

**In Silico Identification of Potential Allosteric Inhibitors of the SARS-CoV-2 Helicase**

Adekunle B. Rowaiye*, Olukemi A. Onuh, Titilayo M. Asala, Amoge C. Ogu, Doofan Bur, Ezinne J. Nwankwo, Uchenna M. Orji, Zainab R. Ibrahim, Jamila Hamza, Adaku L. Ugorji

Department of Medical Biotechnology, National Biotechnology Development Agency, Lugbe, Abuja, Nigeria

ARTICLE INFO

Article history:

Received 08 October 2020

Revised 05 December 2020

Accepted 21 January 2021

Published online 03 February 2021

ABSTRACT

The Coronavirus Disease 2019 (COVID-19) pandemic ravages the globe causing unprecedented health and economic challenges. As the world prospects for a cure, scientists are looking critically at strategic protein targets within the SARS-CoV-2 that have therapeutic significance. The Helicase is one of such targets and it is an enzyme that affects all facets of the SARS-CoV-2 RNA metabolism. This study is aimed at identifying small molecules from natural products that have strong binding affinity with and exhibit inhibitory activity against an allosteric site (Pocket 26) on the SARS-CoV-2 Helicase. The molecular docking simulations of SARS-CoV-2 Helicase (QHD43415-12.pdb) against a library of small molecules obtained from edible African plants was executed using PyRx. Triphenylmethane, which had a docking score of -7.4 kcal/mol on SARS CoV-2 Helicase was chosen as a reference compound. Based on the molecular descriptors of the compounds as provided by PubChem, a virtual screening for oral bioavailability was performed. Further screening for molar refractivity, pharmacokinetic properties, and bioactivity were performed using SwissADME, pkCSM, and Molinspiration webserver respectively. Molecular dynamic simulation and analyses were performed using the Galaxy webserver which uses the GROMACS software. The lead compounds are Gibberellin A12, A20, and A51 obtained from Green peas and the Okra plant. Gibberellin A20 and A51 were predicted to perform better than the standard and Gibberellin A51 showed the greatest inhibitory activity against SARS-CoV-2 Helicase.

Copyright: © 2021 Rowaiye *et al.* This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Keywords: COVID-19, SARS-CoV-2, Helicase, Gibberellin, Molecular Dynamic Simulation.

Introduction

The Coronavirus disease 2019 (COVID-19) pandemic currently affects 213 countries and territories around the world. As of April 15, 2020, the World Health Organization (WHO) has reported about 2,000,000 global cases of COVID-19 with 128,886 deaths.¹ COVID-19 is an infectious disease caused by a 2019-novel coronavirus (SARS-CoV-2) which spreads through the buccal and nasal discharges from infected persons. The disease can be asymptomatic but also present a range of symptoms such as fever, dry cough, fatigue, body pains, sore throat, diarrhea, headache, and loss of smell or taste. In severe cases, patients suffer pneumonia, severe acute respiratory syndrome, multi-organ failure, and death.² The WHO is yet to approve any vaccine or antiviral drug for the prevention or treatment of coronavirus infections. However, CoV-associated pathologies have been treated with some existing broad-spectrum antiviral drugs. Most treatment strategies focus on the management of symptoms and supportive therapy.^{3,4} Coronaviruses (CoVs) have been known to infect a wide variety of mammals and birds causing respiratory and enteric diseases. They are classified into four different genera namely the alpha-, beta-, gamma-, and delta-CoVs. The CoV genome is susceptible to frequent mutations and recombination which can give rise to new strains of varying virulence.⁵ There are seven strains of human CoVs and their major

predilection site is the upper and lower respiratory tract.⁶ CoVs were known to cause milder disease but the previous epidemics of high-morbidity caused by the SARS-CoV in 2003 and MERS-CoV in 2012, highlighted their adaptive potential to the changing environmental conditions and as such they are now classified as “emerging viruses”.⁷ CoVs are enveloped, positive-sense, single-stranded RNA viruses with a genome size ranging between 26.2 and 31.7 kb.⁸ They are named after their crown-like appearance in electron micrographs, which is caused by the club-shaped peplomers that radiate outwards from the viral envelope.^{9,10} SARS-CoV-2 belongs to the β genus and contains Structural and Non-Structural Proteins (NSP). The structural proteins include the Spike, Envelope, Membrane, and Nucleocapsid proteins.¹¹ The spike surface glycoprotein promotes host attachment, and virus-cell membrane fusion during virus infection and it is a major target of neutralizing antibodies.^{12,13}

