         | −0.6      | −1.3      |
| Met165      | −1.0    | −0.2     | 0.1     | −0.1           | −0.6         | −0.6      | −1.2      |
| Complex2    |         |          |         |                |              |           |
| Thr25       | −1.0    | −2.8     | 0.7     | −0.2           | −3.1         | −0.2      | −3.3      |
| Thr45       | −1.5    | −2.3     | 1.2     | −1.0           | −1.0         | −1.7      | −2.7      |
| Met49       | −1.9    | 0.0      | 0.4     | −0.3           | −1.7         | −0.5      | −1.2      |
| His41       | −2.4    | −1.4     | 2.5     | −0.4           | −1.1         | −0.6      | −1.7      |
| Ser46       | −1.7    | −0.3     | 1.0     | −0.3           | −0.6         | −0.7      | −1.3      |
| Gln189      | −1.5    | −0.7     | 1.2     | −0.2           | −0.7         | −0.5      | −1.2      |
| Cys44       | −1.1    | −1.1     | 1.3     | −0.2           | −0.5         | −0.6      | −1.1      |
| Leu27       | −0.9    | −0.2     | 0.2     | −0.1           | −0.9         | −0.1      | −1.0      |
| Complex3    |         |          |         |                |              |           |
| Gln189      | −2.1    | −3.4     | 3.1     | −0.4           | −2.2         | −0.6      | −2.8      |
| Gln192      | −1.1    | −2.1     | 1.6     | −0.1           | −1.4         | −0.3      | −1.7      |
| Met165      | −1.5    | −0.3     | 0.5     | −0.1           | −1.3         | −0.1      | −1.4      |
| Met49       | −1.2    | −0.1     | 0.2     | −0.1           | −1.2         | −0.2      | −1.3      |
| Ala191      | −1.2    | −0.2     | 0.4     | −0.2           | −0.6         | −0.6      | −1.2      |
| His41       | −1.5    | −0.4     | 1.0     | −0.2           | −0.8         | −0.3      | −1.1      |
| Pro168      | −0.8    | −0.1     | 0.1     | −0.2           | −0.8         | −0.2      | −1.0      |

* Energetic contributions from the van der Waals (E_{vdw}) and electrostatic interactions (E_{elec}) as well as polar (G_{pol}) and non-polar solvation energy (G_{pol}) and the total contribution of given residue (G_{total}) for SARS-CoV-2 M^{pro}-food compound complexes are listed. G_{side-chain} and G_{backbone} represent the side chain and backbone contributions. Only residues with |ΔE| ≥ 1.0 kcal/mol are shown. All values are given in kcal/mol.

Table 7

| Acceptor  | Donor      | Avg. Distance (Å) | Occupancy (%) |
|-----------|------------|-------------------|--------------|
| Complex1  |            |                   |              |
| Asn142@OD1| Lig@N1     | 2.85              | 22.17        |
| Gln166@OE2| Lig@N3     | 2.85              | 20.62        |
| Gln189@O  | Lig@O6     | 2.72              | 18.61        |
| Lig@O2    | Ser144@N   | 2.88              | 15.83        |
| Lig@O2    | His163@NE2 | 2.86              | 14.84        |
| Gln166@OE2| Lig@N1     | 2.82              | 13.64        |
| Gln166@OE1| Lig@N1     | 2.83              | 13.58        |
| Ser166@O  | Lig@O6     | 2.74              | 12.90        |
| Gln166@OE1| Lig@N3     | 2.85              | 12.86        |
| Complex2  |            |                   |              |
| Thr25@OG1 | Lig@O12    | 2.73              | 68.60        |
| Thr45@O  | Ligg@O7    | 2.77              | 30.35        |
| Thr24@OG1 | Ligg@O11   | 2.91              | 18.83        |
| His41@O  | Ligg@O7    | 2.77              | 16.33        |
| Gln189@OE1| Ligg@O2    | 2.77              | 15.24        |
| Thr24@OG1| Ligg@O12   | 2.96              | 13.65        |
| Lig@O4   | Asn42@ND2  | 2.91              | 11.25        |
| Lig@O12  | Thr24@OG1  | 2.80              | 10.21        |
| Complex3  |            |                   |              |
| Gln192@O | Ligg@N3    | 2.85              | 17.45        |
| Cys44@O  | Ligg@N3    | 2.92              | 16.56        |
| Lig@N2   | Gln192@NE2 | 2.93              | 15.10        |
| Lig@O2   | Gln189@NE2 | 2.85              | 14.68        |
| Cys44@O  | Ligg@N2    | 2.89              | 12.41        |
| Met165@HG3| Ligg@N2    | 2.78              | 10.11        |
| His41@O | Ligg@O1    | 2.79              | 10.03        |

* The hydrogen bonds are determined by the acceptor … donor distance of ≤3.5 Å and acceptor … H-donor angle of >120°.

