Identification of potential 3CLpro inhibitors-modulators for human norovirus infections: An advanced virtual screening approach

Shovonlal Bhowmick
SilicoScientia Private Limited
Tapan Kumar Mistri
SRM Institute of Science and Technology
Mohammad K. Okla
King Saud University
Ibrahim A. Saleh
Zarqa University
Hamada AbdElgawad
University of Antwerp
Achintya Saha
University of Calcutta
Pritee Chunarkar Patil (preeti.chunarkar@bharatividyapeeth.edu)
Bharati Vidyapeeth Deemed University

Research Article

Keywords: Norovirus, 3CLPro, Molecular docking, Virtual screening, Molecular dynamics, Free energy landscape

Posted Date: November 17th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3614758/v1

Additional Declarations: No competing interests reported.

Identification of potential 3CLpro inhibitors-modulators for human norovirus infections: An advanced virtual screening approach

Shovonlal Bhowmick[1,2#], Tapan Kumar Mistri[3#], Mohammad K. Okla[4], Ibrahim A. Saleh[5], Hamada AbdElgawad[6], Achintya Saha[1**], Pritee Chunarkar Patil[7*]

[1]*Department of Chemical Technology, University of Calcutta, 92, A.P.C. Road, Kolkata, 700009, India*
[2]*SilicoScientia Private Limited, Nagananda Commercial Complex, No. 07/3, 15/1, 18th Main Road, Jayanagar 9th Block, Bengaluru – 56004*
[3]*SRM Institute of Science and Technology, Kattankulathur Campus, SRM Nagar, Potheri, Chennai-603203, India*
[4]*Botany and Microbiology Department, College of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia*
[5]*Faculty of Science, Zarqa University, Zarqa 13110, Jordan*
[6]*Integrated Molecular Plant Physiology Research, Department of Biology, University of Antwerp, 2020 Antwerp, Belgium*
[7]*Department of Bioinformatics, Rajiv Gandhi Institute of IT and Biotechnology, Bharati Vidyapeeth Deemed University, Pune-Satara Road, Pune, India*
[#]Contributed equally, and both can be considered as 1st author

Corresponding authors:
*Pritee Chunarkar Patil (Email: preeti.chunarkar@bharatividyapeeth.edu)
**Achintya Saha (Email: achintya_saha@yahoo.com)

Abstract

The present study aimed to screen small molecular compounds as the human noroviruses (HuNoV) inhibitors/modulators that could be potentially responsible for exhibiting some level of inhibitory activity against HuNoV 3CLPro. The structural similarity-based screening against ChEMBL database is performed against known chemical entities which are presently under pre-clinical trial. Molecules that remained after the similarity search were considered molecular docking using SCORCH and PLANTS. On detailed analyses and comparisons with control molecule, 3 hits (CHEMBL393820, CHEMBL2028556 and CHEMBL3747799) were found to be potential for HuNoV 3CLpro inhibition. The binding interaction analysis revealed several critical amino acids to hold the molecules tightly at the close proximity site of the catalytic residues. Further, three MD simulation study was performed in triplicate to understand the binding stability and potentiality of the proposed molecule towards HuNov 3CLpro. The

binding free energy based on MM-GBSA has revealed their strong interaction affinity with 3CLpro.

Keywords: Norovirus; 3CLPro, Molecular docking; Virtual screening; Molecular dynamics; Free energy landscape

Statements and Declarations

Competing Interests

The authors declare no competing interests.

1. Introduction

The human noroviruses (HuNoV) belong to the Caliciviridae family [1], and are subdivided into seven genogroups (GI to GVII) [2]. HuNoV is one of the major reasons for causing gastroenteritis worldwide [3, 4]. HuNoV infections are considered to be the leading cause of non-bacterial gastroenteritis across many nations, making high mortality and morbidity rates [5]. All age groups are susceptible to getting the HuNoV infection, however, children and the elderly are much more prone to get infection easily and also immunocompromised patients, and hence more likely to develop severe symptoms and so on priority wise are groups of interest for immunization [3, 5]. Fecal-oral transmission, ingesting tainted food or drink, and person-to-person contact are all common ways of spreading HuNoV infections [6]. HuNoV is thought to be accountable for ~699 million illnesses and ~219,000 fatalities annually in the world [7]. According to the estimated During 2013, it was estimated that the disease burden brought on by HuNoV infection causes approximately 1.7–1.9 million ambulatory visits to health centers, 400,000 visits to the emergency room, 56,000 – 71,000 hospitalizations, and 570–800 deaths, and accumulating overall 19–21 million total diseases each year, in a developed country like United States alone [5, 8]. Despite such a devastating and life-taking disease, there are currently no specific approved medicines or chemical compounds or vaccines available for the either treatment or control the HuNoV infection [3, 5]. There are several challenegs in developing potent small drug like molecules or vaccines agasint such virus. Particularly, because of the norovirus's environmental stability, low infectious dosage and high shedding titer, it is very challenging to prevent and control infections with this norovirus [9]. The development of a potent norovirus drug like candidate or vaccine may encounter a number of obstacles. Majorly due to challenge in cultivation of noroviruses has seriously hindered the smooth development of effective and durable vaccinations and other therapeutic options [9].

The NoV encompasses a genetically diverse group of nonenveloped positive-strand RNA viruses that infect many host species including humans and enterics in nature [6]. The genome of HuNoV is comprised of three open reading frames (ORFs) *viz.* ORF1, ORF2 and ORF3 [10]. ORF1 is consisting of non-structural polyproteins which are usually get degraded by the 3C-like protease (3CLpro) enzyme and broken into at least 6 different proteins, and these are crucial for viral replication [11]. ORF2 represents the major viral capsid protein i.e. VP1, whereas, ORF3 encodes the minor capsid protein, VP2 [11]. Both ORF2 and ORF3 are responsible for maintaining the virus structure. The polyprotein is processed by the virally encoded 3CLpro. 3CLpro is consisting of cysteine protease having Cys139-His30-Glu54 catalytic triad which represents to be the catalytic binding cleft [12], and a major substrate

selectivity for a P1 glutamine (or glutamate) residue. The prototypical catalytic triad of HuNoV 3CLPro is located at the interface of a β-sheet domain and a β-barrel domain [13]. Moreover, this protease act as an induced fit enzyme and has an extended binding cleft [14-16]. The substrate specificity of the protease is for a P1 Gln residue (or Gln surrogate) that engages in critical H-bonding interactions with Thr134 and His157 located in close proximity to the active site [17]. As a critically important protease enzyme involve in the life cycle of the norovirus function, the 3CLpro breaks down the lengthy polypeptide chain from viral RNA's ORF1 that cleaves at five sites to release several non-structural proteins [18]. Moreover, it is reported that 3CLpro of HuNoV is an important and critical protein for the virus life cycle which projected as an ideal and effective drug target for the development of potential and therapeutic chemical components for the HuNoV infection [19, 20].

Pharmacoinformatics which represents the computational approach in drug discovery became pivotal and pioneer methodologies to design and identify the potential and effective chemical compounds for any specific drug target [21, 22]. Similarity search of chemical repositories by known or approved molecular fingerprints is widely accepted and used by the scientific community for therapeutic developments for numerous disease models including infectious diseases [23-25]. Moreover, another highly regarded and widely used technique like molecular docking assisted by machine learning (ML) enables to predict and identify the correct binding pose or orientation of the molecules at the active site of the receptor cavity [26-28]. The interaction dynamicity of any small molecules inside the protein target assessed through molecular dynamics (MD) simulation can provide insight into the potentiality and effectivity of the molecules [29, 30]. In particular, some important functional processes are difficult to address experimentally, such as protein-ligand binding orientation in dynamic motion, ligand-induced conformational change in the protein, and the nature of protein folding in the dynamic state may infer through MD simulation analysis. Hence, in a similar notion, the main present work intended to similarity search of the ChEMBL database which contains about 2 million compounds using few known HuNoV inhibitors followed by ML based molecular docking and MD simulation studies to find effective chemical compounds for the management of HuNoV infection. Overall, the present study finding employed through above mentioned techniques have greatly illuminated the understanding of binding interactions of newly identified screened compounds for exploring an array of conceptually-sound approaches toward the development of anti-norovirus therapeutics targeting the HuNoV 3CLpro by means of either inducing or inhibiting or modulating the conventional biological functions of the studied enzyme.

2. Materials and methods

2.1 Fingerprint based similarity search against the ChEMBL database

A similarity search of any chemical database based on known or approved molecules is an important application in drug discovery research. This approach gives the ability to scan chemical compounds to identify similar molecules from the databases to the given query compounds. In order to find the potential 3CLpro inhibitors from the ChEMBL database [31, 32], a set of five chemical compounds (2′-C-methylcytidine, Nitazoxanide, Valopicitabine, Suramin and PPNDs) was considered as query molecules. In particular, the above set of molecules is reported to be in pre-clinical trials for HuNoV infection. A two-dimensional representation of the query molecules is given in Figure S1 (supplementary file). Therefore, any new molecules from the ChEMBL database similar to the above molecules might be crucial and important to manage the HuNoV infection. For the similarity search, extended-connectivity fingerprints-4 (ECFP4) was used. Molecular fingerprints are the depiction of chemical compounds and are widely used for similarity searching, clustering, and classification. The ECFPs are developed to apprehend the molecular features related to the activity of the chemical compound. The simplified molecular input line entry system (SMILES) form of query molecules was collected from the PubChem database. As a target dataset, the entire ChEMBL database was downloaded in SMILES format. The molecules in ChEMBL are manually curated and publicly available bioactive molecules that possess drug-like characteristics. For the similarity search, the Python RDKit was used. Python RDKit is an open-source toolkit for cheminformatics consisting of a number of packages for chemical data analysis. After a similarity search of the ChEMBL database using the five query molecules, all molecules of the target database were arranged according to the ascending order of the Tanimoto coefficient. Molecules found to have a Tanimoto coefficient of more than or equal to 0.6 were considered for further analyses.

