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.
Phytochemicals present in Indian ginseng possess potential to inhibit SARS-CoV-2 virulence: A molecular docking and MD simulation study

Prem Prakash Kushwaha\textsuperscript{a,1}, Atul Kumar Singh\textsuperscript{a,1}, Kumari Sunita Prajapati\textsuperscript{a}, Mohd Shuaib\textsuperscript{a}, Sanjay Gupta\textsuperscript{b}, Shashank Kumar\textsuperscript{a,b}

\textsuperscript{a} Molecular Signaling & Drug Discovery Laboratory, Department of Biochemistry, School of Basic and Applied Sciences, Central University of Punjab, Bathinda, 151401, India
\textsuperscript{b} Department of Urology, Case Western Reserve University, Cleveland, OH, 44106, USA

\textbf{ARTICLE INFO}

\textbf{Keywords:}
Withania somnifera
SARS-CoV-2
\textit{In silico studies}
M\textsubscript{pro}
Phytochemicals
Antiviral

\textbf{ABSTRACT}

Coronaviruses are deadly and contagious pathogens that affect people in different ways. Researchers have increased their efforts in the development of antiviral agents against coronavirus targeting M\textsubscript{pro} protein (main protease) as an effective drug target. The present study explores the inhibitory potential of characteristic and non-characteristic Withania somnifera (Indian ginseng) phytochemicals (n ≈ 100) against SARS-CoV-2 M\textsubscript{pro} protein. Molecular docking studies revealed that certain W. somnifera compounds exhibit superior binding potential (−6.16 to −12.27 kcal/mol) compared to the standard inhibitors (−2.55 to −6.16 kcal/mol) including nelfinavir and lopinavir. The non-characteristic compounds (quercetin-3-rutinoside-7-glucoside, rutin and iso-chlorogenic acid B) exhibited higher inhibitory potential in comparison to characteristic W. somnifera compounds withanolide and withanone. Molecular dynamics (MD) simulation studies of the complex for 100 ns confirm favorable and stable binding of the lead molecule. The MM-PBSA calculation of the last 10 ns of the protein-ligand complex trajectory exhibited stable binding of quercetin-3-rutinoside-7-glucoside at the active site of SARS-CoV-2 M\textsubscript{pro}. Taken together, the study demonstrates that the non-characteristic compounds present in W. somnifera possess enhanced potential to bind SARS-CoV-2 M\textsubscript{pro} active site. We further recommend \textit{in vitro} and \textit{in vivo} experimentation to validate the anti-SARS-CoV-2 potential of these lead molecules.

1. Introduction

COVID-19 is an infectious disease caused by a newly discovered SARS-CoV-2 virus. The number of infections has significantly exceeded globally and threatens human population worldwide. On January 30, 2020 World Health Organization (WHO) announced the COVID-19 epidemic as a public health emergency of international concern. Researchers from all over the world attempted to develop treatment against SARS-CoV-2 infection and identify effective agents to combat this global threat [1]. Studies from several laboratories have identified some SARS-CoV-2 proteins which are potential drug target(s) having ability to block viral-host interaction, virulence and pathogenesis [2, 44]. SARS-CoV-2 main protease (M\textsubscript{pro}), a cysteine protease plays a pivotal role in viral gene expression, transcription, and replication. It mediates its function through a highly complex cascade of proteolytic processing of replicate proteins [3, 4] developed activity-based probes for the M\textsubscript{pro} protein in diagnosing COVID-19 infected patients, providing a framework for drug design against this molecule [4]. Thus, targeting viral M\textsubscript{pro} protein could be a potential and attractive strategy against SARS-CoV-2 antiviral drug development. Moreover, synthetic antiviral drugs show common to life threatening side-effects such as nausea, diarrhea, hepatotoxicity, hyperglycemia etc. [5]. In contrast, natural agents show antiviral potential by inhibiting viral attachment, and targeting critical enzymes (helicase, protease and polymerase) involved in the pathogenesis of the disease [6]. Cost effectiveness and low side effects of phytochemicals make them excellent candidates for antiviral drug discovery.

Accumulating research demonstrates that medicinal plants possess therapeutic potential against microbes and pathogen [7–9]. Recent reports suggest that natural-like products and their metabolites present in
several plant spices exhibit remarkable potential to block SARS-CoV-2 Mpro receptor [10–12]. *Withania somnifera* (Solanaceae) also known as Indian Ginseng is a well-known medicinal plant that possess antiviral properties. Phytochemicals present in *W. somnifera* are known to possess antiviral activity against H1N1 influenza virus, HIV and infectious bursal disease. Studies indicate that the antiviral mode of action of *W. somnifera* compounds include viral protein inhibition, mitigation of viral infection induced inflammation as well as immunomodulatory effects [13–15]. *W. somnifera* possess both characteristic and non-characteristic (present in other plant also) phytochemicals. Recent literature data shows SARS-CoV-2 Mpro inhibitory potential in *W. somnifera* characteristic compounds especially wihanolides [16]. However, the non-characteristic compounds of the plant have not been screened against SARS-CoV-2 Mpro protein. The present study aims to identify lead *W. somnifera* phytochemical from a list of approximately hundred compounds having potential to inhibit the SARS-CoV-2 Mpro protein using computer aided drug discovery approach.

