Elucidating Biophysical Basis of Binding of Inhibitors to SARS-CoV-2 Main Protease by Using Molecular Dynamics Simulations and Free Energy Calculations

Md Fulbabu Sk, Rajarshi Roy, Nisha Amarnath Jonniya, Sayan Poddar, Parimal Kar*

Discipline of Biosciences and Biomedical Engineering, Indian Institute of Technology Indore, Khandwa Road, MP 453552, India.

*Corresponding Author: Parimal Kar | Email: parimal@iiti.ac.in | Phone: +91 731 2438700 (Ext. 550)

Abstract

The recent outbreak of novel “coronavirus disease 2019” (COVID-19) has spread rapidly worldwide, causing a global pandemic. In the absence of a vaccine or a suitable chemotherapeutic intervention, it is an urgent need to develop a new antiviral drug to fight this deadly respiratory disease. In the present work, we have elucidated the mechanism of binding of two inhibitors, namely α-ketoamide and Z31792168 to SARS-CoV-2 main protease (M^{pro} or $3CL^{pro}$) by using all-atom molecular dynamics simulations and free energy calculations. We calculated the total binding free energy (ΔG_{bind}) of both inhibitors and further decomposed ΔG_{bind} into various forces governing the complex formation using the Molecular Mechanics/Poisson-Boltzmann Surface Area (MM/PBSA) method. Our calculations reveal that α-ketoamide is more potent (ΔG_{bind}= - 9.05 kcal/mol) compared to Z31792168 (ΔG_{bind}= - 3.25 kcal/mol) against COVID-19 $3CL^{pro}$. The increase in ΔG_{bind} for α-ketoamide relative to Z31792168 arises due to an increase in the favorable electrostatic and van der Waals interactions between the inhibitor and $3CL^{pro}$. Further, we have identified important residues controlling the $3CL^{pro}$-ligand binding from per-residue based decomposition of the binding free energy. Finally, we have compared ΔG_{bind} of these two inhibitors with the anti-HIV retroviral drugs, such as lopinavir and darunavir. It is observed that α-ketoamide is more potent compared to both lopinavir and darunavir. In the case of lopinavir, a decrease in the size of the van der Waals interactions is responsible for the lower binding affinity compared to α-ketoamide. On the other hand, in the case of darunavir, a decrease in the favorable intermolecular electrostatic and van der Waals interactions contributes to lower affinity compared to α-ketoamide. Our study might help in designing rational anticoronaviral drugs targeting the SARS-CoV-2 main protease.

Introduction

The recent outbreak of "coronavirus disease 2019" (COVID-19) poses a serious public health risk. COVID-19 is caused by the novel coronavirus (nCoV) or SARS-CoV-2, which is a highly contagious and pathogenic virus. The disease has spread worldwide since its outbreak in the city of Wuhan, China, in 2019 (Hui et al. 2020; Wu, Leung & Leung 2020). As of 3rd April 2020, the total number of people diagnosed with the COVID-19 viral infection exceeds over 1 million, with more than 55000 fatalities worldwide. In India, more than 2100 people have been infected with 56 deaths (https://www.mohfw.gov.in/). The number of positive cases is increasing exponentially due to human-to-human transmission. The World Health Organisation (WHO) has declared COVID-19 a global pandemic.

nCoV is a nonsegmented, single-stranded, positive-sense RNA virus that belongs to the beta coronavirus family (Woo, Huang, Lau & Yuen 2010). This gene encodes four structural proteins, namely spike glycoprotein (S), a small envelope protein (E), matrix glycoprotein (M), and nucleocapsid protein (N) (Rota et al. 2003). The S protein attaches with the host receptor by the specific receptor binding (RBD) (Wan, Shang, Graham, Baric & Li 2020), while the N protein binds to RNA in multiple sites to make a helical nucleocapsid structure (Chang, Hou, Chang, Hsiao & Huang 2014). Along with these proteins, chymotrypsin-like cysteine protease ($3CL^{pro}$) is an essential protein for maintaining the viral life cycle by cleaving the essential polyprotein PP1A to individual functional components. $3CL^{pro}$ contains 9 α-helices and 13 β-strands (see Figure 1 A), creating three domains, domain I (residues 8-101), domain II (residues 102-184), and domain III (residues 201-306). The domains II and III are connected by a long loop (residues 185-200). The active site of this protease is located between domains I and II, and each domain contributes a single residue to the catalytic dyad, His41 and Cys145, respectively. $3CL^{pro}$ plays a key role in mediating viral replication and transcription making it an attractive drug target for COVID-19 treatment.

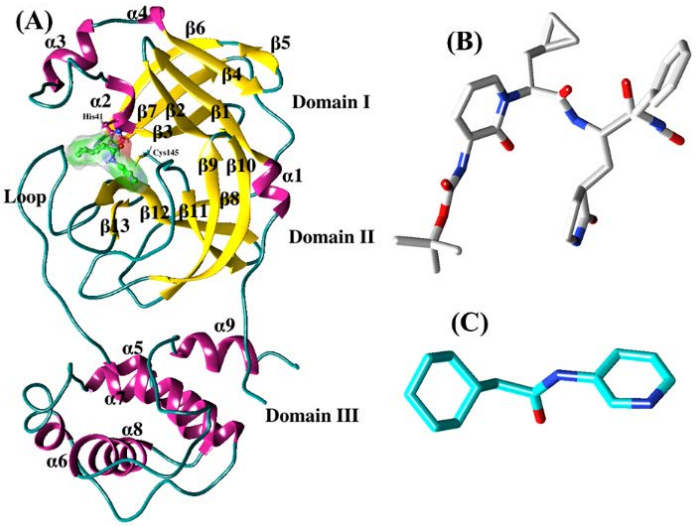

Figure 1: (A) The ribbon representation of SARS-CoV-2 $3CL^{pro}$ (PDB code: 5R84). The inhibitors, α-ketoamide (B) and Z31792168 (C) are represented as a stick model. The catalytic dyad residues (H41 and C145) are shown in ball and stick representation.

Recently, Jin and coworkers have designed a modified peptide inhibitor for this protease and solved the X-ray crystal structure of the complex (Jin et al. 2020). The availability of the crystal structure of $3CL^{pro}$ opens up a lot of opportunities to develop new antiviral drugs. This structure is employed to computationally evaluate the repurposing of different FDA approved drugs against COVID-19 (Wang 2020). Several studies have predicted the binding strength of darunavir, indinavir, and other HIV protease drugs against $3CL^{pro}$ (Lin, Shen, He, Li & Guo 2020; Sang, Tian, Meng & Yang 2020).

Recently, Zhang and coworkers have designed a poten inhibitor, α-ketoamide [tert-butyl (1-((S)-1-(((S)-4-(benzylamino)-3,4-dioxo-1-((S)-2-oxopyrrolidin-3-yl)butan-2-yl)amino)-3-cyclopropyl-1-oxopropan-2-yl)-2-oxo-1,2-dihydropyridin-3-yl)carbamate] (see Figure 1B) targeting $3CL^{pro}$ and resolved the crystal structure of $3CL^{pro}$/α-ketoamide (Zhang et al. 2020). This inhibitor shows a promising binding as well as a pronounced lung tropism, and can be administered by inhalative route (Zhang et al. 2020). In the present work, we have studied the mechanism of binding of this inhibitor to SARS-CoV-2 $3CL^{pro}$ using molecular dynamics simulations in conjunction with the popular molecular mechanics/Poisson-Boltzmann surface area scheme. We have also considered another inhibitor, Z31792168 (2-cyclohexyl-N-pyridin-3-yl-ethanamide) in our study (see Figure 1C) and compared with α-ketoamide. Finally, the binding free energy of these two inhibitors are compared with the anti-HIV retroviral drugs, such as lopinavir and darunavir.

2. Materials and Methods

2.1 Simulation Protocol

The initial coordinates for our molecular dynamics simulations were obtained from the X-ray crystallographic structure of the SARS-CoV-2 $3CL^{pro}$ complexed with the inhibitors α-ketoamide (PDB: 6Y2G) and Z31792168 (PDB: 5R84) (Berman et al. 2002; Zhang et al. 2020). The missing residues in $3CL^{pro}$ were added by using the Modeller (Fiser & Šali 2003; Webb & Sali 2014) plugin in UCSF Chimera software (Pettersen et al. 2004). The protonation states of the charged residues were determined using the Propka 3.1 webserver (Olsson, Søndergaard, Rostkowski & Jensen 2011). All simulations were executed by the *pmemd.cuda* mudule of AMBER18 (Case 2018) package and analyses were done by the Cpptraj module (Roe & Cheatham III 2013). We used the latest AMBER ff14SB force field (Maier et al. 2015) to describe the protein structure and the updated generalized Amber force field (GAFF2) (Wang, Wolf, Caldwell, Kollman & Case 2004) is used to assign parameters to small molecules. All the missing hydrogen atoms were added by the *Leap* module of AMBER (Salomon-Ferrer, Case & Walker 2013). The inhibitors were assigned AM1-BCC (Jakalian, Jack & Bayly 2002) charge, which was calculated by utilizing the Antechamber (Wang, Wang, Kollman & Case 2006) module of AMBER18. The systems were solvated in a truncated octahedron periodic box with an explicit TIP3P (Price & Brooks III 2004) water model and a 10 Å buffer distance was considered from the complex along each side.

A suitable integer number of counterions (Na^+) were added for neutralizing the whole system. The temperature was kept at 300 K and controlled by the Langevin thermostat (Loncharich, Brooks & Pastor 1992). The system pressure was monitored using a Berendsen Barostat (Berendsen, Postma, van Gunsteren, DiNola & Haak 1984) and kept at 1.0 bar. All bond lengths involving hydrogen atoms were constrained by the SHAKE algorithm (Kräutler, Van Gunsteren & Hünenberger 2001). We used a time-step of 2.0 fs for the simulation. The particle mesh Ewald summation (PME) (Darden, York & Pedersen 1993) approach was used to compute the long-range electrostatic interactions. For all cases, the nonbonded cut-off was fixed at 10.0 Å.

Firstly, each complex was optimized using 500 steps of the steepest descent algorithm followed by another 500 cycles of the conjugate gradient scheme. During the minimization, the receptor-inhibitor complexes were restrained to their respective coordinates with a force constant of 2.0

kcal $mol^{-1}Å^{-2}$. Next, we carried out the minimization without applying any harmonic restraint on the solutes to remove any residual steric clashes. After the minimization, the temperature of each solvated system was gradually increased from 0 K to 300 K at the NVT (canonical) ensemble with a force restraint of 2.0 kcal $mol^{-1}Å^{-2}$ acting on the solute atoms. After that, a 50 ps MD simulation with a restraint force constant of 2.0 kcal $mol^{-1}Å^{-2}$ for all solute atoms at constant 1.0 atm pressure was conducted using Berendsen Barostat (Berendsen et al. 1984) at a fixed temperature of 300 K. After 1.0 ns of an equilibration phase, the production simulation was carried for 100 ns at the NPT ensemble, and the Cartesian coordinates were stored every 10 ps. Overall, we accumulated 10000 snapshots corresponding to each production simulation.

