Identification of natural compounds as potent inhibitors of SARS-CoV-2 main protease using combined docking and molecular dynamics simulations

Deeba Shamim Jairajpuria, Afzal Hussian, Khalida Nasreen, Taj Mohammad, Farah Anjum, Md. Tabish Rehman, Gulam Mustafa Hasan, Mohamed F. Alajmi, Md. Imtaiyaz Hassan

A R T I C L E   I N F O
Article history:
Received 16 November 2020
Revised 17 January 2021
Accepted 19 January 2021
Available online 27 January 2021

Keywords:
SARS-CoV-2 main protease
Natural compounds
Virtual high-throughput screening
Molecular dynamics simulation
Small molecule inhibitors

A B S T R A C T
Coronavirus disease 2019 (COVID-19) has emerged from China and globally affected the entire population through the human-to-human transmission of a newly emerged virus called severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The genome of SARS-CoV-2 encodes several proteins that are essential for multiplication and pathogenesis. The main protease (Mpro or 3CL pro ) of SARS-CoV-2 plays a central role in its pathogenesis and thus is considered as an attractive drug target for the drug design and development of small-molecule inhibitors. We have employed an extensive structure-based high-throughput virtual screening to discover potential natural compounds from the ZINC database which could inhibit the Mpro of SARS-CoV-2. Initially, the hits were selected on the basis of their physicochemical and drug-like properties. Subsequently, the PAINS filter, estimation of binding affinities using molecular docking, and interaction analyses were performed to find safe and potential inhibitors of SARS-CoV-2 Mpro. We have identified ZINC02123811 (1-(3-(2,5,9-trimethyl-7-oxo-3-phenyl-7H-furo[3,2-g]chromen-6-yl)propanoyl)piperidine-4-carboxamide), a natural compound bearing appreciable affinity, efficiency, and specificity towards the binding pocket of SARS-CoV-2 Mpro. The identified compound showed a set of drug-like properties and preferentially binds to the active site of SARS-CoV-2 Mpro. All-atom molecular dynamics (MD) simulations were performed to evaluate the conformational dynamics, stability and interaction mechanism of Mpro with ZINC02123811. MD simulation results indicated that Mpro with ZINC02123811 forms a stable complex throughout the trajectory of 100 ns. These findings suggest that ZINC02123811 may be further exploited as a promising scaffold for the development of potential inhibitors of SARS-CoV-2 Mpro to address COVID-19.

1. Introduction
A newly emerged pneumonia outbreak of coronavirus disease 2019 (COVID-19) spread through severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (Asrani et al., 2020b; Huang et al., 2020), now globally affected over 100 million individuals and accounts for over two million deaths worldwide (https://www.worldometers.info/coronavirus/). There are several diagnostic and therapeutic approaches have emerged to handle COVID-19; however, still, no effective therapy has been developed (Asrani et al., 2020a; Graham, 2020). The use of Lopinavir and Ritonavir (Nutho et al., 2020), antiviral drugs (Grein et al., 2020),
chloroquine and hydroxychloroquine (Jakhar and Kaur, 2020), convalescent plasma (Chen et al., 2020), stem cell therapy, etc., have shown some positive outcomes on admitted patients of COVID-19 (Fatima et al., 2020; Kumari et al., 2020). Hence, there is an emergent need to discover potential therapeutic agents to control the pathogenesis of SARS-CoV-2 (Corey et al., 2020).