2.3 Molecular docking simulation

Molecular docking has already been recognized as a prototype of structure-based virtual screening [33-46]. In this approach, potential molecules are identified from large chemical databases and binding interaction affinity is predicted toward the protein targets of interest. The traditional molecular docking approach and scoring functions (SFs) are not comparable with experimental data, hence, a number of improvements are done using the ML approaches. Scoring COnsensus for RMSD-based Classification of Hits (SCORCH) [47], a novel ML scoring function for molecular docking was used to dock the HuNoV 3CLpro target protein

with most structurally similar small molecule compounds screened against the ChEMBL database [31, 32]. The scoring function gives the strength to the docking engines to identify docked orientation that might close resembles a co-crystal ligand pose [47]. The SCORCH program implemented the multiple RMSD labeled docked poses to strengthen the scoring function which is one of the unique and powerful screening strategies. Another docking program, the Protein-Ligand ANT System (PLANTS) [48] is one of the important tools used for molecular docking of identified screened small molecules to the protein target. PLANTS is one of the crucial tools to screen the extra large chemical databases followed by optimization. The stochastic optimization algorithm is called ant colony optimization (ACO) and is used by the PLANTS program. For ligand docking in the active site or any defined site of the protein molecule, an ACO is employed to find a minimum energy conformation of the ligand in the binding site.

In the present study, in particular, for docking execution, both the above mentioned two highly regarded docking programs were implemented. Therefore, the crystal structure of norovirus 3CLpro was obtained from RCSB-PDB [49] having the PDB ID: 5T6F [50]. The above protein structure was found to have a resolution of 1.9Å. The protein structure was prepared by the SCORCH with the help of Autodock tools through the addition of hydrogens and Gasteiger charges. The co-crystal water and other hetero atoms were deleted during protein preparation. Any missing atoms were identified and repaired. Prior to save in .pdbqt format, the type of the atom was assigned to AD4 to the protein structure, as the mandate required for ADV for docking execution. The coordinates (5.677, 65.623 and -7.117 along the x-, y- and z-axes, respectively) of the co-crystal ligand were considered to be the active site for molecular docking. The grid box size was considered to be 40x40x40 Å along x-, y- and z-axes, respectively. The same coordinate was used for docking in SCORCH and PLANTS to dock all the screened ChEMBL molecules inside the HuNoV 3CLpro. The molecules were found after the similarity search prepared for molecular docking following the need for the basic docking principle. The polar hydrogens and Gasteiger charges were added and finally, saved into .pdbqt and .mol2 format for the molecular docking. The binding interactions analysis was explored through Protein-Ligand Interaction Profiler (PLIP), an online web server and it is available at https://plip-tool.biotec.tu-dresden.de/plip-web/plip/index.

2.4 In-silico pharmacokinetics and toxicity analysis

The 'pkCSM', a publicly available web server based *in silico* pharmacokinetics and toxicity properties prediction tool, was used to evaluate the various drug-likeliness parameters of three

identified chemical entities [51]. The highly accepted program 'pkCSM' for *in-silico* drug-likeliness, toxicity and pharmacokinetics-based properties prediction analysis is accessible at http://biosig.unimelb.edu.au/pkcsm/. In order to construct a good prediction model of various pharmacokinetics and toxicity features, the 'pkCSM', in particular, depends on graph-based signatures and mathematical or statistical illustrations of every given substance that are usually validated through highly predictive regression and classification model [51]. Several parameters including a blood-brain barrier (BBB), central nervous system (CNS) permeability, total clearance, AMES and minnow toxicity, skin sensitivity and hepatoxicity were predicted for the proposed molecules. BBB explains the permeability potentiality of the molecule in the brain. It is reported that a low value (< -1) of BBB is poorly distributed in the brain. Similar to the BBB, CNS permeability gives the idea of the penetration of molecules inside the brain. A CNS permeability value of -3 indicates that molecules are unable to penetrate the CNS. The mutagenicity of the molecules can be explored through AMES toxicity. The minnow toxicity is an important parameter to explain the molecules necessary to cause of death of the Flathead Minnows. The hepatotoxicity value of any molecule indicates whether it disrupted the normal function of the liver. Skin sensitization can be explored with skin sensitivity parameters.

2.5 Molecular dynamics simulation

The MD simulation is an essential approach to understand the dynamic nature of the protein-ligand complex. The final proposed molecules and PPNDs bound with 3CLpro were considered for the MD simulation for 100ns of time span. For better insight exploration, the MD simulation of each complex was repeated three times. The MD simulation was carried out in Gromacs2021.1 [52] and installed in a Linux operating system environment with a system configuration of 10th Generation Intel Core i9-10885H and NVIDIA® GeForce RTXTM 2070. The protein topology was generated using the CHARMM36 force field [53], whereas, the ligand topology was developed from SwissParam [54] web server. The entire simulation was carried out with a time step, constant pressure, and constant temperature of 2 fs, 1 atm, and 300 K, respectively. Each of the complexes was immersed in a cubic box system with a minimum distance of 10 Å from the center to the box edge and solvated by the TIP3P water model [55]. The systems were neutralized by the addition of a sufficient number of Na^+/Cl^-. To address the overlap and close contacts between the atoms, the systems were minimized by the steepest-descent algorithm. Further, all prior to transferring the system to the production phase, each system was equilibrated through NVT (constant number of particles, volume, and temperature) followed by NPT (constant number of particles, pressure, and temperature). Finally, the MD

simulation trajectories were analysed to calculate a number of statistical parameters including backbone RMSD, ligand RMSD, root-mean-square fluctuation (RMSF), the radius of gyration (RoG), and the number of inter-molecular H-bond interactions.

2.6 Principal component analysis (PCA) and free energy landscape (FEL) analysis

Free energy landscape (FEL) [56-58] is an important aspect to explore the folding and unfolding of the protein in dynamic states through the assessment of their basins. The relative stability of 3CLpro bound with proposed and PPNDs molecules was explored by the FEL approach. In this regard, the principal component analyses (PCA) [56] was performed for each of the complexes on C-α atoms of 3CLpro. The PCA is known to be crucial for the motions managing the molecular conformational movements throughout the MD simulation [59]. The final FEL for each molecule of all three rounds of MD simulations was calculated by estimating joint probability distribution from the essential plane made by the top two eigenvectors. It is reported that among all eigenvectors, the first few or two signify the most dominant collective motions providing a useful description of sub-conformational structural transitions.

2.7 Binding free energy estimation using MM-GBSA

The MD simulation trajectories were considered to explore binding affinity through the calculation of binding free energy (ΔG_{bind}) using the MM-GBSA approach. This approach has already been considered one of the trusted strategies to explore the binding affinity of any small molecule of interest. From the MD simulation trajectories, the (ΔG_{bind}) of each of the final HuNoV inhibitors along with PPNDs was calculated using the gmx_MMPBSA module [60]. Out of 100000, a total of 2000 frames from the entire trajectory were considered for (ΔG_{bind}) estimation. The stepwise calculation of ΔG_{bind} is given below.

$$\Delta G_{bind} = \left\langle G_{complex} \right\rangle - \left\langle G_{receptor} \right\rangle - \left\langle G_{ligand} \right\rangle \quad (1)$$

In the above expression, $G_{complex}$, $G_{receptor}$ and G_{ligand} can be recognized as the binding energy of protein-ligand complex, receptor and ligand, correspondingly.

Further, ΔG_{bind} may be denoted as

$$\Delta G_{bind} = \Delta H - T\Delta S \quad (2)$$

Where, ΔH signifies the enthalpy of binding, and, $T\Delta S$ indicates the conformational entropy after ligand binding. After the elimination of the entropic term, the anticipated value represents the effective free energy [61, 62].

From equation (2), the ΔH can be divided into simpler form as given below.

$$\Delta H = \Delta E_{MM} + \Delta G_{sol} \quad (3)$$

Here, ΔE_{MM} may be further can be stated as a total of bonded and non-bonded terms.

$$\Delta E_{MM} = \Delta E_{bonded} + \Delta E_{nonbonded} \quad (4)$$

E_{bonded} is the addition of bond stretching, and, angle bending and torsion angle. The combined electrostatic and van der Waals' terms may be expressed as $\Delta E_{nonbonded}$ Both equations are given below.

$$\Delta E_{bonded} = \Delta E_{bond\_length} + \Delta E_{angle} + \Delta E_{dihedral} \quad (5)$$

$$\Delta E_{nonbonded} = \Delta E_{ele} + \Delta E_{vdW} \quad (6)$$

The polar constituent is only present in the solvation energy (ΔG_{sol}) for GB models. The nonpolar (NP) constituent is mostly thought to be proportional to the molecule's total solvent-accessible surface area (SASA). The equations of solvation and non-polar energy can be given below.

$$\Delta G_{sol} = \Delta G_{polar} + \Delta G_{non\text{-}polar} = \Delta G_{GB} + \Delta G_{non\text{-}polar} \quad (7)$$

$$\Delta G_{non\text{-}polar} = NP_{TENSION} + \Delta_{SASA} + NP_{OFFSET} \quad (8)$$

3. Results and discussion

Extensive computational approaches were considered to screen the entire ChEMBL database through ligand-based similarity search, molecular docking, pharmacokinetic and toxicity assessment, and, MD simulation studies. Ligand-based similarity search and screening using molecular docking become the favorite choice to find promising molecules for a given target. The validation of the molecular docking protocol is an essential step prior to using it for the screening of molecular databases. The self-docking approach was considered to validate the current protocol. In this method, the co-crystal bound ligand was re-drawn and it docked at the active site where the co-crystal ligand was originally bound. The best pose after docking is needed to superimpose the co-crystal ligand. If the RMSD between the best pose and co-crystal ligand is to be found < 2Å, it is considered to be the docking protocol will generate the ligand conformation as similar to crystallization. The considered 3CLpro crystal structure consists of Suramin as a co-crystal bound ligand. Suramin was redrawn and docked by replacing the co-crystal bound conformation. A number of best poses from different combinations of docking protocols were extracted and superimposed on the original co-crystal conformation of Suramin and RMSD was recorded. The best-superimposed conformation was revealed with an RMSD of 0.956Å which suggest the validation of the docking protocol and the same protocol was

considered for the screening of the ChEMBL database. The superimposed structure is given in Figure 1.