2.2 Dynamic cross-correlation map (DCCM) analysis

The correlated and non-correlated atomic motions of complex protein residues were computed with the help of DCCM (McCammon 1984; Hünenberger, Mark & Van Gunsteren 1995) analysis using Cpptraj (Roe & Cheatham III 2013) module of AmberTools19. Herein, we avoid the apparent correlations between slow side-chain fluctuations and backbone movements. To reduce the statistical noise, we considered only C_α atomic coordinates of each residue. We avoid initial 10 ns trajectories for system stabilization and used rest 90 ns production simulation trajectories for this analysis.

2.3 Principal component analysis (PCA)

PCA (Ichiye & Karplus 1991) gives us detailed information about residual correlation movements and functional significance of each residue. Similar to DCCM analysis, only Cα atomic coordinates were used for this analysis. The atomic fluctuations of C_α-atoms of each residue form a covariance matrix, and the diagonalization of the covariance matrix gives us eigenvectors and corresponding eigenvalues. The eigenvectors indicate the movements and directions, and eigenvalues are movement strength. These eigenvectors and associate eigenvalues represent the set of principal components (PCs), which may be used to describe the movement characteristics. We also calculate the cosine content via GROMACS (g_covar, g_anaeig, and g_analyze modules) (Hess, Kutzner, Van Der Spoel & Lindahl 2008) of first few PCs to check the statistical convergence significance of each trajectory. The higher value of conformational sampling convergence gives a very lower value of cosine contents. Our first few PC's cosine contents values lie between 0 to 0.6 for each case, which indicates high conformational sampling convergence.

2.4 Free energy landscape

The free energy landscape (FEL) calculations were performed by AmberTools19 Cpptraj module of AMBER18 using the below equation (1) (Frauenfelder, Sligar & Wolynes 1991);

$$G_i = -k_B T \ln\left(\frac{N_i}{N_m}\right) \quad (1)$$

where k_B represents the well-known Boltzmann constant, T is the absolute temperature of each simulation system. N_i is the bin i population and N_m is the population of the most populated bin. So, bins with no population are an artificial barrier as an indication of the lowest probability. To measure the conformational variability of each system in terms of FELs, we considered the first principal components (PC1 and PC2) as a main reaction coordinates for estimating the landscapes.

2.5 Energy calculations

In order to estimate the stabilisation of binding systems, the binding free energies or affinity of all four inhibitors against SARS-CoV-2 $3CL^{pro}$ were calculated by a widely used molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) method (Kollman et al. 2000; Wang, Wang, Kollman & Case 2001; Kar, Seel, Hansmann & Höfinger 2007; Kar, Wei, Hansmann & Höfinger 2007). In this study, 2000 structural frames isolated from the last 20 ns of trajectories of simulations at an interval of 1 ps were applied to run the MM-PBSA calculations. The theory behind the MM-PBSA calculation, we clearly discussed in our WNK1 kinase and HIV-1/PR study (Jonniya & Kar 2019; Jonniya, Sk & Kar 2019; Roy, Ghosh & Kar 2020; Sk, Roy & Kar 2020) and here we used the same. The normal mode analysis (Xu, Shen, Zhu & Li 2011) method was used to calculate the configurational entropy (S_{MM}). Due to high computational cost, we considered only 50 configurations from last 20 ns simulations for entropy calculations. We were also performed the per-residue decomposition of binding energy by molecular mechanics generalize- Born surface area (MM-GBSA) scheme to know the individual residual contributions of each residue. All the parameters used this calculation were developed by Onufreiv and Bashford (Onufriev, Bashford & Case 2004).

2.6 Hydrogen bond criterion

The hydrogen bonds between inhibitors and SARS-CoV-2 residues were analysed using AmberTool19 Cpptraj module (Roe & Cheatham III 2013). The formation of hydrogen bond defined by a distance and an angle cut-off of ≤ 3.5 Å and ≥120º, respectively; between donor (D) and acceptor (A) atom.

3. Results and Discussion

3.1 Structural stability and flexibility analysis

To verify the convergency of our simulations, we estimated the root mean square deviations (RMSDs) of backbone atoms with respect to their energy minimized equilibrated structure, and the temporal RMSD of two complexes and apo-3CL^{pro} is shown in Figure 2 (A). From Figure 2 (A), it is clear that apo 3CL^{pro} reached a stable equilibrium after 30 ns while the 3CL^{pro}/α-ketoamide system reached equilibrium after 40 ns. In the case of 3CL^{pro}/Z31792168, the equilibrium is reached after 60 ns.

Table 1: Average RMSD, SASA and radius of gyration of 3CL^{pro}-inhibitor complexes and apo 3CL^{pro}.

System	Backbone RMSD (Å)	SASA (Å^2)	Radius of gyration (R_g) (Å)
Apo	2.10 (0.01)	14650 (3.5)	22.13 (0.01)
6Y2G	2.25 (0.01)	13969 (3.2)	22.03 (0.01)
5R84	1.95 (0.01)	14208 (3.6)	21.96 (0.01)

The average RMSDs for all simulated systems were calculated and reported in Table 1. The average RMSD was found to vary between 1.95 Å and 2.25 Å for all cases. The highest deviation (2.25 Å) was obtained for 3CL^{pro}/α-ketoamide, while the lowest RMSD (1.95 Å) was obtained for 3CL^{pro}/Z31792168. We also calculated the RMSD of each ligand, and the time evolution of ligand RMSD is shown in Figure 2 (B). As can be seen from Figure 2(B), two inhibitors reached equilibrium after 50 ns of MD simulations. In the case of α-ketoamide, the frequency distribution of RMSD is characterized by a single peak at 2.4 Å while two peaks (0.23 Å and 1.1 Å) are obtained for Z31792168.

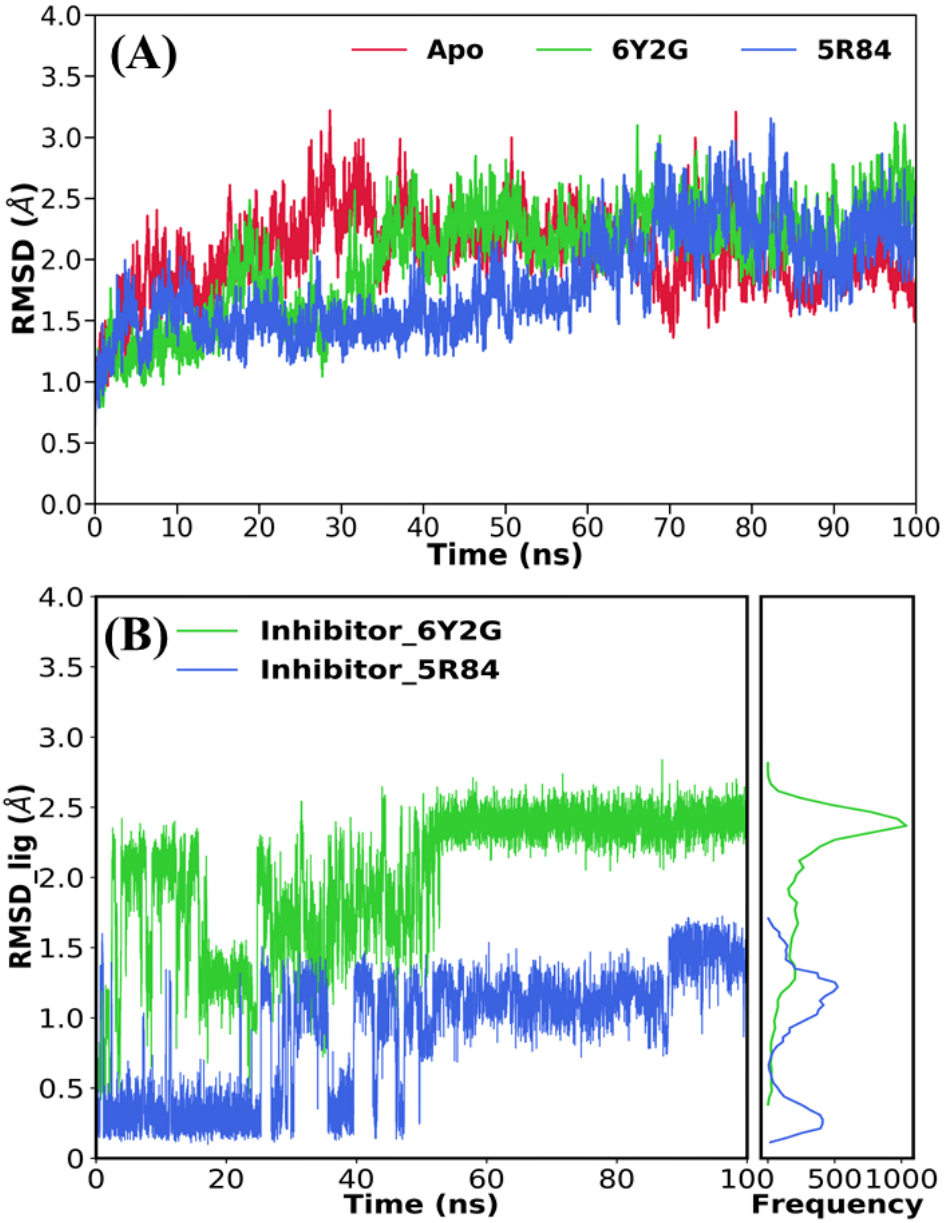

Figure 2: Time evolution of root-mean-square deviations (RMSD) of (A) backbone atoms of $3CL^{pro}$ relative to their respective minimized structure and (B) heavy atoms of inhibitor.

Next, we investigated the flexibility of different parts of $3CL^{pro}$ by calculating the root mean square fluctuations (RMSFs) of C_α-atoms for all systems and is shown in Figure 3. It is evident from Figure 3 that all three systems display more or less similar fluctuations. In the case of apo-$3CL^{pro}$, domain II (residues 102-184) shows slightly higher fluctuations, which gets diminished after the ligand binding. In the case of $3CL^{pro}$/α-ketoamide, a relatively higher value of RMSF is obtained around residue 50 (domain I) compared to apo or $3CL^{pro}$/Z31792168.

