Future Journal of Pharmaceutical Sciences

RESEARCH

Open Access

nCoV-19 therapeutics using cucurbitacin I structural derivatives: an in silico approach

Ram Lal Swagat Shrestha[1], Bishnu Prasad Marasini[2*] and Jhashanath Adhikari Subin[3*]

Abstract

Background Cucurbitacins are present in some common vegetables as secondary metabolites and are used by the plants against harmful microbes. Exploration of this capability of natural product based substances against wide variety of microbes seems relevant due to the ease of availability of the resources and safety. In this regard, considering the current pandemic, the antiviral properties of these molecules with a subset of Cucurbitacin I structural derivatives have been screened. The inhibition potential of the phytochemicals was assessed by the stability of the protein–ligand complex formed with the nucleocapsid protein (PDB ID: 7CDZ) of SARS-CoV-2 by computational methods. The proposition of an alternate antiviral candidate that is cost-effective and efficient relative to existing formulations is the main objective of this work.

Results Server-based molecular docking experiments revealed CBN19 (PubChem CID: 125125068) as a hit candidate among 101 test compounds, a reference molecule (K31), and 5 FDA-approved drugs in terms of binding affinities sorted out based on total energies. The molecular dynamics simulations (MDS) showed moderate stability of the protein-CBN19 complex as implied by various geometrical parameters RMSD, R_g, RMSF, SASA and hydrogen bond count. The ligand RMSD of 3.0 ± 0.5 Å, RMSF of C_α of protein with less than 5 Å, and smooth nature of SASA and R_g curves were calculated for the adduct. The binding free energy (− 47.19 ± 6.24 kcal/mol) extracted from the MDS trajectory using the MMGBSA method indicated spontaneity of the reaction between CBN19 and the protein. The multiple ADMET studies of the phytochemicals predicted some drug-like properties with minimal toxicity that mandate experimental verification.

Conclusions Based on all the preliminary in silico results, Cucurbitacin, CBN19 could be proposed as a potential inhibitor of nucleocapsid protein theoretically capable of curing the disease. The proposed molecule is recommended for further in vitro and in vivo trials in the quest to develop effective and alternate therapeutics from plant-based resources against COVID-19.

Keywords ADMET predictions, Binding free energy, MMGBSA, Molecular docking, Molecular dynamics simulation, Natural products

*Correspondence:
Bishnu Prasad Marasini
bishnu.marasini@gmail.com
Jhashanath Adhikari Subin
subinadhikari2018@gmail.com
Full list of author information is available at the end of the article

Background

After claiming nearly 7 million people, the virus responsible for COVID-19, SARS-CoV-2 is still propagating [1, 2]. Even though the vaccines have helped to contain the pandemic to some extent, effective and safe drugs are still in the experimental phases [3, 4]. One of the alternatives to existing pharmaceuticals is the phytochemical-based therapeutics that need to be designed and developed for proper treatment and cure of the disease in a cost-effective manner. Various natural products have been reviewed in terms of their potential to fight the disease, taking into consideration the specific targets [5]. Cucurbitacins, tetracyclic triterpenoids occur in regularly consumed vegetables like cucumber, gourds, squash, and pumpkins as secondary metabolites and act in the plant's defense mechanisms [6]. The molecular structure of Cucurbitacin I is shown in Fig. 1. These vegetables contain multi-sized nutrients, antioxidants, vitamins, amino acids, and small anti-nutrients [7]. It carries non-trivial ethnobotanical significance and has been in use in the treatment of multiple diseases since ancient times [8–12] with weaker side effects, good safety, and multiple healing pathways [13]. Especially, cucurbitacin B has been found to possess anti-Herpes simplex virus 1 biological activity with IC_{50} of 0.94 ± 0.2 μM [14] and forms the basis for the selection of this class of organic compound with possibly similar pharmacophore. Therefore, the adaptation of local practices by scientific research in developing a new type of compound as a drug with molecular-level understanding seems relevant considering this class of organic compound having known biological activity and the requirement of the present global context.

Out of twenty-nine proteins in the viral genome of SARS-CoV-2, a structural protein, nucleocapsid (N), seems to be a good target for small ligands since it is responsible for RNA encapsidation, transcription and replication of the viral genome [15, 16]. Therefore, its involvement in the assembly, budding, cell cycle regulation, immune system modulation, and cessation of host cell translation makes it one of the prominent functional entities [17, 18]. If a guest molecule binds effectively to it then its normal functioning will be disrupted, which would result in the deceleration or halting of the viral replication, ultimately leading to the cure of the disease caused by it. A recent study involving NMR and mass spectroscopic techniques has shown the protein to be druggable and possesses greater availability [19]. A compound K31 [4-(3-Bromophenyl)-3a,4,5,9b-tetrahydro-3H-cyclopenta[c]quinoline-6-carboxylic acid, PubChem CID: 2872351] has been recently reported to bind to it with EC_{50} of 1.7 ± 0.2 μM [15]. Alectinib has been proposed to be capable of inhibiting nucleocapsid protein by prohibiting phosphorylation from in vitro studies [20]. Using in silico approach, Curcumin, Apigenin, Cinnamic acid, Simeprevir, and Grazoprevir have been found to bind successfully with the nucleocapsid protein in different studies [21–23]. Suramin, an antiparasitic drug, has been proposed based on experimental and in silico studies to disturb the association of N-terminal domain (NTD) with RNA, leading to the cessation of viral replication [24]. These results establish the selected protein as a good therapeutic target and therefore, has been chosen for this work.

The relative strength of the binding of the test molecule and its orientation at the orthosteric site of the receptor can be determined by molecular docking calculations as that performed with Cucurbitacins against multiple target proteins of SARS-CoV-2 [25]. The assessment of the stability of the protein–ligand adduct by monitoring the conservation of the initial docked pose of the ligand at the active pocket by molecular dynamics simulations is required. The spontaneity of the forward reaction inferred from binding free energy changes would help to design a molecule that can effectively inhibit the protein by binding to it in a thermodynamically and geometrically stable manner. The safety and bioavailability of the drug-like candidate from ADMET predictions would help to establish the test compound as a good and safe drug against the disease [26]. The computational approach stands as a viable tool in designing a drug that would prevent later failures in high-throughput experiments and clinical trials. One objective of this work is a quick screening of safe, natural product-based molecules with Cucurbitacin I as a scaffold against SARS-CoV-2. The other is an extension of traditional medicinal knowledge into the modern concept of rational drug design and development with molecular-level understanding.

Since the nucleocapsid protein is of utmost importance in functioning at various stages in the viral life cycle [27] and is druggable, it has been adopted in this research work for exploring its possible interactions with Cucurbitacins I scaffold derivatives. The antiviral capabilities of a

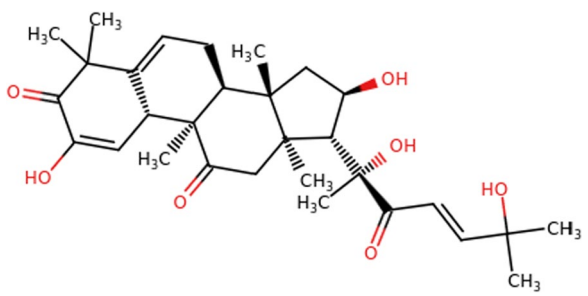

Fig. 1 Molecular structure of a parent molecule, Cucurbitacin I

slightly different structure of Cucurbitacin B have added to the rationale of small molecule selection [14].

Objectives of the work

a. To explore the phytochemicals from regularly used vegetables with known biological activities against a specific viral target by cost-effective computational methods.
b. To identify a drug-like and safe molecule with potential inhibitory capabilities of the functioning of the nucleocapsid protein.
c. To recommend a hit molecule for in vitro and in vivo trials in the course of developing it as an effective nutraceutical/pharmaceutical against SARS-CoV-2.

Experimental/methodology

Ligand and receptor preparation

The parent molecule, Cucurbitacin I (PubChem CID: 5281321) and its substructure derivatives were taken in sdf file format from the PubChem webpage (https://pubchem.ncbi.nlm.nih.gov/, accessed 12 July 2023) [28] and converted to pdb format using *obabel* command (The Open Babel Suite, version 2.3.1 http://openbabel.org, accessed 12 July 2023) [29] and the PyMol program [30]. The N-terminal domain of nucleocapsid protein of SARS-CoV-2 as a drug target (PDB ID: 7CDZ, X-ray diffraction, resolution 1.80 Å, Escherichia coli BL21 expression system) (https://doi.org/https://doi.org/10.2210/pdb7CDZ/pdb) was retrieved in pdb file format from the RCSB website (https://www.rcsb.org/, accessed 4 July 2022) [31], verified and preprocessed for molecular docking.

