Contents lists available at ScienceDirect

Heliyon

Research article

Evaluation of SARS-CoV-2 main protease and inhibitor interactions using dihedral angle distributions and radial distribution function

Tika Ram Lamichhane [a,*], Madhav Prasad Ghimire [a,b,*]

[a] Central Department of Physics, Tribhuvan University, Kathmandu, Nepal
[b] Condensed Matter Physics Research Center (CMPRC), Butwal, Rupandehi, Nepal

ARTICLE INFO

ABSTRACT

In order to evaluate the interactions between a potential drug candidate like inhibitor N3 and the residues in substrate binding site of SARS-CoV-2 main protease (M^{pro}), we used molecular docking and dynamics simulations. The structural features describing the degrees of folding states of M^{pro} formed by beta-barrels and alpha-helices were analyzed by means of root mean square deviation, root mean square fluctuation, radius of gyration, residue velocity, H-bonding, dihedral angle distributions and radial distribution function. All of the residues forming ligand binding domain (LBD) of M^{pro} lie within the allowed region of the dihedral angle distributions as observed from the equilibrating best pose of M^{pro}-N3 system. Sharp peaks of radial distribution function (RDF) for H-bonding atom pairs (about 2 Å radial distance apart) describe the strong interactions between inhibitor and SARS-CoV-2 M^{pro}. During MD simulations, HSE163 has the lowest residue speed offering a sharp RDF peak whereas GLN192 has the highest residue speed resulting a flat RDF peak for the H-bonding atom pairs of M^{pro}-N3 system. Along with negative values of coulombic and Lenard-Jones energies, MM/PBSA free energy of binding contributed by the non-covalent interactions between M^{pro} and N3 has been obtained to be -19.45 ± 3.6 kcal/mol. These physical parameters demonstrate the binding nature of an inhibitor in M^{pro}-LBD. This study will be helpful in evaluating the drug candidates which are expected to inhibit the SARS-CoV-2 structural proteins.

1. Introduction

From the origin of COVID-19 until now, many researches have sought to identify the potential inhibitors of SARS-CoV-2 structural proteins. The SARS-CoV-2 proteins have ligand binding domains (LBD) which are targets for the potential inhibitors [1, 2, 3, 4, 5]. The researchers are working clinically as well as computationally on the different types of potential antiviral drugs in order to search for the effective control of the ongoing pandemic. The best binding drugs like inhibitor N3 [1, 2], α-ketoamide derivative [6], cannabisin A and darunavir [7] execute stable interactions to inhibit the SARS-CoV-2 main protease (M^{pro}). These types of drugs prevent the viral transcription and replication by interfering with the polypeptide transnational activities [3]. Early research focused on the study of covalent inhibitors targeting the catalytic HIS41-CYS145 dyad which may result in off-target side effects and toxicity [8]. Noncovalent inhibitors are beneficial for the easy reaction pathway of ligand binding and dissociation. Some vaccines have been brought into practice along with the clinical trials to

prevent coronavirus transmission. Understanding drug impacts on the molecular level through the analysis of protease-inhibitor interactions is a good way to evaluate effective drugs.

The patients with SARS-CoV-2 are susceptible to pneumonia-induced acute respiratory distress syndrome and multi-organ failure. Human genetics, host immune response, age and comorbidities influence the cytokine release syndrome [9]. The coronavirus consists of a positive-sense, single-stranded RNA genome with about 30 thousand nucleotides which encodes two overlapping polyproteins pp1a and pp1ab. The M^{pro}, also known as $3CL^{pro}$, is a 33.8 kDa cysteine protease that conserves functional polypeptides released by proteolytic processing of pp1a and pp1ab [2]. The key enzyme, SARS-CoV-2 M^{pro} plays an important role in gene transcription activity. The M^{pro}-LBD has the catalytic HIS41-CYS145 dyad in between I and II domains comprising of residues 8-101 and 102-184, respectively, and including an anti-parallel beta-barrel [2, 4]. Flexible alpha-helices and strong beta-barrel structures forming three domains of the M^{pro} are shown in Fig. 1-a. It is an active site of SARS-CoV-2 to evaluate the interaction of potential
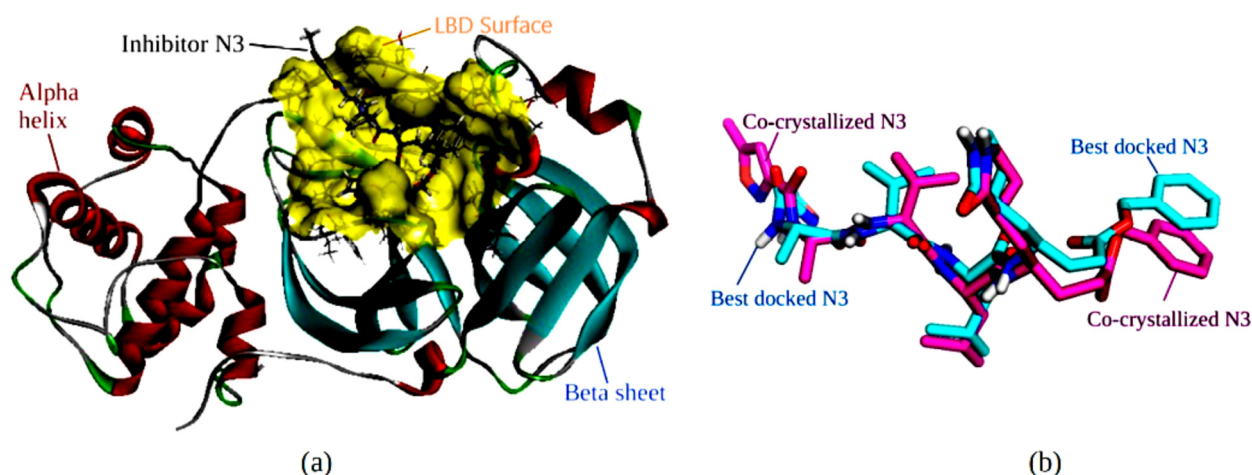

Fig. 1. (a) M^{pro}-LBD surface binding inhibitor N3, (b) Co-crystallized docking of N3 in SARS-CoV-2 main protease (M^{pro}).

drug candidates with the LBD residues. The peptidomimetic inhibitor N3 with well designed side chains fits finely in the substrate binding pocket. The M^{pro}-LBD residues are: MET49, PHE140, ASN142, GLY143, CYS145, HSE163-164, MET165, GLU166, LEU167, PRO168, HSE172, GLN189, ALA191 and GLN192 as identified from the known structure, 6LU7.PDB [4].

