

Discovery of New Hydroxyethylamine Analogs Against 3CL^{pro} Protein Target of SARS-CoV-2: Molecular Docking, Molecular Dynamics Simulation and Structure-Activity Relationship Studies

Sumit Kumar^{1,2}, Prem Prakash Sharma², Uma Shankar³, Dhruv Kumar⁴, Sanjeev K Joshi⁵, Lindomar Pena⁶, Ravi Durvasula⁷, Amit Kumar³, Prakasha Kempaiah⁷, Poonam¹, and Brijesh Rathi^{2,*}

¹*Department of Chemistry, Miranda House, University of Delhi, Delhi-110007.*

²*Laboratory for Translational Chemistry and Drug Discovery, Hansraj College, University of Delhi, Delhi-110007 India*

³*Discipline of Bioscience and Biomedical Engineering, Indian Institute of Technology, Indore, Simrol, Indore-453552, India*

⁴*Amity Institute of Molecular Medicine & Stem Cell Research (AIMMSCR), Amity University Uttar Pradesh, Sec-125, Noida-201313, India*

⁵*Technology Division, Defence Research & Development Organization, HQ, Rajaji Marg, New Delhi-110011*

⁶*Department of Virology, Aggeu Magalhaes Institute (IAM), Oswaldo Cruz Foundation (Fiocruz), 50670-420, Recife, Pernambuco, Brazil.*

⁷*Department of Medicine, Loyola University Stritch School of Medicine, 2160 South First Avenue, Chicago, Illinois 60153, United States*

Abstract

A novel coronavirus, SARS-CoV-2 has caused a recent pandemic called COVID-19 and a severe health threat around the world. In the current situation, the virus is rapidly spreading worldwide, and the discovery of vaccine and potential therapeutics are critically essential. The crystal structure for main protease (M^{pro}) of SARS-CoV-2, 3-chymotrypsin-like cysteine protease (3CL^{pro}) was recently made available and is considerably similar to previously reported SARS-CoV. Due to its essentiality in viral replication, it represents a potential drug target. Herein, computer-aided drug design (CADD) approach was implemented for the initial screening of 13 approved antiviral drugs. Molecular docking of 13 antivirals against 3-chymotrypsin-like cysteine protease (3CL^{pro}) enzyme was accomplished and indinavir was described as a lead drug with a docking score of -8.824 and a XP Gscore of -9.466 kcal/mol. Indinavir possesses an important pharmacophore, hydroxyethylamine (HEA), and thus a new library of HEA compounds (>2500) was subjected to virtual screening that led to 25 hits with a docking score more than indinavir. Exclusively, compound **16** with docking score of -8.955 adhered to drug like parameters, and the Structure-Activity Relationship (SAR) analysis was demonstrated to highlight the importance of chemical scaffolds therein. Molecular Dynamics (MD) simulation studies carried out at 100ns supported the stability of **16** within the binding pocket. Largely, our results supported that this novel compound **16** binds to the domain I & II, and domain II-III linker of 3CL^{pro} protein, suggesting its suitability as strong candidate for therapeutic discovery against COVID-19. Lead compound **16** could pave incredible directions for the design of novel 3CL^{pro} inhibitors and ultimately therapeutics against COVID-19 disease.

Keywords: Coronavirus, COVID-19, SARS-CoV-2, Virtual screening, Hydroxyethylamine, Molecular docking, and Molecular Dynamics simulation.

1. Introduction.

Coronaviruses (CoV) are a large group of viruses consisting of a core of genetic material enveloped with protein spike appearing like a crown, which means “corona” in Latin.¹ Diverse variety of coronaviruses is known, which causes mild respiratory diseases and sometimes gastrointestinal symptoms in different animal species. In humans, four CoV (229E, NL63, OC43 and HC HKU1) are endemic and cause respiratory diseases that can range from a common cold to lung failure, but the disease remains mild in most of the cases.^{2,3} However, other types of CoV i.e., SARS-CoV (Severe Acute Respiratory Syndrome), and MERS-CoV (Middle East Respiratory Syndrome) can cause severe respiratory diseases as identified in China (2002) and Saudi Arabia (2012), respectively.⁴ These viruses are bat-borne in nature and circulate in a range of animals and sometimes transmitted from animals to humans (i.e., SARS-CoV transmitted to humans from civet cats⁵ and MERS-CoV from dromedary camels).⁶ In December 2019, a cluster of pneumonia cases were observed in a group of people associated with seafood and animal market in China.⁷ Subsequently, the outbreak was attributed to a novel CoV related to the SARS virus based on the genetic similarities.⁸ Later, the disease was named as COVID-19 by World Health Organisation (WHO) caused by the virus SARS-CoV-2.⁹

The outbreak started in Wuhan, China and rapidly escalated to other countries. The USA, Italy, Spain and China lead with the highest number of COVID-19 cases officially reported. COVID-19 caused approximately 2,924 deaths, and 85,403 confirmed cases were identified with a mortality ratio of 3.42% till the end of February 2020 across the countries.¹⁰ The number of cases increased suddenly, and the disease was declared a pandemic by WHO on 11 March 2020.¹¹ As of 30 March 2020, a total of 745,308 confirmed cases of COVID-19 have been reported in 177 countries, resulting in 35,307 deaths.¹² A sharp increase in the number of confirmed cases and the number of deaths due to COVID-19 can be clearly observed across the world during the month of March, 2020 as shown in **Fig. 1**.¹³

Major symptoms of this disease includes high fever, and respiratory symptoms (i.e. cough and shortness of breath) whereas in severe cases pneumonia and kidney failure are a major cause for the death.¹⁴ Although, a series of rapid and specific diagnostic tools are available for COVID-19 disease¹⁵, vaccines and SARS-CoV-2 specific therapeutic treatments are not available. Recently, the 3D crystal structure of SARS-CoV-2 main protease, M^{pro}, also called 3CL^{pro}, complexed with an inhibitor N3 was released.¹⁶ 3CL^{pro} enzyme of SARS-CoV-2 processes polyproteins by proteolytic action of replicase enzyme (pp1a and pp1ab) to release the functional polypeptide (**Fig. 2B**). It is a dimeric protein, which contains two asymmetric units designated as protomers. Each protomer is composed of three domains, namely domain I

(residues 8–101), domain II (residues 102–184), and domain III (residues 201–303). Domain III contains five α -helices and connected with domain II through an extended loop region (residues 185–200) (**Fig. 2A**). The 3CL^{pro} has a Cys₁₄₅ and His₄₁ catalytic dyad, and the substrate-binding site is located in a cleft between domain I and II. These features are similar to previously reported protein structures of SARS-CoV.¹⁷⁻²³

The 3CL^{pro} of SARS-CoV-2 is an attractive drug target, and CADD is considered as an undisputable and significant approach to discover antiviral drug candidates. Identically, studies have been adopting a fast identification of drug candidates against SARS-CoV-2 through virtually screening.^{33,34} Antiviral drugs such as HIV protease inhibitors,²⁴ approved drugs such as ribavirin, oseltamivir and few more have been studied against earlier reported for SARS-CoV.²⁵ In this paper, we aim to perform the fast discovery of the potential drug candidates against 3CL^{pro} of SARS-CoV-2 through virtual screening of approved antiviral therapeutics and build a focused library of new potential compounds. Our screening results indicated that indinavir, a known inhibitor of HIV protease, has high potential against main protease, 3CL^{pro} of SARS-CoV-2 based on docking scores. The presence of a high-valued pharmacophore, hydroxyethylamine (HEA) in indinavir remains an incredible core moiety for the design of new antiviral drugs. Encouraged with this findings, we designed a new library of compounds possessing this crucial backbone, HEA, and followed extensive *in silico* studies towards the discovery of potential drugs candidates against SARS-CoV-2.

2. Material and Method.

The protease structure, 3-chymotrypsin-like cysteine protease (3CL^{pro}) enzyme of SARS-CoV-2 (PDB ID: 6LU7) with 2.1Å was obtained from the RCSB site (<http://www.rcsb.org>). The computational work was performed using Schrodinger software.

2.1. Preparation of Protease structure and active site identification.

The SARS-CoV-2 virus protein structure was prepared in the protein preparation wizard and prime module of Schrodinger suite to remove defects such as missing hydrogen atoms, incorrect bond order assignments, charge states, orientations of various groups and missing side chains suite.^{26, 27} Removal of steric clashes and strained bonds/angles were achieved through restrained energy minimization, allowing movement in heavy atoms up to 0.3 Å.

2.2. Preparation of antiviral, other approved drugs and analogs library for virtually screening.

The two libraries were created, firstly for antiviral and other approved drugs such as indinavir, saquinavir, nelfinavir, ritonavir, lopinavir, atazanavir, amprenavir, darunavir, nelfinavir, oseltamivir, tipranavir, fosamprenavir, and galidesvir to select the top-ranked approved drug as reference molecule or positive control. Another library was created from HEA based analogs to compare the efficiency of these compounds as potent analogs. The coordinate files of approved drugs were downloaded from Protein Data Bank at RCSB site (<http://www.rcsb.org>) and PubChem.^{28,29} Over thousand analogs were prepared by the maestro tool. The structure of both approved drugs and analogs were prepared prior to docking by Ligprep. The main objective of Ligprep was to take 2D or 3D structures and produce the corresponding low-energy 3D structures for use by programs such as Glide. All the parameters were kept default except chirality parameters for antivirals and analogs. The chirality was kept default from 3D structure and all combinations of chirality were developed for known antivirals and analogs. The next steps included desalting, generation of tautomer, and possible ionization states at pH 7 ± 2 .³⁰ Schrödinger suite inbuilt Epik module was used to predict the ionization states of all compounds.³¹

2.3. Molecular docking of chemical libraries against target protease active site.

Site specific molecular docking of approved drug (Table 1) and HEA based analogs (Table 2) against protease was performed using the Glide module of Schrödinger suite. Glide tool was used for receptor grid preparation at default parameters.³² The screening of both libraries was performed by Glide at extra precision (XP). The purposes of the XP method are to weed out false positives and to provide a better correlation between excellent poses and good scores.

2.5 ADME Calculations.

Swiss ADME³³ and PKCSM³⁴ were used to calculate ADMET (i.e. Absorption, Distribution, Metabolism, Excretion, and Toxicity) profile of identified hit compounds (**1-25**). The predicted ADMET properties includes Molecular weight, H-bond acceptor, H-bond donors, Predicted octanol/water partition coefficient (MLogP), TPSA (Total Polar Surface Area), Lipinski (druglikeness), Rat LD50, and Hepatotoxicity. The results are shown in Table 3.