Molecular docking calculations

The DockThor web-server (https://dockthor.lncc.br/v2/, accessed 12 July 2023) [32] was used in molecular docking calculations with center coordinates of (−6, 22, 16), discretization of 0.25, and grid size of ($20\times20\times20$ Å^3) of the active pocket. The genetic algorithm search settings involved 1,000,000 evaluations with 750 population size for 24 number of runs deemed suitable for capturing the local minima in a complicated potential energy surface. The affinity prediction was done for each ligand based on the total energy, which was calculated for its different binding modes. It involved inter-molecular, intra-molecular, and torsional energies and were clustered to finally yield the best value [32, 33].

Molecular dynamics simulations and thermodynamics

The docked pose with the best binding affinity was taken for molecular dynamics simulations (MDS) of 90 ns production run at 310 K using the GROMACS version 2021.2 program [34]. The force-field (charmm27) [35] was obtained from the SwissParam server (https://www.swissparam.ch/, accessed 12 July 2023) [36], and all the other parameters as reported by Sharma et al. [37] were adopted. Multiple parameters like root mean square deviation (RMSD), radius of gyration (R_g), root mean square fluctuation (RMSF), solvent accessible surface area (SASA), and number of hydrogen bonds between ligand and residues were determined from MDS trajectory using inbuilt commands of GROMACS. Binding free energy change of complex formation was calculated using MMGBSA method [38] with GB7 option [39] for the best-docked pose using gmx_MMPBSA module (100 frames) [40] in GROMACS and also using the fastDRH server (http://cadd.zju.edu.cn/fastdrh/, accessed 15 July 2023, ID4638) [41]. The differences in the fluctuation of the residues upon ligand binding were determined in terms of RMSF difference from the MDS of 10 ns duration using the CABSflex2 web-server (http://biocomp.chem.uw.edu.pl/CABSflex2/, accessed 15 July 2023) [42].

Pharmacokinetics and pharmacodynamics

The SMILES strings of the molecules were used as input in determining various parameters. The ADMET prediction [43] was made using the pkCSM (https://biosig.lab.uq.edu.au/pkcsm/prediction, accessed 24 October 2023) [44] and the ADMETboost (https://ai-druglab.smu.edu/, accessed 25 October 2023) [45] servers. Various toxicity end-points and the rule of five (RO5) or Lipinski's rule were calculated along with other physicochemical parameters of selected Cucurbitacin I derivatives.

Programs and resources

The structures were visualized using the Biovia Discovery Studio visualizer v21.1.0.20298 [46] and the Avogadro2 v1.93.0 [47] programs. The graphical representations were made with the GNUPLOT 5.2 patch level 8 [48] and the Grace-5.1.25 [49] software. The SMILES strings were converted into 2D molecular structures by the online Marvin JS structure editor (https://chemaxon.com/products/marvin-js). The computational experiments were carried out in different machines with multi-core processors, GPU accelerator, RAM 128 GB, and storage 6 TB. The operating systems were Ubuntu 20.04 LTS and Windows 8. Data analysis, visualization and interpretation were carried out on a PC with minimal configuration.

Results

Binding affinities and orientation of the docked molecule

The 2D representations of selected test molecules and a reference are shown in Fig. 2. The binding affinities of the top five Cucurbitacin I derivatives, a reference compound

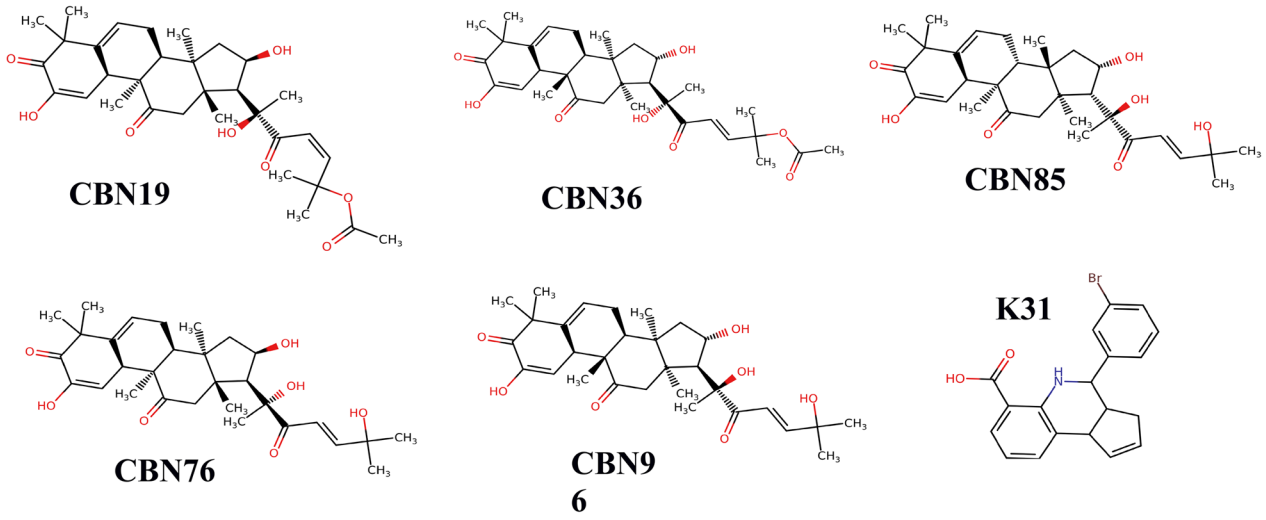

Fig. 2 Molecular structures of the top compounds and a reference

Table 1 Top five Cucurbitacin I derivatives, a reference compound and 7 commercial drugs in terms of binding affinities with NTD of nucleocapsid protein

Molecules (PubChem CID)	Affinity (kcal/mol)	Total energy (kcal/mol)	Commercial drugs	Affinity (kcal/mol)	Total energy (kcal/mol)
CBN19 (125125068)	− 10.27	92.01	Simeprevir	− 10.56	5227.5
CBN36 (162939996)	− 10.20	73.67	Vapreotide	− 10.45	212.33
CBN85 (138107616)	− 10.16	83.82	Atazanavir	− 9.66	82.34
CBN76 (125039415)	− 10.06	73.16	Ivermectin	− 9.13	1375.73
CBN96 (162970907)	− 10.05	68.28	Remdesivir	− 8.84	10.65
			Hydroxychloroquine	− 8.78	43.53
K31 (2872351)	− 8.86	25.67	Chloroquine	− 8.65	16.81

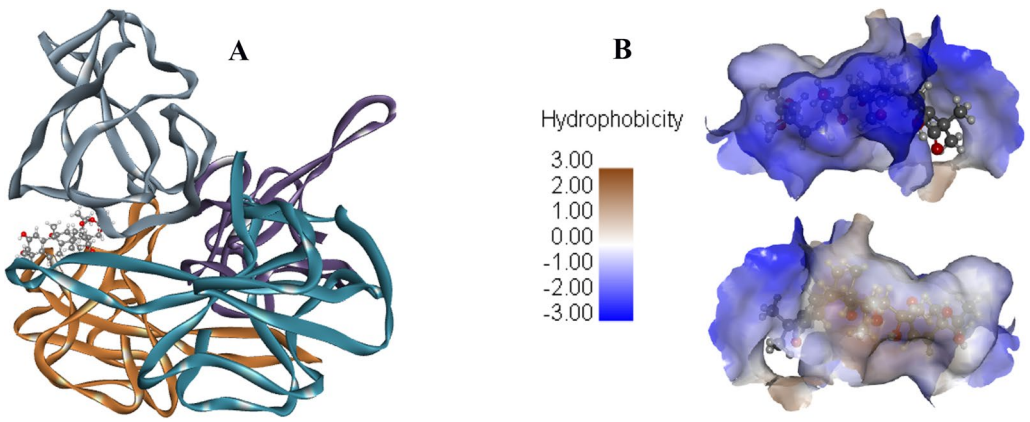

Fig. 3 Docked pose of **A** CBN19 (ball and stick model) with the nucleocapsid protein (ribbon representation) and **B** CBN19 in the active pocket enveloped with hydrophobic surface viewed at 180° apart

(K31) and that of FDA-approved drugs are shown in Table 1.