### 2. Materials and methods

#### 2.1. *Withania somnifera* compound retrieval and preparation

A total of 93 compounds present in *W. somnifera* plant were obtained from different literature and search engine platforms such as PubMed, Google Scholar, Web of Science, Scopus, Chemical Abstracts, Medline, and PubMed Central [17,18]. The structures of these compounds present in *W. somnifera* were prepared using Marvin Sketch software [19]. The 2D or 3D structure of standard compounds against target proteins were retrieved from the NCBI PubChem in. sdf format [20]. Open Babel molecule format converter was used to perform conversion of 2D to 3D conformation and their conversion from. sdf to. mol file [21]. Ligand energy was minimized by applying mmff94 force field and conjugate gradients optimization algorithm using PyRx-Python prescription 0.8 for 200 steps [22].

#### 2.2. Receptor retrieval

3D structure of SARS-CoV-2 main protease (PDB ID: 5RFS) receptor was obtained from protein data bank (https://www.rcsb.org/) [24]. The resolution of the retrieved structures was 1.70 Å.

### 2.3. Receptor preparation

3D structure of SARS-CoV-2 main protease was loaded onto UCSF Chimera for molecular docking preparation [25]. Protein models were cleaned and optimized by removing ligands and other heteroatoms including water. After this step, the energy minimization of protein structures was performed by steepest descent method having 100 steps (step size 0.02 Å), and a conjugate gradient method with 10 steps (step size 0.02 Å) using UCSF Chimera.

#### 2.4. Molecular docking studies

Auto Dock Tools v.1.5.6 (ADT) was used to dock the test ligands on targeted protein [26,27]. Gestgeiger partial charges assigned to the ligands and docking calculations were performed. Polar hydrogen atoms, Kollman charges, and solvation parameters were applied using appropriate Auto Dock tools. Lamarckian Genetic Algorithm (LGA) was used to explore the active binding region in this study. The grid box included the entire binding site of the protein providing enough space for the ligands translational and rotational maneuverability. Docking was performed by keeping the grid dimensions 9.05, −1.87 and 22.09 in X, Y and Z dimension with 10 Å grid spacing for the target protein. For each of the 30 independent runs, a maximum number of 27,000 GA operations generated on a single population of 150 individuals. Operator weights for the rate of crossover, gene mutation, and elitism were set as 0.80, 0.02, and 1, respectively. LigPlot+ (v.1.4.5) and UCSF Chimera (v.1.10.2) online tools were used for protein-ligand interaction visualization [28].

#### 2.5. Molecular dynamics simulation

MD simulations for all atoms were performed for Mpro protein of SARS-CoV-2 before and after the binding of the known inhibitors and identified compounds. GROMOS 54A7 force field was used to perform the molecular dynamics simulation using the GROMACS 5.1.1 package.
### Table 2
DockScore and list of amino residue involved in the interaction with standard inhibitors and *W. somnifera* phytochemicals with $-6.16 <$ cutoff value.

| S. No. | Ligand | Dock Score (kcal/mol) | Hydrogen bond interaction amino acid | Hydrophobic interaction amino acid |
|--------|--------|------------------------|--------------------------------------|-----------------------------------|
| S1     | Nelfinavir | $-6.16$ | Ami142, His41 | Leu141, His163, Cys145, Met165, Glu166, Gln189, His164, Asp187, Arg188, Ser46, Thr26, Met49, Leu27 |
| S2     | Lupinavir  | $-5.33$ | Gly143, Glu166 | Met165, Gln189, Arg188, His41, Thr26, Gly143, Thr24, Cys44, Thr25, Thr45, Ser46, Leu141, Asn142 |

### Table 2 (continued)

| S. No. | Ligand | Dock Score (kcal/mol) | Hydrogen bond interaction amino acid | Hydrophobic interaction amino acid |
|--------|--------|------------------------|--------------------------------------|-----------------------------------|
| 1      | Quercetin-3-rutinoside-7-glucoside | $-12.27$ | Thr25, Thr24, Ser46, Thr26, Ami142, Glu166, His164 | Gly143, Leu141, Thr45, His41, Met165, Arg188, Asn142, Thr190, Pro168 |
| 2      | Rutin    | $-10.16$ | Thr25, Thr24, Gly143, His164, Ami142 | Thr26, Cys45, Arg188, Met165, Asp187, Met49, Gln189, Thr167, Leu141, Asn142 |
| 3      | Isochlorogenic acid B | $-9.08$ | Gly143, His41, Phe40, Glu166 | His163, His164, Met49, Ser46, Gln189, Met165, Leu141, Asn142 |
| 4      | Caffeoyl quinic acid | $-8.22$ | Ami142, Gly143, His164 | Cys145, His41, Glu166, Met165, Gln189, Leu167, Pro168, Gln192 |
| 5      | 4-Deoxyphysolactone | $-7.91$ | Thr190, His41, Gly143, Gln189, Glu166 | Thr26, Thr25, Arg188, Met165, Asp187, Met49, Gln189, Thr190, Pro168 |
| 6      | Sitoinoside IX | $-7.74$ | Thr26, Cys145, Ami142, Leu41, Glu166 | Gly143, Thr25, His164, met49, Gln189, Met165, His163 |
| 7      | Withanolide J | $-7.72$ | Glu166, Gly143 | Ami142, Thr25, Leu27, Cys145, Thr26, His41, Met165, Met49, Gln189, Thr190, pro168 |
| 8      | Withanolide S | $-7.51$ | Gln189, Gly143, His41 | Gln166, Thr190, Arg188, Met165, Pro168, Met49, Thr26, Ami142, Cys145, His164 |
| 9      | Withanolide Q | $-7.45$ | Gly143, Glu166, Thr190 | Met149, Thr25, Leu27, His41, Cys145, Ami142, Gln189, Pro168 |
| 10     | Quercetin  | $-7.44$ | Thr190 | His164, His41, Met49, Met165, Gln192, Pro168, Glu166 |
| 11     | Chlorogenic acid | $-7.41$ | Gly143, Glu166, Thr190 | Cys145, His164, Ami142, Met165, Gln189, Pro168 |
| 12     | Withafusosin E | $-7.25$ | Thr26, Cys44, Glu166 | Gly143, Thr25, Ser46, Thr45, Met49, Gln189 |
| 13     | Withanolide N | $-7.21$ | Gly143, Glu166 | Ami142, Cys145, Leu27, Thr26, Thr25, His41, Met165, Gln189, Thr190, Pro168 |