The genome profiling of SARS-CoV-2 leads to the identification of a few drug targets including, the SARS-CoV-2 main protease (M\textsuperscript{pro}, 3CL\textsuperscript{pro}), that mediates viral replication and transcription together and can be considered as a potential target for therapeutic development (Naqvi et al., 2020; Zhang and Holmes, 2020). Due to its critical role in pathogenesis, combined with the absence of closely related homologues in humans, M\textsuperscript{pro} is acting as an effective target for drug design and development (Zhang et al., 2020). M\textsuperscript{pro} cleaves polyproteins to generate Non-Structural Proteins (NSPs) that form a replicase-transcriptase complex (RTC). M\textsuperscript{pro} exists in NSP5 which releases the majority of NSPs from the polyproteins, is vital for the life cycle of SARS-CoV-2 (V’kovski et al., 2020). The structure of M\textsuperscript{pro} advances the field of modern drug discovery which results in the development of potential lead molecules that block M\textsuperscript{pro} function in cell-based assays (Zhang et al., 2020). SARS-CoV-2 M\textsuperscript{pro} is consists of 306 amino acid residues with a cysteine-histidine catalytic dyad including Cys145 and His41 (Dai et al., 2020; Jin et al., 2020). Structural analysis suggests that these residues including a few others found in the active site pocket of SARS-CoV-2 M\textsuperscript{pro} can be served as a platform for the discovery of its selective inhibitors in the therapeutic management of COVID-19 (Zhang et al., 2020).

In many attempts of ongoing research, new leads are being identified by utilizing advanced computational approaches to screen large chemical libraries (Jin et al., 2020; Shamsi et al., 2020). For the quick development of effective therapeutics, the experimental screening approach alone may not improve lead productivity (Padhi and Tripathi, 2020). Bioinformatics and computational biology play a crucial role in the drug discovery process while employing a structure-based drug design approach such as molecular docking-based virtual high-throughput screening (vHTS) which is currently largely implemented in the modern drug-discovery pipeline to find potential lead molecules among various chemical libraries (Mohammad et al., 2020b; Naqvi et al., 2018; Naqvi and Hassan, 2017). There are many chemical reposito-
from PDB (PDB ID: 6LU7) (Jin et al., 2020). All co-crystallized hetero molecules including N3 inhibitor were removed from the parent structure. Finally, the protein structure was prepared for vHTS in MGL tools by adding hydrogen to polar atoms and assigning appropriate atom types. A library of ~90,000 natural products was downloaded from the ZINC database in processed form.

2.2. Filtration of compounds

All compounds from the ZINC library were filtered out based on their physicochemical and ADMET properties through SwissADME and Discovery Studio Visualizer. First, we have selected compounds based on their physicochemical and drug-like properties satisfying the Lipinski’s rule of five (Lipinski, 2000). We further applied Pan-assay interference compounds (PAINS) filter to avoid compounds with specific patterns with a higher tendency to bind towards multiple biological targets. We further screened the compounds for their carcinogenic patterns and ADMET (Absorption, Distribution, Metabolism, Excretion, and Toxicity) properties. Here, only compounds having well ADMET properties and non-carcinogenic patterns were selected for further docking-based vHTS studies.

2.3. Molecular docking-based vHTS

We performed a molecular docking-based vHTS to filter compounds based on their binding affinities towards SARSCoV-2 M\textsuperscript{pro}. The docking was performed using AutoDock Vina with a grid size of a blind search space as 54, 68 and 62 Å, centralized at −26.34, 12.60 and 58.91 for X, Y and Z coordinates, respectively. The grid spacing was set to 1.00 Å with default docking parameters. The docking results were screened for high binding affinity and then all possible docked conformations were split for each compound which were further analyzed using PyMOL and Discovery Studio Visualizer for their possible interaction towards Mpro. The pounds which were further analyzed using PyMOL and Discovery Studio Visualizer for their possible interaction towards M\textsuperscript{pro}. The polar contacts depicted in PyMOL within 3.5 Å of distance were mapped as close interactions between the compounds and M\textsuperscript{pro}. Discovery Studio Visualizer was utilized to explore detailed interactions formed between the selected compounds with the SARSCoV-2 M\textsuperscript{pro}. Interaction analysis resulted in the identification of three compounds interacting with the critical residues of the binding pocket of SARS-CoV-2 M\textsuperscript{pro}. Here, we have selected only those compounds which were specifically interacting with the binding-site residues including Cys145 of SARS-CoV-2 M\textsuperscript{pro}.

