Structural Analysis of COVID-19 Main Protease and its Interaction with the Inhibitor N3

Tika Ram Lamichhane[1*] and Madhav Prasad Ghimire[1,2*]

[1]*Central Department of Physics, Tribhuvan University, Kirtipur-44613, Kathmandu, Nepal*
[2]*Condensed Matter Physics Research Center, Butwal-11, Rupandehi, Nepal*
**Corresponding authors: tika.lamichhane@cdp.tu.edu.np; madhav.ghimire@cdp.tu.edu.np*

Abstract: *Here, we analyze the structural features of a ligand binding domain (LBD) in COVID-19 main protease (MP) followed by the interactions between the inhibitor N3 and MP-LBD residues through the molecular dynamics simulations. The time based changes in physical parameters that includes root mean square deviation (RMSD), root mean square fluctuation (RMSF), radius of gyration (RG), dihedral distributions, residue velocity, radial distribution function (RDF) and H-bonding signify the degrees of folding states in MP-N3 complex formed by the superimposed β-barrels and flexible α-helices. Sharp and flat RDF peaks observed for the atom pairs dictate the flexibility of MP-LBD residues during their interactions with N3. In spite of larger solvent accessibility of N3, it interacts strongly with the LBD residues resulting in H-bonding. Among the LBD residues, GLU166 is found to have the lowest residue velocity that offers the sharp RDF peaks for three H-bonding atom pairs nearly at 2 Å radial distance, whereas GLY143 has the highest value of residue velocity giving rise to a flat RDF peak for the MP-N3 atom pair. Furthermore, electrostatic and van der Waals interaction energies between N3 and MP-LBD residues are noted to have the negative values. All these parameters explain the binding nature of N3 like inhibitors to the substrate binding sites of COVID-19 main protease. These analysis are expected to be a possible route applicable in drug designing mechanism to restrict the viral replication and transcription of COVID-19.*

Key Words: *Novel corona virus, inhibitor N3, H-bonding, radial distributions, protein-ligand interactions*

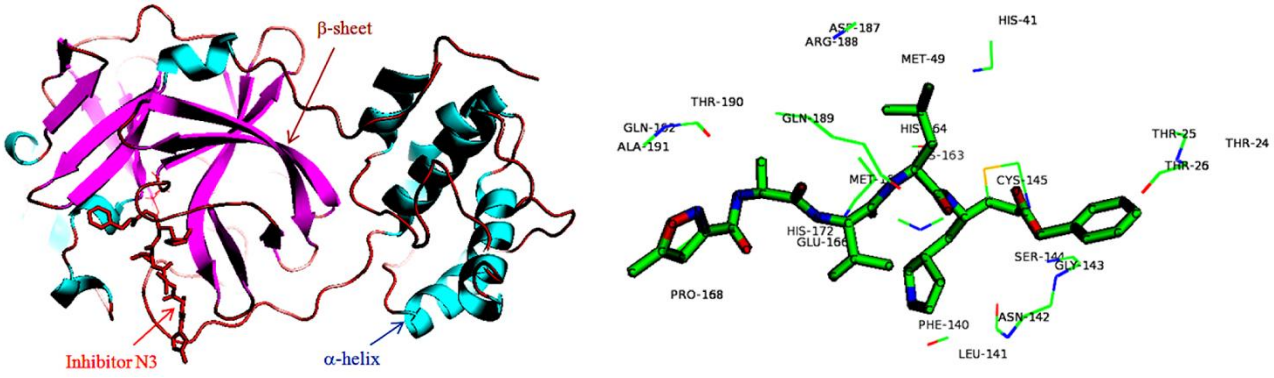
1. Introduction

Very recently, a new type of virus is found that is creating a pandemic situation globally. The said virus was named as SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2) by the International Committee on Taxonomy of Viruses (ICTV) while the World Health Organization (WHO) named it as Corona Virus (COVID-19) based on the guidelines developed by the World Organization for Animal Health (OIE) and the Food and Agriculture Organization of the United Nations (FAO). Hereafter, throughout, we are calling this virus as COVID-19 basing on the name coined by WHO.

To this time of writing, there were no reports known on the discovery of any vaccines or antiviral drugs prescribed for the COVID-19 infection from WHO. However, many pre-clinical and clinical tests are undergoing to develop the vaccines against this worldwide spreading disease. The human coronavirus emerging from zoonotic origin has the substrate-binding sites into its proteases as the conserved drug targets for designing the antiviral inhibitors [1-5]. One of the potential antiviral drugs noted is the inhibitor N3 which is capable of interacting with the coronavirus main protease [1, 2]. This type of drug interferes with the post-translational activities of viral polypeptides and prevents the replication and proliferation of the viruses [3]. Likewise, indole derivatives are the potential drug candidates which also have wide range of applications that includes the probable inhibitory effects on the invading pathogens, and the anticancer efficacy by demonstrating a dose-dependent cytotoxicity against the tested carcinoma cells [6, 7].