2.6. Molecular Dynamics (MD) Simulation.

Extensive 100ns MD simulation was carried out on the complex structure of 6LU7 receptor with the selected approved drug (indinavir) and designed analog (compound **16**) using Desmond software (D. E. Shaw Research, New York, NY, 2015) to access the binding stability of Ligand-6LU7 complex.³⁵ The system was solvated in TIP3P water model and 0.15 M NaCl to mimic a physiological ionic concentration. The full system energy minimization step was done for 100ps. The MD simulation was run for 100ns at 300K temperature, standard pressure

(1.01325 bar), within an orthorhombic box with buffer dimensions $10 \times 10 \times 10 \text{ \AA}^3$ and NPT ensemble. The energy (kcal/mol) was recorded at an interval of 1.2 ps. The protein-ligand complex system was neutralized by balancing the net charge of the system by adding Na^+ or Cl^- counter ions. The Nose-Hoover chain and Martyna-Tobias-Klein dynamic algorithm was used to maintain the temperature of all the systems at 300 K and pressure 1.01325 barr, respectively.

3. Result and Discussion.

The current COVID-19 pandemics urges the immediate therapeutic treatments to reduce the morbidity, and to protect the health of people who are at high risk of infection, particularly when vaccines are still under development. Currently, no specific treatment is available and the only treatment remains is supportive care i.e. oxygen therapy, fluid management, use of approved antiviral developed for Ebola (remdesivir)³⁶, anti-HIV (lopinavir-ritonavir)³⁷ and recently approved antimalaria drugs (hydroxychloroquine)³⁸ as a combination therapy. The combination of azithromycin and hydroxychloroquine was also found to be effective as a treatment against COVID-19.³⁹ In order to develop targeted therapeutics of unknown infection from new virus, initially, we need to design library of compounds based on the available target protein structure and active binding moiety of the protein residues. Although, it is hard to predict the potency and binding specificity from the structure of a chemical compound. The structure-based design of new bioactive molecules involves, as a crucial step, the development of a reaction strategy affording enough chemical diversity and structural variety of potentially bioactive products. As such, not just a single molecule, but a significant part of the chemical space becomes available for the biological investigation that facilitates SAR analysis. Of particular note, finding a potent molecule possessing the drug-like properties remains challenging, and therefore advanced drug discovery involves CADD approaches that plays a crucial role in the early discovery phases. The main protease, 3CL^{pro} of SARS-CoV-2 was presented as an attractive drug target and a rapid identification of drug candidates against COVID-19 was achieved through virtually screening.^{33,34} Encouraged with this and taking the advantage of a recently released crystal structure of 3CL^{pro} from SARS-CoV-2 (PDB code 6LU7)¹⁶, we performed virtual screening for the 13 antiviral drugs. The list of the compounds are presented in Table 1.

3.1. Virtual screening of antiviral drugs and new HEA analogs against SARS-CoV-2 protein 3CL^{pro}

In the beginning, molecular docking of the approved antiviral drugs was carried out with protein 3CL^{pro} using the Glide module of Schrödinger suite.^{32, 40} Docking score, and glide XPGscore (kcal/mol) were considered as the measures of binding affinity to rank the poses of the ligands. Among the listed antiviral drugs, indinavir, which is an approved HIV-1 protease inhibitor, was realized as the best inhibitor based on docking score and XP Gscore of -8.824 and -9.466 kcal/mol, respectively (Table 1, Entry 1). Second rank was secured by another HIV-1 protease inhibitor atazanavir with a docking score of -7.912 (Table 1, Entry 2). Interactions of indinavir with 3CL^{pro} protein are presented in the **Fig. 3**. Indinavir showed H-bond interactions with a residue Gln_189 in domain-I, and Glu_166 in domain-II of 3CL^{pro} protein. It also formed salt-bridge interaction with residue Glu_166 in domain-II (Table 4, Entry 1). Therefore, indinavir possessing an important pharmacophore, HEA was selected as a positive control to build a new library of HEA analogs that could effectively inhibit the main protease 3CL^{pro} of SARS-CoV-2.

Our research studies⁴¹⁻⁴³, have demonstrated HEA as an important pharmacophore.^{44, 45} Together, presence of HEA in indinavir^{46, 47} and our ongoing research towards discovery of HEA analogs, a library of new HEA compounds (>2500) was designed and subjected to virtual screening against 3CL^{pro} protein of SARS-CoV-2. Of the screened >2500 new HEA analogs, 25 displayed a notable docking score over the indinavir as shown in Table 2. Among top-ranked analogs, compound **1** (Table 2, Entry 1) was shown to possess highest docking score (-9.864) and XP Gscore (-10.227 kcal/mol).

Next, to get an insight about the drug-likeness properties, all the 25 top-ranked analogs, were screened for their ADME profiles and the results are depicted in Table 3. Notably, only compound **16** (Table 3, Entry 16) strictly abide to Lipinski's "rule of five" and theoretically met the druglikeness properties. Compound **4** (Table 3, Entry 4) also followed the criteria with a slight deviation in number of rotatable bond (*i.e* 11).⁴⁸ However, compound **4** displayed hepatotoxicity. Based on the favorable physicochemical properties displayed by compound **16** was selected for further studies.

Compound **16** also offered an extra H-bond interaction, and also enhanced hydrophobically packed H-bond³² (Phob EnHB) interaction by -1.5. As depicted in **Fig. 4**, the phenyl group of HEA moiety of compound **16** interacted with Leu_141 and Phe_140 residue by hydrophobic interaction (Table 4, Entry 2). The 4-fluoro-aniline scaffold was found to be packed deeply inside the pocket by hydrophobic interactions with Cys-44 and Cys-145 and pi-pi-interaction with His-41 in domain-I. Residues Cys_145, Ser_144, and Gly_143 interacted with

carboxamide oxygen by H-bond, and His_164 in domain-I with nitrogen of aniline moiety as shown in **Fig. 4**.

SAR analysis was performed to designate the role of scaffolds present in HEA compounds. The lead compound, **16** contains three main functionalities i.e. HEA, anilines derivatives on pocket 1, and aromatic/aliphatic groups on pocket 2 as shown in Table 5. Firstly, the role of HEA backbone was studied, one compound **33** (Table 5, entry 8) that does not contain HEA secured the lowest docking score (i.e. -7.121) highlighting the importance of HEA. Secondly, variations made on pocket 1 were analysed keeping *tert*-butylacetate group at pocket 2 constant. Compound **16** possessing a docking score of -8.955 contains a 4-fluoroaniline at pocket 1 and found to be the best analog among others (i.e **29**, -8.131; **30**, -7.995; and **32**, -7.949) that contains 2-fluoroaniline, 2,4-difluoroaniline and 3-fluoroaniline, respectively (Table 5, entry 4, 5 and 7). The results clearly indicated the significance of the fluorine group at the *para* position of aniline as compared to *ortho* and *meta* positions. Similarly, chemical diversity at pocket 2 were coorelated while 4-fluoroanline at pocket 1 remains constant. Compound **26** possessing a docking score of -8.835 contains 4-trifluoromethylbenzyl moiety at pocket 2, and found to display improved docking score compared to **27** (-8.55), **28** (-8.543) and **31** (-7.965) that contain similar moieties 3-fluoro-4-trifluorophenyl, 4-fluorophenyl and 4-trifluorophenyl moiety, respectively (Table 5, Entry 1-3 and 6). However, **26** was shown to possess lower docking score in coparison to compound **16** that contains a *tert*-butylacetate group at pocket 2, indicating more pronounced effects without fluorinated aromatic components. Therefore, based upon molecular docking, ADME properties and SAR analysis, compound **16** was identified as most suitable analog (Table 2 and 5), and was considered for MD simulation study.

3.2. Molecular Dynamics (MD) Simulation of indinavir, compound 16.

MD simulation at 100ns was performed for indinavir-3CL^{pro} and compound **16**-3CL^{pro}. Un-ligated-3CL^{pro} complex was also considered for MD studies to identify the effect on simulations in the absence of ligands. There are several factors such as conformation of ligand, water molecules, ions, cofactors, ligand protonation state, conformational and solvation entropies that can affect *in silico* predictions in an unexpected pattern. Several reports are available to support the role of MD simulations for the refinement of docking results.^{49, 50} MD simulations were carried out to determine the stability of the interactions of ligand-protein docked complexes. The final structure of simulated complexes exhibited strong stereochemical geometry of the residues as analysed by the Ramachandran map (**Fig. 5**).

The number and percentage of residues in favored, allowed, and outlier regions for these three systems are presented in Table 6. Both, complex of indinavir, and un-ligated 3CL^{pro} possess

two (Gly_251, Thr_224) and three (Gly_11, Asn_84, Asp_187) outlier residues, respectively (Table 6, entry 2-3). Interestingly, there was no outlier residues for **16**-3CL^{pro} complex (Table 6, entry 1). Further, the stability of compound **16**-3CL^{pro} complex was monitored during the entire simulation process where total energy (E), potential energy (E_p), temperature (T), pressure (P), and volume (V) values were computed as shown in **Fig. 6**. The plots for un-ligated and indinavir complexes are displayed in supporting information (Fig. S1). Notably, no significant change was observed in potential energy and total energy and other parameters for compound **16** and indinavir complexes, as the average values with a standard deviation of these parameters were found in same range (supporting information, Table S1-S2). While entire simulations, Root Mean Square Deviation (RMSD) change for C α was also monitored. This parameter measures the global deviation of atoms from a reference status i.e., frame 0. The RMSD plot indicated the fluctuations in the initial conformation of the receptor for all three systems till 30 ns (**Fig. 7**), which later stabilized in production phase with an average values of RMSD_{backbone} (2.015 Å), RMSD_{C α} (1.993 Å) and RMSD_{side-chain} (2.876 Å), for **16**-3CL^{pro} complex. The average RMSD values for backbone, C α , and side-chain for indinavir-3CL^{pro} and un-ligated-3CL^{pro} complexes are depicted in supporting information (Table S3). The RMSD of the both C α and backbone for compound **16**-3CL^{pro} showed fluctuations between the range of 0.766 – 4.304 Å. Protein-RMSF was monitored to assess the local residue flexibility, and ligand-RMSF was examined to study the atom-wise fluctuations in the ligand. When a dynamical system fluctuates over well-defined average position, the RMSD from the average over time can be referred to as RMSF. The RMSF, 'Fit on Protein' trend (brown line; **Fig. 8B**) indicated the ligand fluctuations with respect to the binding site residues present on target protein. Whereas 'Fit on Ligand' trend (pink line, **Fig. 8B**) showed the fluctuations where the ligand in each frame is aligned on the ligand in the first reference frame. Both, compound **16** and indinavir interactions with 3CL^{pro} are shown in **Fig. 8**. Compound **16** interacted with residues (Gly_143, Ser_144, Cys_145, and His_164) upon docking in absence of water. In the presence of water molecules, only one hydrogen bond interaction was persistent with Gly_143, observed during MD simulation. This observation showed the importance of water molecules within binding pocket of protein 3CL^{pro} for ligand **16**. The detailed interaction of compound **16** with binding site residues of 3CL^{pro} were studied and are presented in **Fig. 9**, where interacting residues include Thr_26, Leu_27, His_41, Ser_46, Met_49, Asn_119, Phe_140, Leu_141, Asn_142, Gly_143, Ser_144, Cys_145, His_163, His_164, Met_165, Glu_166, Pro_168, His_172, and Gln_189. The interacting residue with compound **16** are presented with