ZINC02123811 were generated from the PRODRG server and merged to the protein topology to make the M\textsuperscript{pro}-ZINC02123811 complex. Both, M\textsuperscript{pro} and M\textsuperscript{pro}-ZINC02123811 systems were solvated in a cubic box with the SPC (spc216) water model for aqueous simulations. Energy minimization for 1500 steps of the steepest descent method was performed for 1000 ps. The temperature of both systems was consequently increased from 0 to 300 K during the equilibration period. The equilibration period was performed for 100 ps at constant volume under periodic boundary conditions with a stable pressure of 1 bar. The final MD run was performed for 100 ns for both systems, and resulting trajectories were analyzed using the inbuilt utilities of GROMACS as described (Gulzar et al., 2019; Khan et al., 2016; Khan et al., 2017b; Shahbaaz et al., 2018).

2.6. Principal component and free energy landscape analyses

To explore the conformational sampling, atomic motions and structural stability of M\textsuperscript{pro} and M\textsuperscript{pro}-ZINC02123811 complex, principal component (PC) and free energy landscape (FEL) studies were performed by the essential dynamics approach which employs the calculation of the covariance matrix (Altis et al., 2008; Fatima et al., 2019; Gupta et al., 2019; Mohammad et al., 2019b). The following formula was used to calculate the covariance matrix:

$$C_{ij} = <(x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle)>$$

where $x_i/x_j$ is the coordinate of the $i^{th}/j^{th}$ atom of the system, and $\langle \cdot \rangle$ is the ensemble average.

The FELs of M\textsuperscript{pro} and M\textsuperscript{pro}-ZINC02123811 complex were attained using the conformational sampling approach which allows exploring the protein conformations near the native state (Papaleo et al., 2009). FELs were generated to investigate the stability and native states of SARS-CoV-2 M\textsuperscript{pro}, before and after ZINC02123811 binding. The FELs were generated utilizing the following formula:

$$\Delta G(X) = -K_B T \ln P(X)$$

where $K_B$ is the Boltzmann constant, $T$ is the temperature of simulation, and $P(X)$ is the probability distribution of the system along with the PCs.

3. Results

3.1. Filtration of compounds

The physicochemical properties of all the natural compounds present in the library were calculated and analyzed through the SwissADME webservice and Discovery Studio Visualizer. Here we identified a set of 32,902 compounds based on Lipinski’s rule of five (Mol wt. ≤ 500 Da, log P ≤ 5, H-bond donor ≤ 5 and H- bond acceptor ≤ 10), bioavailability score, and PAINS pattern (Lipinski, 2000). The physicochemical properties of the finally selected three compounds satisfying the rule of five, along with the standard SARS-CoV-2 M\textsuperscript{pro} inhibitor N3 are shown in Table 1.

The compounds were also screened out based on their ADMET properties and carcinogenic patterns to identify safe and non-carcinogenic compounds. ADMET properties of the finally selected three compounds and N3 are given in Table 2.

3.2. Molecular docking-based vHTS

In this attempt of molecular docking-based vHTS, the compounds having appreciable binding affinities with SARS-CoV-2 M\textsuperscript{pro} were selected for further analysis. The vHTS analysis results in the identification of 10 compounds from a pool of 32,902 natural
compounds showing a considerable binding score (−9.4 to −9.8 kcal/mol) towards M\textsubscript{pro} (Table 3). We further calculated their mean affinity after performing redocking up to 10 different runs of AutoDock Vina with independent random seeds and found good consistency in the resultant output (Table 3). Further, detailed interaction analysis of the top 10 hits was carried out using the PyMOL and Discovery Studio Visualizer. Here, a total of 90 possible docked conformers were splitted from the out files of the selected hits. From the analysis of all possible docked conformers, we identified three natural compounds that have commonly interacted with the active-site ‘Cys145’ of SARS-CoV-2 M\textsubscript{pro}. The binding pattern of the finally selected three compounds with SARS-CoV-2 M\textsubscript{pro} is illustrating in Fig. 2. Compounds docked to the binding pocket of SARS-CoV-2 M\textsubscript{pro} were checked for their interaction with the functionally important residues of the protein. Detailed interaction of the three selected compounds is shown in Fig. 3 where it is evident that all the selected compounds interact with Cys145 of SARS-CoV-2 M\textsubscript{pro}. The docking poses of the identified compounds and N3 suggests that they could fit inside the substrate-binding pocket of M\textsubscript{pro}.