The compound CBN19 in its docked pose with the receptor is displayed in Fig. 3A. Figure 3B depicts the hydrophobic surface of the active pocket of the protein that interacts with the ligand from two viewing angles 180° apart.

It is evident from the data that the drugs, Simeprevir and Vapreotide lead among all the compounds, and the former, along with Grazoprevir, has been reported by Bhowmik et al. [23] from computational screenings as better candidates. However, the top five test candidates possess better affinities than the other five FDA-approved drugs. The compound CBN19, with the best binding affinity could be considered a hit candidate among the pool of the studied molecules. The highest total energy, however, imply that the reference compound K31, possessed weaker binding (−8.86 kcal/mol) than that of the test compounds and was bound to the N-terminal domain of nucleocapsid protein as inferred experimentally [15]. Apart from the molecules mentioned above, Emetine and Glycyrrhizin have also been found to possess good docking properties against nucleocapsid protein [50, 51]. The computational results show promising properties of the selected molecules derived from a molecular scaffold suitable for further characterization and trials.

Druglikeness and toxicity assessment

The drug-like characteristics in terms of RO5 were found to be acceptable for most of the compounds, inferring good bioavailability and permeability (data presented in supplementary information as ADMET_predictions.xls file). The pharmacokinetics of various Cucurbitacins have been reported by Delgado-Tiburcio et al. [9]. The toxicity in terms of the AMES test, hERG I/II inhibition, hepatotoxicity and skin sensitization showed negative results for the test molecules. The permeability of BBB and CNS are poor or absent, with suitable excretion parameters. The results from the pkCSM server showed that Cucurbitacins possess acceptable drug-like properties with relatively safe end-points.

The ADMET predictions from another server (ADMETboost), to validate the earlier results, were analyzed for different properties. The output, however, showed that Caco2-permeability and aqueous solubility lay outside the acceptable range, whereas all the other properties were predicted to be similar to those from the pkCSM server.

Geometrical stability of complexes

The binding affinity based on a scoring function provides a clue towards the strength of interactions between the ligand and the protein and can be used as a comparative property. It may not necessarily reflect the binding efficacy considering the time the ligand stays at its initially docked position and may not be a measure of the absolute property. To determine the conservation of pose and position of the ligand with time and to determine the geometrical stability of a system, molecular dynamics simulation of 90 ns duration was carried out and various parameters like RMSD, SASA, RMSF, R_g, and hydrogen bond count were extracted from the trajectory. These parameters provide a numerical assessment of the structural stability and are analyzed thoroughly in this section. The RMSD of the ligand and that of the protein are shown in Fig. 4 as subplots A and B, respectively.

The RMSD of ligand CBN19 with respect to the protein backbone (dark curve) was *ca.* 2 Å and was moderately smooth with small fluctuations at 55 ns and 70 ns regions. It remained flat from 75 to 85 ns, similar to the initial part of the trajectory up to 45 ns. The results hint at near preservation of the docked pose during the simulation period. Similar values have been reported for approved

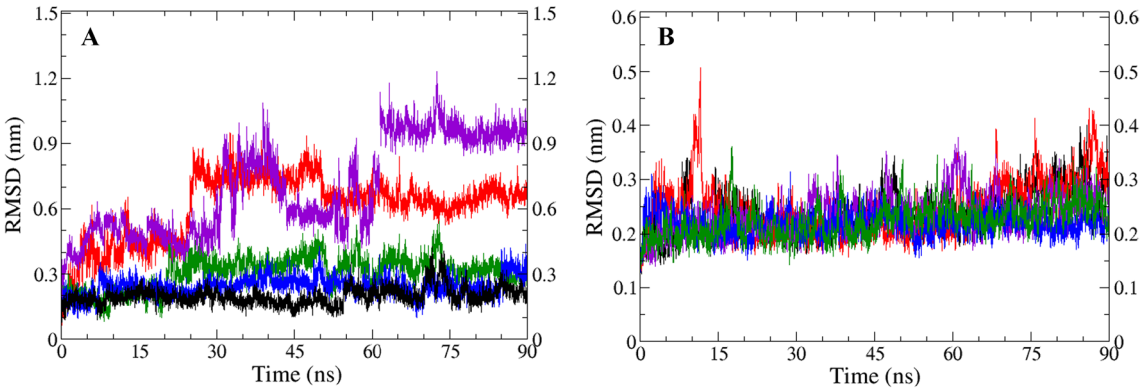

Fig. 4 Multiple parameters extracted from MDS trajectory **A** RMSD of ligands with respect to protein backbone **B** RMSD of protein backbone with respect to protein backbone (CBN19 = dark, CBN36 = red, CBN85 = green, CBN76 = blue, and CBN96 = indigo)

drugs against nucleocapsid CTD [52]. The compounds with CBN36 and CBN96 showed unsteady behavior, both attaining equilibrium after 60 ns with RMSD larger than 6.8 Å. Another compound CBN85 showed an RMSD of *ca.* 4 Å or lower with multiple spikes during the production run. It did not attain steady state, as shown by a deep at 85 ns and necessitates a longer simulation run. The complex of CBN76 showed an RMSD of *ca.* 3 Å with a rise at 85 ns. The unsteady curve implies significant ligand motion at the active site, and the stability could only be confirmed from extended production runs. Multiple curves showed a rise at 70 ns except CNB36, indicating a common factor in the structural modulation. The lack of such a feature in the protein backbone may be attributed to the function of temperature leading to similar coincidental atomic spatial evolution of the docked ligands and may not be relevant in assessing the stability of the adduct. The nature of the ligand RMSD curves suggests the formation of moderately stable protein–ligand adducts, with most of them attaining equilibrium after *ca.* 75 ns.

The multiple RMSD of protein backbone with respect to the protein backbone, as shown in Fig. 4B, are less than 3 Å except that in the case of CBN36, which was sometimes higher in magnitude. The relatively flat and smooth nature of the curves hints at greater stability of the protein structure during the production run. The overall geometry of the receptor and the active pocket can be inferred to remain intact, capable of binding the ligand effectively. Out of five complexes studied, since CBN19 showed comparatively promising stability, other structural features related to it were only derived and analyzed.

The RMSF plots of alpha carbon atoms of the protein in the adduct and in the receptor alone were extracted from two separate MDS trajectories and are given in Fig. 5 as subplots A and B, respectively. Figure 5B shows large fluctuations *ca.* 6.5 Å from residue number 45 to 65 in the two chains out of four of the protein. In the adduct, as shown in Fig. 4A, it is seen to be reduced to *ca.* 5 Å. Similar values have been published for cases with other types of organic molecules against the NTD domain of nucleocapsid protein [53]. A decrease in RMSF in the red curve and an increase in the dark curve in the same region can be seen, which corresponds to two other chains of the protein. Overall, the ligand binding to the protein has resulted in some reduction of RMSF of the alpha carbon atoms of the protein that could be proposed as induction of stability to the system. The presence of large fluctuations in some parts of the structure could be responsible for the wavy nature of the ligand RMSD curve since the residues undergo significant motion during the simulation. The twisted geometry of the anti-parallel beta sheets at the proximity of the docked ligand might also be a determining factor. Large RMSF has been reported for the protein in complex with the hit molecules [23, 54]. The difference in RMSF of the residues from 10 ns server-based MDS as shown in Additional file 1: Fig. S1, reveals the presence of more peaks at the negative side relative to that at the positive side, with the overall value being negative. This implies that the fluctuation of residues is reduced when a ligand binds to it, providing further support for the stability of the adduct from a different approach.

