



UNICA

UNIVERSITÀ
DEGLI STUDI
DI CAGLIARI



Università di Cagliari

UNICA IRIS Institutional Research Information System

This is the Author's accepted manuscript version of the following contribution (in bold face the UNICA contributors):

Integrated docking and enhanced sampling-based selection of repurposing drugs for SARS-CoV-2 by targeting host dependent factors

Amit Kumawat, Sadanandam Namsani, Debabrata Pramanik, Sudip Roy, and Jayant K. Singh

Journal of Biomolecular Structure and Dynamics, 40(20), 9897–9908

The publisher's version is available at:

<https://doi.org/10.1080/07391102.2021.1937319>

When citing, please refer to the published version.

This full text was downloaded from UNICA IRIS <https://iris.unica.it/>

Integrated docking and enhanced sampling-based selection of repurposing drugs for SARS-CoV-2 by targeting host dependent factors

Amit Kumawat^{‡,1,2}, Sadanandam Namsani^{‡,1}, Debabrata Pramanik^{‡,2}, Sudip Roy^{1,} and Jayant K. Singh^{*,1,2}*

¹ Prescience Insilico Private Limited, Old Madras Road, Bangalore 560049, India

² Department of Chemical Engineering, Indian Institute of Technology, Kanpur 208016, India

[‡] Same contribution

*Corresponding authors: Sudip Roy (sudip@prescience.in) and Jayant K. Singh (jayantks@iitk.ac.in)

KEYWORDS

COVID19, molecular dynamics simulation, host proteins, protein-protein interaction, gene regulation, enhanced sampling

ABSTRACT

Since the onset of global pandemic, the most focused research currently in progress is the development of potential drug candidates and clinical trials of existing FDA approved drugs for other relevant diseases, in order to repurpose them for the COVID-19. At the same time, several high throughput screenings of drugs have been reported to inhibit the viral components during the early course of infection but with little proven efficacies. Here, we investigate the drug repurposing strategies to counteract the coronavirus infection which involves several potential targetable host proteins involved in viral replication and disease progression. We report the high throughput analysis of literature-derived repurposing drug candidates that can be used to target the genetic regulators known to interact with viral proteins based on experimental and interactome studies. In this work we have performed integrated molecular docking followed by molecular dynamics (MD) simulations and free energy calculations through an expedite in silico process where the number of screened candidates reduces sequentially at every step based on physicochemical information. We elucidate that in addition to the pre-clinical and FDA approved drugs that targets specific regulatory proteins, a range of chemical compounds (Nafamostat, Chloramphenicol, Ponatinib) binds to the other gene transcription and translation regulatory protein with higher affinity and may harbour potential for therapeutic uses. There is a rapid growing interest in the development of combination therapy for COVID19 to target multiple enzymes/pathways. Our in-silico approach would be useful in generating leads for experimental screening for rapid drug repurposing against SARS-CoV-2 interacting host proteins.

Introduction

With currently no specific antiviral drugs available for COVID19 treatment, there is an urgent need for rapid repurposing of drugs for direct therapeutic approaches.(Zhou, Wang, Tang, Nussinov & Cheng ; Guy, DiPaola, Romanelli & Dutch 2020; Singh et al. 2020) This approach takes advantage of existing information for the approved small molecules and biologics and enables rapid clinical trials or regulatory review. Since the onset of the global pandemic, high throughput screening of drugs has been reported to inhibit the viral components during the early course of infection.(Cao et al. 2020; Mandala et al. 2020; Menéndez, Byléhn, Perez-Lemus, Alvarado & de Pablo 2020; Sacco et al. 2020; Tharappel, Samrat, Li & Li 2020; Trezza, Iovinelli, Santucci, Prischi & Spiga 2020) A large number of drug discovery and repurposing studies are based on the pharmacoepidemiological data available from the coronaviruses family members such as the Middle East Respiratory Syndrome coronavirus (MERS-CoV), the Severe Acute Respiratory Syndrome coronavirus (SARS-CoV-1) and other viruses (Ebola, HIV, Influenza).(Ziebuhr 2004; Du et al. 2009; Abdelrahman, Li & Wang 2020; Gordon et al. 2020; Pardo, Shukla, Chamarthi & Gupte 2020) However, these potential prophylactic measures have certain drawbacks such as strain specific drug resistance and substantial differences in viral protein targets.(Gordon et al. 2020; Plante et al. 2020) SARS-CoV-2 possesses an extremely large RNA genome and a unique RNA replication mechanism, which makes it very challenging to find efficient inhibitors for SARS-CoV-2 in comparison to other viruses. At present, more than 50 vaccine candidates are under clinical trials, and a few are being administered. However, mutations in the genome of SARS-CoV-2 pose a significant challenge in the vaccine development and their efficacy. With the emergence of a number of variant strains, the efficacy of vaccines in the long term has been a debatable question for scientists around the world. In such a situation, it becomes

necessary to identify potential drugs that can inhibit/restrict the viral replications. Interestingly, targeting host factors is an important strategy in overcoming these restrictions and designing host-directed therapies for treating viral infections. With this inception, several in-vivo and in silico studies have characterized the virus-host interaction networks to functionally annotate multiple host proteins implicated in pathways that promote viral pathogenesis.(Li et al. ; Ahmed et al. 2020; Gordon et al. 2020; Zhou et al. 2020) In particular, Gordon *et al.*, in their recent study, identifies 332 high confidence SARS-CoV-2 human protein-protein interactions that aim to understand the hijacking of the host cell during infection.(Gordon et al. 2020) The interactome reveals several host proteins involved in biological processes such as host genetic regulation, innate immune pathways, vesicle transport and translocation, and provides a comprehensive evaluation of small-molecule candidates for drug repurposing.

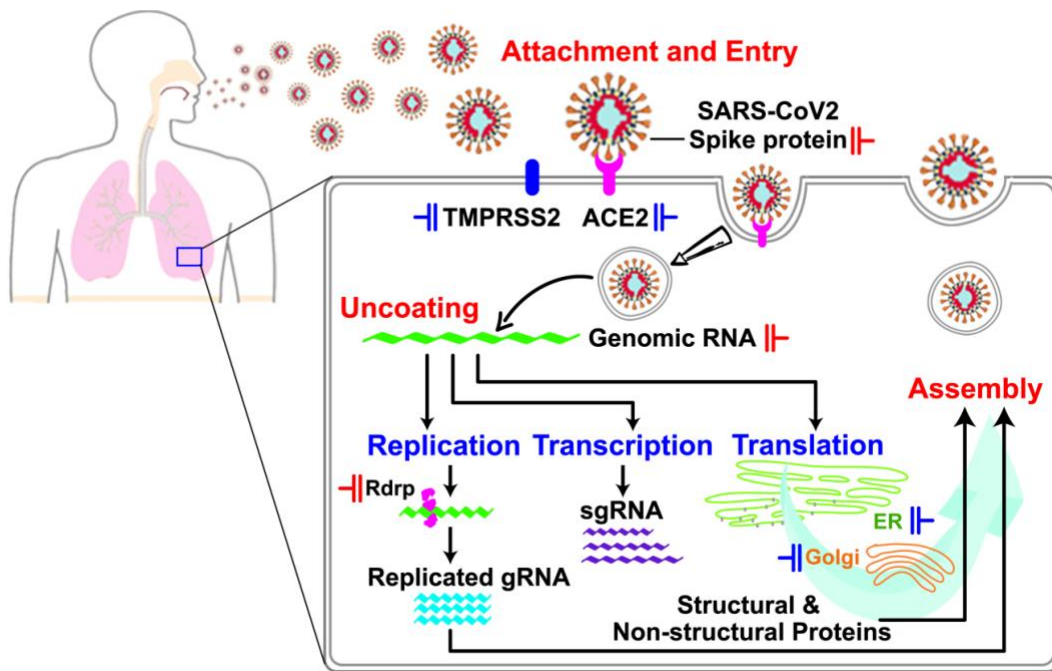


Figure 1. Illustration of therapeutic strategies against COVID-19 through targeting the host and viral proteins as a mechanism for the inhibition of viral transmission. SARS-CoV-2 interacts with Angiotensin-converting enzyme 2 (ACE2) receptor protein on the cell surface and integrates the

virus into the cell. The drug therapy can be targeted at the different host proteins situated in Golgi, ER or gene regulatory pathways involved in the viral replication.

One of the important aspects of viral replication involves the interaction of SARS-CoV-2 with the host machinery. These interactions form a potential target to inhibit SARS-CoV-2 replication and assembly schematically shown in Figure 1. The binding of SARS-CoV-2 at the cellular membrane is mediated by spike (S) glycoprotein of SARS-CoV-2 and the host receptor ACE2. Several experimental and computational studies provide remarkable insights into the structural stability and dynamics of the interacting domains. Interestingly, recent studies reported that the binding of S protein to ACE2 involve distinct conformational changes between open and closed states that are governed by hydrophobic contacts in addition to the hydrogen bonds and electrostatic interactions. This information provides possible target sites for designing and repurposing potential drugs aimed at destabilizing the interactions at the point of virus entry.(Moreira, Chwastyk, Baker, Guzman & Poma 2020; Moreira, Guzman, Boopathi, Baker & Poma 2020; Yang et al. 2020) In this work, we demonstrated an *in silico* robust and novel drug repurposing strategy to counteract the coronavirus infection that involves several potential targetable host proteins involved in viral replication and disease progression. Of all the 66 druggable proteins that are targeted by 69 compounds based on previous interactome studies, (Gordon et al. 2020) we cluster these proteins into different categories based on the biological and molecular functions. We identify 27 proteins based on UniProt annotations that are involved in host gene regulations and host-viral interactions. These proteins can be targeted by more than 30 drugs, as suggested by Gordon and coworkers through knowledge-based search and data mining. (Gordon et al. 2020) Here, we have focused the study on the identification of drugs that can regulate host translation and transcription, and sequentially the viral replication. Host genetic regulatory proteins play an

important role in the viral transmission cycle. The viral genome is transcribed or directly translated to produce structural and non-structural proteins (NSPs) necessary to assemble new viral particles.(V'kovski, Kratzel, Steiner, Stalder & Thiel 2020) An extensive work on the host-viral interactome reveals interactions between several key host factors with the viral proteins that facilitate the control of SARS-CoV-2 over the mRNA processing and translational machinery.(Li et al. ; Gordon et al. 2020) Evidently, these viral mediated alterations in the key genetic regulators can form the primary targets to regulate the polyprotein expression.(Poduri, Joshi & Jagadeesh 2020) Based on the interactome data and the Uniprot annotations for the host proteins, we have shortlisted 13 proteins that constitute the important host genetic regulators (translation and transcription). The proteins were further refined to identify the availability of crystal/cryo-EM structure, binding site information, and structural information for any known drug-bound complex. Out of 13 proteins that are involved in human gene regulation, structural information was available for 9 proteins (Table 1).

