The Mode of Therapeutic Action of 2-deoxy D-glucose: Anti-Viral or Glycolysis Blocker?

Boyli Ghosh[a], Sudip Roy[a], Jayant K Singh[a,b]

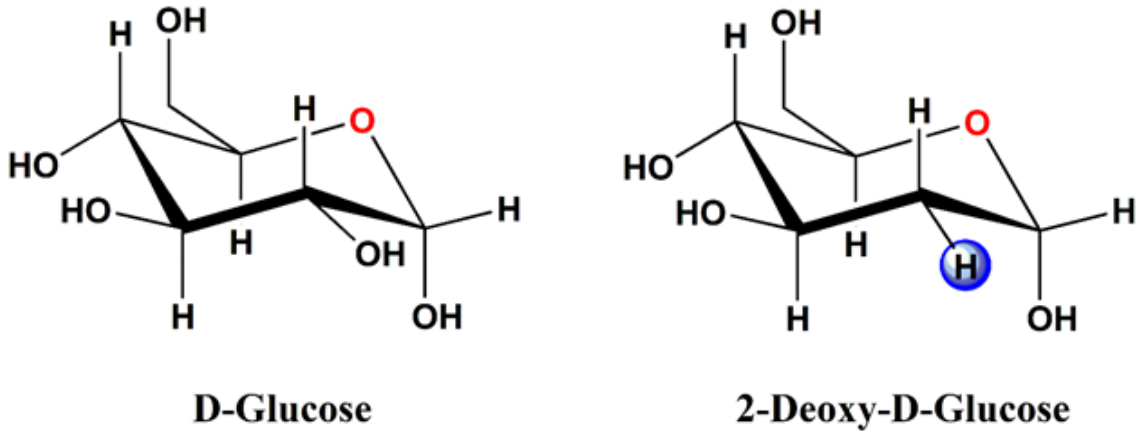
[a]Prescience Insilico Private Limited, Bangalore, India; [b]Department of Chemical Engineering, Indian Institute of Technology, Kanpur, India

Abstract

Recently an anti-COVID-19 therapeutic application of the drug 2-deoxy-D-glucose (2-DG) an analogue of glucose has been developed in collaboration between Institute of Nuclear Medicine and Allied Sciences (INMAS), India, Defence Research and Development Organisation (DRDO), India, and Dr Reddy's Laboratories (DRL), India. As per the reports 2-DG is effective against SARS-COV-2. Publication of phase 2 and phase 3 clinical trial data is pending. However, it has been shown that 2-DG reduces the supplemental oxygen dependence on covid-19 infected patients and make their recovery faster. The present outbreak of Covid-19 infection due to SARS-CoV-2, a virus from the coronavirus family, has become a major menace to human being. As the understanding of the mechanism of the therapeutic action of 2-DG on SARS-CoV-2 infected hosts is missing, in this work we have studied the possible inhibitory interaction of the drug with two different pathways (a) with non-structured viral proteins involved in translation and replication of SARS-CoV-2 and (b) its inhibition mechanism of the glycolysis pathway. We have used our fully automated novel drug designing platform with state-of-the-art free energy of binding calculator PR*in*MTML-ESS to evaluate the role of 2-DG as an antiviral and glycolysis pathway inhibitor in SARS-CoV-2 affected humans. Docking, all atom molecular dynamic simulation and enhanced free energy sampling methods used in PR*in*MTML-ESS have predicted that 2-DG effectively reduced the replication of SARS-CoV-2 in human cell by reducing the glycolytic flux, by competitive inhibition of glucose in binding with the enzyme hexokinase. 2-DG is generally administered in covid patient along with other antivirals and steroid, hence it can be used as a mild clinical therapy which can reduce the viral replication, inflammation when given in the earlier stage of the disease.

Introduction

The present outbreak of Covid-19 infection due to SARS-CoV-2, a virus from the coronavirus family, has become a major threat to human healthcare. This occurrence of SARS-CoV-2 is the third highly pathogenic event and large-scale epidemic affecting the human population. It follows the severe acute respiratory syndrome coronavirus (SARS-CoV) in 2003 and the Middle East respiratory syndrome coronavirus (MERS-CoV) in 2012. SARS-CoV 2 was first reported in December 2019 and after that it has spread all over the world infecting almost 44 million people till date. Moreover, it has been fatal to more than 1.1 million people. The higher rate of transmission of SARS-CoV-2 compared to the SARS-CoV is most probably associated with S-glycoprotein in the receptor-binding domain (RBD) region. The spike S glycoprotein of coronavirus facilitates the binding to the target cell and hence the efficiency of receptor binding domain(RBD)/ angiotensin-converting enzyme 2 (ACE2) is determinant to SARS-COV 2 transmissibility.[1-3] It is a high time for all researchers in drug development to develop or repurpose FDA approved drug or an epitope to circumvent the current situation of pandemic. SARS-CoV2 is an envelope virus and it contains a positive-sense single-stranded RNA (ssRNA) genome (26–32 kb).[4] The genomics data of pathogens are very important to obtain information regarding potential targets relevant for therapy or diagnostics.[5] Hence in order to design and develop drug against such viral pathogens, it is necessary to start with data mining of viral genomes. The SARS-CoV-2 genome is made of less than 30000 nucleotides and contains genes for 29 different proteins and 10 open reading frame (ORF). The ORF1ab constitutes two third of the viral RNA and encodes as many as 16 non-structural proteins.[6] Some of the key proteins encoded by this gene are PL^{pro} (NSP2), $3CL^{pro}$ (NSP5), RdRp (NSP12), and O-methyl transferase (NSP16) which play a vital role in the replication and transcription. The structural proteins such as membrane protein (M), envelope protein (E), spike protein (S), nucleocapsid protein (N) and other auxiliary proteins are encoded in ORF2-10. The RNA gene is packed within the N protein and M, E and S protein make the viral coat.



Scheme 1. Structure of D-glucose and 2-Deoxy-D-Glucose(2-DG)

Further the entry protein i.e Spike protein in involved in host cell recognition and binds specifically to the Angiotensin-converting enzyme 2 (ACE2) mammalian receptor. Therefore, for therapeutic approach all the structured and non-structured viral protein can be considered as potential target. However, the non-structured proteins PL^{pro} (NSP2), $3CL^{pro}$ (NSP5), RNA-dependent RNA polymerase (NSP12), and O-methyl transferase (NSP16) that are responsible for viral replication as well as transcription are more essential to target. Again, the infectivity and transmission capacity of SARS-CoV-2 to host cell depends largely on the S glycoprotein.[2, 7] Preclinical studies have predicted that glucose plays a vital role in replication of SARS-CoV-2. SARS-CoV-2 increases the demand of glucose in cells by upregulating the metabolic processes in host cell. Two such important metabolic processes in human cells are glycolysis and glycosylation.[8] These features may have led to the idea that drugs inhibiting glycolysis and glycosylation might prove efficient in the context of SARS-CoV-2. One of the most widely studied glycolytic inhibitors is 2-deoxyglucose (2-DG), a synthetic glucose analogue in which the hydroxyl group at the second carbon atom is replaced by hydrogen (Scheme 1). Recently an anti-COVID-19 therapeutic application of the drug 2-deoxy-D-glucose (2-DG) has been developed by Institute of Nuclear Medicine and Allied Sciences (INMAS), a lab of Defence Research and Development Organisation (DRDO), in collaboration with Dr Reddy's Laboratories (DRL). 2-DG is found to be effective against SARS-COV-2 in in-vitro studies and based on clinical trial it has been shown that 2-DG reduces the supplemental oxygen dependence on covid-19 infected patients and make their recovery faster (https://pib.gov.in/PressReleasePage.aspx?PRID=1717007). Hence, to

understand the mechanism in two different competitive pathways of therapeutic action of 2-deoxy D-glucose on SARS-CoV-2 infected hosts, we need to take into consideration its interaction with non-structured viral proteins involved in translation and replication of SARS-CoV-2 on one hand and its inhibition mechanism of the glycolysis pathway. In this work, we have used our fully automated integrated platform **Prescience *in silico* Multi-Target Multi-Ligand Enhance Sampling Screening** (PR*in*MTML-ESS) to determine the mode of therapeutic action of 2-DG in SARS-CoV-2 treatment. PR*in*MTML-ESS is a combined computational approach where docking, molecular dynamic simulation and the free energy calculations using enhanced sampling methods are used to explore the energetics of target-ligand complexation associated with the protein's or an enzyme's binding site (including the ligand) in an explicit solvent. To explore the antiviral property of 2-DG its binding interaction has been studied using our platform with the non-structured viral protein PL^{pro} (NSP2), $3CL^{pro}$ (NSP5), RNA-dependent RNA polymerase (NSP12), and O-methyl transferase (NSP16) which plays a major role in replication/translation of virus. The effect of 2-DG on the glycolysis process that converts glucose into ATP has been probed using the same combined computational approach and targeting the enzyme hexokinase that phosphorylates glucose to glucose-6-phosphate. Due the structural similarities between 2-DG and glucose (Scheme 1), it is expected that 2-DG would act as a competitive inhibitor of glucose metabolism and thus might strongly affect the SARS-CoV-2 virus replication and activity, which largely depends on the ATP generated from glycolysis in human cells.