green color, while the orange and blue bands indicate protein secondary structures i.e. helices and β -strands, respectively (**Fig. 9**). These binding site residues possessed RMSF value of $<2\text{\AA}$. The distribution of residues index of protein in secondary structure elements (SSE) was also studied (**Fig. S3**; supporting information). There are two panels in SSE i.e. top panel and bottom panel, where the top panel reports the percent SSE composition from helices and strands over time, and the bottom panel monitors each residue and its SSE assignment over the time. SSE percentage for compound **16**, indinavir, and un-ligated protein with values, 44.45%, 44.74%, and 44.45%, respectively, indicated negligible while 100ns simulation period (Table S4; supporting information). Compound **16** interacted with Leu_27, His_41, Met_49, and Met_165 residues of protein forming hydrophobic interactions, and with Asn_142, Gly_143, Ser_144, Cys_145, His_163, His_164, and Glu_166 formed H-bond-water interactions. There was no detectable ionic interaction for **16**-3CL^{pro} complex as depicted in **Fig. 10**. The total number of contacts formed by protein with the ligand **16** over the course of the trajectory was presented in the top panel and the bottom panel (**Fig. 11**) that showed the residues interacted with the ligand in each trajectory frame. It was observed that some protein residues formed more than one specific contact with the ligand (presented with a darker shade of orange). Overall, six properties were analysed to explain the stability of the compound **16** in the 3CL^{pro} receptor during the simulation of 100ns: a) ligand RMSD - Root mean square deviation of a ligand with respect to the reference conformation (typically the first frame was used as the reference and it is regarded as time $t=0$); b) radius of Gyration (rGyr) - it is used for measuring the 'extendedness' of a ligand, and is equivalent to its principal moment of inertia; c) NS34 is not capable of intramolecular hydrogen bonding, so this plot appears empty; d) Molecular Surface Area (MolSA) - molecular surface calculation with 1.4\AA probe radius; e) Solvent Accessible Surface Area (SASA) - surface area of a molecule accessible by a water molecule; f) Polar Surface Area (PSA) - Solvent accessible surface area in a molecule contributed only by oxygen and nitrogen atoms as shown in **Fig. 12**.

From the graph, it is evident that ligand RMSD remains constant during the simulation process. The overall RMSD of the compound **16** was up to 2.771\AA . In the initial stage, fluctuation was observed from 0ns to 18ns afterwards constant RMSD was observed during the entire simulation process. Fluctuation for Radius of Gyration were noted until 40 ns and subsequently, a stable conformation was acquired throughout the simulation. The radius of gyration throughout the 100ns simulation was ranged from 3.275\AA to 4.524\AA for compound **16**. The SASA plot revealed a fluctuating pattern for initial 42ns and then became stabilized until simulations were completed. MolSA, and PSA plots also indicated the consistency of the

ligand, **16** during the simulation process. Initially, compound **16** showed intramolecular H-bond, which became stabilised after protein interaction and no intramolecular H-bond was observed at the end.

A 2D schematic representation for compound **16** is presented with color coded rotatable bonds (**Fig. 13**). The rotatable torsional bond of compound **16** was accompanied by a dial (radial) plot and bar plots of the same color. The radial plot explains the conformation of the torsion angle throughout the time course of the simulation. The simulation starts from the centre of the radial plot and the time progress was plotted radially towards the outside. The probability density of the torsion angle was summarized by bar plots, which were recapitulating the data on the radial plots. The Y-axis of the bar plots revealed the potential of the rotatable bond, which are expressed in kcal/mol. The radial and bar diagram explained the torsion potential relationships and the conformational strain of compound **16** undergoing to maintain a protein-bound conformation.

4. Conclusion.

COVID-19 is continue to become a global burden on human health and economic losses remains unabated. A rapid progress, particularly in vaccine and therapeutic development, is essential to overcome the further explosion in spread and loss of human life from COVID-19. The availability of a protein target structure is usually helpful in identifying and designing potential drug candidates. Such approaches usually involve explicit molecular docking of ligands into the receptor-binding site, producing a predicted binding mode for each candidate compound. Virtual screening is widely used for structure-based drug design against a known drug target to identify those chemical structures, which are most likely to bind well within the active site of the target protein. Newly released 3CL^{pro} structure of SARS-CoV-2 was considered as potential target for the design of drug candidates and a rapid virtual screening for the libraries of approved drugs. Such approaches are certainly important to get better insight about the targeted therapeutic development for this novel infectious disease. In this article, we have screened approved antiviral drugs against 3CL^{pro}, and indinavir was identified with highest docking score and XP Gscore and thus selected as a positive control for the design of new drug candidates based on HEA scaffold. Molecular docking of new compounds was carried out and a total of 25 analogs were identified with improved or similar docking score as of indinavir. Both, compound **16** and indinavir bind to the domain I, II, and domain II-III linker of targeted protein with H-bond, pi-pi and hydrophobic interactions. Compound **16** and

indinavir were further studied for MD simulation. RMSF for indinavir and **16** complex were below 2 Å. Interestingly, no outlier residue was noted for compound **16-3CL^{pro}** complex as compared to two outlier residues found in indinavir-3CL^{pro} complex. Moreover, presence of water molecules within binding site of protein suggested stability of **16-3CL^{pro}** complex. RMSD and Ligand-RMSF percentage for C α demonstrated that stability of **16-3CL^{pro}** complex and a protein-bound conformation was confirmed by torsional analysis, suggesting suitability as a potent candidate for further exploration as a candidacy as an explorative drug for use against SARS-CoV-2.

***Corresponding Author**

Brijesh Rathi, PhD

brijeshrathi@hrc.du.ac.in; brathi@luc.edu

Acknowledgement.

Authors are grateful to Hansraj College, University of Delhi and Department of Medicine, Loyola University Medical Center and Stritch School of Medicine for providing support for the Drug Discovery Program. SS acknowledges CSIR, PPS is grateful to DBT, Govt of India for junior research fellowships.

Conflict of Interest.

Authors declare no conflict of interest.

References.

1. Pene, F.; Merlat, A.; Vabret, A.; Rozenberg, F.; Buzyn, A.; Dreyfus, F.; Cariou, A.; Freymuth, F.; Lebon, P., Coronavirus 229E-Related Pneumonia in Immunocompromised Patients. *Clinical Infectious Diseases* **2003**, 37, 929-932.
2. Bergmann, C. C.; Lane, T. E.; Stohlman, S. A., Coronavirus infection of the central nervous system: host–virus stand-off. *Nature Reviews Microbiology* **2006**, 4, 121-132.
3. Herrewegh, A. A. P. M.; Smeenk, I.; Horzinek, M. C.; Rottier, P. J. M.; de Groot, R. J., Feline Coronavirus Type II Strains 79-1683 and 79-1146 Originate from a Double Recombination between Feline Coronavirus Type I and Canine Coronavirus. *Journal of Virology* **1998**, 72, 4508.
4. Lee, H.; Lei, H.; Santarsiero, B. D.; Gatuz, J. L.; Cao, S.; Rice, A. J.; Patel, K.; Szypulinski, M. Z.; Ojeda, I.; Ghosh, A. K.; Johnson, M. E., Inhibitor Recognition Specificity of MERS-CoV Papain-like Protease May Differ from That of SARS-CoV. *ACS Chem Biol* **2015**, 10, 1456-1465.
5. Tsai, K.-C.; Chen, S.-Y.; Liang, P.-H.; Lu, I. L.; Mahindroo, N.; Hsieh, H.-P.; Chao, Y.-S.; Liu, L.; Liu, D.; Lien, W.; Lin, T.-H.; Wu, S.-Y., Discovery of a Novel Family of SARS-CoV Protease Inhibitors by Virtual Screening and 3D-QSAR Studies. *Journal of Medicinal Chemistry* **2006**, 49, 3485-3495.
6. Wang, C.; Xia, S.; Zhang, P.; Zhang, T.; Wang, W.; Tian, Y.; Meng, G.; Jiang, S.; Liu, K., Discovery of Hydrocarbon-Stapled Short α -Helical Peptides as Promising Middle East Respiratory Syndrome Coronavirus (MERS-CoV) Fusion Inhibitors. *Journal of Medicinal Chemistry* **2018**, 61, 2018-2026.
7. Guangdi Li, E. D. C., Therapeutic options for the 2019 novel coronavirus (2019-nCoV). *Nature Reviews Drug Discovery* **2020**, 149-150.
8. Chhikara, B. S.; Rathi, B.; Singh, J.; Poonam, Corona virus SARS-CoV-2 disease COVID-19: Infection, prevention and clinical advances of the prospective chemical drug therapeutics. *Chemical Biology Letters* **2020**, 7.
9. Sumit, K.; Poonam; Brijesh, R., Coronavirus Disease COVID-19: A New Threat to Public Health. *Current Topics in Medicinal Chemistry* **2020**, 20, 1-2.
10. World Health Organisation: Coronavirus disease (COVID-2019) situation reports. Situation Report 40. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200229-sitrep-40-covid-19.pdf?sfvrsn=1ba62e57_10 (29 February 2020),
11. World Health Organisation: Coronavirus disease (COVID-2019) situation reports. Situation Report 51. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200311-sitrep-51-covid-19.pdf?sfvrsn=1ba62e57_10 (11 March 2020),
12. Dong, E. D., H.; Gardner, L., An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases* **2020**.
13. World Health Organisation. Coronavirus disease (COVID-2019) situation reports. Situation Report 41-70. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports> (1-30 March 2020),
14. Wei, H.; Yin, H.; Huang, M.; Guo, Z., The 2019 novel coronavirus pneumonia with onset of oculomotor nerve palsy: a case study. *Journal of Neurology* **2020**.
15. Corman, V. M.; Landt, O.; Kaiser, M.; Molenkamp, R.; Meijer, A.; Chu, D. K. W.; Bleicker, T.; Brünink, S.; Schneider, J.; Schmidt, M. L.; Mulders, D. G. J. C.; Haagmans, B. L.; van der Veer, B.; van den Brink, S.; Wijsman, L.; Goderski, G.; Romette, J.-L.; Ellis, J.; Zambon, M.; Peiris, M.; Goossens, H.; Reusken, C.; Koopmans, M. P. G.; Drosten, C., Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Euro Surveill* **2020**, 25, 2000045.