The R_g of protein in protein-CBN19 complex (dark curve of Fig. 6A) of *ca.* 2.5 nm is similar to that reported by Suravajhala et al. [21] and is smaller than that reported by Ribeiro-Filho et al. [55]. SASA (brown curve of Fig. 6A) of *ca.* 273 nm^2 is slightly outside the range 230–260 nm^2 as reported for the Curcumin molecule and nucleocapsid receptor complex [21]. Both the geometrical parameters suggest inherent intactness (no significant

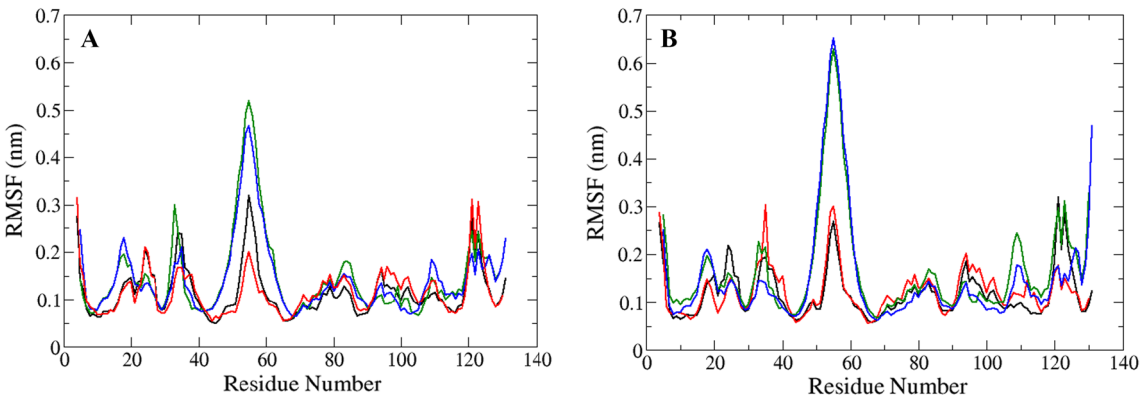

Fig. 5 RMSF of amino acid residues of protein in **A** nucleocapsid-CBN19 adduct and **B** apo structure of nucleocapsid protein derived from two separate 90 ns MDS (the four colors correspond to the four chains of the receptor and NOT to the ligands as in the other figures)

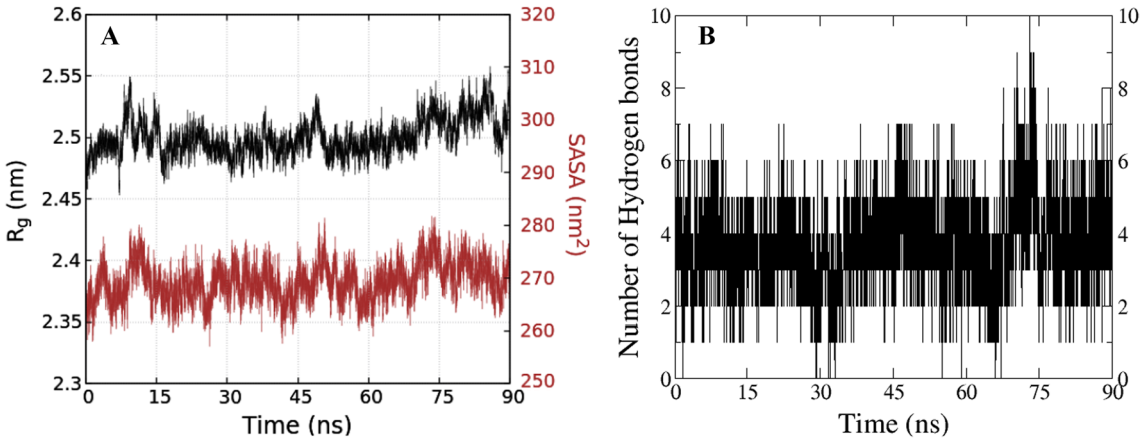

Fig. 6 **A** Radius of gyration (dark) of protein and SASA (brown) of protein in the adduct **B** Variation of hydrogen bond count in the adduct during 90 ns production run

expansion or contraction) of the protein in the adduct during the production run.

The number of hydrogen bonds between the receptor and the ligand in the adduct during the MDS changed from null up to 10, as shown in Fig. 6B, with maximum frames having 4 or 5. The number of hydrogen bonds stayed up to 5 ± 1 at the equilibrated part of the trajectory and contributed significantly to the non-covalent binding of the ligand CBN19 with the protein. It might be the causative factor in the appearance of a smooth curve (dark color) after *ca.* 77 ns up to 85 ns in the RMSD plot, Fig. 4A. The presence of small bumps along the RMSD curve of CBN19 might be due to the fluctuations from the protein backbone. A rise at *ca.* 70 ns in the hydrogen bond count to 10 is reflected by multiple curves (RMSD, SASA and R_g) showing change in ligand orientation and the protein backbone. Similar hydrogen bond count and other parameters have been reported by Dhankhar et al. [53], and therefore, the reliability of the results can be justified. Nonetheless, the lack of appreciable change in the number of hydrogen bonds during the simulation indicated a near-steady nature of interactions between the ligand and the amino acid residues.

The snapshots of nucleocapsid-CBN19 adduct at the active pocket at different time intervals are shown in Fig. 7. The center of mass of the ligand and the structure of the protein backbone are seen to be nearly conserved (comparing snapshots at 1 ns and 90 ns) with a slight variation in the orientation of the ligand. The progressive display of the ligand at the orthosteric site of the receptor during the MDS shows the absence of translational motion. The twisted anti-parallel beta sheets (yellow) seem to show some geometrical changes along with the loop (green) attached to it. The disappearance of the alpha helix structure at the left side of the frame at 30 ns could be correlated to a decrease in hydrogen bond count and is reflected by slightly increased RMSD. The pose of the ligand seems to be maintained with minor variation, which might be due to a change in the location of the protein backbone as a function of temperature. Similar geometrical analysis have been reported on the docked complex of quinoline derivative with M^{pro} of SARS-CoV-2 using snapshots by Singh et al. [56]. Also, Nguyen et al. have studied the druggable targets of dengue and malaria using multiple frames at different times during MDS [57]. The snapshots have revealed that the active site of the protein is capable of holding the docked ligand for sufficient time, which could lead to the disruption or deceleration of metabolic reactions.

Spontaneity of the adduct formation reaction

The binding free energy change (ΔG_{BFE}) of complex formation from the protein and the ligand was calculated to be − 47.19 ± 6.24 kcal/mol from the MDS trajectory and − 34.06 kcal/mol from the server-based approach. The spontaneity of the reaction was revealed by the negative values obtained from both the methods. It showed the feasibility of the nucleocapsid-CBN19 adduct formation and the usefulness of the MMGBSA method in estimating the free energies. The multiple components of the free energy changes are shown in the Additional file 2: Table S1.

The in silico results obtained from MDS in terms of RMSD, SASA, R_g, RMSF and hydrogen bond count have established the effectiveness of molecular docking in providing adequate models and parameters reflecting the geometrical stability of the system studied. The thermodynamic parameter has complemented this finding. A comprehensive suggestion would be that the ligand and the protein could form a moderately stable compound in