### Table 1
List of identified compounds and their physiochemical properties.

| S. No. | Compound ID   | Mol wt. (Da) | Rotatable bond | H-bond acceptor | H-bond donor | LogP | Lipinski Violation |
|--------|---------------|--------------|----------------|-----------------|--------------|------|-------------------|
| 1      | ZINC02123811  | 486.56       | 5              | 5               | 1            | 4.78 | 0                 |
| 2      | ZINC02128147  | 488.49       | 6              | 7               | 3            | 4.01 | 0                 |
| 3      | ZINC02161101  | 472.60       | 3              | 3               | 2            | 3.98 | 0                 |
| 4      | N3-ILP        | 680.79       | 17             | 9               | 5            | 2.08 | 2                 |

### Table 2
List of identified compounds and their ADMET properties.

| Compound ID   | Absorption | GI absorption (%) | Water Solubility | Distribution | Metabolism | Excretion | Toxicity | BBB/CNS permeation | CYP2D6Inh/Subs | OCT2 substrate | AMES/ skin sens |
|---------------|------------|-------------------|------------------|--------------|------------|-----------|----------|---------------------|----------------|----------------|-----------------|
| ZINC02123811  | Soluble    | 99.08             | No               | No           | No         | No        | No       | No                  |                |                |                 |
| ZINC02128147  | Soluble    | 69.67             | No               | No           | No         | No        | No       | No                  |                |                |                 |
| ZINC02161101  | Soluble    | 90.69             | No               | No           | No         | No        | No       | No                  |                |                |                 |
| N3-ILP        | Soluble    | 57.88             | No               | No           | No         | No        | No       | No                  |                |                |                 |

### Table 3
Binding affinities of the selected compounds in 10 different runs of AutoDock Vina with independent random seeds. R1, R2, R3, ... R10 shows replicates of AutoDock Vina run.

| S. No. | Compound ID   | Affinity (kcal/mol) |
|--------|---------------|---------------------|
|        | R1            | R2                  | R3                  | R4                  | R5                  | R6                  | R7                  | R8                  | R9                  | R10                 | Mean                |
| 1      | ZINC02161101  | −9.8                | −9.4                | −8.9                | −9.4                | −9.9                | −9.8                | −9.4                | −9.5                | −9.9                | −9.4                | −9.5                |
| 2      | ZINC02113993  | −9.8                | −9.8                | −9.8                | −9.2                | −9.8                | −9.9                | −8.9                | −8.7                | −8.8                | −9.5                | −9.4                |
| 3      | ZINC02123811  | −9.7                | −9.8                | −9.6                | −9.6                | −9.7                | −9.4                | −9.6                | −9.7                | −9.7                | −9.6                | −9.6                |
| 4      | ZINC02125386  | −9.6                | −9.6                | −8.4                | −9.5                | −8.2                | −9.6                | −9.6                | −9.6                | −9.6                | −9.7                | −9.3                |
| 5      | ZINC02113878  | −9.6                | −9.5                | −9.6                | −9.6                | −9.6                | −9.6                | −9.6                | −9.5                | −9.5                | −9.5                | −9.6                |
| 6      | ZINC02111016  | −9.5                | −8.9                | −9.5                | −9.5                | −9.5                | −9.4                | −9.5                | −9.5                | −9.5                | −9.5                | −9.4                |
| 7      | ZINC02123668  | −9.5                | −8.6                | −9.5                | −9.5                | −9.4                | −9.5                | −8.5                | −9.5                | −9.5                | −9.5                | −9.3                |
| 8      | ZINC02128147  | −9.5                | −9.5                | −9.5                | −8.2                | −9.5                | −9.5                | −9.5                | −9.5                | −9.5                | −9.5                | −9.3                |
| 9      | ZINC02111094  | −9.4                | −9.1                | −9.4                | −9.2                | −9.3                | −9.1                | −9.3                | −9.4                | −9.4                | −9.1                | −9.3                |
| 10     | ZINC02112091  | −9.4                | −9.1                | −8.4                | −9.4                | −8.4                | −9.3                | −9.4                | −8.2                | −9.5                | −8.4                | −9.0                |