Result and Discussions

In our study, we have considered the possibility of 2-deoxy D-glucose (2-DG) working as an anti-viral drug and its role of a competitive inhibitor of glycolysis pathway. A key stage of the SARS-CoV-2 life cycle is the replication of the viral genome within the infected cells. It is a complex process involving the action of several viral and host proteins in order to perform RNA polymerization, proofreading and final capping. To investigate the antiviral property of 2-DG we have selected four viral proteins which are crucial actors of the replicatory machinery of SARS-CoV-2 as targets. The crystal structural data for these 4 viral proteins PL^{pro} (PDB ID: 6wuu)[9], $3CL^{pro}$ (PDB ID: 6lu7)[10], RNA-dependent RNA polymerase (PDB ID: 7bv1)[11], and O-methyl transferase (PDB ID: 6wkq)[12] of SARS-CoV-2 are available now. The interaction of 2-DG with all the four antiviral proteins was computationally investigated using PR*in*MTML-ESS platform

(https://www.prescience.in/prins3) which uses the combination of molecular docking at the binding site of protein followed by all atom molecular dynamic simulation of the best-docked pose in water for determining the stability of the protein-ligand bound structure and finally enhanced free energy sampling for better understand the binding interaction of the ligand with the protein. The same methodology has been used for screening large number of new chemical candidates for $3CL^{pro}$ in our earlier work[13] and the multi-target multi-ligand approach has been used in our high throughput analysis of literature derived repurposing drug candidates for SAR-CoV-2 that can be used to target the genetic regulators known to interact with viral proteins based on experimental and interactome studies[14] In our study, we have initially analysed the interaction of 2-DG with four viral proteins 3C-like protease ($3CL^{pro}$), non-structural protein 16 (Nsp16), papain-like cysteine protease (PL^{pro}) and RNA-dependent RNA polymerase (RdRp) which take part in replication and translation mechanism. Establishment of the viral replication and transcription complex (RTC) that includes, amongst others, RNA-processing and RNA-modifying enzymes and an RNA proofreading function essential for maintaining the integrity of the >30kb coronavirus genome[15] is crucial for virus replication and thus a promising target for antivirals against SARS-CoV-2. One of such target is $3CL^{pro}$ which resides in non-structural protein-5 (nsp5). $3CL^{pro}$ releases majority of the non-structural proteins from polyprotein and is crucial for viral life cycle. It's another role is inhibition of interferon signalling.[16] PL^{pro} releases non-structural protein (nsp1, nsp2, nsp3) and the amino terminus of nsp4 from the polyproteins pp1a and pp1ab. PL^{pro} helps SARS-CoV-2 in evading the host innate immune responses by stripping ubiquitin and ISG15 from host-cell proteins. Therefore, targeting PLpro with antiviral drugs may have an advantage for inhibiting viral replication and also inhibiting the dysregulation of signalling cascades in infected cells that may lead to cell death in surrounding, uninfected cells.[17] 2'-O-methyltransferase plays an essential role in immune evasion. Nsp16 achieves this functionality by mimic its human homolog, CMTr1 which methylates mRNA to enhance translation efficiency. However, Nsp16 requires a binding partner, Nsp10, to activate its enzymatic activity unlike CMTr1.[18] The most important part of the coronavirus RTC is RdRP residing in nsp12 and this viral protein is suggested as a promising drug target as it is a crucial enzyme in the virus life cycle both for replication of the viral genome but also for transcription of sgRNAs.[16] Molecular docking was started with automated generation of gridbox at binding site of the viral proteins PL^{pro} (binding region residues: Asp-164, Gly-163, Cys-111, His-272,

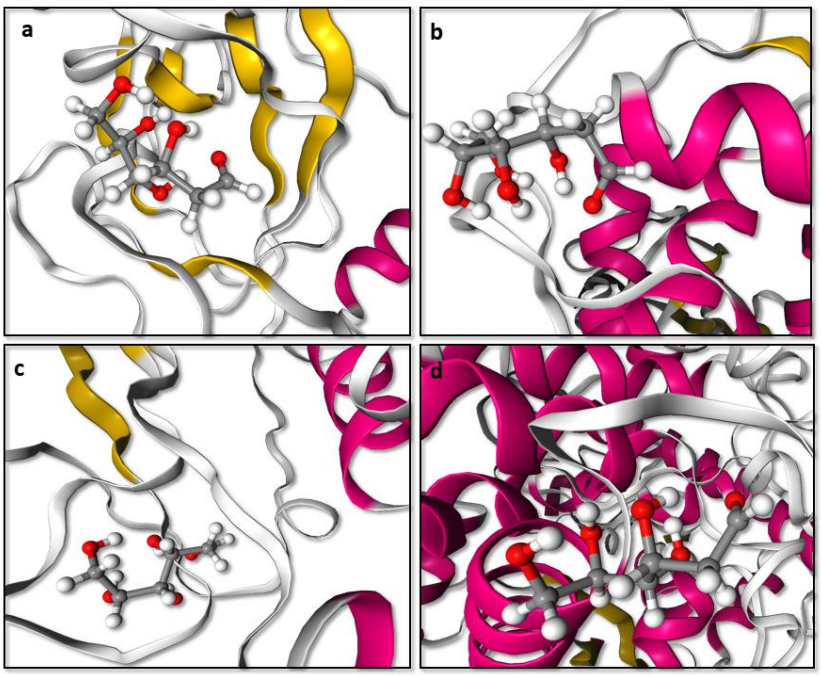


Figure 1. Best docked pose resulted from molecular docking calculation for a. $3CL^{pro}$-2-DG(6lu7-2-DG) b. PL^{pro}-2-DG(6wuu-2-DG) c. NSp-16-2-DG(6wkq-2-DG) d. RdRp-2-DG(7bv1-2-DG) from PR*in*MTML-ESS platform.

Gly-271, Tyr-268) obtained from a study that revealed the inhibitory mechanisms and determined the crystal structure of inhibitor VIR250 and VIR251 in complex with SARS-CoV-2 PL^{pro}.[9] A study by Yang et al. determined the crystal structure of $3CL^{pro}$ of SARS-CoV-2 in complex with the inhibitor (N3).[10] They have predicted the bonding site of inhibitors (binding region residues: Cys-145 to His-41 cavity) which we have used for the grid generation. Yin *et al.* determined the crystal structure of polymerase bound to RNA and to the drug Remdesivir. The binding site information of RdRP (Gly-683, Ser-682, Asp-684, Asp-760, Ser-759, Ala-685, Tyr-689, Ser-814, Gln-815) has been obtained from this study and used for grid generation.[11] Rosas-Lemus *et al.* unravelled the crystal structures 2'-O-methyltransferase, the nsp16-nsp10 heterodimer, in complex with an inhibitor Sinefungin.[12] Sinefungin binds to Nsp-16 at the binding site (ASN-6899, ASP-6897, ASP-6928, ASN-6841, CYS-6913, TYR-6930) which has been used for grid generation. In all the protein ligand combination the best docked pose was found be residing at that particular binding site, whose information we have obtained from crystal structure (Figure 1). The binding energy of the best docked pose for $3CL^{pro}$-2-DG(6lu7-2-DG), NSp-16-2-DG(6wkq-2-DG), PL^{pro}-2-DG(6wuu-2-DG) and RdRp-2-DG(7bv1-2-DG) are listed in Table 1. Docking scores were found to be -2.35 kcal/mol for $3CL^{pro}$-2-DG(6lu7-2-DG)