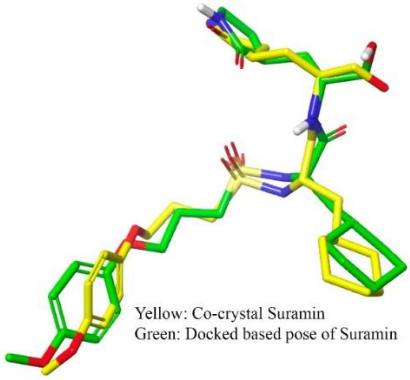

Figure 1. Superimposition of the co-crystal and best-docked pose of Surmin

3.1 Fingerprint similarity-based virtual screening for potential inhibitor/modulator identification

The employment of advanced cheminformatics knowledge makes it possible to use and comprehend diversity by defining molecules as molecular fingerprints and vectorizing their structural properties [63-65]. Those molecular fingerprints can be used to quickly compare similarities and serve as the foundation for research on the relationships between structure and activity, virtual screening, and the creation of chemical space maps [23, 66]. Screening of a large number of molecules using the molecular fingerprinting similarity search become an essential and effective tool for computational drug discovery [67]. In a similar notion, the present study intended to find some promising molecules for inhibiting or modulating the activity of HuNoV 3CLpro protein. The complete workflow of the employed study has been depicted in Figure 2.

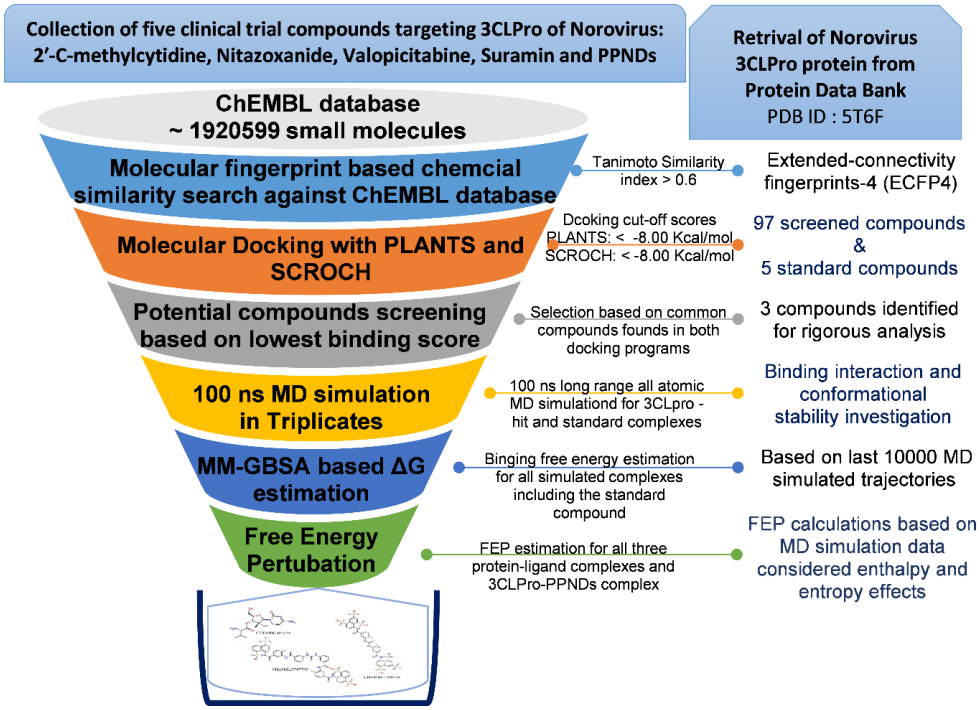

Figure 2. Computational workflow for identification of potential inhibitors/modulators for HuNoV 3CLpro protein based on fingerprint similarity search, molecular docking and MD-simulation and FEL studies

In particular, the ChEMBL chemical database was screened through molecular fingerprints against five known potential chemical entities for managing the HuNoV infection. Five chemical entities namely Nitazoxanide [68], Methylcytidine, Valopicitabine, PPNDs and Suramin are currently under pre-clinical trials for the HuNoV infection treatment [69]. All five molecules mentioned above were considered to be standard compounds and taken as the query compounds for the similarity search against the ChEMBL database. The SMILES representation of the five standard molecules was collected from PubChem. ECFP4 fingerprint [70] is available in Python RDKit and it was used for similarity search against ChEMBL molecules. On successful completion of the similarity search, all molecules were arranged according to the Tanimoto coefficient. Molecules found to have a Tanimoto coefficient of 0.6 and above were considered for further assessment. A total of 97 molecules were found to have a Tanimoto coefficient of 0.6 and above. A list of the above 97 molecules is given in Table S1 (Supplementary data). Further, the above 97 molecules were considered for molecular docking study using two docking engine tools such as ML based ADV implemented in SCORCH [47] and PLANTS [48]. The binding energies from SCORCH and PLANTS of each of the molecules were recorded and explored, and, it is given in Table S1 (Supplementary data). Both the binding energies were arranged in descending order and the top ten molecules from each set were collected. After analyzing binding energies from both docking programs, it was found

three molecules retained or found as common in both docking programs. In particular, three molecules (CHEMBL393820, CHEMBL2028556 and CHEMBL3747799) were found to be common in the top ten molecules from SCORCH and PLANTS docking programs. Hence, the common three molecules from both docking studies were considered to be the best promising molecules for HuNoV 3CLpro inhibitors/modulators and subsequently more extensive and rigorous computational analyses were conducted on them. Moreover, similar types of analysis have also been employed for standard PPNDs and used for comparison purposes. Two dimensional (2D) chemical representation of the selected three hit molecules is depicted in Figure 3.

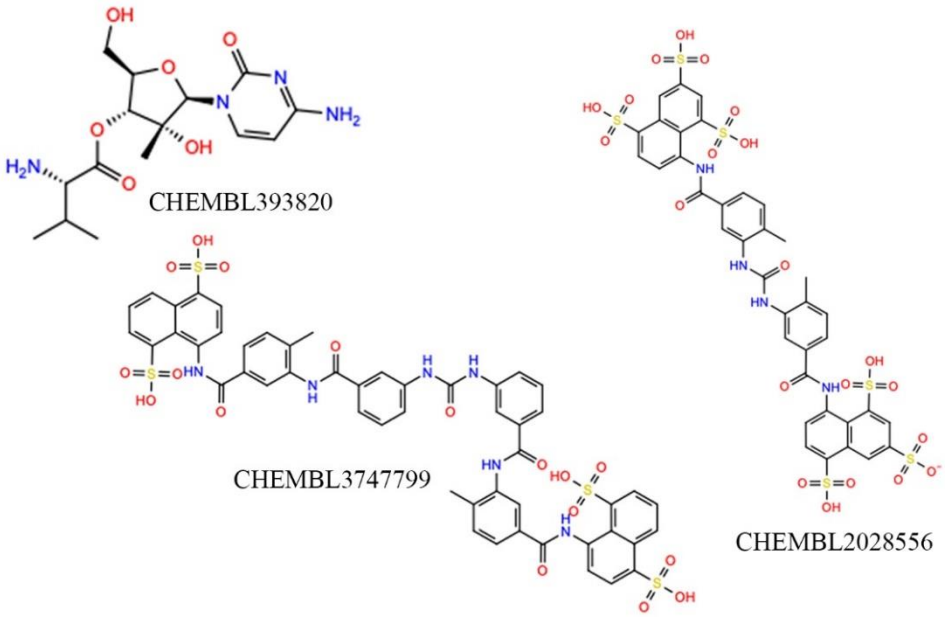

Figure 3. Two-dimensional representation of selected 3CLpro molecules

3.2 Selection of control molecule

To select the control molecule for the comparison of the data, considered five query molecules were docked in both SCORCH integrated via Vina and PLANTS engines with similar parameters as ChEMBL molecules docked. The binding energy of each molecule from both docking engines was collected and it is given in Table 1. All the molecules were shown negative binding affinity which suggests the strong affinity of each molecule towards the 3CLpro. On close observations, it can be seen that among all five, PPNDs was found to have a better affinity towards 3CLpro. Hence, PPNDs was considered to be the standard control molecule in the current study.

Table 1. The binding energy of standard query molecules

Molecule	Binding energy (kcal/mol)	
	SCORCH	PLANTS
2′-C-methylcytidine	-5.305	-63.544
Nitazoxamine	-6.575	-56.240
Valopicitabine	-6.394	-74.532
Suramin	-9.125	-48.503
PPNDs	-8.001	-80.499

3.3 Molecular interaction profile analysis through docking simulation

A molecular docking approach was employed to investigate the possible intermolecular binding interaction formed between the HuNoV 3CLpro and screened-out compounds from the ChEMBL database based on most structural similar characteristics with standard five compounds. Two different docking tools namely SCORCH and PLANTS - standalone programs were used for docking execution with compounds showing > 6.0 Tanimoto similarity search score, which resulted in the selection of three best potential inhibitors/modulators *viz.* CHEMBL393820, CHEMBL2028556, and CHEMBL3747799.

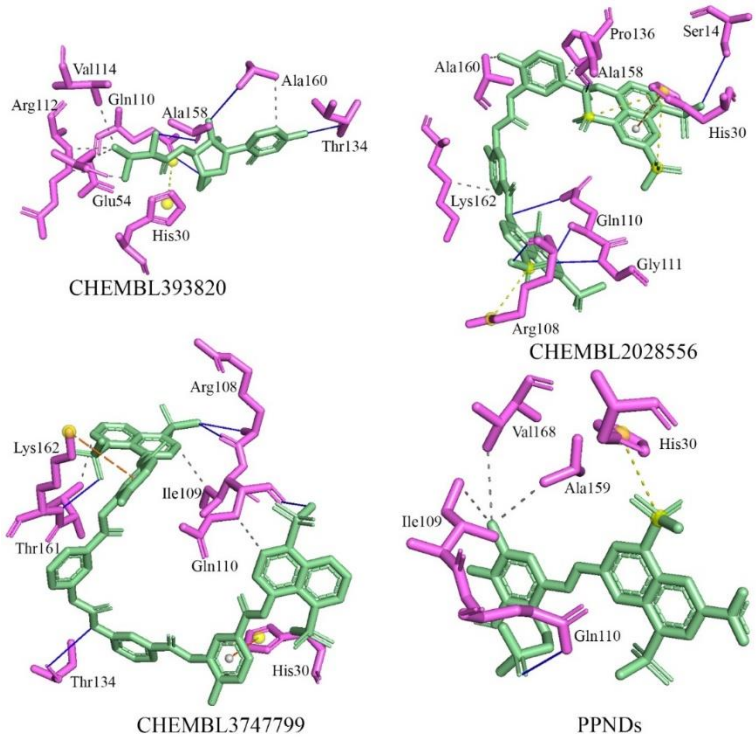

Figure 4. Intermolecular binding interactions of the proposed identified three inhibitors/modulators with HuNoV 3CLpro. A number of important ligand-binding amino acid residues of HuNoV 3CLpro protein are labeled in pink color. H-bond interactions, hydrophobic contacts, salt bridge and π-cation interactions are highlighted in blue, dash grey, dash yellow, and dash dark orange color, respectively.