Among this list of proteins, the host transcriptional regulator proteins such as Bromo-domain containing proteins (BRD2/BRD4) exhibit antiviral activity upon inhibition through activation of innate immunity pathways. (Filippakopoulos et al. 2010; Pierre et al. 2011; Reich et al. 2018) DNA methylation by DNMT1 is another important component of gene regulation. Viral infections identify these DNA methylation patterns to induce endocytosis development. SARS-CoV-2 meticulously exploits this epigenetic mechanism for the production of ACE2 enzyme (virus receptor on the host's lung epithelial cells). (Schäfer & Baric 2017; Tutuncuoglu et al. 2020) Interestingly, therapeutic targeting of the host translation machinery (initiation and elongation factors; EF1A1, IF4E, IF4E2) may also prevent viral protein production and inhibit the viral replication. (Soukarieh et al. 2016)

Table 1. Function based classification and selection of host gene regulatory proteins from 332 high confidence host-pathogen protein-protein interactions (PPIs) between SARS-CoV-2 and human proteins based on previous interactome study. (Gordon et al. 2020) The table highlights the preclinical or FDA approved drugs that target the host proteins and the structural information for these proteins.

INDEX	Gene regulation	Uniprot	Gene-Name	Article Drugs	Structure (PDB ID)
1	BRD2	P25440	BRD2_HUMAN	JQ1, RVX-208, ABBV-744, CPI-0610, dBET6, MZ1	3ONI
2	BRD4	O60885	BRD4_HUMAN	JQ1, RVX-208, ABBV-744, CPI-0610, dBET6, MZ1	3MXF
3	CSNK2A2	P19784	CSK22_HUMAN	TMCB, Silmitasertib	6HMB
4	DNMT1	P26358	DNMT1_HUMAN	Azacitidine, XL413	4WXX
5	EEF1A1	P68104	EF1A1_HUMAN	Ternatin-4-(DA3)	1SYW
6	EIF4E	P06730	IF4E_HUMAN	4E2RCat	1IPC
7	EIF4E2	O60573	IF4E2_HUMAN	Zotatifin-(eFT226)	2JGB
8	HDAC2	Q92769	HDAC2_HUMAN	Apicidin, Valproic-Acid	4LY1
9	MNK2	Q9HBH9	MKNK2_HUMAN	Tomivosertib-(eFT-508)	6CK6
10	MRPS27	Q92552	RT27_HUMAN	Chloramphenicol, Linezolid, Tigecycline	-
11	EIF4G1	Q04637	IF4G1_HUMAN	4E2RCat	-
12	EIF4H	Q15056	IF4H_HUMAN	Zotatifin-(eFT226)	-
13	LARP1	Q6PKG0	LARP1_HUMAN	Rapamycin, Sapanisertib	-
15	NUP98	P52948	NUP98_HUMAN	Verdinexor	-

COVID-19 has a triphasic course where the patients undergo a transition from having mild respiratory and systemic symptoms (cold, cough and fever) to hyperinflammation of lung tissues, also referred to as cytokine storm syndrome resulting in acute respiratory distress syndrome (ARDS) or multiorgan failure. (Mehta et al. 2020; Tang et al. 2020) These manifestations namely

inflammation, immune dysfunction and coagulopathy, often associated with neoplasia, in patients with severe SARS-CoV-2 infections give a rationale for the testing of anticancer drugs, (El Bairi et al. 2020; Saini et al. 2020) and in addition use of immunosuppressive drugs to treat the hyperinflammatory phase of COVID19 that exhibits a high level of cytokines such as IL-6, IL-7 and TNF- α .(El Bairi et al. 2020; Liu, Li, Zhou, Guan & Xiang 2020; Saini et al. 2020) Hence, several immunosuppressive and anticancer drugs such as tocilizumab, dexamethasone, selinexor, imatinib and sirolimus are being investigated under clinical trials for their possible role toward effective therapy for COVID19. Interestingly, coagulant abnormalities have been observed in patients with severe COVID19 and a range of anticoagulant drugs are being investigated as a part of antithrombotic therapy.(Saini et al. 2020) Here, we have performed an extensive in silico drug repurposing exercise for these proteins for a list of selective drugs that are anticancer compounds, anti-coagulants, antibiotics (bacterial infections) or immunosuppressive in nature. Interestingly, we have observed that these drugs are chemically aromatic and heterocyclic nitrogen-containing moieties with a significant number of compounds containing halogen atoms. The inclusion of electron-rich nitrogen or the halogens (predominant for steric effects) in these heterocyclic ring structures favor intermolecular interactions with the proteins such as hydrogen bonds, dipole-dipole interactions, hydrophobic effects, van der Waals and π -stacking interactions that can improve pharmacological features.(Marcelo Zaldini, Suellen Melo, Diogo Rodrigo, Walter Filgueira de Azevedo & Ana Cristina Lima 2010; Wilcken, Zimmermann, Lange, Joerger & Boeckler 2013; Poma, Chwastyk & Cieplak 2015; Pennington & Moustakas 2017)

To understand and screen the drugs that could interact and bind with host proteins, we have assembled an integrated in silico approach consisting of molecular docking followed by molecular dynamics (MD) simulations and enhanced sampling free energy calculations. We followed the

protocol designed and applied in our previous work, with some modifications for the identification of antiviral drugs targeting the viral proteins.(Sadanandam, Debabrata, Mohd Aamir, Sudip & Jayant 2020) Our approach greatly reduces the time and cost of repurposing by in silico identification of lead compounds through an expedited automated process where the number of screened candidates reduces sequentially at every step based on physicochemical information. The drug screening process was performed based on multi-protein against multi-drugs approach with very robust free energy calculations from enhanced sampling. Prior studies have attempted to repurpose drugs that are implicated in similar diseases such as cancers, rheumatoid disorders and blood coagulation disorders. (Saini et al. 2020) We examined the physical nature (polar/non-polar) of interactions between the binding site of the proteins and drugs using molecular docking studies. Based on the initial findings, several candidate drugs exhibit higher docking binding energy and hydrogen bonded interactions as compared to the control protein-drug complex. However, virtual screening tools like molecular docking are known to introduce inaccuracies in the binding predictions due to the lack of sampling and exclusion of various factors such as protein dynamics and solvent electrostatics. Our high throughput analysis that includes free energy barrier calculations predicts various compounds (FDA-approved or clinically approved) that complements the earlier works in the drug repurposing strategies for COVID19.

Materials and methods

Rapid Drug repurposing Protocol

In this work, we have designed an integrated molecular docking and simulation protocol for rapid drug repurposing strategy (Figure 2). At first, molecular docking has been used to study the binding

affinities for the drugs and host transcriptional regulator proteins. The molecular docking also provides a reasonable well minimized conformation and binding pose for initiating a molecular dynamics (MD) simulation. Further, the stability of the obtained docked poses has been evaluated using MD simulations. MD simulations provide the interaction energies and the stability of the drugs at the binding site of the protein. After MD simulations, the drugs with significant interactions based on hydrogen bonding and non-bonded interactions with the binding sites of the target proteins have been selected for well-tempered metadynamics (wt-metaD) simulations. The binding free energy for each drug with the target proteins calculated from wt-metaD and enhanced sampling of drugs at the binding site deterministically indicates the best possible repurposing drugs. Further, we have analyzed all the trajectories from MD and wt-metaD to understand specific interactions between the drugs and residues at the binding sites. Additionally, we have calculated the drugs' residence time at the binding site using the wt-metaD simulations trajectories.

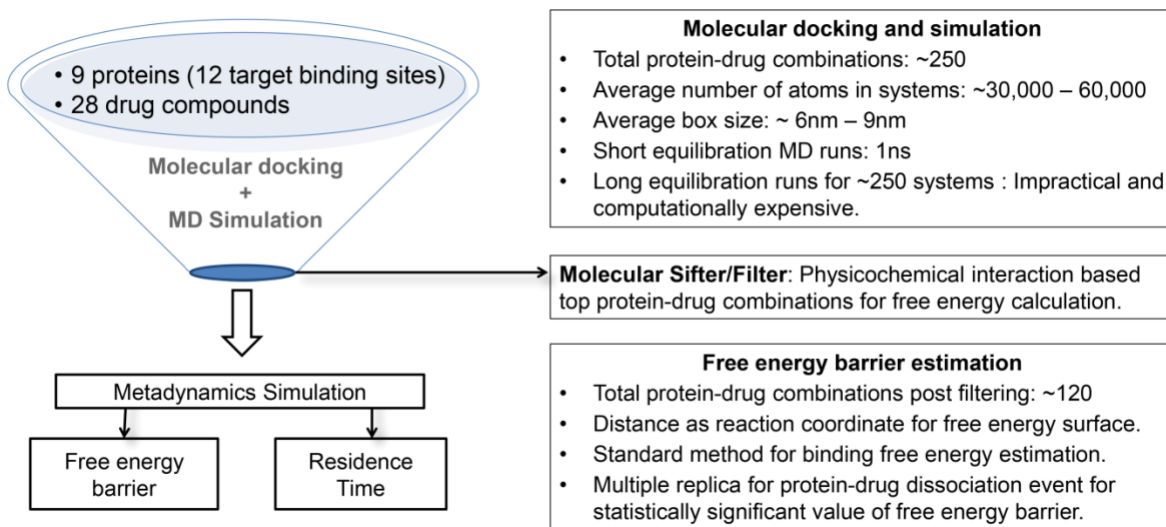


Figure 2. Illustration of the protocol that performs integrated molecular docking followed by molecular dynamics (MD) simulations and free energy calculations through an accelerated in silico

process where the number of protein-drug systems is reduced at every step based on protein-ligand interactions.

