Reappraisal of Trifluperidol against NSP-3 protein: Potential therapeutic for COVID-19

Ajita Pandey (ajitanu5@gmail.com)
GBU

Research Article

Keywords: SARS-CoV-2, NSP3, Molecular docking, Molecular dynamic simulation, FDA drugs, PCA, ADRP

Posted Date: August 5th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-52706/v1

Reappraisal of Trifluperidol against NSP-3 protein: Potential therapeutic for COVID-19

Ajita Pandey*

School Of Biotechnology, Gautam Buddha University, Greater Noida, U.P., India

Running Title: *Identification of anti-COVID-19 compounds from FDA approved drugs*

***Corresponding author:**
Ajita Pandey
School Of Biotechnology,
Gautam Buddha University,
Greater Noida, 201312,
U.P., India
Email: ajitanu5@gmail.com

Abstract

Novel coronavirus disease 2019 (COVID-19) is a highly infectious disease that is caused by the recently discovered severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2). Because there are no specific vaccines or drugs for SARS-CoV-2, drug repurposing may be a promising approach. SARS-CoV-2 has a positive-sense RNA genome that encodes non-structural proteins (Nsps), which are essential for viral replication in the host cell. Non-structural protein 3 (Nsp3) is a multidomain protein and is the largest protein of the replicase complex. Nsp3 contains an ADP-ribose phosphatase (ADRP) domain, also called the macrodomain, which interferes with the host immune response. In the present study, we used computational regression methods to target the ADRP domain of Nsp3, using FDA-approved drugs. We virtually screened 2,892 FDA-approved drugs, using a combination of molecular docking and scoring functions. Saquinavir and trifluperidol were identified as potential leads and were further investigated using molecular dynamics simulation (MDS) to predict the stability and behavior of the ADRP-drug complexes. Analysis of root mean square deviation, root mean square fluctuation, radius of gyration, solvent accessible surface area and number of hydrogen bonds showed that the ADRP-trifluperidol complex is more stable than the ADRP-saquinavir complex. The screening and the MDS results suggest that trifluperidol is a novel inhibitor of the ADRP domain of Nsp3. Trifluperidol could, therefore, potentially be used to help control the spread of COVID-19, either alone or in combination with antiviral agents. Further *in-vitro* and *in-vivo* experiments are necessary to confirm our *in silico* results.

Keywords: SARS-CoV-2; NSP3; Molecular docking; Molecular dynamic simulation; FDA drugs; PCA; ADRP

1. Introduction

In December 2019, a number of cases of pneumonia occurred in Wuhan, Hubei Province, China, with the first patient being hospitalized on 12 December [1]. Physician Li Wenliang was the first to suspect that these cases of pneumonia were caused by a coronavirus and, on 31 December 2019, the Chinese Center for Disease Control and Prevention and the Chinese office of the World Health Organization (WHO) officially confirmed the existence of a new coronavirus. The new virus, which was named severe acute respiratory syndrome

coronavirus-2 (SARS-CoV-2), causes a highly infectious disease, termed novel coronavirus disease 2019 (COVID-19), which often presents with pneumonia-like symptoms. SARS-CoV-2 is closely related to other coronaviruses, such as SARS-CoV and bat and pangolin coronaviruses [2]. SARS-CoV-2 spreads very rapidly and COVID-19 quickly became a global pandemic. As of 1st August 2020, the WHO has confirmed 17,897,378 cases of COVID-19 worldwide, with 685,703 deaths (https://www.worldometers.info/coronavirus/). COVID-19 is now spreading very rapidly in India, with thousands of new cases reported daily. As of 1st August 2020, more than 1.7 million confirmed cases and 37,403 deaths (https://www.covid19india.org/) have been reported. At present there are no effective vaccines or drugs to treat COVID-19.

Like other coronaviruses, SARS-CoV-2 has a positive-sense RNA genome and encodes several structural and non-structural proteins (Nsps). The structural proteins include the envelope, glycoprotein, nucleocapsid and membrane proteins, together with other accessory proteins [3]. The ORF1a and ORF1ab open reading frames are translated to produce two polyproteins, pp1a and pp1ab, which are cleaved by proteases encoded by ORF1a to yield the Nsps [4]. The latter polyprotein results from a ribosomal frameshift that enables continuous translation of ORF1a along with ORF1b. The polyprotein pp1a contains two viral proteases, a papain-like protease (PLpro, encoded within Nsp3) and a 3C-like main protease (Mpro, encoded by Nsp5). These two viral proteases play key roles in the post translation processing of the two polyproteins. The 16 Nsps that are formed by cleavage form a large membrane-bound replicase complex [5]. The multidomain protein Nsp3, which is the largest component in the replicase assembly, consists of an ADP-ribose phosphatase (ADRP) domain, also known as the macrodomain, an N-terminal Nsp3a domain, a PLpro domain, a marker domain, a SARS-unique domain, an RNA binding domain, a Y-domain and a transmembrane domain (https://coronavirus3d.org). Thirty years ago, the ADRP domain, which was initially known as the X domain, was shown, using bioinformatics techniques, to be a conserved and unique domain in the genomes of the *Coronaviridae, Togaviridae* and *Hepeviridae* families [6]. The ADRP domain is involved in various pathways, including posttranslational modification of proteins and ADP-ribose metabolism. The Nsp3 protein removes the 1" phosphate group from Appr-1"-p in *in vitro* assays, confirming its phosphatase activity [7]. It is believed that the ADRP domain plays a key role in altering innate immunity. Studies to investigate the role of this domain in compromising the immune response showed that virus with a mutated macrodomain replicated poorly in bone marrow-derived macrophages, which

are the primary cells involved in mounting an innate immune response [8]. Virus containing an inactivated macrodomain was also shown to be sensitive to pretreatment with interferon [9]. These studies confirm that the ADRP domain plays a crucial role in disease pathogenesis and suggest that inhibition of this domain should reduce viral burden and facilitate recovery [10].

Numerous studies have been conducted using different coronavirus proteins as drug targets [11–13] but, to the best of our knowledge, no studies have looked at the proteins involved in modifying host innate immunity. In this study, therefore, we choose to identify compounds that interact with the ADRP domain as potential antiviral agents. We virtually screened 2892 FDA-approved drugs, using the ADRP domain in the ADPr centric grid. Using a variety of computational methods, we found trifluperidol as potential hit, and could be repurposed to treat COVID-19. Our comprehensive methodology is shown in Figure 1.

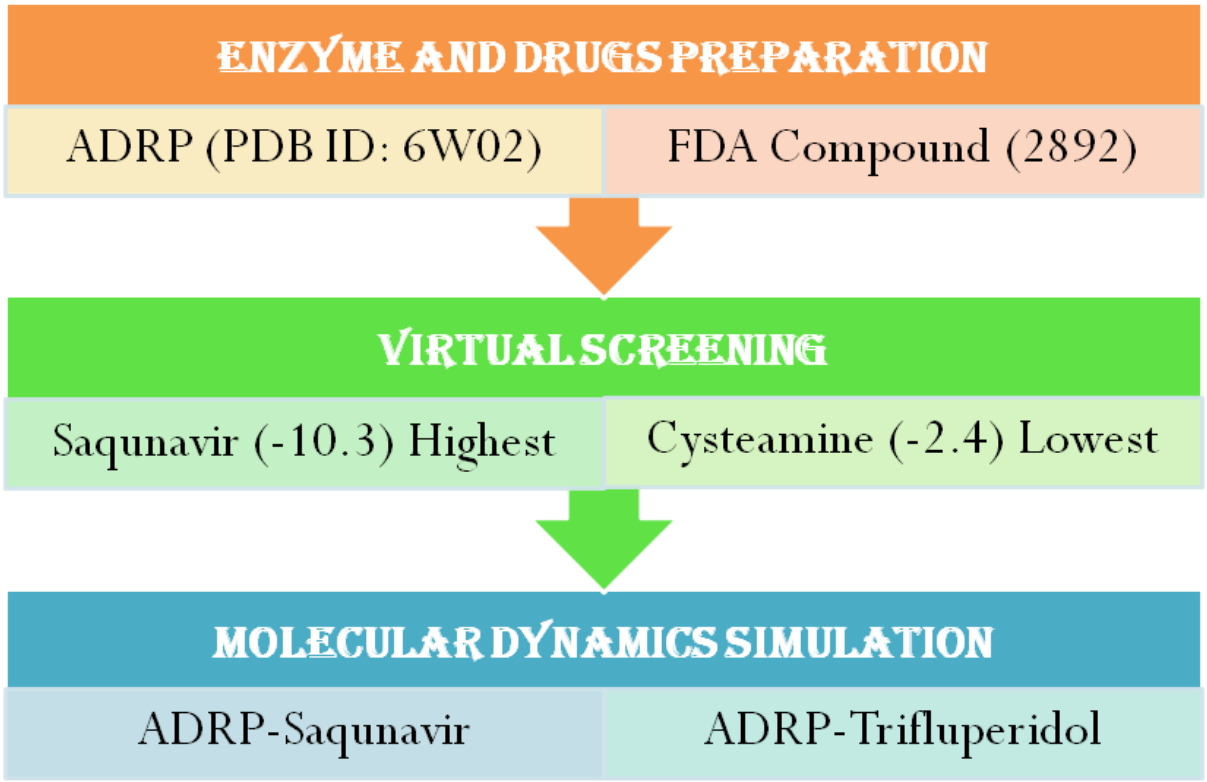


Figure 1: Complete workflow of methodology.

2. Experimental procedures

2.1 Protein preparation

The ADRP domain (PDB ID: 6W02, X-ray, 1.5 Å), which is a subunit of Nsp3, was retrieved from the Protein Data Bank (PDB) and used for protein preparation [10]. Co-crystallization of the ADRP domain with adenosine-5-diphosphoribose (ADPr) revealed interactions with the key catalytic residues. We removed all heteroatoms, water molecules and other unnecessary crystal stabilizers and then prepared the protein using Chimera 1.13.2 [14]. The protein was imported into Chimera and then minimized with the Amber ff99SB force-field, using the 100 steepest descent steps with 10 conjugant gradient steps to obtain the lowest and most stable conformation of the protein. The step size for both methods was set at 0.02 Å. The minimized lowest energy conformation and prepared structure were then used for virtual screening.

2.2 Ligand preparation

All 2892 FDA-approved drugs were retrieved from the DrugBank database (https://www.drugbank.ca/), which contains all FDA-approved drugs, as well as experimental and withdrawn compounds [15]. The same compounds are also available in the ZINC database [16] in *mol2* file format, and are ready to use without preparation. We retrieved the FDA-approved compounds in 3D SDF format and then converted them into *.mol2* file format using Open Bable software [17]. These *.mol2* files were then converted into *.pdbqt* file format using a Python script. During assignment of Gastieger charges, all hydrogen atoms and atomic radii were added in the course of ligand preparation. These converted and prepared ligands were then used for virtual screening.