The computational methods such as molecular docking and dynamics simulations are widely used to identify the binding efficacy of the potential drug and its stability into the target viral receptors. Here, after the co-crystallized docking of inhibitor N3 into the M^{pro}-LBD (Fig. 1-b), structural features of the SARS-CoV-2 M^{pro}-LBD have been evaluated by means of physical parameters such as root mean square deviation [10, 11], root mean square fluctuation [11], radius of gyration [11, 12], radial distribution function [13, 14], dihedral angle distributions [15, 16] and solvent accessibility [17, 18]. Furthermore, the M^{pro}-N3 interactions are evaluated by means of time-based changes of H-bonds, residue speed, coulombic and Lenard-Jones energies, and binding energy calculations during MD simulations [19]. The free energy of binding of the M^{pro}-inhibitor system is obtained by an end-point method, i.e. Molecular Mechanics / Poisson Boltzmann Surface Area (MM/PBSA) calculations [20, 21]. This type of study is important to explore the nature of interactions between the potential drug candidate and the viral structural protein. As a biophysical perspective, the concept of clinical data modeling for the trendline of coronavirus cases and deaths, and SARS-CoV-2 structural features are shortly reviewed by Lamichhane and Ghimire, 2020 [22].

2. Methodology

2.1. Molecular docking

The primary structure of SARS-CoV-2 M^{pro}-N3 system was taken from the crystallographic data, 6LU7.PDB [4]. The protein was prepared with PyMOL deleting everything except amino acid residues and it was protonated by adding hydrogen atoms into it [23]. Thus prepared M^{pro} structure was uploaded to Swiss PDB Viewer and it was optimized selecting all residues using gromos96 43B1 force field [24]. The structure of inhibitor N3 obtained from 6LU7.PDB was cross-checked with PyMOL, protonated and then uploaded to ACD/ChemSketch 12.0 for its geometrical optimization [25]. The electrically neutral N3 with 99 atoms has the mass of 682.81 amu. The ligand moiety, N3 was docked non-covalently in the co-crystallized form of its original position in 6LU7.PB with the resulting RMSD 1.66 Å as determined from the Dock-RMSD tool (Fig. 1-b) [26]. The docking software used was AutoDock 4.2 [27]. The AutoDock Tool performs grid-based energy evaluation using Lamarckian genetic algorithm [28]. During the docking simulations,

side chains and torsional bonds of N3 were rotated freely by keeping the receptor M^{pro} system rigid. Gasteiger charges [29] were computed by adding polar hydrogens to M^{pro}. Fifty runs of docking simulations were performed with 300 population size, 2,500,000 evaluations and 27,000 generations. To generate the docking box of actual mapping, grid point spacing of 0.375 Å, coordinates of central grid points (x = -9.204 Å, y = 14.323 Å, z = 67.590 Å) and grid (x, y, z) points of (64, 90, 70) were employed in the receptor M^{pro}-LBD. To obtain the best docking pose of N3 in the M^{pro} system, the process was revised five times. The best docked conformer of main protease-inhibitor (M^{pro}-N3) was then subjected to nanoscale molecular dynamics (NAMD) simulations for the purpose of evaluating the stable interactions by means of equilibration techniques.

2.2. MD simulations

A protein structure file (PSF) of M^{pro}-N3 system was generated using topology files and NAMD simulations were performed using parameter files of CHARMM 36 force fields [30, 31, 32]. The PSF contains an information of atom types, atomic masses and partial atomic charges. The missing H-atoms in protein data bank (PDB) file of M^{pro}-N3 system were assigned automatically during the PSF generation. The parameter file contains the values of numerical constants which are used to calculate bonded and non-bonded energies of CHARMM potential function [19]. The N3 was obtained from the best pose conformer of M^{pro}-N3 system after the molecular docking simulations. The missing hydrogen atoms were added to N3 and it was subjected to SwissParam server [33] to generate the related parameter and topology files. The related topology files were used to generate PSF of both M^{pro} and inhibitor N3, and both structures were merged to form the full PSF and PDB files of M^{pro}-N3 complex with the help of MergeStructs plugin-1.1 of visual molecular dynamics (VMD) [34]. This method preserves all of the original bonds and existing coordinates in both M^{pro} and N3 molecules so that the original pose of N3 remains unchanged while liganded in the M^{pro}-LBD. Thus prepared M^{pro}-N3 structure assigning with all H-atoms was fully solvated in a periodic box of 41040 TIP3P-water molecules. The system was neutralized with 0.15 mol/L concentration of Na$^+$ and Cl$^-$ ions in order to provide a cellular environment. The software packages used were VMD-1.9.3 for structure generation and analysis, and NAMD-2.12 for geometrical optimization, equilibration run and production run of MD simulations.

The M^{pro}-N3 water box had the cell basis vectors of 71.38 Å, 86.70 Å and 79.88 Å with center coordinates (-25.83, 12.78, 58.05 Å). Particle mesh Ewald (PME) was active with Ewald coefficient 0.25, grid spacing 1.0 Å and grid dimensions (72, 88, 80 Å). The NAMD simulations were performed in NPT ensemble using Langevin piston temperature,

pressure and damping coefficient as 300 K, 1.0 bar and 1 ps^{-1}, respectively. The MD simulations of M^{pro}-N3 system were performed using a time step of 2 fs and constraining the H-atoms with rigid bonds. The coulombic and van der Waals interactions were evaluated using 1-4 scaling and distance related parameters: 10 Å, 12 Å and 14 Å for switching, cut-off and pair-list distances, respectively. The number of steps implemented for energy minimization or geometrical optimization of M^{pro}-N3 system were 3000 CG steps. The system was subjected to 50 ns equilibrating simulation at the constant temperature 300 K. Verlet algorithm [35] was used to calculate the atomic trajectories. With the help of NAMD-Energy Plugin-1.4, the interaction energies (electrostatic and van der Waals) were evaluated. The residue speed or displacement per frame were observed by means of VMD Timeline-2.3 [34]. The time-based changes of the physical parameters (root mean square deviation, root mean square fluctuation, residue velocity, dihedral angle distributions, radial distribution function, solvent accessible surface area and H-bonding) describe the structural properties of protein-ligand systems as studied in the previous researches [36, 37, 38, 39]. These tools were also used to evaluate the structural changes of the M^{pro}-inhibitor system during the MD simulations. All the interaction visualization analysis were performed by PyMOL molecular visualization tools [23], VMD tools [34], BIOVIA discovery studio visualizer [40] and LIGPLOT$^+$ [41].

The radial distribution function (RDF) denoted by $g(r)$ gives the probability of finding the positions of an atom-pair within a specified range of radial distance during MD simulations. The normalized RDF for the atom pairs j, k within a spherical shell having radius r and thickness dr is defined by Equation (1) [13, 14, 36].

$$g_{j,k}(r) = \frac{N_{j,k}(r)/4\pi r^2}{\sum_r (N_{j,k}(r)/4\pi r^2)} \tag{1}$$

where $N_{j,k}(r)$ is the number of occurrences of atom pairs between r and $r + dr$.

Likewise, solvent accessible surface area (SASA) is also observed to know the structural properties and nature of interactions of a biomolecule. The SASA is an area described by the solvent molecule when it rolls around the molecular complex approaching to van der Waals contacts of the atoms. It is calculated by using Equation (2) [17, 18, 36].

$$SASA = \sum_i (R/\sqrt{R^2 - Z_i^2}).D.L_i \tag{2}$$

where R is obtained by adding radius of solvent molecule and van der Waals radius of the atom, Z_i is the distance between sphere center and section i, D is the distance between two consecutive i^{th} and $(i + 1)^{th}$ sections, and L_i is arc length computed on the section i.