The molecular docking simulation based binding interaction profile of the top three identified hit compounds and standard PPNDs has been depicted in Figure 4. The binding conformation of the lowest docking score of CHEMBL393820 establishes four numbers of H-bond interaction with different residues (Gln110, Thr134, Ala158 and Ala160) of 3CLpro protein. Moreover, a number of hydrophobic contacts were found for CHEMBL393820 with residues Glu54, Gln110, Arg112, Val114, and Ala160 of 3CLpro protein. In addition, a salt bridge interaction was also noticed between residue His30 and CHEMBL393820. Among all the identified final hit compounds, CHEMBL2028556 establishes a maximum number of intermolecular interaction formations between several residues of 3CLpro and revealed a total of 6 H-bonds, 3 hydrophobic contacts, 3 salt bridge interactions, and 1 π-cation interaction (Figure 4). Amino acid residues Ser14, Arg108, Gln110, Gly111, and Ala158 of HuNoV 3CLPro formed H-bond interaction with CHEMBL2028556. Hydrophobic contacts was established with residues Ala160, Lys162 and Pro136, and, CHEMBL2028556. Two basic amino acid residues His30 and Arg108 were found to form salt-bridge interaction with CHEMBL2028556. Moreover, π-cation interaction was also noted in association with the involvement of residue His30. The identified hit compound CHEMBL3747799 was found to form five numbers of H-bond interaction with three consecutive amino acid residues Arg108, Ile109, and Gln110 (Figure 4). Two other amino acid residues Thr134 and Thr161 also established H-bond interaction with novel hit CHEMBL3747799. Residues Ile109, Gln110 and Thr161 established hydrophobic contacts with CHEMBL3747799. Two basic amino acid residues (His30 and Lys162) were found to create two numbers of π-cation interaction with CHEMBL3747799. Furthermore, to get a better insight into the 3CLPro protein interacting with the standard compound PPNDs, the critical analysis revealed a comparatively less number of H-bond interactions with the amino acid residue of 3CLpro (Figure 4). Especially, one residue Gln110 engagement was found for establishing H-bond interaction with PPNDs. In addition to H-bond interaction, three numbers of hydrophobic contacts were also found for considered standard PPNDs with hydrophobic residues Ile109, Ala159, and Val168. A salt bridge interaction was also formed between residue His30 and PPNDs. Moreover, for all compounds, observed H-bond distances were found within the range of 1.45 - 3.06 Å, whereas hydrophobic distances were measured within 1.69 - 3.29 Å.

Interestingly, two residues Gln110 and His30 were found to be common amino acid residues that participated in intermolecular interaction formation, as observed from critical interaction analysis obtained through the PLIP program. Undoubtedly, amino acid residue His30 is one of

the important catalytic residues of Norovirus 3CLpro protein [71], and such residue involvement in interaction association with novel identified hit compounds may deliver protease inhibitory potential.

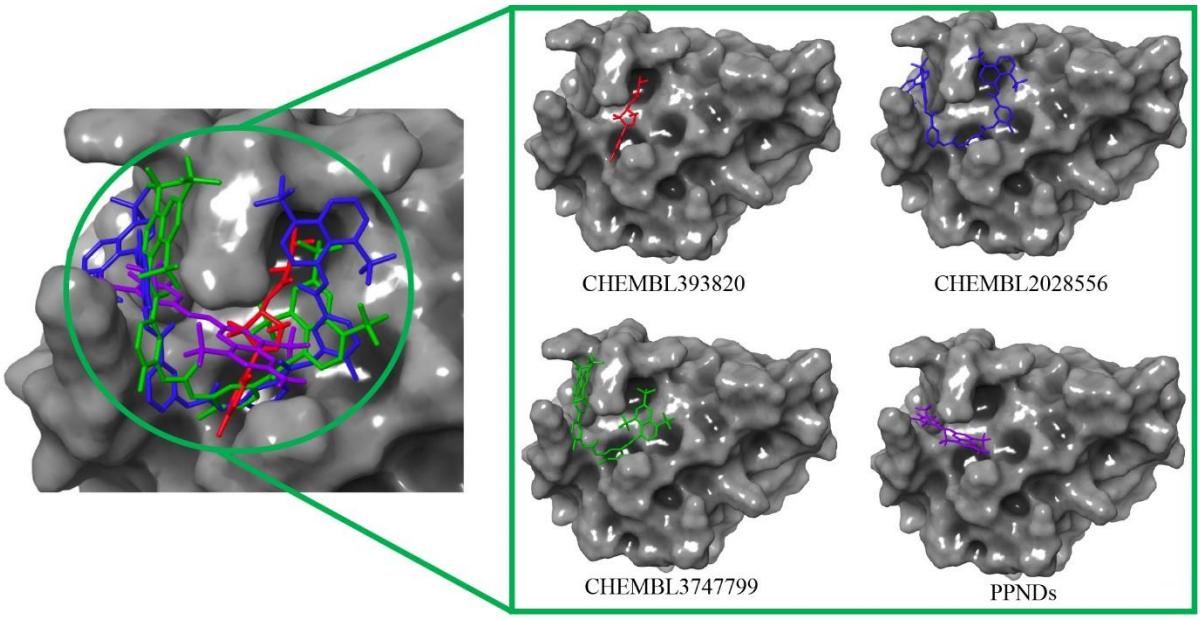

Figure 5. Binding orientation of the proposed molecules inside the 3CLpro receptor cavity

Overall, binding orientations or interaction modes of the identified three hit compounds revealed a relatively similar kind of interaction profile to that of the standard compound PPNDs, at the close conjunction of the catalytic site or pocket of the HuNoV 3CLPro (Figure 5). Molecular docking simulations were performed to investigate the potentiality of small drug-like candidate molecules, which can engage at the catalytic site of 3CLpro of HuNoV site and/or close the vicinal cavity in a way that would disrupt the inhibition process of the 3CLpro, leading to the modulation of bio physicochemical interactions upon binding of novel hit compounds.

Moreover, according to previously published work, H-bond interactions between HuNoV 3CLpro and the studied ‘compound 8B’ shows the association of amino acid residues Thr134, His157, Ala158 and Ala160 of HuNoV 3CLpro [17]. Such evidence is strongly adjudged with the potentiality of the identified screen compounds for showing similar types of H-bonding interactions, and hence could possibly serve as potent inhibitors/modulators of HuNoV 3CLpro. Another study from the same research group, also highlighted the H-bonds interaction network involving the backbone of the studied inhibitor ‘compound 13’ and amino acid residues Gln110, Ala158, and Ala160 of HuNoV 3CLpro that serve to position the inhibitor correctly at the active site clearly evident [17]. Surprisingly, our present study also revealed

similar types of intermolecular interaction affinity for most of the identified compounds (CHEMBL3747799 and CHEMBL2028556) which ordinarily engage in H-bonding interactions at the close vicinity of the catalytic binding cleft of the HuNoV 3CLpro protein.

3.4 Pharmacokinetics and toxicity assessment

The pharmacokinetic and toxicity assessment of CHEMBL393820, CHEMBL2028556 and CHEMBL3747799 were calculated and these are in Table 2. The BBB and CNS permeability of CHEMBL393820, CHEMBL2028556 and CHEMBL3747799 was found to be -1.150 and -3.536, -3.755 and -3.755, and, -3.104 and -4.574, respectively that suggested the low distribution in the brain and CNS. The total clearance of CHEMBL393820, CHEMBL2028556 and CHEMBL3747799 was predicted to be in the range of -3.755 to 0.406. Not a single molecule was found to have AMES, hepatotoxicity and skin sensitivity except CHEMBL393820 shown positive hepatotoxicity, which may need structural optimization for better pharmacokinetic efficacy. The minnow toxicity of the selected molecules was found to be more than 1.560 which undoubtedly revealed that there is no cause of death of Flathead Minnows by administrating these molecules. Overall the data obtained from pkCSM suggested that selected molecules were found to pass almost all the explored pharmacokinetic and toxicity parameters, however, may also need structural modification or optimization for gaining better pharmacological attention.

Table 2. Pharmacokinetics and toxicity parameters of CHEMBL393820, CHEMBL2028556 and CHEMBL3747799

Parameters	CHEMBL393820	CHEMBL2028556	CHEMBL3747799
BBB permeability	-1.150	-3.755	-3.104
CNS permeability	-3.536	-3.755	-4.574
Total Clearance	0.405	-3.755	-1.245
AMES toxicity	No	No	No
Hepatotoxicity	Yes	No	No
Skin Sensitisation	No	No	No
Minnow toxicity	3.797	3.921	1.570

3.5 Molecular dynamics simulation of the 3CLpro and identified ligand complexes

The dynamicity of each selected molecule along with the considered standard compound (PPNDs) bound with 3CLpro was assessed using the triplicate 100ns MD simulation. The relative stability and the deviations from the reference frames were explored through a number of statistical parameters including 3CLpro backbone RMSD, ligand RMSD, RMSF, RoG, number of hydrogen bonds and FEL.

3.5.1 3CLpro backbone RMSD

The protein backbone RMSD is one of the parameters to understand the stability of the protein-ligand complexes during the MD simulation. High deviation from the reference backbone indicates low stability. Steady variations portray the stability and compactness of the complex. From all three rounds of MD simulation, the 3CLpro backbone RMSD was calculated and it is given in Figure 6. For more insightful exploration, the MD simulation was repeated three times and statistical data were calculated for each round of simulation. The average of 3CLpro backbone RMSD from all three rounds of the simulation was calculated and portrayed in Figure 6. It can be seen that in Round #3, the 3CLpro backbone bound with CHEMBL393820 deviated a bit at the beginning and afterward achieved steadiness.