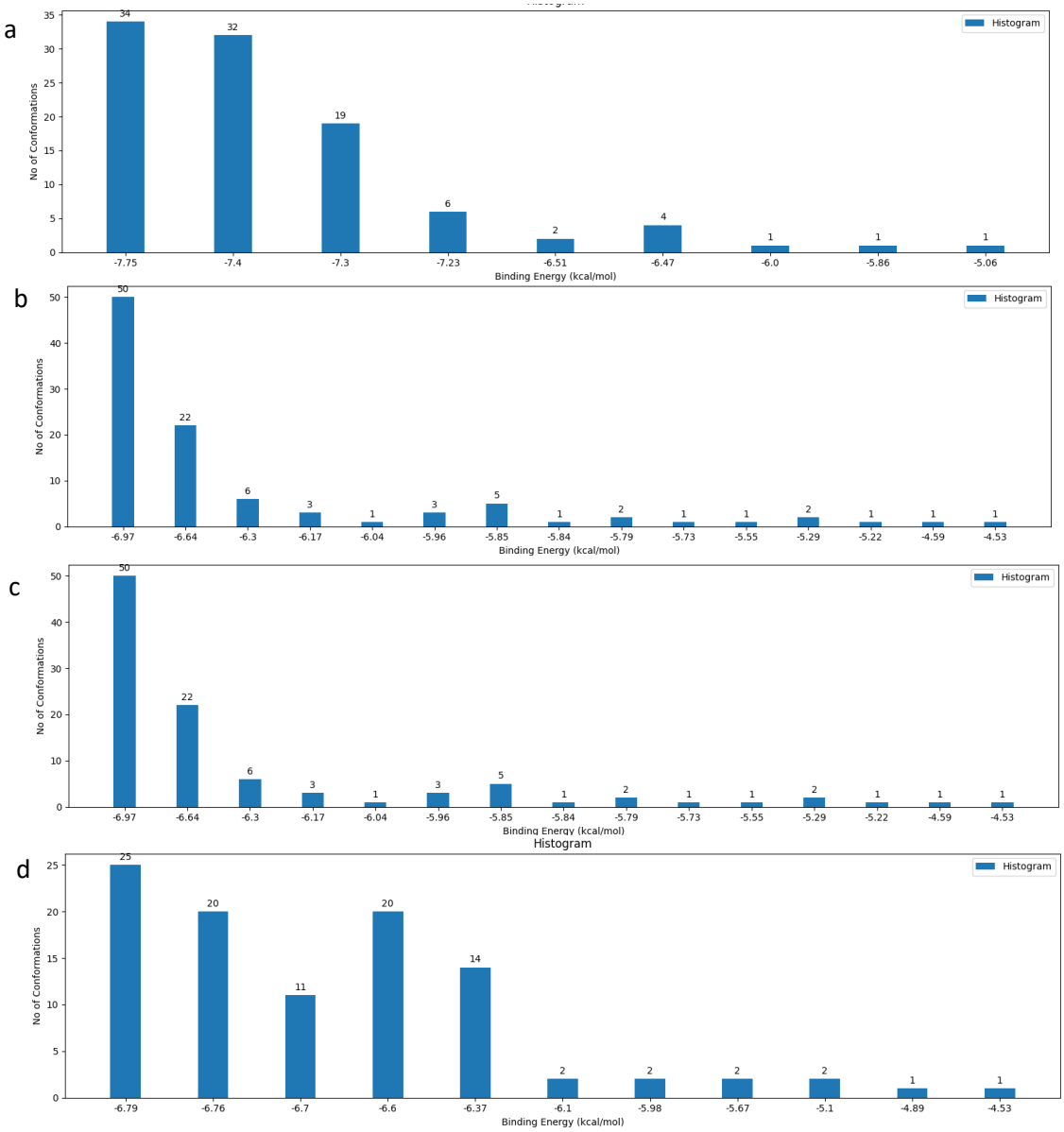


Figure 2. The histogram plot of conformers after molecular docking for a. 3CL^{pro}-2-DG(6lu7-2-DG) b. PL^{pro}-2-DG(6wuu-2-DG) c. NSp-16-2-DG(6wkq-2-DG) d. RdRp-2-DG(7bv1-2-DG) generated from PR*in*MTML-ESS platform.

combination, -1.7 kcal/mol for NSp-16-2-DG(6wkq-2-DG) combination, -1.77 kcal/mol for PL^{pro}-2-DG(6wuu-2-DG) and -1.26 kcal/mol for RdRp-2-DG(7bv1-2-DG).

Docking Energy and Docking Score are plotted for all the protein-ligand combination in the Docking Analysis part of PR*in*MTML-ESS platform. The histogram plots for all the combinations indicates that the lowest energy configuration is the highest in number in the clusters (Figures 2). The combination 3CL^{pro}-2-DG have the lowest docking energy and docking score having three hydrogen bonding interaction of the ligand, two between the OH group of 2-DG with the carbonyl group of the surrounding residue Phe-140 and

one between OH group of 2-DG with the amine nitrogen of surrounding residue Gly-143.

The hydrogen bonding and hydrophobic interactions of all 2-DG with the four viral proteins at their lowest energy configuration are shown in Figures 3. To select the combinations for molecular dynamic simulation in PR*in*MTML-ESS a screening is done based on the docking energy. Since the lowest energy configuration for all the protein-ligand combination, we have set a cutoff score of -6.00 kcal/mol which selects all the four protein-ligand combinations for further MD simulations. The lowest energy configuration of the docked structure was used as the starting structure for performing MD simulation for apprehending the stability of the interaction of the ligand with respect to the protein binding site in explicit water. In the PR*in*MTML-ESS platform for quantified the interactions between the amino acids in the binding pocket and the ligand several analysis has been carried out. The average root mean square

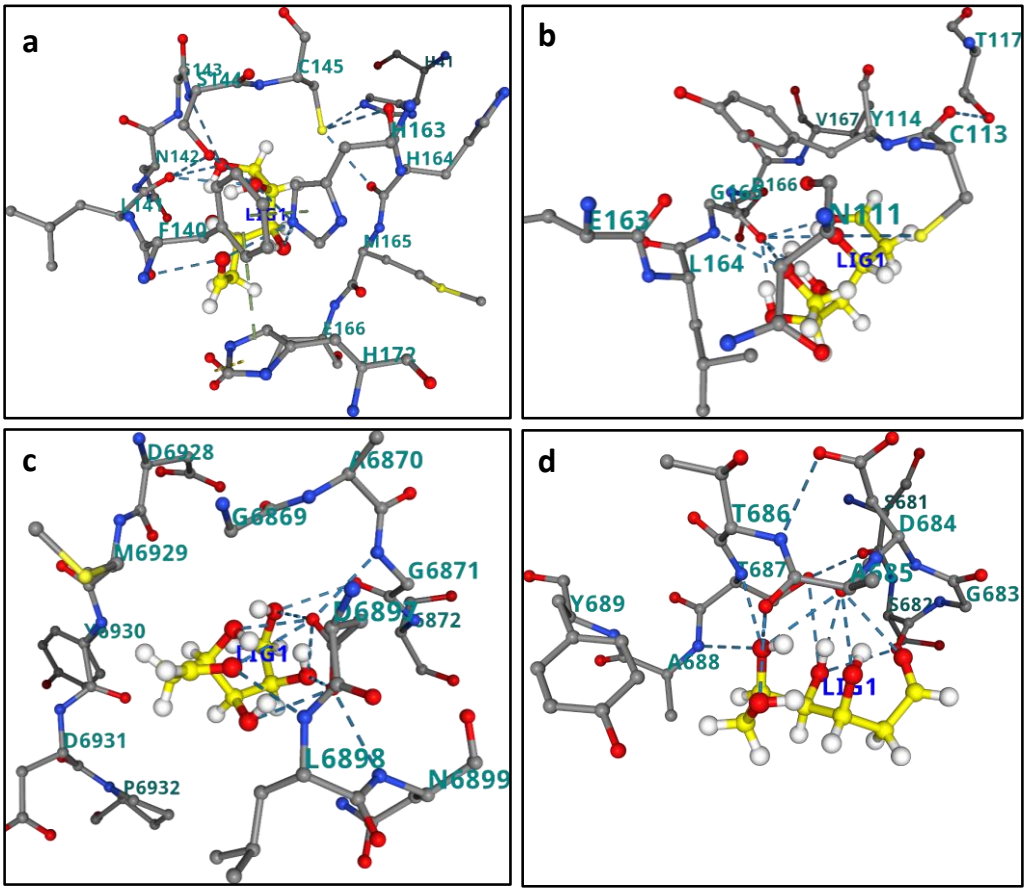


Figure 3. The ligand interaction showing hydrogen bonding and hydrophobic interaction at the best docked pose a. 3CL^{pro}-2-DG(6lu7-2-DG) b. PL^{pro}-2-DG(6wuu-2-DG) c. NSp-16-2-DG(6wkq-2-DG) d. RdRp-2-DG(7bv1-2-DG) generated from PR*in*MTML-ESS platform.

Table 1. The docking energy, docking score, MD non-bonded total energy, free energy barriers for dissociation, average RMSD values, Centre of Mass(COM) distance

Target	Docking Energy (kcal/mol)	Docking Score (kcal/mol)	MD non-bonded total energy (kJ/mol)	Free Energy barrier (kJ/mol)	Average RMSD (nm)	COM Distance (nm)	Interacting Residue	Residue-ligand interacting groups
NSp16	-7.71	-1.7	-197.7	29.2	0.13	0.47	Asp-6897 Asp-6928 Gly-6871	OH-C=O OH-C=O OH-C=O
RdRp	-6.79	-1.26	-66.2	26.4	0.14	0.16	Ser-682	OH-C=O
3CLPro	-7.75	-2.35	-97.1	---	0.11	1.27	---	---
PLPro	-6.97	-1.77	-10.3	---	0.16	1.08	---	---

deviation (RMSD) of the target protein binding pocket residues is calculated between the initial MD configuration (docked structure) and the MD equilibrated structure (final frame of MD simulation). Average RMSD of the final frame of MD simulated structure in the four combinations is found to vary between 0.11 nm to 0.16 nm (Table 1). Average RMSD value lower than 0.2 nm clearly shows that all the protein ligand combinations are stable. The hydrogen bonding section plots the number of hydrogen bonds between the ligand and the residues at the binding site of the target. NSp16-2-DG(6wkq-2-DG) is found to have the highest number of hydrogen bonding interaction, six followed by two hydrogen bonding in RdRp-2-DG(7bv1-2-DG) and no hydrogen bonding in PL^{pro}- 2-DG(6wuu-2-DG), and $3CL^{pro}$-2-DG(6lu7-2-DG) combination (Figure 4). In the Interaction Energy section, the non-bonded interaction energies, coulomb and van der Waals interaction energies, between the ligand and protein are plotted for each combination in the platform.