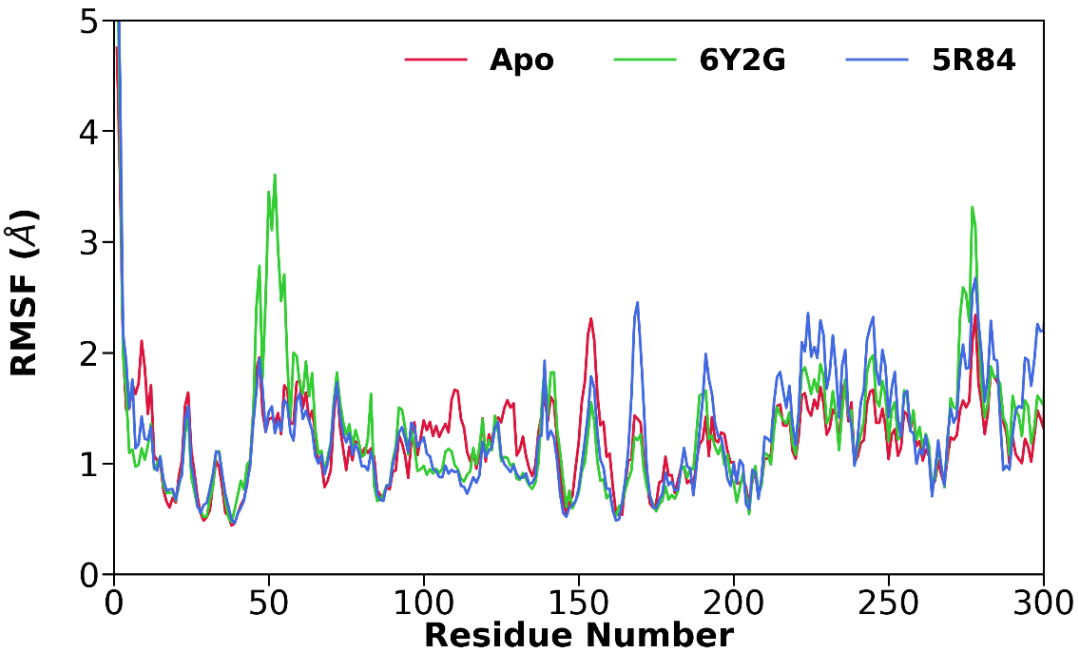

Figure 3: The root-mean-square fluctuations (RMSFs) of C_α atoms for apo (red), $3CL^{pro}$/α-ketoamide (green), and $3CL^{pro}$/Z31792168 (blue).

The structural compactness of each system was analyzed by estimating the radius of gyration (R_g) from their respective MD trajectories, and the average values are reported in Table 1. A similar R_g is obtained for all systems. Finally, the solvent-accessible surface area (SASA) that indicates the degree of solvent exposure was also calculated and reported in Table 1. It is evident from Table 1 that SASA values vary between 13969 $Å^2$ and 14650 $Å^2$ for all systems. The highest value (14650 $Å^2$) was obtained for apo, while the lowest SASA value (13969 $Å^2$) was reported for $3CL^{pro}$/α-ketoamide. An intermediate SASA value (14208 $Å^2$) was obtained for $3CL^{pro}$/Z31792168.

3.2 Dynamic cross-correlation analysis

To elucidate the effect of inhibitor binding on the internal dynamics of $3CL^{pro}$, the cross-correlation matrix was calculated by using the coordinates of C_α atoms from MD trajectories, and the dynamic cross-correlation map (DCCM) is displayed in Figure 4. Overall, after inhibitor binding the apparent anti-correlation motions are observed in both complexes.

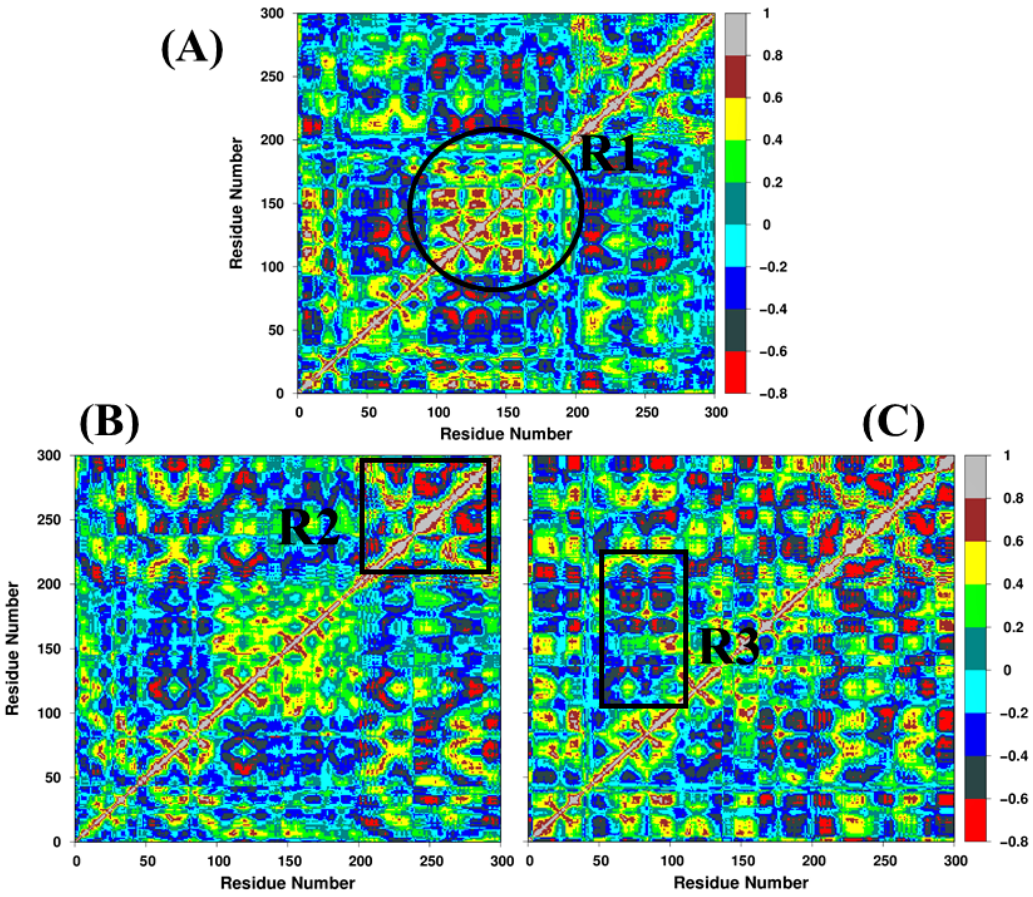

Figure 4: Cross-correlation matrices of the fluctuations of the coordinates for C_α atoms around their mean positions after the equilibrium of MD simulation. The extent of correlated motions and anti-correlated motions are color-coded (1 = highly correlated; 0 = neutral; -1 = anti-correlated).

For Apo system, the correlation map is different and the domain I and domain II shows highly anti-correlated movement but domain III has lower strength of anti-correlation movement. The region R1, which is represents the binding cavity area on the map, indictes that after inhibitor binding the little bit anti-correlation motion has increased and diagonals are highly correlated. Similarly, for 5R84 complex, after inhibitor binding induces more anti-correlation motion as compared to 6Y2G complex, see region R1. The region R2 indicates the residual motion of domain III for all cases and it is evident that after inhibitors binding the anti-correlation movements increases for both complex. Similarly, inhibitor bindings produced different motion modes at domaun I, see region R3. The anti-correlation modes are diminishes for the case of 6Y2G complex and create a stable environment for inhibitor binding to their respective cavity.

3.3 Principal component analysis

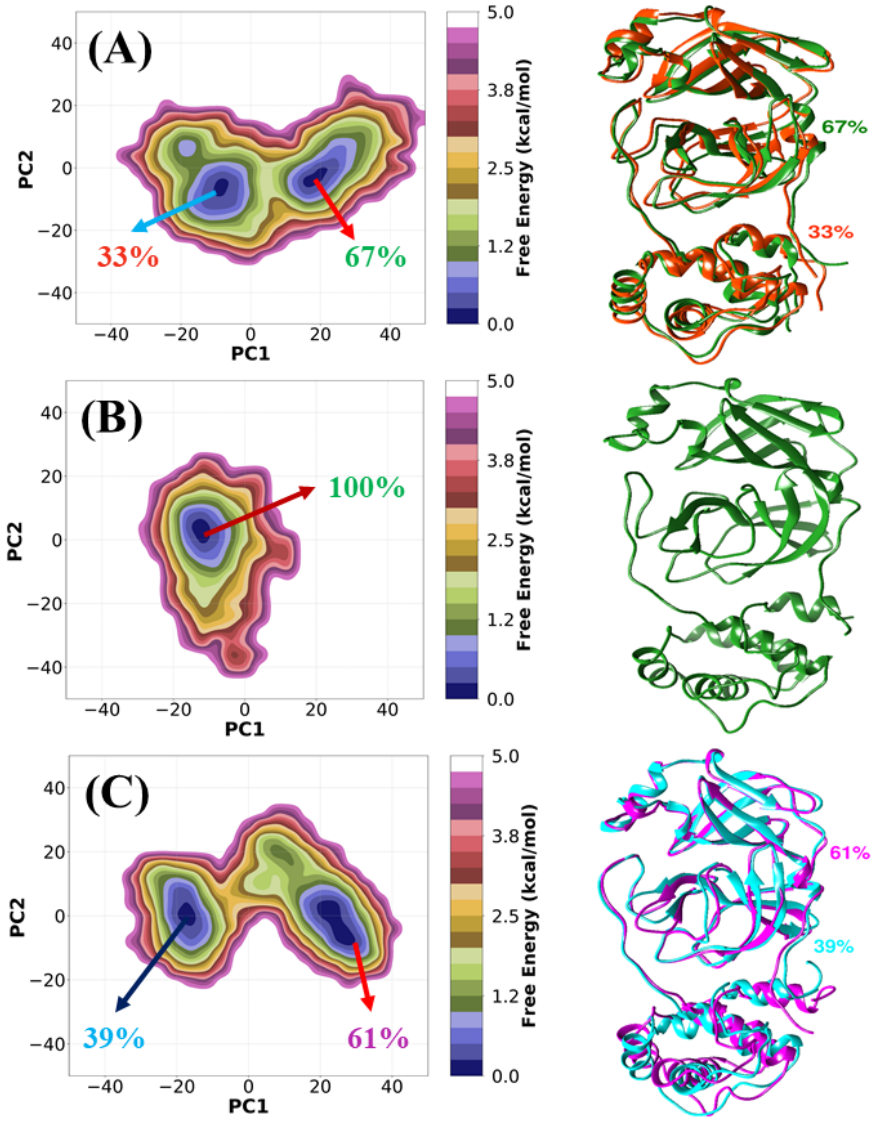

Figure 5: Two dimensional free energy landscapes (FELs) generated by projecting the principal components, PC1 and PC2 of apo (A), 6Y2G (B) and 5R84 (C) complexes in MD simulation at 300 K and the representative structures. The representative structures are shown on its right panel by K-means clustering analysis.

To further characterize the effect of inhibitor binding on the dynamics of $3CL^{pro}$ and extract the structural variations in detail, PCA was applied to the coordinates of the backbone atoms of inhibitors bounded and apo form $3CL^{pro}$. The PCA results in terms of eigenvectors (3N, 3 x 306=918) versus eigenvalues obtained by diagonalization of the atomic fluctuations covariance matrix for complexes and apo form $3CL^{pro}$ are shown in Figure S1. The collective motions of the localized fluctuations can be defined by the first few principal components. The first 10

eigenvectors capture ~87%, ~88% and ~89% of the total motion in apo, 6Y2G and 5R84 complexes.