2.3 Structure-based virtual screening

Structure-based virtual screening (SBVS) is a very powerful computational technique for sorting compounds on the basis of binding affinity [18, 19]. Here, we used SBVS to predict which compounds might bind to the ADRP domain of Nsp3. The crystal structure of ADPr bound to the ADRP domain showed that the binding site comprised residues Ala21, Asp22, Ile23, Ala38, Asn40, Lys44, His45, Gly46, Gly47, Gly48, Val49, Leu126, Ser128, Ala129, Gly130, Ile131, Phe132, Ala154, Phe156 and Leu160. A centric grid box towards ADPr was then prepared on the basis of these catalytic residues. The protein structure of the ADRP domain was prepared using MGL Tools [20]. We added hydrogen atoms and Kollmaan

charges during protein preparation. The prepared structure of the ADRP domain was then converted into *.pdbqt* file format and used for the virtual screening. We used Autodock Vina [21], which is widely used for virtual screening, in this study. The centric grid size in the ADPr-binding cavity was 28, 26, and 52 Å, centralized at 3.941, 5.589 and 22.719 for X, Y, and Z coordinates, respectively. The exhaustiveness and grid spacing were set to 8 and 1.00 Å, respectively, for screening. Compounds with the best binding were shortlisted on the basis of binding affinity and binding pose. We manually analyzed the top few compounds and selected only those compounds that bound to the catalytic residues and had greater binding affinity than the substrate for further analysis.

2.4 Analysis of docking complex

The docking complex was analyzed using Chimera 1.13.2 and Discovery Studio Visualizer and the 5 Å residues were selected to illustrate ADRP-drug interactions. Chimera was also used to generate the charged potential surface of the protein to show how the ligands bind in the deep cavity of the ADRP domain. A detailed 2D interaction diagram was generated using Discovery Studio Visualizer, which displays various interactions, such as hydrogen bonds, interactions with halogen and alkyl groups and van der Waals interactions. This interaction analysis was carried out to check whether or not our predicted drugs bind to the catalytic residues.

2.5 Conformational analysis

Molecular dynamics simulation (MDS) is widely used to investigate the conformational dynamics and stability of protein-ligand complexes [22, 23] and can describe atomic level changes over time, following ligand binding. Here, we used MDS to track atomic changes and to predict the stability of the protein-ligand complexes. The two drug complexes (ADRP-saquinavir and ADRP-trifluperidol) and the substrate complex ADRP-ADPr were used for 100 ns MDS analysis using Gromacs [24]. The topology of the ligands was generated using the ProDRG server [25] and protein topology was generated with the GROMOS 9653a6 force-field [26], using Gromacs. All of the systems were placed in a dodecahedron box and solvated using the SPC water model. The systems were then neutralized by addition of 0.15 nM Na^+ and Cl^- ions and used for energy minimization. The energy minimization removed all steric hindrances and clashes of systems that appeared after addition of solvent and ions. NVT (Number of Particles, Volume and Temperature) and NPT (Number of Particles, Pressure and Temperature) simulations of 100 ps were then carried out to fix the volume,

temperature and pressure of all of the systems. These equilibrated systems were then used for the final production run of 100 ns, and the trajectories was recorded in 2 fs interval.

2.6 Analysis of MDS

The trajectories were preprocessed using the *gmx trjconv* tool before analysis. The artifacts and periodic boundary condition errors were removed from the trajectory and then the processed trajectories were used for further analysis. Various type of analysis was carried out to predict the dynamics of the systems. The *gmx rms, gmx rmsf, gmx gyration, gmx sasa* and *gmx hbond* functions were used to analyze root mean square deviation, root mean square fluctuation, solvent accessible surface area and hydrogen bonds, respectively. Principal component analysis (PCA) was carried out using the *gmx covar* and *gmx anaeig* tools of Gromacs to understand the correlated motions that are induced after ligand binding. The trajectories were visualized using Chimera [27] and Visual Molecular Dynamics software [28].

3 Results and discussion

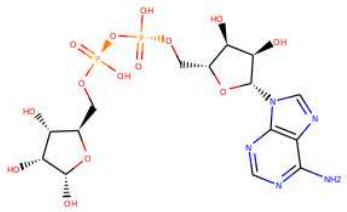
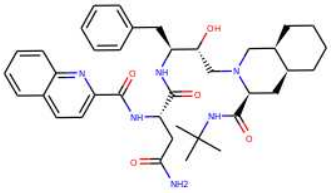
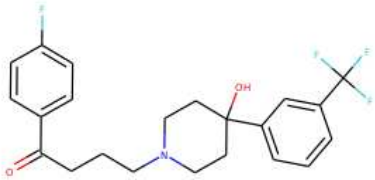
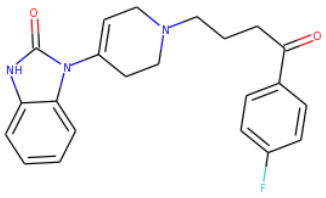
3.1 Virtual screening

Virtual screening was carried out to identify compounds that bind to the ADRP macrodomain subunit of the SARS-CoV-2 Nsp3 enzyme. The binding energies of the 2892 FDA-approved drugs were between −10.3 and −2.5 Kcal/mol. In the virtual screen, saquinavir had the highest binding energy (−10.3 Kcal/mol) and Cysteamine had the lowest binding energy (−2.5 Kcal/mol). We selected the top 20 compounds that showed higher binding affinity than the substrate (Table 1). These top 20 compounds showed binding affinities in the range −10.3 to −9.5 Kcal/mol, which is higher than that of the control compound ADPr (−9.1 Kcal/mol). These top 20 compounds are already FDA-approved drugs and are used in a wide range of therapeutic settings (Table 1). We then selected the top four of these 20 compounds on the basis of binding affinity and analyzed them in more detail. The detailed interaction analysis showed that all these compounds bind to the key catalytic residues located within the deep binding groove of the protein. Drug names, detailed interactions and binding affinities are provided in Table 2.

Table 1: List of FDA-approved drugs showing remarkable binding affinity to the ADRP macrodomain of NSP3.

Sr. No.	ZINC ID	Drug Bank ID	Drug Name	Binding Affinity (Kcal/mol)	Therapeutics
1.	ZINC03914596	DB01232	Squanavir	-10.3	Saquinavir is an HIV-1 protease inhibitor.
2.	ZINC00538505	DB13552	Trifluperidol	-10.2	It is used in the treatment of psychoses including mania and schizophrenia.
3.	ZINC01481815	DB01609	Deferasirox	-10	It is used as an iron chelator.
4.	ZINC19796080	DB00450	Droperidol	-10	Used to maintain the patient in a calm state of neuroleptanalgesia with indifference to surroundings.
5.	ZINC00968279	DB00197	Troglitazone	-9.9	Used for Type II diabetes mellitus. But now it is withdrawn.
6.	ZINC11616581	DB01252	Mitiglinide	-9.8	Used for Type II diabetes mellitus.
7.	ZINC01481956	DB01267	Paliperidone	-9.7	Used for neurological disorders.
8.	ZINC01489478	DB01261	Sitagliptin	-9.7	Used for Type II diabetes mellitus.
9.	ZINC05844792	DB04861	Nebivolol	-9.6	Used for kidney disorder
10.	ZINC01996117	DB00496	Darifenacin	-9.5	Used for the treatment of urinary incontinence.
11.	ZINC03817234	DB04835	Celsentri	-9.5	Used in the HIV medication.
12.	ZINC03830974	DB01167	Itraconazole	-9.5	Used for the fungal infection.
13.	ZINC03869855	DB00266	Dicumarol	-9.5	Used as an oral anticoagulant agent.
14.	ZINC06716957	DB04868	Nilotinib	-9.5	Act as a tyrosine kinase inhibitor and act as a possible medication for chronic myelogenous leukemia (CML).
15.	ZINC00523926	DB01120	Gliclazide	-9.4	Used for the treatment of non-insulin-dependent diabetes mellitus (NIDDM).
16.	ZINC00601317	DB01501	Difenoxin	-9.4	Used as an anti-diarrheal drug.
17.	ZINC03831258	DB04823	Oxyphenisatine	-9.4	A laxative that undergoes enterohepatic circulation.
18.	ZINC18098320	DB00878	Chlorhexidine	-9.4	Act as an antimicrobial agent and generally used by dentist.
19.	ZINC01550477	DB01259	Lapatinib	-9.4	It is an anti-cancer drug.
20.	ZINC03871723	DB04794	Bifonazole	-9.3	It is an azole antifungal drug.

Table 2: The table showing the common name of drug, its structure, binding affinity from Autodock Vina and interacting residues.

Drug Name	**Structure**	**Binding energy**	**Residues**
ADPr		-9.1	Ala21, Asp22, Ile23, Ala38, Asn40, Lys44, His45, Gly46, Gly47, Gly48, Val49, Leu126, Ser128, Ala129, Gly130, Ile131, Phe132, Ala154, Phe156 and Leu160
Saquinavir		-10.3	Asp22, Ile23, Ala38, Asn40, Gly46, Gly48, Gly47, Val49, Ala52, Pro125, Leu126, Ser128, Ala129, Gly130, Ile131, Phe132, Ala154, Asp157 and Phe156
Trifluperidol		-10.2	Asp22, Ile23, Ala38, Asn40, Gly46, Gly47, Gly48, Val49, Gly51, Ala52, Pro125, Leu126, Ser128, Ala129, Gly130, Ile131, Phe132, Ala154, Val155, Phe156 and Asp157
Deferasirox		-10.0	Ala38, Gly48, Val49, Pro125, Leu126, Ser128, Ala129, Gly130, Ile131, Phe132, Gly133, Val155, Phe156, Asp157 and Leu160
Droperidol		-10.0	Ile23, Ala38, Gly48, Val49, Pro125, Leu126, Ser128, Ala129, Gly130, Ile131, Phe132, Ala154, Val155, Phe156, Asp157 and Leu160

3.2 Analysis of interactions

We selected the top four compounds and compared in detail their interactions with that of ADPr. The results are described individually below.