The total non-bonded interaction energy was found to -197.7 kJ/mol for NSp16-2-DG(6wkq-2-DG), hence the most favourable interaction of ligand 2-DG occurs with O-methyl transferase (NSp-16) of SARS-CoV-2. Total non-bonded interaction energy for $3CL^{pro}$-2-DG(6lu7-2-DG), RdRp-2-DG(7bv1-2-DG) and PL^{pro}-2-DG(6wuu-2-DG) are found to be -97.1 kJ/mol, -66.2 kJ/mol and -10.3 kJ/mol respectively (Table 1). These energy values predict that only in NSp16(6wkq) the ligand 2-DG is stable and tends to be in the binding site. However, for the other three viral proteins $3CL^{pro}$, RdRp and PL^{pro} the interaction energy is much lower to stabilise the ligand in the binding cavity.

The next section of MD analysis is the centre of mass distance which calculates the distance between the centre of mass of the residues in the binding site with the centre of mass of the ligand. Average centre of mass distance if larger than 1 nm would predict the ligand has escaped the binding pocket after MD simulation.

The average centre of mass (COM) distance plot shows that the combinations NSp16-2-DG(6wkq-2-DG) and RdRp-2-DG(7bv1-2-DG) have value 0.47 nm and 0.16 nm respectively, hence lies within the binding site cavity at the equilibrated structure after MD simulation. (Table 1) However, in both the combination $3CL^{pro}$-2-DG(6lu7-2-DG) and PL^{pro}-2-DG(6wuu-2-DG) the average COM distance are 1.27 nm and 1.08 nm respectively, hence here the ligand 2-DG have escaped the binding sites because of less interaction strength with the binding site residues of protein. The final equilibrated structure of the four combinations and visualization of the trajectory is consistent with the COM distance results.

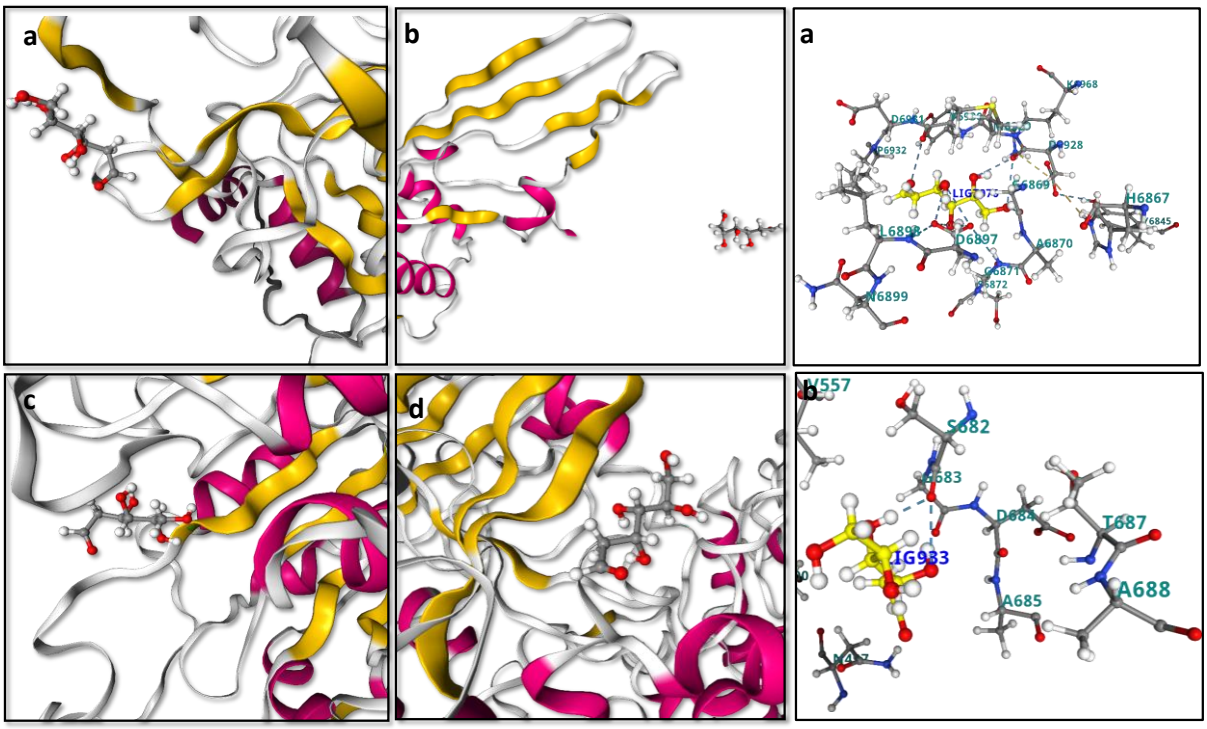


Figure 4. [A] The equilibrated structure after MD simulation for the combinations a. $3CL^{pro}$-2-DG(6lu7-2-DG) b. PL^{pro}-2-DG(6wuu-2-DG) c. NSp-16-2-DG(6wkq-2-DG) d. RdRp-2-DG(7bv1-2-DG) generated from PR*in*MTML-ESS platform (Left). [B] The ligand interaction with the binding residues showing hydrogen bonding and hydrophobic interaction in a. NSp-16-2-DG(6wkq-2-DG) and b. RdRp-2-DG(7bv1-2-DG) (Right).

The equilibrated structure of NSp16-2-DG(6wkq-2-DG) have three hydrogen bonding interaction between 2-DG and the binding residues Asp-6897, Asp-6928 and Gly-6871. (Figure 4). The equilibrated structure of RdRp-2-DG(7bv1-2-DG) has two hydrogen bonds with Ser-682. Our platform generates a combination score based on the average RMSD, internal energy and hydrogen bonding data obtained from analysis of MD simulation. Based on this score the selection of combinations for running the enhanced sampling free energy simulation is done. Among the four combination we have selected NSp16-2-DG(6wkq-2-DG) and RdRp-2-DG(7bv1-2-DG) combination based on their combination score 0.51 and 0.55 respectively, since only in these two combinations the ligand is still in the binding cavity after MD simulation. The sampling around the binding sites in instance of an MD simulation is not sufficient to predict the stability of the ligand in the binding cavity of protein, as conformations might get stuck in local minima. Therefore, meta-dynamics an enhanced sampling method of simulation[19] for quantitatively predicting ligand binding energy and analysis of changes in the conformation of ligands is important to ascertain the most stable (bound) protein-ligand complex and predicting the inhibition property of the ligand. The equilibrium structures obtained from MD simulations were used as the starting configurations in the enhanced sampling free energy simulations (FES). The average free energy of dissociation obtained from 5 independent dissociation simulations for the combination NSp16-2-DG(6wkq-2-DG) and RdRp-2-DG(7bv1-2-DG) are 29.1 kJ/mol and 26.4 kJ/mol (Table 1). For clarity and understanding of the free energy surfaces the individual free energy surfaces of all the independent runs and the average surface is shown in Figure 5. The free energy surfaces predict that the surfaces have multiple local minima and one global minimum for the ligand bound at the binding cavity. This structural characteristic represents multiple interactions between the ligands and residues of the binding site. The free energy of dissociation of the ligand 2-DG is of same order for both the viral protein NSp16(6wkq) and RdRp(7bv1), however the combination NSp16-2-DG(6wkq-2-DG) has the maximum ligand dissociation free energy. However previous studies show that antiviral drugs like Carfilzomib, Eravacycline, Valrubicin, Lopinavir, Elbasvir and Ritonavir interacts with SARS-COV-2 viral proteins $3CL^{pro}$, PL^{pro}, RdRp and spike protein where total binding free energy ranges from -79.4 kJ/mol to -168.1 kJ/mol with the van der Waal