In general, the two largest principal components (PC1 and PC2) account for more than 70% of the overall fluctuations for apo form and inhibitor-bound $3CL^{pro}$. The two dimensional free energy landscapes (FELs) of the apo and inhibitor bound $3CL^{pro}$ for the first two principal component (PC) is represented in Figure 5. It is evident from 5, the pattern of a basin on the FEL are different for the different systems. It is notable that the conformational space sampled by inhibitor-bound 6Y2G system is more restricted, whereas the apo and 5R84 systems sample wider conformational spaces, see in Figure S2. These observations suggest that the apo and 5R84 systems increasing the flexibility, which agrees the RMSF data (Figure 3). The conformational space of apo and 5R84 complex showed a two dispersed free energy basin, whereas in the case of 6Y2G complex showed a single and stable global free energy minimum confined within a basin on the FEL, as is evident from Figure5. Porcupine plots are shown in Figure S3, which generated via first mode that shows the largest collective motions of Cα atoms which has been mapped onto the average structure using VMD (Humphrey, Dalke & Schulten 1996) software Normal Modes Analysis plugin. To visualize the detailed movements for each protein, corresponding structures were selected based on K-means (Hartigan & Wong 1979) clusters of PCA distributions. The structures were obtained from cluster centers and superimposed with each other. Overall, it is observed that the mobility is higher for apo and 5R84 $3CL^{pro}$ complex compared to 6Y2G $3CL^{pro}$ complex.

3.4 Dihedral principal component analysis

To further discover the conformational alterations of loop (residues 185-200) between domain II (residues 102-184) and domain III (residues 201-306), dihedral PCA (dPCA) was performed based on the dihedral angles (ϕ_i,ψ_i) of the peptide group in loop. Cartesian coordinates based PCA does not provie the fully correct internal and overall movements and dPCA gives us the information about internal and overall motions pattern.

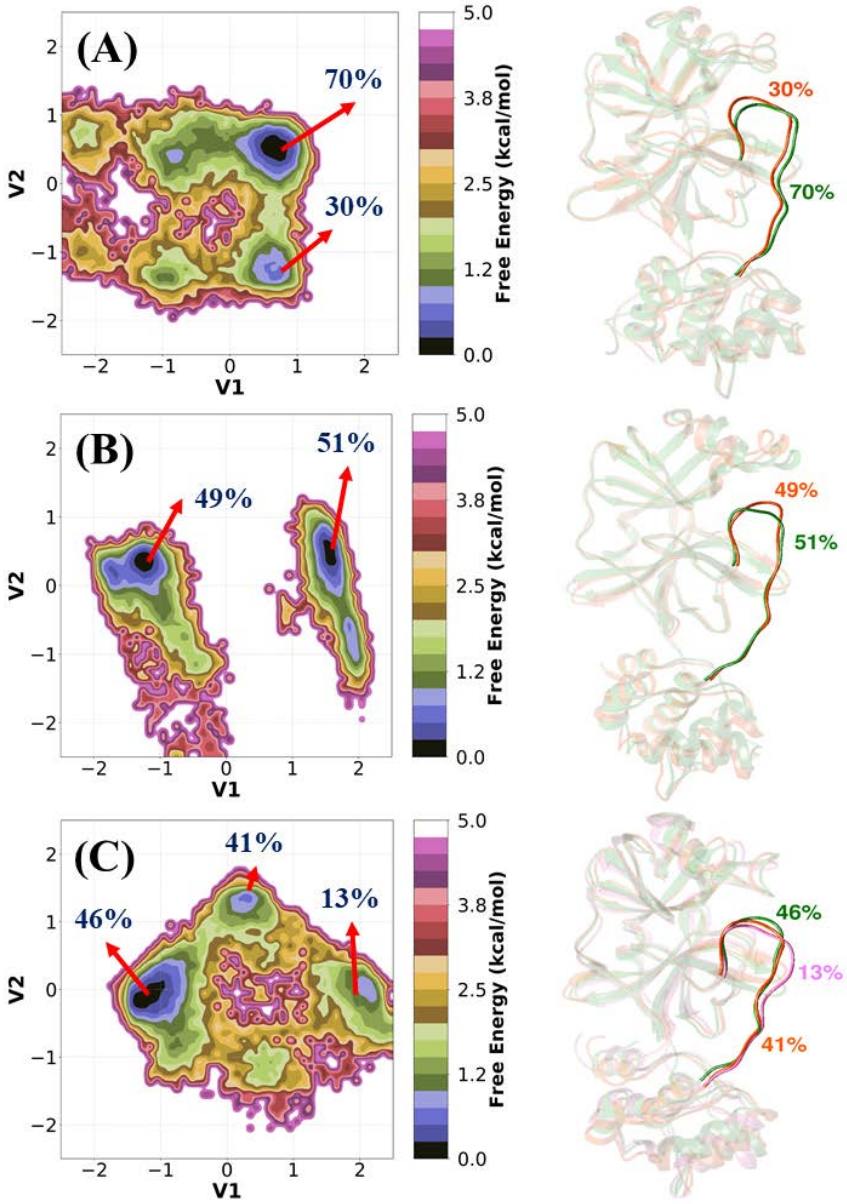

Figure 6: The two dimensional free energy landscapes (FELs) of (A) apo, (B) 6Y2G complex and (C) 5R84 $3CL^{pro}$ complex. FEL obtained from dPCA and their representative structures are shown on its right panel by K-means clustering analysis.

From Figure 6, it is evident that dPCA indicates diverse loop conformations for each systems. The apo $3CL^{pro}$ has one global minimum (70%) and one local minimum (30%). Similarly, 6Y2G complex has two almost equiprobable free energy basins (51% and 49%). The conformational space of 5R84 is quite wider as we see in Figure 6 (C), which contained one global free energy basin (46%) and two local free energy minima of 41% and 13%. In this analysis we confirmed that this loop play a important role in inhibitor binding as well as defined

the dynamical nature of 3CLpro structures. Thus, inhibitor-induced fluctuation of this loop in 3CLpro is worth paying more attention in the next-generation drug design.

3.5 Binding affinity of inhibitors with 3CL^{pro}

To understand the biophysical basis of recognition of inhibitors (I_{6Y2G} and I_{5R84}) with 3CL ^{Pro}, molecular dynamic simulations in conjunction with Molecular-Mechanics Poisson-Boltzmann Surface Area (MM-PBSA) scheme were computed in AMBER18. It provides different individual components contribution to total binding free energy such as ΔE_{vdW}, ΔE_{elec}, ΔG_{pol}, ΔG_{np}, and $T\Delta S$. For binding free energy calculations of both complexes, 2000 structures were taken from the stable region, and 50 structures were considered for the entropy calculations.

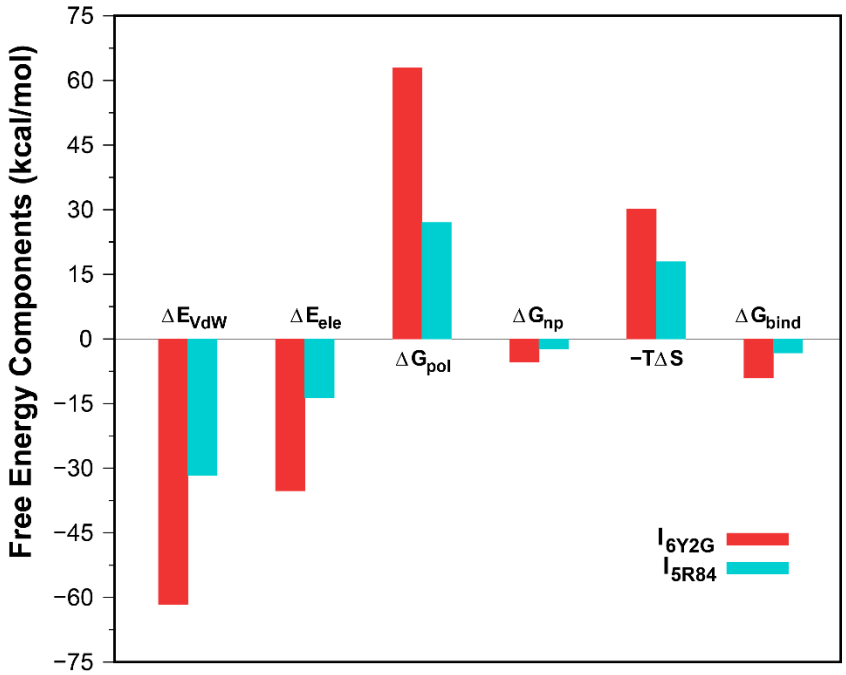

Figure 7 : Energy components (kcal/mol) for the binding of inhibitors to the 3CL^{pro}. ΔE_{vdW}, van der Waals energy; ΔE_{elect}, electrostatics energy in gas phase; ΔG_{polar}, polar solvation energy; $\Delta G_{nonpolar}$, nonpolar solvation energy; $T\Delta S_{MM}$, configurational entropy contribution and ΔG_{bind}, total binding energy.

A summary of binding components in the binding free energy of inhibitors with 3CL^{Pro} is shown graphically in Figure 7**,** and the data are listed in Table 2. According to Table 2 the estimated binding free energy of I_{6Y2G} is higher (-9.05 kcal mol^{-1}) compared to I_{5R84} (-3.25 kcal mol-1) with the 3CL^{pro}. It is evident from Table 2 that van der Waal (ΔE_{vdW}), intermolecular electrostatic interactions (ΔE_{elec}) and non-polar solvation energy (ΔG_{np}) favored the binding, where van der Waals interactions play a significant role in the binding of the

complexes. For which the I_{6Y2G} complex exhibits a higher value being -61.63 kcal mol^{-1} than I_{5R84} complex (-31.61 kcal mol^{-1}). Similarly, the electrostatic interaction favored more in I_{6Y2G} complex (-35.23 kcal mol^{-1}) than I_{5R84} complex (-13.64 kcal mol^{-1}). However, the polar solvation energy (ΔG_{pol}) and configurational entropy ($T\Delta S$) disfavor the complex formation.

Table 1 suggests that the total polar contribution ($\Delta E_{elec} + \Delta G_{pol}$) of the inhibitor I_{5R84} is less unfavorable with 3CL^{pro} (13.44 kcal mol^{-1}) compared to the inhibitor I_{6Y2G} (27.71 kcal mol^{-1}). Moreover, the disfavoring component of the configurational entropy exhibits low value being 17.92 kcal mol^{-1} for the inhibitor I_{5R84} compared to the inhibitor I_{6Y2G} (30.18 kcal mol^{-1}). However, the potency of the inhibitor I_{6Y2G} against 3CL^{pro} is comparatively higher than the inhibitor I_{5R84}. It is due to the total non-polar components ($\Delta E_{vdW} + \Delta G_{np}$) coming from the hydrophobic residues in the binding pocket that contributes more to the inhibitor I_{6Y2G} being -66.94 kcal mol^{-1} compared to the inhibitor I_{5R84} (-34.6 kcal mol^{-1}).

Table 2: Energetic components of the binding free energy for SARS-CoV-2-inhibitors complexes calculated using MM-PBSA (kcal/mol). Standard errors of the mean (SEM) are provided in parentheses.