Fig. 2. Structural representation of docked compounds in the binding pocket of SARS-CoV-2 M\textsubscript{pro}. (A) Cartoon representation of M\textsubscript{pro} with the selected three compounds along with co-crystallized inhibitor N3. (B) Surface potential view of M\textsubscript{pro} binding pocket occupied by the selected compounds and N3.
The compounds binding to critical residues of the substrate-binding pocket including Cys145, Met165, and Glu166 may hinder the substrate accessibility to Mpro thus its inhibition.

3.3. PASS analysis: biological activity predictions

The exploration of biological activities of the selected compounds through the PASS analysis resulted in similar kinds of biological activities. The reference compound, N3 showed to have SARS-CoV-2 Mpro inhibitory potential, validating the results predicted. The compounds ZINC02123811, ZINC02128147 and ZINC02161101 have shown predictions for antithrombotic, antineoplastic, and antiviral potential, with Pa ranging from 0.340 to 0.552 when Pa > Pi. Table 4 shows the biological properties of all three compounds with higher Pa. Here, based on the specific interactions with SARS-CoV-2 Mpro and biological properties showing antiviral potential. Finally, we have selected one compound for MD simulation studies.

3.4. MD simulations

The finally selected compound, ZINC02123811 in complex with SARS-CoV-2 Mpro along with the free state of SARS-CoV-2 Mpro were subjected to all-atom MD simulations for 100 ns. The various systematic and structural parameters were calculated to analyze the stability and dynamics of SARS-CoV-2 Mpro before and after ZINC02123811 binding.

3.4.1. Structural deviations and compactness

To investigate the structural deviations and dynamics of a protein structure, root-mean-square deviation (RMSD) has been utilized (Dahiya et al., 2019; Kuzmanic and Zagrovic, 2010). We also

![Fig. 3. 2D structural representation of SARS-CoV-2 Mpro residues interacting to the compound (A) ZINC02123811, (B) ZINC02128147, and (C) ZINC02161101.](image)

| S. No. | Compound ID   | Pa   | Pi    | Biological Activity                                      |
|--------|---------------|------|-------|----------------------------------------------------------|
| 1      | ZINC02123811  | 0.552| 0.101 | CDP-glycerol glycerophosphotransferase inhibitor         |
|        |               | 0.406| 0.051 | Antithrombotic                                           |
|        |               | 0.340| 0.053 |                                                          |
| 2      | ZINC02128147  | 0.543| 0.105 | CDP-glycerol glycerophosphotransferase inhibitor         |
|        |               | 0.494| 0.113 | Phosphatase inhibitor                                    |
|        |               | 0.354| 0.073 | Antithrombotic                                           |
| 3      | ZINC02161101  | 0.490| 0.080 | Nicotinic alpha2beta2 receptor antagonist                |
|        |               | 0.410| 0.099 | Antineoplastic                                           |
|        |               | 0.385| 0.167 | Nicotinic alpha4beta4 receptor agonist                   |
| 4      | N3-ILP        | 0.477| 0.003 | Antiviral, SARS-CoV-2 Mpro inhibitor                     |
calculated the time-evolution of RMSDs for M\(^{\text{pro}}\) and M\(^{\text{pro}}\)-ZINC02123811 complex during the simulations and found average value as 0.22 nm and 0.25 nm, respectively. The RMSD plot shows both free M\(^{\text{pro}}\) and M\(^{\text{pro}}\)-ZINC02123811 complex are stable throughout the simulation (Fig. 4A). However, a minor increase in random fluctuations of up to 0.15 nm is observed between 0 and 40 ns in the case of complexed M\(^{\text{pro}}\). But, after 40 ns, the plot is showing stable and equilibrated RMSD throughout the simulation. Variations in RMSD of the M\(^{\text{pro}}\)-ZINC02123811 complex reduces after 75 ns and stabilized throughout the trajectory as compared with free M\(^{\text{pro}}\).