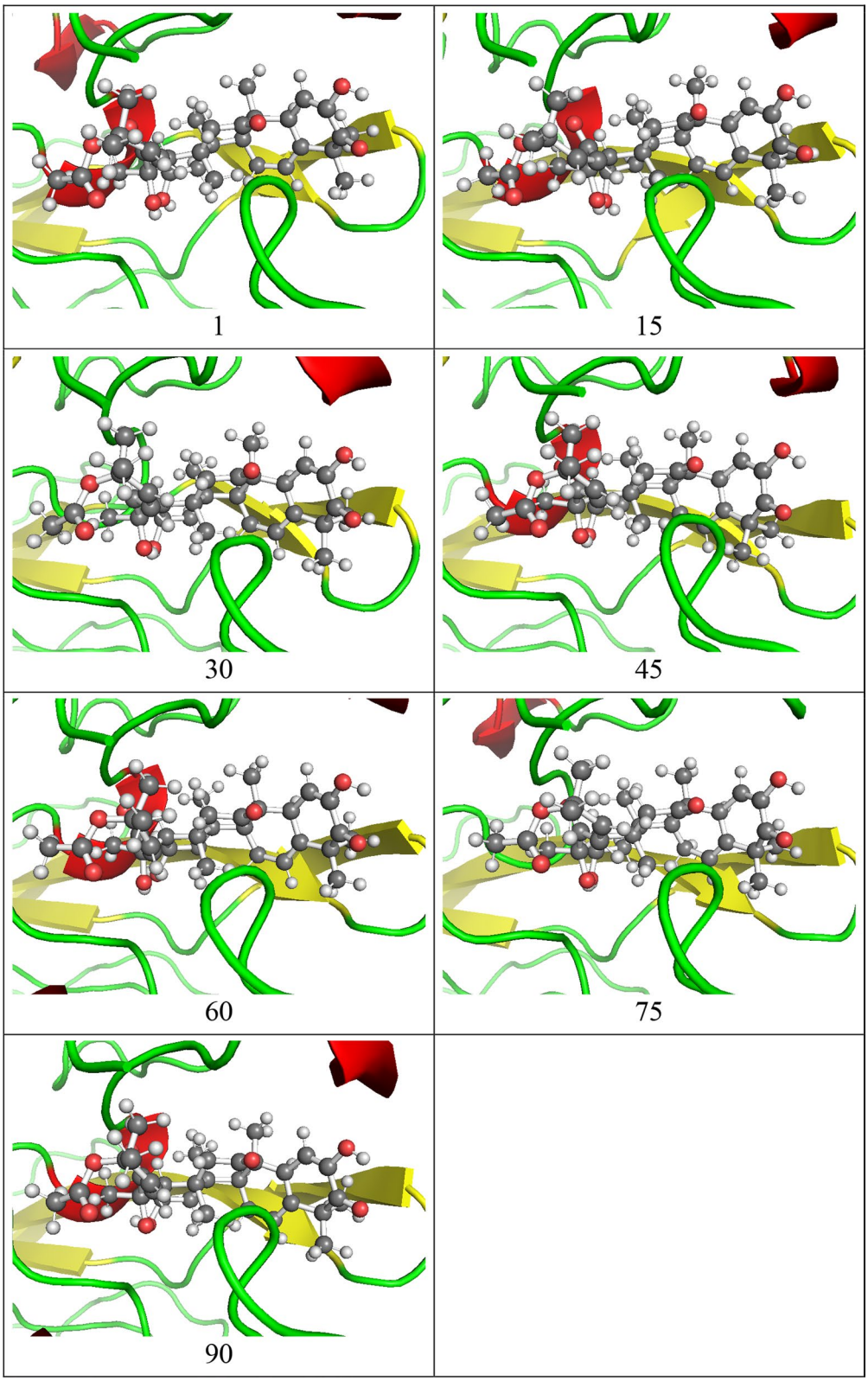

Fig. 7 Snapshots of the active region of the protein-CBN19 complex at different times during MDS (the numbers represent times in nanoseconds)

which the docked pose is fairly conserved, that may be suitable for effective inhibition of the protein functioning.

Molecular level analysis of non-covalent interactions in the adduct

The cumulative interactions between the ligand and the amino acid residues of the protein are presented in Fig. 8 in terms of 2D representation. An earlier figure, Fig. 3A, shows the docked position of the ligand at the orthosteric site of the receptor calculated by blind docking and verified again by 5 different ligands. The findings of nondeep cavities unsuitable as a well-defined active region in the NTD domain for RNA binding have added to the ambiguity of the active site [58]. One side of the pocket is completely hydrophilic (exterior), and the other is hydrophobic (interior). Therefore, the lack of bonding involving electrostatics is seen at the hydrophobic part of the protein and is compensated by the involvement of multiple residues with more than 25 van der Waals interactions.

The molecular level understanding of the interactions constitutes the knowledge of key amino acid residues involved. There is a presence of disruptive interaction (acceptor-acceptor) from residue THR101 with simultaneous associative interaction through very weak hydrogen bonding at 4.37 Å with different oxygen atoms of the ligand. Another weak hydrogen bonding with THR30 at 3.92 Å and alkyl-related interaction with ARG49 at 5.10 Å are also present. Different sets of key residues have been reported to be involved in binding Ceftriaxone sodium and RNA separately by Luan et al. [58]. The poses with the best binding affinities sorted out based on total energies seem to be acceptable as these form stable complexes with a few non-covalent interactions good enough for inhibiting the protein's functioning. It could be compared with the performance of Ceftriaxone sodium, reported to be an inhibitor of NTD of nucleocapsid protein. For comparative purposes the 2D representation of the adduct formed by the receptor with the reference compound K31 depicting various types of interactions is shown in Fig. 9.

The involvement of a common amino acid residue ARG49 with alkyl related hydrophobic interactions and THR101, ILE111, and VAL112 with other different types of binding interactions in both the adducts support the inference that the orthosteric pocket was occupied by the ligands. The presence of larger number of van der Waals bonding with the hit candidate compared to that with the reference ligand could be attributed to better binding affinity and stability of the complex. The 2D representations of other adducts are shown in Additional file 1: Figs. S2–S5.

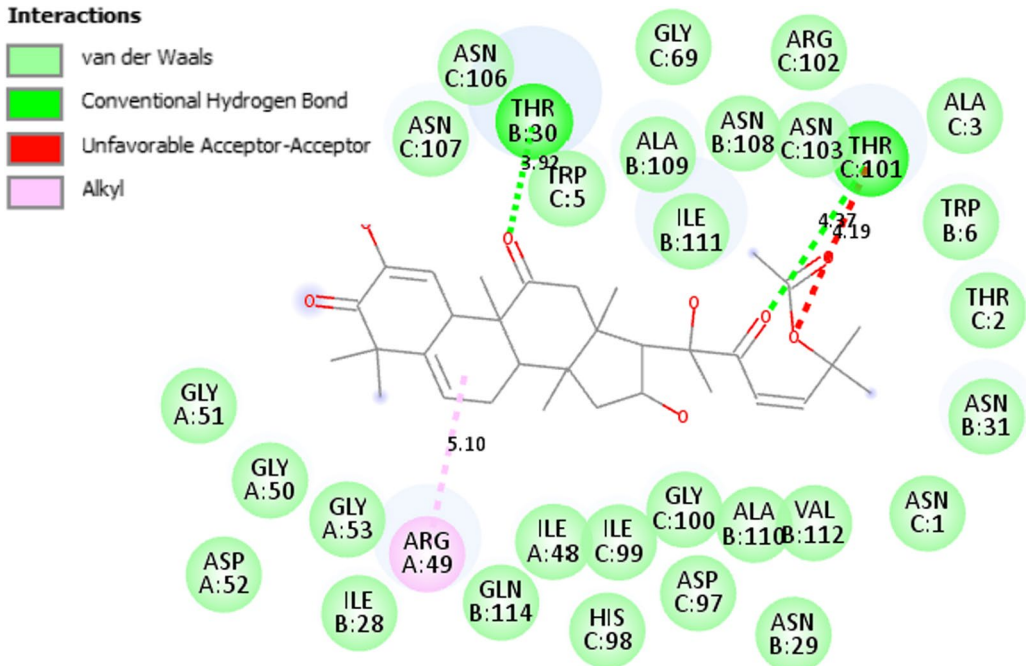

Fig. 8 Protein-CBN19 interactions in the 2D projection of the complex with different types of interactions and distances (Å)