One of the NSP which is the Helicase enzyme is a motor protein that utilizes the energy derived from nucleotide hydrolysis to unwind double-stranded nucleic acids into two single-stranded nucleic acids along the 5' – 3' direction.¹⁴ Helicases are not only involved in the unwinding of nucleic acids during recombination, replication, and repair. Recent studies have also shown that helicases are involved in several other biological processes such as transcription, translation, RNA stability, mRNA splicing, mRNA export, packaging of nucleic acids into virions, and mitochondrial gene expression.^{15,16} Their validity as drug targets was confirmed when compounds that inhibit a helicase encoded by Herpes Simplex Virus was shown to block viral replication and disease progression in animal models.¹⁷

The SARS-CoV-2 Helicase has 596 amino acids and adopts a triangular pyramid shape comprising of 5 domains.¹⁸ They include two RecA-like domains at the C-terminal Helicase core, the beta-barrel domain, the N-terminal Zinc Binding Domain, and the interconnecting Stalk domain.¹⁸ The active site of SARS-CoV-2 Helicase has residues which are involved in NTP hydrolysis.¹⁹

*Corresponding author. E mail: adekunlerowaiye@gmail.com
Tel: +2348023287179

Citation: Rowaiye AB, Onuh OA, Asala TM, Ogu AC, Bur D, Nwankwo EJ, Orji UM, Ibrahim ZR, Hamza J, Ugorji AL. *In Silico* Identification of Potential Allosteric Inhibitors of the SARS-CoV-2 Helicase. Trop J Nat Prod Res. 2021; 5(1):165-177. doi.org/10.26538/tjnpr/v5i1.22

Official Journal of Natural Product Research Group, Faculty of Pharmacy, University of Benin, Benin City, Nigeria.

This study is aimed at identifying potential inhibitors of an allosteric site (Pocket 26) on the SARS-CoV-2 Helicase²⁰ and a known helicase inhibitor, Triphenylmethane which is a basic structural component of synthetic dyes is used as the reference compound.²¹ However, other allosteric sites abound, one of which is pocket 25 which the antiviral drug, Darunavir inhibits.²² Natural compounds,²³ synthetic compounds,²⁴ and re-purposed drugs²² have been empirically proven or computationally predicted to be small-molecule helicase inhibitors.²⁵

Materials and Methods

Preparation, analysis and validation of target protein structure

The 3D structure of SARS-CoV-2 helicase in Protein Data Bank (pdb) format (ID: QHD43415_12.pdb) was obtained from I-TASSER online server with an estimated Template Modelling (TM) score of 0.99. This webserver is used for prediction of protein structure and function.²⁶ The architecture of the target protein was revealed by the Volume, Area, Dihedral Angle Reporter (VADAR 1.8) webserver. The target was analysed using the Ramachandran plot obtained from the MolProbity web server.²⁷

Ligand preparation

One thousand and forty eight (1,048) compounds obtained from natural products such as spices, edible fruits, and vegetables were downloaded from *PubChem*.²⁸ Lipinski's Rule of Five and Veber's rules were used for the pre-screening of all the compounds in the library i.e. molecular weight ≤ 500 , hydrogen bond donor (HBD) ≤ 5 , hydrogen bond acceptor (HBA) ≤ 10 , $\log P \leq 5$, polar surface area (PSA) ≤ 140 , and rotatable bonds ≤ 10 .²⁹ Their 3D structures and that of the reference compound, Triphenylmethane (PubChem CID 10614) were downloaded from *PubChem* in sdf format.²⁸

Molecular docking and virtual screening

The reference compound, and 1,048 other Lipinski and Veber rule-compliant compounds were uploaded to the virtual screening software, *PyRx* (Python Prescription) 0.8 using the *Open Babel* plug-in tool.³⁰ The ligands were subjected to energy minimization and then transformed from structure-data file (sdf) to Protein Data Bank, Partial Charge, & Atom Type (pdqqt) format in preparation for molecular docking. All ligands and the reference compound were docked against the target protein, SARS-CoV-2 Helicase using *AutoDock Vina* plug-in tool in *PyRx*.³¹ The grid parameters for docking with the target protein were set at: Centre X = 79.7763, Y = 79.8336, Z = 79.8336 and Dimensions (Angstrom): X = 89.9405, Y = 65.6233, Z = 77.6795. For stable conformation, Universal Force Field (UFF) and the conjugate gradient descent was used as the energy minimization parameter and optimization algorithm respectively.