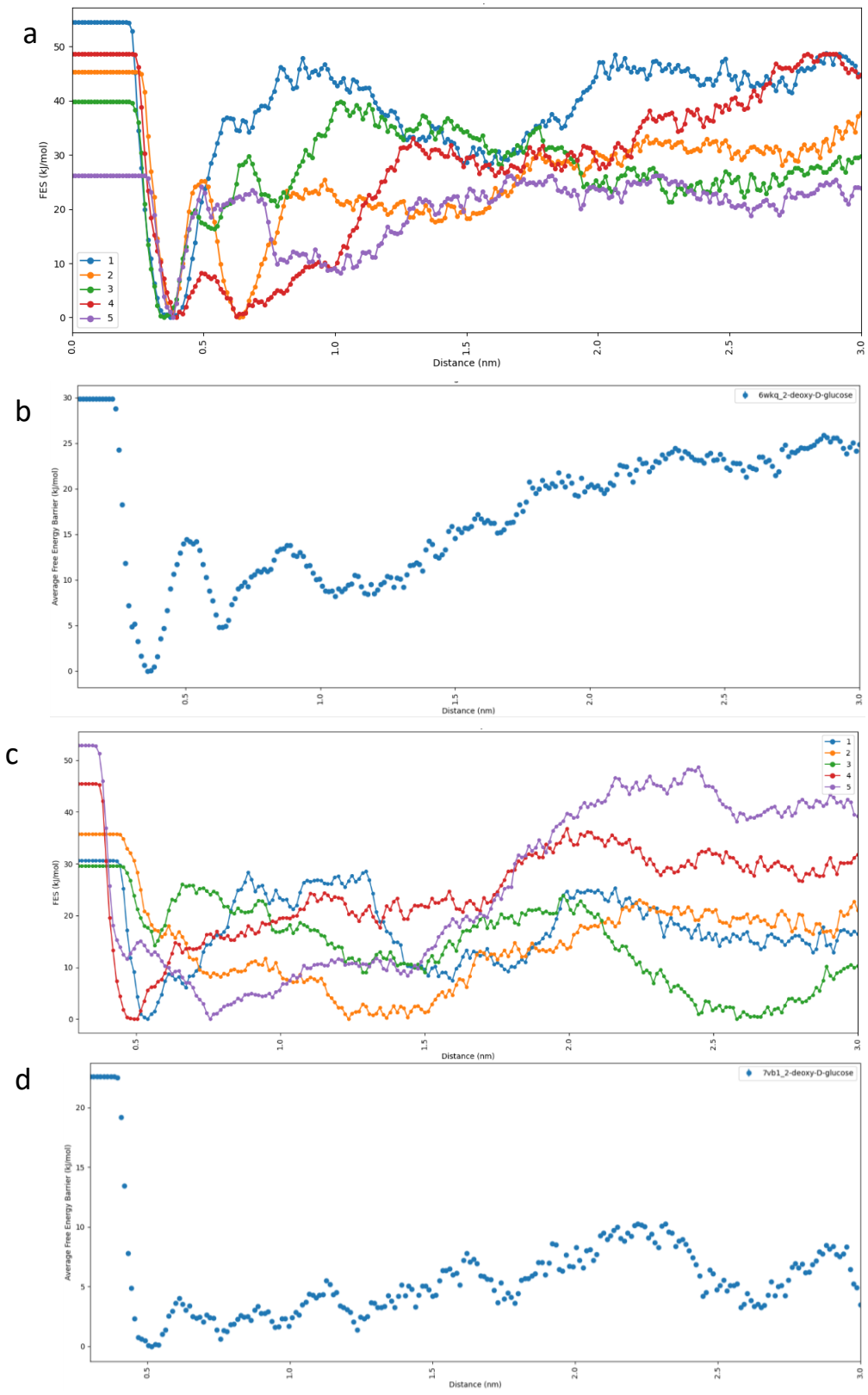
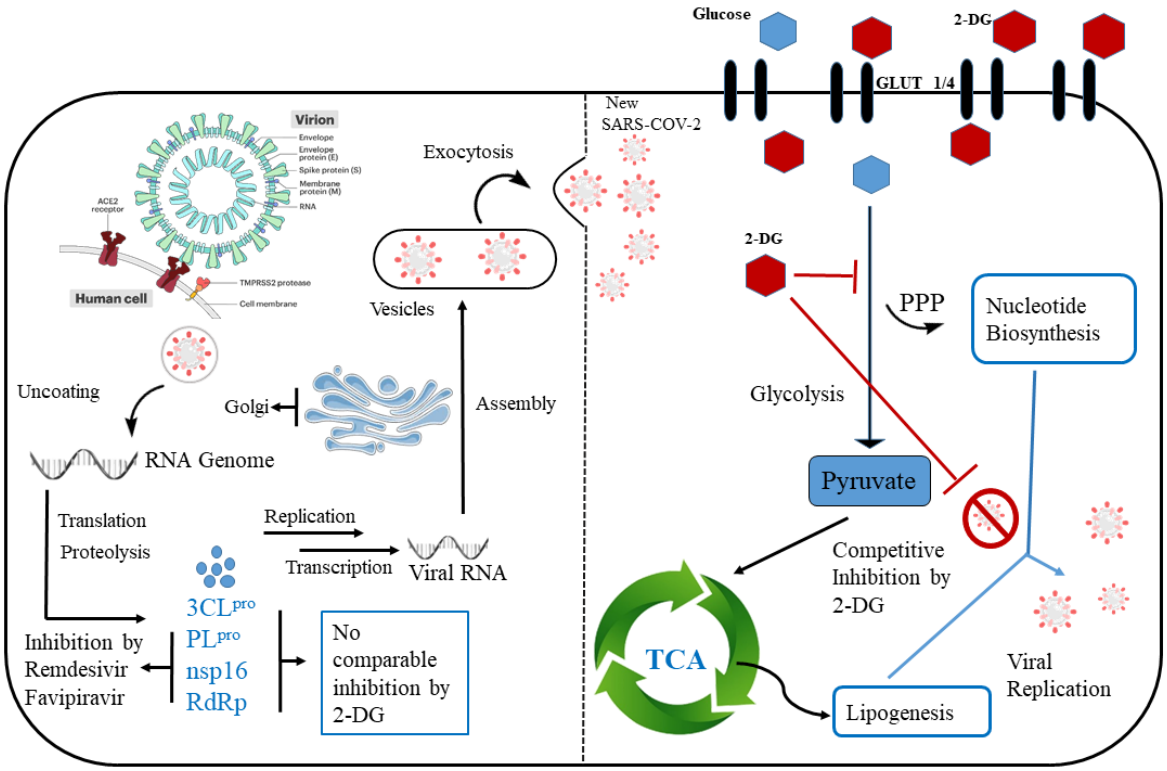


Figure 5. a. Free energy surface obtained from five Individual dissociation simulations for the combination NSp16-2-DG(6wkq-2-DG) b. Average free energy surface for the combination NSp16-2-DG(6wkq-2-DG) c. Free energy surface obtained from five Individual dissociation simulations for the combination RdRp-2-DG(7bv1-2-DG) b. Average free energy surface for the combination NSp16-2-DG(6wkq-2-DG) d. Average free energy surface for the combination RdRp-2-DG(7bv1-2-DG)



Scheme 2. a. Inhibition of viral proteins taking part in replication and transcription of SARS-COV 2 b. Glycolytic inhibition via 2-deoxy-D-glucose (2-DG)

energy components ranging from -133 kJ/mol to -255.1 kJ/mol when calculated using molecular dynamics and MM-GBSA based scoring function.[20] Since 2-DG shows a lower value of binding energy to all the viral proteins compared to most of the antiviral drugs it is challenging to predict if 2-DG plays a role of antiviral drug for inhibiting the SARS-CoV-2 viral proteins. The free energy barrier of dissociation of the 2-DG from both the viral proteins Nsp16 and RdRp is extremely low and can hardly predict any inhibition property of 2-DG. We also carried a 50 ns simulation for all the four combination and aligned with the enhanced free energy sampling results we found in all the case 2-DG have escaped the binding cavity at the final equilibrated structure.

On the other hand, several studies have shown glucose plays a major role in proliferation of SARS-CoV-2.[21] This is one of the reasons why obese and diabetic patients with uncontrolled blood glucose level are more prone to develop a severe form of COVID-19. Due to the structural similarity between 2-DG and glucose (Scheme 1), it is expected that 2-DG would act as the competitive inhibitor of glucose metabolism and would strongly effect the metabolic processes like glycolysis and glycosylation which are dependent on glucose. When SARS-CoV-2 attacks our cells, they co-opt both of these metabolic

Table 2. The docking energy, MD non-bonded total energy, free energy barriers for dissociation, average RMSD values, Centre of Mass(COM) distance