2.3. MM/PBSA method of binding free energy

The binding free energy of M^{pro}-N3 complex was calculated by using MM/PBSA method in CaFE1.0 plugin [20]. The CaFE uses NAMD trajectories supporting CHARMM force field parameters. The end-point energy method, MM/PBSA consists of three terms: molecular mechanics (MM) for gas-phase energy, Poisson-Boltzmann (PB) for polar solvation energy and surface area (SA) for nonpolar solvation energy [21]. The free energy of binding, ΔG_{bind} can be estimated by using Equation (3) [20].

$$\Delta G_{bind} = \Delta H - T\Delta S = \left\langle \Delta G_{gas} + \Delta G_{sol}^{polar} + \Delta G_{sol}^{nonpolar} - T\Delta S \right\rangle \tag{3}$$

where the MM term, i.e. gas-phase energy (ΔG_{gas}) was obtained from non-bonded (coulombic and Lennard-Jones) interactions between inhibitor and protein. The PB term, i.e. ΔG_{sol}^{polar} was calculated by setting interior dielectric constant 2 and exterior dielectric constant 80 in APBS program [42] interfaced to CaFE. The SA term, i.e. $\Delta G_{sol}^{nonpolar}$ was calculated by setting an offset 0.92 kcal/mol and surface tension 0.00542 kcal/mol/Å^2. Here, the SA term has an approximate linear relation with

SASA. The entropy term $T\Delta S$ is ignorable for a larger ligand [43] and also for a group of structurally similar compounds [20]. Due to inconsistency in the results by roughly entropic approximation, CaFE excludes the entropy term and sums the results from MM, PB and SA terms averaged over the saved snapshots of simulated system.

3. Results and discussion

The molecular docking based analysis explains that the antiviral drugs against SARS-CoV-2 reveal strong interactions with higher binding affinity or smaller inhibition constant. From our observations, the inhibitor N3 has binding energy of -9.53 kcal/mol, inhibition constant of 103.5 nM and the best pose RMSD of 1.66 Å with respect to its co-crystallized form in the M^{pro}-LBD (6LU7.PDB). To evaluate the binding stability of the docked complexes, MD simulations are performed and the equilibrated structures are analyzed using the different tools as explained in the methodology section.

The electrically neutral M^{pro}-N3 solvated water box has the mass of 283297 amu consisting of 45903 atoms, 32196 bonds, 22385 angles, 12738 dihedrals, 890 impropers, 304 cross terms, 16179 H-groups and 94305 degrees of freedom. These parameters collectively describe the conformational and binding properties of the receptor-ligand system [19]. The protein structure and stability depends on the number of H-bonding residues in the protein backbone and side chains. As a binary two dimensional matrix, Fig. 2-a shows the contact map of H-bonding amino acid residues obtained by machine learning technique [38] after 50 ns equilibrations of SARS-CoV-2 M^{pro} system. The contact map is invariant of rotations and translations which distinguishes nature of H-bonding interactions between main chains and/or side chains as indicated by dark, blue and red points in Fig. 2-a. Here, alpha-helices, antiparallel beta-sheet and parallel beta-sheet appear adjacent to diagonal, cross-diagonal and parallel to diagonal, respectively. In this way, root mean square deviation (RMSD), root mean square fluctuation (RMSF) and radius of gyration (RG) describe the conformational stability of the protein backbone forming alpha- and beta-strands and side chain residues. Their time-based changes explain the degree of folding states and compactness of the system by acting as the reaction coordinates. The average RMSD of the viral receptor M^{pro} is 2.40 ± 0.34 Å and that of the inhibitor N3 in LBD is 1.41 ± 0.32 Å provided with standard deviation (SD) as the error value. The small fluctuations seen in the RMSD plots (Fig. 2-b) are due to the minutely changing number of H-bonds among the protein residues during the equilibration run. The mean RG is 22.70 ± 0.20 Å for M^{pro} and 6.61 ± 0.25 Å for N3. The RMSF of each residue has been shown in Fig. 2-c whose value lies between 0.3 Å and 1.5 Å. The RMSF of terminal residues SER1 and GLN306, and the intermediate residues TYR154 and ASP155 has larger values than the other residues of showing higher flexibility in the M^{pro} system. Thus, the average measures of RMSD, RMSF and RG resulting in small fluctuations indicate the conformational stability of protein-drug systems during MD simulations.

The M^{pro} residue speed (displacement/frame) ranges from 0.2 Å/frame for CYS160 to 2.51 Å/frame for TYR154 as demonstrated in Fig. 3-a. In case of H-bonding residues, the speed seems to be slightly fluctuating over the time frame of simulation (20 ps/frame) as shown in Fig. 3-b. As demonstrated by Ligplot (Fig. 4), the stable H-bonding residues interacting with N3 at the donor-acceptor distance of about 3 Å are PHE140, ASN142, HSE163, HSE164, GLU166 and GLN192 whose speeds are measured to be 0.87 ± 0.38, 0.95 ± 0.41, 0.70 ± 0.31, 0.73 ± 0.31, 0.80 ± 0.35 and 1.00 ± 0.45 Å/frame, respectively. Here, each of these average values does not exceed 1 Å/frame that signifies the stability of N3 in the viral M^{pro}-LBD.

The hydrogen bonds existing among the amino acid residues preserve the protein conformation. Due to its strict rules governed by dipole-dipole interactions, even breaking of a strong H-bond between M^{pro} helices influences the activities of viral gene transcription. The strong interactions between an inhibitor and M^{pro} residues are ex-
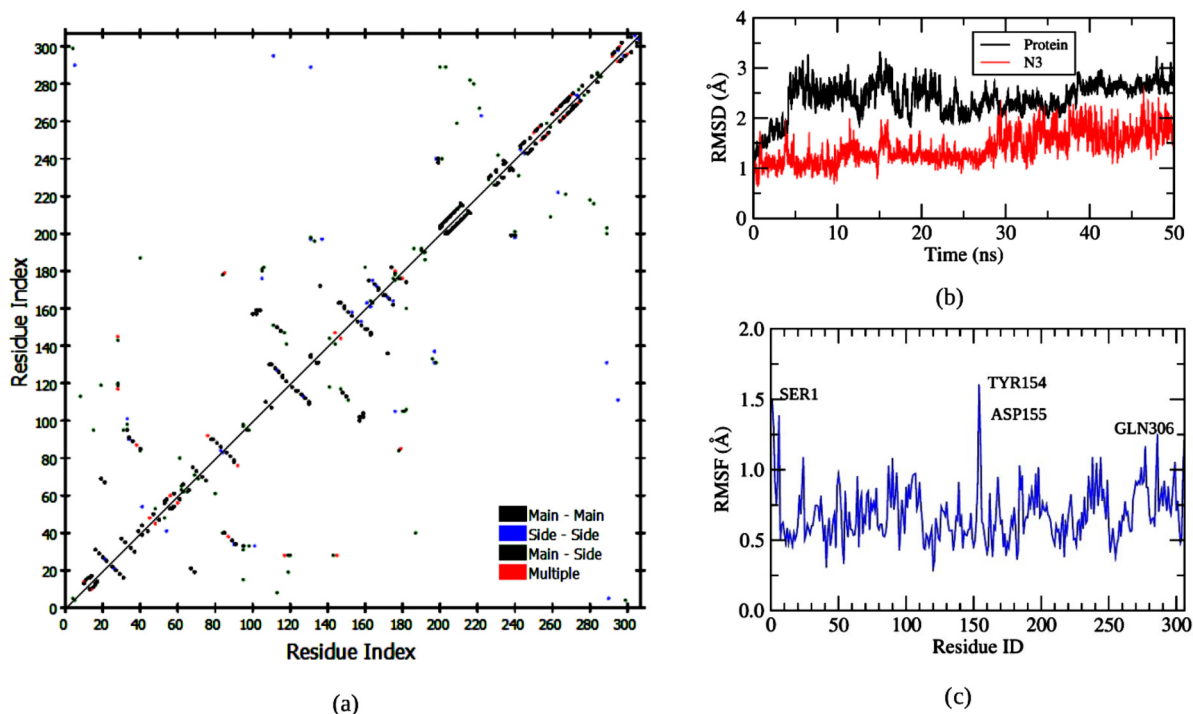
(a)