16. Liu, X. Z., B.; Jin, Z.; Yang, H.; Rao, Z., The crystal structure of COVID-19 main protease in complex with an inhibitor N3. . **2020**.
17. Wang, F.; Chen, C.; Tan, W.; Yang, K.; Yang, H., Structure of Main Protease from Human Coronavirus NL63: Insights for Wide Spectrum Anti-Coronavirus Drug Design. *Sci Rep* **2016**, 6, 22677.
18. Ren, Z.; Yan, L.; Zhang, N.; Guo, Y.; Yang, C.; Lou, Z.; Rao, Z., The newly emerged SARS-like coronavirus HCoV-EMC also has an "Achilles' heel": current effective inhibitor targeting a 3C-like protease. *Protein Cell* **2013**, 4, 248-250.
19. Xue, X.; Yu, H.; Yang, H.; Xue, F.; Wu, Z.; Shen, W.; Li, J.; Zhou, Z.; Ding, Y.; Zhao, Q.; Zhang, X. C.; Liao, M.; Bartlam, M.; Rao, Z., Structures of Two Coronavirus Main Proteases: Implications for Substrate Binding and Antiviral Drug Design. *Journal of Virology* **2008**, 82, 2515.
20. Yang, H.; Yang, M.; Ding, Y.; Liu, Y.; Lou, Z.; Zhou, Z.; Sun, L.; Mo, L.; Ye, S.; Pang, H.; Gao, G. F.; Anand, K.; Bartlam, M.; Hilgenfeld, R.; Rao, Z., The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. *Proceedings of the National Academy of Sciences of the United States of America* **2003**, 100, 13190-13195.
21. Anand, K.; Palm, G. J.; Mesters, J. R.; Siddell, S. G.; Ziebuhr, J.; Hilgenfeld, R., Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra alpha-helical domain. *EMBO J* **2002**, 21, 3213-3224.
22. Mukherjee, P.; Shah, F.; Desai, P.; Avery, M., Inhibitors of SARS-3CLpro: Virtual Screening, Biological Evaluation, and Molecular Dynamics Simulation Studies. *Journal of Chemical Information and Modeling* **2011**, 51, 1376-1392.
23. Liu, Z.; Huang, C.; Fan, K.; Wei, P.; Chen, H.; Liu, S.; Pei, J.; Shi, L.; Li, B.; Yang, K.; Liu, Y.; Lai, L., Virtual Screening of Novel Noncovalent Inhibitors for SARS-CoV 3C-like Proteinase. *Journal of Chemical Information and Modeling* **2005**, 45, 10-17.
24. Yamamoto, N.; Yang, R.; Yoshinaka, Y.; Amari, S.; Nakano, T.; Cinatl, J.; Rabenau, H.; Doerr, H. W.; Hunsmann, G.; Otaka, A.; Tamamura, H.; Fujii, N.; Yamamoto, N., HIV protease inhibitor nelfinavir inhibits replication of SARS-associated coronavirus. *Biochemical and Biophysical Research Communications* **2004**, 318, 719-725.
25. Tan, E. L. C.; Ooi, E. E.; Lin, C.-Y.; Tan, H. C.; Ling, A. E.; Lim, B.; Stanton, L. W., Inhibition of SARS coronavirus infection in vitro with clinically approved antiviral drugs. *Emerg Infect Dis* **2004**, 10, 581-586.
26. Schrödinger Release 2020-1 (2020) Protein Preparation Wizard, Schrödinger, LLC, New York.
27. Schrödinger Release 2020-1: Prime, Schrödinger, LLC, New York, NY. **2020**.
28. Chang, Y. T., Y.; Lee, K.; Chen, T.; Hsiao, Y.; Chang, H.; Hsieh, T.; Su, C.; Wang, S.; Yu, J.; Shih, S.; Lin, Y.; Lin, Y.; Tu, Y.E.; Tung, C.; Chen, C. , Potential Therapeutic Agents for COVID-19 Based on the Analysis of Protease and RNA Polymerase Docking. *Preprints* **2020**, 2020020242.
29. Bedi, R. K.; Patel, C.; Mishra, V.; Xiao, H.; Yada, R. Y.; Bhaumik, P., Understanding the structural basis of substrate recognition by Plasmodium falciparum plasmepsin V to aid in the design of potent inhibitors. *Sci Rep* **2016**, 6, 31420-31420.
30. Schrödinger Release 2020-1: LigPrep, Schrödinger, LLC, New York, NY. **2020**.
31. Schrödinger Release 2020-1: Epik, Schrödinger, LLC, New York. **2020**.
32. Schrödinger Release 2020-1: Glide, Schrödinger, LLC, New York, NY. **2020**.
33. Daina, A.; Michielin, O.; Zoete, V., SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci Rep* **2017**, 7, 42717-42717.

34. Pires, D. E. V.; Blundell, T. L.; Ascher, D. B., pkCSM: Predicting Small-Molecule Pharmacokinetic and Toxicity Properties Using Graph-Based Signatures. *Journal of medicinal chemistry* **2015**, *58*, 4066-4072.
35. Schrödinger Release 2020-1: Desmond Molecular Dynamics System, D. E. Shaw Research, New York, NY, 2020. Maestro-Desmond Interoperability Tools, Schrödinger, New York, NY. **2020**.
36. Holshue, M. L.; DeBolt, C.; Lindquist, S.; Lofy, K. H.; Wiesman, J.; Bruce, H.; Spitters, C.; Ericson, K.; Wilkerson, S.; Tural, A.; Diaz, G.; Cohn, A.; Fox, L.; Patel, A.; Gerber, S. I.; Kim, L.; Tong, S.; Lu, X.; Lindstrom, S.; Pallansch, M. A.; Weldon, W. C.; Biggs, H. M.; Uyeki, T. M.; Pillai, S. K., First Case of 2019 Novel Coronavirus in the United States. *New England Journal of Medicine* **2020**, *382*, 929-936.
37. Cao, B.; Wang, Y.; Wen, D.; Liu, W.; Wang, J.; Fan, G.; Ruan, L.; Song, B.; Cai, Y.; Wei, M.; Li, X.; Xia, J.; Chen, N.; Xiang, J.; Yu, T.; Bai, T.; Xie, X.; Zhang, L.; Li, C.; Yuan, Y.; Chen, H.; Li, H.; Huang, H.; Tu, S.; Gong, F.; Liu, Y.; Wei, Y.; Dong, C.; Zhou, F.; Gu, X.; Xu, J.; Liu, Z.; Zhang, Y.; Li, H.; Shang, L.; Wang, K.; Li, K.; Zhou, X.; Dong, X.; Qu, Z.; Lu, S.; Hu, X.; Ruan, S.; Luo, S.; Wu, J.; Peng, L.; Cheng, F.; Pan, L.; Zou, J.; Jia, C.; Wang, J.; Liu, X.; Wang, S.; Wu, X.; Ge, Q.; He, J.; Zhan, H.; Qiu, F.; Guo, L.; Huang, C.; Jaki, T.; Hayden, F. G.; Horby, P. W.; Zhang, D.; Wang, C., A Trial of Lopinavir–Ritonavir in Adults Hospitalized with Severe Covid-19. *New England Journal of Medicine* **2020**.
38. Yao, X.; Ye, F.; Zhang, M.; Cui, C.; Huang, B.; Niu, P.; Liu, X.; Zhao, L.; Dong, E.; Song, C.; Zhan, S.; Lu, R.; Li, H.; Tan, W.; Liu, D., In Vitro Antiviral Activity and Projection of Optimized Dosing Design of Hydroxychloroquine for the Treatment of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). *Clinical Infectious Diseases* **2020**.
39. Gautret, P.; Lagier, J.-C.; Parola, P.; Hoang, V. T.; Meddeb, L.; Mailhe, M.; Doudier, B.; Courjon, J.; Giordanengo, V.; Vieira, V. E.; Dupont, H. T.; Honoré, S.; Colson, P.; Chabrière, E.; La Scola, B.; Rolain, J.-M.; Brouqui, P.; Raoult, D., Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial. *International Journal of Antimicrobial Agents* **2020**, 105949.
40. Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L., Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *Journal of Medicinal Chemistry* **2004**, *47*, 1750-1759.
41. Poonam; Gupta, Y.; Gupta, N.; Singh, S.; Wu, L.; Chhikara, B. S.; Rawat, M.; Rathi, B., Multistage inhibitors of the malaria parasite: Emerging hope for chemoprotection and malaria eradication. *Medicinal Research Reviews* **2018**, *38*, 1511-1535.
42. Upadhyay, C.; Chaudhary, M.; De Oliveira, R. N.; Borbas, A.; Kempaiah, P.; Poonam; Rathi, B., Fluorinated scaffolds for antimalarial drug discovery. *Expert Opinion on Drug Discovery* **2020**, 1-14.
43. Neha Sharma, A. V., Poonam, Prakasha Kempaiah, Brijesh Rathi, Chemical libraries targeting Liver Stage Malarial infection. *Chemical Biology Letters; Vol 6, No 1 (2019)* **2019**.
44. Singh, S.; Rajendran, V.; He, J.; Singh, A. K.; Achieng, A. O.; Vandana; Pant, A.; Nasamu, A. S.; Pandit, M.; Singh, J.; Quadiri, A.; Gupta, N.; Poonam; Ghosh, P. C.; Singh, B. K.; Narayanan, L.; Kempaiah, P.; Chandra, R.; Dunn, B. M.; Pandey, K. C.; Goldberg, D. E.; Singh, A. P.; Rathi, B., Fast-Acting Small Molecules Targeting Malarial Aspartyl Proteases, Plasmepsins, Inhibit Malaria Infection at Multiple Life Stages. *ACS Infectious Diseases* **2019**, *5*, 184-198.
45. Kumar Singh, A.; Rajendran, V.; Singh, S.; Kumar, P.; Kumar, Y.; Singh, A.; Miller, W.; Potemkin, V.; Poonam; Grishina, M.; Gupta, N.; Kempaiah, P.; Durvasula, R.; Singh, B. K.; Dunn, B. M.; Rathi, B., Antiplasmodial activity of hydroxyethylamine analogs: Synthesis,

biological activity and structure activity relationship of plasmepsin inhibitors. *Bioorganic & medicinal chemistry* **2018**, 26, 3837-3844.

46. Bhargava, S.; Adhikari, N.; Amin, S. A.; Das, K.; Gayen, S.; Jha, T., Hydroxyethylamine derivatives as HIV-1 protease inhibitors: a predictive QSAR modelling study based on Monte Carlo optimization. *SAR and QSAR in Environmental Research* **2017**, 28, 973-990.

47. Ghosh, A. K.; Williams, J. N.; Kovala, S.; Takayama, J.; Simpson, H. M.; Walters, D. E.; Hattori, S.-i.; Aoki, M.; Mitsuya, H., Potent HIV-1 protease inhibitors incorporating squaramide-derived P2 ligands: Design, synthesis, and biological evaluation. *Bioorganic & Medicinal Chemistry Letters* **2019**, 29, 2565-2570.

