RESEARCH ARTICLE

Mechanistic insights of key host proteins and potential repurposed inhibitors regulating SARS-CoV-2 pathway

Debabrata Pramanik[1] ⬤ | Aiswarya B. Pawar[2] | Sudip Roy[2] | Jayant Kumar Singh[1,2] ⬤

[1]Department of Chemical Engineering, Indian Institute of Technology Kanpur, Kanpur, India

[2]Prescience Insilico Private Limited, Bangalore, India

Correspondence
Sudip Roy, Prescience Insilico Private Limited, Old Madras Road, Bangalore 560049, India.
Email: sudip@prescience.in

Jayant Kumar Singh, Department of Chemical Engineering, Indian Institute of Technology Kanpur, Kanpur, India.
Email: jayantks@iitk.ac.in

Abstract

The emergence of pandemic situations originated from severe acute respiratory syndrome (SARS)-CoV-2 and its new variants created worldwide medical emergencies. Due to the non-availability of efficient drugs and vaccines at these emergency hours, repurposing existing drugs can effectively treat patients critically infected by SARS-CoV-2. Finding a suitable repurposing drug with inhibitory efficacy to a host-protein is challenging. A detailed mechanistic understanding of the kinetics, (dis)association pathways, key protein residues facilitating the entry–exit of the drugs with targets are fundamental in selecting these repurposed drugs. Keeping this target as the goal of the paper, the potential repurposing drugs, Nafamostat, Camostat, Silmitasertib, Valproic acid, and Zotatifin with host-proteins HDAC2, CSK22, eIF4E2 are studied to elucidate energetics, kinetics, and dissociation pathways. From an ensemble of independent simulations, we observed the presence of single or multiple dissociation pathways with varying host-proteins-drug systems and quantitatively estimated the probability of unbinding through these specific pathways. We also explored the crucial gateway residues facilitating these dissociation mechanisms. Interestingly, the residues we obtained for HDAC2 and CSK22 are also involved in the catalytic activity. Our results demonstrate how these potential drugs interact with the host machinery and the specific target residues, showing involvement in the mechanism. Most of these drugs are in the preclinical phase, and some are already being used to treat severe COVID-19 patients. Hence, the mechanistic insight presented in this study is envisaged to support further findings of clinical studies and eventually develop efficient inhibitors to treat SARS-CoV-2.

KEYWORDS
dissociation pathways, host proteins, inhibitors, metadynamics, molecular dynamics

1 | INTRODUCTION

Over the past two decades, the virus that causes coronavirus disease has turned out to be lethal across the world. Earlier it was considered as a mild disease infecting immunocompetent people. Over the years, the virus has evolved into being fatal, with a higher transmissibility rate. Severe acute respiratory syndrome (SARS-CoV-2),[1] the virus that causes COVID-19, has become the worst and most uncontrollable public health crisis in a century. After the first outbreak, the virus mutated into more infectious variants, triggering multiple pandemic waves. The second pandemic wave primarily targets the younger population aggressively. Unlike the first wave, 60%–70% of the infected people in the second wave required critical care and hospitalization. A statement released by the Centers for Disease Control and Prevention (CDC) in July 2021 claimed that the Delta variant has the highest spread cases and severe outcomes even among vaccinated people.

The Delta variant has shown to be twice as powerful as previous variants, but further details are still not understood. Generally, on-set of infection, a person can experience milder flu-like symptoms,[2,3] further developing into more acute severe respiratory failure[4] and other multi-organ complications. Many repurposing drugs have been applied for COVID-19 treatment in this emergency, although the FDA does not yet approve them.

Researchers are scrambling to unravel how the virus hijacks the host and eventually affects the host regulatory mechanism. The viral genome is complex and encodes structural and non-structural proteins.[5,6] The viral proteins utilize the host proteins from their entry to transmission within host cells for their life cycle.[7,8] Identifying virus-host protein interactors in the complex pool of sequential pathways is vital to study the viral pathogenesis and host tolerant environment for viruses. In particular, the Delta variant with multiple mutations in the spike protein binds to the host angiotensin-converting enzyme 2 (ACE2)[9] receptor, which is present outside the human throat and lung cells, 2–4 times stronger. Recent proteomics approaches have revealed that viral proteins can enhance or lower the activity of hundreds of human kinases involved in cellular activities and immune response.[10] So, the possibility of targeting pro-viral cellular signaling could make the host immune system resistant to viral replication. It is crucial when the host has antibodies developed for SARS-CoV-2 by vaccination or recovery from COVID-19 infection and is still susceptible to new variants.[11]

The Caesin Kinase 2 (CK2),[12,13] Histone deactylease (HDAC2),[14,15] and eIF4E2 host proteins are essential for the normal functioning of cells, involved in cell cycle checkpoints, metabolic pathways, epigenetic regulators, mRNA translation pathways. These proteins get perturbed in the cancerous cells. Previous experimental studies have highlighted that cancer cells and viral pathogens follow similar molecular mechanisms for apoptosis and evade host immune response. Coronaviruses take over these machinery pathways. For example, viral protein Nsp1 recruits the host proteins and reduces the overall host protein translation in the cell by 70%.[16]

Recent mass-spectrometric experiments reported that an entry of a virus to the host target induces a significant change in protein phosphorylation and disrupts the host cytoskeletal organization.[17] Additionally, pharmacological studies in human cell lines have indicated the drug Silmitasertib (currently in phase 2 trials) to have high inhibitory activity of CK2 kinases, suggesting the role of these kinases in the SARS-CoV-2 life cycle.[17] Previous studies[18–20] on SARS-CoV have shown that drugs Nafamostat[18,19] and Camostat[21] are potential candidates to stop viral entry by binding to transmembrane protease serine 2 (TMPRSS2) proteases.[22–24] However, even at higher concentrations, Camostat showed 65% inhibition of protease binding, and 35% virus entry takes place via the endosomal cathepsin pathway.[25] Although, the role of these drugs on intracellular proteins that help regulate viral protein processing is still not clear. Similarly, the Zotatifin drug has shown to be effective in the translation machinery of the host, where viruses rely on host-dependent mRNA translation. Repurposing of these drugs can be an effective way in terms of cost and time for treating COVID-19.

Our recent study[26] followed an exhaustive screening of FDA-approved drugs against multiple host targets by using high throughput computational screening[27] utilizing molecular docking, all-atom molecular dynamics simulations including solvent and state-of-the-art free energy calculations. The multi-target multi-ligand enhanced sampling approach provides highly selective off-target drug combinations with stable affinities. Interestingly, we found that drugs with high binding free energies were already in the clinical trial phase for COVID-19. Especially the drug Silmitasertib targeting CK2, which showed good binding free energies in our previous work,[26] is already in phase II clinical trials. Similarly, the drug Zotatifin targeting the eIF4E elongation factor is in clinical trial phase I. We found that some off-targets showed good binding free energies but have not been previously explored experimentally.[28] However, thermodynamics and kinetics are essential in the early stages of rational drug design, but molecular determinants could describe the modulation of binding and unbinding kinetics and disclosing their pathways.[29] In Figure 1, we have schematically presented the background of this study and outlined the key goals of this work. In this paper, we have considered the targets and off-targets combinations[30] having good binding free energies and subjected to performing a combined all-atom molecular dynamics and enhanced sampling simulations. We explore the dissociation pathways of the drugs from the host proteins by considering the essential interacting residues involved in (un)binding processes and quantitatively estimated the probabilities of dissociation through different pathways, and showed dominant unbinding pathways. The information obtained from this study will further help in enhancing our knowledge on drug target mechanisms and designing better therapeutics.