The novel COVID-19 is found to cause a mild to acute respiratory failure or multi-organ injuries as the hallmarks of cytokine release syndrome in humans influenced by the host genetics, age, and comorbidities as reported recently [8]. The virus genome consists of nearly thirty thousand nucleotides whose replicase gene encodes two overlapping polyproteins. The extensive proteolytic processing of the polyproteins releases functional polypeptides which are predominantly conserved by a 33.8 kDa main protease or 3C-like protease [2]. The COVID-19 main protease (MP) is a key enzyme which plays significant role in the corona-viral replication and transcription. It has a CYS-HIS catalytic dyad and three domains with a substrate binding site in between the first (residue: 8-101) and second (residue: 102-184) domain of β-barrel structure [2, 4]. The MP with strongly arranged antiparallel β-forms and more flexible α-helixes (Figure 1-a) comprise of the active zone of the virus so that it seems to be an attractive target for the antiviral drug design. As an active ligand of MP, inhibitor N3 is a potential drug to combat COVID-19. The peptide-like inhibitor N3 is capable of making a strong covalent bonding between C_β of vinyl group and S_γ of CYS145-A [2]. N3 makes H-bonds with the surrounding

residues of ligand binding domain (LBD) in the pharmacophore. The peptidomimetic N3, with its Michael acceptor and well-designed side chains (Figure 1-b), fits well into the conserved MP-LBD. The interacting amino acid residues such as THR24-26, HIS41, MET49, PHE140, LEU141, GLY143, SER144, CYS145, HIS163-164, MET165, GLU166, PRO168, HIS172, ASP187, ARG188, GLN189, THR190, ALA191 and GLN192 shown in Figure 1-b are found to surround the inhibitor N3 within a distance of 4 Å.



(a) (b)

Fig. 1 (a) Structure of COVID-19 main protease (MP) liganded with inhibitor N3, and (b) structure of N3 and MP amino acid residues within 4 Å range of N3 visualized by the PyMOL molecular graphics [9].

The structural analysis of the COVID-19 main protease has been performed by means of root mean square deviation (RMSD) [10, 11], root mean square fluctuation (RMSF) [11], radius of gyration (RG) [11, 12], radial distribution function (RDF) [13], dihedral angles distribution [14, 15], solvent accessible surface area (SASA) [16, 17], infrared spectral density (IRSD) [18], and its interaction with the inhibitor N3 in terms of varying H-bonds, residue velocity, electrostatic as well as van der Waals interactions and internal energy fluctuations using nanoscale molecular dynamics (NAMD) simulations [19]. This analysis will be useful to compare the biophysical properties of COVID-19 with the other existing viruses and to observe the properties of drugs on how they interact with the main protease.

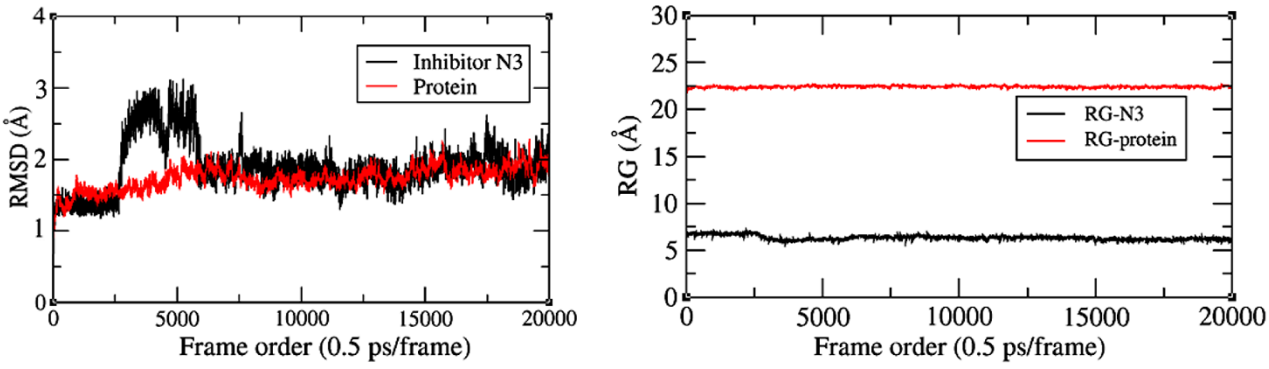
2. Methodology

The primary structure of COVID-19 main protease with an inhibitor N3 was taken from the crystallographic data code 6LU7.PDB [4]. In order to perform NAMD simulations of a fully solvated MP-N3 system, the related topology and parameters were obtained from the CHARMM force fields [20-22]. The topology file contains the information about name and type of atom, mass and charge of each atom in every residue so that necessary protein structure file (PSF) can be generated. Construction of PSF automatically assigns coordinates to the missing H-atoms in the protein data bank (PDB) file. The parameter file contains numerical constants that are required to calculate all the bonded and non-bonded energy terms in the CHARMM energy function [19]. Structure of N3 was extracted from 6LU7.PDB, hydrogenated and parameterized in order to describe its molecular geometry, partial atomic charges, atomic masses, and numerical constants needed to evaluate forces and energies [22]. The electrically neutral N3, treated as a ligand moiety, consists of 101 atoms with a total mass of 684.83 amu including 52 H-atoms. By using the topology files, the PSF of MP-N3 complex was generated with *psfgen* package of visual molecular dynamics (VMD) [23]. It preserves the original coordinates and the existing bonds in MP and N3 molecules. Thus prepared structure was solvated with a total of 23112 water molecules (TIP3P) under the spherical boundary conditions forming a nanodroplet neutralized with the Na^+ and Cl^- ions in 0.15 mol/L concentrations. It closely resembles the cellular environment. The MD simulations were performed with the NAMD-2.12 specifically for energy minimization, equilibration and production runs, and for the structural analysis we use VMD (version 1.9.3).