(b)

(c)

Fig. 2. (a) Contact map of the H-bonding residues in SARS-CoV-2 M^{pro}, (b) RMSD of protein backbone and inhibitor N3, (c) RMSF of M^{pro} residues.
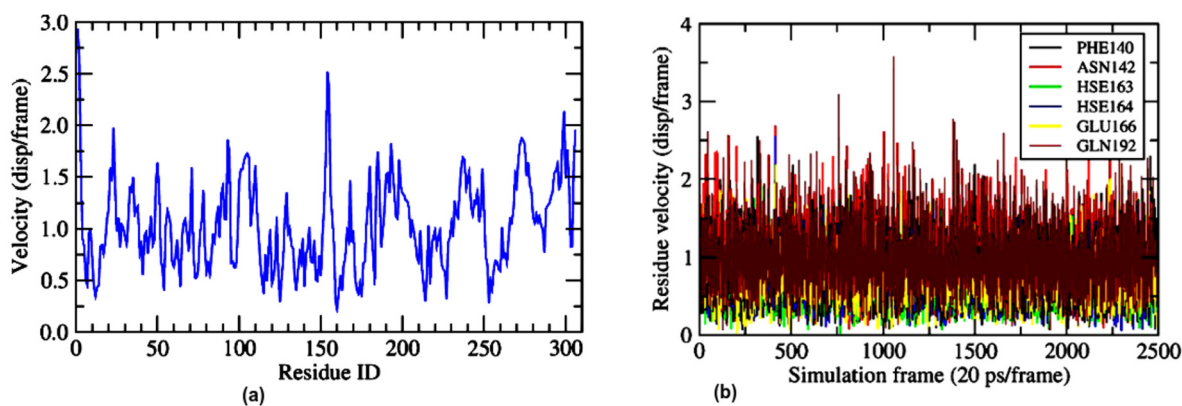

(a)

(b)

Fig. 3. (a) Residue velocities of SARS-CoV-2 main protease (M^{pro}) and (b) changing velocities of H-bonding residues in M^{pro}-LBD.

pected to restrict the replication phenomena. The H-bonds distribution between polar N and O atoms by the M^{pro}-N3 interactions is 6 ± 1 throughout the simulation time (Fig. 5-b). The number of H-bonds was estimated by assuming donor-acceptor distance 3.5 Å and cut-off angle 30°. Even the energy due to thermal fluctuation makes the H-bonds distribution change. From in-silico docking followed by 50 ns equilibration runs of M^{pro}-N3 best pose, we have found that the M^{pro}-LBD residues forming stable H-bonds with N3 are PHE140, ASN142, HSE163, HSE164, GLU166 and GLN192 (Fig. 4). According to Jin et al. (2020) [2], there are seven stable H-bonds while binding the computer designed drug N3 by the SARS-CoV-2 M^{pro}-LBD residues.

The sharp and flat RDF peaks, as shown in Fig. 5-a, are obtained by the radial distributions of H-bonding atom pairs between M^{pro} and inhibitor N3 during MD simulations. The sharper RDF peaks suggest more potent interactions. A strong H-bonding atom pair: HE2(HSE163)-O7(N3) results a sharp RDF peak in about 2.10 ± 0.55 Å radial distance and a weaker pair: O(GLN192)-H2(N3) results a flat RDF in about 4.20 ± 2.65 Å radial distance. Here, the strongly interacting residue GLU166 forms two stable H-bonds with the inhibitor N3 resulting the sharp RDF peaks. Among the H-bonding M^{pro}-LBD residues, the sharpest RDF

peak for HSE163 is also associated with the smallest speed of 0.70 ± 0.31 Å/frame during the equilibrating simulation.