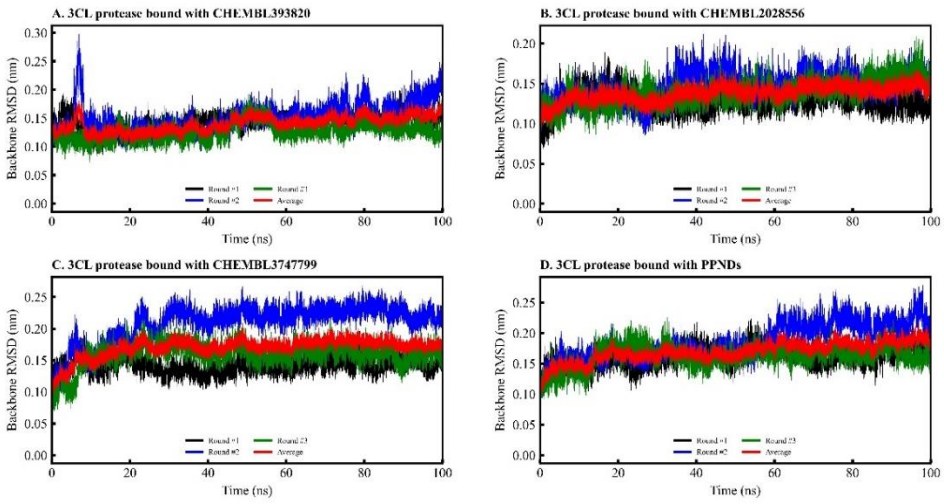

Figure 6. 3CLpro backbone RMSD bound with selected molecules and PPNDs. The red color graph is the average of all three rounds of data.

In the case of 3CLpro backbone bound with CHEMBL3747799, both Round #1 and #2 deviated almost in the same range of RMSD values, while, the backbone in Round #3 was found to increase a bit at the beginning and achieved compactness at around RMSD of 0.20

nm. For three rounds of MD simulation, 3CLpro backbone bound with both CHEMBL2028556 and PPNDS were found to deviate in an almost similar manner. The average backbone RMSD of three rounds of MD simulations for each complex was calculated and found a very compact steady variation throughout the simulation. Hence, it can be postulated that the 3CLpro backbone was quite stable and compact throughout all three rounds of MD simulation.

3.5.2 Ligand RMSD

Deviation of the ligands bound with 3CLpro was checked by calculation of the ligand RMSD from all three rounds of MD simulation and these data are given in Figure 7. The average RMSD from all three rounds of simulation for each complex was calculated and given in Figure 7. It can be observed that CHEMBL393820 and PPNDs were almost compact and stable throughout the simulation with some variation. In the beginning, PNDSs was seen to have a little bit of high RMSD in Round #2 and afterward, it remained consistent. During all three rounds of simulations, CHEMBL2028556 was seen to fluctuate at the beginning and after about 70ns it was found to be equilibrated. Interestingly, CHEMBL3747799 was found to be almost equilibrated for all three rounds of simulation with different RMSD values. The possible reason behind stability achievement with different RMSD might be the starting of the simulation with different velocities. In order to get more insight into the deviation of the molecules, the average RMSD of all three simulations for each molecule was calculated and plotted in Figure 7 (red lines). It can be seen that the average graph clearly was shown consistency and a small deviation from the beginning to the end of the simulation. The above observations undoubtedly indicated that with small exceptions, all molecules remained stable in the dynamic states.

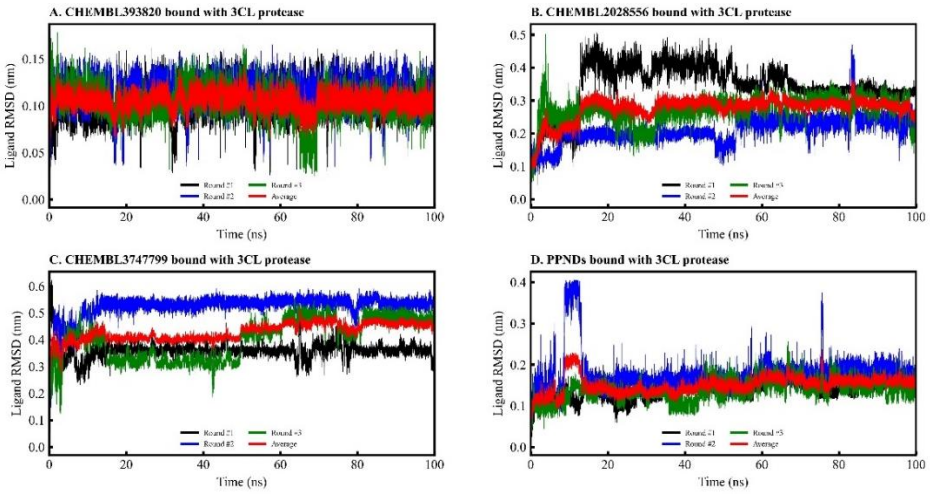

Figure 7. RMSD of CHEMBL393820, CHEMBL2028556, CHEMBL3747799 and PPNDs bound with 3CLpro

3.5.3 Residual fluctuation analysis

The role of individual amino acids are playing a crucial role to hold the ligand and give complex stability in dynamic states. The fluctuation of the 3CLpro amino acids was explored through RMSF and it is given in Figure 8. From all three rounds of simulation for each complex, it can be seen that all amino acids of 3CLpro fluctuated in a similar pattern. The average RMSF from three rounds of simulations was calculated and given in Figure 8. The average (red lines) RMSF was also found to fluctuate in similarly to the other three rounds of simulations. Apart from the extreme two ends, amino acids from about 90 to 120 were seen to fluctuate a bit more in comparison to the other part of the protein. The possible cause of such a bit high fluctuation in the above region might be the loss and reform of the binding interaction with the active site amino acids.

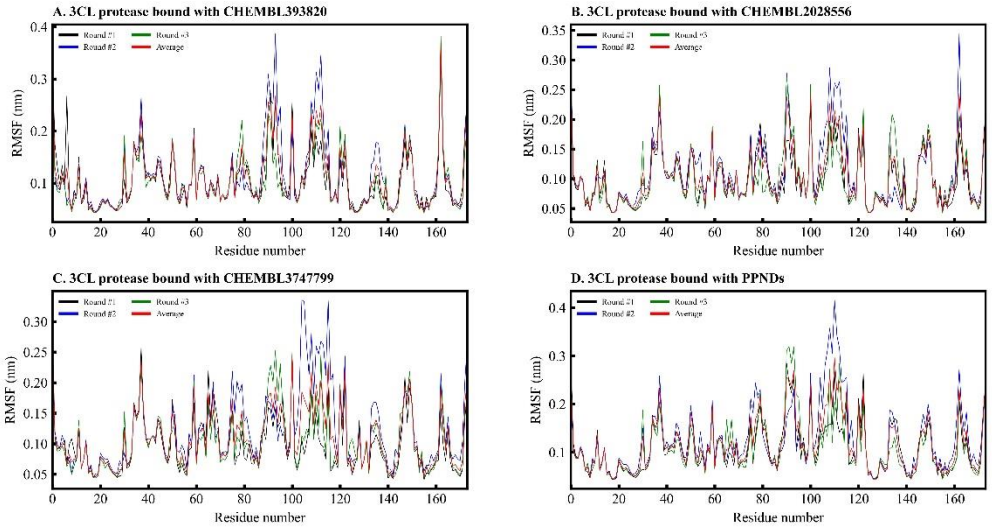

Figure 8. Root-mean-square fluctuation of 3CLpro amino acids bound with CHEMBL393820, CHEMBL2028556, CHEMBL3747799 and PPNDs

3.5.4 Radius of gyration

The compactness of the system during the MD simulation was assessed through RoG calculation and it is given in Figure 9. For more insightful observation, the average RoG of all three independent simulations for each complex was calculated and plotted in Figure 9. The RoG data of the proposed molecules bound in 3CLpro clearly indicates the compactness of the systems for all three rounds of simulation. It can be seen that for all three proposed molecules along with PPNDs, the average RoG deviated around 1.52 nm. Not a single simulation was

found to show abnormal RoG deviation. Hence, such a consistent variation of the RoG in all three rounds of MD simulation for each of the proposed molecules definitely explained that no unusual opening of the protein happened in the dynamic states.

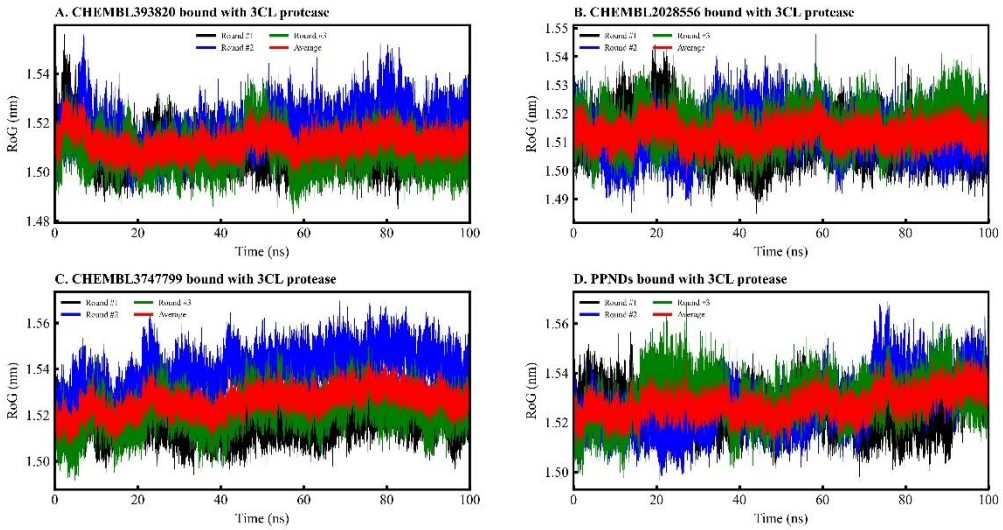

Figure 9. The radius of gyration of 3CLpro bound with CHEMBL393820, CHEMBL2028556, CHEMBL3747799 and PPNDs

3.5.5 Hydrogen bond assessment

Along with other non-covalent bonds, the hydrogen bonds between protein and small molecule play a crucial role to hold the small molecules inside the active site of the protein. In the dynamic states, a number of HBs are lost and reformed due to orientational changes. Therefore, it is essential to explore the number of HBs in each of the frames obtained from the MD simulation trajectory to understand the stability between protein and small molecules. The HBs for each frame for all three rounds of MD simulation of all three proposed molecules along with PPNDs were calculated and these are given in Figure 10. It can be seen CHEMBL2028556, CHEM3747799 and PPNDs formed a significant number of HBs with 3CLpro. It can be noted that CHEMBL393820 formed fewer HBs in comparison to others and it might be due to the smaller size of the molecule. Although the number of HBs is a whole number, an average HBs calculated and it is given in Figure 10. Except for CHEMBL393820, the average HBs graph of all other molecules was shown between 2 to 4. In the case of CHEMBL393820, with some exceptions, the average HBs remained within 2. Hence, the above observations clearly suggested that all proposed molecules successfully retained some HBs in dynamics states and hold them inside the receptor cavity.