1.1 | Computational methodologies

We applied a rigid docking protocol to generate the docked complex using the Autodock suite (Autodock 4.2.6).[31] We performed docking on the crystal structures of the proteins (Table 1), and missing loops, if present, have been modeled using Modeller (9.24).[32] We first obtained minimum energy structures for proteins and ligands (Figure 2) separately and then performed docking to get the docked complex. The docking sites have been selected by either using the inhibitor-bound position in the crystal structure or finding the binding sites using the Autoligand tool.[33] We took the conformation corresponding to the minimum docking score for all-atom molecular dynamics (MD) using GROMACS-2018.[34] Charmm27 force field parameters with cmap corrections have been used to model bonded, nonbonded interactions for proteins.[35–37] To generate ligand interaction parameters, we optimized the ligands' geometry employing GAUSSIAN 09[36] using a semi-empirical method at the PM6 level, followed by density functional theory (DFT) optimization using M06 functionals and 6-311g (d,p) basis sets. To account bulk solvent effect, we employed the PCM method. The partial charges have been calculated using the CHELPG method[38] on this DFT-optimized structure. Following this, we used the swissparam web service[37] to obtain charmm compatible parameters. We first solvated the docked
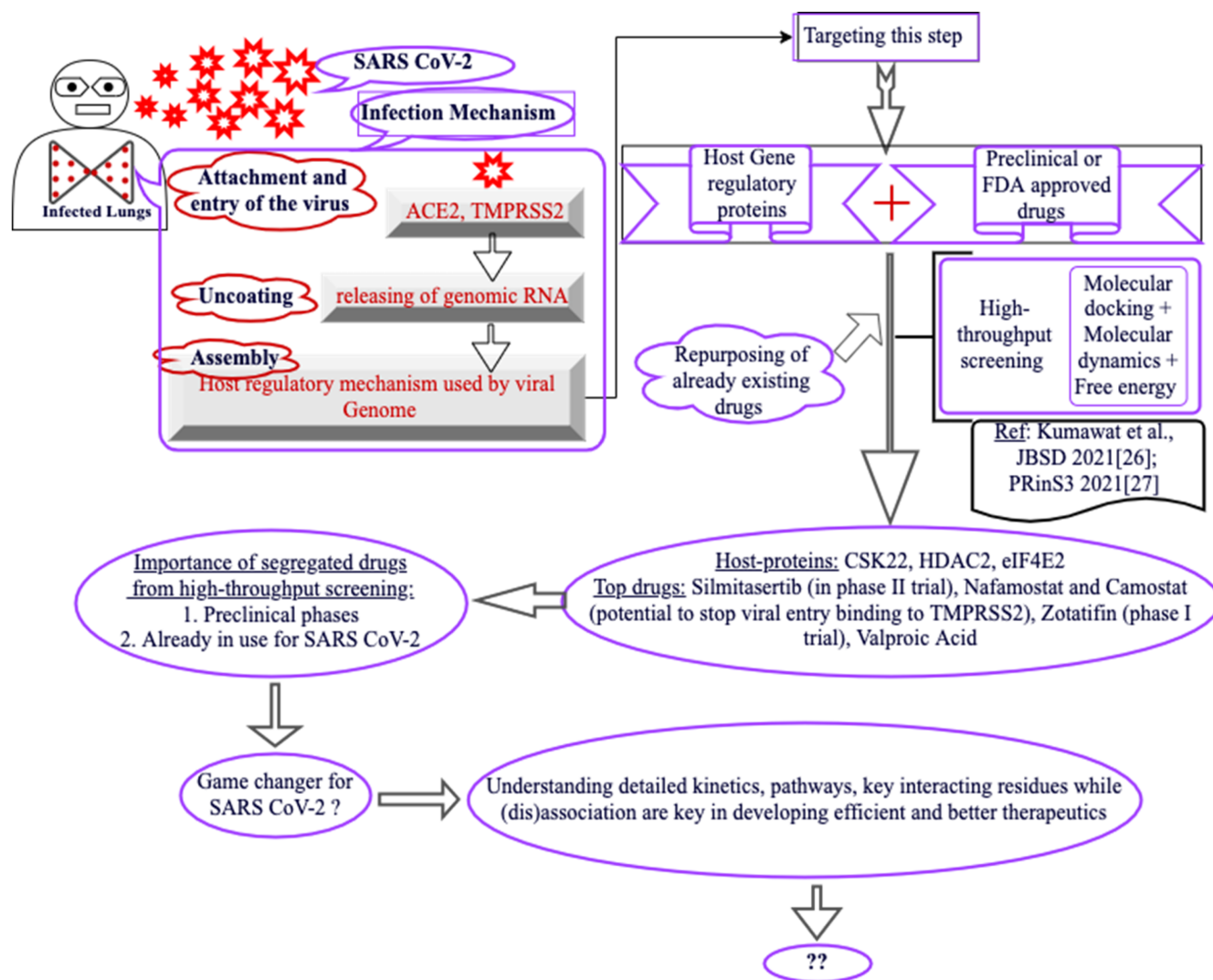
FIGURE 1 The background and the goal of the work.

TABLE 1 The host target protein and drug combinations studied in this paper, and their role of specific pathways involved in host cells.

Target	Pathway	Structure (PDB ID)	Drug
eIF4E2	Translation	2JGB	Nafamostat
eIF4E2	Translation	2JGB	Camostat
eIF4E2	Translation	2JGB	Zotatifin
HDAC2	Epigenetic factor	4LY1	Nafamostat
HDAC2	Epigenetic factor	4LY1	Valproic acid
CSK22	Transcription	6HMB	Silmitasertib

complex in TIP3P water and then added Na+ and Cl− ions to neutralize the system. Periodic boundary conditions were applied in all directions to mimic the bulk property. After setting up the initial system, we performed energy minimization using the steepest descent method to remove the overlap of the solute with the water molecules. It was followed by NVT equilibration for duration of 500 ps at 300 K using a modified Berendsen thermostat.[39] Later, a 1 ns NPT equilibration simulation was followed after NVT simulation at 1 atm using Parrinello-Rahman[40] barostat. The Particle-Mesh Ewald (PME)[41] method was employed to calculate the long-range electrostatic interaction and a cut-off value of 1 nm for the short-range interaction. We used the LINCS algorithm[42] and a 2 fs step to integrate the equation of motion. The production run was executed in the NPT ensemble for 100 ns. We analyzed these trajectories and calculated the bound state properties of the protein-ligand systems. We calculated the contact frequency between the protein and ligand to extract the crucial residues involved in the interaction at the bound state, considering the minimum distance criterion choosing a cut-off of 0.4 nm. Further, we pulled out the interaction fingerprint for the selective protein residues (i.e., that remain in contact for more than 75% of simulation time) obtained from the contact frequency.[43]

To explore the ligand dissociation[29,44–47] from the protein binding pocket, we employed metadynamics (metaD)[48] and its variant well-tempered metadynamics (wt-metaD)[49] using PLUMED 2.5.4[50] and GROMACS 2018.[34] In a metaD simulation, a history-dependent
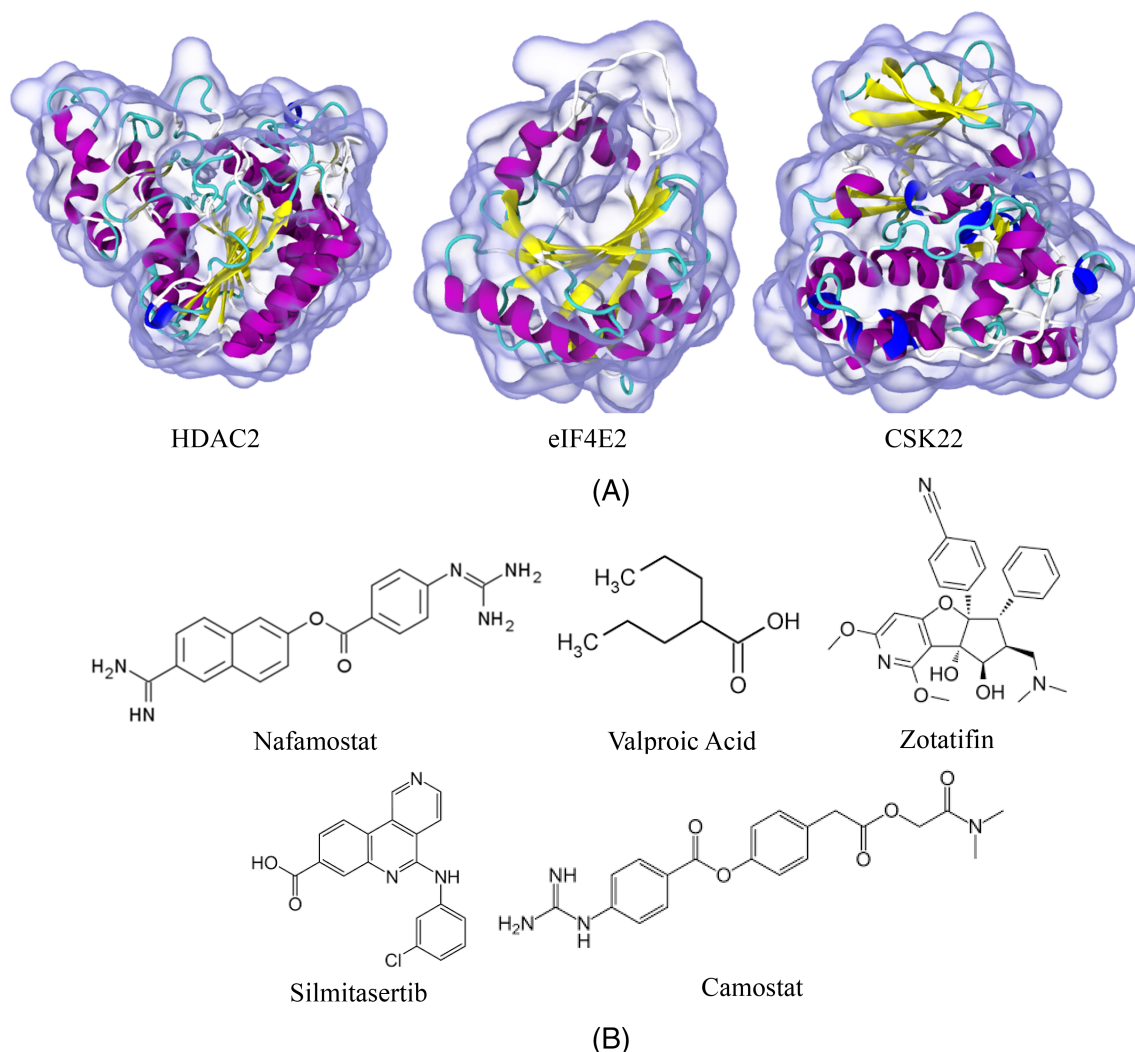
FIGURE 2 (A) Instantaneous snapshots of the secondary structures of HDAC2, eIF4E2 and CSK22 protein targets. (B) 2D chemical structures of Nafamostat, Valproic acid, Zotatifin, Silmitasertib, and Camostat

TABLE 2 Average free energy barriers, residence times and dissociation rates are shown for six target–drug combinations.