3.2.1 ADRP-ADPr

We analyzed the crystal structure of the ADRP-ADPr complex and then compared this with the binding of the docked ligands. The ADRP-ADPr complex showed 12 hydrogen bonds between ADPr and various residues in the ADRP domain and was also stabilized by other interactions between the ligand and ADRP domain. The residues that are involved in binding of ADPr to ADRP are Ala21, Asp22, Ile23, Ala38, Asn40, Lys44, His45, Gly46, Gly47, Gly48, Val49, Leu126, Ser128, Ala129, Gly130, Ile131,Phe132, Ala154, Phe156 and Leu160.

3.2.2 ADRP-saquinavir

Saquinavir, an anti-viral drug that is used to control HIV, was the top compound in our virtual screen. Saquinavir showed higher binding affinity than the control compound ADPr, indicating that it can bind competitively in the active site and inhibit the function of the ADRP macrodomain of Nsp3. The complex, which has a binding affinity of −10.3 Kcal/mol, shows interactions between saquinavir and various key catalytic residues and is also stabilized by several other interactions. Gly130 and Leu126 form hydrogen bonds with saquinavir, and Val49 and Phe132 form π and σ interactions with saquinavir. Other residues, including Asp22, Ile23, Ala38, Asn40, Gly46, Gly48, Gly47, Ala52, Pro125, Ser128, Ala129, Ile131, Ala154, Asp157 and Phe156, are also involved in the interaction between saquinavir and the ADRP domain. Saquinavir binds to the key catalytic residues that also participate in ADPr binding, indicating that saquinavir can inhibit the activity of the ADRP macrodomain and can potentially inhibit the Nsp3 protein of COVID-19. We found that saquinavir interacts with several residues that are also involved in ADPr binding in the crystal structure, showing that the drug is binding in the ADPr binding cavity and can act as a competitive inhibitor. The detail is shown in Figure 2.

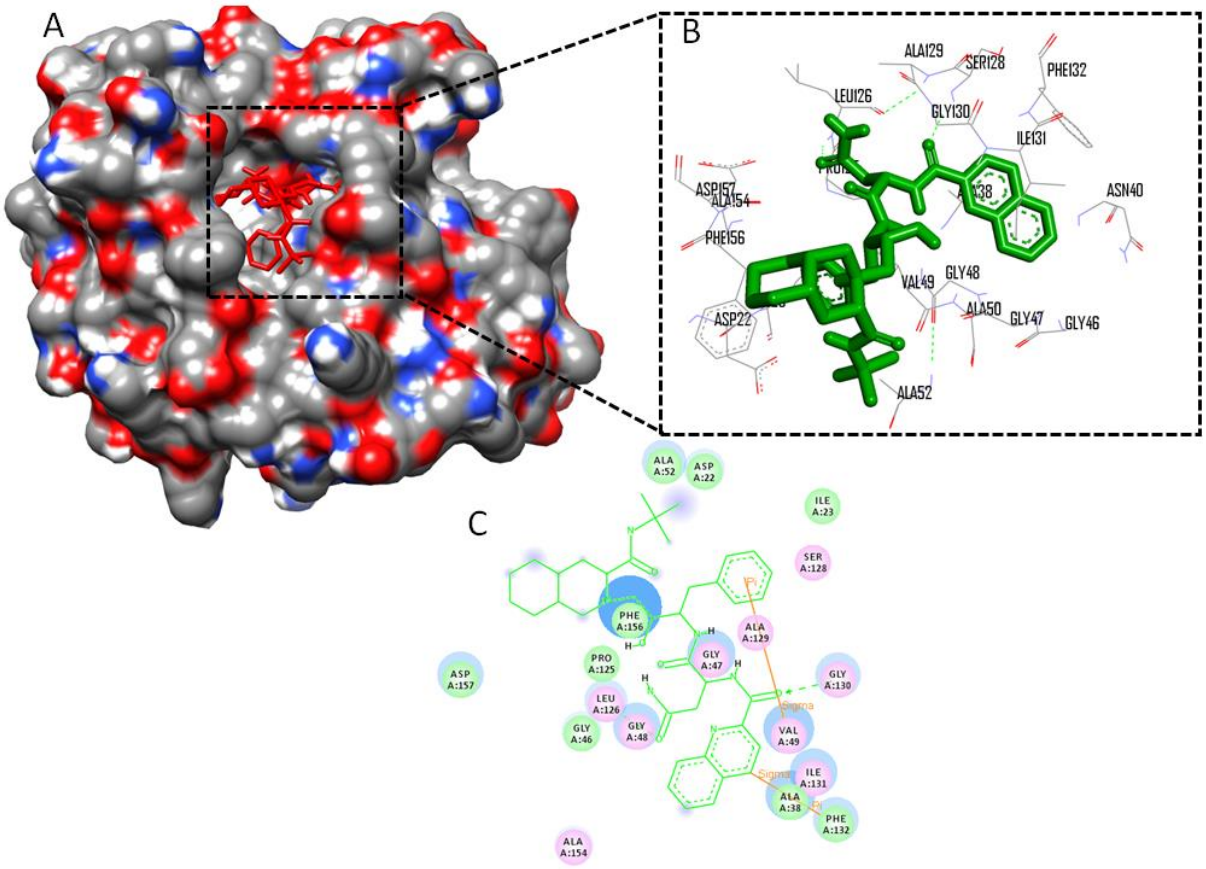


Figure 2: Binding pattern of Saquinavir with ADRP. (A) Surface view of ADRP-Saquinavir complex. (B) Residue interaction diagram of Saquinavir with ADRP. (C) 2D interaction diagram of ADRP-Saquinavir interaction.

3.2.3 ADRP-trifluperidol

Trifluperidol was the second best hit in the virtual screen. The ADRP-trifluperidol complex, which has a binding affinity of −10.2 Kcal/mol, is stabilized by three hydrogen bonds and several hydrophobic interactions. Gly48, Val49 and Phe156 form hydrogen bonds with trifluperidol and a π-π interaction was seen with Ile131. The complex was also stabilized by interactions with Asp22, Ile23, Ala38, Asn40, Gly46, Gly47, Gly51, Ala52, Pro125, Leu126, Ser128, Ala129, Gly130, Phe132, Ala154, Val155 and Asp157. Trifluperidol binds to the key catalytic residues and shows higher binding affinity than ADPr, indicating that it too is a potential inhibitor of the ADRP macrodomain and can render the Nsp3 protein inactive. Several residues that interact with trifluperidol are also involved in ADPr binding in the crystal structure, showing that the drug is binding in the ADPr binding cavity and can act as a competitive inhibitor. The detail is shown in Figure 3.

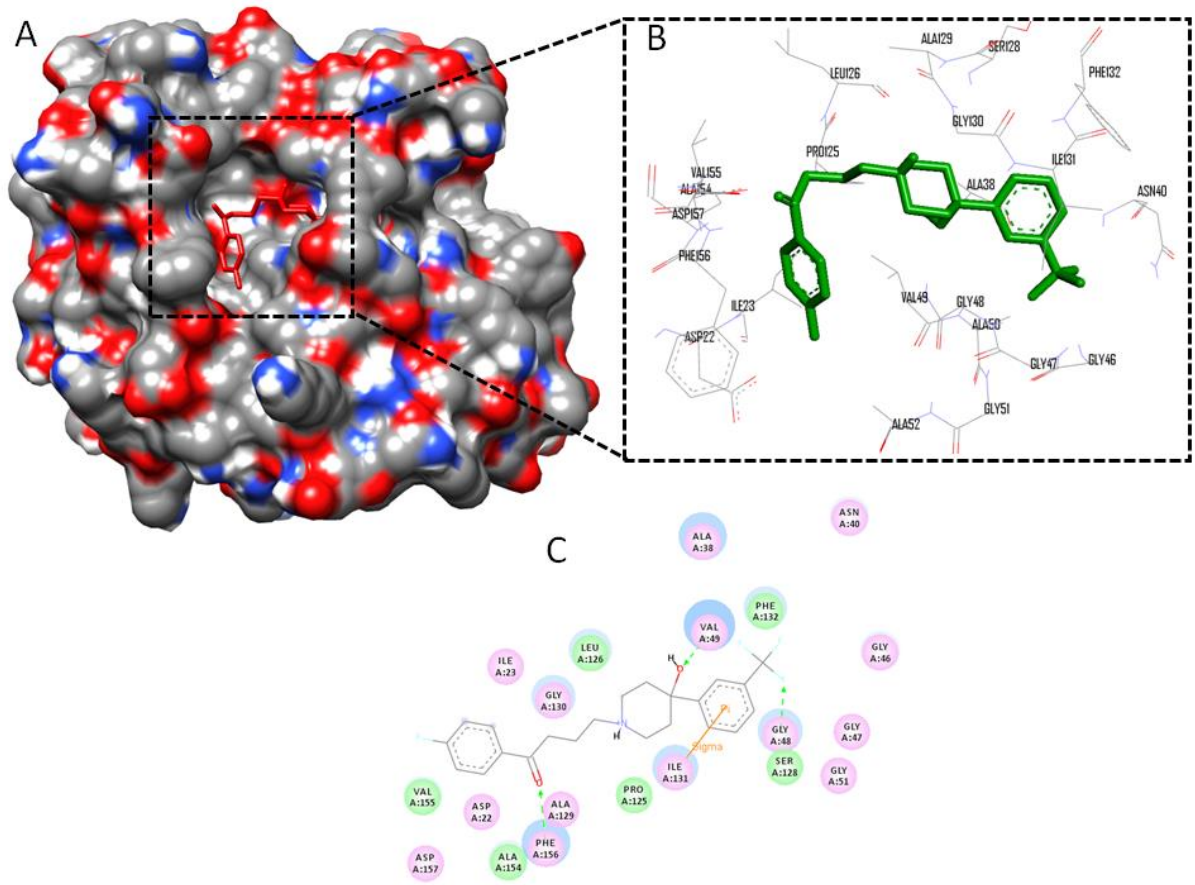


Figure 3: Binding pattern of Trifluperidol with ADRP. (A) Surface view of ADRP-Trifluperidol complex. (B) Residue interaction diagram of Trifluperidol with ADRP. (C) 2D interaction diagram of ADRP- Trifluperidol interaction.

3.2.4 ADRP-deferasirox

The ADRP-deferasirox complex, which has a binding affinity of −10.0 Kcal/mol and is among the top hits in the virtual screen, is stabilized by various interactions, including one hydrogen bond with Gly130 and one π-π interaction with Ile131. Other residues, including Ala38, Gly48, Val49, Pro125, Leu126, Ser128, Ala129, Phe132, Gly133, Val155, Phe156, Asp157 and Leu160, stabilize the complex through various interactions. Deferasirox also binds to the key catalytic residues, indicating that it can also act as a good inhibitor of the ADRP macrodomain of Nsp3.