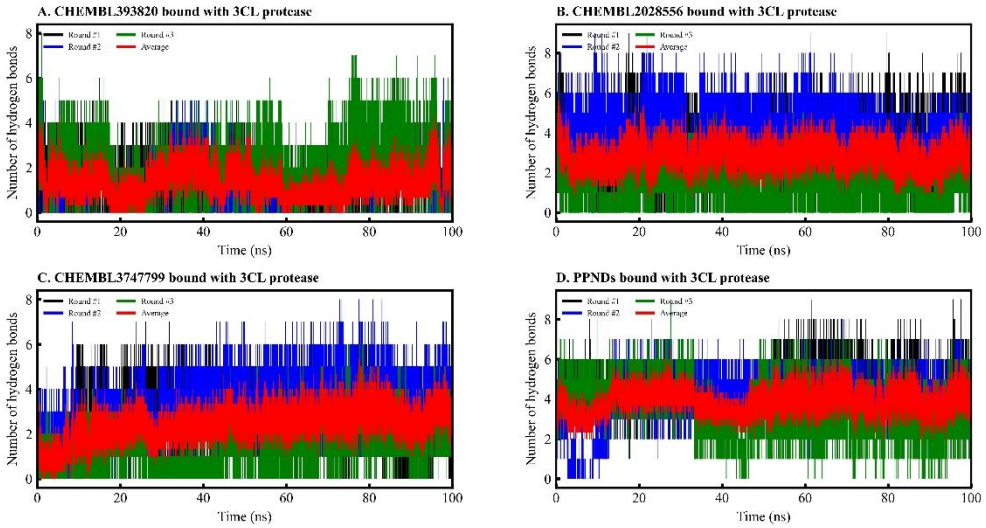

Figure 10. Number of hydrogen bonds between 3CLpro, and, CHEMBL393820, CHEMBL2028556, CHEMBL3747799 and PPNDs

3.5.6 Free energy landscape

The conformational transitions of 3CLpro bound with proposed molecules were explored using the FEL estimation and given in Figure 11. It is the representation of Gibbs free energy against two eigenvectors and it might be crucial to understand the stability of the protein conformations. In Figure 11, the blue color represents the minimum energy area, whereas, the red areas indicate the maxima. On close observations, it can be seen that no clear large minima were found for 3CLpro bound with PPNDs in all three rounds of simulations. Moreover, the greater distances between the minima of the 3CLpro-PPNDs complex explained the higher transition states that are unfavorable for the stability of the molecule. The 3CLpro bound with CHEMBL393820 and CHEMBL2028556 were found to show clear flat deep minima in Round #2 and #3, respectively. The complex between 3CLpro and CHEMBL3747799 was found to form evident and prominent large minima in both Round #1 and #3. The above observations clearly suggested that 3CLpro bound with all three proposed molecules were shown low energy states and stable clusters in comparison to the 3CLpro bound with PPNDs.

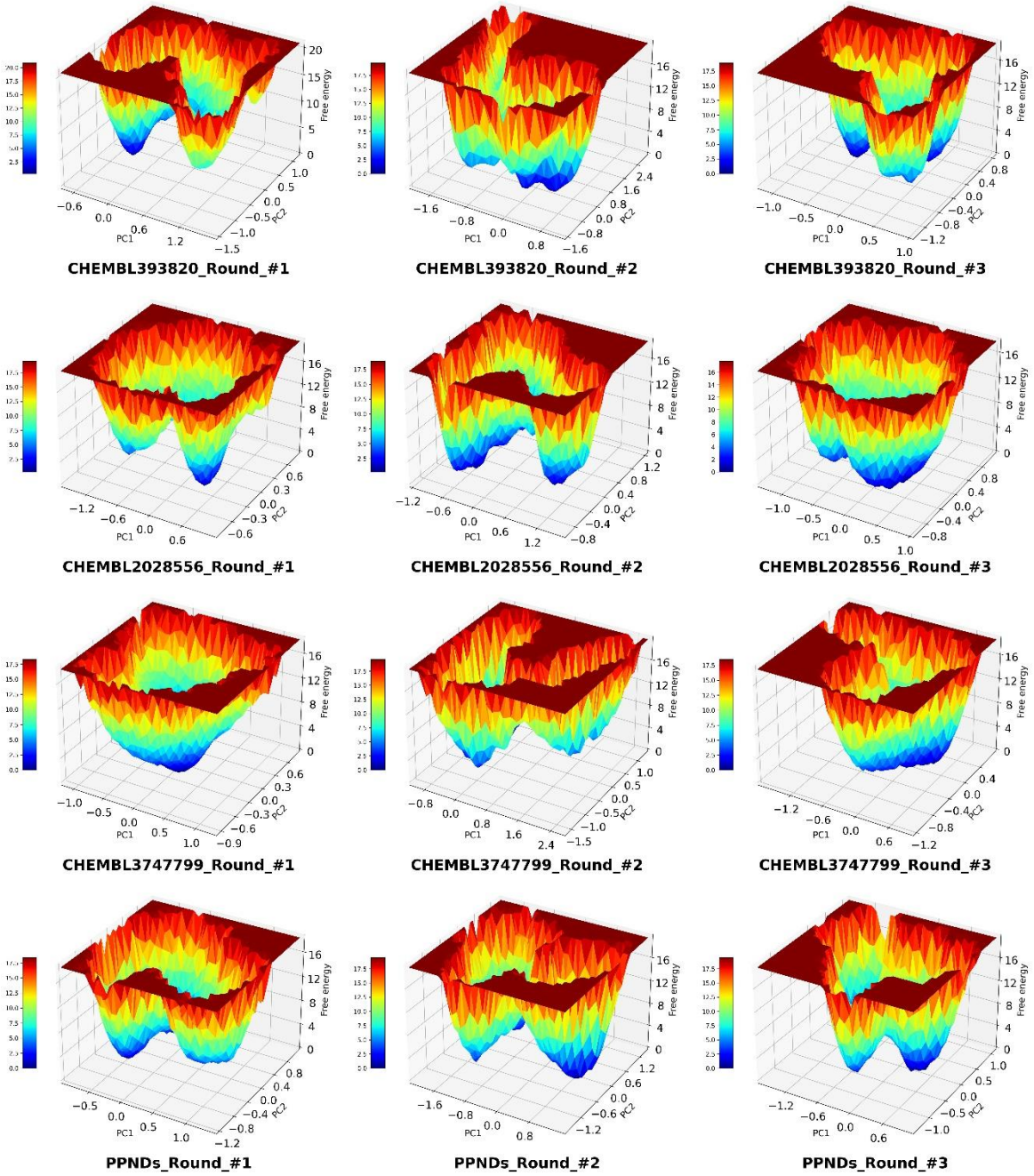

Figure 11. Free energy landscape between PC1 and PC2 of 3CLpro bound with CHEMBL393820, CHEMBL2028556, CHEMBL3747799 and PPNDs

3.5.7 Binding free energy using MM-GBSA

The binding free energy (ΔG_{bind}) of each of the proposed small molecules along with PPNDs was calculated using the MM-GBSA approach and it is given in Table 3. To confirm and validate, the ΔG_{bind} of all molecules from each of the three rounds of MD simulation was calculated. It can be seen that CHEMBL393820 shows binding affinity towards 3CLpro by ΔG_{bind} of -12.904, -5.931 and -21.154 kcal/mol for Round #1, #2 and #3, respectively.

Similarly, ΔG_{bind} in Round #1, #2 and #3 for CHEMBL2028556, CHEMBL3747799 and PPNDs was found to be -29.055, -39.984 and -2.844 kcal/mol, -35.951, -33.593 and -27.244 kcal/mol, and, -35.983, -23.245 and -28.483 kcal/mol, respectively. The average ΔG_{bind} was calculated and it was found to be -13.330, -30.961, -32.263 and -29.237 kcal/mol for CHEMBL393820, CHEMBL2028556, CHEMBL3747799 and PPNDs, respectively. In particular, CHEMBL393820 was found to show the lowest affinity in comparison to other molecules but not a single round of simulation found positive ΔG_{bind}. Both CHEMBL2028556 and CHEMBL3747799 were revealed to exert a bit higher binding affinity towards 3CLpro in comparison to the standard molecule, PPNDs. From the binding free energy analyses, it can be postulated that CHEMBL2028556 and CHEMBL3747799 might be crucial 3CLpro inhibitors, while, CHEMBL393820 needs some optimization to improve the potentiality.