In the solvated MP-N3 droplet, it has a radius of 40.86 Å with the center coordinates (-26.00, 12.55, 59.18 Å), an exponent 2 for the first potential to be applied and the force constant of 10 kcalmol^{-1}Å^{-2}. The Langevin dynamics was active with the piston temperature of 310 K and damping coefficient of 1 ps^{-1}. The simulation parameter used was 2 fs/step implementing the constrained vibrations with the rigid bonds of H-atoms. The multilevel summation method (MSM) was implemented with a grid spacing of 2.5 Å for the evaluation of electrostatic force. In order to evaluate the electrostatic and van der Waals interactions, the force field related term used was 1.0 (general scaling range is 1-4), and switching, cut-off and pair-list distances were set to 10, 12 and 14 Å, respectively. The initial energy of the MP-N3 system was minimized through 3000 CG steps and it was equilibrated up to 10 ns at the body scale temperature of 310 K. During MD simulations, the atomic trajectories were calculated by using the Verlet algorithm [24]. The coulombic and van der Waals interaction energies between inhibitor N3 and MP were calculated by using NAMD energy plugin -1.4 and the IRSD was evaluated by using VMD plugin-1.3. This plugin uses a graphical interface to compute the spectral densities from the time series data of atomic trajectories created with the measure dipole command in VMD. Custom spectral densities were computed from the auto-correlated Fourier transform. By using VMD Timeline-2.3, crude velocity analog, i.e. per residue inter-frame displacement of MP residues and inhibitor N3 were calculated during MD simulations. As explained in the structural analysis of the protein-hormone system [25]; RMSD, RMSF, RDF, Ramachandran plots, SASA, distance, velocity and number of H-bond distributions were used as the tools to analyze the conformational properties of MP-N3 system.

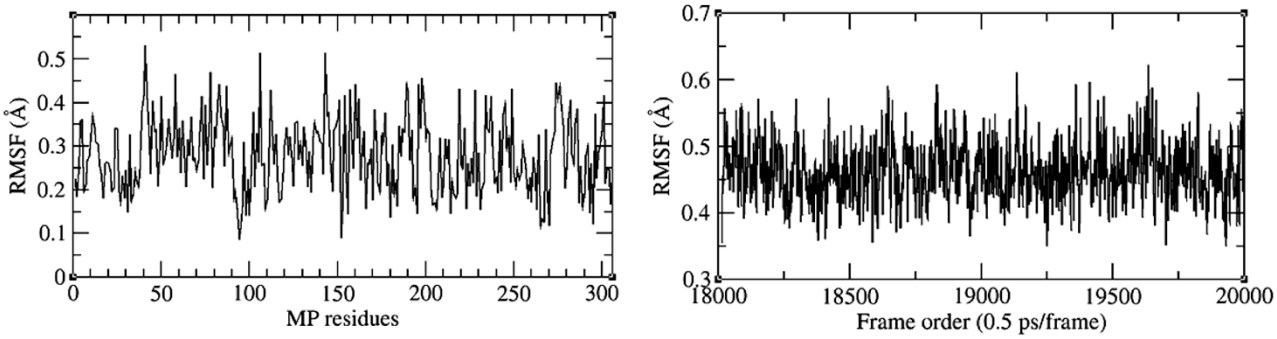
3. Results and Discussion

As a structural basis of MP-N3 solvated system, the electrically neutral nanodroplet has a total mass of 174682 amu including 27949 atoms, 20250 bonds, 16415 angles, 12746 dihedrals, 891 impropers, 304 cross terms, 10171 H-groups and 58363 degrees of freedom. The changes in structure and interatomic interactions are found to depend on such conformational parameters of the system [18]. Since the physical terms such as RMSD, RMSF and RG are the average measures of the backbone atoms or residues of superimposed α- and β-strands, they act as the reaction coordinates to signify the degree of folding states. During the 10 ns long equilibrations, the mean ± SD in RMSDs of MP and N3 are found to be 1.73 ± 0.16 Å and 1.89 ± 0.36 Å, respectively as listed in Table 1. The average RMSD is quite small, however its fluctuating value (see Figure 2-a) gives the signature of breaking and reforming of H-bonds among the MP-N3 residues. The mass weighted RMS distance of the particles from the center of mass, RG measures the compactness of the system. Our observation shows that average RG of MP is 22.41 ± 0.10 Å and that of N3 is 6.34 ± 0.23 Å. Almost constant values of RG are found on MP and N3 as shown in Figure 2-b.



(a) (b)

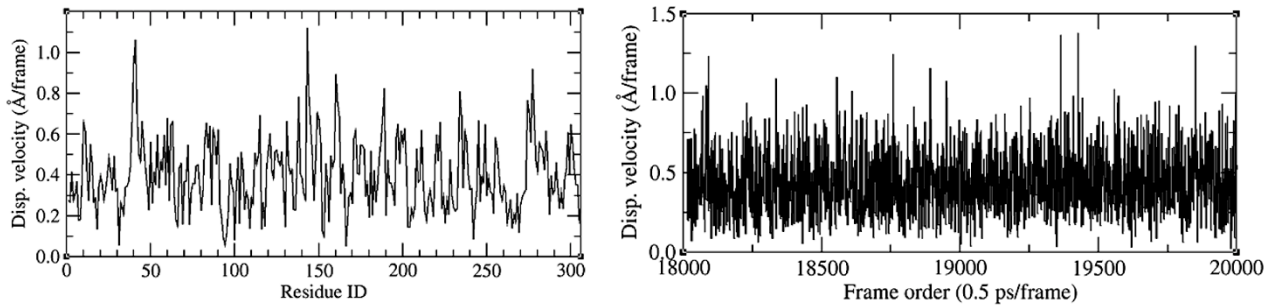
Fig. 2 (a) RMSDs and (b) RGs of main protease and inhibitor N3 during equilibrating simulations of MP-N3 nanodroplet.



(a) (b)

Fig. 3 RMSF of (a) MP residues in the last frame, and (b) inhibitor N3 in MP-LBD during the last 1ns in MD simulations.