3.2.5 ADRP-droperidol

Droperidol has a binding affinity of −10.0 Kcal/mol, calculated using Autodock Vina, and binds to various key catalytic residues, indicating that it also binds in the substrate binding

cavity. Droperidol forms only one hydrogen bond, with Gly130, and other interactions with Ile23, Ala38, Gly48, Val49, Pro125, Leu126, Ser128, Ala129, Ile131, Phe132, Ala154, Val155, Phe156, Asp157 and Leu160 play a role in stabilizing this complex. Again, we found that several residues involved in binding droperidol are also involved in ADPr binding in the crystal structure, indicating that the drug is binding in the ADPr binding cavity and can act as a competitive inhibitor.

From all of these analyses, we selected only two drugs (saquinavir and trifluperidol) for further analysis because they show good binding affinity for ADRP and bind in the deep groove containing the key catalytic residues. We compared the ADRP-saquinavir and ADRP-trifluperidol complexes with the ADRP-ADPr complex in the 100 ns MDS to analyze the stability of the protein-ligand complexes.

3.3 Conformational analysis

The natural substrate (ADPr) and the two drugs with the highest binding affinity (saquinavir and trifluperidol) were used for MDS studies to investigate the binding mechanism, conformational dynamics and stability of the ADRP-ligand complexes. Three systems (ADRP-ADPr, ADRP-saquinavir and ADRP-trifluperidol) were prepared and used for the 100 ns MDS studies. Each system produced stable trajectories, which were then used for analysis. All the analyses were performed after the system attained equilibrium. Root mean square deviation (RMSD), root mean square fluctuation (RMSF), radius of gyration (Rg), number of hydrogen bonds, SASA, PCA and binding free energy were calculated and analyzed in detail.

3.3.1 Stability analysis

Deviation of the protein backbone from its initial state was calculated to determine structural stability. The calculated RMSD values were plotted for the whole duration of the 100 ns simulation of each system. The RMSD describes the conformational changes of a given protein over time. The complexes showed few differences at each time step, but the overall RMSD of each complex at any point of time was < 0.26 nm. The average RMSD values for the ADRP domain complexed with ADPr, trifluperidol and saquinavir were 0.147 nm, 0.164 nm and 0.175 nm, respectively. The RMSD values for each complex are thus comparable. The ADRP-ADPr complex showed the lowest average RMSD value of all the complexes, showing that the ADRP-substrate complex is slightly less stable than the ADRP-drug

complexes. The overall patterns of the RMSD (Figure 4A) are similar for the predicted drugs and the co-crystallized control ligand.

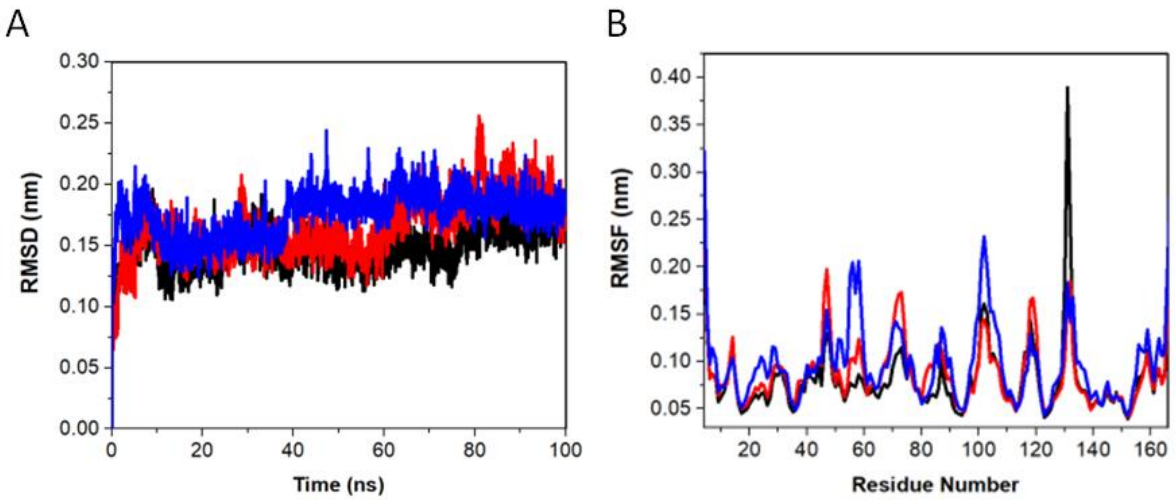


Figure 4: Stability and Flexibility analysis. (A) RMSD of Cα backbone of ADRP and ligand complexes for 100 ns. (B) RMSF of Cα atoms of ADRP and ligand complexes. The black, red and blue represent ADRP-ADPr, ADRP-Trifluperidol ADRP-Saquinavir respectively.

The RMSD for each ligand was also calculated to predict ligand fluctuation in the binding pocket during the simulation. The average RMSD value of ADPr, trifluperidol and saquinavir were 0.31 nm, 0.23 nm and 0.30 nm, respectively. Slight conformational deviations in the ligand RMSD values of ADPr and saquinavir were observed, although each ligand formed a stable complex with ADRP during the simulation. Trifluperidol had a lower RMSD than saquinavir or ADPr, and saquinavir also showed a slightly lower RMSD than the control ligand. The overall RMSD results of the ADRP-ligand complexes and the unbound ligands showed that all trajectories reached equilibrium from the initial point of simulation and produced a stable trajectory throughout the analysis. In both analyses we saw that trifluperidol had a lower RMSD value than saquinavir and is thus the more stable protein-ligand complex.

3.3.2 Flexibility analysis

RMSF values were calculated to investigate changes inflexibility of the protein after ligand binding. RMSF values should be high for well-organized structures, such as α-helixes and β-sheets and low for loosely organized structures, such as turns, coils and loops. The average RMSF values for ADRP complexes with ADPr, trifluperidol and saquinavir were 0.082 nm, 0.088 nm and 0.1 nm, respectively. These values, which are different for each protein-ligand

complex, clearly indicate that ligand binding induces conformational changes in the ADRP domain (Figure 4B). The ADRP-ADPr complex showed the highest RMSF peak for residues 130 to 132, whereas the highest RMSF values for the ADRP-saquinavir complex were between residues 54–59 and 100–103, with a high deviation in the peak in the C-terminal region. The ADRP-trifluperidol complex showed high RMSF values for residues 45–47, 70–73 and 117–119. As we have already seen in the docking section, some of these residues belong to the catalytic core of the ADRP domain. It is widely acknowledged that the native function of ADRP, or any other enzyme, requires a specific conformation and, as we can see from the RMSF results, drug binding alters the conformation and induces indrances in the native dynamics of the protein. This means that ADRP cannot perform its native phosphatase activity, which may affect survival of the virus in the host because of inactivation of the Nsp3 enzyme. The RMSF results allow us to conclude that these drugs may inhibit ADRP activity and that the ADRP-trifluperidol complex is more stable than the ADRP-saquinavir complex.

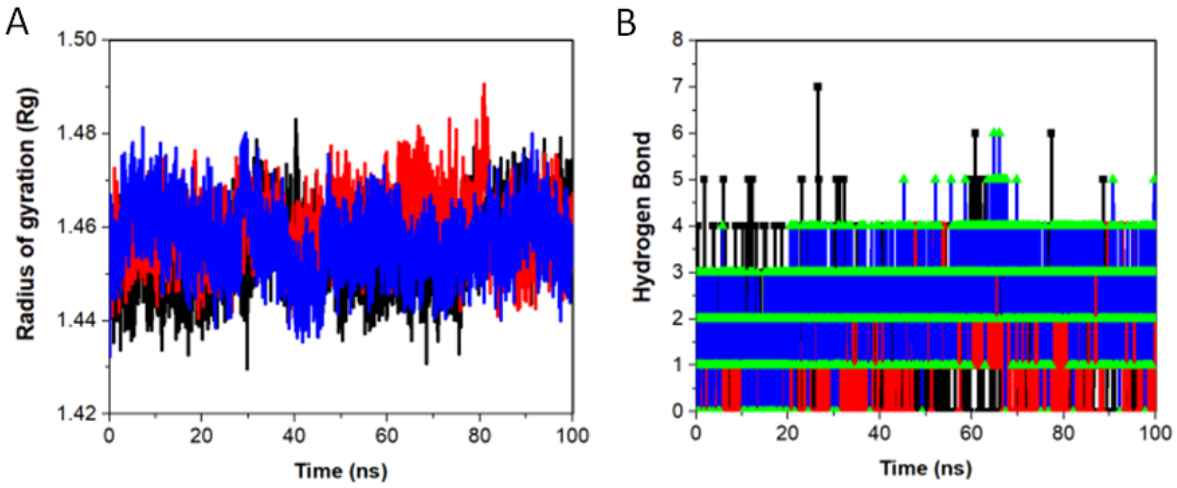


Figure 5: Compactness and Interaction analysis. (A) Plot of Radius of gyration *vs.* time for ADRP with ligands. **(B)** Number of hydrogen bonds between protein and ligands. The black, red and blue represents ADRP-ADPr, ADRP-Trifluperidol ADRP-Saquinavir respectively.

3.3.3 Compactness analysis

Rg is the best parameter to describe the compactness of a protein after ligand binding. Here, we predicted Rg values to investigate changes in compactness after ligand binding. It is assumed that smaller Rg values represent tightly packed protein structures, and vice versa.

The average Rg values for ADRP complexes with ADPr, trifluperidol and saquinavir were 1.454 nm, 1.459 nm and 1.456 nm, respectively (Figure 5A). The pattern of Rg values is very similar for all the ADRP-ligand complexes, with only minor differences. This indicates that all the ADRP-ligand complexes are stable and compact throughout the simulation and that ligand binding does not induce major changes in the compactness of the ADRP domain. The average Rg values are in the order ADRP-ADPr complex <ADRP-saquinavir <ADRP-trifluperidol, showing that all the complexes are stable and that saquinavir and trifluperidol could be used as lead compounds in the design of ADRP domain inhibitors.

3.3.4 Interaction analysis

Hydrogen bonds are very important and transient interaction in protein-ligand complexes. Since they provide an indication of the stability of the protein-ligand complex, we also calculated the number of hydrogen bonds in each complex (Figure 5B). As shown in Figure 5B the ADRP-ADPr complex has the highest number of hydrogen bonds, and the ADRP-saquinavir complex has more hydrogen bonds than the ADRP-trifluperidol complex while average number of hydrogen bonds for ADRP-ADPr, ADRP-trifluperidol and ADRP-saquinavir was 2, 2 and 3. All the complexes thus have a good number of hydrogen bonds and are stable within the ADRP domain binding cavity throughout the whole simulation.