Table 3. Binding free energy of CHEMBL393820, CHEMBL2028556, CHEMBL3747799 and PPNDs towards HuNoV 3CLpro

Molecules	**Round #1**		**Round #2**		**Round #3**		**Average**	
	ΔG_{bind}	**Std. Dev.**	**ΔG_{bind}**	**Std. Dev.**	**ΔG_{bind}**	**Std. Dev.**	**ΔG_{bind}**	**Std. Dev.**
CHEMBL393820	-12.904	±5.156	-5.931	±7.839	-21.154	±5.015	-13.330	±6.003
CHEMBL2028556	-29.055	±8.007	-39.984	±7.338	-23.844	±8.537	-30.961	±7.961
CHEMBL3747799	-35.951	±8.545	-33.593	±6.300	-27.244	±8.295	-32.263	±7.713
PPNDs	-35.983	±5.087	-23.245	±8.360	-28.483	±9.293	-29.237	±7.580

4. Conclusions

In the current work, a ligand-based similarity search was carried out to find potential chemical compounds against 3CLpro for the management of HuNoV infections. The potential chemical entities in pre-clinical trials for the treatment of HuNoV infections were considered for a similarity search of the entire ChEMBL database. High-similarity molecules found from ChEMBL were considered for the molecular docking studies using two docking engines such as SCORCH and PLANTS. The binding energy from both the docking studies was analyzed and three molecules (CHEMBL393820, CHEMBL2028556 and CHEMBL3747799) were found to be common within the top ten molecules from both docking studies, hence, proposed as potential 3CLpro inhibitors/modulators for HuNoV infections. The binding interaction analyses revealed a number of critical amino acids to form several non-covalent bonds between proposed molecules and 3CLpro. The binding interactions of PPNDs were compared and found a number of common amino acids to interact with the proposed molecule. Three rounds of MD simulation for each of the molecules bound with 3CLpro clearly explained their stability and

consistency in the dynamic states. High negative binding free energy towards 3CLpro has also suggested the potentiality of the molecules. The free energy landscape for each complex was explored from all three rounds of MD simulation and it showed the stability of the complexes. Hence, the proposed molecules might be important candidates for the management and treatment of HuNoV infections subjected to *in-vitro*/*in-vivo* validation.

CRediT author statement

Shovonlal Bhowmick: Conceptualization, Data curation, Methodology, original draft, Writing - review & editing, Validation; Visualization; **Tapan Kumar Mistri:** Data curation, Methodology, original draft, Writing - review & editing; **Mohammad K. Okla:** Conceptualization, Writing - review & editing; **Ibrahim A. Saleh:** Funding acquisition, Writing - review & editing; **Hamada AbdElgawad:** Data curation, Methodology, original draft, Writing - review & editing; **Achintya Saha:** Conceptualization, Data curation, Methodology, original draft, Writing - review & editing, Supervision; **Pritee Chunarkar Patil:** Conceptualization, Data curation, Methodology, original draft, Writing - review & editing, Supervision.

Acknowledgment

The authors extend their appreciation to the Researchers Supporting Project number (RSP2023R374) King Saud University, Riyadh, Saud Arabia.

Consent to Participate

The work of the research article is executed and organized by all authors together.

Consent for Publication

All authors are willing to publish this research work.

References

[1] P.A. White, Evolution of norovirus, Clin Microbiol Infect, 20 (2014) 741-745.
[2] A.C. Galasiti Kankanamalage, P.M. Weerawarna, A.D. Rathnayake, Y. Kim, N. Mehzabeen, K.P. Battaile, S. Lovell, K.O. Chang, W.C. Groutas, Putative structural rearrangements associated with the interaction of macrocyclic inhibitors with norovirus 3CL protease, Proteins, 87 (2019) 579-587.
[3] K.V. Todd, R.A. Tripp, Human Norovirus: Experimental Models of Infection, Viruses, 11 (2019).

[4] E. Robilotti, S. Deresinski, B.A. Pinsky, Norovirus, Clin Microbiol Rev, 28 (2015) 134-164.
[5] C.P. Campillay-Veliz, J.J. Carvajal, A.M. Avellaneda, D. Escobar, C. Covian, A.M. Kalergis, M.K. Lay, Human Norovirus Proteins: Implications in the Replicative Cycle, Pathogenesis, and the Host Immune Response, Front Immunol, 11 (2020) 961.
[6] J. Van Dycke, M. Puxeddu, G. La Regina, E. Mastrangelo, D. Tarantino, J. Rymenants, J. Sebastiani, M. Nalli, J. Matthijnssens, J. Neyts, R. Silvestri, J. Rocha-Pereira, Discovery of a Novel Class of Norovirus Inhibitors with High Barrier of Resistance, Pharmaceuticals (Basel), 14 (2021).
[7] M. Shah, A. Hall, Global Disease Burden of Foodborne Illnesses Associated With Norovirus, 2017, pp. 3-19.
[8] A.J. Hall, B.A. Lopman, D.C. Payne, M.M. Patel, P.A. Gastanaduy, J. Vinje, U.D. Parashar, Norovirus disease in the United States, Emerg Infect Dis, 19 (2013) 1198-1205.
[9] K.F. Azim, M. Hasan, M.N. Hossain, S.R. Somana, S.F. Hoque, M.N.I. Bappy, A.T. Chowdhury, T. Lasker, Immunoinformatics approaches for designing a novel multi epitope peptide vaccine against human norovirus (Norwalk virus), Infection, Genetics and Evolution, 74 (2019) 103936.
[10] O. Ebenezer, N. Damoyi, M. Shapi, Predicting New Anti-Norovirus Inhibitor With the Help of Machine Learning Algorithms and Molecular Dynamics Simulation-Based Model, Frontiers in chemistry, 9 (2021) 753427.
[11] E. Behmard, A. Ghasemian, E. Barzegari, A. Farjadfar, A. Kouhpayeh, S. Najafipour, Advanced simulations and screening to repurposing a 3C protease inhibitor against the rupintrivir-resistant human norovirus-induced gastroenteritis, Journal of Molecular Graphics and Modelling, 118 (2023) 108345.
[12] K.M. Muzzarelli, B. Kuiper, N. Spellmon, J. Brunzelle, J. Hackett, F. Amblard, S. Zhou, P. Liu, I.A. Kovari, Z. Yang, R.F. Schinazi, L.C. Kovari, Structural and Antiviral Studies of the Human Norovirus GII.4 Protease, Biochemistry, 58 (2019) 900-907.
[13] Y. Kim, A.C. Galasiti Kankanamalage, K.-O. Chang, W.C. Groutas, Recent Advances in the Discovery of Norovirus Therapeutics, Journal of Medicinal Chemistry, 58 (2015) 9438-9450.
[14] M.E. Hardy, T.J. Crone, J.E. Brower, K. Ettayebi, Substrate specificity of the Norwalk virus 3C-like proteinase, Virus Res, 89 (2002) 29-39.
[15] R.J. Hussey, L. Coates, R.S. Gill, P.T. Erskine, S.F. Coker, E. Mitchell, J.B. Cooper, S. Wood, R. Broadbridge, I.N. Clarke, P.R. Lambden, P.M. Shoolingin-Jordan, A structural study of norovirus 3C protease specificity: binding of a designed active site-directed peptide inhibitor, Biochemistry, 50 (2011) 240-249.
[16] Z. Muhaxhiri, L. Deng, S. Shanker, B. Sankaran, M.K. Estes, T. Palzkill, Y. Song, B.V. Prasad, Structural basis of substrate specificity and protease inhibition in Norwalk virus, J Virol, 87 (2013) 4281-4292.
[17] V.C. Damalanka, Y. Kim, A.C. Galasiti Kankanamalage, A.D. Rathnayake, N. Mehzabeen, K.P. Battaile, S. Lovell, H.N. Nguyen, G.H. Lushington, K.O. Chang, W.C. Groutas, Structure-guided design, synthesis and evaluation of oxazolidinone-based inhibitors of norovirus 3CL protease, Eur J Med Chem, 143 (2018) 881-890.
[18] S. Weerasekara, A.M. Prior, D.H. Hua, Current tools for norovirus drug discovery, Expert Opin Drug Discov, 11 (2016) 529-541.
[19] D. Takahashi, Y. Kim, S. Lovell, O. Prakash, W.C. Groutas, K.-O. Chang, Structural and inhibitor studies of norovirus 3C-like proteases, Virus Research, 178 (2013) 437-444.
[20] V.C. Damalanka, Y. Kim, A.C. Galasiti Kankanamalage, G.H. Lushington, N. Mehzabeen, K.P. Battaile, S. Lovell, K.O. Chang, W.C. Groutas, Design, synthesis, and evaluation of a

novel series of macrocyclic inhibitors of norovirus 3CL protease, Eur J Med Chem, 127 (2017) 41-61.
[21] G. Sliwoski, S. Kothiwale, J. Meiler, E.W. Lowe, Jr., Computational methods in drug discovery, Pharmacol Rev, 66 (2014) 334-395.
[22] S. Ekins, J. Mestres, B. Testa, In silico pharmacology for drug discovery: applications to targets and beyond, Br J Pharmacol, 152 (2007) 21-37.
[23] I. Muegge, P. Mukherjee, An overview of molecular fingerprint similarity search in virtual screening, Expert Opin Drug Discov, 11 (2016) 137-148.
[24] M.A. Skinnider, C.A. Dejong, B.C. Franczak, P.D. McNicholas, N.A. Magarvey, Comparative analysis of chemical similarity methods for modular natural products with a hypothetical structure enumeration algorithm, Journal of Cheminformatics, 9 (2017) 46.
[25] M.C. Hutter, Differential Multimolecule Fingerprint for Similarity Search─Making Use of Active and Inactive Compound Sets in Virtual Screening, Journal of Chemical Information and Modeling, 62 (2022) 2726-2736.
[26] M. Simončič, M. Lukšič, M. Druchok, Machine learning assessment of the binding region as a tool for more efficient computational receptor-ligand docking, Journal of Molecular Liquids, 353 (2022) 118759.
[27] P.H.M. Torres, A.C.R. Sodero, P. Jofily, F.P. Silva-Jr, Key Topics in Molecular Docking for Drug Design, Int J Mol Sci, 20 (2019).
[28] C. Yang, E.A. Chen, Y. Zhang, Protein-Ligand Docking in the Machine-Learning Era, Molecules, 27 (2022).
[29] S.A. Hollingsworth, R.O. Dror, Molecular Dynamics Simulation for All, Neuron, 99 (2018) 1129-1143.
[30] A. Hospital, J.R. Goni, M. Orozco, J.L. Gelpi, Molecular dynamics simulations: advances and applications, Adv Appl Bioinform Chem, 8 (2015) 37-47.
[31] A. Gaulton, A. Hersey, M. Nowotka, A.P. Bento, J. Chambers, D. Mendez, P. Mutowo, F. Atkinson, L.J. Bellis, E. Cibrián-Uhalte, M. Davies, N. Dedman, A. Karlsson, M.P. Magariños, J.P. Overington, G. Papadatos, I. Smit, A.R. Leach, The ChEMBL database in 2017, Nucleic Acids Res, 45 (2017) D945-d954.
[32] A. Gaulton, L.J. Bellis, A.P. Bento, J. Chambers, M. Davies, A. Hersey, Y. Light, S. McGlinchey, D. Michalovich, B. Al-Lazikani, J.P. Overington, ChEMBL: a large-scale bioactivity database for drug discovery, Nucleic Acids Res, 40 (2012) D1100-1107.
[33] S. Bhowmick, N.A. AlFaris, A.J. Zaidan, A.L. ZA, P.C. Patil, T.S. Aldayel, S.M. Wabaidur, A. Saha, Identification of bio-active food compounds as potential SARS-CoV-2 PLpro inhibitors-modulators via negative image-based screening and computational simulations, Comput Biol Med, 145 (2022) 105474.
[34] S. Bhowmick, S.A. Alissa, S.M. Wabaidur, R.V. Chikhale, M.A. Islam, Structure-guided screening of chemical database to identify NS3-NS2B inhibitors for effective therapeutic application in dengue infection, J Mol Recognit, 33 (2020) e2838.
[35] S. Bhowmick, R.D. Chorge, C.S. Jangam, L.D. Bharatrao, P.C. Patil, R.V. Chikhale, M.A. Islam, Identification of potential cruzain inhibitors using de novo design, molecular docking and dynamics simulations studies, J Biomol Struct Dyn, 38 (2020) 4005-4015.
[36] S. Bhowmick, K. Roy, A. Saha, Exploring CIP2A modulators using multiple molecular modeling approaches, J Biomol Struct Dyn, 40 (2022) 1048-1063.
[37] S. Bhowmick, A. Saha, N.A. AlFaris, A.L. JZ, A.L. ZA, T.S. Aldayel, S.M. Wabaidur, M.A. Islam, Identification of potent food constituents as SARS-CoV-2 papain-like protease modulators through advanced pharmacoinformatics approaches, J Mol Graph Model, 111 (2022) 108113.