To explore the residual flexibility in M\(^{\text{pro}}\) in the free-state and upon ZINC02123811 binding, the average fluctuation of all residues was considered and plotted as root-mean-square fluctuation (RMSF). The RMSF plot showed several residual fluctuations in M\(^{\text{pro}}\) in different regions. These fluctuations were found to be stable and minimized upon ZINC02123811 binding with the progression of simulation at region spanning from N- to C- termini (Fig. 4B).

The conformational stability of M\(^{\text{pro}}\) before and after ZINC02123811 binding was also evaluated by calculating the radius of gyration (\(R_{g}\)) of both systems. The average \(R_{g}\) values for M\(^{\text{pro}}\) before and ZINC02123811 binding were estimated to be the same as calculated 2.19 nm. \(R_{g}\) plot suggested no major changes in the packing of M\(^{\text{pro}}\) when bound with ZINC02123811. An initial fluctuation until 15 ns of MD trajectories might occur due to packing adjustment of M\(^{\text{pro}}\), but thereafter, the \(R_{g}\) became stable and equilibrated throughout the simulation suggesting complex stability (Fig. 4C).

The average solvent-accessible surface area (SASA) values for M\(^{\text{pro}}\) and M\(^{\text{pro}}\)-ZINC02123811 complexes were found to be 148.38 nm\(^2\), and 147.46 nm\(^2\), respectively. The SASA plot is showed to have a similar pattern of equilibration in the case of both the systems. A slight decrease in the average SASA might be owing to tighter packing of the M\(^{\text{pro}}\) upon ZINC02123811 binding (Fig. 4D).

3.4.2. Dynamics of hydrogen bonds

To further examine the stability of M\(^{\text{pro}}\) before and after ZINC02123811 binding, the time-evolution of hydrogen bonds (H-bonds) formed within 0.35 nm during the simulation was explored. In M\(^{\text{pro}}\), the average number of intramolecular H-bonds before and after ZINC02123811 binding was estimated to be 216 and 214, respectively (Fig. 5A). We also plotted the probability distribution function (PDF) of the H-bonds for both systems (Fig. 5B).

3.5. Principal component and free energy landscape analyses

Principal components analysis (PCA) is a useful approach to extract the dominant modes in a protein motion. It helps to identify the configurational space of the protein that contains a few degrees of freedom while the motion occurs. We have performed PCA to explore the conformational sampling of the M\(^{\text{pro}}\) and M\(^{\text{pro}}\)-ZINC02123811.
ZINC02123811 complex via studying their collective motions through the essential dynamics approach (Amadei et al., 1993). The conformational sampling of M\textsuperscript{pro} and M\textsuperscript{pro-ZINC02123811} complex in the essential subspace is portrayed in Fig. 6A. The projection shows the conformational sampling of M\textsuperscript{pro} along with the EV-1 and EV-2 projected by the protein C\textsuperscript{a} atoms. We found that the M\textsuperscript{pro-ZINC02123811} complex occupied the same conformational subspace as M\textsuperscript{pro} in the free-state. A little decrease was observed at both EVs in the case of the complex with no overall shift of the motion (Fig. 6B).

The FEL analysis provides an atomic resolution of a protein–ligand bound system, possible binding transition states, and metastable states, which can be useful in designing inhibitor. To study the conformational stability and native states of M\textsuperscript{pro} and M\textsuperscript{pro-ZINC02123811} complex, the FELs were generated using the first two PCs. The contoured FELs of M\textsuperscript{pro} and M\textsuperscript{pro-ZINC02123811} complex are illustrated in Fig. 7. While exploring the plots, a deeper blue is portentous to the conformational states with lower energy near to native states. We observed that M\textsuperscript{pro} is having only a single global minimum confined within three local basins. Similarly, M\textsuperscript{pro} in presence of ZINC02123811 acquires different states with multiple minima showing three local basins with different conformational motions (Fig. 7B).