From the RMSF study, we notice that each of the MP residues fluctuates minutely from its mean position during the atomic vibrations for all the residues lying within 0.5 Å (Figure 3-a). The RMSF of N3 in its binding domain is found to be 0.46 ± 0.04 Å as a mean value of its fluctuations in the last 1 ns of equilibrating runs (Figure 3-b). The observation describes the prominent flexibility of N3 in MP-LBD. It is to be noted that the physical factors like RMSD, RMSF and RG are useful techniques to predict the conformational stability of the drugs like inhibitor N3 in the ligand binding domain. Throughout the simulations, the velocities of the MP residues are distributed as shown in Figure 4-a where the smallest value is 0.05 Å/frame for GLU166 and the largest value is 1.12 Å/frame for GLY143. Likewise, almost constant values for the displacement per unit time-frame or the velocity of N3 were obtained as shown in Figure 4-b with a values of 0.44 ± 0.20 Å/frame. Here, single frame refers to 0.5 ps of the simulation time. It should be noted that GLY is a single amino acid residue which is allowed in the white region of the Ramachandran plot (see Figure 5). GLY existing without side chain has no steric hindrance with the surrounding residues because of which it has a larger displacement per frame.



(a) (b)

Fig. 4 (a) Displacement per frame or velocity of MP residues in the last frame and (b) the varying velocity of N3 during frame to frame equilibrations of MPN3 system.

Figure 5 depicts the Ramachandran plots for torsion angle distributions of MP residues, in which right twisted antiparallel β-barrels lies in the blue region I, right-handed α-helices lies in the blue region II and the left-handed α-helices lie in the green region III [14, 15], while the GLY residues with no side chains have no steric clashes so that they are allowed in the white region IV. Here, the variation in ϕ, ψ along the α-helix is induced by the backbone H-bonding constraints. During equilibrating run, the torsion angles (ϕ, ψ) of LBD residues CYS145 and GLY143 are changing as shown by the dark-line spots in Figures 5-a,b where (-50.84, 145.14°) and (66.45°, 24.02°) refer to the red spot in the last frame, respectively.

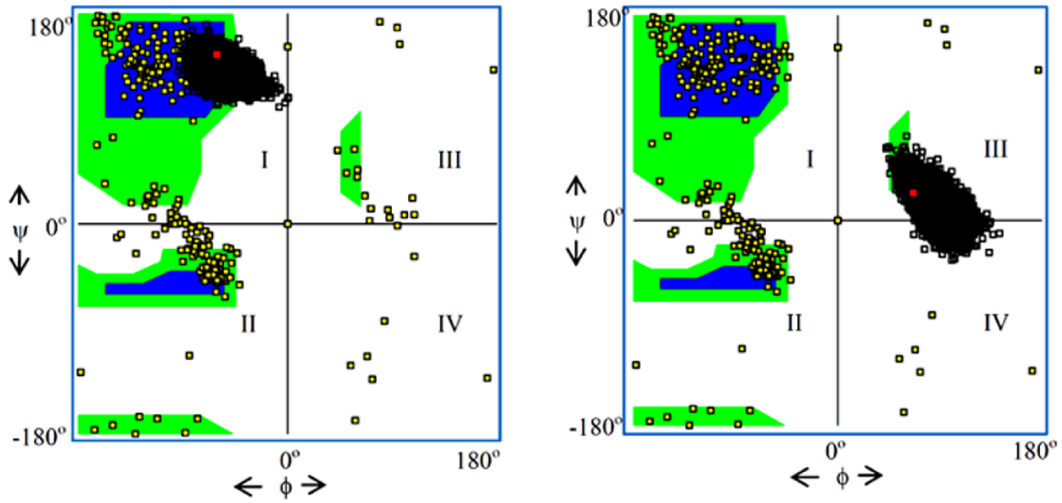


Fig. 5 Ramachandran plots for (ϕ, ψ) distributions of MP-N3 residues. The dark line spots show the variations of dihedral angles during the simulation with the red spot in the last frame for (a) CYS145 (ϕ = -50.84°, ψ = 145.14°) and (b) GLY143 (ϕ = 66.45°, ψ = 24.02°). The point in the y-axis refers to SER1 (ϕ = 0°, ψ = 151.90°).

Table 1: Structural parameters (mean ± SD) of COVID-19 main protease (MP) and inhibitor N3 where (*) refers to the range for the MP residues and (†) refers to the number of H-bonds between N3 and MP residues.

Molecule	*RMSD(Å)*	*RG(Å)*	*RMSF(Å)*	*velocity(Å/frame)*	*SASA(Å²)*	*H-bonds*
COVID-19 MP	1.73 ± 0.16	22.41 ± 0.10	0.10 - 0.52*	0.05 - 1.15*	173.32 - 358.93*	199 ± 9
N3	1.89 ± 0.36	6.34 ± 0.23	0.46 ± 0.04	0.44 ± 0.20	997.03 ± 15.05	6 ± 2†

We now consider the SASA which measures the surface accessibility of a biomolecule to its solvents. As depicted in Table 1, the SASA of MP residues lies within the range of 173.32 to 358.93 Å^2 related to GLY2 and ARG105, respectively (see Figure 6-a). As reported by Ausaf Ali *et al.* (2014), the SASA of amino acid residues varies from 0 to 300 Å^2 [16]. The probe radius used in calculating the SASA is 1.4 Å, which approximately equals the radius of a water molecule. The inhibitor N3 being located in the LBD of MP β-barrels has a much larger SASA of 997.03 ± 15.05 Å^2 indicating the larger surface area accessible to the water molecules. However, it offers hydrophobic interactions with the LBD residues [1], and interacts strongly with the water molecules which is coulombic in nature.