**Fig. 1.** Analysis of *W. somnifera* characteristic, non-characteristic phytochemicals and standard inhibitor versus binding energy of targeted SARS-CoV-2 M·mut protein. NCWS-Non-characteristic *W. somnifera* phytochemicals, CWS-Characteristic *W. somnifera* phytochemicals.
Fig. 2. List of *Withania somnifera* phytochemicals lead SARS-CoV-2 main protease inhibitors with less than -6.00 kcal/J docking score cut-off value [34,35,37–39,41, 43,45].
Ligand topologies and other parameters generated using PRODRG server and combined with the protein topology to make simulation complexes. All systems were solvated using the SPC water model in a cubic box. Energy minimization was achieved to get rid of steric clashes by using 1500 steps of the steepest descent approach. Following energy minimization the temperature of all systems was increased from 0 to 300 K in equilibration phase of 100 ps at a constant volume and stable pressure of 1 bar. Pressure of the systems was maintained by using Parrinello-Rahman barostat [19]. Conservation of bond lengths was achieved using LINear Constraint Solver (LINCS) [29] algorithm. Long range interactions were computed using Particle Mesh Ewald (PME) [30]. The final MD run was set to 100,000 ps for all six systems. Trajectory analysis was performed using various GROMACS utilities such as gmx rms, gmx rmsf, gmx gyrate, gmx sasa, gmx hbond and gmx do_dssp for the analysis of rmsd, rmsf, radius of gyration, solvent accessible surface area, hydrogen bonds and secondary structure respectively. Principal component analysis (PCA) was performed to determine the correlated motions of proteins. For PCA analysis, gmx covar module of GROMACS was utilized for the generation and diagonalization of covariance matrix. Further gmx anaeig was used to determine the overlap between the principal components. Free energy surface (FES) was generated to show the energy of conformations adopted by the protein throughout the 100 ns MD simulation. gmx sham module was applied for the generation of FES [31].

Fig. 3. Superimposed docking pose, and interaction of test compounds at SARS-CoV-2 main protease active site (A) Docking pose of quercetin-3-rutinoside-7-glucoside (yellow), rutin (green), quercetin (red) and nelfinavir (blue). (B) Docking pose of sitoindoside IX (pale yellow), chlorogenic acid (lime green), iso-chlorogenic acid B (cyan), caffeoyl quinic acid (hotpink), 4-Deoxyphysalac-tone (orange) and nelfinavir (blue). (C) Withanolide F (lightblue), somnifera withanolide (grey), somniferawithanolide (purple), somniferanolide (wheat), withanosomiferanolide (firebrick), withanosomniferiolide (deapteal), somnifericin (chartreuse), withanolide Q (orange), 2,3-didehydrosomnifericin (chocolate), sominone (violet), withanolide N (forest), withanolide J (palecyan), (18R)-withaphysalin F (pink) and nelfinavir (blue) (D) Surface structure of Mpro protein interacted with lead compound quercetin-3-rutinoside-7-glucoside. (E) Surface structure of Mpro protein interacted with lead compound sitoindoside IX. Green color represents the mentioned amino acid involved in hydrogen bonding (G) Surface structure of Mpro protein interacted with quercetin-3-rutinoside-7-glucoside. (H) Surface structure of Mpro protein interacted with nelfinavir. Green color represents the amino acid involved in hydrogen bonding.
The binding free energies of the protein-ligand complexes of Mpro-lopinavir and Mpro-quercetin-3-rutinoside-7-glucoside complexes were estimated by Molecular Mechanics-Poisson-Boltzmann Solvent-Accessible surface area (MMPBSA) method using g-mmpbsa package [32]. g-mmpbsa is a widely used method to determine the free energy of biomolecular interactions. Binding free energy was determined by using following equations:

\[ \Delta G_{\text{bind}} = G_{\text{Prot}}(\text{complex}) + G_{\text{Lig}}(\text{complex}) - G_{\text{Prot}} + G_{\text{Lig}}(\text{free}) \]

Where \( \Delta G_{\text{bind}} \) signifies the energy of protein-ligand complex, \( G_{\text{Prot}}(\text{free}) \) and \( G_{\text{Lig}}(\text{free}) \) signify the individual energy of protein and ligand respectively. The MMPBSA calculation was performed for last 10 ns (90–100 ns) of each simulation trajectories.