**Occupancy** (to evaluate the stability and strength of the hydrogen bonds) is defined as the percentage of simulation time that a specific hydrogen bond exists, the hydrogen bonds occurring less than 10% of simulations are not shown.

![Fig. 8. Time evolution of the number of hydrogen bonds between three food compounds and SARS-CoV-2 M^{pro} with respect to their initial conformations.](Image)

3.3. Hydrogen bonding and hydrophobic interactions analysis

We performed production trajectory-based hydrogen bonds analysis for three food compound complexes to complement the binding free
Fig. 9. The Mpro-food compound interactions profile for (A) complex1, (C) complex2, and (E) complex3. The food compounds are shown in balls and sticks. Hydrogen bonds are depicted in green dotted lines, and red semicircles and dotted lines are involved in hydrophobic interactions. Each interacting residual position in the binding pocket is shown in the right panel for (B) complex1, (D) complex2, and (F) complex3. The blue color residues are involved in hydrophobic interactions, and green color residues are involved in hydrogen bond formations. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)
energy analysis, and more 10% occupancy hydrogen bonds are reported in Table 7. The time evolution of hydrogen bonds for three complexes are also shown in Fig. 8. As suggested by Fig. 8, we observed that complex1 has a relatively more significant number of total numbers of hydrogen bonds than the other two complexes. Furthermore, complex3 hydrogen bond time spectra have high dynamics. In complex1, critical residues involved in the hydrogen bonding are Asn142, Glu166, Gln189, Ser144, and Ser46. We found two stable hydrogen bonds with more than 20% occupancy between ligand1 (7) and Mpro (Asn142@OD1-Lig@N1 and Glu166@OE2-Lig@N3) see in Table 7.

In the case of complex2, both Thr25 and Thr45 formed a hydrogen bond with ligand2 (17) with an occupancy of 86.60% (Thr25@OG1-Lig@O12) and 30.35% (Thr45@O-Lig@O7), respectively. Thr24, His41, Gln189, and Asn142 also formed hydrogen bonds with ligand2 (17) during our simulations with an occupancy range from 10% to 18%. Finally, in complex3, the hydrogen bonds’ residues are Gln192, Cys44, Gln189, Met165, and His41. Both Gln192 and Cys44 form two hydrogen bonds with ligand3 (85) with an occupancy range of 12.41%–17.45%. On the other hand, His41 and Met165 form hydrogen bonds with an occupancy of ~10%, Met165@HG3-Lig@N2, and His41@O-Lig@O1, respectively.

Finally, we supplemented the above results by analyzing the final production simulation conformation with LigPlot + software for each complex. The interacting residues, both hydrophobic and hydrogen, are shown in Fig. 9, and side by side, its corresponding position in the protein structure is also shown in the right panel. Hydrogen bonds are depicted in green dotted lines, and red semicircles and dotted lines are involved in hydrophobic interactions. For complex1, Fig. 9(A and B) displayed, 13 hydrophobic interactions with His41, His164, His163, Gln189, Met165, Leu167, Pro168, His172, Ser139, Phe140, Leu141, Asn142 and Cys145 (blue color residues in Fig. 9(B)). Moreover, we also found four stable hydrogen bonds in the final hydrogen conformation, shown in green color in Fig. 9(B). This large number of hydrophobic and hydrogen bonds interactions account for the high stability and good binding affinity of ligand1 (7) to SARS-CoV-2 Mpro. Ligand2 (17) formed hydrophobic interactions with Ser46, Cys44, His41, Asp187, Cys145, Gln189, and Met49 and hydrogen bonds with Thr24, Thr25 and Thr45, (see in Figure (C, D)). Finally, Figure (E, F) shows that complex3 formed hydrogen interactions with Thr45, Ser46, Met49, His41, Arg188, Ala191, Leu167, Met165, His164 and Glu166 and four hydrogen bonds with Gln189, Thr190, Gln192, and Cys44. Overall, ligand1(7) and ligand3 (85) have a higher binding affinity against Mpro compared to ligand2 (17) due to a larger number of hydrophobic and hydrogen bonds interactions.

4. Conclusion

In conclusion, for the first time, a food database has been used in unique way to identify hits for COVID-19. The present docking studies along with robust support from molecular dynamics simulations indicated that two derivatives of spermidine exhibited high potential for binding to the active site of COVID-19 Mpro. The key structural features for the inhibition of COVID-19 main protease of N1,N10-dicoumaroyl spermidine are: nitrogen atom from amine group and phenolic hydroxyl groups. Study has revealed that spermidines, as food constituents, are interesting target for the development of a drug or natural healing of COVID-19 disease.