Molecular modelling and docking

The Autodock suite (Autodock 4.2.6) was used to perform molecular docking analysis of the compounds for different proteins as the docking targets. (Morris et al. 2009) The docking was performed onto the available crystal structures for different proteins. The missing loop regions in the selective proteins were modelled using Modeller (Modeller 9.24).(Webb & Sali 2016) The docking sites were identified based on either availability of inhibitor-bound crystal structures or ligand site prediction using the Autoligand tool in the Autodock suite.(Harris, Olson & Goodsell 2008) The protein structures and chemical compounds were energy minimised before docking analysis. To perform the docking of the compounds onto the identified ligand binding sites, grid-based ligand docking was performed. Ligand centered grid maps were generated using the Autogrid program with a spacing of 0.375Å and grid dimensions of 60×60×60 Å³. The best-docked structures were identified based on the docking energy and population distribution of the lowest-energy clusters from the conformational clustering histogram in Autodock.

Molecular Dynamics Simulations

We have performed molecular dynamics simulations for different protein-ligand complexes using GROMACS-2018. (Abraham et al. 2015) The best-docked ligand conformations for different systems were selected as the starting structure for the simulations. In addition, the simulations were also performed for selective protein-ligand crystal structures as control systems, where the ligand can differ from the list of screening compounds. The charmm27 force field with cmap corrections was used for the proteins.(MacKerell, Banavali & Foloppe 2001) The geometry

optimizations of all the ligands (Table S1) were performed using a semi-empirical method at the PM6 level, followed by geometry optimization using density functional theory (DFT) with the M06 functional and 6-311g (d,p) basis set. To account for the bulk solvent effects, the PCM method is used. Further, the partial atomic charges for the ligands are computed by fitting the electrostatic potential using the CHELPG method as implemented in the Gaussian09 code. These charges are computed for the optimized structures using a single point calculation at the DFT with the M06 functional with 6-311g (d,p) basis set and used water as the solvent. The forcefield parameters for ligand molecules were obtained using swissparam webservice which performs automated assignment of parameters by analogy and is compatible with charmm force field. (Zoete, Cuendet, Grosdidier & Michielin 2011) All the structures were solvated using the TIP3P water model and simulated with periodic boundary conditions. The systems were neutralized by adding counter ions, Na⁺ in the negatively charged system and Cl⁻ in the positively charged system. The structures were energy minimized using steepest descent algorithm. The energy minimization was evaluated based on the negative value of the potential energy and the maximum force is no greater than 1000 kJ/mol/nm. This was followed by NVT equilibration using modified Berendsen thermostat (Bussi, Donadio & Parrinello 2007) for 500 ps and NPT equilibration using Parrinello-Rahman barostat for 1 ns. (Parrinello & Rahman 1981) For the production run, the temperature was controlled through velocity rescaling at 300K with a time constant of 0.1 ps and pressure was kept constant at 1 bar. The cutoff for short-range interactions was 1.0 nm, and the long-range electrostatic interactions were calculated using Particle-Mesh Ewald (PME) method.(Essmann et al. 1995) The bonds were constrained using the LINCS algorithm.(Hess, Bekker, Berendsen & Fraaije 1997) We have performed equilibrium MD simulations of these systems for 1ns to obtain energetically favourable protein-ligand conformations post rigid docking

protocol. The non-bonded interaction energies and hydrogen bond analysis were performed using the in-built GROMACS utilities.

Metadynamics simulations

Further, we have performed free energy calculations using enhanced sampling based technique metadynamics and its variant well-tempered metadynamics.(Barducci, Bussi & Parrinello 2008) We have used software packages PLUMED 2.5.4 combined with GROMACS 2018 as the MD engine.(Laio & Parrinello 2002) We have taken the final coordinates from the equilibrium MD simulations as the initial starting conformations for metadynamics simulations. Since the dissociation event of a ligand from the binding pocket of a protein takes sufficiently long time, and this process is separated by high energy barriers, so unbiased brute force MD simulations would not be sufficient to observe dissociation events within a short simulation time. Thus, to make the dissociation process faster and to study these kind of complex systems, external bias is added to the system which accelerates the dissociation processes. Thus, using enhanced biased sampling, we study protein-ligand systems using available limited resources and within a short time. To address the length and time scale issues for complex biomolecular systems, various methods have been developed over the years to calculate free energies for these systems of interest. From these huge lists of enhanced sampling techniques, we used metadynamics (metaD) and its variant well-tempered metadynamics (wt-metaD). The reaction coordinate (s) is a linear or non-linear combination of the few selected order parameters of the systems which are functions of the coordinates of the system. Here, we choose distance (d) between the center of mass - center of mass of the ligand heavy atoms and protein binding pocket as the reaction coordinate for biasing in the simulations. We add a gaussian bias and once the system converges, we can extract the unbiased Boltzmann probability of the distribution of RC by adding up the deposited hills over the

course of the simulation. In the wt-metaD simulations, the Gaussian hills were deposited at every 500 steps (1ps), with hill height in the range of 1.5-2.0 kJ/mol, width 0.02-0.1 nm, biasfactor 10-15, and temperature at 300K. For each ligand-protein system, we performed minimum 10 independent runs starting from different initial coordinates and velocities for each copy and then averaged over probabilities to get the averaged free energy.

Further to support the discernment of ligand-protein binding nature, ligand residence time is estimated using the wt-metaD simulations data. The residence time (RT) of a ligand within a free energy basin can be estimated using the equation 1:

$$RT_{calc} = \alpha(t) \times t \quad (1)$$

In the above equation t is the wt-metaD simulation time required to observe the transition from free energy basin A to basin B and $\alpha(t)$ is the acceleration factor which is defined using the equation 2.

$$\alpha(t) = \frac{1}{t} \int_0^t e^{\beta V(s(R), t')} dt' \quad (2)$$

where the $\alpha(t)$ is running average of the bias applied over the metadynamics run within in the time t , $V(s(R), t')$ is the bias at time t in the metadynamics simulation and β is the Boltzmann factor. The simulation time necessary for the transition from the energy basin, t , is determined from the derivative of $\alpha(t)$ versus t plot.

$$\frac{d \alpha}{dt} = \frac{1}{t} \left[e^{\beta V(s,t)} - \frac{1}{t} \int_0^t e^{\beta V(s(R), t')} dt' \right] \quad (3)$$

The derivative exhibits abrupt change whenever the ligand crosses an energy barrier and goes into a new state. The sign change is used to identify the transition time (Figure S1), which is required

to calculate the residence using equation 1. This transition time is used to get the drug residence time.

Results and discussion

Structure selection and binding site determination

The dependence of virus on the host translational machinery has introduced constraints that are central to the viral biology and has led to an alternative and complementary strategy to target the host factors essential for viral replication. Translation factors play a critical role in protein synthesis, and dysregulation of their activity have been implicated in a broad range of disorders, including cancers, cell transformations, etc. Several small-molecule inhibitors have been proposed as therapeutic potential, and few are clinically approved that exhibit antitumor activity and modulate gene expression. Hence, several crystal structures of gene regulation proteins (eIF, kinases, histone deacetylases) with the inhibitors have been reported. The potential ligand-binding site in the target protein is the prerequisite and critical information for any drug repurposing strategy. Our approach is to computationally screen drugs based on the inhibitor binding site information obtained from the existing protein-drug crystal structures.

We have selected the crystal structures for proteins in the inhibitor bound state from the protein data bank. The preferential criterion for the selection was to identify the protein structures co-crystallised with drugs that are suggested for repurposing through an extensive literature and knowledge-based search.(Gordon et al. 2020) Interestingly, we could find structures for Casein kinase 2 (CSK22), MAP kinase 2 (MNK2) and bromodomains (BRD2/BRD4) proteins in complex with Silmitasertib, Tomivosertib and JQ1 respectively (PDB: 6HMB, 2AC3, 3ONI, 3MXF). The ligand binding sites for histone deacetylase (HDAC2), translation initiation factor eIF4E were

obtained based on co-crystal structures with inhibitors (PDB: 4LY1, 1IPC2). These inhibitor bound structures were used as a control set for further analysis. We have also used the AutoLigand tool (Harris et al. 2008) to predict the binding site in the translation factors (EF1A1 and eIF4E2) and DNA methyltransferase 1 (DNMT1). Two ligand binding sites were predicted for EF1A1 (PDB: 1SYW) and one site for eIF4E2 (PDB: 2JGB). Further, two ligand binding sites were defined for DNMT1 based on the ligand bound crystal structure (PDB: 3SWR) and zinc finger region based on the function annotation from UniProt. These sites were further analysed to understand the electrostatic nature of the binding cavity before proceeding with the molecular docking protocol (Figure S2). Several of these binding sites in proteins (HDAC2, MNK2, DNMT1) exist as deep concave pockets that maximise the protein-ligand interactions. The core of the binding cavity consists of hydrophobic residues or polar residues that could positively contribute to the binding of organic or charged molecules. Similarly, there are binding sites (eIF4E, EF1A1) that are relatively flat protein surfaces consisting of both polar and apolar residues.

Molecular docking and MD simulation

Of the several in silico approaches that aim at identifying an existing drug for targeting clinically relevant targets, molecular docking has proven to be a powerful tool and could be reliably used for starting structure generation for MD. Docking of all the 12 binding sites corresponding to 9 proteins (Table 1) has been carried out with 28 drugs (Table S1). These drugs were derived from the suggested list of drugs using interactome studies for proteins involved in gene regulation. In addition, we have also included drugs that target the known protein-viral interactions based on UniProt functional annotations. The docking was performed using Autodock software, and the docking energy for the most populated cluster of docked structures was determined for more than 250 protein-drug combinations. Figure 3a shows the heat map representation of these docking

energies. We have observed that the docking energy difference between the best-docked conformations for different protein-drug complexes is less than 1 kcal/mol. However, the molecular docking protocol incorporates only the static interactions excluding the effect of solvent and ions in the system, and thus it does not count the entropic effect. It is very well established that thermodynamics and conformational dynamics play crucial roles in the interaction and stability of the protein-ligand systems. Hence, the docking protocol was followed by MD simulations of ligand-docked structures.