Components	α-ketoamide	Z31792168	Lopinavir[R](Wang 2020)	Darunavir[R](Sang et al. 2020)
ΔE_{vdW}	-61.63 (0.10)	-31.61 (0.06)	-20.09 (0.63)	-41.32
ΔE_{elec}	-35.23 (0.13)	-13.64 (0.09)	-52.46 (0.33)	-5.80
ΔG_{pol}	62.94 (0.13)	27.08 (0.09)	46.58 (0.56)	29.01
ΔG_{np}	-5.31 (0.00)	-2.99 (0.00)	-4.59 (0.02)	-4.75
ΔG_{solv} [a]	57.63 (0.13)	24.09 (0.09)	41.99 (0.56)	24.26
$\Delta G_{pol + elec}$ [b]	27.71 (0.18)	13.44 (0.13)	-5.88 (0.65)	23.21
ΔE_{MM} [c]	-96.86 (0.16)	-45.25 (0.11)	-30.56 (0.71)	-47.12
$^{-}T\Delta S$ [d]	30.18 (0.81)	17.92 (0.88)	-23.93 (0.01)	NA
ΔG_{Total} [e]	-39.23 (0.11)	-21.16 (0.01)	-30.56	-22.86
ΔG_{Bind} ^{sim}	-9.05 (0.82)	-3.25 (0.88)	-6.63 (0.28)	NA
IC_{50} ^{Exp}	0.67 ± 0.18 μM	NA	NA	NA

[a]$\Delta G_{solv} = \Delta G_{np} + \Delta G_{pol}$, [b]$\Delta G_{pol + elec} = \Delta E_{elec} + \Delta G_{pol}$, [c]$\Delta E_{MM} = \Delta E_{vdW} + \Delta E_{elec}$, [d]$\Delta S$ = configuration entropy, [e]$\Delta G = E_{vdW} + \Delta E_{elec} + \Delta G_{np} + \Delta G_{pol}$, ΔG_{Bind} ^{sim} = $E_{vdW} + \Delta E_{elec} + \Delta G_{np} + \Delta G_{pol}$ - (TΔS), R1 and R5 are references from which data are taken

Since no approved drugs have been reported so far for the effective treatment of the SARS-CoV-2 virus, and with an urgent need for an epidemic situation, it has been a choice to use some of the available retroviral drugs. Hence in our study, the binding affinity of inhibitors I_{6Y2G} and I_{5R84} were further evaluated and compared with the FDA approved anti-HIV1 protease inhibitors, such as Lopinavir and Darunavir which has been reported as potent drugs against $3CL^{pro}$ of SARS-CoV-2. Recently the molecular recognition of the Lopinavir against COVID-19 $3CL^{pro}$ has been reported with the MMPBSA scheme (Wang 2020), where the estimated binding free energy of Lopinavir with $3CL^{pro}$ was found to be -6.63 kcal mol^{-1} and showed that the electrostatic interaction (-52.46 kcal mol^{-1}) favored most compared to the van der Waal interaction (-20.09 kcal mol^{-1}). Similarly, one more recent study using the MM-PBSA scheme suggests that the Darunavir can be a potent inhibitor against SARS-CoV-2 $3CL^{pro}$ (Sang et al. 2020).

Table 3: Decomposition of binding free energy into contributions from individual residues[a]

Residue	E_{vdW}	E_{elec}	G_{pol}	G_{np}	G_{side_chain}	$G_{backbone}$	G_{total}
$3CL^{pro}$/α-ketoamide							
Met165	-2.95	-1.82	1.30	-0.23	-2.03	-1.67	-3.70
Leu27	-2.05	-0.47	0.23	-0.18	-1.79	-0.68	-2.47
His164	-1.80	-4.15	3.70	-0.09	-0.31	-2.03	-2.34
His41	-2.04	-2.49	2.43	-0.22	-2.18	-0.14	-2.32
Glu166	-3.10	-4.14	5.99	-0.51	-0.37	-1.39	-1.76
Cys145	-1.86	-0.56	1.26	-0.16	-1.08	-0.24	-1.32
Pro168	-1.07	-0.02	0.13	-0.23	-0.79	-0.40	-1.19
His163	-0.54	-2.98	2.56	-0.06	-0.96	-0.06	-1.02
$3CL^{pro}$/Z31792168							
Met165	-2.11	-1.07	0.82	-0.20	-1.47	-1.09	-2.56
His163	-0.37	-2.86	1.74	-0.03	-1.43	-0.09	-1.52
His164	-0.95	-2.24	2.12	-0.08	-0.23	-0.92	-1.15

[a] Energetics contributions from the van der Waals (E_{vdW}) and electrostatic interactions (E_{elec}) as well as polar (G_{pol}) and nonpolar solvation energy (G_{np}) and the total contribution of given residue (G_{total}) for SARS-CoV-2-inhibitor complexes are listed. G_{side_chain} and $G_{backbone}$ represent the side chain and backbone contributions. Only residues with $|\Delta G| \geq 1.0$ kcal/mol are shown. All values are given in kcal/mol. The standard error of free energy estimated as well as of individual components are not larger than 0.4 kcal/mol.

It shows that for Darunavir, van der Waal interaction being higher (-41.32 kcal mol^{-1}) than electrostatic interaction (-5.80 kcal mol^{-1}). As shown in Table 2, if we compare the binding free energy among all inhibitors without considering the entropy part. Our study reports that the

binding affinity follows the order, α-ketoamide > Lopinavir > Darunavir > Z31792168 against COVID-19 $3CL^{pro}$. Hence, the inhibitor, α-ketoamide may be considered as a lead compound in the discovery of drugs against COVID-19.

Further to explore the key residues involved in the binding, ligand-residue free energy decomposition was performed using the MMGBSA scheme. Along with the electrostatic and van der Waal interactions, the solvation effect was taken into account using the Generalized Born model.(Onufriev, Bashford & Case 2004) The ligand-reside MM-GBSA interaction energy was computed in AMBER18. A hotspot residue is considered in the MM-GBSA interaction energy when it is higher than -1.0 $kcal.mol^{-1}$. The free energy decomposition of ligand-residue is shown in Table 3 and plotted in the Figure 7 . It is evident from Table 2 that a large number of hotspot residues contributing to the binding is seen in the $3CL^{pro}$/α-ketoamide complex, including the catalytic dyad C145 and H41 compared to the $3CL^{pro}$/Z31792168 complex, which also suggests the higher binding affinity for the inhibitor α-ketoamide against $3CL^{pro}$. Significant residues of the $3CL^{pro}$/α-ketoamide complex include M165, L27, H164, H41, E166, C145, P168, and H163 from the substrate-binding region of domains I and II, suggesting that these residues from $3CL^{pro}$ play a significant role in drug binding. Overall, these hotspot residues of the inhibitor-protein can facilitate the rational drug design of novel selective inhibitor against COVID-19.

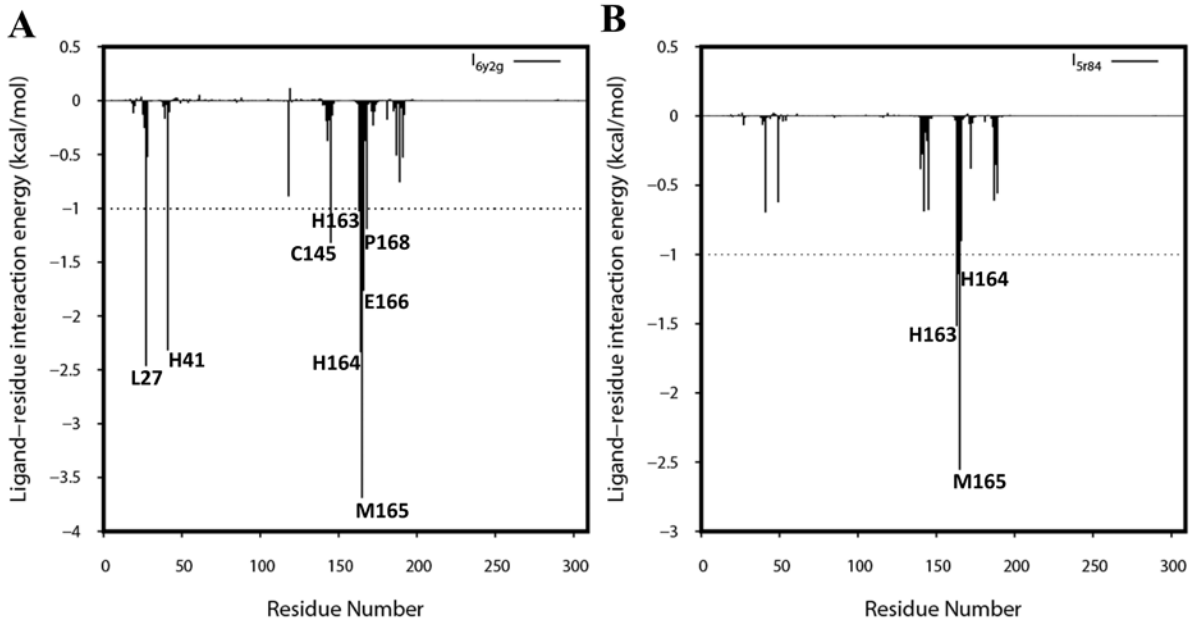

Figure 8: Decomposition of Δ*G* into contributions from individual residues for (A) $3CL^{pro}$/α-ketoamide and (B) $3CL^{pro}$/Z31792168.

Subsequently, to compliment the above results, we have performed the hydrogen bond (H-bond) analysis using the Cpptraj (Roe & Cheatham III 2013) module in AMBER for both the complexes on the MD trajectories, and their occupancies are reported in Table 4. It is observed from Table 4 that residues H41 (~ 43 %), E166 (~ 43 %), and H164 (~ 42 %) form > ~ 40 % H-bond with the inhibitor I_{6Y2G} during the simulations, explaining their high intermolecular electrostatic interactions in binding. However, in the case of the I_{5R84} complex, only E166 form the highest H-bond occupancy of ~ 35 %. Besides, the number of H-bond with the simulations time is shown in Figure S4, which also ensures that the highest number of H-bond persists for the $3CL^{pro}$/α-ketoamide complex compared to $3CL^{pro}$/Z31792168, suggesting the strong bonding of the inhibitor α-ketoamide with $3CL^{pro}$.

Table 4: The hydrogen bonds formed between SARS-CoV-2 and inhibitor and the corresponding average distance and percent determined using the MD trajectories in the MD simulations.