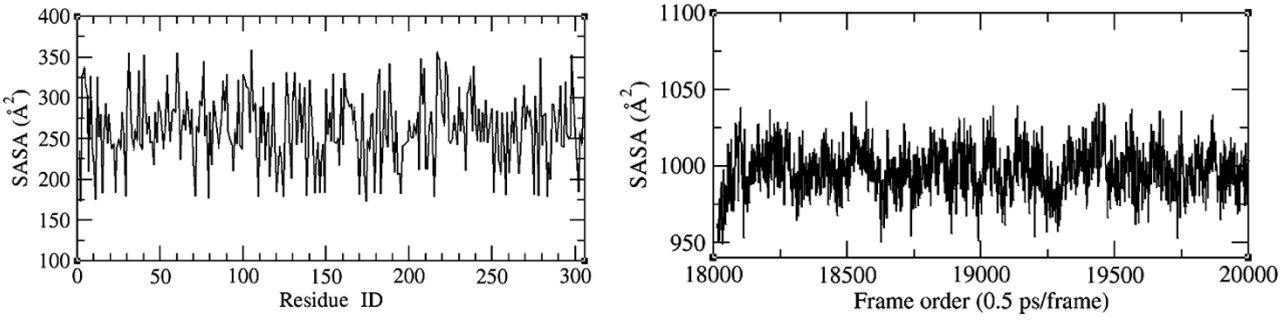
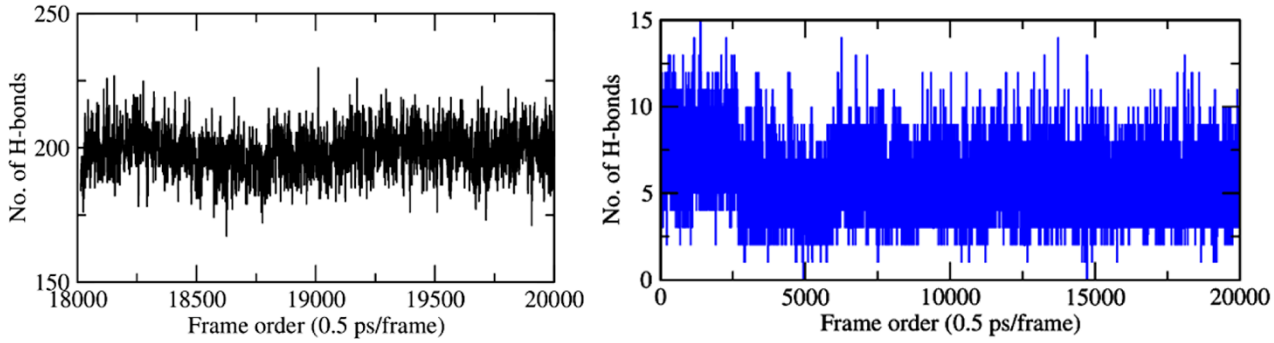


Fig. 6 SASA of (a) MP residues in the last frame, and (b) inhibitor N3 in the last 1ns of the equilibration runs.



(a) (b)

Fig. 7 Number of H-bonds (a) among MP residues, and (b) between MP-LBD residues and N3 w.r.t. time-frame of simulations.

As known, H-bonding, often described as a strong electrostatic dipole-dipole interaction, maintains the globular conformations of proteins. It generally provides strict geometrical rules in the molecule. Usually, breaking of H-bonds among the MP residues makes the virus unstable towards gene transcription. The purpose of virus-drug interactions is mainly to restrict the viral activity towards cell invasion, replication and transcription. As shown in Table 1 and Figure 7-a, H-bonds among the MP residues fluctuates within the range of 199 ± 9 during the simulations. The number of H-bonds between N3 and MP-LBD residues are distributed by 6 ± 2 (Figure 7-b) including the polar atoms N and O, respectively. The calculations were performed by adjusting the donor-acceptor distance of 3.5 Å and a cut-off angle of 30°. The H-bonds are switched on or off with energies that lie in the range of thermal fluctuations. The MP-LBD residues permissible to H-bonding with inhibitor N3 are GLY143, HSE163, HSE164, GLU166 and THR190. As explained recently by Jin *et al.* (2020) [2], the atom pairs forming H-bonds are: O2(N3)-HN(GLU166), H4(N3)-O(THR190), O4(N3)-HN(GLY143), O8(N3)-HE2(HSE163), H9(N3)-O(GLU166), H41(N3)-OE1(GLU166) and H44(N3)-O(HSE164). The mean distances between the atoms in such pairs are measured to be 1.93 ± 0.16, 3.84 ± 0.92, 5.48 ± 0.83, 2.10 ± 0.26, 2.18 ± 0.50, 2.73 ± 0.96 and 2.04 ± 0.24 Å, respectively. The changing radial distances throughout the simulation are described by the RDF plots (see Figure 8) where the minutely varying distances execute the sharp peaks and largely varying distances yield the flat peaks. The smaller magnitude of RDF extends to a larger radial distances. For example, the highly interacting atom pair: H31(N3)-HG1(CYS145) has a flat RDF peak about the radial distance of 3.47 ± 0.77 Å whereas H-bonding atom pair: O2(N3)-HN(GLU166) has the very sharp RDF peak about the radial distance of 1.93 ± 0.16 Å. The initial and final structures obtained by the atomic interactions among N3 and MP-LBD residues are shown in Figure 9-a,b. Here, the most stable residue GLU166 develops 3 H-bonds with N3. The position of GLY143 with no side chains seems to be highly fluctuating as explained above by the largest displacement per frame (Figure 5-a) and the flattest RDF peak for its H-bonding atom pair (Figure 8).