The docking results were exported in comma-separated values (csv) format to Microsoft Excel for filtering. Only ligands that had binding affinity scores less than that of the reference compound, Triphenylmethane (-7.4 kcal/mol) were selected. The prediction of pharmacokinetic properties, molar refractivity, and bioactivity of all ligands were performed using *pkCSM*, *SwissADME*, and *Molinspiration* respectively.³²⁻³⁴

Analysis of the binding site

Docked poses of all the front-runner compounds were superimposed with the target protein using the *Pymol* software.³⁵ The resultant structures were evaluated using the *Protein-Ligand Interaction Profiler (PLIP)* webserver.³⁶ All binding pockets of SARS-CoV-2 Helicase were analysed with *Fpocket* online server.³⁷ The three-dimensional depictions of the best docked complexes were analysed using hydrogen bonds, salt bridges, and other protein-ligand interactions.³⁵

Molecular dynamics simulations (MDS) and analyses

The Galaxy (versions 2019.1 and 2019.1.4) supercomputing server which uses the GROMACS software was used to perform the MDS of the Apo and Holo structures of SARS-CoV-2 Helicase.³⁸ The

LigParGen server was used for the ligand parameterization of the lead compounds with OPLS-AA/1.14*CM1A as force field parameter.³⁹ A 2-nanoseconds MDS was carried out for all the Apo and Holo structures with 1,000,000 steps after solvation, energy minimization, and equilibration (NVT and NPT). MDS analyses of trajectory parameters such as the Root Mean Square Deviation of atomic positions (RMSD), per residue Root Mean Square Fluctuation (RMSF) of protein backbone, Principal Component Analysis (PCA), and Dynamical Cross-Correlation Matrix (DCCM), were determined using the BIO3D tool on the Galaxy super-computing platform.⁴⁰

Results and Discussion

Structural analysis, validation and preparation of SARS-CoV-2 Helicase (QHD43415_12.pdb)

The Apo structure of SARS-CoV-2 Helicase (QHD43415_12.pdb) has 601 amino acids with the following constituent secondary structures: α helix 27%; beta sheets 31%; Coil 41%; and Turns 15%. (Figure 1). The Total Accessible Solvent Area (ASA) is 26083.1Å². The geometry of SARS-CoV-2 Helicase (QHD43415_12.pdb) reveals 6.88% poor rotamers, 82.79% favored rotamers, 4.67% Ramachandran outliers, 85.48% Ramachandran favored, 3.34% Carbon Beta deviations $>0.25\text{\AA}$, 0.00% bad bonds, and 0.81% bad angles (Figure 2). The Peptide omegas of SARS-CoV-2 Helicase (QHD43415_12.pdb) include 0.00% Cis Prolines and 2.83% Twisted Peptides. The low-resolution criteria include 6.70% CaBLAM outliers and 0.34% CA Geometry outliers.

Chemoinformatic profile of ligands

The application of high-throughput computer-assisted methods to predict the relationship between the structure, chemical properties, and the biological activity of a compound is indeed a valuable tool in the field of drug design and discovery.⁴¹ These drug-like properties of compounds would impart largely on their bioavailability and increase cellular uptake of biomolecules within the body. The molecular descriptors of such compounds are well described by the Lipinski (RO5), Veber, and Ghose rules. Put together, these rules state that hydrogen bond acceptors should be ≤ 10 , hydrogen bond donors should be ≤ 5 ; Log P should be ≤ 5 , molecular weight should be ≤ 500 g/mol; the polar surface area should be $\leq 140\text{\AA}^2$, molar refractivity should be between 40-130 cm³, and the number of rotatable bonds should be < 10 .⁴²⁻⁴⁴

Our results prove that Gibberellin (GA)12, GA20, and GA51 had no violation of the Lipinski (RO5), Ghose, and Veber rules suggesting that they possess good drug permeability (Figure 3, Table 1). Specifically, the standard has a Log P value that violates the Lipinski rule. It is more lipophilic than all the lead compounds suggesting that it has the greatest absorbability across lipid membranes. The standard and the lead compounds can penetrate the Blood-Brain Barrier because their TPSA values are less than 90 angstroms squared.⁴⁵ Measured by saturation, the molecular complexity of organic molecules is an important property in computational drug discovery. The lead compounds have their fraction of carbons in the sp³ hybridization more than 0.25 while that of the standard is lower. This suggests the standard has the lowest saturation and hence the least molecular stability.⁴⁶