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.

Acknowledgements

The authors are thankful to the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University, Riyadh, KSA, for its support of this research through the Research Grant No. 21-13-18-070. PK acknowledges the research support by the Department of Biotechnology, Govt. of India (Grant number BT/RLF/Re-entry/40/2014, Ramalingaswami Re-entry Fellowship). MFS receives PhD fellowship from DST-INSPIRE, Govt. of India.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chemolab.2021.104394.

References

[1] L. Zhang, D. Lin, X. Sun, U. Curth, C. Drosten, L. Sauererhing, S. Becker, K. Rox, R. Hilgenfeld, Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketamide inhibitors, Science 368 (2020) 409–412, https://doi.org/10.1126/science.abb4050.
[2] C.W. Wu, Y. Liu, Y. Yang, P. Zhang, W. Zhong, Y. Wang, Q. Wang, Y. Xu, M. Li, X. Li, M. Zheng, L. Chen, H. Li, Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods, Acta Pharm. Sin. B 10 (2020) 766–788, https://doi.org/10.1016/j.apsb.2020.02.008.
[3] V.H. Masand, V. Rastija, M.K. Patil, A. Gandhi, A. Chapolikar, Extending the identification of structural features responsible for anti-SARS-CoV-2 activity of peptide-type compounds using QSAR modeling, SAR QSAR Environ. Res. 31 (2020) 643–654.
[4] V.H. Masand, S. Akasapu, A. Gandhi, V. Rastija, M.K. Patil, Structure features of peptide-type SARS-CoV-2 main protease inhibitors: quantitative structure activity relationship study, Chemometr. Intell. Lab. Syst. (2020) 206.
[5] R.J. Khan, R.K. Jha, G.M. Amera, M. Jain, E. Singh, A. Pathak, R.P. Singh, J. Muthukumaran, A.K. Singh, Targeting SARS-CoV-2: a systematic drug repurposing approach to identify promising inhibitors against 3C-like protease and Z'-α-chole methyltransferase, J. Biomet. Struct. Dyn. (2020) 1–14.
[6] R. Chilamakuri, S. Agarwal, COVID-19: characteristics and therapeutics, Cells (2021) 10.
[7] A.J. Mulholland, R.E. Amaro, COVID-19 : computational chemists meet the moment, J. Chem. Inf. Model. 60 (2020) 5724–5726.
[8] N. Tripathi, N. Tripathi, M.K. Goshish, COVID-19: Inflammatory Responses, Structure-Based Drug Design and Potential Therapeutics, Molecular Diversity, 2021.
[9] T. Kirby, New variant of SARS-CoV-2 in UK causes surge of COVID-19, The Lancet Respiratory Medicine (2021) e20–e21, https://doi.org/10.1016/S2213-2600(21)00005-9.
[10] Z. Jin, X. Du, Y. Xu, Y. Deng, M. Liu, Y. Zhao, B. Zhang, X. Li, L. Zhang, C. Peng, Y. Du, E. 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, Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors, Nature 582 (2020) 289–293.
[11] S. Jeon, M. Ko, J. Lee, I. Choi, S.Y. Byun, S. Park, D. Shim, S. Kim, Identification of Antiviral Drug Candidates against SARS-CoV-2 from FDA-Approved Drugs, Antimicrobial Agents and Chemotherapy, vol. 64, 2020.
[12] L. Fu, F. Ye, Y. Feng, F. Yu, Q. Wang, Y. Wu, C. Zhao, H. Sun, B. Huang, P. Niu, H. Song, Y. Shi, X. Li, W. Tan, J. Qi, G.F. Gao, Both Beceprevir and GC376 efficaciously inhibit SARS-CoV-2 by targeting its main protease, Nat. Commun. 11 (2020).
[13] P. Chowdhury, In silico investigation of phytoconstituents from Indian medicinal herb ‘Tinospora cordifolia (glpy)’ against SARS-CoV-2 (COVID-19) by molecular dynamics approach, J. Biomet. Struct. Dyn. (2020) 1–18.
[14] Y. Zhou, Y. Hou, J. Shen, Y. Huang, W. Martin, F. Cheng, Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2, Cell Discov 6 (2020) 14.

Drafting the manuscript: V. H. Masand, V. Rastija, P. Kar, M.E.A. Zaki.

Revising the manuscript critically for important intellectual content: M.E.A. Zaki.

Category 3.

Approval of the version of the manuscript to be published (the names of all authors must be listed): V. H. Masand, V. Rastija, P. Kar, M.E.A. Zaki, M. Fulbabu.
[15] F. Wang, C. Chen, W. Tan, K. Yang, H. Yang, Structure of main protease from human coronavirus NL63: insights for wide spectrum anti-coronavirus drug design, Sci. Rep. 6 (2016) 25677.