Structure (PDB ID)	Ligand	Free energy (kcal/mol)	Residence time (min)	k_{off} (1/min)	log(k_{off})
HDAC2 (4LY1)	Nafamostat	−12.3 (0.44)	1.4E+08	7.1E−09	−8.15
HDAC2 (4LY1)	Valproic acid	−12.6 (0.33)	1.5E+11	6.7E−12	−11.17
eIF4E2 (2JGB)	Nafamostat	−4.1 (0.36)	8.3E+06	1.2E−07	−6.92
eIF4E2 (2JGB)	Camostat	−3.2 (0.29)	6.3E+06	1.6E−07	−6.8
eIF4E2 (2JGB)	Zotatifin	−2.9 (0.26)	3.7E+06	2.7E−07	−6.57
CSK22 (6HMB)	Silmitasertib	−11.5 (0.33)	4.3E+12	2.3E−13	−12.64

bias is deposited sequentially, prohibiting the system from revisiting already visited portions in the phase space. Thus, the system eventually escapes from the energy landscape's minimum to the aqueous medium. In wt-metaD, the height of the deposited bias is gradually decreased to have better and quicker convergences of the system. The final equilibrium coordinates from the MD simulation were taken as the initial structures in wt-metaD simulations. In a metaD or wt-

metaD simulation, the bias is added to the system along a reaction coordinate. A good reaction coordinate will have the following two criteria: (i) it should show a timescale separation between the slowest motion in the system with that of the fastest motion, and (ii) it should be able to demarcate all the relevant metastable states. Here we take distance (distance between the center of masses of the ligand heavy atoms and the center of masses of the protein-binding pocket) as the

reaction coordinate (RC) (see Figure S1). Once the simulation converges, the unbiased Boltzmann distribution of the RC is extracted by adding the deposited hills. The following biasing parameters were used in wt-metaD simulations: bias deposition frequency of 500 steps (1 ps), Gaussian hill height 1.5–2.0 kJ/mol, hill width 0.02–0.1 nm, bias factor 10–15, and temperature 300 K. For each system we performed several sets of independent wt-metaD simulations (20 simulations each for systems having single dissociation pathways, and 35 simulations each for those having multiple dissociation pathways) starting from a new initial configuration. Once the simulation converges, we extract the independent free energy adding up the deposited hills using the sum_hills plugin available in PLUMED. Once we have the free energy, we convert it to probability. Then we average over probabilities and later re-calculate the averaged free energy from averaged

probability ($F = -kT \ln P$) (F—free energy, k—Boltzmann constant, T—temperature, P—probability).[26,51]

To estimate the dissociation rate, we first calculate the residence time over independent trajectories for a system. To calculate the unbiased acceleration time (t_{acc}) from a biased simulation, we use the following relation: $t_{acc} = \int_0^t e^{\beta V(t)}(\delta t)$. Here, $\beta = \frac{1}{kT}$, and k is Boltzmann constant, T is temperature, $V(t)$ is the deposited bias as a function of time. δt is the unbiased integration time step or a multiple of δt.

Further to explore the crucial protein residues while dissociating the ligand from the binding pocket, we calculated the minimum distance between the heavy atoms of the ligand and protein residues and repeated this exercise for all the independent runs, as shown in Figure S3. Following this, we extracted only those specific residues of the protein for which, at least for a run, it has a minimum ligand and protein residue distance <0.2 nm (shown in Figure S4). Then we calculated the normalized probabilities for these specific residues for the distance between the residues and the ligand heavy atoms and repeated this exercise for all six systems as shown in Figures S5–S10.
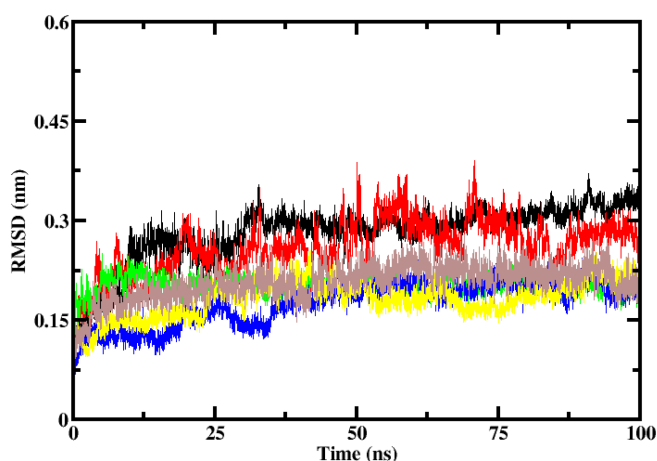

FIGURE 3 Root mean square deviation of target eIF4E2 complexed to different drugs Nafamostat (black), Camostat (red), Zotatifin (green), target HDAC2 complexed to Nafamostat (blue), Valproic acid (yellow), and target CSK22 Silmitasertib (brown).

2 | RESULTS AND DISCUSSION

Silmitasertib, Zotatifin showed high binding free energies with CK2 and eIF4E proteins in clinical trials for COVID-19 (NCT04668209, NCT04632381). In addition, we identified drugs that potentially interact with one or more off-targets with higher affinities. On-set of infection, the proteins associated with transcription, translation mechanisms are expressed. So, it is advantageous to use drugs that can target multiple host proteins at different stages of the mechanism. Such off-target combinations, eIF4E2 with Nafamostat, Camostat, and HDAC2 with Valproic acid, Nafamostat having good binding affinities, are listed in Table 2.
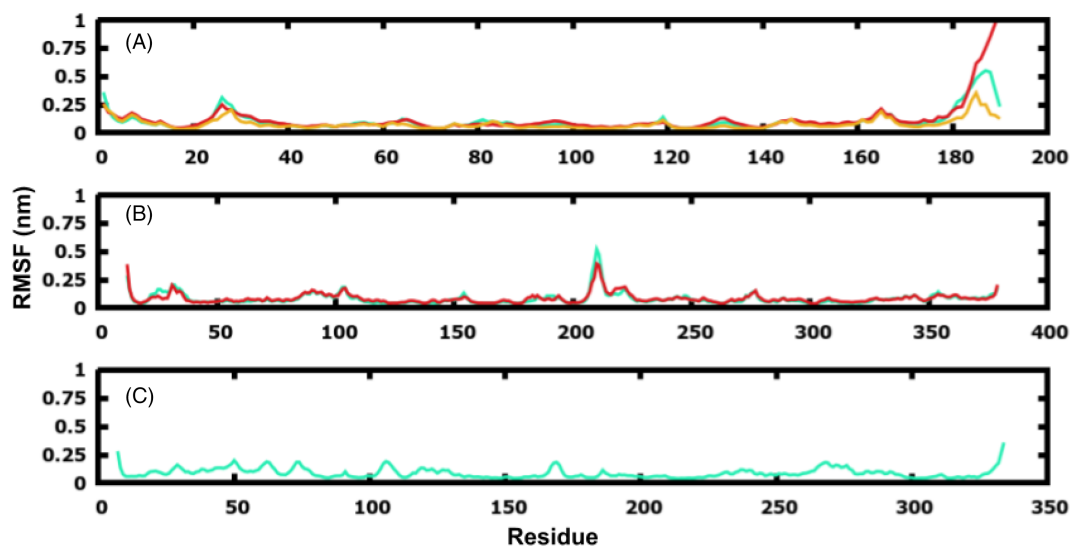

FIGURE 4 Root mean square fluctuation of protein backbone (A) eIF4E complexed to drug Nafamostat (cyan), Camostat (red), Zotatifin (orange). (B) HDAC2 complexed to Nafamostat (cyan), Valproic acid (red) (C) CSK22 complexed to Silmitasertib (cyan).

2.1 | Conformational stability of the host target–drug complexes

We performed atomistic molecular dynamics simulations on six drug-protein combinations (Table 1). We calculated the backbone root mean square deviation (RMSD) and root mean square fluctuation (RMSF) to evaluate the conformational stability. Figures 3 and 4 show the RMSD and RMSF of different targets bound to Nafamostat, Camostat, Zotatifin, Silmitasertib, and Valproic acid. In all cases, the RMSD was less than 0.45 nm suggesting the conformation over 100 ns has been less deviated and stable. The eIF4E2-Zotatifin and CSK2-Silmatersitib complexes showed more stable conformations ranging between 0.2–0.25 nm.

Additionally, the HDAC2 protein complex with Nafamostat and Valproic acid displayed a stable conformation with a value less than 0.15 nm for the first 50 ns. The values deviated after 50 ns, but still, conformations remained within the stable range of 0.2–0.25 nm. The eIF4E2-Camostat complexes showed the highest deviations for the rest of the complexes. It shows the highest variation at 50 and 75 ns. To understand protein regions that contribute to the conformational changes or stability, we calculated the RMSF at the residue level. The mobility of residues was observed for all six drug-protein combinations, as shown in Figure 4. The high fluctuations observed on the terminal ends of proteins can be neglected due to their high mobility. Overall, fewer residue-level fluctuations were observed for all six candidates, suggesting stable nature. Although high fluctuations were observed in residue PHE210 in both Nafamostat and Valproic acid, this particular residue was conserved in HDAC2 for both these drugs. This conserved PHE210 was present at the gateway in the binding pocket of HDAC2 and selectively interacted with ligands.
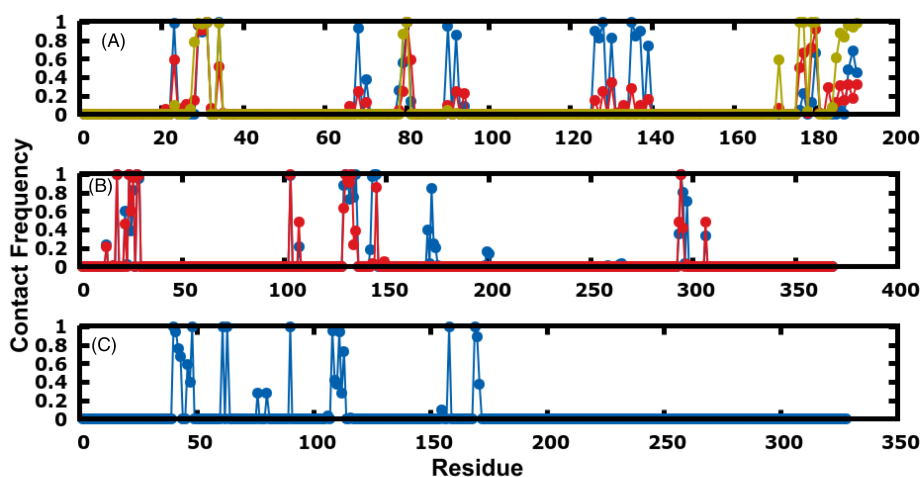