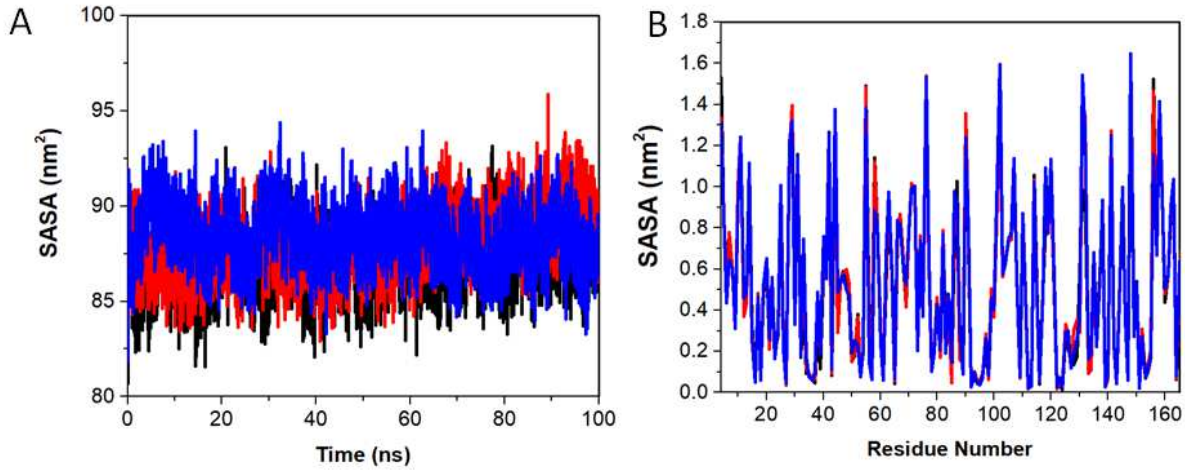


Figure 6: Solvent accessible surface area. (A) SASA value *vs.* time for all the complex for 100 ns. **(B)** The residue *vs.* SASA value for ADRP-ligand complex. The black, red and blue represents ADRP-ADPr, ADRP-Trifluperidol ADRP-Saquinavir respectively.

3.3.5 Solvent accessible surface area analysis

We also calculated values of SASA, which represent the area that is accessible to the solvent. We predicted SASA values to investigate ligand-induced changes in the ADRP domain (Figure 6A). The average values of SASA of both drugs were different from that of the control ligand ADPr, although the differences were not statistically significant. The average values of SASA for ADRP-ADPr, ADRP-trifluperidol and ADRP-saquinavir were 86.94 nm^2, 88.04 nm^2 and 88.39 nm^2, respectively. The SASA of the ADRP-trifluperidol complex was thus smaller than that of the ADRP-saquinavir complex. We can thus conclude that both of the ADRP-drug complexes are stable, and that ADRP-trifluperidol is more stable than ADRP-saquinavir.

We also calculated the residue SASA value, which indicates the SASA value on the basis of residues instead of time (Figure 6B). The average values for ADRP-ADPr, ADRP-trifluperidol and ADRP-saquinavir were 0.52 nm, 0.53 nm and 0.53 nm, respectively. The ADRP-saquinavir complex shows much greater fluctuation in several residues, compared with the other complexes. The ADRP-ADPr complex had a smaller residue SASA value, representing a more stable complex. We also found that the ADRP-trifluperidol complex is more stable than the ADRP-saquinavir complex.

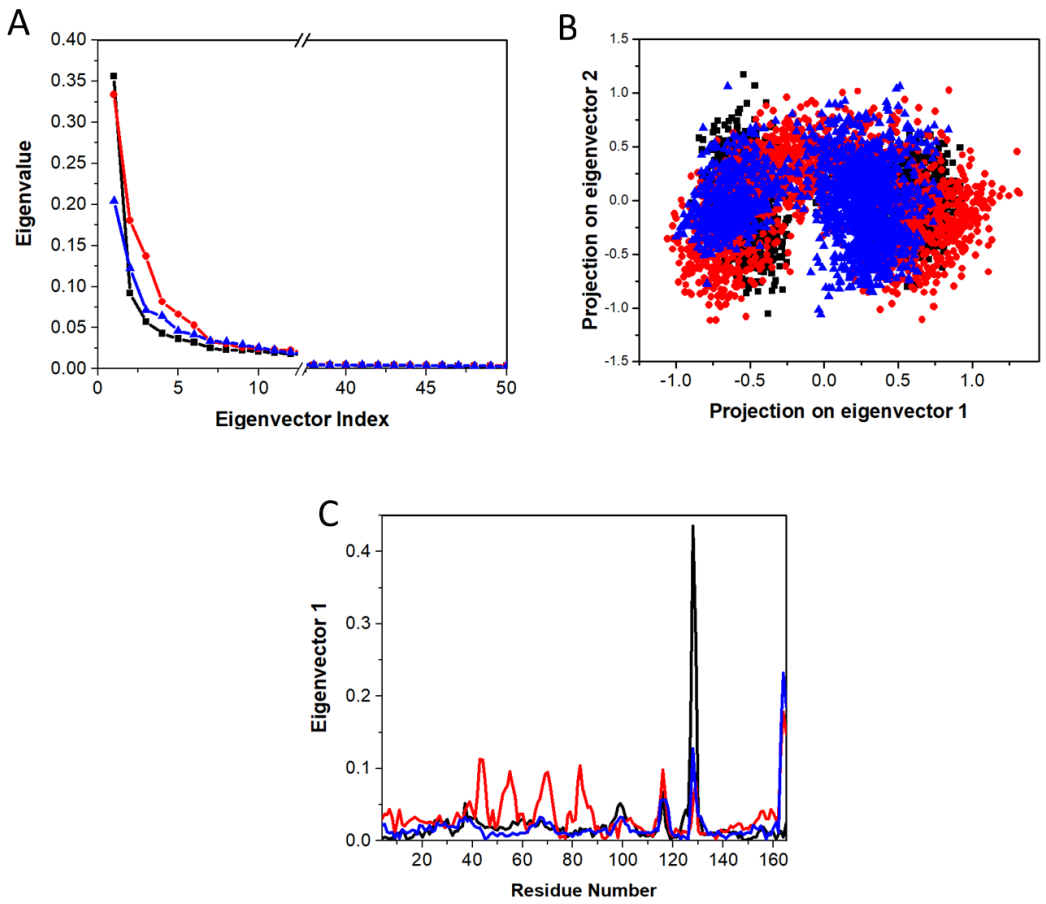


Figure 7: Principal Component analysis. (A) First 50 eigenvectors were plotted *vs.* eigenvalue for ADRP-ligand complexes. **(B)** Projection of the motion of the protein in phase space along PC1 *vs*. PC2 for ADRP-ligand complexes. **(C)** The first eigenvector *vs*. residue to see the residue wise fluctuation. The black, red and blue represent ADRP-ADPr, ADRP-Trifluperidol ADRP-Saquinavir respectively.

3.3.6 Principal component analysis

PCA or essential dynamics analysis was performed to track changes in the correlated motions after ligand binding. Eigenvectors *vs.* eigenvalues were calculated for the whole trajectory and the plots are shown in Figure 7A. We included only the first 50 eigenvectors in the analysis, and for the best representation of the graph, since the first eigenvectors represent the overall dynamics of the protein. The covariance matrix of atomic fluctuations was diagonalized to predict the eigenvalues. The plot shows the eigenvalue corresponding to each eigenvector for the ADRP-ADPr, ADRP-trifluperidol and ADRP-saquinavir complexes. ADRP-ADPr showed the highest value for the first eigenvectors and the ADRP-saquinavir showed the smallest value for first eigenvectors. The first eigenvector value of the ADRP-

trifluperidol complex lies between the values of the other two complexes. For further analysis, we thus selected the first 10 eigenvectors and calculated the percentage motions for all three systems. The ADRP-ADPr, ADRP-trifluperidol and ADRP-saquinavir complexes accounted for 70.29%, 72.60% and 66.31%, respectively, of motions in the whole trajectory for the first 10 eigenvectors. Saquinavir thus induced less correlated motions than the other ligands and we can conclude that, although all the complexes are stable, the ADRP-saquinavir complex is more stable than the ADRP-trifluperidol complex in terms of PCA analysis.

From the above analysis, we can see that the first eigenvectors are very important for characterizing the overall essential dynamics of the protein-ligand complex. We therefore selected the first eigenvectors and plotted these against each other (Figure 7B). The ADRP-saquinavir complex showed a very stable cluster in the phase space, compared with the other complexes; otherwise both complexes showed the same type of pattern for the cluster. This indicates that the ADRP-saquinavir complex is more stable than the ADRP-trifluperidol complex and induces less correlated motions. The ADRP-trifluperidol complex is also stable because it does not show the abrupt pattern and does not cause very high fluctuations in the phase space. PCA analysis thus suggests that both drugs form stable complexes with ADRP and that the ADRP-saquinavir complex is more stable.

We also calculated motions on the basis of residues. The eigRMSF values were calculated only for first eigenvector on the basis of residues and are shown in Figure 7C. The average eigRMSF values for the ADRP-ADPr, ADRP-trifluperidol and ADRP-saquinavir complexes were 0.022 nm, 0.034 nm and 0.021 nm, respectively. Compared with the other complexes, the ADRP-ADPr complex showed a very high eigRMSF value between residues 128–133. Trifluperidol also induced motions in residues 40–48, 50–60, 62–77 and 78–92. These residues are within the catalytic region of the protein and the result indicates that binding of trifluperidol alters the original confirmation of the active site and induces conformational changes in the active site residues. We can, therefore, say that trifluperidol is a good inhibitor of ADRP. The eigRMSF showed a similar pattern of residue fluctuation to the RMSF analysis.

Conclusion

The COVID-19 pandemic is spreading rapidly day-by-day across the globe. SARS-CoV-2 has multiple Nsps, of which Nsp3 is a multidomain complex that regulates RNA

transcription. The macrodomain ADRP plays a key role in this process by removing ADP-ribose from ADP-ribosylated proteins and RNA. The ADRP domain can thus be regarded as a viable drug target and we screened 2892 FDA-approved compounds against the ADRP domain using the SBVS approach. The twenty top energy compounds were selected for further analysis. From these compounds, trifluperidol and saquinavir were chosen for further validation since they showed good binding affinity and interacted with the key catalytic residues of the macrodomain. The complexes of these two compounds with ADRP were compared with the ADP-ADRP complex in 100 ns MDS studies and various parameters, including RMSD, RMSF, Rg, SASA, number of hydrogen bonds, PCA were analyzed. The results suggest that the ADRP-trifluperidol complex is more stable than the ADRP-saquinavir complex. We predict, that trifluperidol could be repurposed as an inhibitor of the catalytic activity of the ADRP domain of the Nsp3 protein to control the spread of COVID-19. We acknowledge that this is a computational study and hope that experimental validation of this drug will be carried out by other scientists.