In biological assays, certain compounds (PAINS: pan assay interference) yield false positive response because they have problematic structural moieties. These compounds are considered promiscuous as they are frequent hitters. From Table 1, the standard and all the lead compounds have no PAIN alerts.⁴⁷ The bioactivity scores of the standard and the lead compounds showed moderate to good activity against ion channel, GCPR, nuclear receptor, kinase, proteases, and other enzyme targets. The greatest enzyme inhibiting activity is seen in GA51.⁴⁸

Pharmacokinetic properties of ligands

Traditional drug discovery process has bottlenecks. However, the application of *in silico* ADMET properties prediction to evaluate

potential leads at early stages of drug development is important in order to eliminate candidates which would have been chemically synthesized and biologically tested.⁴⁹ Multi-parametric optimization strategies such as absorption, distribution, metabolism, excretion/elimination, and toxicity are applied in making and screening compounds in drug discovery.⁴⁹

The absorption parameters that ensure drugs get to their molecular targets include human intestinal absorption (poor: <30%), caco2 permeability (high: > 0.9), water solubility (insoluble: -4.0 Log mol/L), and skin permeability (low: LogKp >-2.5). Triphenylmethane is insoluble in water and therefore is a poor drug candidate (Table 2). All predicted absorption values for the lead compounds are within pharmacological range.^{32,50}

The pharmacokinetic indicators for distribution such as Fraction unbound, Volume of distribution steady state (Low: Log VDss <-0.15; High: Log VDss > 0.45), BBB permeability (permeable: Log BBB > 0.3; poor <: Log BBB <-1), and CNS permeability (permeable Log PS > -2; poor Log PS < -3) for the standard (i.e., Triphenylmethane) and all lead compounds are within acceptable range (Table 2).⁵²

The standard and all lead compounds are not P-glycoprotein I & II inhibitors suggesting that these ATP-dependent cell membrane proteins would continue to pump foreign substances out of cells unhindered (Table 2). Triphenylmethane is a P-glycoprotein substrate which suggests it would be pumped out of the cell if it is not administered with a P-glycoprotein inhibitor.³²

The inhibition of the major isoforms of cytochrome P450 enzyme makes for the toxic accumulation of their substrates. The predicted metabolic behavior of all the lead compounds shows no inhibition of CYP3A4, CYP1A2, CYP2C9, CYP2C19, and CYP2D6 enzymes. However, the standard shows inhibition of CYP1A2 and CYP2C19 enzymes (Table 2). The standard and lead compounds are substrates of CYP3A4 which means that their doses would be affected either by induction or the inhibition of CYP3A4.⁵¹

From Table 2, the predicted excretion values for Total Clearance for the standard and lead compounds are within pharmacological range.³²

All of the lead compounds and the standard are not substrates of Renal Organic Cation Transporter 2 (OCT2) which implies that they will not be eliminated by the protein from the blood into the proximal tubular cell.⁵²

The dose to be administered in the phase 1 of clinical trials is determined by the predicted maximum recommended tolerated dose. From Table 2, the standard has a high value (more than 0.477 log mg/kg/day) while the lead compounds have low maximum recommended tolerated dose (less than 0.477 log mg/kg/day).³²

Two important parameters in drug discovery are the *Tetrahymena pyriformis* and Minnow toxicities which are the dose required to inhibit 50% of the growth of *T. pyriformis* (IGC50), a protozoan bacteria and the minnow fish respectively. For *T. pyriformis*, when the pIGC50 value is greater than -0.5 log Ug/L, a compound is considered toxic. Therefore, the standard and all the leads show antibacterial properties but might not be toxic to human cells.³² Similarly, log LC50 is the log of a compound to cause death of 50% of flathead Minnows. A value less than 0.3 log mM signifies high acute toxicity. From the results, the standard is toxic to Minnows while the lead compounds are not.³²

The cardiotoxic and genotoxic properties of compounds is revealed by the hERG inhibition and AMES predictions respectively. While all the compounds (standard and leads) showed no AMES toxicity, only the standard showed hERG II inhibition properties. This makes it a potentially dangerous drug candidate. The standard also showed dermatotoxic properties.³²

Molecular docking analyses of ligands against SARS-CoV-2 Helicase

In modern drug design and discovery, binding affinity determination is very crucial in order to find a high affinity ligand that would bind to the target protein to inhibit its disease-associated function, catalytic activity or interaction with other molecules.⁵³ This procedure typically begins with screening an initial library of compounds to computationally identify binders of the target protein before continuing with the experimental screening.⁵⁴

From Figure 4 and Table 3, all lead compounds have stronger binding affinity than the standard, Triphenylmethane and hence showed greater potency as drug candidates. GA A51 has the strongest binding affinity of -8.6 Kcal/mol.