MD simulations were performed for all the best-docked structures of protein-drug complexes to equilibrate the systems. These would allow the conformational fluctuations of the drug inside the binding cavity and subsequent interactions with the protein residues at the binding site. We have analyzed the hydrogen bonding interactions and the non-bonded interaction energies for all the complexes. The non-bonded interaction energies for the protein-drug complex have been shown in Figure 3b in the form of a heatmap. It can be observed that few drugs have higher interaction energy as compared to the other drugs for individual proteins (Table S2). It is evident that the interaction energies for the complexes vary from -90 kcal/mol to -10 kcal/mol. The large difference in the electrostatic and van der Waals interaction energies between the systems is accounted for based on additional favorable polar and nonpolar interactions that vary with the chemical nature of the drug itself. To understand these differences in the energies, we have also analyzed the hydrogen bonding sites of drugs in each protein. In certain cases, the given set of drugs exhibit common hydrogen bonding residues in the binding site of a particular protein. For example, K69, H161, D176 in CSK22; Y386, N429 in BRD2; Y97, M105, N140 in BRD4 and so on (See Table S2 for list of hydrogen bonding residues for each protein-drug complex). Thus, it is evident from our analysis that multiple drugs can target the binding site with varying degrees of interactions.

Several of these drugs are FDA approved genetic translation regulators and have favorable pharmacological properties. Thus, these drugs can be promising candidates for repurposing for treating COVID-19 patients in phase 2-3 of the disease.

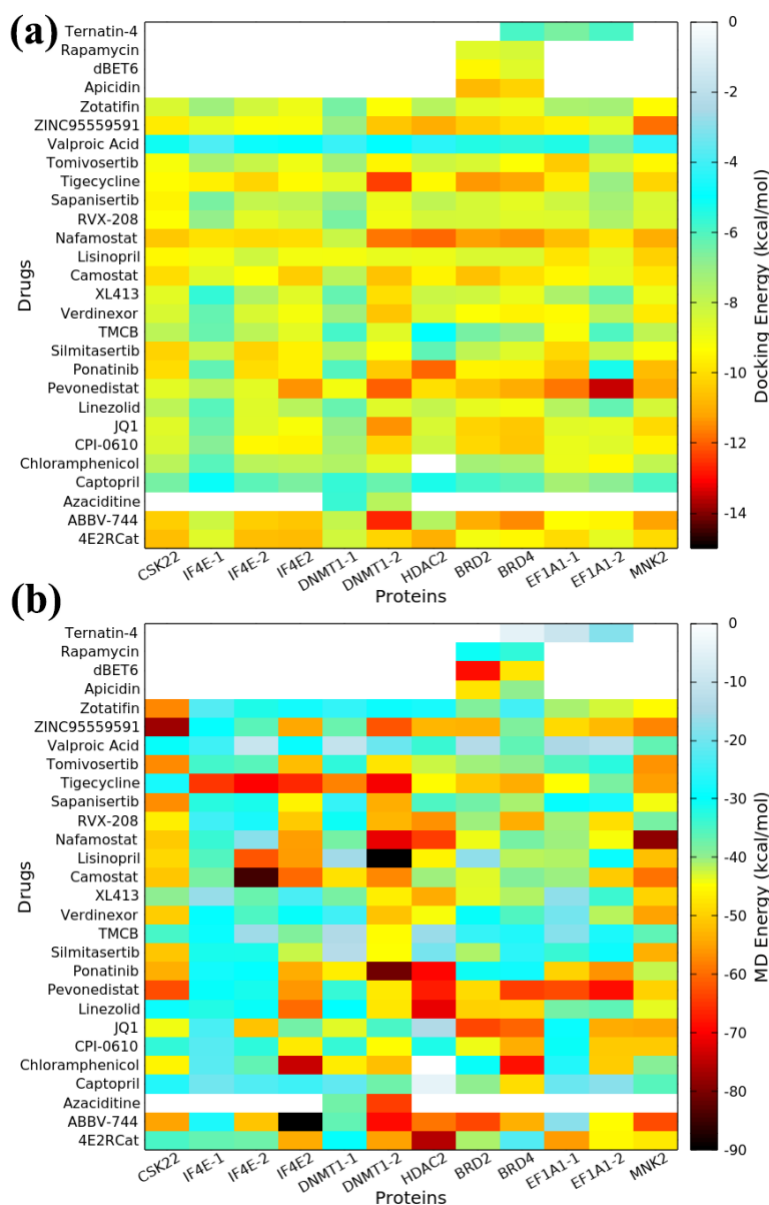


Figure 3. Heat map representation of the binding energy from (a) molecular docking and the electrostatic interaction energy from (b) MD simulation. The figure provides qualitative and quantitative insight into the differential order of interactions between various protein-ligand systems. The docking energy (binding energy) was calculated using Autodock software which

includes the contribution of the desolvation free energy of the ligand, and an estimate of the loss of conformational degrees of freedom of the ligand upon binding, whereas the interaction energy between protein and ligand from MD simulation was calculated using Gromacs utility (gmx energy) as the sum of coulomb and van der Waals interaction energy components. The values are reported in Table S2.

Further, to quantitatively estimate the binding affinity of these drugs, we calculated the dissociation free energy for these protein-drugs from free energy calculations. Interestingly, we find that not all drugs show high/similar binding affinity toward a protein in our preliminary analysis. Hence, we have reduced the number of calculations through the selection of a set of drugs for each protein for free energy calculations. The protein-drugs systems were ranked on the basis of docking energy, non-bonded interaction energies and the hydrogen bonds. The cutoff for each protein was chosen in such a way that the energetics exhibit the electrostatic or hydrophobic complementarity between the binding site of the protein and drugs (Table S2). For example, the van der Waals interaction energy dominates the electrostatic interaction energy for the HDAC2 protein and accounts for the top ranked protein-drugs with respect to docking energy. Similarly, the ranking was performed for all the systems, including the control set of drugs. This exercise reduces the large number of protein-drug combinations that ranked low and have lower interaction energy than the drugs with higher interactions.

Quantitative binding estimation

Following the above selection criterion, we have selected a list of protein-drug combinations shown in Table S1, and for those corresponding combinations, the instantaneous snapshots from the short equilibrium MD simulations are shown in Figure S3. We have calculated the average free

energy barriers for these protein-drug combinations. To calculate the average FESs, we first calculated the probability from independent free energy simulations and then calculated the averaged probability and average FES. The FES barriers are shown in Figure S4. We extracted the FES barriers corresponding to all the drugs for respective proteins and for multiple binding sites for a single protein. Table 2 shows the FES barrier heights for all the selected protein-drugs with respective errors in the calculations. We calculated the barrier heights by subtracting the maximum of the FES value from the global minimum in the FES profile. To understand these quantitative values better, we have plotted these barrier heights in Figure 4 in the heatmap representation. Figure 4 shows the heat map representation for the free energy barriers and compared our results with that obtained from the heat map shown in Figure 3.

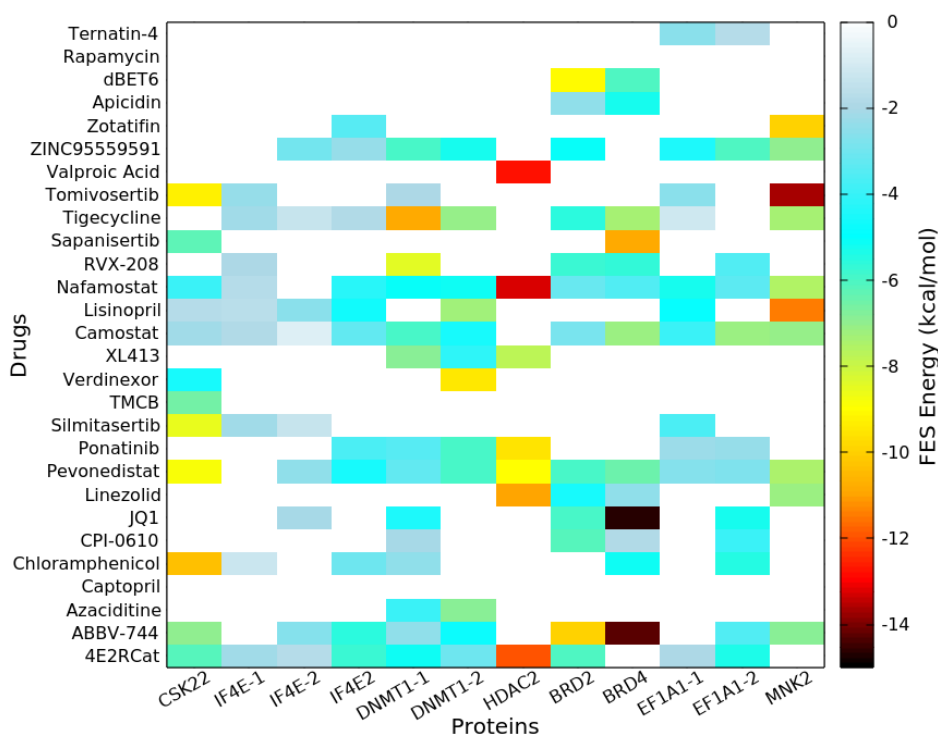


Figure 4. Heat map representation of the free energy barrier for different protein-drug combinations from metadynamics simulations. The color bar ranges from -15 kcal/mol to 0

kcal/mol, as shown on the right-side color bar. The values of free energy barrier and average error are reported in Table 2.

The color bar ranges from -15 kcal/mol to 0 kcal/mol, as shown on the right-side color bar. Interestingly, we find that the drugs with higher binding energy barriers for different proteins include control (crystal structure bound inhibitor, e.g., Tomivosertib in MNK2, Silmitasertib in CSK22) among the top scoring drugs. We must also note that our approach that identifies repurposing drugs with higher binding energy corroborates the known experimental findings with respect to approved or under clinical trial drugs for specific proteins (Valproic Acid in HDAC2, 4E2RCat in eIF4E).