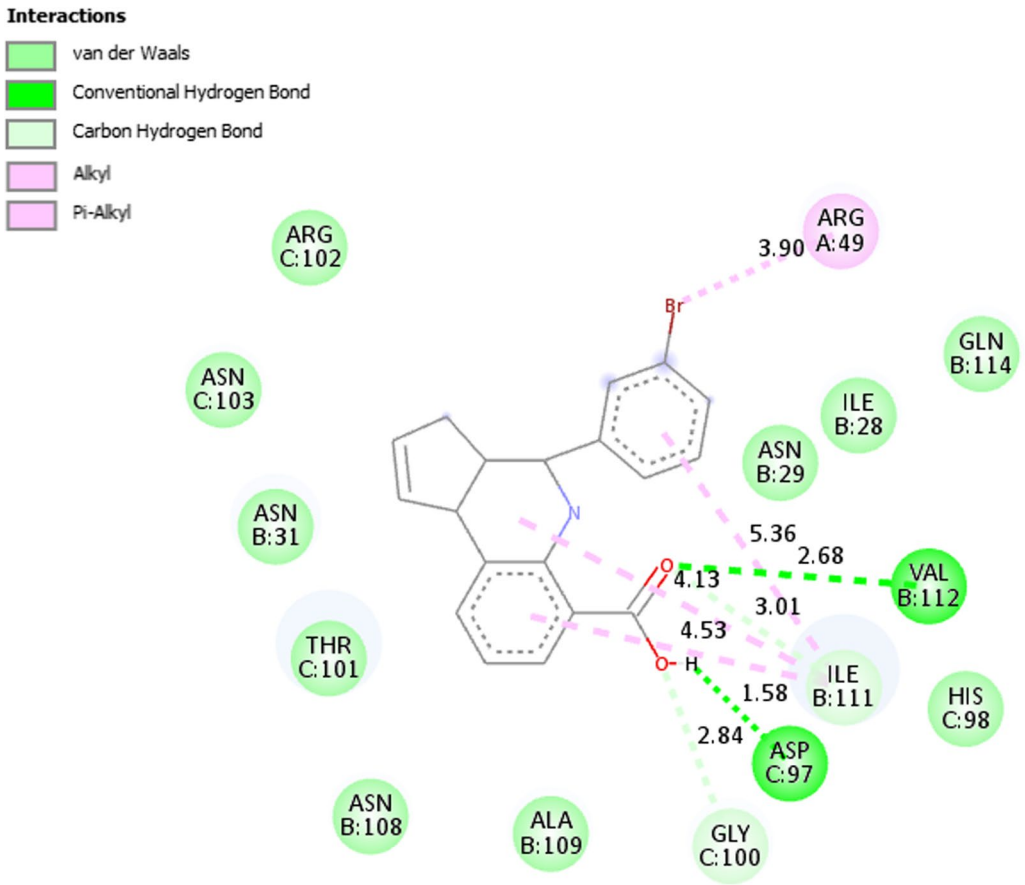

Fig. 9 Protein-K31 interactions in the 2D projection of the complex with different types of interactions and distances (Å)

Discussion

Relative strength of binding of ligands by the protein and stability of adduct

The hit molecule, CBN19, have been found to bind with sufficient strength with the receptor protein relative to other 100 Cucurbitacins, five FDA-approved drugs, and a reference molecule. The strength of interactions is reflected by the binding affinities and the stability of the adduct with time is shown by various geometrical parameters extracted from the MDS. The poses at different times during the MDS production run supported this inference. The presence of weak hydrogen bonding and van der Waals' interactions between the ligand and the amino acid residues seems to be the contributing factor in the conservation of the docked pose at the orthosteric site of the receptor. The molecule could thus be proposed to be a potential inhibitor of the nucleocapsid protein.

Druglikeness and safety

The test compounds seem to have drug-like properties with no observable toxicity. These compounds are present in regularly consumed vegetables, and the administrable dose of each component has to be precisely determined to formulate it as a safe drug candidate that is free from any adverse effects. The proximity of toxic dose amount to its biological activity could be exploited for better therapeutics [59]. The time and dose-dependent cytotoxicity has been reported for Cucurbitacin B [60]. Therefore, the toxicity from the curative effect has to be distinguished by various in vitro and clinical trials [61] and is highly recommended for the future work involving Cucurbitacins.

Validation by thermodynamic parameter

The negative values of binding free energy changes, ΔG_{BFE} obtained from two different methods clearly showed that the adduct formation was spontaneous in nature. It provided additional support for the geometrical stability of the nucleocapsid-CBN19 adduct at the physiological conditions. Therefore, the feasibility of the forward reaction determined from the calculation of the thermodynamic parameter, ΔG_{BFE} strengthened the inferences derived. It supported the molecular docking

methodology, observations derived from it, and the geometrical parameters extracted from the MDS. Similar results have been obtained from the models of protein and approved drugs by Chauhan et al. [52]. The findings of this research is in par with the current literature and thus validates the employment of many types of computational methods.

In the end, the results showed that the server-based molecular docking calculations could provide quick and reliable models that could be used further for the stability assessment. The molecular dynamics trajectory of the adduct formed by CBN19 with the target protein hinted at the sufficient sturdiness of the complex as revealed by multiple parameters. However, pharmacokinetics of the molecule and toxicity predictions obligate mandatory experimental verification for the safety of the hit candidate present in different edible sources. The stability analysis of adducts formed by other different classes of organic molecules with nucleocapsid protein would provide comparative basis against the performance of Cucurbitacins. It is being explored along with the derivatives of compound K31 for finding effective and safe drug-like candidates for fighting COVID-19 and related diseases.

Conclusions

Different computational tools were used in screening the phytochemicals (substructures of Cucurbitacin I) of common vegetables for proposing the potential inhibitors of nucleocapsid protein of SARS-CoV-2. A Cucurbitacin, CBN19, was found to be a hit candidate better than 100 other Cucurbitacins, five approved drugs, and a reference molecule based on ligand binding affinity obtained from molecular docking calculations and from geometrical as well as thermodynamic parameters derived from molecular dynamics simulations. Additional experiments are recommended to confirm the drug-like properties and safety of the proposed molecule in terms of toxicity. In summary, accelerated computational assessment of natural product-based bio-active compounds could be systematically performed in low resource settings for designing therapeutics against COVID-19 and extended to other related diseases in accordance with the modern rational drug discovery process.

Abbreviations

ADMET	Absorption, distribution, metabolism, excretion and toxicity
CID	Compound ID
CTD	C-terminal domain
CBN	Cucurbitacin
EC_{50}	Half maximal effective concentration
FDA	Food and drug administration
hERG	Human ether-a-go-go-related gene
IC_{50}	Half maximal inhibitory concentration
MDS	Molecular dynamics simulation
MMGBSA	Molecular mechanics generalized Born surface area
nCoV	Novel corona-virus
NTD	N-terminal domain
R_g	Radius of gyration
RMSD	Root mean square deviation
RMSF	Root mean square fluctuation
RO5	Rule of five
SARS-CoV-2	Severe acute respiratory syndrome coronavirus-2
SASA	Solvent accessible surface area
SMILES	Simplified molecular-input line-entry system

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s43094-024-00628-y.

Additional file 1. ADMET predictions.

Additional file 2. RMSF difference, 2D projections, and binding free energy changes with its components.

Acknowledgements

Not applicable.

Author contributions

JAS: Conception; calculations; analysis and interpretation of data; drafted the work; RLS: Calculations; managed the resources; revised the work; BPM: Analysis and interpretation of data; supervised the work. All authors read and approved the final manuscript.

Funding

This work did not receive any funding from any sources.

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its supplementary information files. (ADMET_predictions.xls file; SI_Cucurbitacins_1.0.0.docx).

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Department of Chemistry, Amrit Campus, Tribhuvan University, Thamel, Kathmandu 44600, Nepal. [2]Nepal Health Research Council, Government of Nepal, Ramshah Path, Kathmandu 44600, Nepal. [3]Bioinformatics and Chem-informatics Division, Scientific Research and Training Nepal P. Ltd., Kaushaltar, Bhaktapur 44800, Nepal.

Received: 17 November 2023 Accepted: 28 March 2024
Published online: 04 April 2024

References

1. WHO coronavirus (covid-19) dashboard, (2023). https://covid19.who.int/.
2. Dong E, Du H, Gardner L (2020) An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis 20:533–534. https://doi.org/10.1016/S1473-3099(20)30120-1
3. Scholkmann F, May C-A (2023) COVID-19, post-acute COVID-19 syndrome (PACS, "long COVID") and post-COVID-19 vaccination syndrome (PCVS, "post-COVIDvac-syndrome"): similarities and differences. Pathol Res Pract 246:154497. https://doi.org/10.1016/j.prp.2023.154497