FIGURE 5 Contact frequency of the drug bound to target protein was calculated for (A) eIF4E–Nafamostat (blue), Camostat (red) and Zotatifin (orange). (B) HDAC2–Nafamostat (blue), Valproic acid (red) (C) CSK22–Silmitasertib (blue).
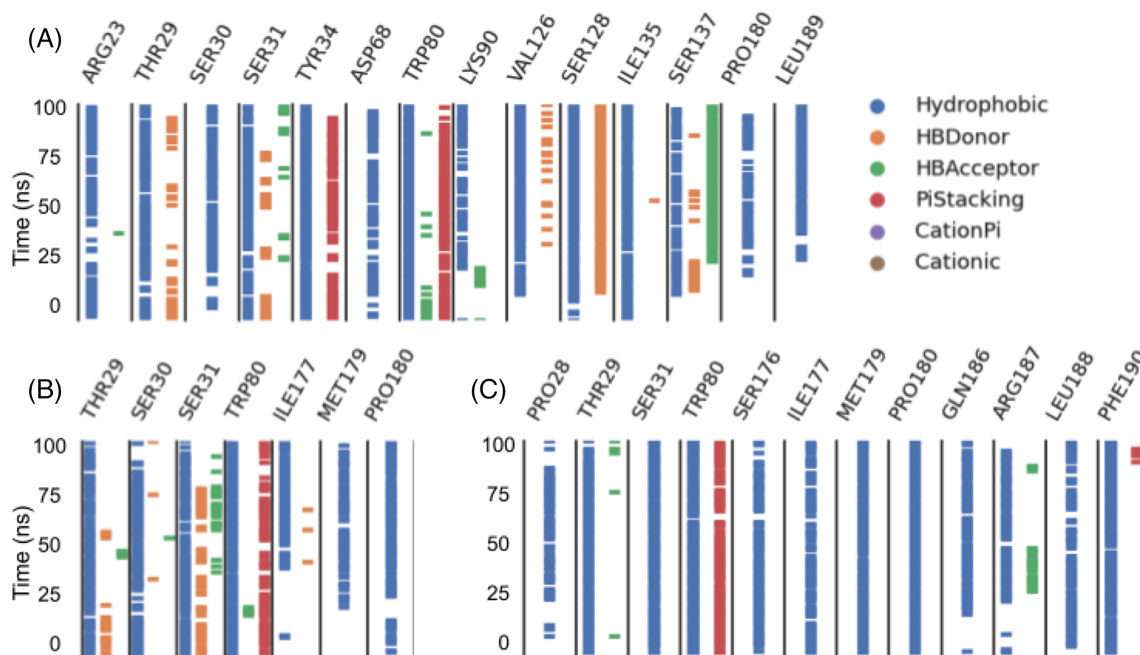

FIGURE 6 Interaction fingerprint for (A) eIF4E–Nafamostat. (B) eIF4E–Camostat, (C) eIF4E–Zotatifin

2.2 | Preferential binding sites for drug–host target proteins

To further understand the molecular details of the protein-drug interactions in a bound state, we calculated the minimum distance between them, as shown in Figure 5. It was interesting to see crucial residues involved in the interactions that remain in contact for more than 75% of the simulation time (shown in Table S1). In the eIF4E translation factor, the residues THR29, SER30, SER31, TRP80 interact with Nafamostat, Camostat, and Zotatifin. However, ILE177, MET179, and PRO180 are present in Camostat and Zotatifin and not in Nafamostat. Similarly, for the HDAC2 epigenetic factor, the residues that interact with Nafamostat and Valproic acid were quite similar. The residues interacting are TYR29, MET35, PRO37, ARG39, ILE40, PHE114, ALA141, GLY142, LEU144, CYS156, GLY305. Interestingly, HIS183 was shown to interact with Nafamostat and not Valproic acid. For CK2-Silmatersitib, the residues LEU46, GLY47, ARG48, GLY49, VAL54, VAL60, LYS69, ILE96, PHE114, ILE117, ASN119, MET164, ILE175, ASP176 interact in the bound state.

We calculated interaction fingerprints in a binary vector to further understand the nature of intermolecular interactions between the target proteins and drugs that remain in contact for more than 75% of simulation time. To understand the interaction aspects, we considered hydrogen-bonding, hydrophobic interactions, $\pi-\pi$ stacking (see Figure S2), and the structural conformations for the protein-drug systems. Figure 6 shows the interaction fingerprints for the eIF4E translational factor. THR29, SER31, TRP80, and PRO180 residues were observed in all three drug complexes involved in hydrophobic interactions. The TRP80 forms $\pi-\pi$ stacking interaction. In Nafamostat and Camostat, SER31 was shown to form transient hydrogen bonding. Interestingly, ILE177 and MET179 are involved in hydrophobic interactions in Camostat transiently and Zotatifin

continuously. Similarly, in Figure 7A,B, the HDAC2 residues TYR29, MET35, ARG39, ILE40, ALA141, GLY143, LEU144, CYS156, GLY305 form hydrophobic interactions with the drug Nafamostat and Valproic acid. There was a transient hydrogen bonding interaction with residues ARG39 in both and TYR308 in Nafamostat and GLY294 in Camostat. We observed the $\pi-\pi$ interactions only in Nafamostat at residues HIS134, HIS135. In CK2-Silmitasertib, the hydrophobic interactions were observed with LEU46, GLY47, ARG48, GLY49, VAL54, LYS69, IL196, PHE114, ILE117, ASN119, MET164, ILE175, ASP176. At the beginning of the simulation, we observed transient hydrogen bonds in LYS69 and ASP176. These hydrogen bonds eventually
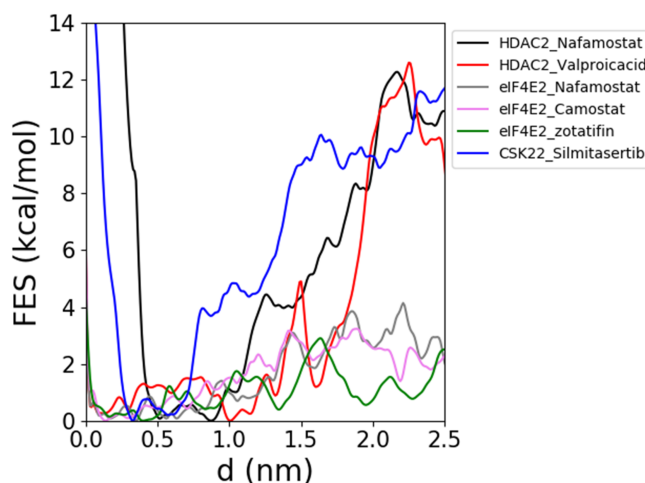

FIGURE 8 'FES' versus 'distance' for six target-ligand combinations. Averaged FES profiles have been shown averaging over 35 independent simulations for HDAC2–Nafamostat, HDAC2–Valproicacid, eIF4E2–Nafamostat, eIF4E2–Camostat and over 20 independent simulations for eIF4E2–Zotatifin, CSK22–Silmitasertib.
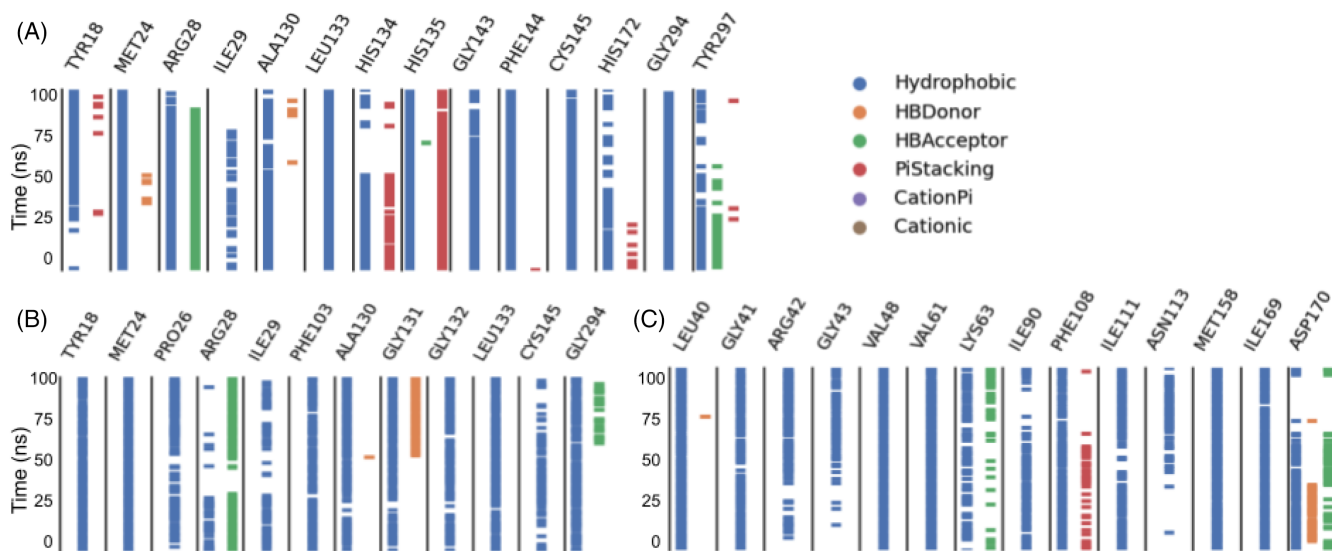

FIGURE 7 Interaction fingerprint for (A) HDAC2–Nafamostat. (B) HDAC2–Valproic acid. (C) CSK22–Silmitasertib

disappear at the end of the simulations. We noticed π–π interactions formed by PHE114 at the start of the simulations.

2.3 | Correlation between thermodynamics and kinetics of drug–target combinations

Further, to correlate the host target–drug complexes' thermodynamic behavior and kinetics, we calculated free energy barriers and dissociation rates, as shown in Table 2. Figure 8 shows the averaged free energy profiles for all six systems. Free energy values have been tabulated in Table 2. Figure 9 shows the 'log(k_{off})' versus "free energy barrier (ΔG)" plot for all the drug–target combinations. For better efficacies of drug molecules, the free energy barriers are suggested to be high, and the rate of dissociation of the drugs is proposed to be low and vice versa. We observe that for systems with higher free energy barriers (low values of ΔG), log(k_{off}) values are lesser, and for low free energy barriers (higher values of ΔG), the log(k_{off}) is higher. We observe such a scenario, especially in the off-target combinations. HDAC2-Nafamostat showed a low dissociation rate, and eIF4E2-Nafamostat displayed high dissociation rates. Specifically, for eIF4E2 combinations, the dissociation rates are higher, suggesting the drugs are bound less tightly and have lesser affinities. The HDAC2 and CSK22 combinations showed lower ΔG values, indicating more stable and energetically favorable at the bound state. As a result, it will take longer to dissociate from the binding pocket to the solvent crossing the transition barrier.