48. Ju, H.; Xiu, S.; Ding, X.; Shang, M.; Jia, R.; Huang, B.; Zhan, P.; Liu, X., Discovery of novel 1,2,3-triazole oseltamivir derivatives as potent influenza neuraminidase inhibitors targeting the 430-cavity. *European Journal of Medicinal Chemistry* **2020**, 187, 111940.

49. And er, M.; Luzhkov, V. B.; Aqvist, J., Ligand binding to the voltage-gated Kv1.5 potassium channel in the open state--docking and computer simulations of a homology model. *Biophys J* **2008**, 94, 820-831.

50. Clau en, H.; Buning, C.; Rarey, M.; Lengauer, T., FlexE: efficient molecular docking considering protein structure variations1 1Edited by J. Thornton. *Journal of Molecular Biology* **2001**, 308, 377-395.

Caption for Figures.

Fig. 1 Number of COVID-19 between 1-30 March, 2020 globally: A) Confirmed cases; B) Death reported (*Source:* WHO. Novel coronavirus (2019-nCoV) situation reports. Situation report – 41-70. 1-30 March, 2020).

Fig. 2. 3CL^{pro} enzyme of SARS-CoV-2. A) showing 3 domains and catalytic residues; B) functioning.

Fig. 3. Indinavir-3CL^{pro} docked complex, showing interaction with crystal structure 6LU7 (brown color) and surrounded binding site residue.

Fig. 4. Compound **16**-3CL^{pro} docked complex, showing interaction with crystal structure 6LU7 (brown color) and surrounded binding site residue.

Fig. 5. Ramachandran plot for 3CL^{pro} in case of: A) Compound **16**-3CL^{pro} complex; B) Indinavir-3CL^{pro} complex; C) un-ligated-3CL^{pro} complex.

Fig. 6. Simulation Quality parameters for Compound **16**-3CL^{pro} complex for 100 ns simulation in terms of total energy (E), potential energy (E_p), pressure (P), temperature (T), and volume (V).

Fig. 7. RMSD plot for C α of 3CL^{pro}: a) **16**-3CL^{pro} complex (blue color); b) indinavir-3CL^{pro} complex (orange color) and c) un-ligated-3CL^{pro} complex (grey color).

Fig. 8. A) Ligand RMSD plot for compound **16** and indinavir and B) Ligand-RMSF plot for compound **16**-3CL^{pro} complex.

Fig. 9. RMSF plot for C α of 3CL^{pro} residues in compound **16**-3CL^{pro} complex, residues are shown in three letter code with green color belong to binding site interacting to compound **16**.

Fig. 10. Histogram (stacked bar chart) showing compound **16**-3CL^{pro} forming H-bonds interactions (green color), hydrophobic interaction (grey violet color), and water bridges (blue color) during 100 ns simulation.

Fig. 11. Timeline representation of the interactions and contacts (H-bonds, Hydrophobic, Water bridges) in compound **16**-3CL^{pro} complex during the 100ns MD simulation.

Fig. 12. Ligand properties during the 100-ns simulations for compound **16**: (A) Ligand RMSD - Root mean square deviation; (B) Radius of Gyration (rGyr); (C) NS34; (D) Molecular Surface Area (MolSA); (E) Solvent Accessible Surface Area (SASA); and (F) Polar Surface Area (PSA).

Fig. 13. Torsional analysis of ligand-**16** NS34 conformations during the 100-ns simulations.

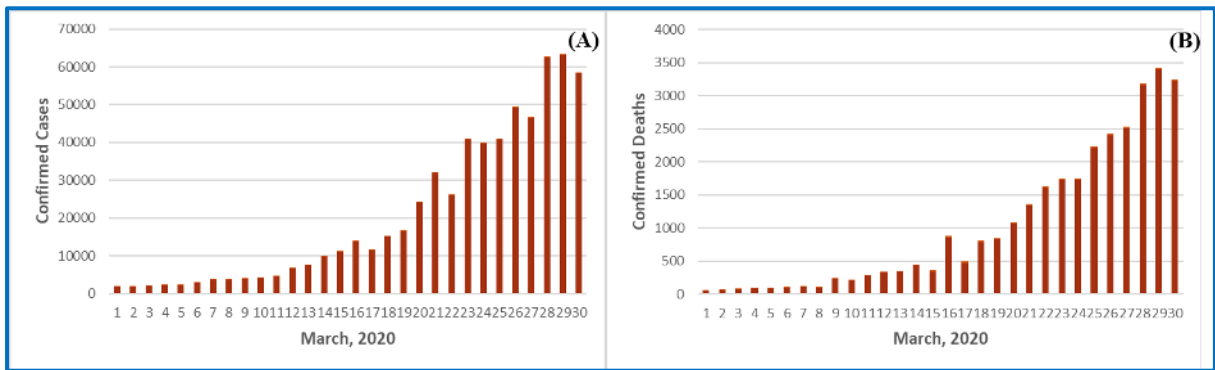


Fig. 1.

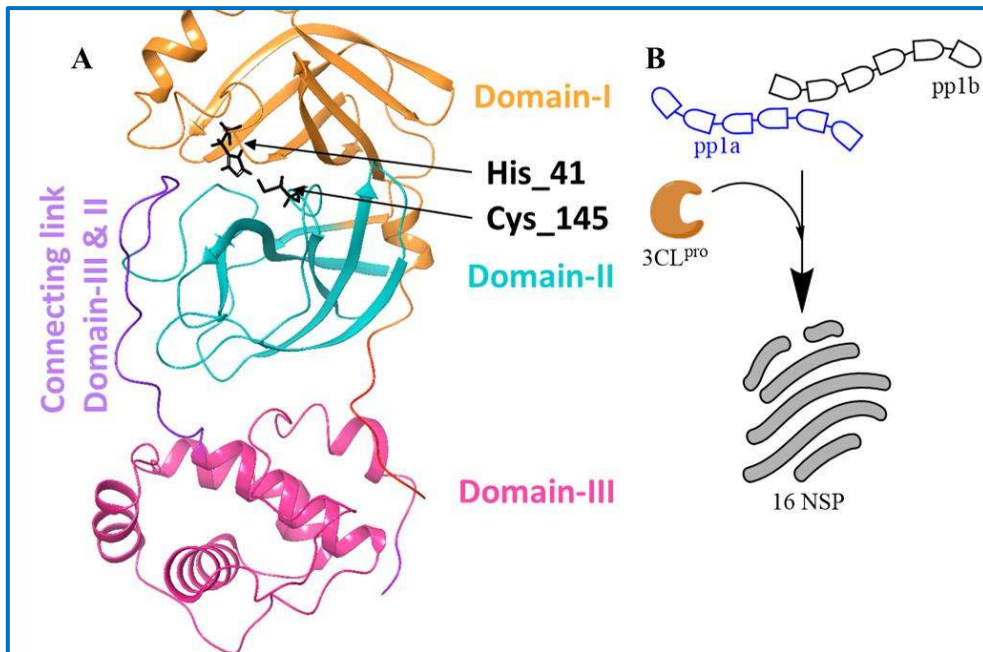


Fig. 2.

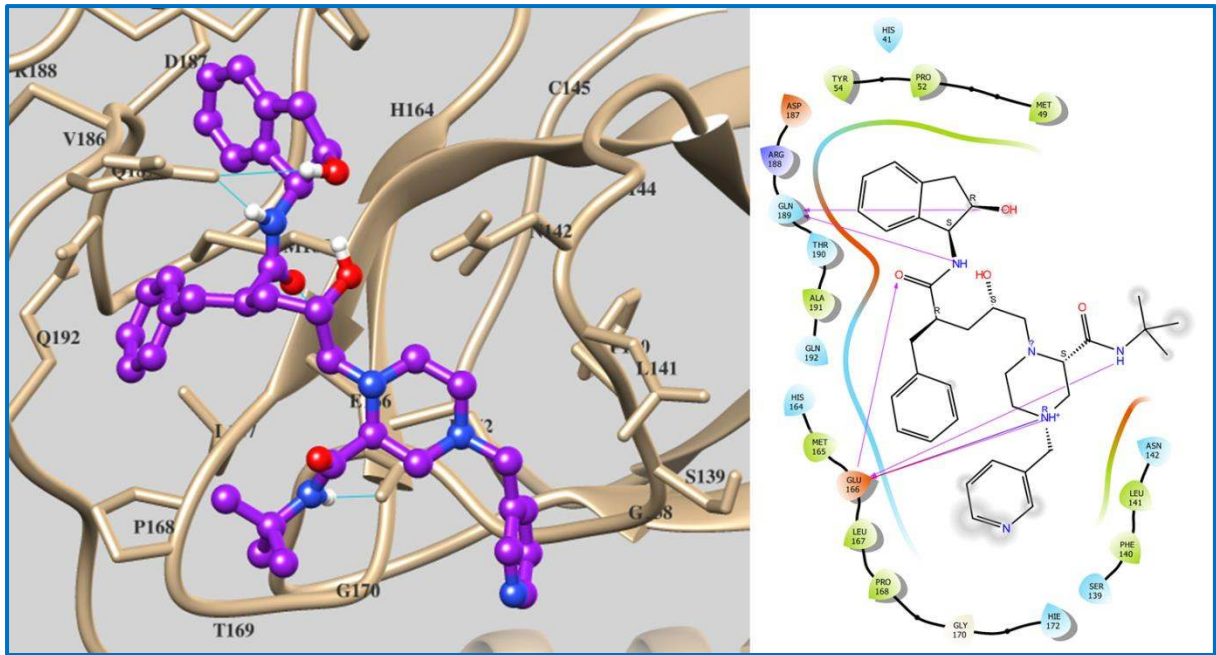


Fig. 3.