Table 2. FES barrier heights and the estimated residence times for all the protein-drug systems.

The average error in the free energy barrier is reported in the parenthesis.

	Ligand	ΔG (kcal/mol)	Residence time (min)		Ligand	ΔG (kcal/mol)	Residence time (min)
B R D 2	dBET6 (L12)	-9.1 (0.50)	2.5E+13	B R D 4	dBET6 (L12)	-6.1 (0.27)	1.2E+04
	ABBV-744 (L11)	-10.0 (0.46)	4.0E+07		ABBV-744 (L11)	-14.3 (0.40)	2.0E+10
	CPI-0610 (L13)	-6.2 (0.45)	2.4E+00		CPI-0610 (L13)	-1.9 (0.38)	8.9E+00
	RVX-208 (L2)	-5.8 (0.55)	2.1E+04		RVX-208 (L2)	-5.7 (0.29)	4.7E+06
	JQ1 (L1)	-6.0 (0.42)	1.4E+03		JQ1 (L1)	-14.7 (0.54)	1.2E+17
	ZINC95559591 (L10)	-5.1 (0.29)	1.7E+06		Pevonedistat (L17)	-6.5 (0.37)	1.1E+16
	Tigecycline (L26)	-5.6 (0.32)	3.7E+05		Tigecycline (L26)	-7.4 (0.53)	2.1E+06
	Linezolid (L27)	-4.6 (0.28)	3.8E+02		Nafamostat (L24)	-3.6 (0.48)	4.2E+09
	Pevonedistat (L17)	-6.0 (0.44)	1.4E+03		Linezolid (L27)	-2.5 (0.38)	4.4E+05
	Apicidin (L5)	-2.5 (0.42)	4.6E+01		Apicidin (L5)	-5.3 (0.37)	4.1E+05
	Nafamostat (L24)	-3.2 (0.35)	9.6E+00		Chloramphenicol (L25)	-5.2 (0.29)	3.8E+03
	Camostat (L23)	-2.9 (0.49)	1.4E+06		Camostat (L23)	-7.2 (0.45)	2.4E+07
	4E2RCat (L19)	-6.1 (0.34)	2.2E-01		Sapanisertib (L14)	-10.9 (0.42)	1.4E+10
e	4E2RCat (L19)	-2.2 (0.38)	1.2E+01	e	4E2RCat (L19)	-1.8 (0.28)	1.2E-03
I	4TPW-33R	-1.0 (0.29)	1.2E+01	I	Camostat (L23)	-0.84 (0.22)	1.1E-03
F	Tigecycline (L26)	-2.2 (0.30)	1.7E+00	F	Tigecycline (L26)	-1.4 (0.40)	1.8E+02
4	Camostat (L23)	-1.9 (0.22)	1.7E-03	4	Lisinopril (L22)	-2.6 (0.28)	5.7E+00
E	Lisinopril (L22)	-1.7 (0.26)	1.4E-03	E	JQ1 (L1)	-2.1 (0.24)	5.2E+00

-1	Tomivosertib (L20)	-2.4 (0.42)	8.7E+00	-2	ABBV-744 (L11)	-2.7 (0.29)	3.6E+00
	Nafamostat (L24)	-1.8 (0.25)	1.0E+01		ZINC95559591 (L10)	-3.0 (0.31)	1.1E-01
	Silmitasertib (L3)	-2.2 (0.25)	4.4E-02		Pevonedistat (L17)	-2.5 (0.30)	7.9E+01
	RVX-208 (L2)	-2.0 (0.22)	1.3E+01		Silmitasertib (L3)	-1.4 (0.32)	1.2E-02
	Chloramphenicol (L25)	-1.3 (0.26)	3.4E-06				
E F 1 A 1 - 1	Ternatin-4 (L18)	-2.6 (0.37)	5.1E+04	E F 1 A 1 - 2	Ternatin-4 (L18)	-1.8 (0.49)	2.7E-006
	Pevonedistat (L17)	-2.7 (0.25)	1.0E+05		Pevonedistat (L17)	-2.8 (0.55)	3.9E+04
	4E2RCat (L19)	-2.0 (0.37)	5.4E+02		JQ1 (L1)	-5.3 (0.34)	1.1E-01
	Ponatinib (L7)	-2.3 (0.26)	8.6E+09		ZINC95559591 (L10)	-6.1 (0.27)	3.8E+05
	Nafamostat (L24)	-5.3 (0.41)	3.8E+13		Chloramphenicol (L25)	-5.5 (0.41)	3.9E+02
	Lisinopril (L22)	-4.9 (0.35)	9.3E+00		Camostat (L23)	-7.2 (0.46)	8.3E+02
	Tigecycline (L26)	-1.2 (0.29)	5.4E+01		ABBV-744 (L11)	-3.6 (0.23)	3.6E-04
	ZINC95559591 (L10)	-4.5 (0.52)	1.4E+08		Nafamostat (L24)	-3.4 (0.24)	7.1E+00
	Camostat (L23)	-4.0 (0.55)	8.8E+08		4E2RCat (L19)	-5.4 (0.34)	1.44E+02
	Tomivosertib (L20)	-2.6 (0.35)	1.4E-03		CPI-0610 (L13)	-4.0 (0.53)	2.5E+02
	Silmitasertib (L3)	-3.7 (0.28)	3.3E-01		Ponatinib (L7)	-2.4 (0.33)	6.2E+03
C S K 2 2	Silmitasertib (L3)	-8.6 (0.56)	1.3E+09	e I F 4 E 2	Zotatifin (L16)	-3.5 (0.31)	1.8E-01
	Pevonedistat (L17)	-8.9 (0.54)	8.4E+05		ABBV-744 (L11)	-5.6 (0.55)	1.1E+12
	ABBV-744 (L11)	-7.0 (0.41)	1.7E+06		Pevonedistat (L17)	-4.6 (0.37)	1.4E+11
	Verdinexor (L28)	-4.6 (0.39)	3.7E+02		Camostat (L23)	-3.3 (0.34)	7.91E+01
	Chloramphenicol (L25)	-10.4 (0.50)	4.1E+02		Tigecycline (L26)	-1.9 (0.46)	4.1E+02
	4E2RCat (L19)	-6.2 (0.31)	9.1E+07		4E2RCat (L19)	-5.8 (0.28)	1.7E+02
	TMCB (L4)	-6.6 (0.55)	1.3E+03		Nafamostat (L24)	-4.3 (0.31)	7.5E+03
	Tomivosertib (L20)	-9.3 (0.38)	5.1E+06		Ponatinib (L7)	-3.7 (0.40)	1.1E+02
	Sapanisertib (L14)	-6.3 (0.55)	3.6E+11		Lisinopril (L22)	-4.7 (0.20)	2.5E-02
	Camostat (L23)	-2.2 (0.28)	3.9E+01		Chloramphenicol (L25)	-3.1 (0.33)	8.6E+06
	Nafamostat (L24)	-4.0 (0.36)	1.4E+05		ZINC95559591 (L10)	-2.4 (0.34)	1.6E+04
	Lisinopril (L22)	-1.8 (0.39)	6.0E-02				
H D A C 2	Valproic Acid (L6)	-12.8 (0.34)	1.3E+10	M N K 2	Tomivosertib (L20)	-13.7 (0.27)	5.5E+04
	4LY1-20Y	-11.4 (0.38)	5.1E+11		Tigecycline (L26)	-7.4 (0.29)	1.8E+04
	6WBZ-TV7	-7.2 (0.52)	2.7E+05		Camostat (L23)	-7.1 (0.38)	2.7E+03
	4E2RCat (L19)	-12.0 (0.18)	2.6E+13		Lisinopril (L22)	-11.5 (0.55)	4.8E+09
	Linezolid (L27)	-11.0 (0.44)	5.3E+16		Nafamostat (L24)	-7.6 (0.55)	1.9E+12
	Ponatinib (L7)	-9.6 (0.54)	1.6E+09		Linezolid (L27)	-7.2 (0.37)	1.3E+04
	Pevonedistat (L17)	-9.0 (0.37)	1.4E+05		Pevonedistat (L17)	-7.5 (0.44)	1.8E+09
	Nafamostat (L24)	-13.3 (0.45)	4.3E+08		ZINC95559591 (L10)	-7.0 (0.33)	1.1E+06
	XL413 (L8)	-7.8 (0.44)	1.3E+05		ABBV-744 (L11)	-6.9 (0.55)	1.9E+07
					Zotatifin (L16)	-10.0 (0.40)	1.2E+03
D N M T	Azacididine (L9)	-4.0 (0.46)	2.4E+09	D N M T	Azacididine (L9)	-6.9 (0.53)	2.6E+12
	XL413 (L8)	-6.9 (0.33)	5.8E+01		XL413 (L8)	-4.2 (0.39)	1.1E+00
	Tigecycline (L26)	-10.9 (0.28)	2.7E+05		ABBV-744 (L11)	-4.8 (0.35)	5.8E+07
	Camostat (L23)	-6.0 (0.27)	2.9E+06		Tigecycline (L26)	-7.1 (0.46)	3.4E+15

1 - 1	Nafamostat (L24)	-5.1 (0.36)	6.6E+03	1 - 2	Nafamostat (L24)	-5.2 (0.34)	2.34E+02
	Chloramphenicol (L25)	-2.5 (0.28)	9.4E-04		Ponatinib (L7)	-6.0 (0.56)	2.4E+07
	ABBV-744 (L11)	-2.5 (0.23)	2.20E+06		Lisinopril (L22)	-7.3 (0.48)	4.7E+10
	Pevonedistat (L17)	-3.3 (0.37)	6.7E+01		ZINC95559591 (L10)	-5.3 (0.50)	2.1E+11
	JQ1 (L1)	-4.5 (0.30)	4.0E-02		Camostat (L23)	-4.6 (0.37)	5.6E+00
	Ponatinib (L7)	-3.5 (0.32)	2.2E+09		4E2RCat (L19)	-3.1 (0.45)	1.7E+02
	ZINC95559691	-6.0 (0.18)	1.6E+04		Verdinexor (L28)	-9.5 (0.49)	4.8E+08
	4E2RCat (L19)	-5.2 (0.34)	6.1E+09		Pevonedistat (L17)	-6.0 (0.31)	3.5E+05
	CPI-0610 (L13)	-2.1 (0.43)	2.6E+01				
	Tomivosertib (L20)	-2.0 (0.35)	1.2E-03				
	RVX-208 (L2)	-8.5 (0.33)	4.3E+06				

The identified list of drugs with higher free energy barriers shows chemical similarity in a few of the cases. For example, BRD2/BRD4 (PDB: 3ONI/3MXF) domain exhibits a higher binding affinity with JQ1, dBET6 and ABBV-744, where JQ1 is a substructure of dBET6 and hence shows higher chemical similarity. However, unlike BRD2/BRD4, with respect to HDAC2 (PDB: 4LY1), the control drug 20Y and Nafamostat, are found to exhibit higher free energy barriers without higher chemical similarity.