Binding site analyses

Hydrogen bonds have ubiquitous influence in nature and play an important role in protein folding, protein-ligand interactions, and catalysis.⁵⁵⁻⁵⁷ By displacing protein-bound water molecules into the solvent, H-bonds also enhances ligand binding affinity.⁵⁸ The length and orientation of the hydrogen bond are two key factors that determine the specificity and direction of ligand binding.⁵⁹

Figure 5 and Table 4 reveal that while the standard has no hydrogen bond, GA12, GA20, and GA51 all have hydrogen bonds with residues that fall within Pocket 26. GA12 forms 3 hydrogen bonds within one residue (LEU235) while GA20 and GA51 form two hydrogen bonds within two residues (PHE24 & PHE133) signifying greater stability.

Table 1: Chemoinformatic properties of standard and lead compounds

| | Triphenylmethane | GA12 | GA20 | GA51 |
|---|---------------------------------|--|--|--|
| Molecular Formula | C ₁₉ H ₁₆ | C ₂₀ H ₂₈ O ₄ | C ₁₉ H ₂₄ O ₅ | C ₁₉ H ₂₄ O ₅ |
| Molecular Weight (g/mol) | 244.30 | 332.40 | 332.40 | 332.40 |
| Log P | 5.30 | 3.90 | 1.20 | 1.70 |
| Hydrogen Bond Acceptors | 0 | 4 | 5 | 5 |
| Hydrogen Bond Donors | 0 | 2 | 2 | 2 |
| # heavy atoms | 19 | 24 | 24 | 24 |
| # rotatable bonds | 3 | 2 | 1 | 1 |
| TPSA (A ^a) | 0 | 74.60 | 83.80 | 83.80 |
| Molar Refractivity | 80.38 | 0.80 | 0.79 | 0.79 |
| Saturation (fraction csp ³) | 0.05 | 90.52 | 86.18 | 86.14 |
| PAIN Alert | 0 | 0 | 0 | 0 |
| GCPR ligand | -0.21 | 0.32 | 0.22 | 0.17 |
| Ion channel modulator | -0.17 | 0.14 | 0.23 | 0.21 |
| Kinase Inhibitor | -0.57 | -0.44 | -0.21 | -0.31 |
| Nuclear receptor ligand | -0.15 | 0.80 | 0.49 | 0.67 |
| Protease inhibitor | -0.40 | 0.12 | 0.09 | 0.16 |
| Enzyme inhibitor | -0.03 | 0.36 | 0.30 | 0.38 |