Target	Ligand	Docking Energy (kcal/mol)	MD non-bonded total energy (kJ/mol)	Free Energy barrier (kJ/mol)	Average RMSD (nm)	COM Distance (nm)	Interacting Residue	Residue-ligand interacting groups
Hexokinase	Glucose	-7.72	-330.2	66.3	3.7	0.50	Gln-291 Lys-173 Asn-235 Thr-172 Asn-208 Glu-260 Asp-209 Glu-294	NH2-OH N-OH NH2-O OH-OH NH2-OH C=O-OH C=O-OH C=O-OH
Hexokinase	2-DG	-7.33	-257.2	52.4	3.6	0.38	Glu-260 Asn-235 Asp-209 Thr-172	C=O-OH NH2-O C=O-OH OH-OH

processes to increase its replication and transcription. SARS-CoV-2 induces an anabolic state in their host cell which causes these infected cells to upregulate their production of energy using glycolysis as compared with their healthy neighbours.[22] 2-DG differs from the glucose by removal of an oxygen atom from the 2-position (Scheme 1). In the glycolysis 2-DG is absorbed by the cell and undergoes phosphorylation at the 6 position to generate 2-DG-6P, in the next step unlike glucose-6-phosphate, 2-DG-6P cannot undergo isomerisation by glucose-6-phosphate isomerase hence the glycolytic flux is reduced (Scheme 2). Therefore, in presence of 2-DG, the rate limiting reaction of glycolysis is ATP-dependent phosphorylation of glucose to form glucose-6-phosphate (G-6-P) and is catalysed by tissue-specific isoenzymes known as hexokinases. A recent study showed that SARS-COV-2 infection induces higher amount of glucose influx and glycolysis in the infected cells, resulting in selective high accumulation of the fluorescent glucose/2-DG analogue in the viral infected cells. They further observed that mannose inhibit the entry of 2-DG analogue at a very low concentration hence predicted that 2-DG entry in virus-infected cells might be manipulating specific mannose transporter or high-affinity glucose transporter, GLUT3, which was found to be increased on SARS-CoV-2 infection.[23] In our study we have tried to explore the competitive inhibition property of 2-DG on the first step of phosphorylation with the aid of hexokinase enzyme after entering the

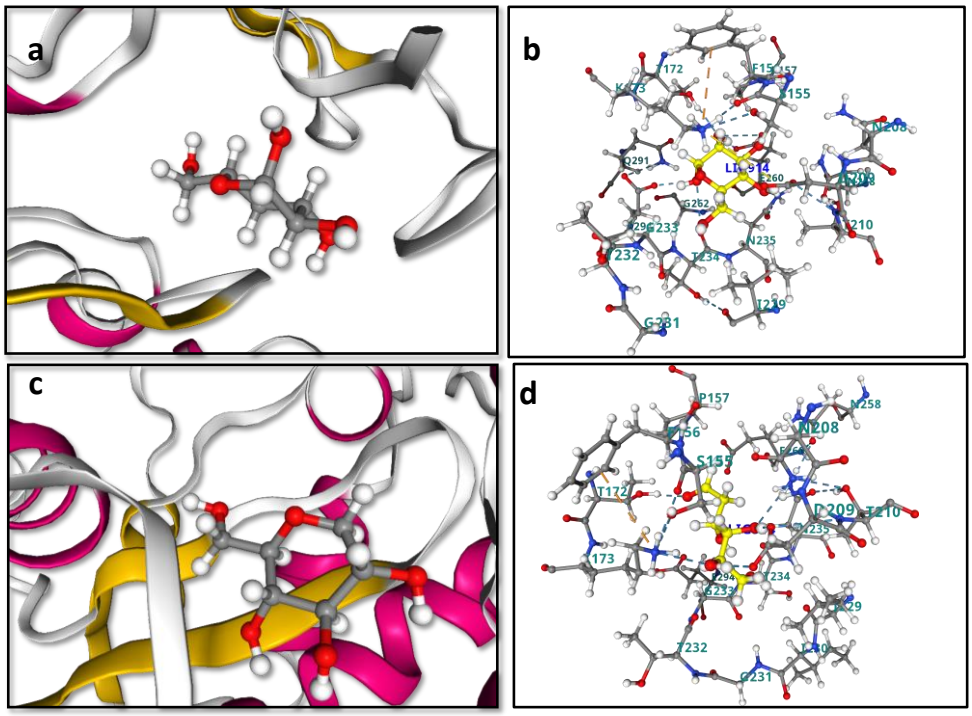


Figure 6. [A] The equilibrated structure after MD simulation for the combinations a. Hexokinase-Glucose(1qha-glucose) c. Hexokinase-2-deoxy-D-glucose(1qha-2-DG) (Left) [B] The ligand interaction with the binding residues showing hydrogen bonding and hydrophobic interaction in b. Hexokinase-Glucose(1qha-glucose) and d. Hexokinase-2-deoxy-D-glucose(1qha-2-DG) (Right).

viral infected cell. The 2-DG and glucose cyclic conformers were docked at the binding site of hexokinase (Thr-210, Thr-232, Ser-155, Lys-173).The binding energy was found to be -7.72 kcal/mol and -7.33 kcal/mol for glucose and 2-DG respectively. The entropic contribution associated with the solvent effect and conformational changes in the docked ligand is not accounted in the docking. The best docked pose and MD simulation is carried out in the next step. The non-bonded total interaction energy in the equilibrated structure were found to be slightly lower in glucose, -330.2 kJ/mol than 2-DG, -257.2 kJ/mol. The Coulomb energy and Van der Walls energy was also found to follow the same trend (Table 2). RMSD was computed for the target binding pocket residues is the initial MD configuration and the MD equilibrated structure for both the 2-DG-hexokinase and glucose-hexokinase structure (Table 2). The Centre of mass distance 0.50nm and 0.38 nm for glucose and 2-DG indicated that the ligands are stable in the binding cavity. Number of hydrogen bonding interaction in the equilibrated structure in glucose and 2-DG are ten and seven respectively, the interacting residues are indicated in Table 2 and Figure 6. To increase the sampling around the binding site of the protein and calculate the more accurate

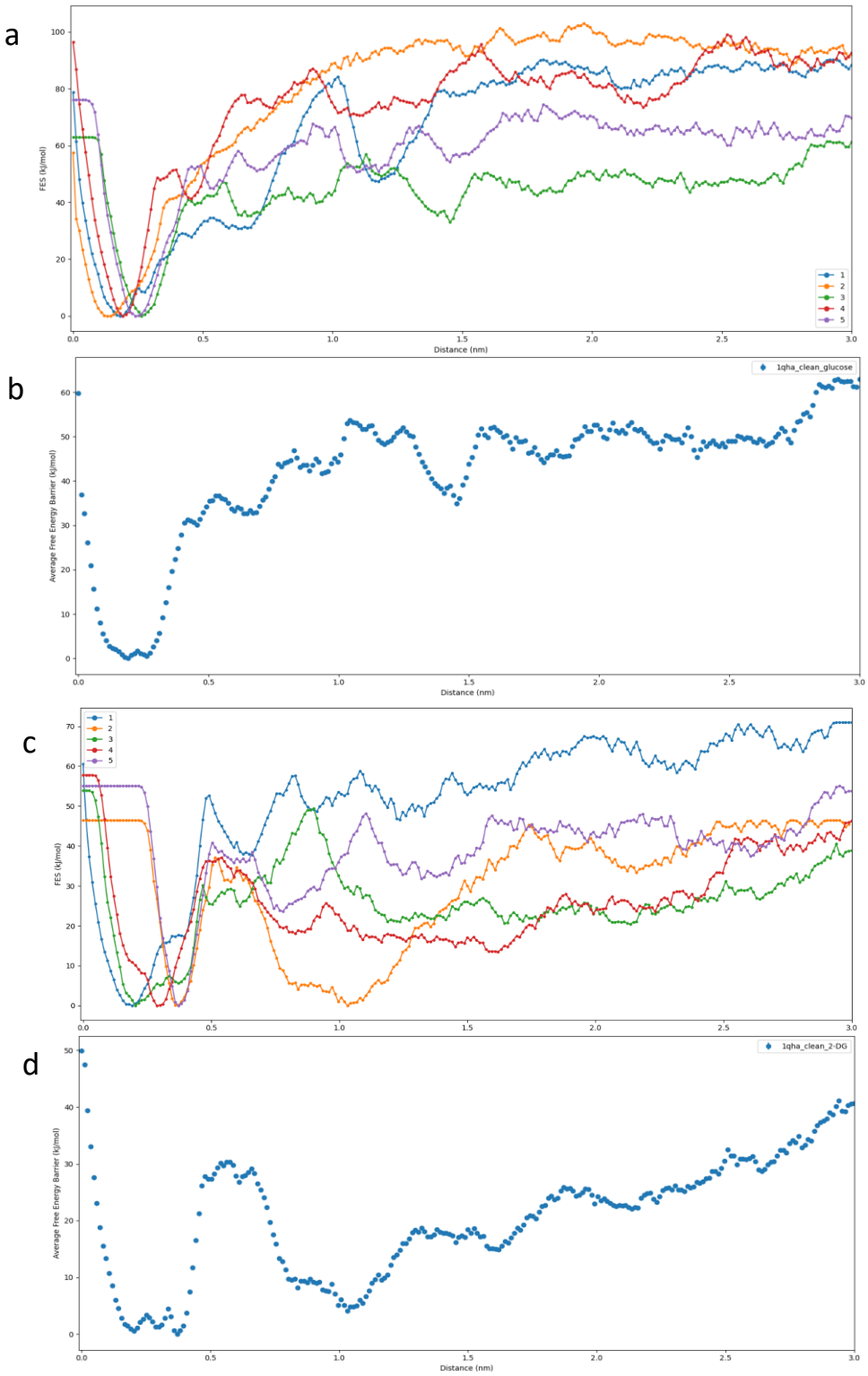
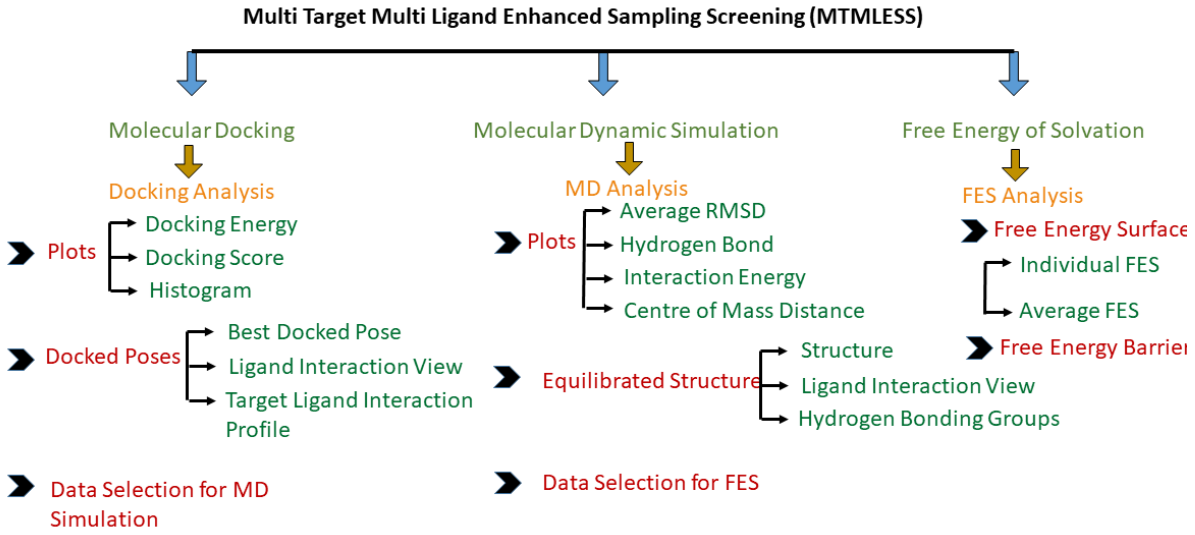