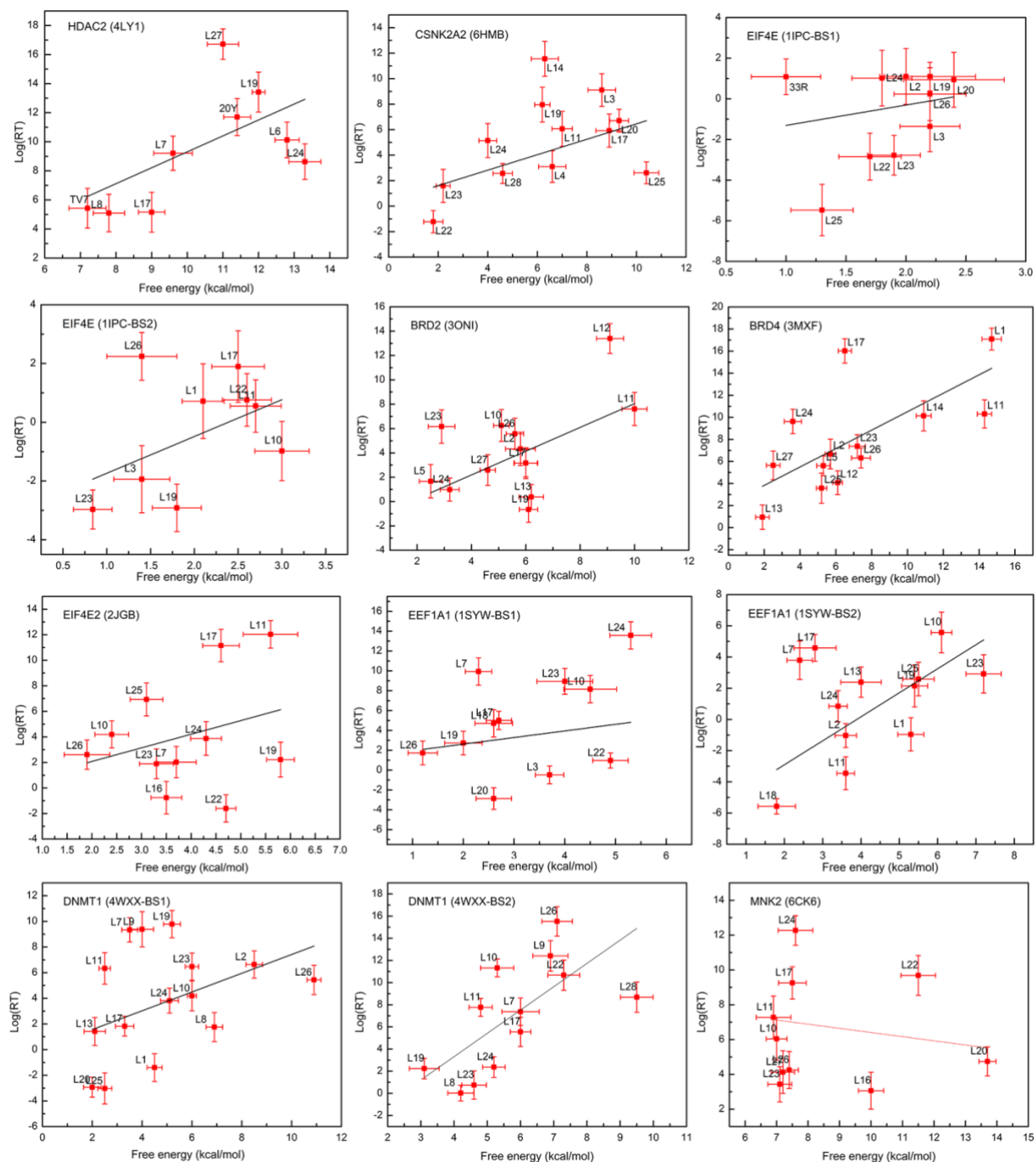


Figure 5. Residence time versus free energy barrier for all the protein-drug systems. Here, the RT data is fitted linear equation ($y = mX + C$). The fitting parameters m and C values are given here: 4LY1 ($m=1.008$ and $C= -1.515$); 6HMB ($m=0.574$ and $C= 0.542$); 1IPC-BS1 ($m=1.008$ and $C= -2.310$); 1IPC-BS2 ($m=1.234$ and $C= -2.996$); 3ONI ($m=0.962$ and $C= -1.605$); 3MXF

($m=0.849$ and $C= 2.026$); 1SYW-BS1 ($m=0.652$ and $C= 1.224$); 1SYW-BS2 ($m=1.530$ and $C= -5.943$); 4WXX-BS1 ($m=0.708$ and $C= 0.152$); 4WXX-BS2 ($m=2.149$ and $C=-5.216$); 6CK6 ($m=-0.233$ and $C= 8.738$).

To corroborate the binding nature evident from free energy barriers, we have also estimated the residence time for the ligands using equations 1 and 2. The residence time, RT, is calculated for all the 10 replicas, and the resultant average values are reported in Table 2. Most of the cases the estimated RT values are found to be high for the ligands, which have shown higher energy barriers in the FES. Furthermore, the RT values are found to increase linearly with increasing barrier height. This linear correlation is clearly shown in the log (RT) vs. FES barrier data plots in Figure 5. The linear relationship is more clearly shown for the proteins-ligand where the energy barriers are not close. We have also observed that the residence time for the control drugs and previously literature-suggested drugs are found to exhibit higher residence time and free energy barriers. The proteins with inhibitor bound crystal structures, i.e., control drugs (20Y:HDAC2, 4E2RCat:IF4E, Tomivosertib: MNK2, Sildenafil: CSK22, JQ1: MXF) were found to be in the group of drugs which exhibit higher residence time. It is evident from the interaction energy calculations from docking, MD, and free-energy calculations that these higher residence times correspond to the increase in the protein-drug non-bonded interactions and subsequently confirm the drug's higher stability at the binding site (Figure S5). Thus, in contrast to the prevailing focus on virtual screening of a large number of drugs through molecular docking followed by MM/PBSA binding energy calculations, we propose a multi-target based drug repurposing strategy with robust free energy calculations using enhanced sampling.

Conclusion

To date, no FDA approved specific drugs have been found for COVID-19. However, a range of antivirals and other drugs have shown inhibitory behavior against SARS-CoV-2 in-vitro and in clinical conditions. These drugs either target the viral-related components or the host proteins that regulate the viral pathogenesis. Here, we have performed an extensive drug repurposing exercise of the host transcription and translation regulatory proteins for a list of selective drugs that are anticancer compounds, anti-coagulants, antibiotics (bacterial infections), or immunosuppressive in nature. In this study, we report the high throughput analysis of these literature-derived repurposing drug candidates that can be used to target the genetic regulators known to interact with viral proteins based on experimental and interactome studies. On the basis of the known/predictor inhibitor binding sites for these proteins, we extrapolated our multiple protein-multiple ligand approach to identify drugs with approved pharmacological properties that can bind favorably and show higher binding energy as compared to the existing/proposed inhibitors. The molecular docking and simulation analysis revealed that the protein-drug complexes possess stable conformations and differential patterns of interaction energies and hydrogen bonds. It is possible that multiple drugs can bind the same binding site/cavity in a protein with varying binding affinity and interactions. Moreover, we elucidate the difference in the binding energy for these protein-drug complexes using free energy barrier calculations. From the perspective of the efficiency the average computational time per system is reasonably less for metadynamics simulations to acquire the binding free energy in comparison to brute force MD. Even in comparison to other enhanced sampling methods like umbrella sampling, free energy perturbation which require sufficiently long time to get the full free energy profile, metadynamics is comparatively faster and provides quick convergence. Thus, our approach is efficient and less costly to get binding free energy for ligand-

protein systems. We must note that in addition to the pre-clinical and FDA approved drugs that target specific regulatory proteins, a range of chemical compounds (Nafamostat, Chloramphenicol, Ponatinib) binds to the other gene transcription and translation regulatory protein with higher affinity and may harbour the potential for therapeutic uses. There is a rapidly growing interest in the development of combination therapy for COVID19 to target multiple enzymes/pathways. Our in-silico approach would be envisaged to expedite the process of generating leads for experimental screening for rapid drug repurposing against SARS-CoV-2 interacting host proteins.

ASSOCIATED CONTENT

Supporting Information.

Binding sites representation, Average value of interaction energy for different protein-ligand systems, Hydrogen bond details.

The following files are available free of charge.

Protein-ligand conformations corresponding to the FES minimum for article-based, control and identified repurposing drugs. (PDF)

Protein-ligand structures from MD simulation that are used for further FE calculations (PDF)

AUTHOR INFORMATION

Corresponding Author

*Sudip Roy (sudip@prescience.in) and * Jayant K. Singh (jayantks@iitk.ac.in)

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript. ‡These authors contributed equally.

Notes

The authors declare no competing financial interests.

ACKNOWLEDGMENT

D.P. thanks IITK for postdoctoral fellowship. A.K. thanks IITK and Prescience for financial support. JKS would like to acknowledge SERB/DST, India for the partial support. Prescience acknowledges the CSR funding of GST IN. Authors thank IITK for the supercomputing facility.