Fig. 6-a shows a plot for dihedral angle distributions which is also known as Ramachandran plot. In this plot, there are 15 amino acid residues of the M^{pro}-LBD whose dihedral angles (ϕ, ψ) distribute mainly in three quadrants. The characteristic features of the Ramachandran plot are such that the right twisted antiparallel beta-barrels lie in the first quadrant ($+\phi$, $+\psi$), right-handed alpha-helices fall in the second quadrant (-ϕ, $+\psi$) and the left-handed alpha-helices fall in the third quadrant (-ϕ, -ψ). Because the residue GLY has no side chains, it performs no steric hindrance with the neighboring residues and offers a larger crude velocity. Due to this fact that it is allowed in the white region of the Ramachandran plot [14, 15]. The constraints imposed by backbone H-bonding induce the changes in dihedral angle distributions of alpha-helix and beta-forms. The torsion angles of M^{pro} residues obtained after in-silico docking and 50 ns equilibrating run are plotted in Fig. 6-b and these values for M^{pro}-LBD residues are listed in Table 1. The dihedral angles of the amino acid residues fluctuate during the MD simulation as given by the ranges of ϕ and ψ in Table 1. However, all
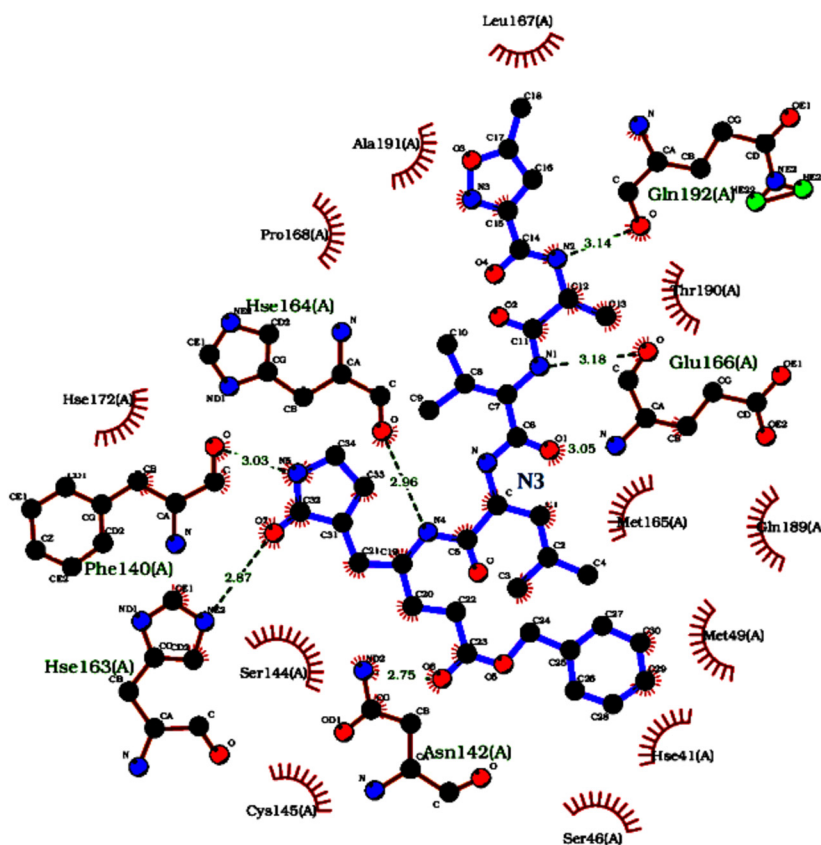
Fig. 4. Ligplot showing M^{pro}-N3 interactions after 50 ns equilibration runs where H-bonds are represented by dashed lines.
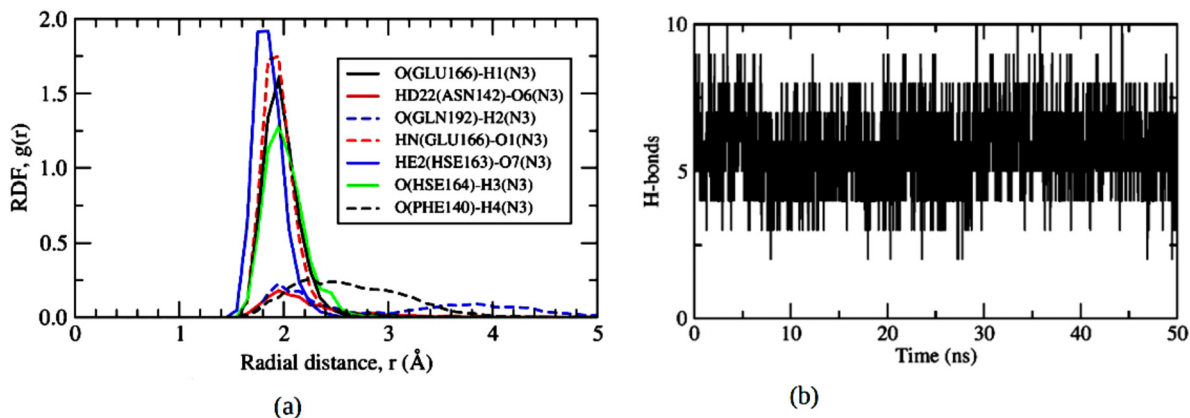

(a)   (b)

Fig. 5. a) Radial distribution functions (RDF) among the H-bonding atom pairs between SARS-CoV-2 main protease (M^{pro}) and inhibitor N3, (b) distributions of number of H-Bonds between M^{pro} and N3 during MD simulations.

H-bonding residues and most of the M^{pro}-LBD residues lie in the allowed region of the Ramachandran plot.

Coulombic interactions in the biomolecules arise from the partial charge distributions in their atoms and these are mainly related to the electrostatic dipole-dipole interactions. The M^{pro}-inhibitor binding efficacy depends on electrostatic and van der Waals contributions of the LBD residues. It is known that the positively charged amine groups are histidine, lysine and arginine, and negatively charged groups are phosphate, carboxyl and glutamate side chains. The negative term in Lenard-Jones potential comes from the induced dipole moments by fluctuating charge density and positive term originates from repulsive effects of the overlapped electron cloud of nearby atoms. NAMD-Energy Plugin calculates the coulombic and van der Waals interaction energies which are observed as -49.74 ± 6.21 kcal/mol and -58.24 ± 4.94

kcal/mol, respectively for the M^{pro}-N3 system. These values indicate the higher binding efficacy of such protein-ligand complex. Fluctuations on the interaction energies as explained by the inclusion of standard deviation in their mean values are also reflected by the time-based plots shown in Fig. 6-c. The interaction energies depend on the type of drug candidates and the nature of point mutation occurred in the viral structural proteins.

SASA measures the surface area of amino acid residues that are accessible to the solvent molecules. The SASA of the LBD residues highlights nonpolar solute-solvent interactions in the drug binding site. As plotted in Fig. 7-a, the SASA of H-bonding residues in M^{pro}-LBD measured in Å^2 are 313.27 ± 7.05 (PHE140), 258.83 ± 4.92 (ASN142), 290.93 ± 4.70 (HSE163), 292.53 ± 4.94 (HSE164), 278.72 ± 5.05 (GLU166), 284.45 ± 5.15 (GLN192). In a protein, the residue specific
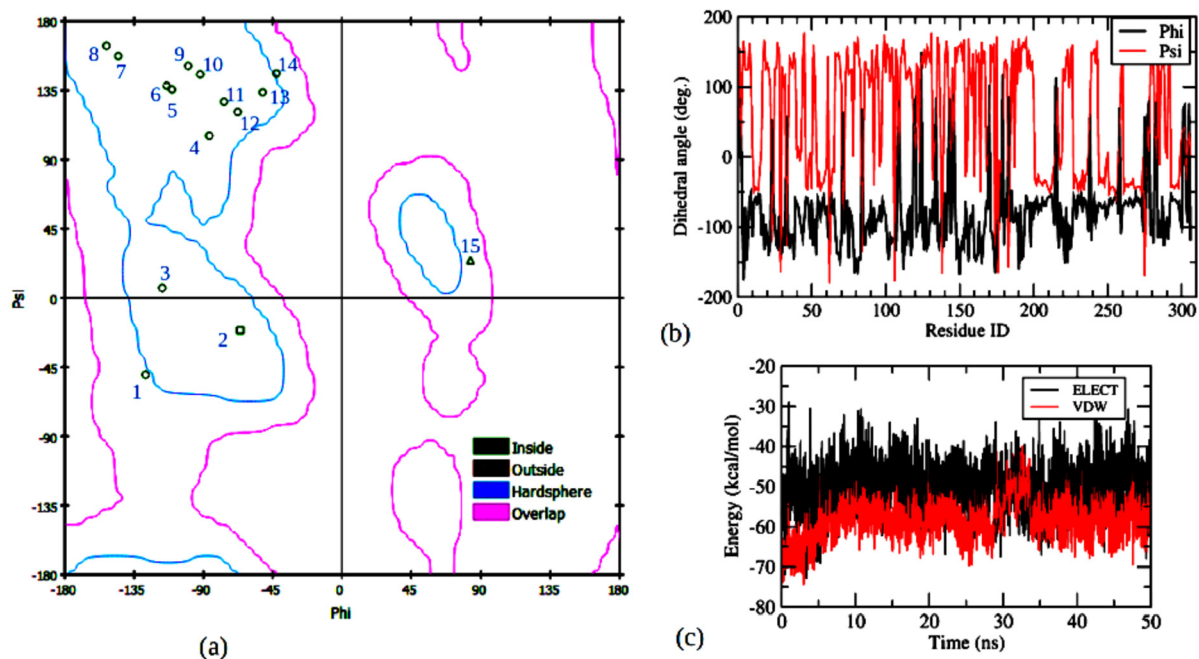
Fig. 6. (a) Ramachandran plot for M^{pro}-LBD residues, (b) dihedral distributions of M^{pro} residues, (c) electrostatic and van der Waals interaction energies between M^{pro} and N3.
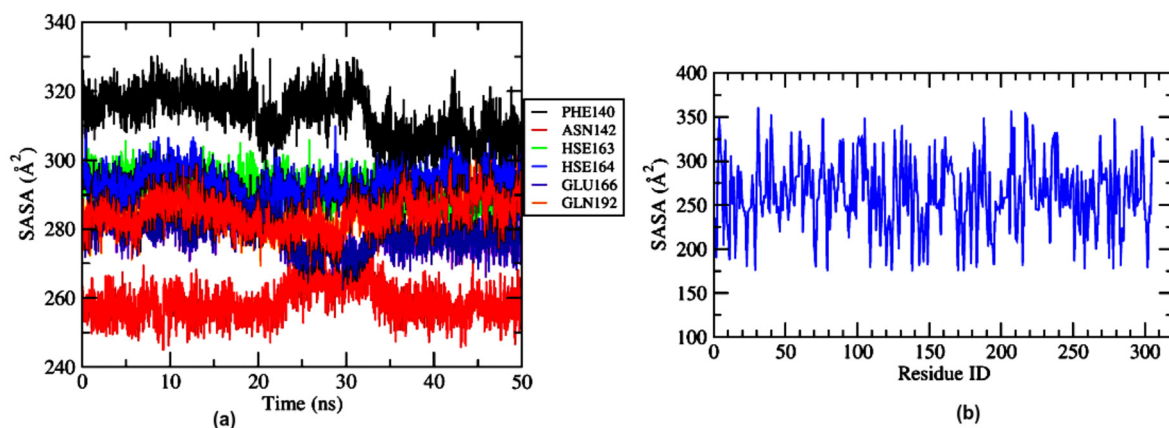