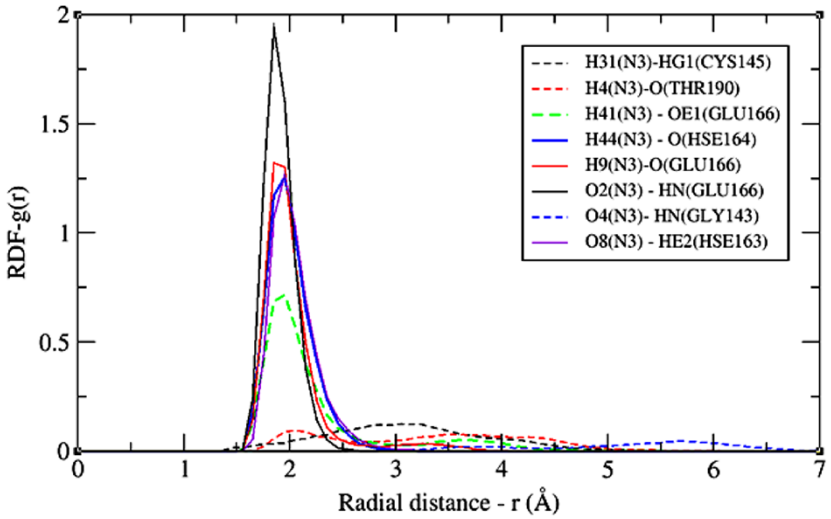
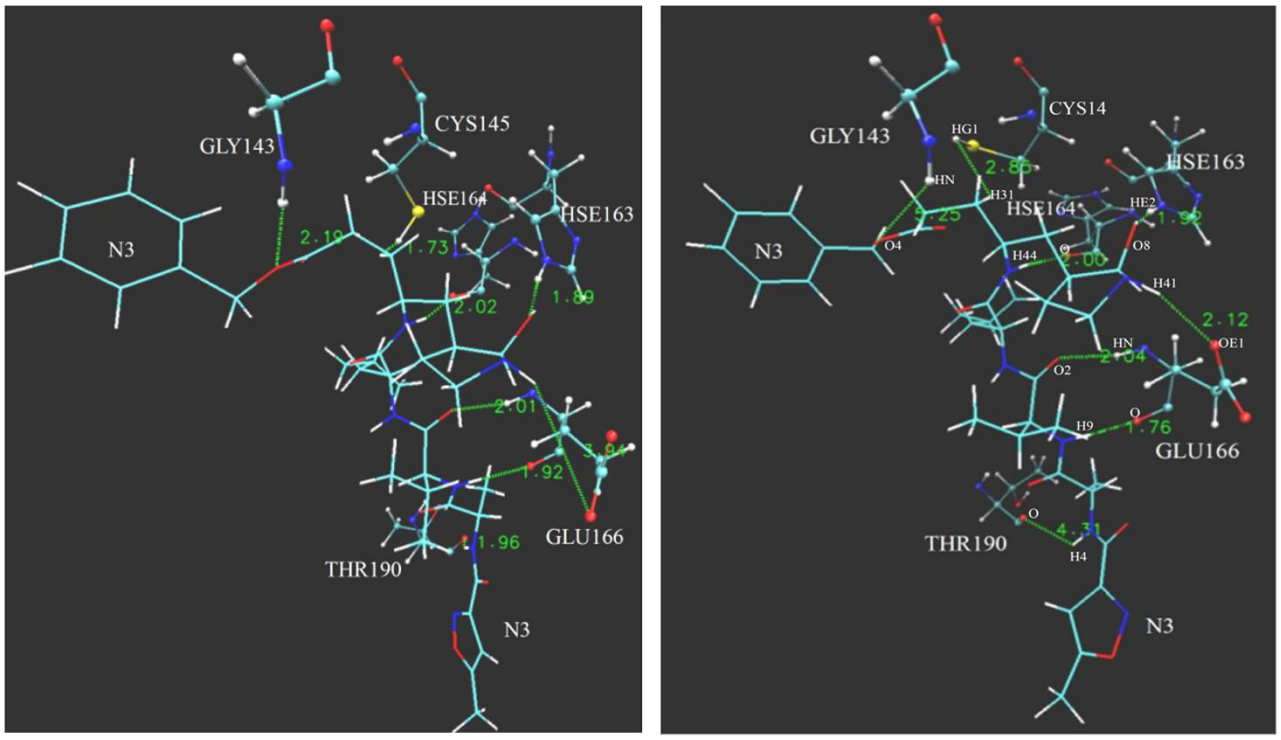


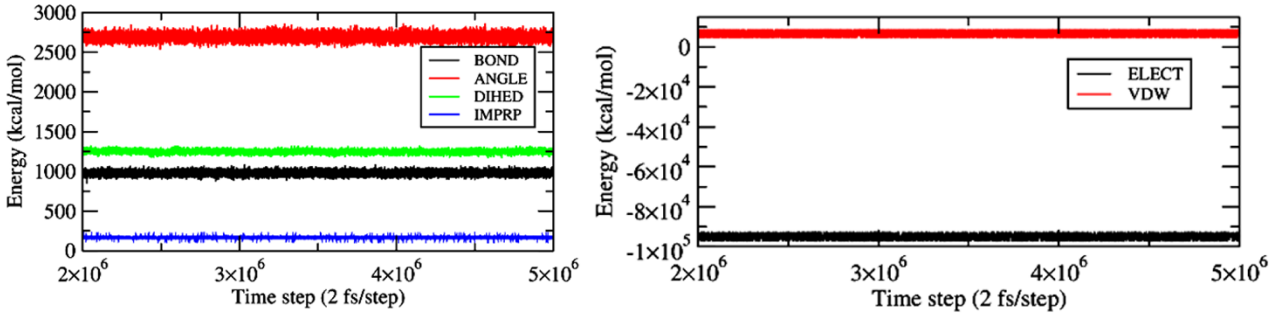
Fig. 8 RDFs of highly interacting atom pairs between N3 and MP-LBD residues that make H-bonds to each other except the atom pair: H31(N3)-HG1(CYS145).