### 3. Result and discussion

A total of 93 phytochemicals were extracted after comprehensive literature survey on *Withania somnifera*. These phytochemicals were further categorized on the basis of characteristic and non-characteristic presence in *W. somnifera* (Supplementary Table). Molecular docking screening of 93 phytochemicals against Mpro protein revealed 23 lead compounds exhibiting higher binding potential than standard/reference compounds. These test phytochemicals showed potential binding energy ranging from \(-6.16\) to \(-12.27\) kcal/mol at Mpro active site, compared to the standard protease inhibitors (nelfinavir, lopinavir, etc.) in the range of \(-2.55\) to \(-6.16\) kcal/mol (Supplementary Table 1). The name and structure of lead compounds (cutoff less than \(-6.16\) kcal/mol binding energy) are shown in Fig. 2. The ADME of the compounds used in the present study are provided in Table 1. Quercetin-3-rutinoside-7-glucoside and rutin showed low binding score (\(-12.27\) and \(-10.16\) kcal/mol respectively) against Mpro protein, compared to nelfinavir (\(-6.16\) kcal/mol). Type of amino acid residue involved in the protein-ligand interactions are summarized in Table 2. The binding energy of the characteristic and non-characteristic compounds were plotted in a grouped manner to compare the binding potential. The result showed that the *W. somnifera* non-characteristic compound have better potential than the characteristic phytochemicals in terms of binding energy at the active site of the SARS-CoV-2 Mpro protein (Fig. 1).

Binding pose of lead phytochemicals at SARS-CoV-2 Mpro active site was predicted and depicted (DockScore less than \(-6.16\) kcal/mol) in Fig. 3. The lead compounds were categorized into three major groups. First group comprised quercetin and its associated derivative structures such as rutin, and quercetin-3-rutinoside-7-glucoside. Second group consists of compounds reported in *W. somnifera* but also found in other plants i.e. non-characteristic compounds. The third group of compounds was *W. somnifera* specific i.e. characteristic compounds. Result showed that all the three groups of *W. somnifera* compounds as well as standard Mpro inhibitor binds in the same pose at the active site of the enzyme (Fig. 3A–C). Amino acids involved in the binding with quercetin-3-rutinoside-7-glucoside, caffeoyl quinic acid and withanolide J as well as standard inhibitor (nelfinavir and lopinavir) are depicted in Fig. 3D–H. Asn142, His41, Cys145, and Glu166 residue of Mpro were involved in hydrogen bond formation with standard inhibitors (nelfinavir and lopinavir). Besides several amino acid residues were involved in hydrophobic interaction with the protease (Table 2). Hydrogen bond length between the quercetin-3-rutinoside-7-glucoside and Thr25, Thr24, Ser46, Thr26, Asn142, Glu166, His164 residues were 2.74; 2.89; 3.00; 3.03 and 2.80; 3.22; 3.12 2.79 and 2.81; and 2.91 Å respectively (Supplementary Figure 1). Gly143, Leu27, Thr45, Met165, His41, Met49, Arg188, Leu167, and Pro168 amino acid residues Mpro protein...
interacted with Quercetin-3-rutinoside-7-glucoside by hydrophobic interaction (Table 2). Rutin showed hydrogen bonding with Thr25, Thr24, Gly143, His164, and Asn142 residue and interacted with Cys145 and Glu166 amino acids via hydrophobic interaction. Further an important compound caffeoyl quinic acid formed four hydrogen bonds (with Asn142, Gly143, His164, and Thr190) and several hydrophobic interactions at the active site of the protein (Supplementary Figure 1 and Table 2). The top three lead molecules (Quercetin-3-rutinoside-7-glucoside, Rutin, Caffeoyl quinic acid) showed H-bond interaction with Gly143, His164 and Asn142 amino acid residues. Besides the lead showed hydrophobic interaction with some of the common amino acids (Table 2). Rutin showed hydrogen bonding with Thr25, Asn142 amino acid residues (Table 2). Type of interaction, hydrogen bond length, amino acids involved in hydrogen bonding and hydrophobic interaction for the lead phytochemicals-M \( \text{pro} \)-active site in comparison to M\( \text{pro} \)-lopinavir complex binding. Further an important compound caffeoyl quinic acid formed four hydrogen bonds (with Asn142, Gly143, His164, and Thr190) and several hydrophobic interactions at the active site of the protein (Supplementary Figure 1 and Table 2). The top three lead molecules (Quercetin-3-rutinoside-7-glucoside, Rutin, Caffeoyl quinic acid) showed H-bond interaction with Gly143, His164 and Asn142 amino acid residues. Besides the lead showed hydrophobic interaction with some of the common amino acids (Supplementary Figure 1 and Table 2) at the active site of the protein.