[38] S. Bhowmick, A. Saha, N.A. AlFaris, A.L. JZ, A.L. ZA, T.S. Aldayel, S.M. Wabaidur, M.A. Islam, Structure-based identification of galectin-1 selective modulators in dietary food polyphenols: a pharmacoinformatics approach, Mol Divers, 26 (2022) 1697-1714.
[39] S. Bhowmick, A. Saha, S.M. Osman, F.A. Alasmary, T.M. Almutairi, M.A. Islam, Structure-based identification of SARS-CoV-2 main protease inhibitors from anti-viral specific chemical libraries: an exhaustive computational screening approach, Mol Divers, 25 (2021) 1979-1997.
[40] P.P. Dike, S. Bhowmick, G.E. Eldesoky, S.M. Wabaidur, P.C. Patil, M.A. Islam, In silico identification of small molecule modulators for disruption of Hsp90-Cdc37 protein-protein interaction interface for cancer therapeutic application, J Biomol Struct Dyn, 40 (2022) 2082-2098.
[41] C.S. Jangam, S. Bhowmick, R.D. Chorge, L.D. Bharatrao, P.C. Patil, R.V. Chikhale, N.A. AlFaris, A.L. JZ, S.M. Wabaidur, M.A. Islam, Pharmacoinformatics-based identification of anti-bacterial catalase-peroxidase enzyme inhibitors, Comput Biol Chem, 83 (2019) 107136.
[42] M.S. Kalbhor, S. Bhowmick, A.M. Alanazi, P.C. Patil, M.A. Islam, Multi-step molecular docking and dynamics simulation-based screening of large antiviral specific chemical libraries for identification of Nipah virus glycoprotein inhibitors, Biophys Chem, 270 (2021) 106537.
[43] P. Parida, S. Bhowmick, A. Saha, M.A. Islam, Insight into the screening of potential beta-lactamase inhibitors as anti-bacterial chemical agents through pharmacoinformatics study, J Biomol Struct Dyn, 39 (2021) 923-942.
[44] R.U. Savale, S. Bhowmick, S.M. Osman, F.A. Alasmary, T.M. Almutairi, D.S. Abdullah, P.C. Patil, M.A. Islam, Pharmacoinformatics approach based identification of potential Nsp15 endoribonuclease modulators for SARS-CoV-2 inhibition, Arch Biochem Biophys, 700 (2021) 108771.
[45] P.B. Shinde, S. Bhowmick, E. Alfantoukh, P.C. Patil, S.M. Wabaidur, R.V. Chikhale, M.A. Islam, De novo design based identification of potential HIV-1 integrase inhibitors: A pharmacoinformatics study, Comput Biol Chem, 88 (2020) 107319.
[46] P.M. Tambe, S. Bhowmick, S.K. Chaudhary, M.R. Khan, S.M. Wabaidur, M. Muddassir, P.C. Patil, M.A. Islam, Structure-Based Screening of DNA GyraseB Inhibitors for Therapeutic Applications in Tuberculosis: a Pharmacoinformatics Study, Appl Biochem Biotechnol, 192 (2020) 1107-1123.
[47] M. McGibbon, S. Money-Kyrle, V. Blay, D.R. Houston, SCORCH: Improving structure-based virtual screening with machine learning classifiers, data augmentation, and uncertainty estimation, J Adv Res, 46 (2023) 135-147.
[48] O. Korb, T. Stützle, T.E. Exner, PLANTS: Application of Ant Colony Optimization to Structure-Based Drug Design, in: M. Dorigo, L.M. Gambardella, M. Birattari, A. Martinoli, R. Poli, T. Stützle (Eds.) Ant Colony Optimization and Swarm Intelligence, Springer Berlin Heidelberg, Berlin, Heidelberg, 2006, pp. 247-258.
[49] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The Protein Data Bank, Nucleic Acids Research, 28 (2000) 235-242.
[50] A.C. Galasiti Kankanamalage, Y. Kim, A.D. Rathnayake, V.C. Damalanka, P.M. Weerawarna, S.T. Doyle, A.F. Alsoudi, D.M.P. Dissanayake, G.H. Lushington, N. Mehzabeen, K.P. Battaile, S. Lovell, K.O. Chang, W.C. Groutas, Structure-based exploration and exploitation of the S(4) subsite of norovirus 3CL protease in the design of potent and permeable inhibitors, Eur J Med Chem, 126 (2017) 502-516.
[51] D.E.V. Pires, T.L. Blundell, D.B. Ascher, pkCSM: Predicting Small-Molecule Pharmacokinetic and Toxicity Properties Using Graph-Based Signatures, Journal of Medicinal Chemistry, 58 (2015) 4066-4072.

[52] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX, 1-2 (2015) 19-25.
[53] J. Huang, A.D. MacKerell, Jr., CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data, Journal of computational chemistry, 34 (2013) 2135-2145.
[54] V. Zoete, M.A. Cuendet, A. Grosdidier, O. Michielin, SwissParam: A fast force field generation tool for small organic molecules, 32 (2011) 2359-2368.
[55] P. Mark, L. Nilsson, Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K, The Journal of Physical Chemistry A, 105 (2001) 9954-9960.
[56] G.G. Maisuradze, A. Liwo, H.A. Scheraga, Relation between Free Energy Landscapes of Proteins and Dynamics, Journal of Chemical Theory and Computation, 6 (2010) 583-595.
[57] D.J. Wales, The energy landscape as a unifying theme in molecular science, Philosophical transactions. Series A, Mathematical, physical, and engineering sciences, 363 (2005) 357-375; discussion 375-357.
[58] A. Amadei, A.B.M. Linssen, H.J.C. Berendsen, Essential dynamics of proteins, 17 (1993) 412-425.
[59] C.C. David, D.J. Jacobs, Principal component analysis: a method for determining the essential dynamics of proteins, Methods Mol Biol, 1084 (2014) 193-226.
[60] M.S. Valdés-Tresanco, M.E. Valdés-Tresanco, P.A. Valiente, E. Moreno, gmx_MMPBSA: A New Tool to Perform End-State Free Energy Calculations with GROMACS, Journal of Chemical Theory and Computation, 17 (2021) 6281-6291.
[61] J.D. Chodera, D.L. Mobley, Entropy-enthalpy compensation: role and ramifications in biomolecular ligand recognition and design, Annu Rev Biophys, 42 (2013) 121-142.
[62] F. Peccati, G. Jimenez-Oses, Enthalpy-Entropy Compensation in Biomolecular Recognition: A Computational Perspective, ACS Omega, 6 (2021) 11122-11130.
[63] P. Carracedo-Reboredo, J. Liñares-Blanco, N. Rodríguez-Fernández, F. Cedrón, F.J. Novoa, A. Carballal, V. Maojo, A. Pazos, C. Fernandez-Lozano, A review on machine learning approaches and trends in drug discovery, Comput Struct Biotechnol J, 19 (2021) 4538-4558.
[64] L. David, A. Thakkar, R. Mercado, O. Engkvist, Molecular representations in AI-driven drug discovery: a review and practical guide, J Cheminform, 12 (2020) 56.
[65] H. Kuwahara, X. Gao, Analysis of the effects of related fingerprints on molecular similarity using an eigenvalue entropy approach, Journal of Cheminformatics, 13 (2021) 27.
[66] A. Cereto-Massagué, M.J. Ojeda, C. Valls, M. Mulero, S. Garcia-Vallvé, G. Pujadas, Molecular fingerprint similarity search in virtual screening, Methods, 71 (2015) 58-63.
[67] A. Capecchi, D. Probst, J.-L. Reymond, One molecular fingerprint to rule them all: drugs, biomolecules, and the metabolome, Journal of Cheminformatics, 12 (2020) 43.
[68] J.-F. ROSSIGNOL, Y.M. EL-GOHARY, Nitazoxanide in the treatment of viral gastroenteritis: a randomized double-blind placebo-controlled clinical trial, Alimentary Pharmacology & Therapeutics, 24 (2006) 1423-1430.
[69] S. Ferla, N.E. Netzler, S. Ferla, S. Veronese, D.E. Tuipulotu, S. Guccione, A. Brancale, P.A. White, M. Bassetto, In silico screening for human norovirus antivirals reveals a novel non-nucleoside inhibitor of the viral polymerase, Sci Rep, 8 (2018) 4129.
[70] D. Rogers, M. Hahn, Extended-connectivity fingerprints, J Chem Inf Model, 50 (2010) 742-754.
[71] D. Takahashi, Y. Kim, S. Lovell, O. Prakash, W.C. Groutas, K.O. Chang, Structural and inhibitor studies of norovirus 3C-like proteases, Virus Res, 178 (2013) 437-444.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarydata.docx