Overall, the drug-like properties including physicochemical and ADMET, higher and specific binding towards the SARS-CoV-2 M\textsuperscript{pro} binding site, and stability during MD simulation studies suggest that ZINC02123811 can act as a potential lead in drug development against SARS-CoV-2 infection. The compound is showing antiviral potential with improved pharmacological properties and considerably high affinity and stability with SARS-CoV-2 M\textsuperscript{pro} hence could be implemented in effective therapeutic development against COVID-19 after required validation.

4. Discussion

Computational approaches are commendable in lead discovery by screening large chemical libraries against predefined drug targets. Here in this study, a library of natural compounds was filtered out based on their physicochemical and ADMET properties to identify safe and effective compounds against SARS-CoV-2 M\textsuperscript{pro}. The finally selected compounds are showing well ADMET properties and drug-likeness. The molecular docking identified that the
selected natural compounds showing a considerable binding affinity towards SARS-CoV-2 MPro. The interaction analysis suggests that all selected compounds occupy the same position where most of the co-crystalized ligands bind. The SARS-CoV-2 MPro cysteine-histidine catalytic dyad including Cys145 and His41, located at the main catalytic center of the protein is responsible for the functional activity of MPro. We found three natural compounds which are commonly interacting with the binding-site residue ‘Cys145’ of SARS-CoV-2 MPro. PASS analysis was carried out to explore the biological properties of the compounds. Based on the PASS analysis, and specific interactions towards SARS-CoV-2 MPro cysteine-histidine catalytic dyad, Cys145 and His41, we have selected ZINC02123811 as a potent compound against SARS-CoV-2 MPro.

The binding of any small chemical compound can make significant conformational changes to a protein structure (Kalita et al., 2020). RMSD analysis shows that the binding of ZINC02123811 with MPro showed to have equilibration in RMSD throughout the simulation time of 100 ns which suggesting stability of the docked complex. The RMSF fluctuations were taken during the simulation for each residue in the backbone of MPro before and after ZINC02123811 binding. The fluctuations were found to be stable and minimized upon ZINC02123811 binding suggesting significant stability of the protein–ligand complex. The Rg of a protein is directly linked to its tertiary structure and thus it is one of the widely employed parameters to study the compactness of a protein structure (Gupta et al., 2020; Naqvi et al., 2018). The Rg plot suggested that MPro was stably folded with ZINC02123811 and behaved like the MPro free. The SASA of a macromolecule is the area that is accessible to its surrounding solvent (Mazola et al., 2015). It has been employed in exploring the folding behavior of proteins under solvent conditions. No switching in SASA was observed throughout the trajectory of 100 ns suggesting a stable complex of MPro and ZINC02123811.

The intramolecular H-bonds within a protein plays a fundamental role in its stability (Hubbard and Kamran Haider, 2001; Khan et al., 2017a; Prakash et al., 2019; Prakash et al., 2018; Shahbaaz et al., 2019). The plot suggests that there is no major change in the number of H-Bonds formed intramolecular within MPro and MPro-ZINC02123811 complex. This analysis suggests that the MPro-ZINC02123811 complex is quite stable throughout the simulation.

The structural dynamics and conformational sampling of a protein can be explored through its phase space performance (Naz et al., 2019; Naz et al., 2018; Papaleo et al., 2009). The projection of conformational sampling of MPro-ZINC02123811 along with the EV-1 and EV-2 projected by the protein Cα atoms is overlapping the stable clusters with phase space of MPro-apo. The PCA analysis including the EV1 RMSF indicates that MPro and its complex with ZINC02123811 are pretty stable during the simulation course. FEL plots also suggest that the binding of ZINC02123811 to MPro affects the size and the location of the sampled essential subspace but with a confined stable global minimum.