[16] Childs, Calder, miles, diet and immune function, Nutrients 11 (2019).

[17] A.J. Teodoro, Bioactive compounds of food: their role in the prevention and treatment of diseases, Oxidative Med. Cell. Longevity 2019 (2019) 1–4.

[18] K.-C. Hsu, Y.-P. Chen, S.-R. Lin, J.-M. Yang, GEMDOCK: a graphical environment of enhancing GEMDOCK using pharmacological interactions and post-screening analysis, BMC Biol. 12 (2011) S33.

[19] D.A. Case, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, I.T.E. Cheatham, D.J. Price, C.L. Brooks III, A modiﬁcation of the Amber biomolecular dynamics force-ﬁeld, J. Comput. Chem. 23 (2002) 1623–1641.

[20] J. Wang, W. Wang, P.A. Kollman, D.A. Case, Antechamber: an accessory software package for molecular mechanical calculations, J. Chem. Inf. Comput. Sci. 42 (2002) 1403.

[21] D.J. Price, C.L. Brooks III, A modiﬁed TIP3P water potential for simulation with Ewald summation, J. Chem. Phys. 121 (2004) 10096–10103.

[22] R. Salomon-Ferrer, D.A. Case, R.C. Walker, An overview of the Amber biomolecular simulation package, Wiley Interdisciplinary Reviews: Computational Molecular Science 3 (2013) 198–210.

[23] V. Krautler, W.F. Van Gunsteren, P.H. Hünenberger, A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations, J. Comput. Chem. 22 (2001) 501–508.

[24] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an N- log(N) method for Ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089–10092.

[25] R.J. Louchart, B.R. Brooks, R.W. Pastor, Langevin dynamics of peptides: the frictional dependence of isomerization rates of N-acetyl-Lalanine-N'-methylamide, Biopolymers: Original Research on Biomolecules 32 (1992) 523–535.

[26] I.J. Berendsen, J.H. Berendsen, W.F. van Gunsteren, A. DiNola, J. Haak, Molecular dynamics with coupling to an external bath, J. Chem. Phys. 81 (1984) 3684–3690.

[27] R.F. Wijesinghe, R. Salomon-Ferrer, J. Shen, C.L. Simmerling, J. Smith, R. SalomonFerrer, J. Swails, R.C. Walker, J. Wang, H. Wei, R.M. Wolf, X. Wu, L. Xiao, D.M. Y, A.P.A. Kollman, AMBER 2018, University of California, San Francisco, 2018.

[28] D.A. Case, K. Marzlin, Y. Miao, F. Termier, M. Rostkowski, J.H. Jensen, PROPKA3: consistent and comprehensive treatment of internal and surface residues in empirical pK a predictions, J. Chem. Theory Comput. 10 (2014) 1145–1157.

[29] A.C. Wallace, R.A. Laskowski, J.M. Thornton, LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions, Protein Eng. Des. Sel. 8 (1995) 1384–1391.

[30] D.R. Roe, T.E. Cheatham III, PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data, J. Chem. Theor. Comput. 9 (2013) 3084–3095.

[31] N.A. Jonniya, M.F. Sk, P. Kar, A comparative study of structural and conformational properties of WNk kinase isoforms bound to an inhibitor: insights from molecular dynamic simulations, J. Biomol. Struct. Dyn. (2020), https://doi.org/10.1080/07391102.2019.1724196. just-accepted.

[32] C. Bartels, M. Karplus, Multidimensional adaptive umbrella sampling: applications to main chain and side chain peptide conformations, J. Comput. Chem. 19 (1998) 1450–1462.

[33] P. Kar, V. Knecht, Mutation-induced loop opening and energetics for binding of tamiflu to influenza N8 neuraminidase, J. Phys. Chem. B 116 (2012) 6137–6149.

[34] J. Wang, W. Wang, P.A. Kollman, D.A. Case, Investigating phosphorylation-induced conformational changes in WNk1 kinase by molecular dynamics simulations, ACS Omega 4 (2019) 17404–17416.

[35] S. Singh, M.F. Sk, A. Sonawane, P. Kar, S. Sadhukhan, Plant-derived natural polyphenols as potential antiviral drugs against SARS-CoV-2 via RNA-dependent RNA polymerase (RdRp) inhibition: an in-silico analysis, J. Biomol. Struct. Dyn. (2020) 1–16, just-accepted.

[36] M.F. Sk, N.A. Jonniya, R. Roy, S. Poddar, P. Kar, Exploring the potency of currently used drugs against HIV-1 protease of subtype D variant by using multiscale simulations, J. Biomol. Struct. Dyn. (2020) 1–21, just-accepted.