Binding couples		MD	
Acceptor	Donor	Avg. Distance (Å)[a]	Occupancy (%)[b]
$3CL^{pro}$/α-ketoamide			
Lig@O40	His41@NE2	2.88	43.12
Lig@O22	Glu166@N	2.90	42.53
His164@O	Lig@N38	2.88	41.76
Lig@O48	His163@NE2	2.83	29.97
Glu166@O	Lig@N23	2.87	24.43
Glu166@OE2	Lig@N49	2.84	23.06
Glu166@OE1	Lig@N49	2.84	13.17
Leu27@HD23	Lig@N36	2.82	10.33
Lue27@HD21	Lig@N36	2.83	10.25
Leu27@HD22	Lig@N36	2.84	09.78
Asn142@OD1	Lig@N36	2.82	09.56
Lig@O41	Gly143@N	2.84	07.98
Lig@H20	Gln189@N	2.85	05.10
$3CL^{pro}$/Z31792168			

Lig@O	Glu166@N	2.87	34.55
His164@O	Lig@N	2.84	25.34
Lig@N1	His163@NE2	2.91	09.19

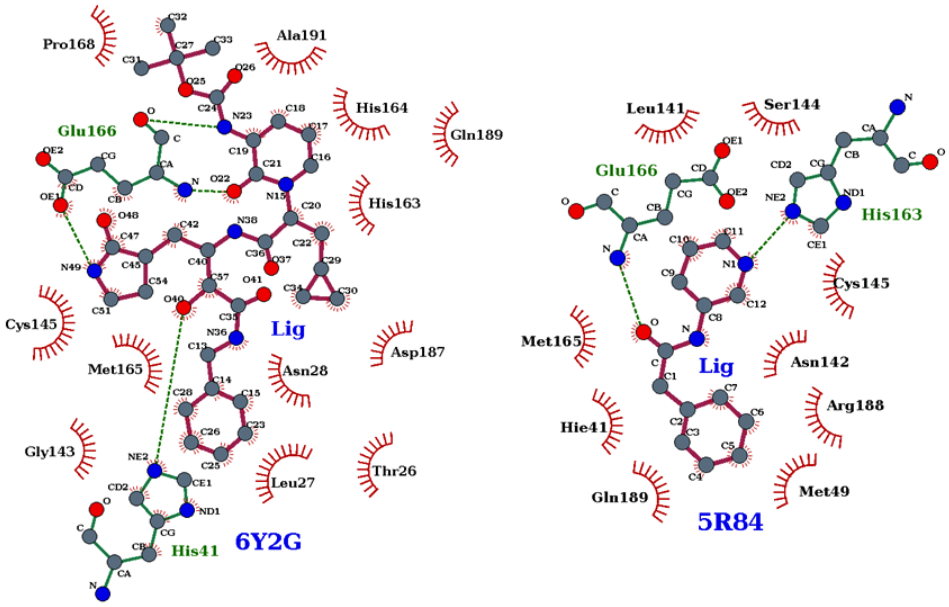

Figure 9: The receptor-ligand interaction profile for (left) α-ketoamide and (right) Z31792168. The plots are generated by using Ligplot+. Hydrogen bonds are shown as green dotted lines. The residues involved in the hydrophobic contacts are shown by red semicircles and residues involved in hydrogen bonds are represented in green. The inhibitors are represented as ball and sticks.

The detailed interactions profile of residues involving H-bond and hydrophobic interactions is also computed using the LigPlot (Wallace, Laskowski & Thornton 1995) program and shown in Figure 9. It suggests that the residue of the catalytic dyad H41 plays a significant role in forming a strong H-bond with inhibitors. Also, the hydrophobic contacts resulting from the hydrophobic residues for both the complexes throughout the simulations were computed and shown in Figure S5. It also suggests the highest number of hydrophobic contacts remain in the simulations for the inhibitor I_{6Y2G} complex in agreement with the above results of binding free energy. Overall, the binding free energy and its decomposition analysis reveal that I_{6Y2G} inhibitor is a more potent lead molecule against the novel CIVID-19 protease ($3CL^{pro}$) in

comparison with FDA approved anti-HIV1 protease inhibitor such as Lopinavir and Darunavir which is being recently focused in the treatment of COVID-19.

4. Conclusion

The recent outbreak of COVID-19 has caused a severe strain in the public health system in many countries. COVID-19 can cause mild to severe illness. The current situation demands an efficacious vaccine or novel antiviral drugs targeting COVID-19. Herein, we have studied the mechanism of binding of two potential inhibitors, namely α-ketoamide and Z31792168 to COVID-19 main protease ($3CL^{pro}$) by using an atomistic molecular dynamics simulation of 100 ns in conjunction with the widely used molecular mechanics/Poisson-Boltzmann surface area (MM/PBSA) scheme. Out study shows that the $3CL^{pro}$-inhibitor complexation is favored by the intermolecular van der Waals and electrostatic interactions as well as nonpolar solvation free energy. We have also shown that the inhibitor α-ketoamide is more potent compared to Z31792168 due to an increased favourable contribution from the intermolecular van der Waals and electrostatic interactions relative to Z31792168. Furthermore, in the case of α-ketoamide, the nonpolar component of the solvation free energy is also slightly more favourable compared to Z31792168. We have also identified the hotspot residues controlling the receptor-ligand binding. Finally, our study also reveals that α-ketoamide has better binding affinity compared to anti-HIV retroviral drugs, such as darunavir and lopinavir. Overall, α-ketoamide can be used as a lead compound in the development of drug targeting SARS-CoV-2 $3CL^{pro}$ and our study might play an useful role for the same.

Acknowledgments

This work was partially supported by the Department of Biotechnology, Govt. of India (grant number BT/RLF/Re-entry/40/2014, DBT-Ramalingaswami Re-entry Fellowship), and Department of Science and Technology (DST), Govt. of India (grant number ECR/2017/000010). R. R. acknowledges the Indian Institute of Technology Indore for financial assistance

Disclosure Statement

No potential conflict of interest was reported by the authors.

References:

Berendsen, H. J., Postma, J. v., van Gunsteren, W. F., DiNola, A. & Haak, J. (1984). Molecular dynamics with coupling to an external bath. *The Journal of chemical physics 81*: 3684-3690.

Berman, H. M., Battistuz, T., Bhat, T. N., Bluhm, W. F., Bourne, P. E., Burkhardt, K., Feng, Z., Gilliland, G. L., Iype, L. & Jain, S. (2002). The protein data bank. *Acta Crystallographica Section D: Biological Crystallography 58*: 899-907.

Case, D. A., Ben-Shalom, I.Y., Brozell, S.R., Cerutti, D.S. , Cheatham, T.E., III, Cruzeiro, V.W.D., Darden, T.A., Duke, R.E., Ghoreishi, D., Gilson, M.K., Gohlke,H., Goetz, A.W., Greene, D., Harris, R., Homeyer, N., Izadi, S., Kovalenko, A., Kurtzman, T., Lee, T.S., LeGrand, S., Li, P., Lin, C., Liu, J., Luchko, T., Luo, R., Mermelstein, D.J., Merz, K.M., Miao, Y., Monard, G., Nguyen, C., Nguyen, H., Omelyan, I., Onufriev, A., Pan, F., Qi, R., Roe, D.R., Roitberg, A., Sagui, C., Schott-Verdugo, S., Shen, J., Simmerling, C.L., Smith, J., Salomon-Ferrer, R., Swails, J., Walker, R.C., Wang, J., Wei, H., Wolf, R.M., Wu, X., Xiao, L., York, D.M., and Kollman, P.A. (2018). AMBER 2018. *University of California, San Francisco*.

Chang, C.-k., Hou, M.-H., Chang, C.-F., Hsiao, C.-D. & Huang, T.-h. (2014). The SARS coronavirus nucleocapsid protein–forms and functions. *Antiviral Research 103*: 39-50.

Darden, T., York, D. & Pedersen, L. (1993). Particle mesh Ewald: An N· log (N) method for Ewald sums in large systems. *Journal of Chemical Physics 98*: 10089-10092.

Fiser, A. & Šali, A. (2003). Modeller: generation and refinement of homology-based protein structure models. *Methods in Enzymology*, Elsevier. *374***:** 461-491.

Frauenfelder, H., Sligar, S. G. & Wolynes, P. G. (1991). The energy landscapes and motions of proteins. *Science 254*: 1598-1603.

Hartigan, J. A. & Wong, M. A. (1979). Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics) 28*: 100-108.

Hess, B., Kutzner, C., Van Der Spoel, D. & Lindahl, E. (2008). GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *Journal of Chemical Theory and Computation 4*: 435-447.

Hui, D. S., I Azhar, E., Madani, T. A., Ntoumi, F., Kock, R., Dar, O., Ippolito, G., Mchugh, T. D., Memish, Z. A. & Drosten, C. (2020). The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health—The latest 2019 novel coronavirus outbreak in Wuhan, China. *International Journal of Infectious Diseases 91*: 264-266.

Humphrey, W., Dalke, A. & Schulten, K. (1996). VMD: visual molecular dynamics. *Journal of molecular graphics 14*: 33-38.

Hünenberger, P., Mark, A. & Van Gunsteren, W. (1995). Fluctuation and cross-correlation analysis of protein motions observed in nanosecond molecular dynamics simulations. *Journal of Molecular Biology 252*: 492-503.

Ichiye, T. & Karplus, M. (1991). Collective motions in proteins: a covariance analysis of atomic fluctuations in molecular dynamics and normal mode simulations. *Proteins: Structure, Function, and Bioinformatics 11*: 205-217.

Jakalian, A., Jack, D. B. & Bayly, C. I. (2002). Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J Comput Chem 23*: 1623-1641.

Jin, Z., Du, X., Xu, Y., Deng, Y., Liu, M., Zhao, Y., Zhang, B., Li, X., Zhang, L. & Peng, C. (2020). Structure of Mpro from COVID-19 virus and discovery of its inhibitors. *bioRxiv*.

Jonniya, N. A. & Kar, P. (2019). Investigating specificity of the anti-hypertensive inhibitor WNK463 against With-No-Lysine kinase family isoforms via multiscale simulations.

Journal of Biomolecular Structure and Dynamics: DOI:10.1080/07391102.07392019.01602079.

Jonniya, N. A., Sk, M. F. & Kar, P. (2019). Investigating Phosphorylation-Induced Conformational Changes in WNK1 Kinase by Molecular Dynamics Simulations. *ACS Omega 4*: 17404-17416.

Kar, P., Seel, M., Hansmann, U. H. & Höfinger, S. (2007). Dispersion Terms and Analysis of Size-and Charge Dependence in an Enhanced Poisson− Boltzmann Approach. *The Journal of Physical Chemistry B 111*: 8910-8918.

Kar, P., Wei, Y., Hansmann, U. H. & Höfinger, S. (2007). Systematic study of the boundary composition in Poisson Boltzmann calculations. *J Comput Chem 28*: 2538-2544.

Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L., Lee, M., Lee, T., Duan, Y. & Wang, W. (2000). Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Accounts of Chemical Research 33*: 889-897.

Kräutler, V., Van Gunsteren, W. F. & Hünenberger, P. H. (2001). A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. *Journal of Computational Chemistry 22*: 501-508.

Lin, S., Shen, R., He, J., Li, X. & Guo, X. (2020). Molecular Modeling Evaluation of the Binding Effect of Ritonavir, Lopinavir and Darunavir to Severe Acute Respiratory Syndrome Coronavirus 2 Proteases. *bioRxiv*.