\( W. \text{somnifera} \) characteristic compounds such as withanolide J \((-7.72\text{ kcal/mol})\) showed tight binding with M\( \text{pro} \) active site in comparison to niflumic acid \((-6.16\text{ kcal/mol})\). It should be noted that Rutin, Caffeoyl quinic acid, Withanolide J, Withanolide Q, Chlorogenic acid, Withanolide N, Somniferawithanolide, Somnifinanolide, Withasomniferanolid, 2,3-Didehydrosomnifericin, and Withasomnifilide showed similar type of interaction with Cys145, His41, Glu166 and Asn142 amino acid residues (Table 2). Type of interaction, hydrogen bond length, amino acids involved in hydrogen bonding and hydrophobic interaction for the lead phytochemicals-M\( \text{pro} \) complex binding are summarized in Table 2 and supplementary figure 1.

The three systems, M\( \text{pro} \)-protein (MP), M\( \text{pro} \)-lopinavir complex (MLC), and M\( \text{pro} \)-quercetin-3-rutinoside-7-glucoside complex (MQC) were subjected to the molecular dynamics (MD) simulations for 100 ns. Dynamic properties of the system such as temperature, pressure, density, and total energy were further calculated to ensure the equilibration and stability of the systems during the simulation period. Result showed that there were no significant alterations in the parameters of free protein and protein bound with standard/lead inhibitor during the simulation (Supplementary Figure 2). Root-mean-square deviation (RMSD) is widely used to get insights of the structural dynamics of proteins. For the assessment of the structural dynamics of M\( \text{pro} \) protein in free and complex states, we evaluated the RMSD of MP, MLC and MQC. The average RMSD of MP, MLC and MQC was \(~0.29\), \(~0.28\), and \(~0.30\) nm respectively (Fig. 4A). Backbone RMSD of all the three systems showed that M\( \text{pro} \) gets stabilized after ligand binding (Fig. 4A). RMSD of M\( \text{pro} \)-lead complexes were very close to the unbound protein. The stable RMSD showed that the ligand bound complex was equilibrated throughout the simulation trajectory and the binding was stable.

To determine the residual vibrations in the free and ligand bound M\( \text{pro} \) protein, the fluctuation of each residue was evaluated as root-mean-square fluctuation (RMSF). Random fluctuations present in M\( \text{pro} \) protein in bound and unbound states were evident from N- to C- termini of the protein (Fig. 4B). The fluctuation in the protein was compared for each residue after the binding of lopinavir and quercetin-3-rutinoside-7-glucoside. The fluctuations were found to be minimum at several residues in the case of MLC and MQC complexes. However there is a slight increase in RMSF of MQC which might be induced by the larger size of bound inhibitor. Another important parameter radius of gyration (Rg) is associated with the structural volume of the protein tertiary structure which indicates the structural compactness and folding behavior in the biological system. Elevated Rg value indicate loose packing of amino acid residues in protein. In the present study, the average values of Rg for MP, MLC and MQC were \(~2.24\), \(~2.22\), and \(~2.25\) nm, respectively (Fig. 4C). The Rg of test complex and unbound protein showed a minor deviation which suggests that the binding of compounds did not affected the compactness of protein structure. SASA is the surface area of proteins which is available for the surrounding water to interact with the protein. It is a fundamental parameter which provides insights into the structural folding-unfolding dynamics of a protein in the aqueous environment. SASA values for MP, MLC and MQC were \(~151.48\), \(~151.76\), and \(~148.93\) nm\(^2\), respectively (Fig. 4D).

---

**Fig. 5.** Hydrogen bond analyses of SARS-CoV-2 M\( \text{pro} \)-protein and protein-ligand complexes during 100 ns simulation. (A) Number of hydrogen bonds in the SARS-CoV2 M\( \text{pro} \)-protein in unbound, bound with lopinavir and quercetin-3-rutinoside-7-glucoside. (B) Hydrogen bond between M\( \text{pro} \)-protein and solvent in unbound state, bound with lopinavir and quercetin-3-rutinoside-7-glucoside. (C) Number of hydrogen bonds formed between M\( \text{pro} \) and lopinavir (D) Number of hydrogen bonds formed between M\( \text{pro} \)-protein and quercetin-3-rutinoside-7-glucoside. Unbound protein parameters are depicted in black color. Parameters for M\( \text{pro} \)-lopinavir and M\( \text{pro} \)-quercetin-3-rutinoside-7-glucoside complex are represented in red and green color respectively.
Intramolecular hydrogen bonding network of a protein is an essential parameter in the determination of its three dimensional structure and conformation. For the assessment of stability of Mpro protein in unbound and bound (with standard/lead inhibitor) conformations, number of intramolecular hydrogen bonds were calculated. MQC had an average of 226 hydrogen bonds in comparison to MLC and MP which had 219 and 215 hydrogen bonds (Fig. 5A). Based on this observation, it is evident that binding of quercetin-3-rutinoside-7-glucoside to Mpro protein stabilizes its internal conformation. Result showed no significant changes in hydrogen bonding network of Mpro protein in unbound and ligand (standard/lead inhibitor) bound state (Fig. 5A). Furthermore number of hydrogen bonds in Mpro protein remained similar throughout the trajectory which also suggests minimal fluctuation in the protein structure. Similarly the hydrogen bonding interaction of the complexes (MP, MLC, and MQC) with surrounding solvent showed stable pattern (5B). MQC had an average of 510 hydrogen bonds with surrounding solvent in comparison to the 527 and 546 hydrogen bonds formed by MLC and MP respectively. Assessment of the hydrogen bonds between the protein and ligand provides direction and specificity of the interaction which are important criterion in protein-ligand binding. To get an insight of stability of binding among MLC and MQC, hydrogen bonds between protein and inhibitor were calculated. Binding of quercetin-3-rutinoside-7-glucoside to Mpro protein encompasses significantly higher number of hydrogen bonds (≈6) in comparison to lopinavir binding (≈1.5) (Fig. 5C and D). These results indicate the stable binding of quercetin-3-rutinoside-7-glucoside to the Mpro protein and was comparable to lopinavir binding.