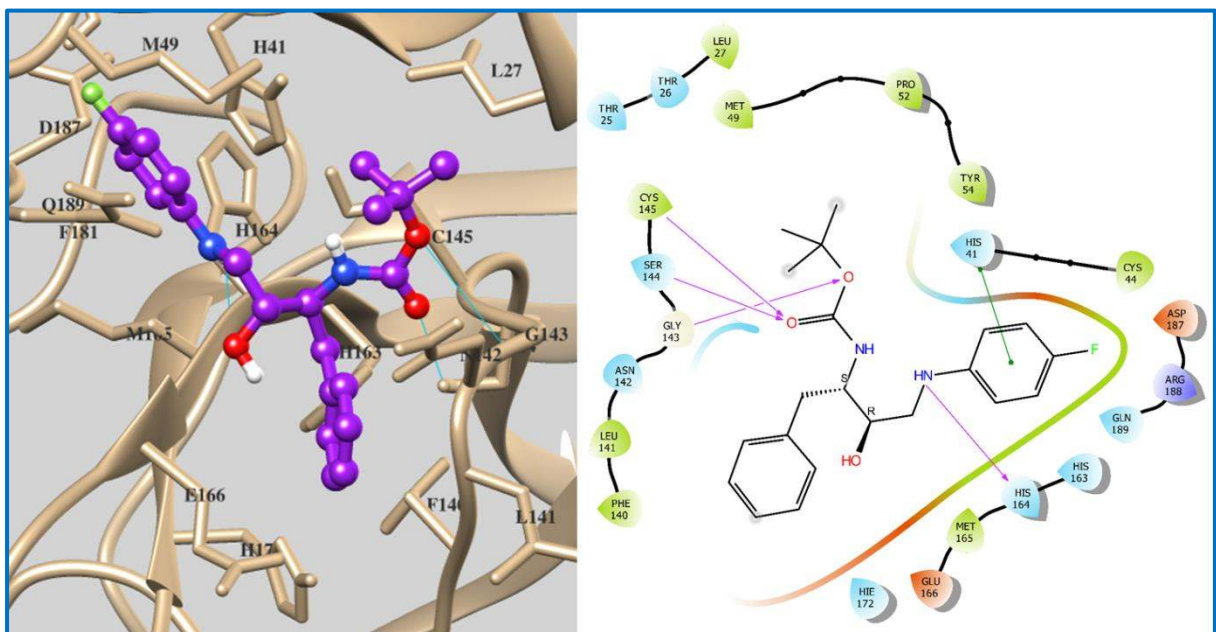


Fig. 4.

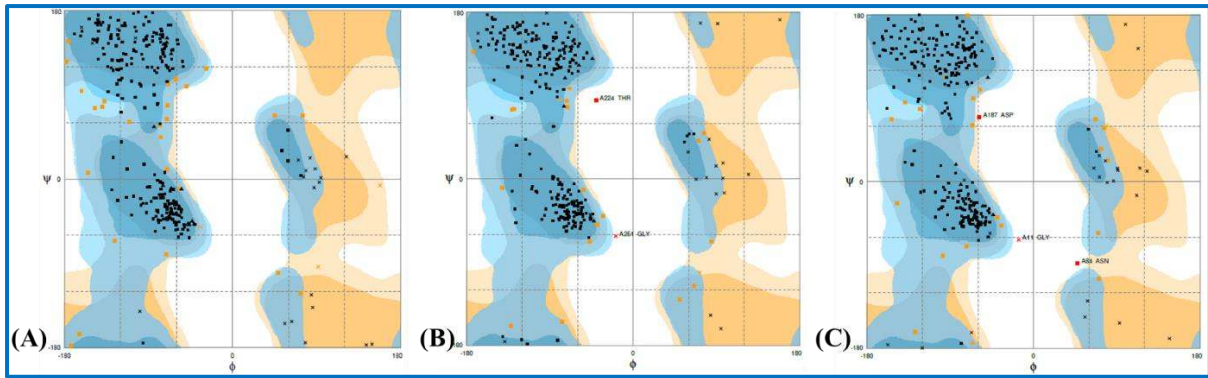


Fig. 5

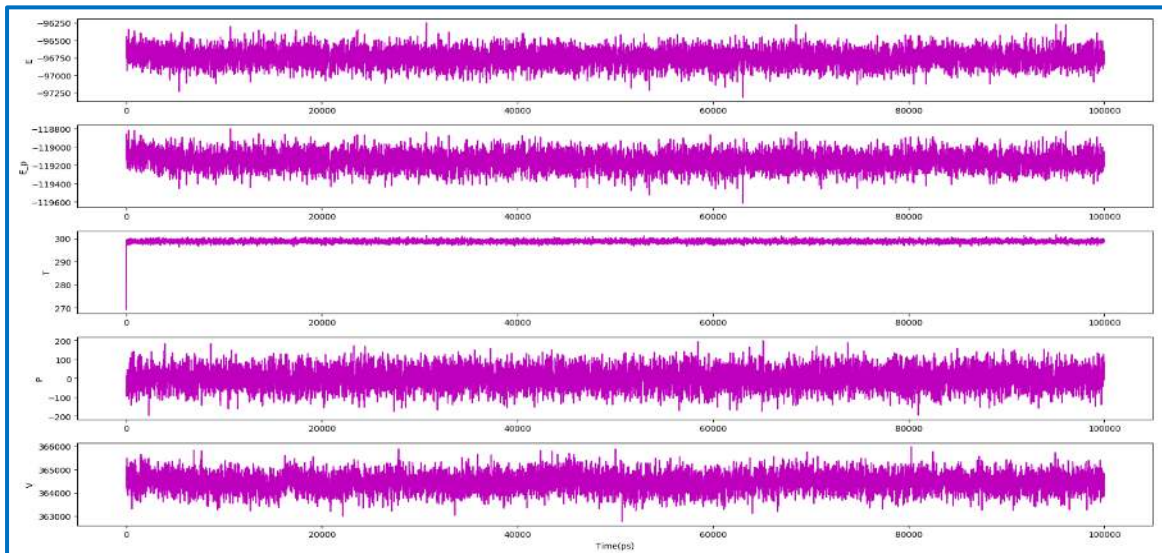


Fig. 6.

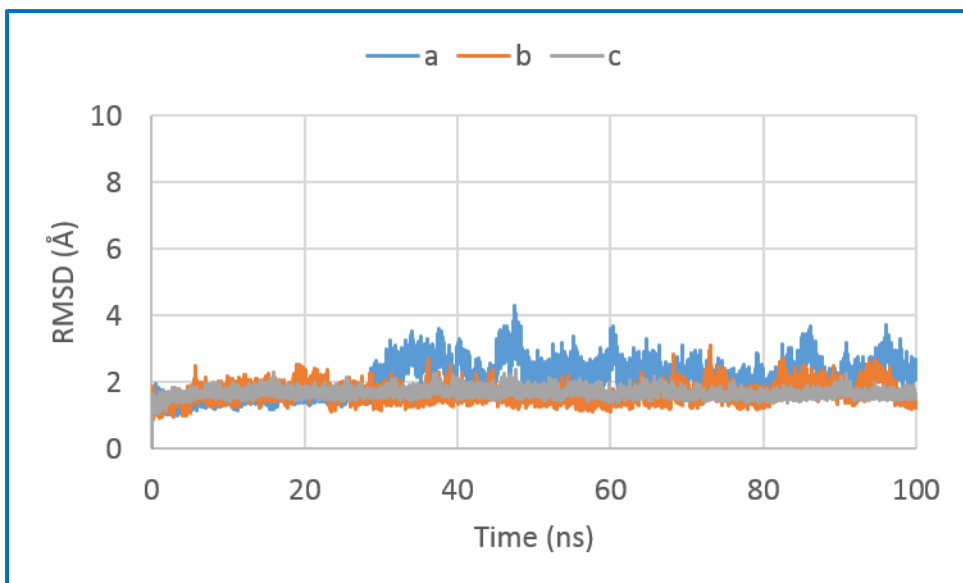


Fig. 7.

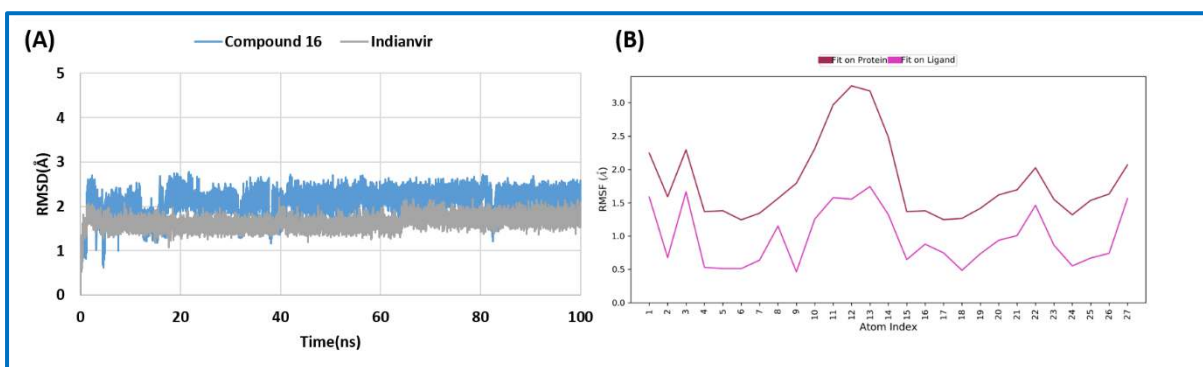


Fig. 8.

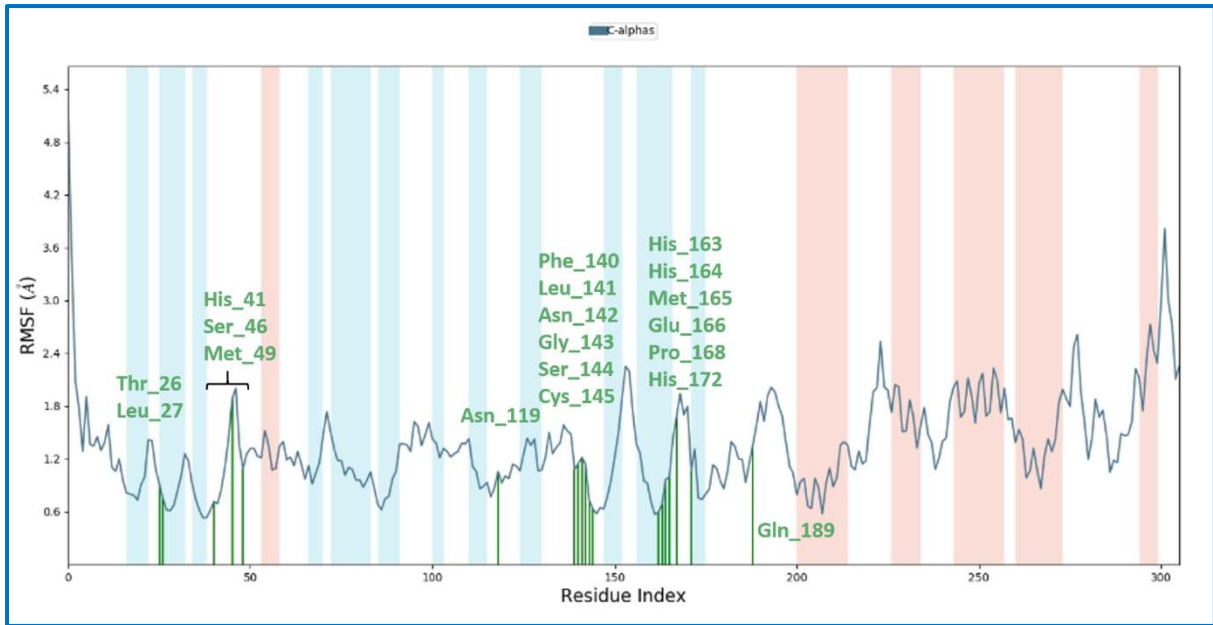


Fig. 9.