Fig. 7. (a) SASA of H-bonding residues in M^{pro}-LBD, (b) SASA of all amino acid residues in SARS-CoV-2 M^{pro}.

Table 1. Dihedral angles (ϕ, ψ measured in degree) of M^{pro}-LBD residues after in-silico docking of N3 and 50 ns equilibrations of M^{pro}-N3 system given with the range of fluctuating ϕ and ψ during its MD simulations.

| SN	MP-LBD residues	ϕ	ψ	range of ϕ	range of ψ
1	HSE164	-127.46	-49.99	-90.72 to -163.16	-13.49 to -87.28
2	PRO168	-65.97	-20.91	-40.36 to -85.71	-33.09 to 6.60
3	MET49	-116.65	6.43	-57.04 to 9.35	-51.85 to 36.36
4	GLN189	-86.08	105.43	-29.76 to -109.35	-18.20 to 150.66
5	HSE172	-110.30	135.52	-68.06 to 148.52	115.88 to 176.91
6	GLU166	-113.90	138.20	-86.03 to -151.09	109.70 to 177.13
7	GLN192	-145.22	157.31	-24.62 to -180.00	102.93 to 180.00
8	MET165	-152.86	163.93	-106.45 to -168.61	125.17 to 180.00
9	LEU167	-99.83	150.87	-54.14 to -140.90	138.44 to 180.00
10	HSE163	-92.00	145.46	-57.68 to -108.96	125.36 to 162.22
11	CYS145	-76.48	127.60	-26.93 to -81.61	116.26 to 150.94
12	ALA191	-67.64	120.91	-30.23 to -129.27	8.53 to 156.45
13	PHE140	-51.39	133.67	-18.36 to -158.56	116.89 to 172.07
14	ASN142	-42.49	146.01	-23.26 to -84.67	118.13 to 161.63
15	GLY143	83.17	24.48	54.60 to 105.96	-9.27 to 42.54
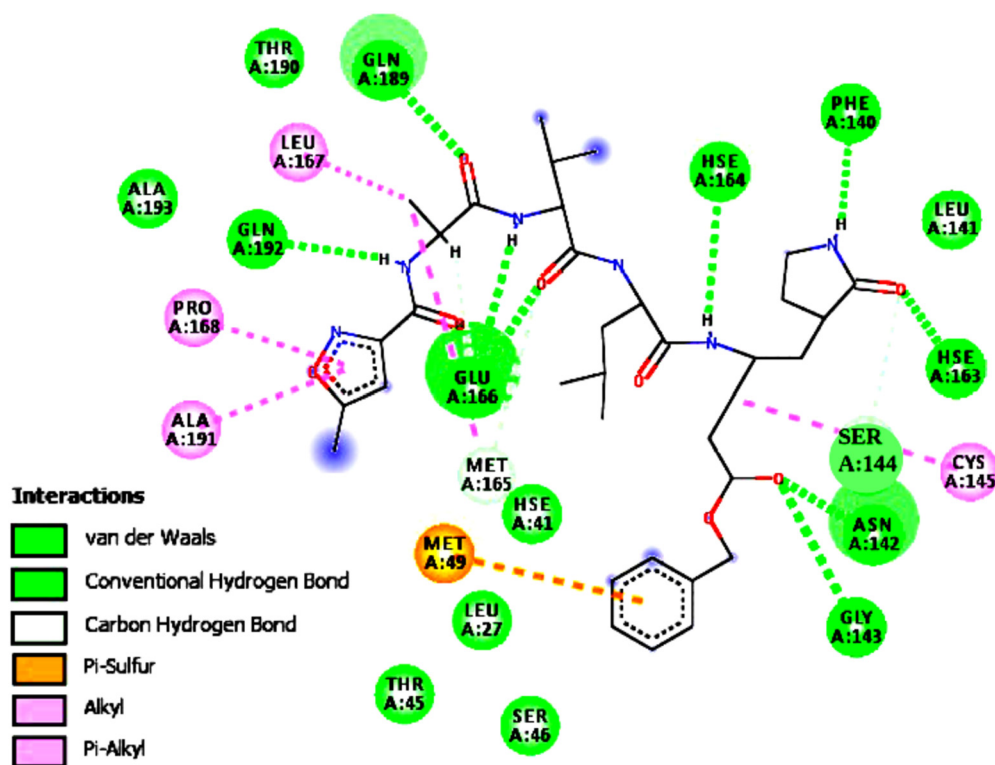
Fig. 8. Two dimensional representation of interacting residues in M^{pro}-N3 system after 50 ns equilibration.

SASA ranges from 0 to 300 Å^2 [17]. The M^{pro} residues have SASA ranging from 174.53 Å^2 for GLY138 to 360.48 Å^2 for TRP31 and such variation can be seen in Fig. 7-b. Here, the probe radius used to find SASA is 1.4 Å which is equivalent to the radius of a solvent molecule. Since the drug like N3 relies on the shallow binding cleft of M^{pro}, it has larger surface accessibility to the solvent molecules and executes strong electrostatic interactions. The inhibitor also gets hydrophobic interactions with SARS-CoV-2 M^{pro} forming H-bonds with the LBD residues [1].