Protein exerts its function by collective motions of atoms. Principle component analysis (PCA) was performed to understand the conformational dynamics of the MP, MLC and MQC by utilizing their collective motions using essential dynamics method. The conformational sampling along the eigenvector 1 and eigenvector 2 was projected by backbone atoms of Mpro protein in unbound and standard/lead inhibitor bound complex (Fig. 6A–D). The results showed MQC possess less collective motions in comparison to MP and MLC suggesting stable conformation.

To visualize of minimum energy state of the MP, MLC, and MQC,
next we studied the free energy surface (FES) against the PC1 (Radius of gyration, Rg) and PC2 (Root mean square deviation, RMSD) principal components (Fig. 7 A–C). The size and shape of the minimum energy area (shown in blue color) illustrated the stability of the protein in unbound and ligand bound conformation (Fig. 7 A–C). FES plot of MQC was more compact and had dense energy minima in comparison to the MP and MLC indicating the stability of the complex.

The percentage of secondary structure content in a protein is an important parameter in the study of its structural behavior. There were no significant change in the secondary structure content of MP, MLC and MQC (Fig. 8 A–C). However binding of lead phytochemical to the Mpro protein appears to result in a modest increase in the residue number in structured region. The involvement of more amino residues at the binding site might be responsible for the stability of the ligand bound conformations of targeted protein, compared to unbound conformation (Fig. 8 C).

Protein-ligand interaction energy was calculated to find out the ligand interaction and stability of the complexes (MLC and MQC). Interaction energy between protein and inhibitor mainly consist of two short range interactions known as Coulombic short range interaction energy (CsrIE) and Lennard Jones short range interaction energy (LJsrIE) [23]. Interaction energy plot of MLC and MQC clearly indicate that the high interaction energy for the lead phytochemical bound Mpro protein complex in comparison to Mpro-lopinavir complex (Fig. 9 A and B). The CsrIE and LJsrIE for MLC and MQC were ~-47.92 kJ/mol and ~-237.25 kJ/mol and ~-212.26 and ~-149.32 kJ/mol respectively.

The last 10 ns (90–100 ns) of the trajectory obtained from 100 ns MD simulations of the ligand bound Mpro were utilized to estimate the binding free energy of the MLC and MQC. Binding free energy (Δ\(E_{\text{binding}}\)), van der Waals energy (\(E_{\text{vdw}}\)), electrostatic energy (\(E_{\text{elec}}\),

![Fig. 8.](image)

**Fig. 8.** DSSP analysis for the secondary structure fluctuations as a function of time from 0 to 100 ns for SARS-CoV-2 Mpro protein in unbound and ligand bound state at 300 K (A) SARS-CoV-2 Mpro unbound (B) SARS-CoV-2 Mpro complexed with lopinavir (C) SARS-CoV-2 Mpro complexed with quercetin-3-rutinoside-7-glucoside.

![Fig. 9.](image)

**Fig. 9.** Short range interaction analysis of Mpro-lopinavir and Mpro-quercetin-3-rutinoside-7-glucoside complex. (A) Short range Coulombic interaction energy of Mpro-lopinavir and Mpro-quercetin-3-rutinoside-7-glucoside complex. (B) Short range Lennard Jones interaction energy of Mpro-lopinavir and Mpro-quercetin-3-rutinoside-7-glucoside complex. Parameters for Mpro-lopinavir and Mpro-quercetin-3-rutinoside-7-glucoside complex are depicted in black and red color respectively. CsrIE = coulombic short range interaction energy; LJsrIE = lennard jones short range interaction energy.

| Table 3 | MM-PBSA calculations of binding free energy for Mpro-lopinavir and Mpro-quercetin-3-rutinoside-7-glucoside complex. |
|---------|------------------------------------------------------------------------------------------------------------------|
| Type of Binding energy | Binding energy values (Mpro-Lopinavir complex) | Binding energy values (Mpro-quercetin-3-rutinoside-7-glucoside) |
| \(\Delta E_{\text{binding}}\) (kJ/mol) | \(-195.525 \pm 0.481\) | \(-67.767 \pm 0.536\) |
| SASA (kJ/mol) | \(-25.823 \pm 0.052\) | \(-22.548 \pm 0.044\) |
| \(\Delta E_{\text{polar solvation}}\) (kJ/mol) | \(-31.008 \pm 0.288\) | \(-89.499 \pm 0.892\) |
| \(\Delta E_{\text{van der Waals}}\) (kJ/mol) | \(-279.994 \pm 0.485\) | \(-191.207 \pm 0.538\) |