(a) (b)

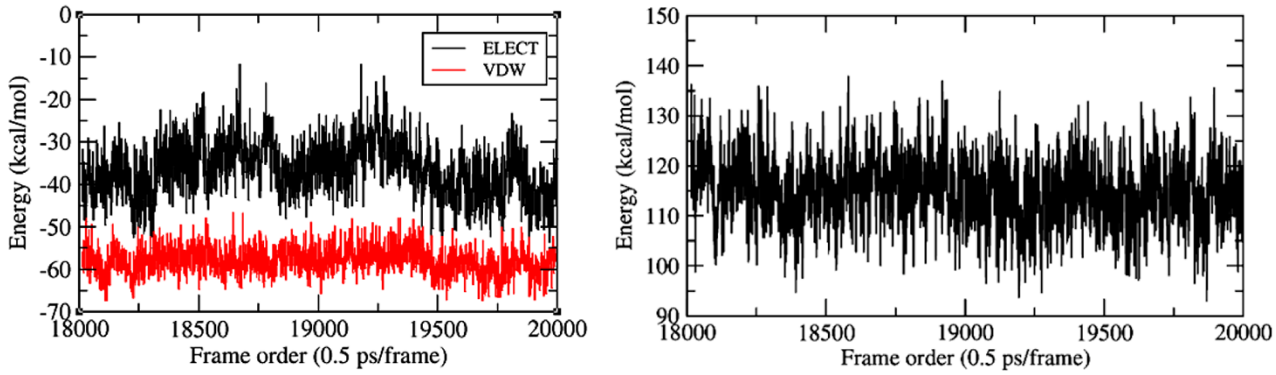
Fig. 9 MP-LBD residues interacting with the ligand moiety N3 are shown with H-bonding and distances for (a) initial structure and (b) final structure after 10 ns of equilibrating simulations. The H-bonding atoms are named (except for H31-HG1) as shown in (b).

The distributions of energy terms in CHARMM energy function of MP-N3 nanodroplet are shown in Figure 10. The mean ± SD of bond, angle, dihedral and improper energies are 977.63 ± 26.34, 2693.38 ± 40.95, 1246.67 ± 20.66 and 164 ± 8.99 in kcal/mol, respectively. In this way, electrostatic and van der Waals energies are -95034.83 ± 202.76 and 6748.95 ± 132.39 kcal/mol, respectively. These energies contribute to the internal energy of the system. The minutely fluctuating energy levels support the stability of MP-N3 system over the course of MD simulations after its geometrical optimization.



(a) (b)

Fig. 10 (a) Bonded and (b) non-bonded energy distributions in MP-N3 solvated system.



(a) (b)

Fig. 11 (a) Electrostatic and van der Waals interaction energies and (b) internal energy of the inhibitor N3 bounded to the substrate binding site of main protease in COVID-19 virus

In this study of MP-N3, like other molecules, the electrostatic interactions arise due to charge-charge, charge-dipole and dipole-dipole arrangements between the partially charged atoms, ligand functional groups and protein side chains. Here amine groups, LYS, ARG, and HIS are positively charged while carboxyl group, phosphate groups and glutamate side chains are negatively charged [26]. Van der Waals interactions controls the binding event between MP and N3. Its attractive part arises from the induced dipole pairs formed by the charge density fluctuations in non-covalent atom groups and repulsive part arise from the electronic repulsions in the partly overlapping electron cloud of the closest atoms. The dipole moments of polar MP-LBD residues affect their interactions with the inhibitor N3. The electrostatic and van der Waals interaction energies between N3 and LBD residues are calculated to be -36.24 ± 6.49 and -57.91 ± 3.26 kcal/mol, respectively. The negative interaction energies show that the inhibitor N3 is tightly bounded to MP of the novel COVID-19 virus. The internal energy of N3 molecule bounded to MP has been observed to be 114.81 ± 7.28 kcal/mole, and the fluctuating values of interaction and internal energy along the frames of simulation are reflected in Figures 11-a,b. The nature of mutation in MP-LBD and type of inhibitor or drugs used against the COVID-19 virus could make changes on the observed energy levels.

4. Summary

To investigate the structural properties of N3 liganded COVID-19 main protease (MP) through MP-N3 interactions, we perform NAMD simulations. Here, the inhibitor N3 is adapted by the substrate binding site of MP formed by the stable β-barrels and more flexible α-helices. The conformations of MP and N3 are such that the minutely fluctuating RMSDs do not exceed 2 Å which is within the limit of globular stability. RGs of MP and N3 which measure the compactness are almost constant with the respective values of 22.41 ± 0.10 Å and 6.34 ± 0.23 Å. Further, N3 has a very

small displacement per unit time-frame or crude velocity, i.e. 0.44 ± 0.20 Å/frame while interacting with the MP-LBD residues. Among the LBD residues, GLU166 which makes three stable H-bonds with N3 is the least flexible while GLY143 is the most flexible residue. GLY143 with no side chains has no steric hindrance so that its torsion angles are distributed in the white region of Ramachandran plot. The flexibility of residues is described by RMSF, residue speed, inter-atomic distance and RDF of atom pairs. The average RMSF of each MP residue along with N3 lies below 0.5 Å. The RDFs for three H-bonding atom pairs between GLU166 and N3 have sharp peaks ranging from 1.5 to 3 Å of radial distances. In this way, the possible H-bonding atom pair between GLY143 and N3 has a flat RDF peak with larger radial distance 5.48 ± 0.83 Å. The stability of MP-N3 complex depends on ~200 H-bonds among the MP residues and ~7 H-bonds between N3 and LBD residues. The SASA of each amino acid residue lies within the range of 173.32 - 358.93 $Å^2$ that describes the surface accessibility of MP with water molecules. In spite of larger solvent accessibility of N3, it gets hydrophobic interactions with the LBD residues resulting in van der Waals interaction energy of -57.91 ± 3.26 kcal/mol and electrostatic interaction energy of -36.24 ± 6.49 kcal/mol. In conclusions, the ligand moiety N3 is strongly interacting with the LBD residues of COVID-19 main protease. The physical parameters explained in this research are expected to be very useful tools for the discovery of drugs that restrict the viral replication and transcription.