4. Provost P (2023) The blind spot in COVID-19 vaccination policies: under-reported adverse events. Int J Vaccine Theory Pract Res 3:707–726. https://doi.org/10.56098/ijvtpr.v3i1.65
5. Rafiq A, Jabeen T, Aslam S, Ahmad M, Ashfaq UA, Mohsin NUA, Zaki MEA, Al-Hussain SA (2023) A comprehensive update of various attempts by medicinal chemists to combat COVID-19 through natural products. Molecules 28:4860. https://doi.org/10.3390/molecules28124860
6. Díaz MTB, Font R, Gómez P, Río Celestino MD (2020) Summer squash. Nutritional composition and antioxidant properties of fruits and vegetables. Elsevier, New York, pp 239–254. https://doi.org/10.1016/B978-0-12-812780-3.00014-3
7. Kaur S, Panghal A, Garg MK, Mann S, Khatkar SK, Sharma P, Chhikara N (2019) Functional and nutraceutical properties of pumpkin: a review. Nutr Food Sci 50:384–401. https://doi.org/10.1108/NFS-05-2019-0143
8. Jing N, Tweardy DJ (2005) Targeting Stat3 in cancer therapy. Anticancer Drugs 16:601–607. https://doi.org/10.1097/00001813-200507000-00002
9. Delgado-Tiburcio EE, Cadena-Iñiguez J, Santiago-Osorio E, Ruiz-Posadas LDM, Castillo-Juárez I, Aguiñiga-Sánchez I, Soto-Hernández M (2022) Pharmacokinetics and biological activity of cucurbitacins. Pharmaceuticals 15:1325. https://doi.org/10.3390/ph15111325
10. Blaskovich MA, Sun J, Cantor A, Turkson J, Jove R, Sebti SM (2003) Discovery of JSI-124 (cucurbitacin I), a selective Janus kinase/signal transducer and activator of transcription 3 signaling pathway inhibitor with potent antitumor activity against human and murine cancer cells in mice. Can Res 63:1270–1279
11. Alsayari A (2014) Anticancer and Antiviral Activities of Cucurbitacins Isolated From Cucumis Prophetarum var. Prophetarum Growing in the Southwestern Region of Saudi Arabia, Electronic Theses and Dissertations. 1985. https://openprairie.sdstate.edu/etd/1985
12. Jing S, Zou H, Wu Z, Ren L, Zhang T, Zhang J, Wei Z (2020) Cucurbitacins: bioactivities and synergistic effect with small-molecule drugs. J Funct Foods 72:104042. https://doi.org/10.1016/j.jff.2020.104042
13. Hassan ST, Masarčíková R, Berchová K (2015) Bioactive natural products with anti-herpes simplex virus properties. J Pharm Pharmacol 67:1325–1336. https://doi.org/10.1111/jphp.12436
14. Hassan STS, Berchová-Bímová K, Petráš J, Hassan KTS (2017) Cucurbitacin B interacts synergistically with antibiotics against Staphylococcus aureus clinical isolates and exhibits antiviral activity against HSV-1. S Afr J Bot 108:90–94. https://doi.org/10.1016/j.sajb.2016.10.001
15. Royster A, Ren S, Ma Y, Pintado M, Kahng E, Rowan S, Mir S, Mir M (2023) SARS-CoV-2 nucleocapsid protein is a potential therapeutic target for anticoronavirus drug discovery. Microbiol Spectrum 11:e01186-e1223. https://doi.org/10.1128/spectrum.01186-23
16. Peng Y, Du N, Lei Y, Dorje S, Qi J, Luo T, Gao GF, Song H (2020) Structures of the SARS-CoV-2 nucleocapsid and their perspectives for drug design. EMBO J 39:e105938. https://doi.org/10.15252/embj.2020105938
17 Vkovski P, Kratzel A, Steiner S, Stalder H, Thiel V (2021) Coronavirus biology and replication: implications for SARS-CoV-2. Nat Rev Microbiol 19(3):155–170. https://doi.org/10.1038/s41579-020-00468-6
18. Wu W, Cheng Y, Zhou H, Sun C, Zhang S (2023) The SARS-CoV-2 nucleocapsid protein: its role in the viral life cycle, structure and functions, and use as a potential target in the development of vaccines and diagnostics. Virol J 20:6. https://doi.org/10.1186/s12985-023-01968-6
19. Padroni G, Bikaki M, Novakovic M, Wolter AC, Rüdisser SH, Gossert AD, Leitner A, Allain FH (2023) A hybrid structure determination approach to investigate the druggability of the nucleocapsid protein of SARS-CoV-2. Nucleic Acids Res 51:4555–4571. https://doi.org/10.1093/nar/gkad195
20. Yaron TM, Heaton BE, Levy TM, Johnson JL, Jordan TX, Cohen BM, Kerelsky A, Lin T-Y, Liberatore KM, Bulaon DK, Kastenhuber ER, Mercadante MN, Shobana-Ganesh K, He L, Schwartz RE, Chen S, Weinstein H, Elemento O, Piskounova E, Nilsson-Payant BE, Lee G, Trimarco JD, Burke KN, Hamele CE, Chaparian RR, Harding AT, Tata A, Zhu X, Tata PR, Smith CM, Possemato AP, Tkachev SL, Hornbeck PV, Beausoleil SA, Anand SK, Aguet F, Getz G, Davidson AD, Heesom K, Kavanagh-Williamson M, Matthews D, tenOever BR, Cantley LC, Blenis J, Heaton NS (2020) The FDA-approved drug Alectinib compromises SARS-CoV-2 nucleocapsid phosphorylation and inhibits viral infection in vitro. bioRxiv. https://doi.org/10.1101/2020.08.14.251207
21. Suravajhala R, Parashar A, Choudhir G, Kumar A, Malik B, Nagaraj VA, Padmanaban G, Polavarapu R, Suravajhala P, Kishor PBK (2021) Molecular docking and dynamics studies of curcumin with COVID-19 proteins. Netw Model Anal Health Inf Bioinform 10:44. https://doi.org/10.1007/s13721-021-00312-8
22. Husain I, Ahmad R, Siddiqui S, Chandra A, Misra A, Srivastava A, Ahamad T, Mohd F, Khan Z, Siddiqi A, Trivedi S, Upadhyay A, Gupta AN, Srivastava B, Ahmad S, Mehrotra S, Kant S, Mahdi AA, Mahdi F (2022) Structural interactions of phytoconstituent(s) from cinnamon, bay leaf, oregano, and parsley with SARS-CoV -2 nucleocapsid protein: a comparative assessment for development of potential antiviral nutraceuticals. J Food Biochem 46:e14262. https://doi.org/10.1111/jfbc.14262
23. Bhowmik D, Nandi R, Jagadeesan R, Kumar N, Prakash A, Kumar D (2020) Identification of potential inhibitors against SARS-CoV-2 by targeting proteins responsible for envelope formation and virion assembly using docking based virtual screening, and pharmacokinetics approaches. Infect Genet Evol 84:104451. https://doi.org/10.1016/j.meegid.2020.104451
24. Guo C, Xu H, Li X, Yu J, Lin D (2023) Suramin disturbs the association of the N-terminal domain of SARS-CoV-2 nucleocapsid protein with RNA. Molecules 28:2534. https://doi.org/10.3390/molecules28062534
25. Kapoor N, Ghorai SM, Kushwaha PK, Shukla R, Aggarwal C, Bandichhor R (2020) Plausible mechanisms explaining the role of cucurbitacins as potential therapeutic drugs against coronavirus 2019. Inf Med Unlock 21:100484. https://doi.org/10.1016/j.imu.2020.100484
26. Van De Waterbeemd H, Gifford E (2003) ADMET *in silico* modelling: towards prediction paradise? Nat Rev Drug Discovery 2:192–204. https://doi.org/10.1038/nrd1032
27. Louten J (2016) Virus replication. Essential Human Virology. Elsevier, New York, pp 49–70. https://doi.org/10.1016/B978-0-12-800947-5.00004-1
28. Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, Li Q, Shoemaker BA, Thiessen PA, Yu B, Zaslavsky L, Zhang J, Bolton EE (2023) PubChem 2023 update. Nucleic Acids Res 51(2023):D1373–D1380. https://doi.org/10.1093/nar/gkac956
29. O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR (2011) Open Babel: an open chemical toolbox. J Cheminform 3:33. https://doi.org/10.1186/1758-2946-3-33
30. Yuan S, Chan HS, Hu Z (2017) Using PyMOL as a platform for computational drug design. Wiley Interdiscipl Rev Comput Mol Sci 7:e1298. https://doi.org/10.1002/wcms.1298
31. Berman HM (2000) The protein data bank. Nucleic Acids Res 28:235–242. https://doi.org/10.1093/nar/28.1.235
32. Guedes IA, Barreto AMS, Marinho D, Krempser E, Kuenemann MA, Sperandio O, Dardenne LE, Miteva MA (2021) New machine learning and physics-based scoring functions for drug discovery. Sci Rep 11:3198. https://doi.org/10.1038/s41598-021-82410-1
33. Guedes IA, Costa LSC, Dos Santos KB, Karl ALM, Rocha GK, Teixeira IM, Galheigo MM, Medeiros V, Krempser E, Custódio FL, Barbosa HJC, Nicolás MF, Dardenne LE (2021) Drug design and repurposing with DockThor-VS web server focusing on SARS-CoV-2 therapeutic targets and their non-synonym variants. Sci Rep 11:5543. https://doi.org/10.1038/s41598-021-84700-0
34. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, Lindahl E (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 1–2:19–25. https://doi.org/10.1016/j.softx.2015.06.001
35. Brooks BR, Brooks CL, Mackerell AD, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caflisch A, Caves L, Cui Q, Dinner AR, Feig M, Fischer S, Gao J, Hodoscek M, Im W, Kuczera K, Lazaridis T, Ma J, Ovchinnikov V, Paci E, Pastor RW, Post CB, Pu JZ, Schaefer M, Tidor B, Venable RM, Woodcock HL, Wu X, Yang W, York DM, Karplus M (2009) CHARMM: the biomolecular simulation program. J Comput Chem 30:1545–1614. https://doi.org/10.1002/jcc.21287
36. Zoete V, Cuendet MA, Grosdidier A, Michielin O (2011) SwissParam: a fast force field generation tool for small organic molecules. J Comput Chem 32:2359–2368. https://doi.org/10.1002/jcc.21816
37 Sharma BP, Adhikari Subin J, Marasini BP, Adhikari R, Pandey SK, Sharma ML (2023) Triazole based Schiff bases and their oxovanadium(IV) complexes: Synthesis, characterization, antibacterial assay, and computational assessments. Heliyon 9(4):e15239. https://doi.org/10.1016/j.heliyon.2023.e15239
38 Hou T, Wang J, Li Y, Wang W (2011) Assessing the Performance of the MM/PBSA and MM/GBSA Methods. 1. The accuracy of binding free energy