Fig. 8 shows the equilibrated M^{pro}-LBD residues interacting with N3 where the residues and the bond with green color represent van der Waals interactions and conventional H-bonds, respectively. In this way, C-H bonds, pi-Sulpher, alkyl and pi-alkyl residues are distinguished with separate colors. The characteristic features of M^{pro}-LBD such as interpolated charge of receptor atoms, H-bondig with donor and acceptor atoms, hydrophobicity and basic as well as acidic ionizability can be understood with color codes as shown in Fig. 9. The colored surfaces of M^{pro}-LBD in Fig. 9-c indicate the hydrophobicity of amino acid residues representing hydrophilic (blue) and hydrophobic (brown) interactions.

The calculations based on MM/PBSA method of binding free energy assist to obtain the important biophysical information of drug binding mechanism in the protein-enzyme systems. The MM/PBSA free energy of M^{pro}-N3 binding has been obtained to be -19.45 ± 3.6 kcal/mol for which the MM contributions from coulombic and Lennard-Jones interactions are -47.06 ± 5.01 kcal/mol and -61.16 ± 3.82 kcal/mol, respectively. Here, the PB term is 95.54 ± 4.77 kcal/mol as polar solvation energy and the SA term is -6.77 ± 0.24 kcal/mol as nonpolar solvation energy. Due to the changing nature of dielectric constant (ε) with the protein residues, the MM/PBSA results may be more accurate for slightly higher ε whose default value is taken to be unity in MM electrostatic calculation [44]. Though there are some evidences of poor correlation between MM/PBSA results and experimental data for the systems like SARS-CoV-2 M^{pro} having shallow binding cleft [45, 46], our result is in a close agreement with the MM/PBSA free energy values for the non-covalent ligands reported by Muhtar et al. (2021) [47]. According to Gupta et al. (2020), the MM/PBSA free energy for di-

hydroergotamine is -17.9 kcal/mol and that for midostaurin is -16.2 kcal/mol. Due to the higher binding affinity, these were reported as potential drug candidates in SARS-CoV-2 M^{pro} [43].

The N3 like peptidomimetic inhibitors form the most extensive and stable H-bonds and hydrophobic interactions with the LBD residues resulting the decreased flexibility of SARS-CoV-2 M^{pro}. The evaluation based on the results obtained from hydrogen bonding analysis, radial distribution of donor-acceptor pairs, dihedral angle distributions of the LBD residues and free energy calculations suggests that the SARS-CoV-2 M^{pro} stands as one of the most promising targets for the potential drug candidates because no human protease has been recognized as such sort of analogous cleavage procedures.

4. Conclusion

The SARS-CoV-2 M^{pro} organized by stable beta-barrels and flexible alpha-helices strongly adapts N3 like drugs in its binding domain. The small fluctuations seen in RMSD and RG values explain the conformational stability of M^{pro}-inhibitor system during the molecular docking and dynamics simulations. The RMSD of both receptor M^{pro} and inhibitor N3 lies below 3 Å and the RG values are 22.70 ± 0.20 Å and 6.61 ± 0.25 Å, respectively. The RG with small standard deviation measures the compactness of the protein-ligand system. Also, the RMSF values of M^{pro} residues lie below 1.5 Å. Residue speed or displacement/frame is the least (0.2 Å/frame) for CYS160 and the greatest (2.51 Å/frame) for TYR154. The sharpness of RDF curve about small radial distance of H-bonding atom pairs reflects the nature of strong interactions between inhibitor and SARS-CoV-2 M^{pro}. A way of evaluating the binding mechanism of an inhibitor in the M^{pro}-LBD is to study the nature of RDF curves. In the M^{pro}-N3 binding, HSE163 has a sharp RDF for the strong H-bonding atom pairs within 2.10 ± 0.55 Å radial distance, whereas GLN192 has a flat RDF for the weak H-bonding atom pairs within 4.20 ± 2.65 Å radial distance. Being adapted in the M^{pro}-LBD, N3 acquires hydrophobic interactions along with coulombic energy -49.74 ± 6.21 kcal/mol and van der Waals energy -58.24 ± 4.94 kcal/mol. The MM/PBSA free energy of M^{pro}-N3 binding has been calculated to be -
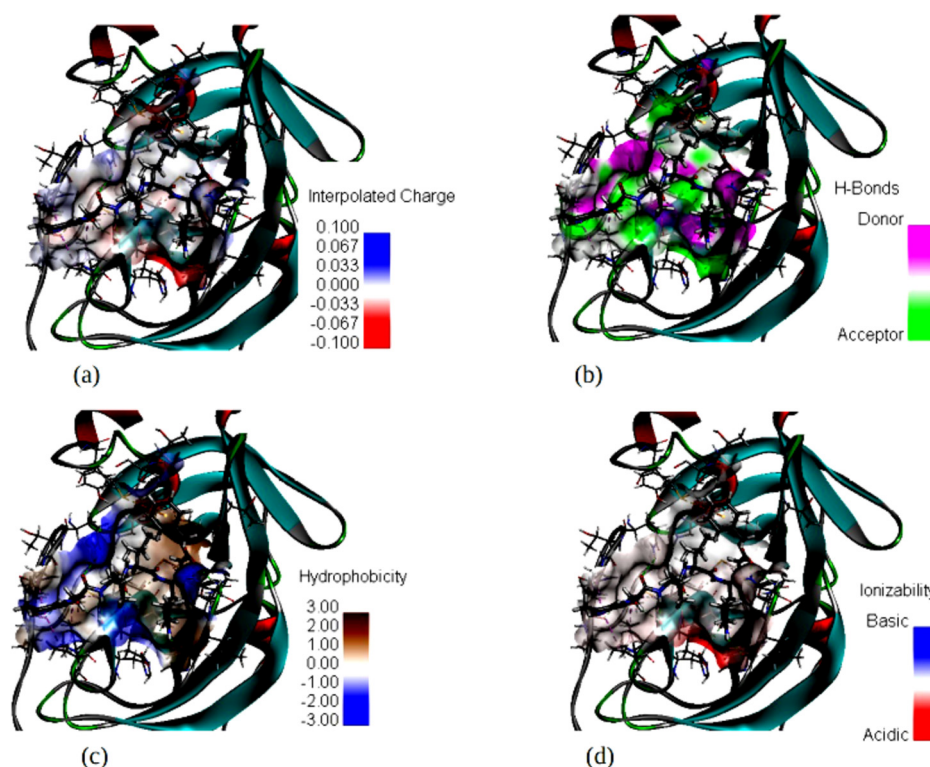
Fig. 9. Characteristics of M^pro receptor binding inhibitor N3 where the color coded surfaces show (a) interpolated charge of receptor atoms, (b) H-bondig with donor and acceptor atoms, (c) hydrophobicity, and (d) basic and acidic ionizability.