Figure 7. a. Free energy surface obtained from five Individual dissociation simulations for the combination Hexokinase-Glucose(1qha-glucose) b. Average free energy surface for the combination Hexokinase-Glucose(1qha-glucose) c. Free energy surface obtained from five Individual dissociation simulations for the combination Hexokinase-2-deoxy-D-glucose(1qha-2-DG) d. Average free energy surface for the combination Hexokinase-2-deoxy-D-glucose(1qha-2-DG)

binding energy of the ligand considering the changes in conformation of the ligand the equilibrium structure obtained from MD simulations was used as the starting configuration in the enhanced wt-metaD simulation. The average free energy barrier of dissociation obtained from enhanced wt-metaD simulations was found to be 66.3 kJ/mol and 52.4 kJ/mol for glucose and 2-DG (Table 2). Hence, it can be predicted that since the free energy barrier is comparable and of same order in both glucose and 2-DG, they have the similar binding affinity to hexokinase enzyme. Therefore, the phosphorylation step of conversion of glucose to glucose-6-phosphate is inhibited in presence of 2-DG at a certain concentration. Thus, competitive inhibition is caused by 2-DG, which structurally similar to glucose and can combine at the same binding site of hexokinase. For clearly understanding the free energy surface the average and the individual surfaces are plotted in PR*in*MTML-ESS platform as shown in Figure 7. The free energy surfaces show jagged and uneven surfaces arising due to the multiple interaction of the ligand at the binding site. The MD simulation was extended to 50 ns keeping all other parameters same as the 2 ns simulation. However negligible change in the ligand configuration and interaction of the ligand with the binding residues at the final equilibrated structure.

Computational Method

In this study for carrying out all the calculations and generating the figure the Prescience *in silico* Multi Target Multi Ligand Enhanced Sampling Screening (PR*in*MTML-ESS) has been used. In PR*in*MTML-ESS the process flow is in-build and fully automated, so no manual intervention is required from the user-side. The process flow here consists of molecular docking for docking the ligands in by automated grid generation at the target binding site. To explore the stability of the target-ligand complex all-atom molecular dynamic simulation is carried out in the next step and enhanced free energy sampling at the last stage for final analysis of the binding property of the ligand with the target (protein/enzyme/DNA-RNA). Efficient scoring algorithms are used after Docking and MD simulation for screening at each stages. All molecular docking calculation and MD simulations are performed using Autodock 4 and GROMACS-5.1.4 simulation package.[24-25] The enhanced sampling using metadynamics[19] (metaD) and its variant well-tempered metadynamics[26] (wt-metaD) using



Scheme 2. PR*in*MTML-ESS platform structure of integrated computational methods and analysis

Plumed 2.3.0[27] patched with MD engine GROMACS 5.1.4 integrated within the PR*in*MTML-ESS platform. Enhanced sampling-based method metadynamics developed by Laio et al.[19], is widely used in simple and complex molecules due to its advantage over free energy perturbation and thermodynamic integration methods which are computationally more expensive for larger complex system like protein-ligand, protein-protein binding. A time-dependent bias is included to the system in a metadynamic simulation along with some deposited bias like a suitably chosen reaction coordinate(s) which will eventually push the complex away from the minimum energy state, so that the system doesn't get trapped at a local minima for sufficiently long time. Since this method is independent on the choice of reaction coordinate and not very sensitive to the precise choice of biasing parameters (except in the case where the parameters are chosen to be too high or low), any rudimentary reaction coordinate can bias the system helping it to escape the local minima within a small time and generate a qualitative free energy surface (FES). Metadynamics simulation uses a history-dependent bias which inhibits the system from repeated revisiting previously visited regions of the phase space. Another, widely used scoring method over molecular dynamic simulation is the molecular mechanics Poisson-Boltzmann surface area (MMPBSA) which uses approximations to calculate enthalpic and entropic contributions using implicit continuum solvent hence less accurate in comparison to enhanced free energy sampling method used in our study. Additionally, as an alternative to MMPBSA we have performed all-atom simulation in explicit solvent medium including dynamics of solvent, solute and

ions. Considering the effectiveness and computational viability of enhanced free energy sampling method, it has been used as the final scoring method for calculating free energy barrier for dissociation in PR*in*MTML-ESS platform. In-house codes have been used for generating the analysis in each section of computational methods used. A scheme of the methodologies and analysis used in PR*in*MTML-ESS platform are shown in Scheme 2. The structure of ligands used in this study has been downloaded from PubChem (https://pubchem.ncbi.nlm.nih.gov/). The crystal structure of the viral proteins of SAR-COV2 PLpro (PDB ID: 6wuu), 3CLpro (PDB ID: 6lu7), RNA-dependent RNA polymerase (PDB ID: 7bv1) and the Hexokinase enzyme (PDB ID: 1QHA) are obtained from RCSB Protein Data Bank(https://www.rcsb.org/). All the protein and enzyme structures were cleaned (removing ligands, ions, water molecules) and missing loop was modelled using our PR*in*Bio platform. A ligand conformational search is carried out generating an automated grid box of the dimension 60*60*60 Å by taking the binding site centroid as a grid center with a spacing of 0.375 Å. The docking was conducted the Lamarckian genetic algorithm (LGA), and a total of 100 GA-LA hybrid runs for performing the conformational search. In the MD simulation the proteins and enzyme were modelled using the CHARMM27 force field[28] parameters. For generating the CHARMM27 force field for all the ligands, SwissParam[29] was used and the force field generation is automated in PR*in*MTML-ESS platform. The target-ligand systems were solvated in water and equilibrated using MD simulations at room temperature. The systems were first equilibrated using an NVT ensemble at 300 K for 1 ns and extended to the NPT ensemble at 300 K and 1 atm for another 2 ns. The temperature and pressure during the simulations were maintained using a velocity rescaling thermostat and Parrinello-Rahman barostat, respectively. A time step of 1 fs was used to integrate the equation of motion, and a non-bonded cutoff of 10Å was used to perform the MD simulations. We have used a bias V(s,t) in the form of Gaussians with every 500 steps (1 ps) deposition pace with a Gaussian hill-height of 2.0 kJ/mol, width of σ (0.1 nm), bias factor of 15, and temperature (T) of 300 K. For smooth convergence of the system in a wt-metaD the amplitude of the bias is tuned accordingly. Here a tempering factor ΔT (Equation 1) is used to adjust the height of the hills, and henceforth a smooth convergence of the free energy landscape is attained.

$$F(s) = -\frac{T+\Delta T}{T}V(s,t) + C(t) \quad (1)$$

Once the system converges, free energy F(s) (Eq. (1))[30] can be extracted by adding the deposited hills along the biased reaction coordinates. The centre of mass distance between the heavy atoms of the ligands and the protein backbone in the surrounding area of the binding pocket is considered as the reaction coordinates in the free energy sampling, since the aim of the calculation is taking into account the ligand dissociation from the binding site. In this study 5 independent simulations are done and an averaging is done for each combination to obtain better sampling and statistically reliable results.