To check the experimental efficacy of these drugs, we compared the dissociation free energy with the reported activity (IC50 values). The IC50 value measures the amount of a drug required for 50% inhibition of its target. For Silmitasertib–CK2 inhibitor, we observe an excellent correlation between the obtained free energy value (−11.5 [0.33] kcal/mol) and the reported activity[52] IC50 value 1 nM. For Zotatifin–eIF4E2, reported[53] IC50 value 2 nM and in simulation we obtained free energy of −2.9 (0.26) kcal/mol. For Valproic Acid and its target inhibitor HDAC2, we observe a low free energy (−12.6 [0.33] kcal/mol), whereas its reported activity[54] is 1.02–2.15 mM. For Nafamostat, we studied two off-targets, HDAC2 and eIF4E2, for which we get barriers of −12.3 (0.44) kcal/mol and −4.1 (0.36) kcal/mol, respectively. The reported activity[25] for Nafamostat with its on-target serine protease one inhibitor (TMPRSS2) is 55 (7) nM. Hence, for Nafamostat, we also observe a good agreement for the selected off-target combinations. For Camostat–eIF4E2 (off-target) we got free energy of −3.2 (0.29) kcal/mol and activity[25] of 142 (31) nM. Hence, comparing free energy from simulations with the activity (IC50) of the reported drug, we observe an overall excellent correlation. Clearly, Silmitasertib, Zotatifin, and Nafamostat show high efficacy as
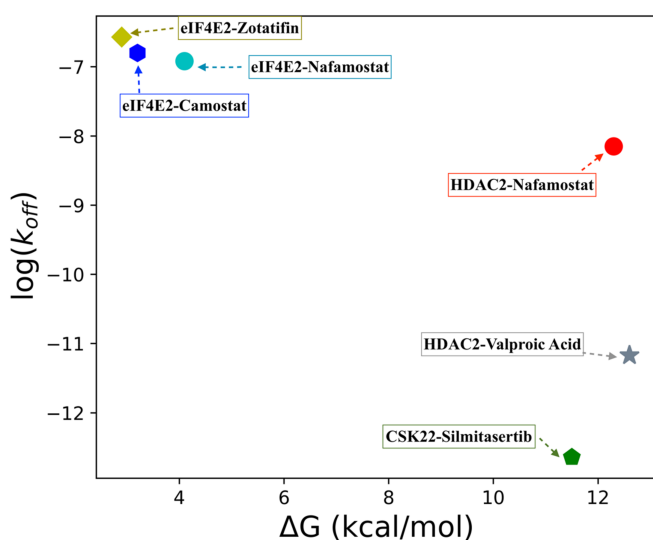

FIGURE 9 'Log(k_{off})' versus 'free energy barrier (ΔG)' has been plotted here for various target–ligand combinations as shown by separate colors and symbols. Different drugs have been shown in separate symbols

Target	PDB ID	Drug	Interacting residues
eIF4E2	2JGB	Nafamostat	TYR34, TRP80, LYS90, ARG94, SER128
eIF4E2	2JGB	Camostat	THR29, SER30, MET79, TRP80, GLU81, ARG130, MET179
eIF4E2	2JGB	Zotatifin	SER30, SER31, TRP80, SER176, PRO180, ARG187
HDAC2	4LY1	Nafamostat	ARG39, GLY142, LEU144, HIS146, GLY154, PHE155, CYS156, HIS183, GLY306
HDAC2	4LY1	Valproic Acid	ARG39, ILE40, GLY142, LEU144, HIS146, PHE155, LEU276
CSK22	6HMB	Silmitasertib	GLY47, ARG48, GLY49, SER52, VAL54, LYS69, GLU82, LEU86, ILE96, LYS159, MET164, ASP176, SER195

TABLE 3 List of target protein residues interacting with the ligand from wt-metaD simulations from all independent runs for each system.

Note: The target residues having probability of $> = 80\%$ and with the distance <0.3 nm have been chosen. The table shows interaction sites and corresponding protein residues of the target for HDAC2–Nafamostat, eIF4E2–Nafamostat, eIF4E2–Camostat, CSK22–Silmitasertib, eIF4E2–Zotatifin and HDAC2–Valproic acid for dissociation pathways.

FIGURE 10 Interacting residues having probability more than 80% and distance <0.3 nm for six systems. Orange colors show ligands, yellow residues with probability 1 and lime colors are with probability >0.8 and distance <0.3 nm. (A) HDAC2–Nafamostat, (B) eIF4E2–Nafamostat, (C) eIF4E2–Camostat, (D) CSK22–Silmitasertib, (E) eIF4E2–Zotatifin, (F) HDAC2–Valproic acid
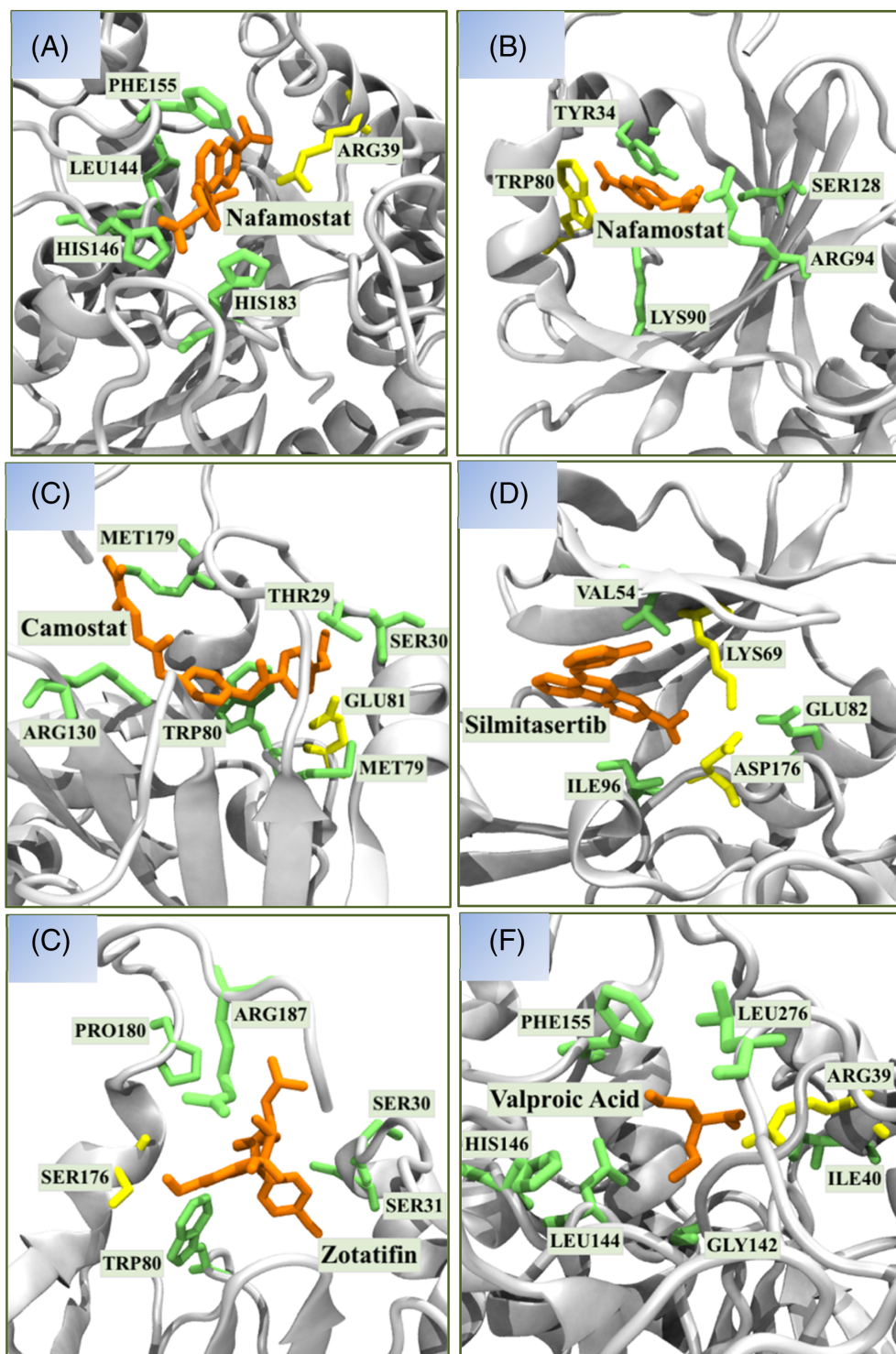
inhibitors. It is noteworthy that the drug Silmitasertib is already in the trial phase for SARS-CoV-2. Although we see good agreement for all the drugs, for Valproic Acid, we see its activity to be very high, contrary to the simulated free energy. We speculate that since Valproic Acid is a small molecule, it might interact more strongly with other targets in the cell assay than HDAC2, hence might be the possible reason behind its higher value of activity (∼ in the mM range).