Loncharich, R. J., Brooks, B. R. & Pastor, R. W. (1992). Langevin dynamics of peptides: The frictional dependence of isomerization rates of N-acetylalanyl-N′-methylamide. *Biopolymers: Original Research on Biomolecules 32*: 523-535.

Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E. & Simmerling, C. (2015). ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *Journal of Chemical Theory and Computation 11*: 3696-3713.

McCammon, J. (1984). Protein dynamics. *Reports on Progress in Physics 47*: 1.

Olsson, M. H., Søndergaard, C. R., Rostkowski, M. & Jensen, J. H. (2011). PROPKA3: consistent treatment of internal and surface residues in empirical p K a predictions. *Journal of Chemical Theory and Computation 7*: 525-537.

Onufriev, A., Bashford, D. & Case, D. A. (2004). Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins: Structure, Function, Bioinformatics 55*: 383-394.

Onufriev, A., Bashford, D. & Case, D. A. (2004). Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins: Structure, Function, and Bioinformatics 55*: 383-394.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of Computational Chemistry 25*: 1605-1612.

Price, D. J. & Brooks III, C. L. (2004). A modified TIP3P water potential for simulation with Ewald summation. *The Journal of chemical physics 121*: 10096-10103.

Roe, D. R. & Cheatham III, T. E. (2013). PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *Journal of Chemical Theory and Computation 9*: 3084-3095.

Rota, P. A., Oberste, M. S., Monroe, S. S., Nix, W. A., Campagnoli, R., Icenogle, J. P., Penaranda, S., Bankamp, B., Maher, K. & Chen, M.-h. (2003). Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science 300*: 1394-1399.

Roy, R., Ghosh, B. & Kar, P. (2020). Investigating Conformational Dynamics of Lewis Y Oligosaccharides and Elucidating Blood Group Dependency of Cholera Using Molecular Dynamics. *ACS Omega*.

Salomon-Ferrer, R., Case, D. A. & Walker, R. C. (2013). An overview of the Amber biomolecular simulation package. *Wiley Interdisciplinary Reviews: Computational Molecular Science 3*: 198-210.

Sang, P., Tian, S.-H., Meng, Z.-H. & Yang, L.-Q. (2020). Insight derived from molecular docking and molecular dynamics simulations into the binding interactions between HIV-1 protease inhibitor s and SARS-CoV-2 3CLpro.

Sk, M. F., Roy, R. & Kar, P. (2020). Exploring the potency of currently used drugs against HIV-1 protease of subtype D variant by using multiscale simulations. *Journal of Biomolecular Structure and Dynamics*.

Wallace, A. C., Laskowski, R. A. & Thornton, J. M. (1995). LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein engineering, design and selection 8*: 127-134.

Wan, Y., Shang, J., Graham, R., Baric, R. S. & Li, F. (2020). Receptor recognition by the novel coronavirus from wuhan: an analysis based on decade-long structural studies of SARS coronavirus. *Journal of Virology 94*.

Wang, J. (2020). Fast Identification of Possible Drug Treatment of Coronavirus Disease-19 (COVID-19) Through Computational Drug Repurposing Study.

Wang, J., Wang, W., Kollman, P. A. & Case, D. A. (2001). Antechamber: an accessory software package for molecular mechanical calculations. *J. Am. Chem. Soc 222*: U403.

Wang, J., Wang, W., Kollman, P. A. & Case, D. A. (2006). Automatic atom type and bond type perception in molecular mechanical calculations. *Journal of Molecular Graphics Modelling 25*: 247-260.

Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. (2004). Development and testing of a general amber force field. *J Comput Chem 25*: 1157-1174.

Webb, B. & Sali, A. (2014). Comparative protein structure modeling using MODELLER. *Current protocols in bioinformatics 47*: 5.6. 1-5.6. 32.

Woo, P. C., Huang, Y., Lau, S. K. & Yuen, K.-Y. (2010). Coronavirus genomics and bioinformatics analysis. *Viruses 2*: 1804-1820.

Wu, J. T., Leung, K. & Leung, G. M. (2020). Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *The Lancet 395*: 689-697.

Xu, B., Shen, H., Zhu, X. & Li, G. (2011). Fast and accurate computation schemes for evaluating vibrational entropy of proteins. *J Comput Chem 32*: 3188-3193.

Zhang, L., Lin, D., Sun, X., Curth, U., Drosten, C., Sauerhering, L., Becker, S., Rox, K. & Hilgenfeld, R. (2020). Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. *Science*.

Supporting Information

Elucidating Biophysical Basis of Binding of Inhibitors to SARS-Cov-2 Main Protease by Using Molecular Dynamics Simulations and Free Energy Calculations

Md Fulbabu Sk, Rajarshi Roy, Nisha Amarnath Jonniya, Sayan Poddar, Parimal Kar*

Discipline of Biosciences and Biomedical Engineering, Indian Institute of Technology Indore, Khandwa Road, MP 453552, India.

*Corresponding Author: Parimal Kar | Email: parimal@iiti.ac.in | Phone: +91 731 2438700 (Ext. 550)

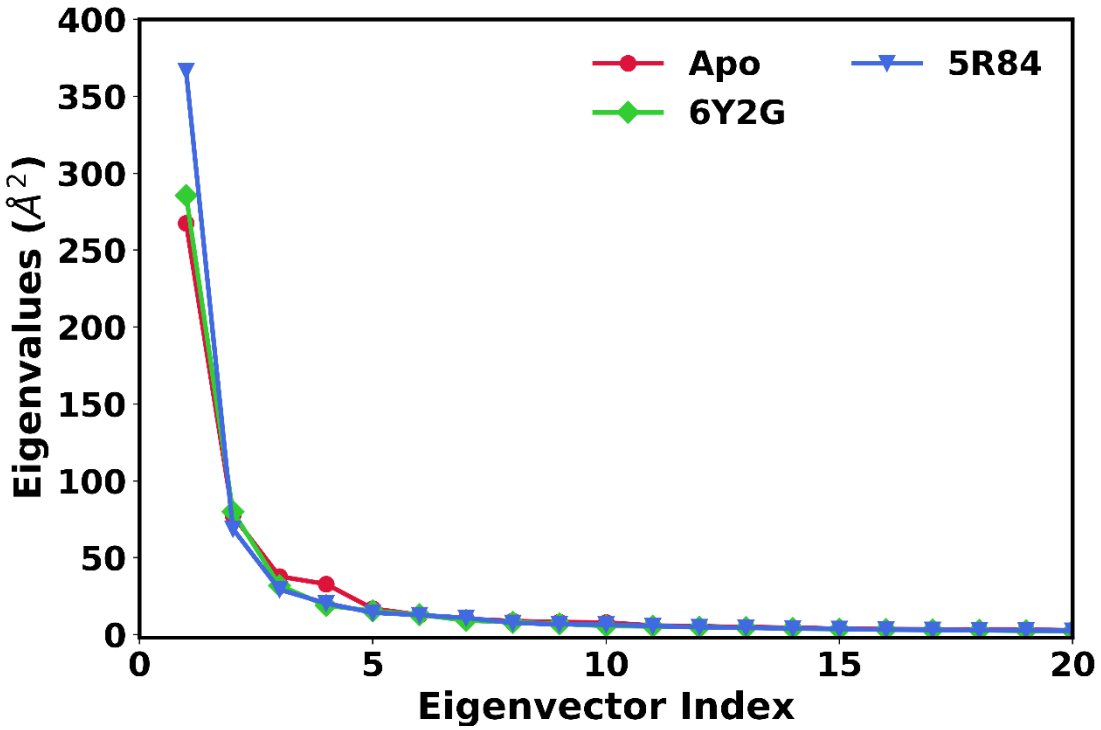

Figure S1: Comparison of the eigenvalues plotted against the corresponding eigenvector indices from the Cα covariance matrix constructed from the MD trajectories of apo and complexes.

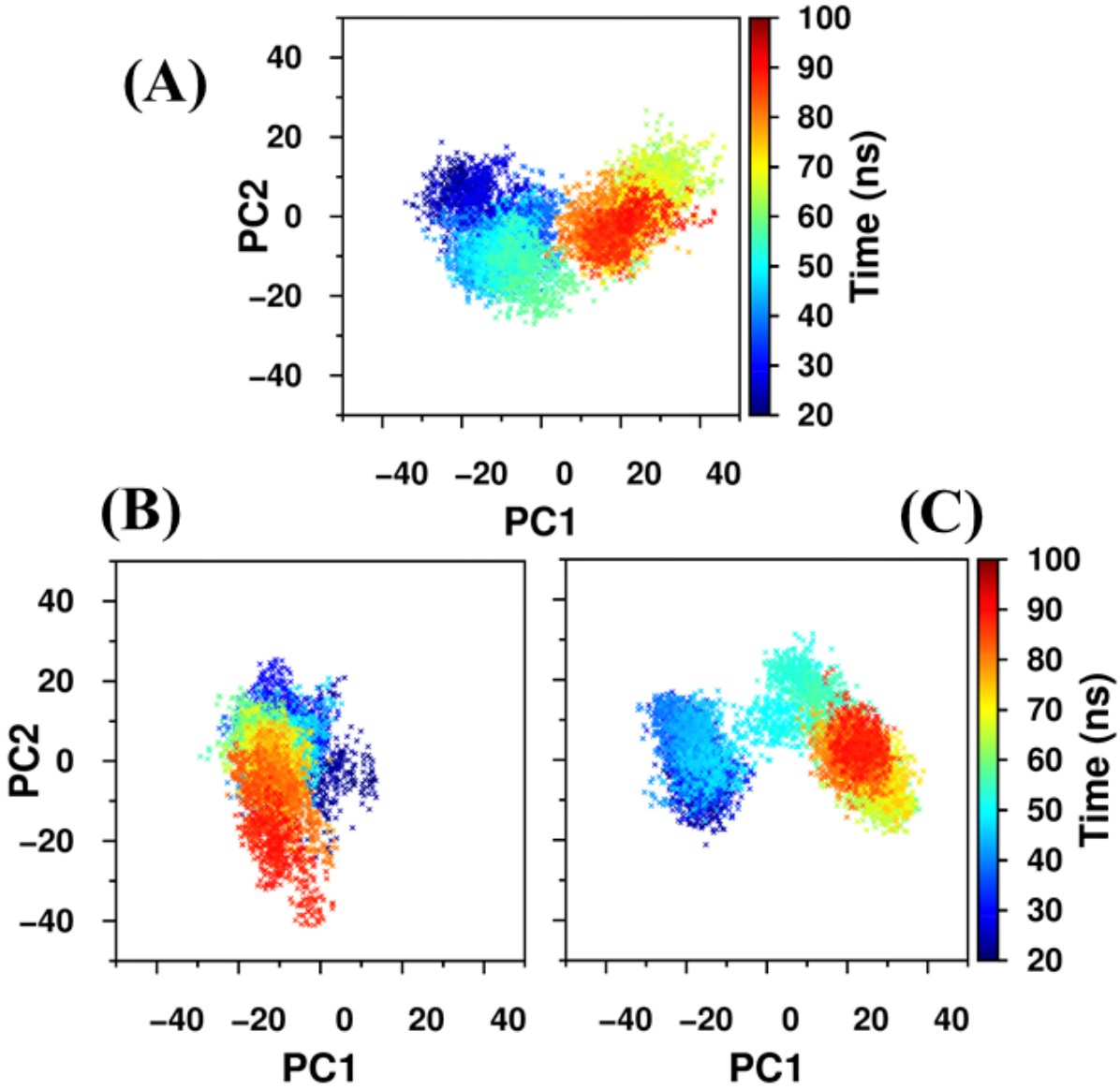

Figure S2: The maps of PC1 and PC2 during MD simulation of apo form of $3CL^{pro}$ and complex and colour codes represents the simulation time.