Conflict of interest: None

Author contribution: AP performed all the experiments and drafted the manuscript.

References

[1] Zheng J. SARS-CoV-2: an Emerging Coronavirus that Causes a Global Threat. Int J Biol Sci. 16(10): 1678–1685 (2020).

[2] Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, *et al.* Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. The Lancet. 395(10223): 497–506 (2020).

[3] Wu F, Zhao S, Yu B, Chen Y-M, Wang W, Song Z-G, *et al.* A new coronavirus associated with human respiratory disease in China. Nature. 579(7798): 265–269 (2020).

[4] Cui J, Li F, Shi Z-L. Origin and evolution of pathogenic coronaviruses. Nat Rev Microbiol. 17(3): 181–192 (2019).

[5] Baez-Santos YM, St John SE, Mesecar AD. The SARS-coronavirus papain-like protease: structure, function and inhibition by designed antiviral compounds. Antiviral Res. 115: 21–38 (2015).

[6] Lee HJ, Shieh CK, Gorbalenya AE, Koonin EV, La Monica N, Tuler J, *et al.* The complete sequence (22 kilobases) of murine coronavirus gene 1 encoding the putative proteases and RNA polymerase. Virology. 180(2): 567–582 (1991).

[7] Saikatendu KS, Joseph JS, Subramanian V, Clayton T, Griffith M, Moy K, *et al.* Structural basis of severe acute respiratory syndrome coronavirus ADP-ribose-1''-

phosphate dephosphorylation by a conserved domain of nsP3. Structure. 13(11): 1665–1675 (2005).

[8] Grunewald ME, Chen Y, Kuny C, Maejima T, Lease R, Ferraris D, *et al.* The coronavirus macrodomain is required to prevent PARP-mediated inhibition of virus replication and enhancement of IFN expression. PLoS Pathog. 15(5): e1007756 (2019).

[9] Kuri T, Eriksson KK, Putics A, No Value] Züst R, No Value] Snijder EJ, Davidson AD, *et al.* The ADP-ribose-1''-monophosphatase domains of SARS-coronavirus and human coronavirus 229E mediate resistance to antiviral interferon responses. Journal of General Virology. 92: 1899–1905 (2011).

[10] Michalska K, Kim Y, Jedrzejczak R, Maltseva NI, Stols L, Endres M, *et al.* Crystal structures of SARS-CoV-2 ADP-ribose phosphatase (ADRP): from the apo form to ligand complexes. bioRxiv. 2020.05.14.096081 (2020).

[11] Bahadur Gurung A, Ajmal Ali M, Lee J, Abul Farah M, Mashay Al-Anazi K. Structure-based virtual screening of phytochemicals and repurposing of FDA approved antiviral drugs unravels lead molecules as potential inhibitors of coronavirus 3C-like protease enzyme. J King Saud Univ Sci. (2020).

[12] Padhi AK, Shukla R, Tripathi T. Rational Design of the Remdesivir Binding Site in the RNA-dependent RNA Polymerase of SARS-CoV-2: Implications for Potential Resistance. bioRxiv. 2020.06.27.174896 (2020).

[13] Shamsi A, Mohammad T, Anwar S, AlAjmi MF, Hussain A, Rehman MT, *et al.* Glecaprevir and Maraviroc are high-affinity inhibitors of SARS-CoV-2 main protease: possible implication in COVID-19 therapy. Biosci Rep. 40(6) (2020).

[14] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem. 25(13): 1605–1612 (2004).

[15] Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, *et al.* DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res. 46(D1): D1074–D1082 (2018).

[16] Sterling T, Irwin JJ. ZINC 15 – Ligand Discovery for Everyone. J Chem Inf Model. 55(11): 2324–2337 (2015).

[17] O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR. Open Babel: An open chemical toolbox. J Cheminform. 3: 33 (2011).

[18] Shukla R, Munjal NS, Singh TR. Identification of novel small molecules against GSK3β for Alzheimer's disease using chemoinformatics approach. J Mol Graph Model. 91: 91–104 (2019).

[19] Pathak RK, Gupta A, Shukla R, Baunthiyal M. Identification of new drug-like compounds from millets as Xanthine oxidoreductase inhibitors for treatment of Hyperuricemia: A molecular docking and simulation study. Comput Biol Chem. 76: 32–41 (2018).

[20] Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, *et al.* AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. J Comput Chem. 30(16): 2785–2791 (2009).

[21] Trott O, Olson AJ. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of Computational Chemistry. 31(2): 455–461 (2009).
[22] Shukla R, Shukla H, Kalita P, Tripathi T. Structural insights into natural compounds as inhibitors of Fasciola gigantica thioredoxin glutathione reductase. J Cell Biochem. 119(4): 3067–3080 (2018).
[23] Shukla H, Shukla R, Sonkar A, Tripathi T. Alterations in conformational topology and interaction dynamics caused by L418A mutation leads to activity loss of Mycobacterium tuberculosis isocitrate lyase. Biochem Biophys Res Commun. 490(2): 276–282 (2017).
[24] Abraham MJ, Murtola T, Schulz R, Pall S, Smith JC, Hess B, *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX. 1–2: 19–25 (2015).
[25] Schuittelkopf AW, van Aalten DMF. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. Acta Crystallogr D Biol Crystallogr. 60(Pt 8): 1355–1363 (2004).
[26] Oostenbrink C, Villa A, Mark AE, van Gunsteren WF. A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. J Comput Chem. 25(13): 1656–1676 (2004).
[27] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem. 25(13): 1605–1612 (2004).
[28] Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Mol Graph. 14(1): 33–38, 27–28 (1996).
[29] Kumari R, Kumar R, Lynn A. g_mmpbsa--a GROMACS tool for high-throughput MM-PBSA calculations. J Chem Inf Model. 54(7): 1951–1962 (2014).

Figures

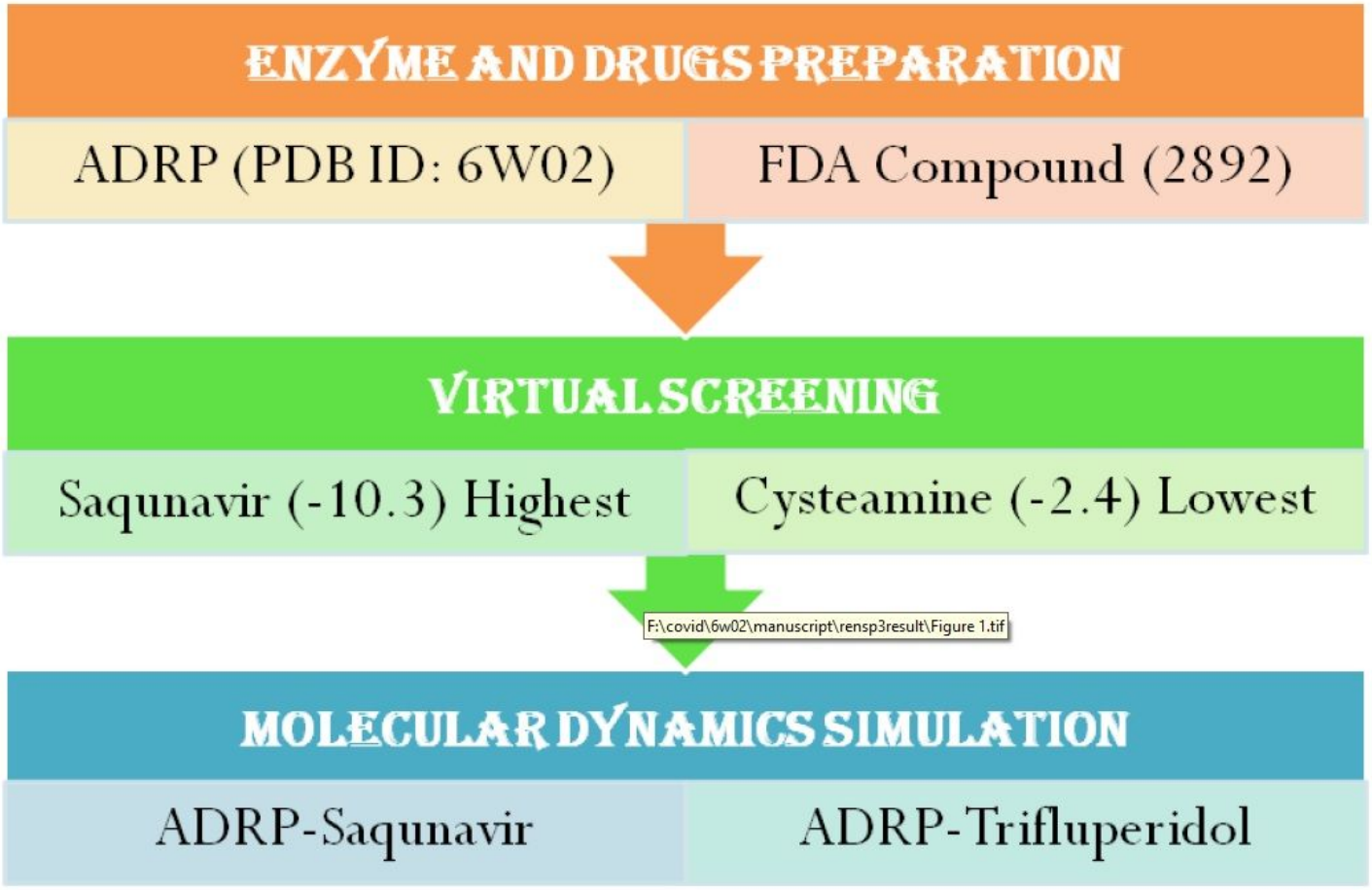


Figure 1

Complete workflow of methodology.