REFERENCES

Abdelrahman, Z., Li, M. & Wang, X. (2020). Comparative Review of SARS-CoV-2, SARS-CoV, MERS-CoV, and Influenza A Respiratory Viruses. *Frontiers in Immunology 11*: 2309.

Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B. & Lindahl, E. (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX 1-2*: 19-25.

Ahmed, S. S. S. J., Paramasivam, P., Raj, K., Kumar, V., Murugesan, R. & Ramakrishnan, V. (2020). Regulatory Cross Talk Between SARS-CoV-2 Receptor Binding and Replication Machinery in the Human Host. *Frontiers in Physiology 11*: 802.

Barducci, A., Bussi, G. & Parrinello, M. (2008). Well-tempered metadynamics: A smoothly converging and tunable free-energy method. *Physical Review Letters 100*: 1-4.

Bussi, G., Donadio, D. & Parrinello, M. (2007). Canonical sampling through velocity rescaling. *The Journal of Chemical Physics 126*: 014101.

Cao, L., Goreshnik, I., Coventry, B., Case, J. B., Miller, L., Kozodoy, L., Chen, R. E., Carter, L., Walls, A. C., Park, Y.-J., Strauch, E.-M., Stewart, L., Diamond, M. S., Veessler, D. & Baker, D. (2020). De novo design of picomolar SARS-CoV-2 miniprotein inhibitors. *Science* 370: 426.

Du, L., He, Y., Zhou, Y., Liu, S., Zheng, B.-J. & Jiang, S. (2009). The spike protein of SARS-CoV—a target for vaccine and therapeutic development. *Nature Reviews Microbiology* 7: 226-236.

El Bairi, K., Trapani, D., Petrillo, A., Le Page, C., Zbakh, H., Daniele, B., Belbaraka, R., Curigliano, G. & Afqir, S. (2020). Repurposing anticancer drugs for the management of COVID-19. *European Journal of Cancer* 141: 40-61.

Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H. & Pedersen, L. G. (1995). A smooth particle mesh Ewald method. *The Journal of Chemical Physics* 103: 8577-8593.

Filippakopoulos, P., Qi, J., Picaud, S., Shen, Y., Smith, W. B., Fedorov, O., Morse, E. M., Keates, T., Hickman, T. T., Felletar, I., Philpott, M., Munro, S., McKeown, M. R., Wang, Y., Christie, A. L., West, N., Cameron, M. J., Schwartz, B., Heightman, T. D., La Thangue, N., French, C. A., Wiest, O., Kung, A. L., Knapp, S. & Bradner, J. E. (2010). Selective inhibition of BET bromodomains. *Nature* 468: 1067-1073.

Gordon, D. E., Hiatt, J., Bouhaddou, M., Rezelj, V. V., Ulferts, S., Braberg, H., Jureka, A. S., Obernier, K., Guo, J. Z., Batra, J., Kaake, R. M., Weckstein, A. R., Owens, T. W., Gupta, M., Pourmal, S., Titus, E. W., Cakir, M., Soucheray, M., McGregor, M., Cakir, Z., Jang, G., O'Meara, M. J., Tummino, T. A., Zhang, Z., Foussard, H., Rojic, A., Zhou, Y., Kuchenov, D., Hüttenhain, R., Xu, J., Eckhardt, M., Swaney, D. L., Fabius, J. M., Ummadi, M., Tutuncuoglu, B., Rathore, U., Modak, M., Haas, P., Haas, K. M., Naing, Z. Z. C., Pulido, E. H., Shi, Y., Barrio-Hernandez,

I., Memon, D., Petsalaki, E., Dunham, A., Marrero, M. C., Burke, D., Koh, C., Vallet, T., Silvas, J. A., Azumaya, C. M., Billesbølle, C., Brilot, A. F., Campbell, M. G., Diallo, A., Dickinson, M. S., Diwanji, D., Herrera, N., Hoppe, N., Kratochvil, H. T., Liu, Y., Merz, G. E., Moritz, M., Nguyen, H. C., Nowotny, C., Puchades, C., Rizo, A. N., Schulze-Gahmen, U., Smith, A. M., Sun, M., Young, I. D., Zhao, J., Asarnow, D., Biel, J., Bowen, A., Braxton, J. R., Chen, J., Chio, C. M., Chio, U. S., Deshpande, I., Doan, L., Faust, B., Flores, S., Jin, M., Kim, K., Lam, V. L., Li, F., Li, J., Li, Y.-L., Li, Y., Liu, X., Lo, M., Lopez, K. E., Melo, A. A., Moss, F. R., Nguyen, P., Paulino, J., Pawar, K. I., Peters, J. K., Pospiech, T. H., Safari, M., Sangwan, S., Schaefer, K., Thomas, P. V., Thwin, A. C., Trenker, R., Tse, E., Tsui, T. K. M., Wang, F., Whitis, N., Yu, Z., Zhang, K., Zhang, Y., Zhou, F., Saltzberg, D., Hodder, A. J., Shun-Shion, A. S., Williams, D. M., White, K. M., Rosales, R., Kehrer, T., Miorin, L., Moreno, E., Patel, A. H., Rihn, S., Khalid, M. M., Vallejo-Gracia, A., Fozouni, P., Simoneau, C. R., Roth, T. L., Wu, D., Karim, M. A., Ghousaini, M., Dunham, I., Berardi, F., Weigang, S., Chazal, M., Park, J., Logue, J., McGrath, M., Weston, S., Haupt, R., Hastie, C. J., Elliott, M., Brown, F., Burness, K. A., Reid, E., Dorward, M., Johnson, C., Wilkinson, S. G., Geyer, A., Giesel, D. M., Baillie, C., Raggett, S., Leech, H., Toth, R., Goodman, N., Keough, K. C., Lind, A. L., Klesh, R. J., Hemphill, K. R., Carlson-Stevermer, J., Oki, J., Holden, K., Maures, T., Pollard, K. S., Sali, A., Agard, D. A., Cheng, Y., Fraser, J. S., Frost, A., Jura, N., Kortemme, T., Manglik, A., Southworth, D. R., Stroud, R. M., Alessi, D. R., Davies, P., Frieman, M. B., Ideker, T., Abate, C., Jouvenet, N., Kochs, G., Shoichet, B., Ott, M., Palmarini, M., Shokat, K. M., García-Sastre, A., Rassen, J. A., Grosse, R., Rosenberg, O. S., Verba, K. A., Basler, C. F., Vignuzzi, M., Peden, A. A., Beltrao, P. & Krogan, N. J. (2020). Comparative host-coronavirus protein interaction networks reveal pan-viral disease mechanisms. *Science*: eabe9403.

Gordon, D. E., Jang, G. M., Bouhaddou, M., Xu, J., Obernier, K., White, K. M., O'Meara, M. J., Rezelj, V. V., Guo, J. Z., Swaney, D. L., Tummino, T. A., Hüttenhain, R., Kaake, R. M., Richards, A. L., Tutuncuoglu, B., Foussard, H., Batra, J., Haas, K., Modak, M., Kim, M., Haas, P., Polacco, B. J., Braberg, H., Fabius, J. M., Eckhardt, M., Soucheray, M., Bennett, M. J., Cakir, M., McGregor, M. J., Li, Q., Meyer, B., Roesch, F., Vallet, T., Mac Kain, A., Miorin, L., Moreno, E., Naing, Z. Z. C., Zhou, Y., Peng, S., Shi, Y., Zhang, Z., Shen, W., Kirby, I. T., Melnyk, J. E., Chorba, J. S., Lou, K., Dai, S. A., Barrio-Hernandez, I., Memon, D., Hernandez-Armenta, C., Lyu, J., Mathy, C. J. P., Perica, T., Pilla, K. B., Ganesan, S. J., Saltzberg, D. J., Rakesh, R., Liu, X., Rosenthal, S. B., Calviello, L., Venkataramanan, S., Liboy-Lugo, J., Lin, Y., Huang, X.-P., Liu, Y., Wankowicz, S. A., Bohn, M., Safari, M., Ugur, F. S., Koh, C., Savar, N. S., Tran, Q. D., Shengjuler, D., Fletcher, S. J., O'Neal, M. C., Cai, Y., Chang, J. C. J., Broadhurst, D. J., Klippsten, S., Sharp, P. P., Wenzell, N. A., Kuzuoglu-Ozturk, D., Wang, H.-Y., Trenker, R., Young, J. M., Cavero, D. A., Hiatt, J., Roth, T. L., Rathore, U., Subramanian, A., Noack, J., Hubert, M., Stroud, R. M., Frankel, A. D., Rosenberg, O. S., Verba, K. A., Agard, D. A., Ott, M., Emerman, M., Jura, N., von Zastrow, M., Verdin, E., Ashworth, A., Schwartz, O., d'Enfert, C., Mukherjee, S., Jacobson, M., Malik, H. S., Fujimori, D. G., Ideker, T., Craik, C. S., Floor, S. N., Fraser, J. S., Gross, J. D., Sali, A., Roth, B. L., Ruggero, D., Taunton, J., Kortemme, T., Beltrao, P., Vignuzzi, M., García-Sastre, A., Shokat, K. M., Shoichet, B. K. & Krogan, N. J. (2020). A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* 583: 459-468.

Guy, R. K., DiPaola, R. S., Romanelli, F. & Dutch, R. E. (2020). Rapid repurposing of drugs for COVID-19. *Science* 368: 829.

Harris, R., Olson, A. J. & Goodsell, D. S. (2008). Automated prediction of ligand-binding sites in proteins. *Proteins: Structure, Function, and Bioinformatics* 70: 1506-1517.

Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. (1997). LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry* 18: 1463-1472.

Laio, A. & Parrinello, M. (2002). Escaping free-energy minima. *Proceedings of the National Academy of Sciences* 99: 12562.