![Diagram](image)
polar solvation energy ($\Delta E_{\text{polar}}$), and SASA were estimated (Supplementary figure 3, Table 3 and Fig. 10A). Binding energy for both the complexes was stable throughout the analyzed simulation trajectory (Supplementary figure 3). Further contribution of each residue was evaluated by decomposing the total binding free energy into per residue contribution energy (Fig. 10B and C). In binding of lopinavir to M\(^{\text{pro}}\) protein amino acid residues Thr25, Leu27, Cys44, Met49, Cys145, His164, and Glu189 contributed significantly in binding. Furthermore in binding of quercetin-3-rutinoside-7-glucoside to the M\(^{\text{pro}}\) protein amino acid residues Leu27, Thr45, Leu141, Gly143, Cys145, Leu167, and Pro168 contributed significantly in binding (Supplementary table 2).

The M\(^{\text{pro}}\) protein has been considered as one of the best drugable target for the anti-SARS-CoV-2 drug discovery. The protease activity of the protein is required to process viral protein in the host and thus aid in the viral replication and virulence. Our study revealed the potential binding of the \emph{W. somnifera} phytochemicals at the active site of the target SARS-CoV-2 M\(^{\text{pro}}\) protein in comparison to the known standard inhibitor (Supplementary Table 1). The lead phytochemicals showed better interaction with the targeted protein which was revealed by hydrogen bond formation, hydrophobic interaction and docking pose of the compounds (Supplementary Figure 1 and Table 2). Interaction of the \emph{W. somnifera} phytochemicals with the key amino acids (Asn142, Gly143, Cys145, and His164 amino acid residues) of the protein indicates that they have ability to lock the active site of the targeted protein. Further, the MD simulations results including (RMSD, RMSF, Rg, SASA and hydrogen bonding pattern) showed the dynamically stable and potential binding of lead \emph{W. somnifera} phytochemical in comparison to the standard M\(^{\text{pro}}\) protein inhibitor (Figs. 4 and 5). Binding free energy calculations of the lead-protein complex showed that the amino residue such as Cys145 contributes significantly in the overall binding energy of the complex (Fig. 10). Taken together, the results of the present study indicates the \emph{W. somnifera} phytochemicals possess SARS-CoV-2 M\(^{\text{pro}}\) protein inhibition potential and could serve as potential source of the lead molecules for anti-SARS-CoV-2 drug discovery. Further, the \textit{in vitro} and \textit{in vivo} experimentations are required to assess the viral replication and virulence by inhibiting the M\(^{\text{pro}}\) protein by \emph{W. somnifera} phytochemicals.

4. Conclusion

Natural products are composed of several bioactive constituents and are recognized as multiple targeting agents. This property of the natural compounds makes them better candidates for novel drug discovery. In the present study, we identified the non-characteristic and characteristic phytochemicals from \emph{W. somnifera} which possess differential potential against SARS-COV-2 M\(^{\text{pro}}\) protein in molecular docking studies. The 100ns molecular dynamic and MMPBSA analysis establishes the energetically and structurally stable complex formation ability of the lead phytochemicals with the SARS-CoV-2 M\(^{\text{pro}}\) protein. Overall, we conclude that non-characteristic compound of \emph{W. somnifera} either alone or in combination with the characteristic phytochemicals has ability to target M\(^{\text{pro}}\) and might be useful in the future anti-SARS-CoV-2 drug discovery.

Declaration of competing interest

The authors declare no conflict of interest.

Acknowledgement

PPK and MS acknowledge financial support from Indian Council of Medical Research (ICMR), India in the form of ICMR-Senior Research fellowships (File No. 5/3/8/82/ITR-F/2020 and File No. 5/3/8/80/ITR-F/2020-ITR respectively). SK acknowledges University Grants Commission, India and Department of Science and Technology, India for providing financial support in the form of UGC-BSR Research Start-Up Grant [No. F.30-372/2017 (BSR)] and DST-SEBB Grant [EEQ/2016/000350] respectively. SK acknowledges Central University of Punjab, Bathinda, India for providing Research Seed Money Grant [GP-25]. AKS
and KSP acknowledge CSIR-India and DBT-India funding agencies respectively for providing financial assistance in the form of Junior Research Fellowship. SK also acknowledges DST-FIST India, for providing financial support to the Department of Biochemistry, Central University of Punjab, Bathinda, India.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.micpath.2021.104954.