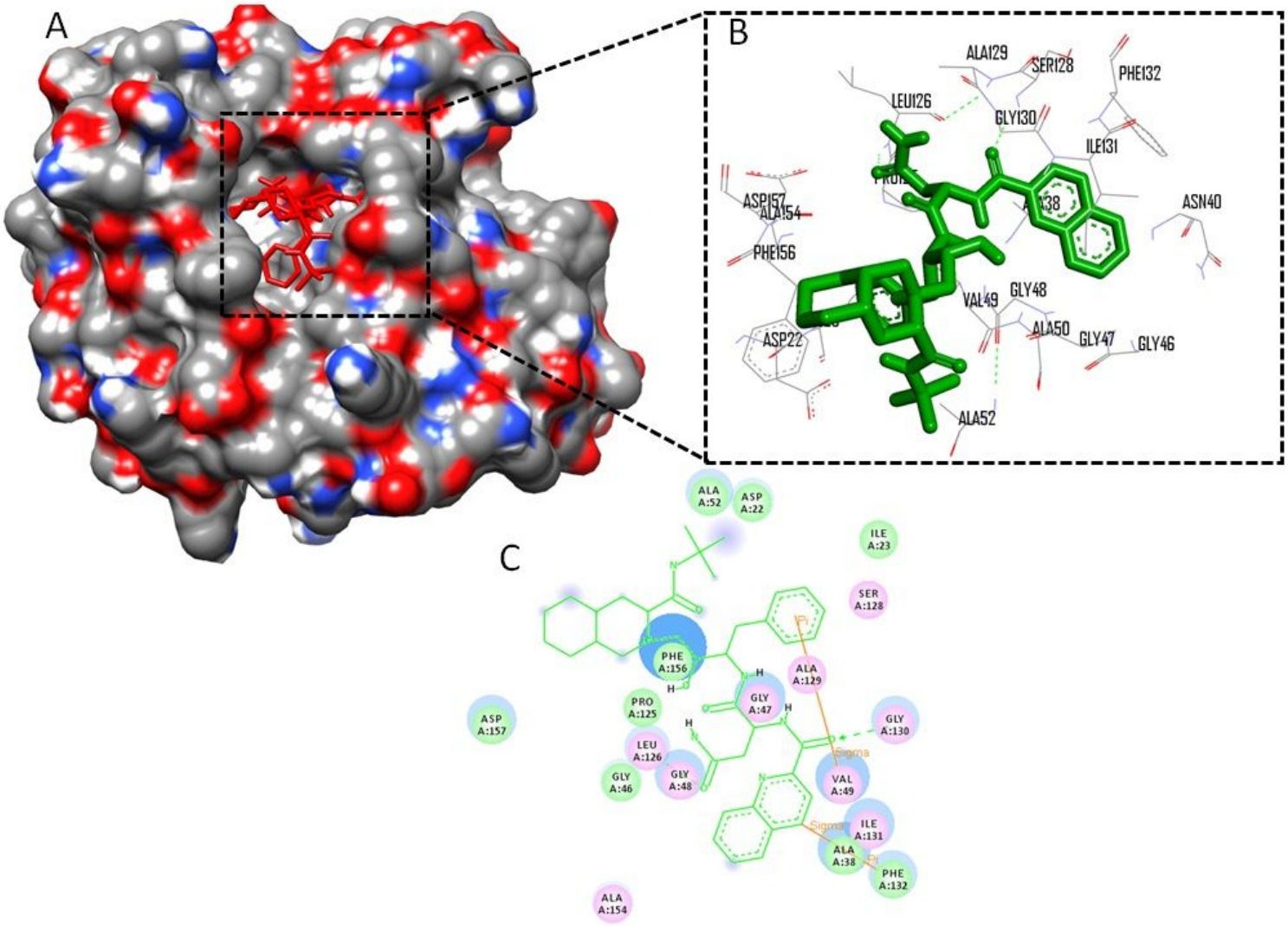


Figure 2

Binding pattern of Saquinavir with ADRP. (A) Surface view of ADRP-Saquinavir complex. (B) Residue interaction diagram of Saquinavir with ADRP. (C) 2D interaction diagram of ADRP-Saquinavir interaction.

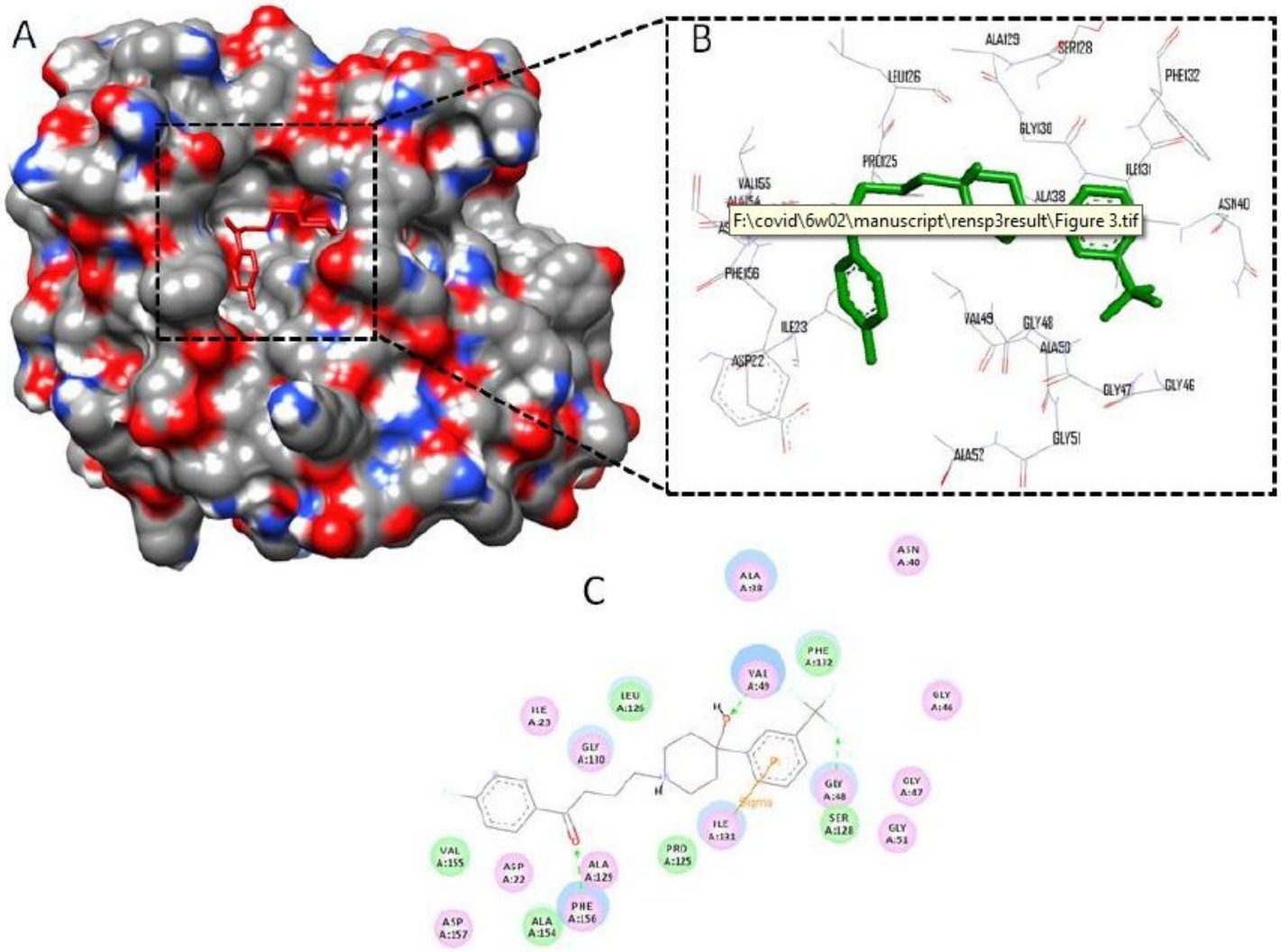


Figure 3

Binding pattern of Trifluperidol with ADRP. (A) Surface view of ADRP- Trifluperidol complex. (B) Residue interaction diagram of Trifluperidol with ADRP. (C) 2D interaction diagram of ADRP- Trifluperidol interaction.

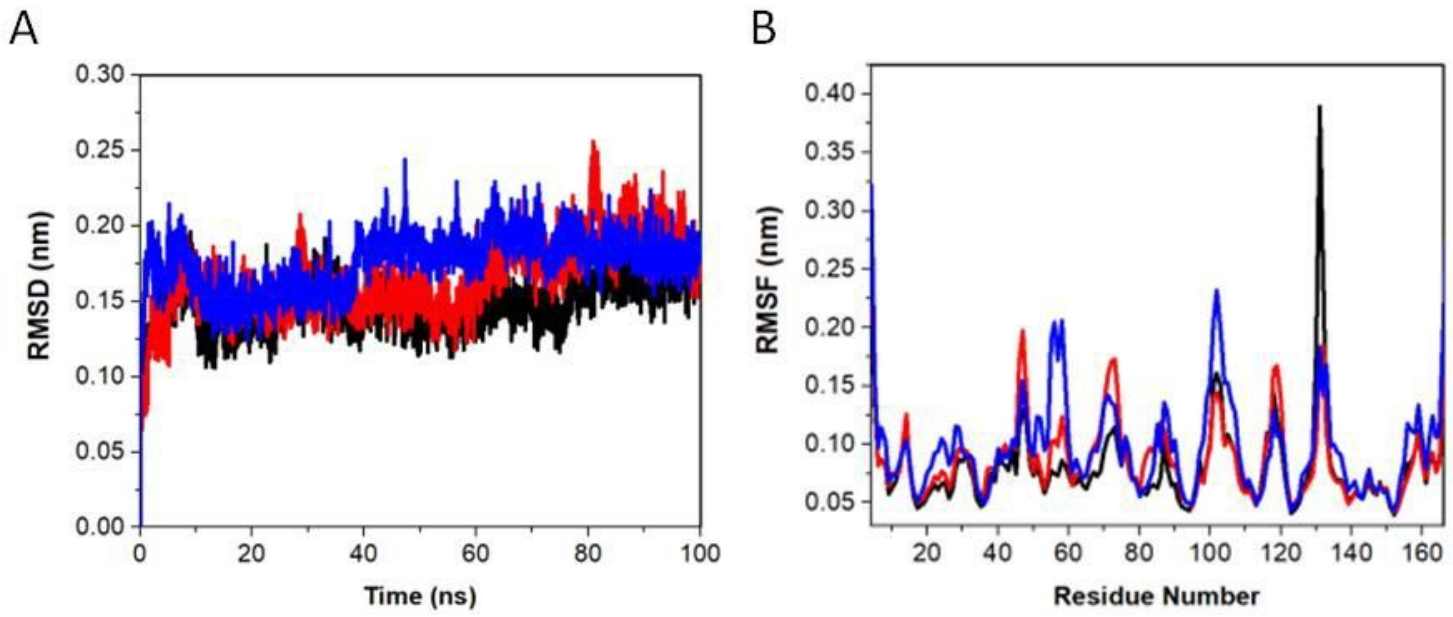


Figure 4

Stability and Flexibility analysis. (A) RMSD of Cα backbone of ADRP and ligand complexes for 100 ns. (B) RMSF of Cα atoms of ADRP and ligand complexes. The black, red and blue represent ADRP-ADPr, ADRP-Trifluperidol ADRP-Saquinavir respectively.