Overall analysis suggests that computational methods would play a significant role in the design and development of potential therapeutic molecules to address COVID-19. MPro plays a significant role in mediating replication and transcription and subsequent pathogenesis of SARS-CoV-2. In a recent study, Wu et al. (2020) have demonstrated that MPro of SARS-CoV-2 inhibits the IFN induction, which may reduce antiviral responses in infected cells and thus be considered as a novel target for potential therapeutic intervention on SARS-CoV-2 infection. In another study, using a combination of structure-based virtual and high-throughput screening, Jin et al. (2020) assayed a large number of compounds and identified potential MPro inhibitors with half-maximal inhibitory concentration (IC₅₀) values in the range of 0.67–21.4 μM. In a similar study, Douangamath et al. (2020) screened a larger set of the electrophile and non-covalent fragments using combined mass spectrometry and X-ray approach against the MPro of SARS-CoV-2. Several of these compounds bind to the active site pocket of enzymes and offered promising antiviral activity. Our results demonstrate the efficacy of our in-silico screening strategy, which offers a rapid discovery of potential drug leads for COVID-19. The identified compounds are showing better drug-like properties and similar fashion of binding pattern as compared to the reported MPro inhibitor N3, in-silico. The identified compound ZINC02123811 is also shown to have antiviral potential in PASS analysis making it a potent lead scaffold for drug development against COVID-19.

5. Conclusion

With the emergence of the COVID-19 pandemic, a quick drug development against SARS-CoV-2 is immediately needed. Targeting SARS-CoV-2 MPro with natural compounds is an attractive strategy for antiviral therapy. We employed a structure-based drug discovery approach and identified a natural compound, ZINC02123811 (1-(3-(2,5,9-trimethyl-7-oxo-3-phenyl-7H-furo[3,2-g]chromen-6-yl)propargyl)piperidine-4-carboxamide) which is showing antiviral potential with improved pharmacological properties and considerably high affinity and stability with SARS-CoV-2 MPro. The MD simulation study suggests the formation of a highly stable complex of MPro with ZINC02123811. Altogether, this study provides a strong indication that ZINC02123811 might be further employed as a lead to develop potent and selective inhibitors of SARS-CoV-2 MPro for the therapeutic management of COVID-19 after required clinical validations.

Declaration of Competing Interest

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

Acknowledgments

MHI acknowledges Council of Scientific and Industrial Research (Grant No. 27/0368/20-EMR1) and Indian Council of Medical Research for financial support (Grant No. ISRM/12(22)/2020). Authors sincerely thank to the Department of Science and Technology, Government of India for the FIST support (FIST program No. SR/FST/LSI-541/2012). MFA and AH extend their sincere appreciation to the Deanship of Scientific Research at King Saud University for its funding of this research through the Research Group Project No. RGP-150. Authors sincerely acknowledge Natural Science Foundation of Sichuan Province of China (Grant No. 21JH0266). DSJ sincerely acknowledge the support of CMMS, Arabian Gulf University, Kingdom of Bahrain (E006-PI-10/20).

References

Abraham, M.J., Murtolà, T., Schulz, R., Páll, S., Smith, J.C., Hess, B., Lindahl, E., 2015. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 1, 19–25.
Altis, A., Otten, M., Nguyen, P.H., Hegger, R., Stock, G., 2008. Construction of the free energy landscape of biomolecules via dihedral angle principal component analysis. J. Chem. Phys. 128, 060620.
Amadei, A., Linsen, A.B., Berendsen, H.J., 1993. Essential dynamics of proteins. Prot. Struct. Funct. Bioinf. 17, 412–425.
Asrani, P., Eappen, M.S., Chia, C., Haug, G., Weber, H.C., Hassan, I., Sohal, S.S., 2020a. Diagnostic approaches in COVID-19: clinical updates. Exp. Rev. Respir. Med. 1, 1–16.
Asrani, P., Hasan, G.M., Sohal, S.S., Hassan, M.I., 2020b. Molecular basis of pathogenesis of coronaviruses: a comparative genomics approach to