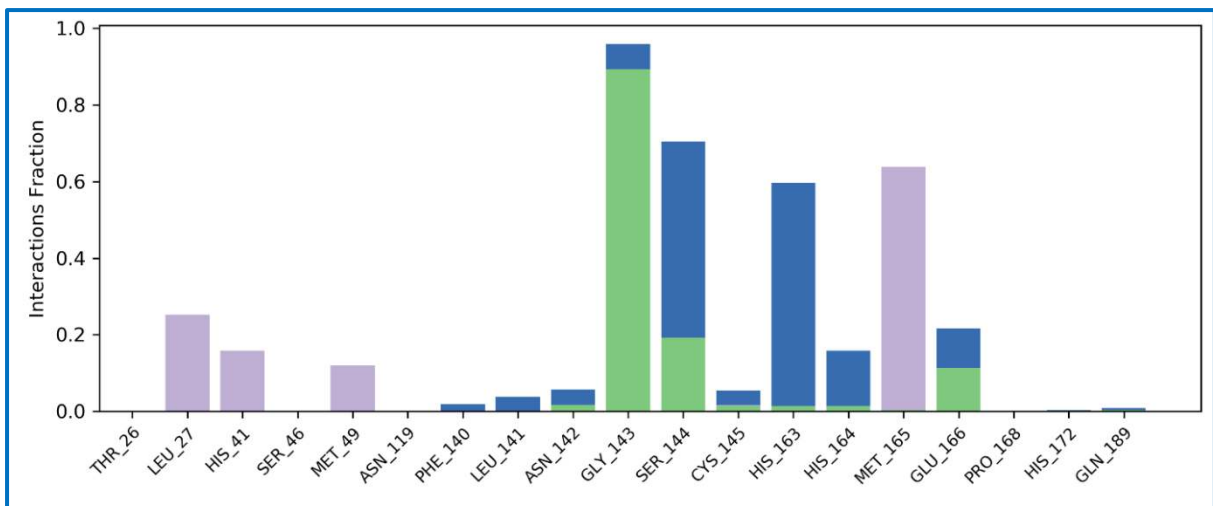


Fig. 10.

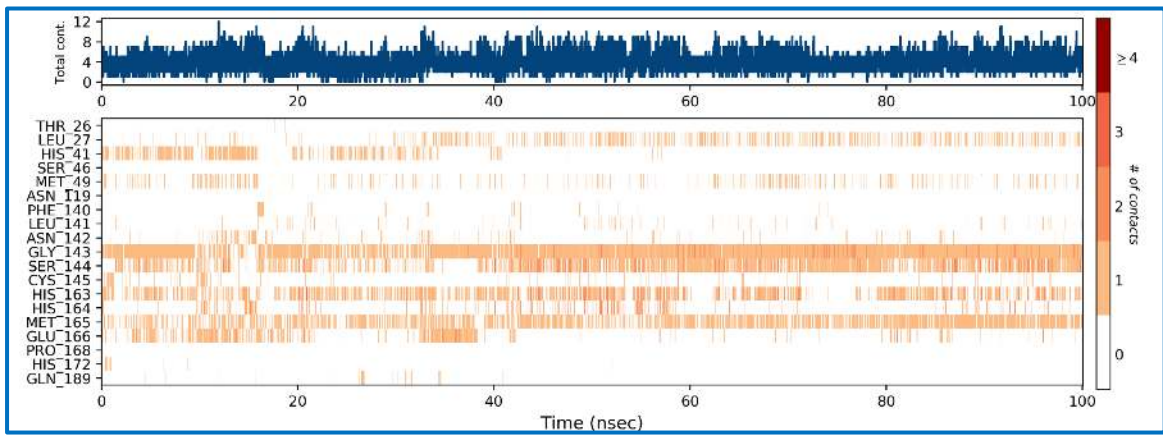


Fig.11.

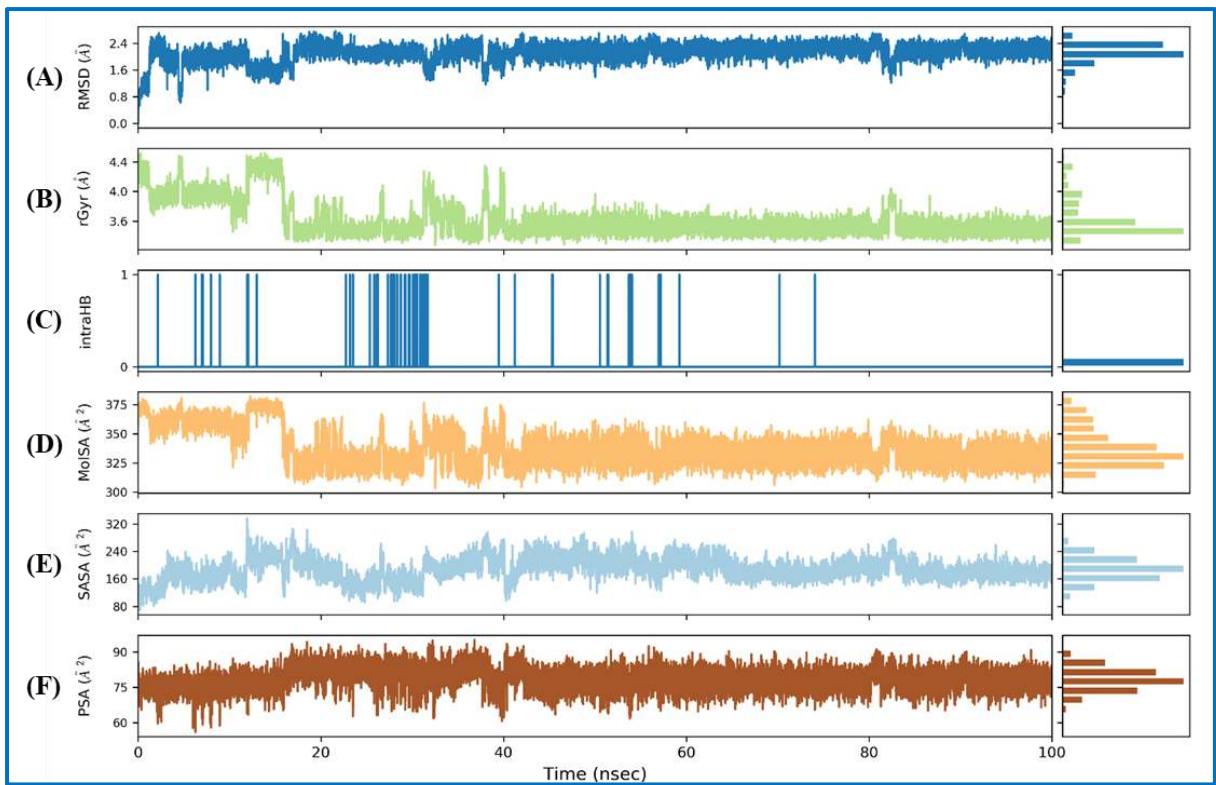


Fig. 12.

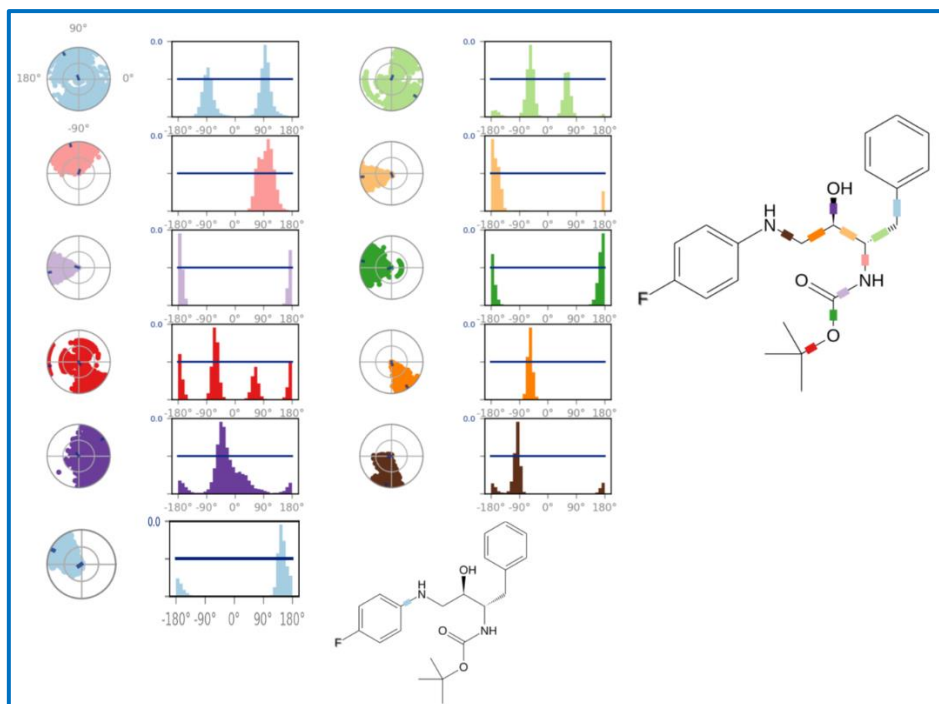
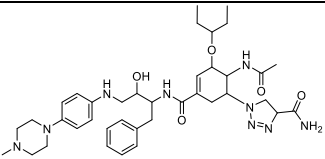
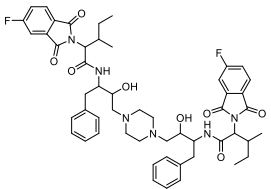
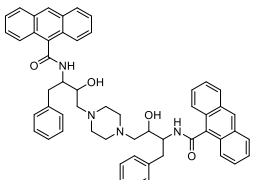
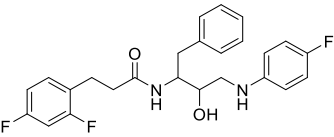
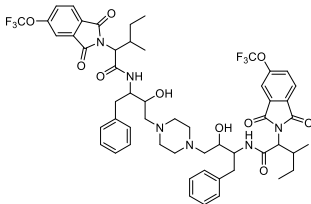
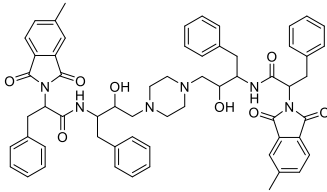
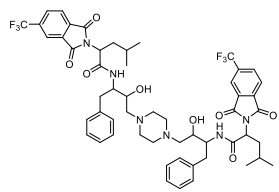


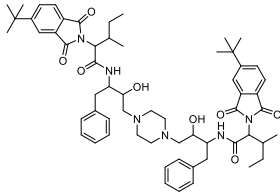
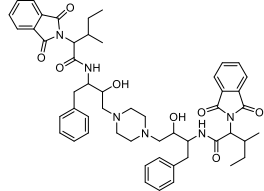
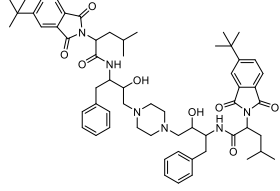
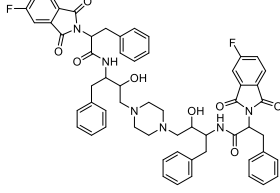
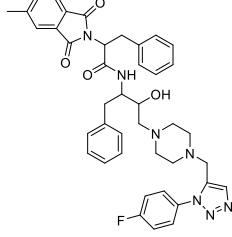
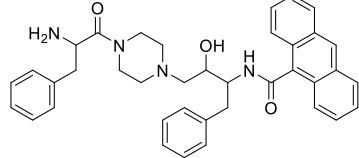
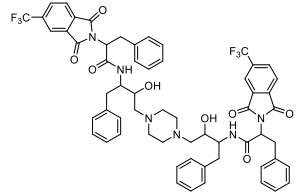
Fig. 13.