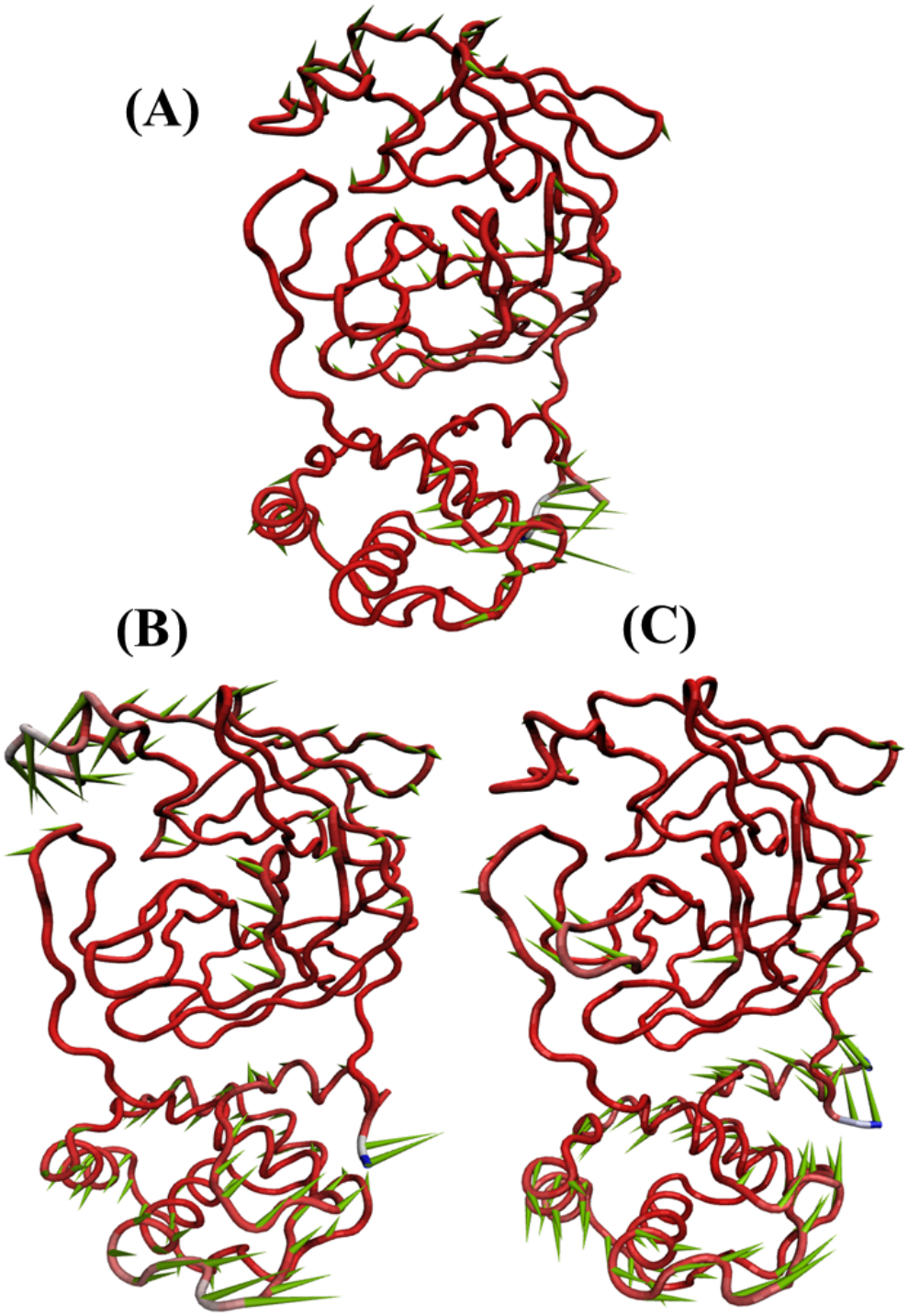

Figure S3: Porcupine plots showing prominent motions for (A) apo, (B) 6Y2G and (C) 5R84. Green represent eigenvector showing the direction of prominent movements. Length of the eigenvectors represents the magnitude of the movements.

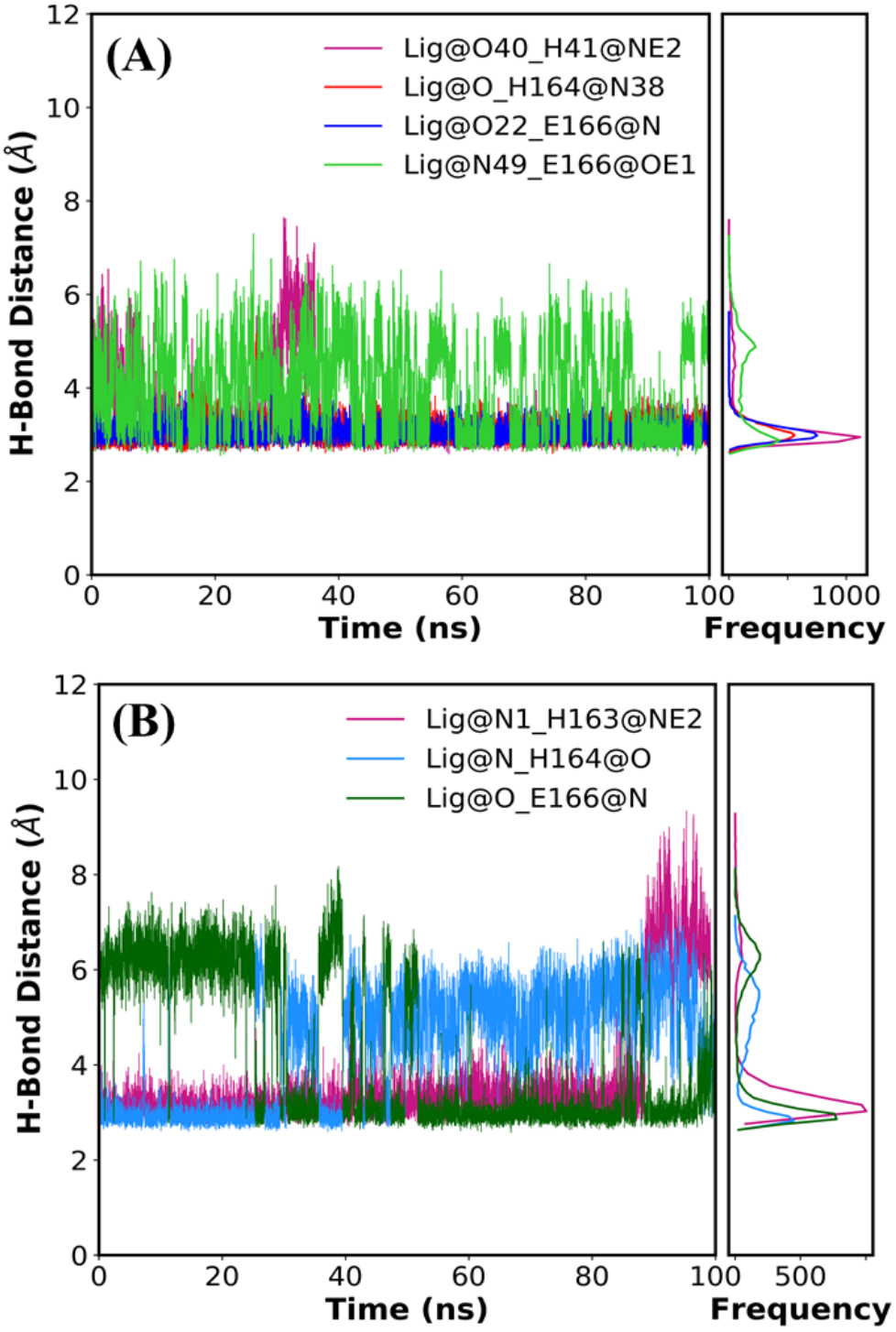

Figure S4: Time evolution of hydrogen bond distances and probability distribution of (A) 6Y2G and (B) 5R84 complexes.

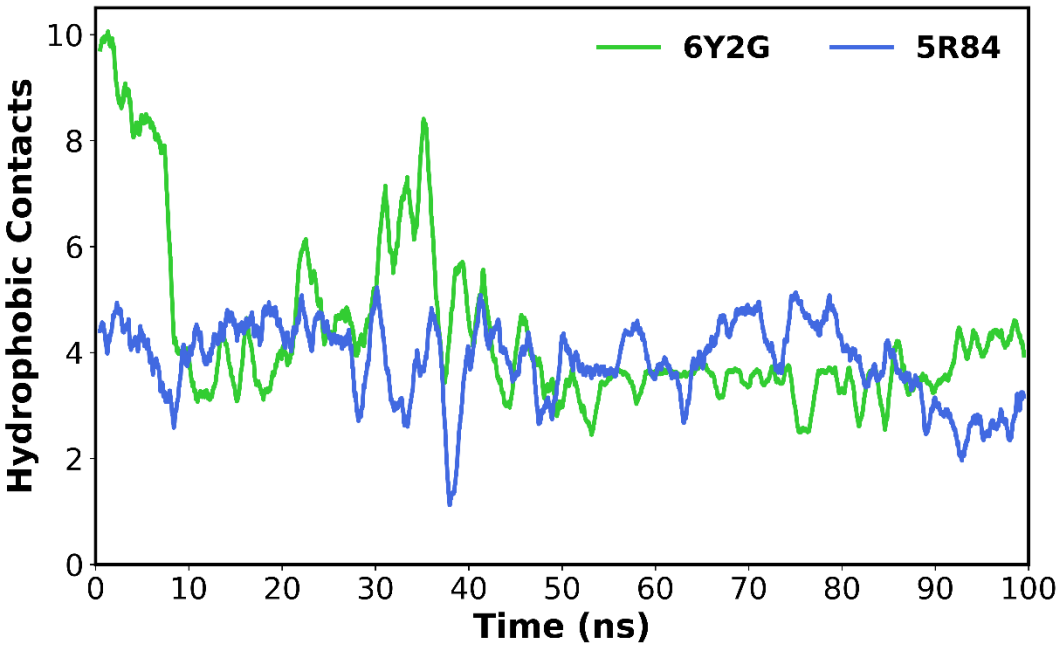

Figure S5: The time evolution of hydrophobic contacts of I_{6Y2G} and I_{5R84}.