Conclusion

This computational study aims to understand the role of 2-deoxy-D-glucose as an antiviral and glycolysis pathway inhibitor in SAR-CoV-2 affected human body using PR*in*MTML-ESS platform. Using this platform molecular docking, all-atom molecular dynamics simulations with enhanced free energy simulation was performed and screening was done in each stage using effective scoring algorithm. Here we have explored the inhibition property of 2-DG with the four SAR-CoV-2 viral proteins $3CL^{pro}$, PL^{pro}, NSp-16 and RdRp which are important for transmission and replication of SARS-CoV-2. 2-DG has very low dissociation barrier with O-methyl-transferase (Nsp16) and RdRp in free energy simulation. Hence it can be predicted that 2-DG can barely inhibit any viral protein taking part in replication and translation of SARS-CoV-2. On the other hand, PR*in*MTML-ESS platform predicts 2-DG binds strongly to hexokinase, the enzyme responsible for phosphorylation of glucose to glucose-6-phosphate in glycolysis. Though glucose has higher binding affinity to hexokinase, comparable binding affinity of 2-DG may lead to competitive inhibition of the glycolysis, hence the glycolytic flux is reduced which in turn reduces the replication and transcription of SARS-CoV-2 virus in human cells. Therefore, our study predicts that 2-DG though doesn't show any comparable inhibition for non-structured protein associated with translation and replication of SAR-CoV-2 can be used as a clinical therapy in COVID-19 infected patient for competitive inhibition of glycolysis process in virus infected cells, thereby reducing the rate the replication of virus.

Reference

1. Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Krüger, N.; Herrler, T.; Erichsen, S.; Schiergens, T. S.; Herrler, G.; Wu, N.-H.; Nitsche, A., SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *cell* **2020,** *181* (2), 271-280. e8.
2. Lan, J.; Ge, J.; Yu, J.; Shan, S.; Zhou, H.; Fan, S.; Zhang, Q.; Shi, X.; Wang, Q.; Zhang, L., Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **2020,** *581* (7807), 215-220.
3. Koopmans, M.; Haagmans, B., Assessing the extent of SARS-CoV-2 circulation through serological studies. *Nat Med* **2020,** *26* (8), 1171-2.
4. Woo, P. C.; Huang, Y.; Lau, S. K.; Yuen, K.-Y., Coronavirus genomics and bioinformatics analysis. *Viruses* **2010,** *2* (8), 1804-1820.
5. Sadee, W., Genomics and personalized medicine. Elsevier: 2011.
6. Helmy, Y. A.; Fawzy, M.; Elaswad, A.; Sobieh, A.; Kenney, S. P.; Shehata, A. A., The COVID-19 pandemic: a comprehensive review of taxonomy, genetics, epidemiology, diagnosis, treatment, and control. *Journal of clinical medicine* **2020,** *9* (4), 1225.
7. Yuan, M.; Wu, N. C.; Zhu, X.; Lee, C.-C. D.; So, R. T.; Lv, H.; Mok, C. K.; Wilson, I. A., A highly conserved cryptic epitope in the receptor binding domains of SARS-CoV-2 and SARS-CoV. *Science* **2020,** *368* (6491), 630-633.
8. Bagdonaite, I.; Wandall, H. H., Global aspects of viral glycosylation. *Glycobiology* **2018,** *28* (7), 443-467.
9. Rut, W.; Lv, Z.; Zmudzinski, M.; Patchett, S.; Nayak, D.; Snipas, S. J.; El Oualid, F.; Huang, T. T.; Bekes, M.; Drag, M., Activity profiling and crystal structures of inhibitor-bound SARS-CoV-2 papain-like protease: A framework for anti–COVID-19 drug design. *Science advances* **2020,** *6* (42), eabd4596.
10. Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C., Structure of M pro from SARS-CoV-2 and discovery of its inhibitors. *Nature* **2020,** *582* (7811), 289-293.
11. Yin, W.; Mao, C.; Luan, X.; Shen, D.-D.; Shen, Q.; Su, H.; Wang, X.; Zhou, F.; Zhao, W.; Gao, M., Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. *Science* **2020,** *368* (6498), 1499-1504.
12. Rosas-Lemus, M.; Minasov, G.; Shuvalova, L.; Inniss, N. L.; Kiryukhina, O.; Brunzelle, J.; Satchell, K. J., High-resolution structures of the SARS-CoV-2 2′-O-methyltransferase reveal strategies for structure-based inhibitor design. *Science Signaling* **2020,** *13* (651).
13. Namsani, S.; Pramanik, D.; Khan, M. A.; Roy, S.; Singh, J. K., Metadynamics-based enhanced sampling protocol for virtual screening: case study for 3CLpro protein for SARS-CoV-2. *Journal of Biomolecular Structure and Dynamics* **2021**, 1-16.
14. Kumawat, A.; Namsani, S.; Pramanik, D.; Roy, S.; Singh, J. K., Integrated docking and enhanced sampling based selection of repurposing drugs for SARS-CoV-2 by targeting host dependent factors. **2020**.
15. Gorbalenya, A. E.; Enjuanes, L.; Ziebuhr, J.; Snijder, E. J., Nidovirales: evolving the largest RNA virus genome. *Virus research* **2006,** *117* (1), 17-37.
16. V'kovski, P.; Kratzel, A.; Steiner, S.; Stalder, H.; Thiel, V., Coronavirus biology and replication: implications for SARS-CoV-2. *Nature Reviews Microbiology* **2020**, 1-16.
17. Báez-Santos, Y. M.; John, S. E. S.; Mesecar, A. D., The SARS-coronavirus papain-like protease: structure, function and inhibition by designed antiviral compounds. *Antiviral research* **2015,** *115*, 21-38.
18. Vithani, N.; Ward, M. D.; Zimmerman, M. I.; Novak, B.; Borowsky, J. H.; Singh, S.; Bowman, G. R., SARS-CoV-2 Nsp16 activation mechanism and a cryptic pocket with pan-coronavirus antiviral potential. *Biophysical Journal* **2021**.
19. Laio, A.; Parrinello, M., Escaping free-energy minima. *Proceedings of the National Academy of Sciences* **2002,** *99* (20), 12562-12566.

20. Murugan, N. A.; Kumar, S.; Jeyakanthan, J.; Srivastava, V., Searching for target-specific and multi-targeting organics for Covid-19 in the Drugbank database with a double scoring approach. *Scientific reports* **2020,** *10* (1), 1-16.
21. Codo, A. C.; Davanzo, G. G.; de Brito Monteiro, L.; de Souza, G. F.; Muraro, S. P.; Virgilio-da-Silva, J. V.; Prodonoff, J. S.; Carregari, V. C.; de Biagi Junior, C. A. O.; Crunfli, F., Elevated glucose levels favor SARS-CoV-2 infection and monocyte response through a HIF-1α/glycolysis-dependent axis. *Cell metabolism* **2020,** *32* (3), 437-446. e5.
22. Gualdoni, G. A.; Mayer, K. A.; Kapsch, A.-M.; Kreuzberg, K.; Puck, A.; Kienzl, P.; Oberndorfer, F.; Frühwirth, K.; Winkler, S.; Blaas, D., Rhinovirus induces an anabolic reprogramming in host cell metabolism essential for viral replication. *Proceedings of the National Academy of Sciences* **2018,** *115* (30), E7158-E7165.
23. Bhatt, A. N.; Kumar, A.; Rai, Y.; Kumari, N.; Vedagiri, D.; Harshan, K. H.; Chinnadurai, V.; Chandna, S., Glycolytic inhibitor 2-Deoxy-D-glucose attenuates SARS-CoV-2 multiplication in host cells and weakens the infective potential of progeny virions. *bioRxiv* **2021**.
24. Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E., GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015,** *1*, 19-25.
25. Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D., GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **2013,** *29* (7), 845-854.
26. Barducci, A.; Bussi, G.; Parrinello, M., Well-tempered metadynamics: a smoothly converging and tunable free-energy method. *Physical review letters* **2008,** *100* (2), 020603.
27. Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; Raiteri, P.; Donadio, D.; Marinelli, F.; Pietrucci, F.; Broglia, R. A., PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Computer Physics Communications* **2009,** *180* (10), 1961-1972.
28. MacKerell Jr, A. D.; Banavali, N.; Foloppe, N., Development and current status of the CHARMM force field for nucleic acids. *Biopolymers: Original Research on Biomolecules* **2000,** *56* (4), 257-265.
29. Zoete, V.; Cuendet, M. A.; Grosdidier, A.; Michielin, O., SwissParam: a fast force field generation tool for small organic molecules. *Journal of computational chemistry* **2011,** *32* (11), 2359-2368.
30. Clark, A. J.; Tiwary, P.; Borrelli, K.; Feng, S.; Miller, E. B.; Abel, R.; Friesner, R. A.; Berne, B. J., Prediction of protein–ligand binding poses via a combination of induced fit docking and metadynamics simulations. *Journal of chemical theory and computation* **2016,** *12* (6), 2990-2998.