2.4 | Pathways and mechanism of drug dissociations from the binding pocket of the host proteins

To explore the complete dissociation pathways from its bound pose and uncover the essential interacting residues that work as gate-keepers facilitating the (dis)association mechanism, we analyzed the crucial interacting residues from the dissociation trajectories for all six
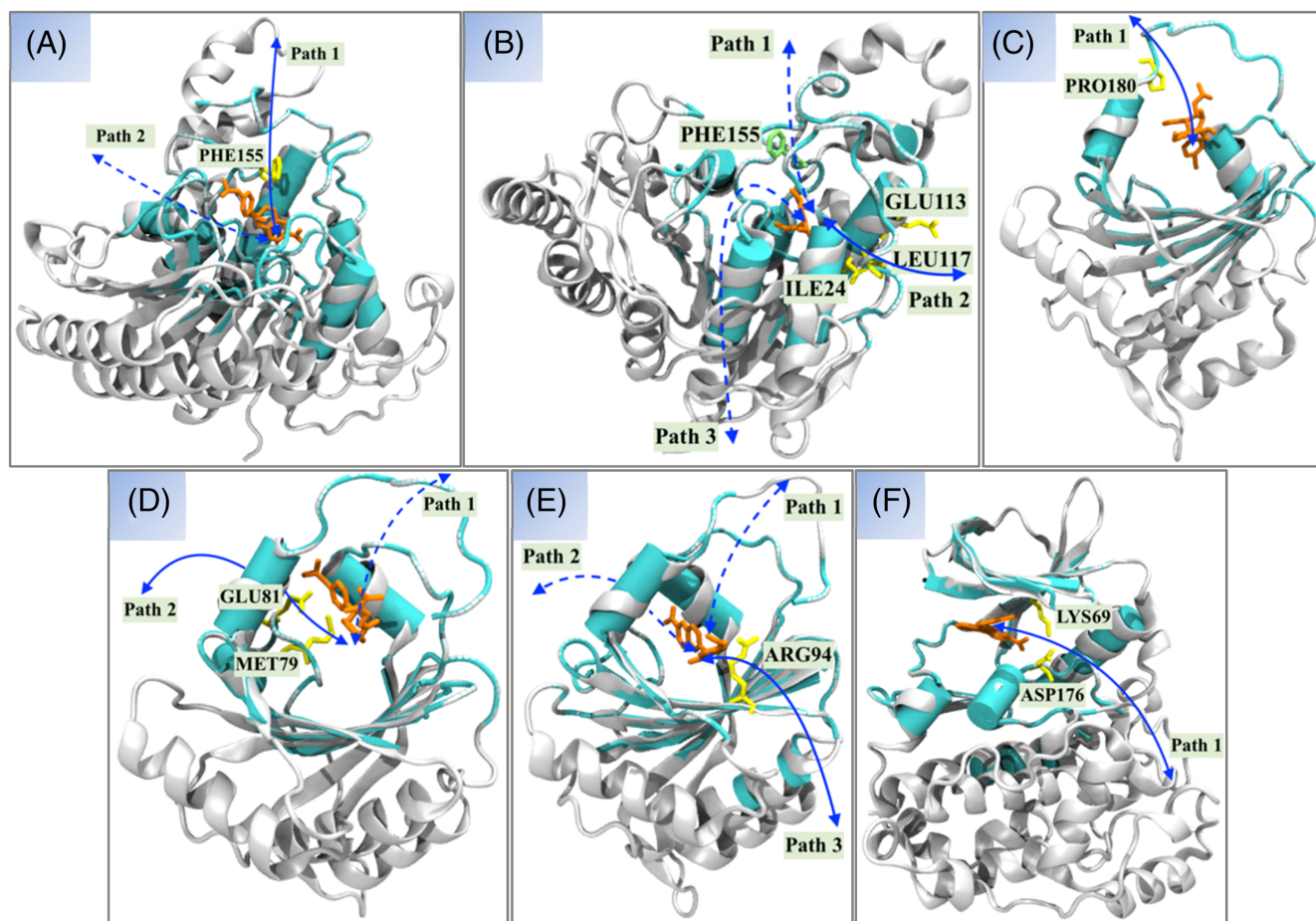
FIGURE 11 Schematic representation of the ligand dissociation pathways for (A) HDAC2–Nafamostat, (B) HDAC2–Valproic acid, (C) eIF4E2–Zotatifin, (D) eIF4E2–Camostat, (E) eIF4E2–Nafamostat, (F) CSK22–Silmitasertib. Ligands have been shown in orange color. Solid lines show most probable pathway and other dashed lines show lesser probable pathways. Different pathways have been demonstrated by notation "path 1", "path 2" and "path 3." Key gateway residues of the protein have been shown in yellow colors corresponding to the main dominating pathway only

combinations. To select the important one from the list of all the residues, we chose a minimum distance cut-off of <0.3 nm and a normalized probability of 0.8 (see Figures S5–S10). In Table 3, we showed the list of dominant residues for HDAC2–Nafamostat, eIF4E2–Nafamostat, eIF4E2–Camostat, CK2–Silmitasertib, eIF4E2–Zotatifin, and HDAC2–Valproic Acid while dissociating from the bound state to the solvent environment. We observed ARG39 (HDAC2–Nafamostat), TRP80 (eIF4E2–Nafamostat), GLU81 (eIF4E2–Camostat), LYS69 and ASP176 (CK2–Silmitasertib), and ARG39 (HDAC2–Valproic Acid) with the highest probability and with a distance cut-off of <0.2 nm, whereas the residue SER176 (for eIF4E2-Zotatifin) with probability 0.9, distance cut-off <0.3 nm. In Figure 10 (a-f), we showed instantaneous snapshots of the host protein-ligand binding pocket highlighting the key residues (ligand (orange), residues with probability 1 (yellow), other key residues (lime)). In the case of both HDAC2 combinations, for Nafamostat and Valproic Acid, common residues ARG39, GLY142, LEU144, HIS146, PHE155 were shown to be interacting. Interestingly for eIF4E2 combinations, TRP80 was critical in interactions with all three drugs (Nafamostat, Camostat, and Zotatifin).

Understanding the complete dissociation pathways for a ligand from its bound state and knowledge about the central residues facilitating the (dis)association mechanism is vital in accurate, fast, and effective drug development purposes. To have a rigorous understanding, we analyzed the wt-metaD trajectories of the dissociation events for all the systems. We explored all possible pathways for a ligand to dissociate from the bound state. In Figure 11 (A–F) we have schematically shown all the dissociation pathways for (A) HDAC2–Nafamostat, (B) HDAC2–Valproic acid, (C) eIF4E2–Zotatifin, (D) eIF4E2–Camostat, (E) eIF4E2–Nafamostat and (F) CSK22–Silmitasertib. Solid lines show the most probable pathway, while other dashed lines show lesser probable pathways. Different pathways have been demonstrated by the notation "Path 1", "Path 2", and "Path 3." Key gateway residues of the protein have been shown in yellow colors corresponding to the main dominating pathway only. Figure 12A–F explicitly demonstrates instantaneous snapshots for all six systems and possible pathways. For HDAC2–Nafamostat, we observe two dissociation pathways as shown in Figure 11A and Figure 12A (i and ii). Along 'path 1' (shown in Figures 11A and 12A (i), PHE155 works as the gatekeeper while for
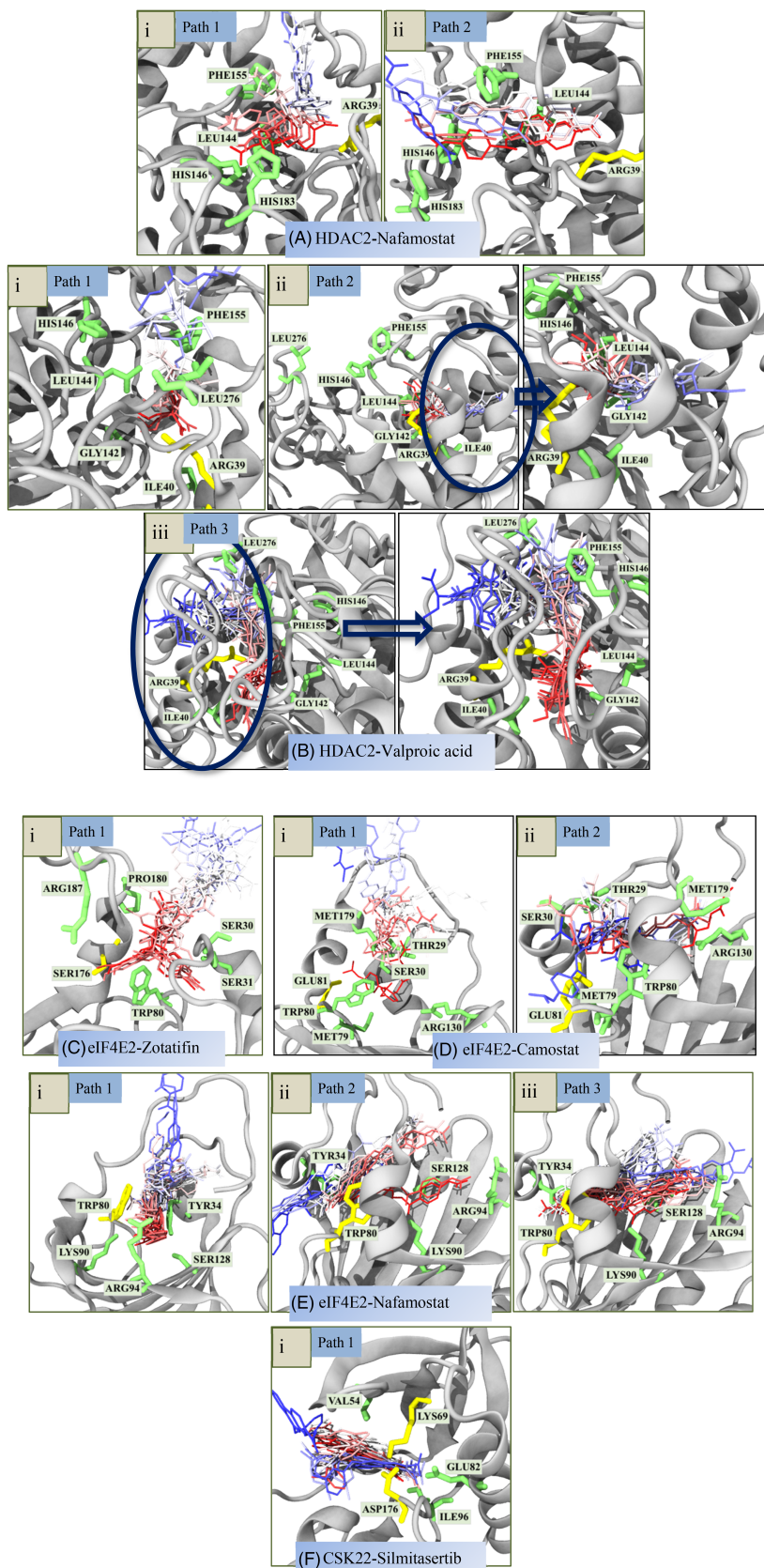
FIGURE 12 Explicit demonstration of the ligand dissociation pathways for (A) HDAC2–Nafamostat (i and ii), (B) HDAC2–Valproic acid (i, ii, and iii), (C) eIF4E2–Zotatifin (I), (D) eIF4E2–Camostat (i and ii), (E) eIF4E2–Nafamostat (i, ii and iii), (F) CSK22–Silmitasertib (i). The time evolution of the dissociation events for the ligands have been shown in Timestep representation with licorice drawing method and the colors denote: red (bound state) → pink → white → light blue → blue (unbound state).