References

[1] M.T. ul Qamar, S.M. Alqbatani, M.A. Alamri, L.L. Chen, Structural basis of SARS-CoV-2 3CLpro and anti-COVID-19 drug discovery from medicinal plants, J. Pharm. Anal. 10 (4) (2020) 313–319.
[2] M. Prajapati, P. Sarma, N. Shekhawat, P. Arvi, S. Sinha, H. Kaur, S. Kumar, N. Arnold, A. Bhattacharyya, H. Kumar, S. Bansal, P. Avti, S. Sinha, H. Kaur, S. Kumar, Identification of potential natural inhibitors of SARS-CoV-2 main protease by molecular docking and simulation studies, J. Biomol. Struct. Dyn. (1–12) (2020), 2020.
[3] P.P. Kushwaha, P.S. Vardhan, P. Kumari, A.G. Mtewa, S. Kumar, Bioactive lead compounds and targets for the development of antimalarial drugs, in: Proceedings of the 105th Indian Science Congress, Molecular Biology techniques, Molecules 24 (22) (2019) 1–22.
[4] V. Singh, B. Singh, A. Sharma, K. Kaur, A.P. Gupta, R.K. Salar, V. Hallan, P.K. Pati, Identification of FDA approved drugs and nucleoside analogues as potential SARS-CoV-2 Mpro inhibitors: in-silico drug discovery, J. Korean Med. Sci. 31 (2021) 205–426.
[5] R. A. Laskowski, M.B. Swindells, LigPopt: multiple ligand-protein interaction diagrams for drug discovery, J. Chem. Inf. Model. 51 (2011) 2778–2786.
[6] B. Hess, H. Bekker, H.J. Berendsen, J.G. Fraaije, LINCS: a linear constraint solver for molecular simulations, J. Comput. Chem. 18 (12) (1997) 1463–1472.
[7] S. Darden, T. York, L. Pedersen, Particle mesh Ewald: an N congery of potential SARS-CoV-2 Mpro inhibitors: in silico study, Comp. boil. Med. 130 (1–2021).
[8] A. Bhattacharyya, H. Kumar, S. Bansal, B. Medhi, Drug targets for corona virus: a molecular dynamics puzzled out potential SARS-CoV-2 A1pp domain inhibitor: an in silico study, Molecules 24 (2021) 1–22.
[9] R. Kumar, R. Kumar, Open Source Drug Discovery Consortium and Lyma, A., g, mmpbsa: a GROMACS tool for high-throughput MM-PBSA calculations, J. Chem. Inf. Model. 54 (7) (2014) 1951–1961.
[10] M. Ali, M. Shuaib, S.H. Ansari, Withanolides from the stem bark of Withania somnifera, Phytochemistry 44 (6) (1997) 1163–1168.
[11] J.A. Duke, Database of Phytochemical Constituents of GRAS Herbs and Other Economic Plants, CRC Press, 1992.
[12] S. Shibuyama, A. Gevorkyan, U. Yoo, S. Tim, K. Dzhangiryan, J.D. Scott, Crystallogr. D Biol. Crystallogr. 58 (6) (2002) 899–907.
[13] P.P. Kushwaha, P.S. Vardhan, P. Kumari, A.G. Mtewa, S. Kumar, Bioactive lead compounds and targets for the development of antimalarial drugs, in: Proceedings of the 105th Indian Science Congress, Molecular Biology techniques, Molecules 24 (22) (2019) 1–22.
[14] R. A. Laskowski, M.B. Swindells, LigPopt: multiple ligand-protein interaction diagrams for drug discovery, J. Chem. Inf. Model. 51 (2011) 2778–2786.
[15] B. Hess, H. Bekker, H.J. Berendsen, J.G. Fraaije, LINCS: a linear constraint solver for molecular simulations, J. Comput. Chem. 18 (12) (1997) 1463–1472.
[16] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an N congery of potential SARS-CoV-2 Mpro inhibitors: in silico study, Molecules 24 (2021) 1–22.
[17] N. Poolsup, N. Suksomboon, P.D.M. Kurnianta, K. Deawjaroen, Effects of curcumin on glycemic control and lipid profile in prediabetes and type 2 diabetes mellitus: a systematic review and meta-analysis, PLoS One 14 (4) (2019) 1–18.
[18] A. Tiwari, P.P. Kushwaha, S. Kumar, Future scope for Withania somnifera in anticancer drug discovery, in: Proceedings of the 105th Indian Science Congress, Section of New Biology (Including Biochemistry, Biophysics & Molecular Biology and Biotechnology) at, vol. 180, Manipur University, Imphal, 2018, 3-7 January, 2018.
[19] M. Farrinello, A. Rahman, Polymeric transitions in single crystals: a new molecular dynamics method, J. Appl. Phys. 52 (12) (1981) 7182–7190.
[20] Q.X.S. Xie, Exploiting PubChem for virtual screening, Expt Opin. Drug Discov. 5 (12) (2010) 1205–1220.
[21] N.M. O’Boyle, M. Banck, C.A. James, C. Morley, T. Vandermeersch, G. R. Hutchison, Open Babel: an open chemical toolbox, J. Cheminf. 3 (1) (2011) 1–14.
[22] S. Dallakyan, PyRx-python Prescription V. 0.8, The Scripps Research Institute, 2010.
[23] A.M. Londhe, C.G. Gadhe, S.M. Lim, A.N. Pae, Investigation of molecular details of Keap1-Nrf2 inhibitors using molecular dynamics and umbrella sampling techniques, Molecules 24 (22) (2019) 1–22.