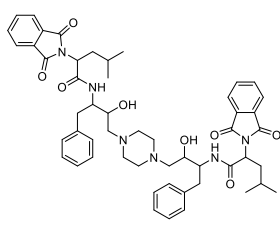
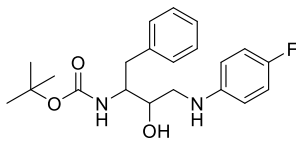
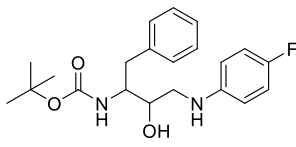
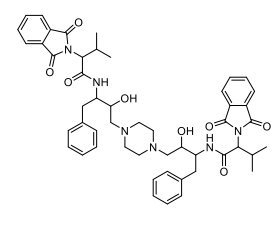
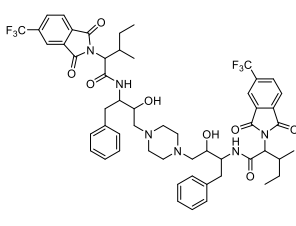
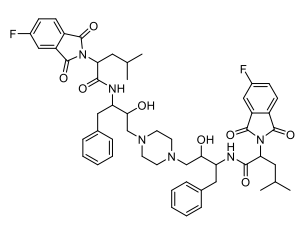
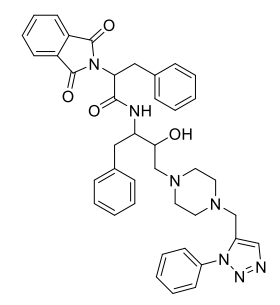
Table 1. Docking score and XP Gscore (kcal/mol) for antiviral drugs.

Entry	Drug	Docking score	XP Gscore (kcal/mol)
1	Indinavir	-8.824	-9.466
2	Atazanavir	-7.912	-7.92
3	Remdesvir	-7.804	-7.804
4	Amprenavir	-7.747	-7.747
5	Saquinavir	-7.455	-7.468
6	Ritonavir	-7.422	-7.422
7	Lopinavir	-7.041	-7.041
8	Darunavir	-7.028	-7.028
9	Nelfinavir	-6.73	-6.744
10	Oseltamivir	-5.825	-5.907
11	Tipranavir	-5.64	-5.778
12	Fosamprenavir	-5.309	-6.578
13	Galidesvir	-4.967	-6.322

Table 2. Docking score and XP Gscore (kcal/mol) for hit HEA molecules (**1-25**).

Entry No	Compound	docking score	XP Gscore
1	 1	-9.864	-10.227
2	 2	-9.832	-9.851
3	 3	-9.791	-9.81
4	 4	-9.713	-9.713
5	 5	-9.685	-9.704
6	 6	-9.446	-9.465
7	 7	-9.424	-9.443

8		8	-9.294	-9.313
9		9	-9.281	-9.3
10		10	-9.24	-9.259
11		11	-9.229	-9.248
12		12	-9.198	-9.203
13		13	-9.097	-9.402
14			-9.017	-9.036

15	<p style="text-align: center;">14</p> 		
16	<p style="text-align: center;">15</p> 	-8.974	-8.993
17	<p style="text-align: center;">16</p> 	-8.955	-8956
18	<p style="text-align: center;">17</p> 	-8.954	-8.973
19	<p style="text-align: center;">18</p> 	-8.952	-8.971
20	<p style="text-align: center;">19</p> 	-8.935	-11.053
20	<p style="text-align: center;">20</p> 	-8.927	-8.932

21		-8.899	-8.918
22		-8.893	-8.911
23		-8.849	-9.048
24		-8.83	-8.84
25		-8.825	-8.843

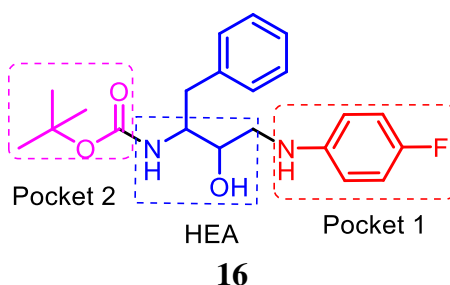
Table 3. Physicochemical properties (ADME) of best compounds.

Entry No.	Compound	MW	nON	nOHNH	Nrot	TPSA	mLogP	LD50	Hept.	nVio.
1	1	717.9	8	5	18	177.2	0.47	2.88	Yes	2
2	2	935.1	12	4	22	179.9	2.83	2.48	Yes	2
3	3	821.0	6	4	16	105.1	4.74	2.48	Yes	1
4	4	442.5	5	3	11	61.4	4.68	2.32	Yes	1
5	5	1067	18	4	26	198.3	1.96	2.47	Yes	2
6	6	995.1	10	4	22	179.9	3.23	2.48	Yes	2
7	7	1035	16	4	24	179.9	3.17	2.45	Yes	2
8	8	1011	10	4	24	179.9	3.61	2.46	Yes	2
9	9	899.1	10	4	22	179.9	2.41	2.45	Yes	2
10	10	1011	10	4	24	179.9	3.61	2.47	Yes	2
11	11	1003	12	4	22	179.9	3.35	2.48	Yes	2
12	12	715.8	9	2	14	123.9	3.41	2.52	Yes	2
13	13	600.7	5	3	12	98.9	3.18	2.44	Yes	1
14	14	1103	16	4	24	179.9	3.66	2.48	Yes	2
15	15	899.1	10	4	22	179.9	2.41	2.45	Yes	2
16	16	374.4	4	3	10	70.6	3.31	2.79	No	0
17	17	871.0	10	4	20	179.9	2.1	2.44	Yes	2
18	18	1035	16	4	24	179.9	3.17	2.46	Yes	2
19	19	935.1	12	4	22	179.9	2.83	2.46	Yes	2
20	20	683.8	8	2	14	123.9	2.89	2.51	Yes	2
21	21	1135	18	4	16	198.3	2.46	2.48	Yes	2
22	22	927.1	10	4	22	179.9	2.72	2.45	Yes	2
23	23	524.6	5	3	10	98.9	2.3	2.41	Yes	1
24	24	652.8	6	3	15	111.2	3.14	2.35	Yes	1
25	25	1079	10	4	24	179.9	4.08	2.48	Yes	2

MW = molecular weight (g/mol); nON = no. of hydrogen bond acceptor; nOHNH = no. of hydrogen bond donors; Nrot = no. of rotatable bonds; TPSA = total polar surface area; MLogP = Predicted octanol/water partition coefficient ; LD50 = Oral Rat Acute Toxicity; Hept. = Hepatotoxicity; nVio. = no. of Lipinski violation.

Table 4. List of interacting residues of 3CL^{pro} with indinavir and lead compound **16**.

Entry No	Compound	H-bond Residues	Salt bridge residues	Pi-pi-Interaction residues	Hydrophobic Residues	Domain
1.	Indinavir	Glu_166, Gln_189	Glu_166	NA	Met_49, Pro_52, Tyr_54, Phe_140, Leu_141, Met_165, Leu_167, Pro_168, Gly_170 Ala_191	Domain-I Domain-II Domain-II-III Linker
2.	16	Gly_143, Ser_144, Cys_145, His_164	NA	His_41	Leu_27, Cys_44, Met_49, Pro_52, Tyr_54, Phe_140, Leu_141, Gly_143, Cys_145, Met_165	

Table 5. SAR analysis.

Entry No	Compound	docking score	XP GScore (kcal/mol)
1	 26	-8.835	-8.836
2	 27	-8.55	-8.551

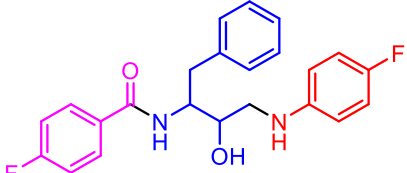
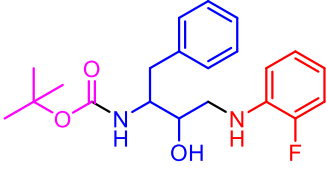
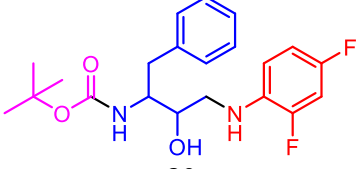
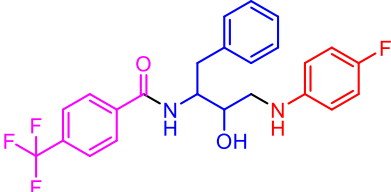
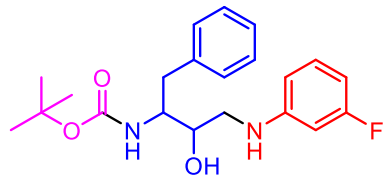
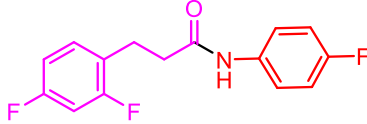
3	 <p>28</p>	-8.543	-8.544
4	 <p>29</p>	-8.131	-8.131
5	 <p>30</p>	-7.995	-7.995
6	 <p>31</p>	-7.965	-7.965
7	 <p>32</p>	-7.949	-7.949
8	 <p>33</p>	-7.121	-7.121

Table 6. Number and percentage of residues exist in favoured, allowed and outlier region for simulated system.

Entry No	System	Number and percentage of residues in favoured region	Number and percentage of residues in allowed region	Number and percentage of residues in the outlier
1.	Compound 16-3CL ^{pro}	274 (90.1%)	30 (9.9%)	0.0
2.	Indinavir-3CL ^{pro}	282 (92.8%)	20 (6.6%)	2 (0.7%)
3.	Un-ligated 3CL ^{pro}	280 (92.1%)	21 (6.9%)	3 (1.0%)

Graphical Abstract

Novel Inhibitors of SARS-CoV-2 main protease 3CL^{pro}