calculations based on molecular dynamics simulations. J Chem In Model 51:69–82. https://doi.org/10.1021/ci100275a
39. Mongan J, Simmerling C, McCammon JA, Case DA, Onufriev A (2007) Generalized born model with a simple, robust molecular volume correction. J Chem Theory Comput 3:156–169. https://doi.org/10.1021/ct600085e
40. Valdés-Tresanco MS, Valdés-Tresanco ME, Valiente PA, Moreno E (2021) gmx_MMPBSA: a new tool to perform end-state free energy calculations with GROMACS. J Chem Theory Comput 17:6281–6291. https://doi.org/10.1021/acs.jctc.1c00645
41. Wang Z, Pan H, Sun H, Kang Y, Liu H, Cao D, Hou T (2022) fastDRH: a webserver to predict and analyze protein–ligand complexes based on molecular docking and MM/PB(GB)SA computation. Brief Bioinform 23:bbac201. https://doi.org/10.1093/bib/bbac201
42 Kuriata A, Gierut AM, Oleniecki T, Ciemny MP, Kolinski A, Kurcinski M, Kmiecik S (2018) CABS-flex 2.0: a web server for fast simulations of flexibility of protein structures. Nucleic Acids Res 46:W338–W343. https://doi.org/10.1093/nar/gky356
43. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (2012) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 64:4–17. https://doi.org/10.1016/j.addr.2012.09.019
44. Pires DEV, Blundell TL, Ascher DB (2015) pkCSM: predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. J Med Chem 58:4066–4072. https://doi.org/10.1021/acs.jmedchem.5b00104
45. Tian H, Ketkar R, Tao P (2022) ADMETboost: a web server for accurate ADMET prediction. J Mol Model 28:408. https://doi.org/10.1007/s00894-022-05373-8
46. Biovia DS (2021) Discovery studio visualizer v21. 1.0. 20298, San Diego: Dassault Systèmes
47. Hanwell MD, Curtis DE, Lonie DC, Vandermeersch T, Zurek E, Hutchison GR (2012) Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. J Cheminform 4:17. https://doi.org/10.1186/1758-2946-4-17
48. T. Williams, C. Kelley, C. Bersch, H.-B. Bröker, J. Campbell, R. Cunningham, D. Denholm, G. Elber, R. Fearick, C. Grammes, gnuplot 5.2, An interactive plotting program. Available Online: http://www.Gnuplot.Info/Docs_5.2 (2017).
49. P. Turner, V. XMGRACE, 5.1. 19, Center for Coastal and Land-Margin Research, Oregon Graduate Institute of Science and Technology, Beaverton, OR. 2 (2005).
50. Ren P, Shang W, Yin W, Ge H, Wang L, Zhang X, Li B, Li H, Xu Y, Xu EH, Jiang H, Zhu L, Zhang L, Bai F (2022) A multi-targeting drug design strategy for identifying potent anti-SARS-CoV-2 inhibitors. Acta Pharmacol Sin 43:483–493. https://doi.org/10.1038/s41401-021-00668-7
51. Muhseen ZT, Hameed AR, Al-Hasani HMH, Ahmad S, Li G (2021) Computational determination of potential multiprotein targeting natural compounds for rational drug design against SARS-COV-2. Molecules 26:674. https://doi.org/10.3390/molecules26030674
52. Chauhan A, Avti P, Shekhar N, Prajapat M, Sarma P, Bhattacharyya A, Kumar S, Kaur H, Prakash A, Medhi B (2021) Structural and conformational analysis of SARS CoV 2 N-CTD revealing monomeric and dimeric active sites during the RNA-binding and stabilization: insights towards potential inhibitors for N-CTD. Comput Biol Med 134:104495. https://doi.org/10.1016/j.compbiomed.2021.104495
53. Dhankhar P, Dalal V, Singh V, Tomar S, Kumar P (2022) Computational guided identification of novel potent inhibitors of N-terminal domain of nucleocapsid protein of severe acute respiratory syndrome coronavirus 2. J Biomol Struct Dyn 40:4084–4099. https://doi.org/10.1080/07391102.2020.1852968
54. Rao TV (2022) Screening of SARS-CoV-2 nucleocapsid (N) protein inhibitors as potential drugs for Sars-Cov-2. Biomed J Sci Tech Res. https://doi.org/10.26717/BJSTR.2022.42.006779
55. Ribeiro-Filho HV, Jara GE, Batista FAH, Schleder GR, Costa Tonoli CC, Soprano AS, Guimarães SL, Borges AC, Cassago A, Bajgelman MC (2022) Structural dynamics of SARS-CoV-2 nucleocapsid protein induced by RNA binding. PLOS Comput Biol 18:e1010121. https://doi.org/10.1371/journal.pcbi.1010121
56. Singh VK, Chaurasia H, Kumari P, Som A, Mishra R, Srivastava R, Naaz F, Singh A, Singh RK (2022) Design, synthesis, and molecular dynamics simulation studies of quinoline derivatives as protease inhibitors against SARS-CoV-2. J Biomol Struct Dyn 40:10519–10542. https://doi.org/10.1080/07391102.2021.1946716
57. Nguyen PTV, Nguyen GLT, Thi Đinh O, Duong CQ, Nguyen LH, Truong TN (2022) In search of suitable protein targets for anti-malarial and anti-dengue drug discovery. J Mol Struct 1256:132520. https://doi.org/10.1016/j.molstruc.2022.132520
58. Luan X, Li X, Li Y, Su G, Yin W, Jiang Y, Xu N, Wang F, Cheng W, Jin Y (2022) Antiviral drug design based on structural insights into the N-terminal domain and C-terminal domain of the SARS-CoV-2 nucleocapsid protein. Sci Bull 67:2327–2335. https://doi.org/10.1016/j.scib.2022.10.021
59. Kaushik U, Aeri V, Mir S (2015) Cucurbitacins: an insight into medicinal leads from nature. Pharmacogn Rev 9:12. https://doi.org/10.4103/0973-7847.156314
60. Qu Y, Cong P, Lin C, Deng Y, Li-Ling J, Zhang M (2017) Inhibition of paclitaxel resistance and apoptosis induction by cucurbitacin B in ovarian carcinoma cells. Oncol Lett 14:145–152. https://doi.org/10.3892/ol.2017.6148
61. Njoroge GN, Newton LE (1994) Edible and poisonous species of Cucurbitaceae in the Central Highlands of Kenya. J East African Nat Hist 83:101–115. https://doi.org/10.2982/0012-8317(1994)83[101:EAPSOC]2.0.CO;2

Publisher's Note