Acknowledgments

We thank Dr. Pramod Aryal, Central Department of Biotechnology, Tribhuvan University (TU) for the fruitful discussion on the influence of inhibitors in COVID-19. Prof. Binil Aryal, Central Department of Physics (CDP), TU is gratefully acknowledged for providing the necessary facilities to complete this work. The computations were performed on the supercomputer machine (CDP,TU) supported by the Alexander von Humboldt Foundation, Germany under the equipment grants program and HERP-TU, Nepal.

References

1. Wang, F. *et al.* Structure of main protease from human coronavirus NL63: insights for wide spectrum anti-coronavirus drug design. *Scientific Reports* **6**, 22677 (2016).
2. Jin, Z. *et al.* Structure of M^{pro} from COVID-19 virus and discovery of its inhibitors. *Nature* (2020). https://doi.org/10.1038/s41586-020-2223-y
3. Zhavoronkov, A. *et al.* Potential COVID-2019 3C-like protease inhibitors designed using generative deep learning approaches. *Insilico Medicine Hong Kong Ltd A* 307: E1 (2020).
4. Liu, X. *et al.* The crystal structure of COVID-19 main protease in complex with an inhibitor N3. http://www.rcsb.org/structure/6LU7 (2020).
5. Lan, J. *et al.* Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* (2020). https://doi.org/10.1038/s41586-020-2180-5
6. Tha, S. *et al.* Prospects of indole derivatives as methyl transfer inhibitors: antimicrobial resistance managers. *BMC Pharmacology and Toxicology* **21**, 1-11 (2020).
7. Prasad, K. S. *et al.* Indole moiety induced biological potency in pseudo-peptides derived from 2-amino-2-(1H-indole-2-yl) based acetamides: Chemical synthesis, in vitro anticancer activity and theoretical studies. *Journal of Molecular Structure* **1217**, 128445 (2020).
8. Acharya, D. *et al.* Dysregulation of type I interferon responses in COVID-19. *Nature Reviews Immunology* (2020). https://doi.org/10.1038/s41577-020-0346-x
9. Schrodinger, L. L. C. The PyMOL molecular graphics system. *Version* 1.8 (2015).
10. Walser, R. *et al.* Molecular dynamics simulations of a double unit cell in a protein crystal: volume relaxation at constant pressure and correlation of motions between the two unit cells. *Proteins: Structure Function and Bioinformatics* **48**, 327–340 (2002).
11. Riniker, S. *et al.* Solvating atomic level fine-grained proteins in supra-molecular level coarse-grained water for molecular dynamics simulations. *European Biophysics Journal* **41**, 647–661 (2012).
12. Hu, C.Y. *et al.* Trimethylamine N-oxide influence on the backbone of proteins: an oligoglycine model. *Proteins: Structure Function and Bioinformatics* **78**, 695–704 (2010).

13. Levine, B. G. *et al.* Fast analysis of molecular dynamics trajectories with graphics processing units-radial distribution function histogramming. *Journal of Computational Physics* **230**, 3556–3569 (2011).
14. Hollingsworth, S.A. *et al.* A fresh look at the Ramachandran plot and the occurrence of standard structures in proteins. *Biomolecular Concepts* 1, 271–283 (2010).
15. Ho, B. K. *et al.* The Ramachandran plots of glycine and pre-proline. *BMC Structural Biology* (2005). http://www.biomedcentral.com/1472-6807/5/14
16. Ausaf Ali, S. *et al.* A review of methods available to estimate solvent-accessible surface areas of soluble proteins in the folded and unfolded states. *Current Protein and Peptide Science* **15**, 456–476 (2014).
17. Lee, B. *et al.* The interpretation of protein structures: estimation of static accessibility. *Journal of Molecular Biology* **55**, 379–400 (1971).
18. Rekik, N. *et al.* Equivalence between the classical and quantum IR spectral density approaches of weak H-bonds in the absence of damping. *Journal of Physical Chemistry A* **122**, 2108–2115 (2018).
19. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry* **26**, 1781–1802 (2005).
20. MacKerell, A. D. *et al.*All-atom empirical potential for molecular modeling and dynamics studies of proteins. *Journal of Physical Chemistry B* **102**, 3586–3616 (1998).
21. MacKerell, A. D. *et al.*Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *Journal of Computational Chemistry* **25**, 1400–1415 (2004).
22. Zoete, V. *et al.* SwissParam: a fast force field generation tool for small organic molecules. *Journal of Computational Chemistry* **32**, 2359–2368 (2011).
23. Humphrey, W. *et al.* VMD-visual molecular dynamics. *Journal of Molecular Graphics* **14**, 33–38 (1996).
24. Verlet, L. Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Physical Review* **159**, 98–103 (1967).
25. Lamichhane, T. R. *et al.* Structural changes in thyroid hormone receptor-beta by T3 binding and L330S mutational interactions. *AIMS Biophysics* **7**, 27-40 (2020).
26. Bronowska, A. K. Thermodynamics of ligand-protein interactions: implications for molecular design. *Thermodynamics-Interaction Studies-Solids, Liquids and Gases*. IntechOpen (2011).