Li, J., Guo, M., Tian, X., Wang, X., Yang, X., Wu, P., Liu, C., Xiao, Z., Qu, Y., Yin, Y., Wang, C., Zhang, Y., Zhu, Z., Liu, Z., Peng, C., Zhu, T. & Liang, Q. Virus-Host Interactome and Proteomic Survey Reveal Potential Virulence Factors Influencing SARS-CoV-2 Pathogenesis. *Med.*

Liu, B., Li, M., Zhou, Z., Guan, X. & Xiang, Y. (2020). Can we use interleukin-6 (IL-6) blockade for coronavirus disease 2019 (COVID-19)-induced cytokine release syndrome (CRS)? *Journal of Autoimmunity*: 102452.

MacKerell, a. D., Banavali, N. & Foloppe, N. (2001). Development and current status of the CHARMM force field for nucleic acids. *Biopolymers* 56: 257-265.

Mandala, V. S., McKay, M. J., Shcherbakov, A. A., Dregni, A. J., Kolocouris, A. & Hong, M. (2020). Structure and drug binding of the SARS-CoV-2 envelope protein transmembrane domain in lipid bilayers. *Nature Structural & Molecular Biology*.

Marcelo Zaldini, H., Suellen Melo, T. C., Diogo Rodrigo, M. M., Walter Filgueira de Azevedo, J. & Ana Cristina Lima, L. (2010). Halogen Atoms in the Modern Medicinal Chemistry: Hints for the Drug Design. *Current Drug Targets* 11: 303-314.

Mehta, P., McAuley, D. F., Brown, M., Sanchez, E., Tattersall, R. S., Manson, J. J. & Collaboration, H. A. S. (2020). COVID-19: consider cytokine storm syndromes and immunosuppression. *Lancet (London, England)* 395: 1033.

Menéndez, C. A., Byléhn, F., Perez-Lemus, G. R., Alvarado, W. & de Pablo, J. J. (2020). Molecular characterization of ebselen binding activity to SARS-CoV-2 main protease. *Science Advances* 6: eabd0345.

Moreira, R. A., Chwastyk, M., Baker, J. L., Guzman, H. V. & Poma, A. B. (2020). Quantitative determination of mechanical stability in the novel coronavirus spike protein. *Nanoscale* 12: 16409-16413.

Moreira, R. A., Guzman, H. V., Boopathi, S., Baker, J. L. & Poma, A. B. (2020). Characterization of structural and energetic differences between conformations of the SARS-CoV-2 spike protein. *Materials* 13: 5362.

Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S. & Olson, A. J. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of Computational Chemistry* 30: 2785-2791.

Pardo, J., Shukla, A. M., Chamarthi, G. & Gupte, A. (2020). The journey of remdesivir: from Ebola to COVID-19. *Drugs in context* 9: 2020-2024-2014.

Parrinello, M. & Rahman, A. (1981). Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics* 52: 7182-7190.

Pennington, L. D. & Moustakas, D. T. (2017). The Necessary Nitrogen Atom: A Versatile High-Impact Design Element for Multiparameter Optimization. *Journal of Medicinal Chemistry* 60: 3552-3579.

Pierre, F., Chua, P. C., O'Brien, S. E., Siddiqui-Jain, A., Bourbon, P., Haddach, M., Michaux, J., Nagasawa, J., Schwaebe, M. K., Stefan, E., Vialettes, A., Whitten, J. P., Chen, T. K., Darjania, L., Stansfield, R., Anderes, K., Bliesath, J., Drygin, D., Ho, C., Omori, M., Proffitt, C., Streiner, N., Trent, K., Rice, W. G. & Ryckman, D. M. (2011). Discovery and SAR of 5-(3-chlorophenylamino)benzo[c][2,6]naphthyridine-8-carboxylic acid (CX-4945), the first clinical stage inhibitor of protein kinase CK2 for the treatment of cancer. *Journal of Medicinal Chemistry* 54: 635-654.

Plante, J. A., Liu, Y., Liu, J., Xia, H., Johnson, B. A., Lokugamage, K. G., Zhang, X., Muruato, A. E., Zou, J., Fontes-Garfias, C. R., Mirchandani, D., Scharton, D., Bilello, J. P., Ku, Z., An, Z., Kalveram, B., Freiberg, A. N., Menachery, V. D., Xie, X., Plante, K. S., Weaver, S. C. & Shi, P.-Y. (2020). Spike mutation D614G alters SARS-CoV-2 fitness. *Nature*.

Poduri, R., Joshi, G. & Jagadeesh, G. (2020). Drugs targeting various stages of the SARS-CoV-2 life cycle: Exploring promising drugs for the treatment of Covid-19. *Cellular Signalling* 74: 109721.

Poma, A. B., Chwastyk, M. & Cieplak, M. (2015). Polysaccharide–protein complexes in a coarse-grained model. *The Journal of Physical Chemistry B* 119: 12028-12041.

Reich, S. H., Sprengeler, P. A., Chiang, G. G., Appleman, J. R., Chen, J., Clarine, J., Eam, B., Ernst, J. T., Han, Q., Goel, V. K., Han, E. Z. R., Huang, V., Hung, I. N. J., Jemison, A., Jessen, K. A., Molter, J., Murphy, D., Neal, M., Parker, G. S., Shaghafi, M., Sperry, S., Staunton, J., Stumpf,

C. R., Thompson, P. A., Tran, C., Webber, S. E., Wegerski, C. J., Zheng, H. & Webster, K. R. (2018). Structure-based Design of Pyridone–Aminal eFT508 Targeting Dysregulated Translation by Selective Mitogen-activated Protein Kinase Interacting Kinases 1 and 2 (MNK1/2) Inhibition. *Journal of Medicinal Chemistry* 61: 3516-3540.

Sacco, M. D., Ma, C., Lagarias, P., Gao, A., Townsend, J. A., Meng, X., Dube, P., Zhang, X., Hu, Y., Kitamura, N., Hurst, B., Tarbet, B., Marty, M. T., Kolocouris, A., Xiang, Y., Chen, Y. & Wang, J. (2020). Structure and inhibition of the SARS-CoV-2 main protease reveals strategy for developing dual inhibitors against M^{pro} and cathepsin L. *Science Advances*: eabe0751.

Sadanandam, N., Debabrata, P., Mohd Aamir, K., Sudip, R. & Jayant, S. (2020). Potential Drug Candidates for SARS-CoV-2 Using Computational Screening and Enhanced Sampling Methods. *ChemRxiv*.

Saini, K. S., Lanza, C., Romano, M., de Azambuja, E., Cortes, J., de las Heras, B., de Castro, J., Lamba Saini, M., Loibl, S., Curigliano, G., Twelves, C., Leone, M. & Patnaik, M. M. (2020). Repurposing anticancer drugs for COVID-19-induced inflammation, immune dysfunction, and coagulopathy. *British Journal of Cancer* 123: 694-697.

Schäfer, A. & Baric, R. S. (2017). Epigenetic Landscape during Coronavirus Infection. *Pathogens (Basel, Switzerland)* 6: 8.

Singh, T. U., Parida, S., Lingaraju, M. C., Kesavan, M., Kumar, D. & Singh, R. K. (2020). Drug repurposing approach to fight COVID-19. *Pharmacological reports : PR*: 1-30.

Soukarieh, F., Nowicki, M. W., Bastide, A., Pöyry, T., Jones, C., Dudek, K., Patwardhan, G., Meullenet, F., Oldham, N. J., Walkinshaw, M. D., Willis, A. E. & Fischer, P. M. (2016). Design of nucleotide-mimetic and non-nucleotide inhibitors of the translation initiation factor eIF4E: Synthesis, structural and functional characterisation. *European Journal of Medicinal Chemistry* 124: 200-217.

Tang, Y., Liu, J., Zhang, D., Xu, Z., Ji, J. & Wen, C. (2020). Cytokine Storm in COVID-19: The Current Evidence and Treatment Strategies. *Frontiers in Immunology* 11: 1708-1708.

Tharappel, A. M., Samrat, S. K., Li, Z. & Li, H. (2020). Targeting Crucial Host Factors of SARS-CoV-2. *ACS Infectious Diseases* 6: 2844-2865.

Trezza, A., Iovinelli, D., Santucci, A., Prischi, F. & Spiga, O. (2020). An integrated drug repurposing strategy for the rapid identification of potential SARS-CoV-2 viral inhibitors. *Scientific Reports* 10: 13866.

Tutuncuoglu, B., Cakir, M., Batra, J., Bouhaddou, M., Eckhardt, M., Gordon, D. E. & Krogan, N. J. (2020). The Landscape of Human Cancer Proteins Targeted by SARS-CoV-2. *Cancer Discovery* 10: 916.

V'kovski, P., Kratzel, A., Steiner, S., Stalder, H. & Thiel, V. (2020). Coronavirus biology and replication: implications for SARS-CoV-2. *Nature Reviews Microbiology*.

Webb, B. & Sali, A. (2016). Comparative Protein Structure Modeling Using MODELLER. *Current Protocols in Bioinformatics* 54: 5.6.1-5.6.37.

Wilcken, R., Zimmermann, M. O., Lange, A., Joerger, A. C. & Boeckler, F. M. (2013). Principles and Applications of Halogen Bonding in Medicinal Chemistry and Chemical Biology. *Journal of Medicinal Chemistry* 56: 1363-1388.

Yang, J., Petitjean, S. J., Koehler, M., Zhang, Q., Dumitru, A. C., Chen, W., Derclaye, S., Vincent, S. P., Soumillon, P. & Alsteens, D. (2020). Molecular interaction and inhibition of SARS-CoV-2 binding to the ACE2 receptor. *Nature communications* 11: 1-10.

Zhou, Y., Hou, Y., Shen, J., Huang, Y., Martin, W. & Cheng, F. (2020). Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2. *Cell Discovery* 6: 14.

Zhou, Y., Wang, F., Tang, J., Nussinov, R. & Cheng, F. Artificial intelligence in COVID-19 drug repurposing. *The Lancet Digital Health*.

Ziebuhr, J. (2004). Molecular biology of severe acute respiratory syndrome coronavirus. *Current opinion in microbiology* 7: 412-419.

Zoete, V., Cuendet, M. A., Grosdidier, A. & Michielin, O. (2011). SwissParam: A fast force field generation tool for small organic molecules. *Journal of Computational Chemistry* 32: 2359-2368.