'path 2' (Figure 12A (ii) HIS183. Out of 35 independent dissociation events, we observed 60% dissociations along the pathway shown in Figure 11 (a) 'path 1', whereas 40% along 'path 2' shown in Figure 12A (ii). Hence, the dissociation pathway along the PHE155 residue is the most probable pathway facilitating the (dis)association mechanism for the HDAC2-Nafamostat system. For HDAC2-Valproic acid (Figures 11B and 12B, the ligand is deeply embedded inside the binding cavity surrounded by ARG39, ILE40, GLY142, LEU144. While dissociating, it reaches the top of the surface and dissociates interacting through PHE155 (Figure 12B i), 'path 1'. It also follows two more distinct paths between the helices, one interacting with ILE24, LEU117, GLU113 (Figures 11B and 12B (ii), 'path 2', and in another, 'path 3', the ligand reaches the top between LEU276 and PHE155. Then it dissociates interacting with TYR308, ILE310, PHE155, PRO34 (Figure 12B (iii). Thus, in 35 independent dissociation events, we observe three distinct pathways for the HDAC2-Valproic acid system. Out of these three pathways, we noticed the path shown in (Figures 11B and 12B (ii) in solid line 'path 2' to be statistically most probable with 52% compared to Figure 12B (i) 'path 1' with 37% and Figure 12B (iii) 'path 3' with 11%. Hence, for HDAC2, we observe two different dominant pathways to be present for Nafamostat and Valproic acid. For eIF4E2–Zotatifin, out of 20 independent runs, we observe a single pathway (Figures 11C and 12C (i) with PRO180 as the gatekeeper residue. For eIF4E2–Camostat and eIF4E2–Nafamostat, we performed 35 independent runs in both systems. For eIF4E2–Camostat, MET179, THR29, SER30 are at the gateway (pathway Figures 11D and 12D (i) with 40%, 'path 1' and GLU81, MET79 at the gateway (for pathway Figures 11D and 12D (ii) with 60%, 'path 2'. For eIF4E2–Nafamostat (Figures 11E and 12E), TRP80, TYR34 residues (pathway Figure 12E (i) 'path 1' with 14%, between helices through TRP80, TYR34 residues (pathway Figure 12E ii) with 9% 'path 2', residue ARG94 along with the beta-sheet (pathway Figures 11E and 12E (iii) with 77%, 'path 3' shown in solid line in Figure 11E. Hence, with target eIF4E2, we observe different dissociation pathways as dominant for respective ligands. For zotatifin we see 'path 1' (Figure 11C) as the dominant pathway, while for Camostat 'path 2' (Figure 11D) and for Nafamostat 'path 3' (Figure 11E) are more specific to dissociation. For CSK22–Silmitasertib (Figure 11F and 12F (i), out of 20 independent dissociation events, we observe a single dissociation pathway having LYS69 and ASP176 residues to be crucial for dissociation at the gateway. Thus, we explored various dissociation pathways for three targets with different important target inhibitors and showed dominant dissociation pathways with specific interacting residues facilitating the dissociation events. It might be useful in developing potential target-specific inhibitors with pharmaceutical importance.

3 | CONCLUSION

Designing new therapeutic approaches or repurposing existing drugs requires understanding viral pathogenesis, host targets and pathways that mediate the infection. Various interactome studies have highlighted detailed information on crucial host regulatory proteins for which repurposed drugs are shown to be effective. It is in line with our previous study,[26] where we have done an exhaustive screening of existing drugs and human targets using molecular docking, molecular dynamics simulations, and free energy calculations. From this study,[26] we were able to filter out potential drug-target combinations that not only have better docking scores, good free energy, and residence time but they are in clinical phase trials. Here, we took those top listed host-targets–drugs and investigated to enhance our understanding of the residue-level details, kinetics, and dissociation pathways from the targets. Using all-atom MD and well-tempered metadynamics, we performed 20–35 independent dissociation pathway simulations for each system to get statistically reliable results.

Analyzing 100 ns long all-atom trajectories, we explored the stable interactions and key residues involved at the bound state. We observed multiple dissociation pathways with varying target-drug systems and quantified the probability of those pathways for respective systems from the dissociation trajectories. We have explicitly shown the most probable pathway and key residues present at the pathway, facilitating the dissociation mechanism. For HDAC2-Valproicacid, we observe the dominating pathway to be the one with ILE24, LEU117, GLU113 residues present at the gateway. Similarly, PHE155 for HDAC2-Nafamostat; ARG94 for eIF4E2-Nafamostat; MET79 and GLU81 for eIF4E2-Camostat; PRO180 for eIF4E2-Zotatifin; LYS69 and ASP176 for CSK22-Silmitasertib are the dominant pathways and associated key gatekeeper residues facilitating the (un)binding mechanism. Since the association of a ligand from the aqueous medium to the protein binding pocket is entropically favorable and occurs in the long timescale, we did not perform here association simulations for the ligands. However, we expect that the ligands will possibly follow the same pathways as in the case of the dissociation mechanism. Our extensive study showed that host target HDAC2 has deep concave pockets that maximize the protein-drug interactions. The predominant pathway taken by both HDAC2 combinations is from the catalytic region referred to as a lipophilic tube, which includes the residues PHE155 and HIS183.

Similarly, CSK22-Silmaterstib, the residues LYS69, and ASP176 are mediators involved in interaction with the drug molecule. Here we observe LYS69 and ASP176 facilitating the dissociation of the drug. Our results demonstrate how these crucial drugs interact with the host machinery and the specific target residues, showing involvement in the mechanism. Mutation of these important residues and further investigation with the ligands to understand the effects of mutation on thermodynamics and kinetics will be an essential aspect to study in the future for these systems. From the comparative efficacy between the free energy barriers and reported activity (IC50), we observe overall a good correlation for studied drugs with on- and off-target combinations and propose Silmitasertib and Zotatifin to be a potential candidates for the SARS-CoV-2 inhibition with very high efficacy.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Debabrata Pramanik https://orcid.org/0000-0002-3432-4498
Jayant Kumar Singh https://orcid.org/0000-0001-8056-2115

REFERENCES

[1] P. Calligari, S. Bobone, G. Ricci, A. Bocedi, *Viruses* **2020**, *12*, 445. https://doi.org/10.3390/v12040445

[2] L. Wang, Y. Wang, D. Ye, Q. Liu, *Int. J. Antimicrob. Agents* **2020**, *55*, 105948. https://doi.org/10.1016/j.ijantimicag.2020.105948

[3] M. C. Grant, L. Geoghegan, M. Arbyn, Z. Mohammed, L. McGuinness, E. L. Clarke, R. G. Wade, *PLoS One* **2020**, *15*, e0234765. https://doi.org/10.1371/journal.pone.0234765

[4] J. Ziebuhr, *Curr. Opin. Microbiol.* **2004**, *7*, 412.

[5] M. Prajapat, P. Sarma, N. Shekhar, A. Prakash, P. Avti, A. Bhattacharyya, H. Kaur, S. Kumar, S. Bansal, A. R. Sharma, B. Medhi, *Indian J. Pharm.* **2020**, *52*, 142.

[6] M. A. Shereen, S. Khan, A. Kazmi, N. Bashir, R. Siddique, *J. Adv. Res.* **2020**, *24*, 91.

[7] Y. J. Tan, S. G. Lim, W. Hong, *Antivir. Res.* **2005**, *65*, 69.

[8] Y. Ye, B. G. Hogue, *J. Virol.* **2007**, *81*, 3597.

[9] A. A. Al-Karmalawy, M. A. Dahab, A. M. Metwaly, S. S. Elhady, E. B. Elkaeed, I. H. Eissa, et al., *Front. Chem.* **2021**, *9*, 661230. https://doi.org/10.3389/fchem.2021.661230

[10] S. Liu, L. Zhu, G. Xie, B. W.-Y. Mok, Z. Yang, S. Deng, S. Y. Lau, P. Chen, P. Wang, H. Chen, Z. Cai, *CCS Chem.* **2021**, *4*, 559.

[11] J. Shastri, S. Parikh, V. Aggarwal, S. Agrawal, N. Chatterjee, R. Shah, P. Devi, P. Mehta, R. Pandey, *Front. Med.* **2021**, *8*, 737007. https://doi.org/10.3389/fmed.2021.737007

[12] D. Lindenblatt, A. Nickelsen, V. M. Applegate, J. Hochscherf, B. Witulski, Z. Bouaziz, C. Marminon, M. Bretner, M. le Borgne, J. Jose, K. Niefind, *ACS Omega* **2019**, *4*, 5471.