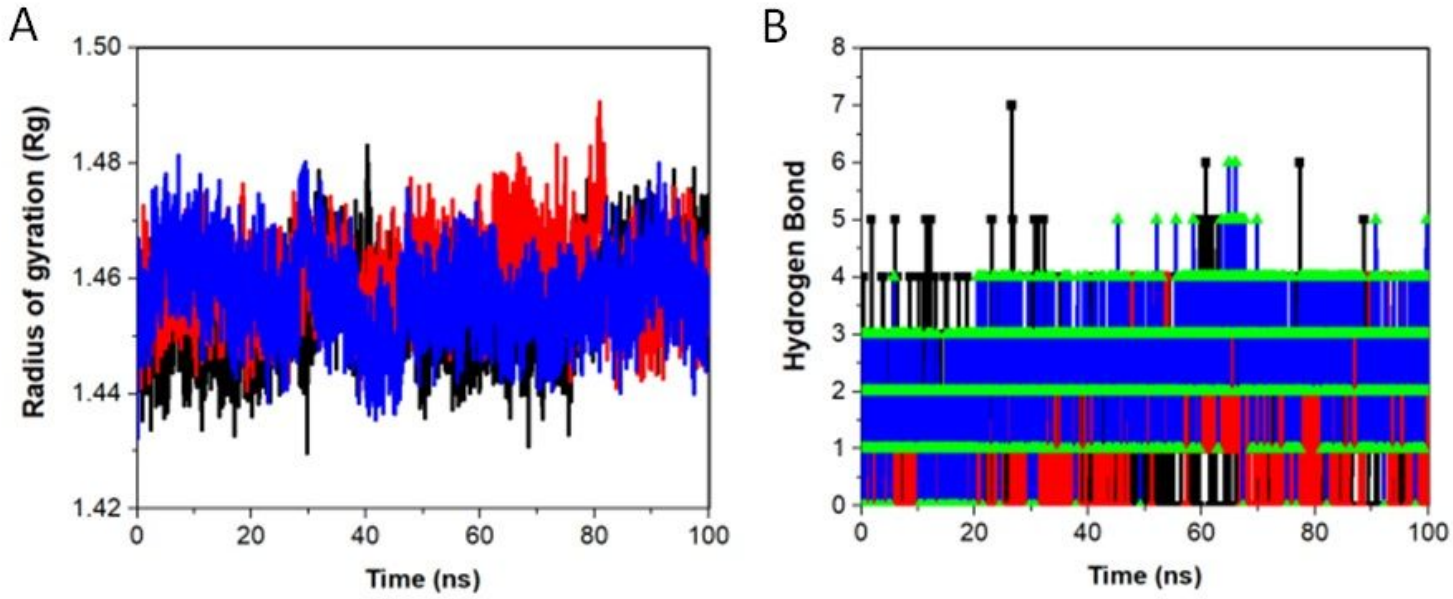


Figure 5

Compactness and Interaction analysis. (A) Plot of Radius of gyration vs. time for ADRP with ligands. (B) Number of hydrogen bonds between protein and ligands. The black, red and blue represents ADRP-ADPr, ADRP-Trifluperidol ADRP-Saquinavir respectively.

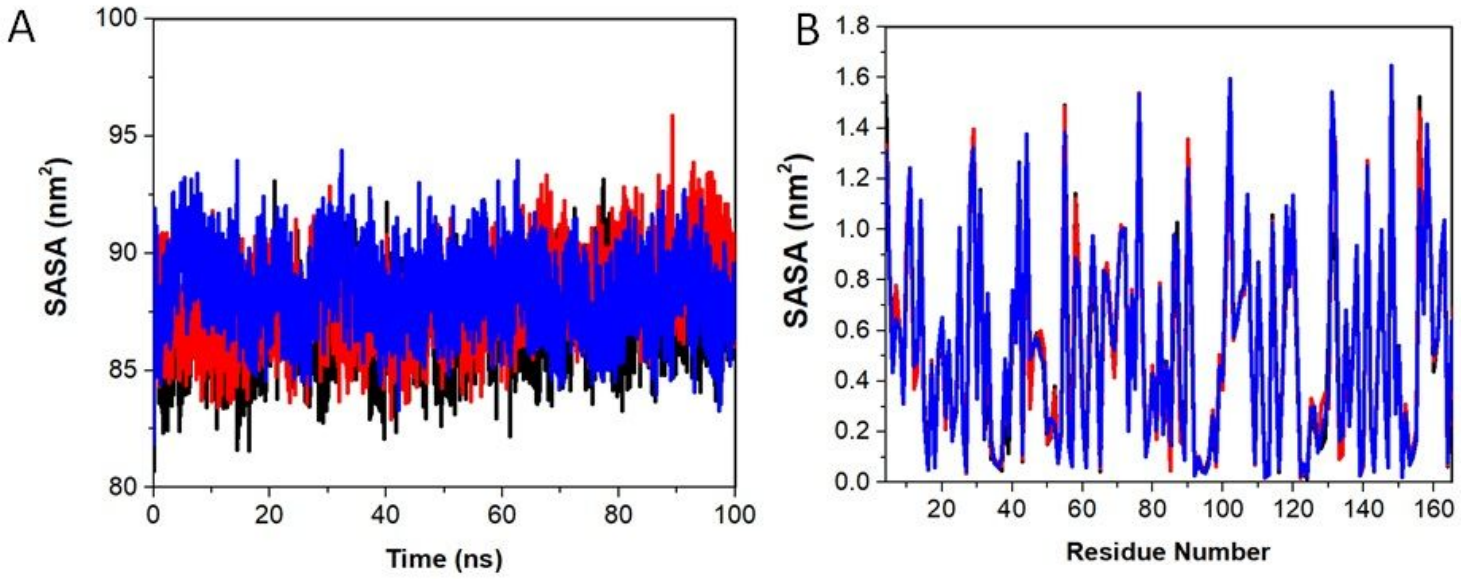


Figure 6

Solvent accessible surface area. (A) SASA value vs. time for all the complex for 100 ns. (B) The residue vs. SASA value for ADRP-ligand complex. The black, red and blue represents ADRP-ADPr, ADRP-Trifluperidol ADRP-Saquinavir respectively.

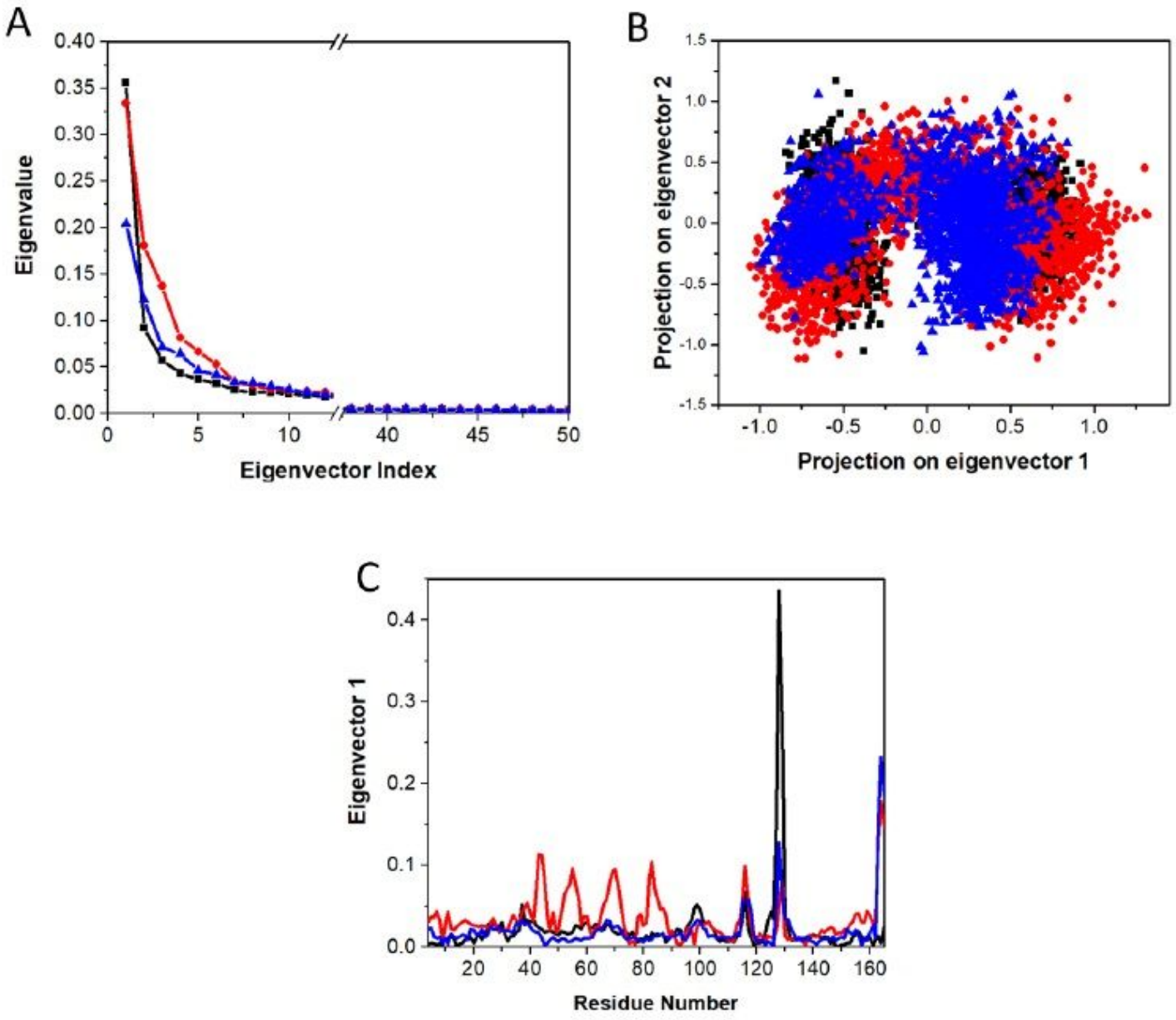


Figure 7

Principal Component analysis. (A) First 50 eigenvectors were plotted vs. eigenvalue for ADRP-ligand complexes. (B) Projection of the motion of the protein in phase space along PC1 vs. PC2 for ADRP-ligand complexes. (C) The first eigenvector vs. residue to see the residue wise fluctuation. The black, red and blue represent ADRP-ADPr, ADRP-Trifluperidol ADRP-Saquinavir respectively.