19.45 ± 3.6 kcal/mol as the contribution of non-covalent interactions. In conclusion, the physical parameters explained in this research will be useful tools for evaluating the binding mechanism of potential drug candidates that inhibit SARS-CoV-2 structural proteins.

Declarations

Author contribution statement

Tika Ram Lamichhane: Conceived and designed the experiments; Analyzed and interpreted the data; Wrote the paper.

Madhav Prasad Ghimire: Contributed reagents, materials, analysis tools or data; Wrote the paper.

Data availability statement

Data will be made available on request.

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

References

[1] F. Wang, C. Chen, W. Tan, K. Yang, H. Yang, Sci. Rep. 7 (2016) 22677.
[2] Z. Jin, X. Du, Y. Xu, Y. Deng, M. Liu, Y. Zhao, B. Zhang, X. Li, L. Zhang, C. Peng, Y. Duan, Nature 9 (2020) 1.
[3] A. Zhavoronkov, B. Zagribelnyy, A. Zhebrak, V. Aladinskiy, T. Verentiev, Q. Vanhaelen, D.S. Bezrukov, D. Polykovskiy, R. Shayakhmetov, A. Filimonov, M. Bishop, Insilico Med. Hong Kong Ltd. A 307 (2020) E1.
[4] X. Liu, B. Zhang, Z. Jin, H. Yang, Z. Rao, Protein DataBank, http://www.rcsb.org/structure/6LU7, 2020.
[5] J. Lan, J. Ge, J. Yu, S. Shan, H. Zhou, S. Fan, Q. Zhang, X. Shi, Q. Wang, L. Zhang, X. Wang, Nature (2020).
[6] Y. Kumar, H. Singh, C.N. Patel, J. Inf. Public Heal. 13 (2020) 1210.
[7] S.T. Ngo, N. Quynh Anh Pham, L. Le Thi, D.H. Pham, V.V. Vu, J. Chem. Inf. Model. 60 (2020) 577.
[8] A.T. Ton, F. Gentile, M. Hsing, F. Ban, A. Cherkasov, Mol. Inf. 39 (2020) 2000028.
[9] D. Acharya, G. Liu, M.U. Gack, Nat. Rev. Immunol. (2020).
[10] R. Walser, P.H. Hunenberger, W.F. van Gunsteren, Proteins, Struct. Funct. Bioinform. 48 (2002) 327.
[11] S. Riniker, A.P. Eichenberger, W.F. van Gunsteren, Eur. Biophys. J. 41 (2002) 647.
[12] C.Y. Hu, G.C. Lynch, H. Kokubo, B.M. Pettitt, Proteins, Struct. Funct. Bioinform. 78 (2010) 695.
[13] B.G. Levine, J.E. Stone, A. Kohlmeyer, J. Comput. Phys. 230 (2011) 3556.
[14] H. Gohlke, M. Hendlich, G. Klebe, J. Mol. Biol. 295 (2000) 337.
[15] S.A. Hollingsworth, P.A. Karplus, Biomol. Concepts 1 (2010) 271.
[16] B.K. Ho, R. Brasseur, BMC Struct. Biol. (2005), http://www.biomedcentral.com/1472-6807/5/14.
[17] S. Ausaf Ali, I. Hassan, A. Islam, F. Ahmad, Curr. Protein Pept. Sci. 15 (2014) 456.
[18] B. Lee, F.M. Richards, J. Mol. Biol. 55 (1971) 379.
[19] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kal, K. Schulten, J. Comput. Chem. 26 (2005) 1781.
[20] H. Liu, T. Hou, Bioinformatics 32 (2016) 2216.
[21] P.A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, Acc. Chem. Res. 33 (2000) 889.
[22] T.R. Lamichhane, M.P. Ghimire, Tribhuvan Univ. J. 34 (2020) 1.
[23] L.L.C. Schrodinger, The PyMOL Molecular Graphics System, Version 1.8, 2015.

[24] N. Guex, M.C. Peitsch, Electrophoresis 18 (1997) 2714.
[25] ACD/ChemSketch, ACD/ChemSketch for Academic and Personal Use, 2018, ACD/ Labs.com.
[26] E.W. Bell, Y. Zhang, J. Cheminform. (2019) 11.
[27] G.M. Morris, R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell, A.J. Olson, J. Comput. Chem. 16 (2009) 2785.
[28] G.M. Morris, D.S. Goodsell, R.S. Halliday, R. Huey, W.E. Hart, R.K. Belew, A.J. Olson, J. Comput. Chem. 19 (1998) 1639.
[29] J. Gasteiger, M. Marsili, Tetrahedron 36 (1980) 3219.
[30] A.D. MacKerell Jr, D. Bashford, M.L. Bellott, R.L. Dunbrack Jr, J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, J. Phys. Chem. B102 (1998) 3586.
[31] J. Huang, A.D. MacKerell Jr, J. Comput. Chem. 34 (2013) 2135.
[32] J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B.L. De Groot, H. Grubmüller, A.D. MacKerell, Nat. Methods 14 (2017) 71.
[33] V. Zoete, M.A. Cuendet, A. Grosdidie, O. Michielin, J. Comput. Chem. 32 (2011) 2359.
[34] W. Humphrey, A. Dalke, K. Schulten, J. Mol. Graph. 14 (1996) 33.
[35] L. Verlet, Phys. Rev. 159 (1967) 98.
[36] T.R. Lamichhane, H.P. Lamichhane, AIMS Biophys. 7 (2020) 27.
[37] T.R. Lamichhane, S. Paudel, B.K. Yadav, H.P. Lamichhane, J. Biol. Phys. 45 (2019) 107.
[38] T.R. Lamichhane, H.P. Lamichhane, AIMS Biophys. 5 (2018) 245.
[39] A.K. Bronowska, Thermodynamics-Interaction Studies-Solids, Liquids and Gases, IntechOpen, 2011.
[40] BIOVIA, Dassault Systemes, Discovery Studio Visualizer, Version 20.1.0.19295, Dassault Systemes, San Diego, 2020.
[41] R.A. Laskowski, M.B. Swindells, J. Chem. Inf. Model. 51 (2011) 2778.
[42] N.A. Baker, D. Sept, S. Joseph, M.J. Holst, J.A. McCammon, Proc. Natl. Acad. Sci. USA 98 (2001) 10037.
[43] A. Gupta, Z.X. Zhou, ACS Comb. Sci. 22 (2020) 826.
[44] S. Genheden, U. Ryde, Expert Opin. Drug Discov. 10 (2015) 449.
[45] S.T. Ngo, N.M. Tam, M.Q. Pham, T.H. Nguyen, J. Chem. Inf. Model. 61 (2021) 2302.
[46] Z. Li, X. Li, Y.Y. Huang, Y.Y. Wu, R. Liu, L. Zhou, Y. Lin, D. Wu, L. Zhang, H. Liu, X. Xu, K. Yu, Y. Zhang, J. Cui, C.G. Zhan, X. Wang, H.B. Luo, Proc. Natl. Acad. Sci. USA 117 (2020) 27381.
[47] E. Muhtar, M. Wang, H. Zhu, Future Virol. 16 (2021) 507.