[13] L. R. Cruz, I. Baladrón, A. Rittoles, P. A. Díaz, C. Valenzuela, R. Santana, M. M. Vázquez, A. García, D. Chacón, D. Thompson, G. Perera, A. González, R. Reyes, L. Torres, J. Pérez, Y. Valido, R. Rodriguez, D. M. Vázquez-Bloomquist, M. Rosales, A. C. Ramón, G. V. Pérez, G. Guillén, V. Muzio, Y. Perera, S. E. Perea, ATENEA-Co-300 Group, *ACS Pharmacol Transl Sci.* **2021**, *4*, 206.

[14] F. F. Wagner, M. Weïwer, S. Steinbacher, A. Schomburg, P. Reinemer, J. P. Gale, A. J. Campbell, S. L. Fisher, W. N. Zhao, S. A. Reis, K. M. Hennig, M. Thomas, P. Müller, M. R. Jefson, D. M. Fass, S. J. Haggarty, Y. L. Zhang, E. B. Holson, *Bioorg. Med. Chem.* **2016**, *24*, 4008.

[15] F. Yang, N. Zhao, D. Ge, Y. Chen, *RSC Adv.* **2019**, *9*, 19571.

[16] L. K. Clark, T. J. Green, C. M. Petit, *J. Virol.* **2021**, *95*, e02019-20. https://doi.org/10.1128/jvi.02019-20

[17] M. Bouhaddou, D. Memon, B. Meyer, K. M. White, V. V. Rezelj, M. Correa Marrero, B. J. Polacco, J. E. Melnyk, S. Ulferts, R. M. Kaake, J. Batra, A. L. Richards, E. Stevenson, D. E. Gordon, A. Rojc, K. Obernier, J. M. Fabius, M. Soucheray, L. Miorin, E. Moreno, C. Koh, Q. D. Tran, A. Hardy, R. Robinot, T. Vallet, B. E. Nilsson-Payant, C. Hernandez-Armenta, A. Dunham, S. Weigang, J. Knerr, M. Modak, D. Quintero, Y. Zhou, A. Dugourd, A. Valdeolivas, T. Patil, Q. Li, R. Hüttenhain, M. Cakir, M. Muralidharan, M. Kim, G. Jang, B. Tutuncuoglu, J. Hiatt, J. Z. Guo, J. Xu, S. Bouhaddou, C. J. P. Mathy, A. Gaulton, E. J. Manners, E. Félix, Y. Shi, M. Goff, J. K. Lim, T. McBride, M. C. O'Neal, Y. Cai, J. C. J. Chang, D. J. Broadhurst, S. Klippsten, E. de Wit, A. R. Leach, T. Kortemme, B. Shoichet, M. Ott, J. Saez-Rodriguez, B. R. tenOever, R. D. Mullins, E. R. Fischer, G. Kochs, R. Grosse, A. García-Sastre, M. Vignuzzi, J. R. Johnson, K. M. Shokat, D. L. Swaney, P. Beltrao, N. J. Krogan, *Cell* **2020**, *182*, 685.

[18] D. Bhowmik, R. D. Sharma, A. Prakash, D. Kumar, *J. Mol. Struct.* **2021**, *1233*, 130094. https://doi.org/10.1016/j.molstruc.2021.130094

[19] K. Doi, M. Ikeda, N. Hayase, K. Moriya, N. Morimura, H. Maehara, et al., *Crit. Care* **2020**, *24*, 20.

[20] J. Ramakrishnan, S. Kandasamy, A. Iruthayaraj, S. Magudeeswaran, K. Chinnasamy, K. Poomani, *Appl. Biochem. Biotechnol.* **2021**, *193*, 1909.

[21] D. E. Escalante, D. M. Ferguson, *Med. Chem. Res.* **2021**, *30*, 399.

[22] N. M. Anand, D. H. Liya, A. K. Pradhan, N. Tayal, A. Bansal, S. Donakonda, A. K. Jainarayanan, *PLoS One* **2021**, *16*, 1.

[23] S. K. Behera, N. Vhora, D. Contractor, A. Shard, D. Kumar, K. Kalia, A. Jain, *Sci. Rep.* **2021**, *11*, 1.

[24] A. Carino, F. Moraca, B. Fiorillo, S. Marchianò, V. Sepe, M. Biagioli, C. Finamore, S. Bozza, D. Francisci, E. Distrutti, B. Catalanotti, A. Zampella, S. Fiorucci, *Front. Chem.* **2020**, *8*, 1.

[25] T. Hempel, L. Raich, S. Olsson, N. P. Azouz, A. M. Klingler, M. Hoffmann, S. Pöhlmann, M. E. Rothenberg, F. Noé, *Chem. Sci.* **2021**, *12*, 983.

[26] A. Kumawat, S. Namsani, D. Pramanik, S. Roy, J. K. Singh, *J. Biomol. Struct. Dyn.* **2021**, 1. https://doi.org/10.1080/07391102.2021.1937319

[27] PRinS3, Version v1.0.0, Release 2021, Prescience Insilico Pvt. Ltd., India. **2021**.

[28] Y. Sixto-López, M. Bello, J. Correa-Basurto, *J. Comput. Aided Mol. Des.* **2020**, *34*, 857.

[29] P. Tiwary, V. Limongelli, M. Salvalaglio, M. Parrinello, *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, E386. https://doi.org/10.1073/pnas.1424461112

[30] A. Sauvat, F. Ciccosanti, F. Colavita, M. Di Rienzo, C. Castilletti, M. R. Capobianchi, et al., *Cell Death Dis.* **2020**, *11*, 656.

[31] G. M. Morris, H. Ruth, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, et al., *J. Comput. Chem.* **2009**, *30*, 2785.

[32] B. Webb, A. Sali, *Curr. Protoc. Bioinformatics* **2016**, *2016*, 5.6.1.

[33] R. Harris, A. J. Olson, D. S. Goodsell, *Proteins Struct. Funct. Genet.* **2008**, *70*, 1506.

[34] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, *SoftwareX* **2015**, *1-2*, 1.

[35] A. D. MacKerell, N. Banavali, N. Foloppe, *Biopolymers* **2000**, *56*, 257. DOI: 10.1002/1097-0282(2000)56:43.0.CO;2-W

[36] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, *Gaussian 09, Revision A. 02*, Gaussian, Inc, Wallingford, CT **2009**, p. 2015.

[37] V. Zoete, M. A. Cuendet, A. Grosdidier, O. Michielin, *J. Comput. Chem.* **2011**, *32*, 2359.

[38] C. M. Breneman, K. B. Wiberg, *J. Comput. Chem.* **1990**, *11*, 361.

[39] G. Bussi, D. Donadio, M. Parrinello, *J. Chem. Phys.* **2007**, *126*, 14101. https://doi.org/10.1063/1.2408420

[40] M. Parrinello, A. Rahman, *J. Appl. Phys.* **1981**, *52*, 7182.

[41] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, *J. Chem. Phys.* **1995**, *103*, 8577.

[42] B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, *J. Comput. Chem.* **1997**, *18*, 1463.

[43] C. Bouysset, S. Fiorucci, *J. Cheminform.* **2021**, *13*, 72. https://doi.org/10.1186/s13321-021-00548-6

[44] Z. Smith, D. Pramanik, S. T. Tsai, P. Tiwary, *J. Chem. Phys.* **2018**, *149*, 234105. https://doi.org/10.1063/1.5064856

[45] D. Pramanik, Z. Smith, A. Kells, P. Tiwary, *J. Phys. Chem. B* **2019**, *123*, 3672.

[46] P. Tiwary, *J. Phys. Chem. B* **2017**, *121*, 10841. https://doi.org/10.1021/acs.jpcb.7b09510

[47] J. M. L. Ribeiro, S. T. Tsai, D. Pramanik, Y. Wang, P. Tiwary, *Biochemistry* **2019**, *58*, 156.

[48] A. Laio, M. Parrinello, *Proc. Natl. Acad. Sci.* **2002**, *99*, 12562. https://doi.org/10.1073/PNAS.202427399

[49] A. Barducci, G. Bussi, M. Parrinello, *Phys. Rev. Lett.* **2008**, *100*, 20603. https://doi.org/10.1103/PhysRevLett.100.020603

[50] M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R. A. Broglia, M. Parrinello, *Comput. Phys. Commun.* **2009**, *180*, 1961.

[51] S. Namsani, D. Pramanik, M. A. Khan, S. Roy, J. K. Singh, *J. Biomol. Struct. Dyn.* **2021**, 1. https://doi.org/10.1080/07391102.2021.1892530

[52] R. Battistutta, G. Cozza, F. Pierre, E. Papinutto, G. Lolli, S. Sarno, S. E. O'Brien, A. Siddiqui-Jain, M. Haddach, K. Anderes, D. M. Ryckman, F. Meggio, L. A. Pinna, *Biochemistry*. **2011**, *50*, 8478. https://www.selleckchem.com/products/cx-4945-silmitasertib.html

[53] J. T. Ernst, P. A. Thompson, C. Nilewski, P. A. Sprengeler, S. Sperry, G. Packard, T. Michels, A. Xiang, C. Tran, C. J. Wegerski, B. Eam, N. P. Young, S. Fish, J. Chen, H. Howard, J. Staunton, J. Molter, J. Clarine, A. Nevarez, G. G. Chiang, J. R. Appleman, K. R. Webster, S. H. Reich, *J. Med. Chem.* **2020**, *63*, 5879. https://www.medchemexpress.com/zotatifin.html

[54] M. Shoji, I. Ninomiya, I. Makino, J. Kinoshita, K. Nakamura, K. Oyama, H. Nakagawara, H. Fujita, H. Tajima, H. Takamura, H. Kitagawa, S. Fushida, S. Harada, T. Fujimura, T. Ohta, *Int. J. Oncol.* **2012**, *40*, 2140.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: D. Pramanik, A. B. Pawar, S. Roy, J. K. Singh, *J. Comput. Chem.* **2022**, *43*(18), 1237. https://doi.org/10.1002/jcc.26888