Table 2: Pharmacokinetic properties of front-runner compounds

| | Triphenylmethane | GA12 | GA20 | GA51 |
|--|------------------|-------|-------|-------|
| Water solubility (log mol/L) | -6.88 | -2.89 | -2.64 | -2.73 |
| Caco2 permeability (log Papp in 10 ⁻⁶ cm/s) | 1.54 | 1.05 | 1.19 | 1.14 |
| Human Intestinal absorption (% Absorbed) | 98.55 | 100 | 98.91 | 100 |
| Skin Permeability (log Kp) | -2.71 | -2.74 | -2.74 | -2.74 |
| P-glycoprotein substrate (Yes/No) | Yes | No | No | No |
| P-glycoprotein I inhibitor (Yes/No) | No | No | No | No |
| P-glycoprotein II inhibitor (Yes/No) | No | No | No | No |
| VDss (human) (log L/kg) | 0.26 | -1.32 | -0.83 | -0.97 |
| Fraction unbound (human) (Fu) | 0.16 | 0.22 | 0.42 | 0.29 |
| BBB permeability (log BB) | 0.85 | 0.07 | -0.21 | -0.09 |
| CNS permeability (log PS) | -1.11 | -2.08 | -0.29 | -2.41 |
| CYP2D6 substrate (Yes/No) | No | No | No | No |
| CYP3A4 substrate (Yes/No) | Yes | Yes | Yes | Yes |
| CYP1A2 inhibitor (Yes/No) | Yes | No | No | No |
| CYP2C19 inhibitor (Yes/No) | Yes | No | No | No |
| CYP2C9 inhibitor (Yes/No) | No | No | No | No |
| CYP2D6 inhibitor (Yes/No) | No | No | No | No |
| CYP3A4 inhibitor (Yes/No) | No | No | No | No |
| Total Clearance (log ml/min/kg) | 0.20 | 0.46 | 0.42 | 0.42 |
| Renal OCT2 substrate (Yes/No) | No | No | No | No |
| AMES toxicity (Yes/No) | No | No | No | No |
| Max. Tolerated dose (human) (log mg/kg/day) | 0.67 | 0.42 | 0.37 | -0.14 |
| hERG I inhibitor (Yes/No) | No | No | No | No |
| hERG II inhibitor (Yes/No) | Yes | No | No | No |
| Oral Rat Acute Toxicity (LD50) (mol/kg) | 2.03 | 2.46 | 2.05 | 2.10 |
| Oral Rat Chronic Toxicity (log mg/kg_bw/day) | 0.82 | 2.07 | 2.14 | 2.36 |
| Hepatotoxicity (Yes/No) | Yes | Yes | No | Yes |
| Skin Sensitization (Yes/No) | Yes | No | No | No |
| <i>T. Pyriformis</i> toxicity (log ug/L) | 0.43 | 0.29 | 0.29 | 0.29 |
| Minnow toxicity (log mM) | 0.06 | 0.31 | 1.96 | 1.16 |

In terms of angle formed by hydrogen bonds, GA12 forms two strong (greater than 130°) and one weak (less than 130°) hydrogen bonds with the target protein. GA20 and GA51 form one weak and one strong hydrogen bond each.⁶⁰ In terms of the donor to acceptor distance, GA12 forms two moderate (2.5-3.2 Å) and one weak (3.2-4.0 Å) hydrogen bonds with the target protein. GA20 forms two moderate hydrogen bonds while GA51 forms one moderate and one weak hydrogen bond.⁶⁰

The strength and stability of the protein-ligand complexes is enhanced by the presence of hydrophobic interactions and salt bridges.⁶¹ From Table 5, the standard forms 4 hydrophobic interactions and one strong salt bond.⁶²

GA12, GA20 and GA51 forms 2 hydrophobic interactions and no salt bridge. These suggest that for the other protein-ligand interactions, the standard have a slightly more atom-efficient binding than the lead compounds.^{63,64}

Put together, the lead compounds show a stronger interaction with SARS-CoV-2 Helicase as revealed by the hydrogen bonds and binding affinity scores.

Molecular dynamics simulation analyses

As seen in Figures 1 and 6, comparing the crystal structure with the simulated apo and holo structures suggests that there is an unfolding of the alpha helix at residues 64, 65, and 66 during the molecular dynamic simulation.⁶⁵

Root mean square deviation of atomic positions (RMSD)

RMSD is the measure of similarity between a reference and a target structure. It measures the variations in the distances between atoms in two superimposed protein structures. In protein structure prediction, RMSD is used for analysing protein stability and conformational changes. It describes the similarity of conformers.^{66,67}

Figure 7 and Table 6 suggest that there was a gradual increase in the RMSD of the simulated Apo protein relative to the crystal structure as the production time increased. It peaked at Frame 9 (2.42 Å) and thereafter stabilized. Throughout the trajectory, the total RMSD for the Apo structure is 41.05 Å while the average is 1.95 Å. While the RMSD values of the Apo protein appear to have stabilized with the

simulation time, data suggests that with more simulation time, the RMSD values for all the holo structures would increase. Of all the holo structures, the Helicase-GA51 complex has the greatest deviation from the reference structure. This is seen in the highest values of the total and average RMSD. Helicase-GA12 and Helicase-Triphenylmethane complexes have lower total and average RMSD values than the Helicase-GA20. The Helicase-GA51 complex showed the greatest deviation to the right (most peaks within 2.00 - 2.49 Å). The Helicase-GA20 complex also showed greater deviation to the right than the Helicase-GA12 complex which in turn was greater than the Helicase-Triphenylmethane complex (Figure 8 and Table 6). Put together, ligand binding with GA51 and GA20 appear to cause greater deviation from the reference structure than Triphenylmethane. The least deviation was induced by GA12. The results predict GA51 to cause the greatest structural deviation of all the holo structures.

Root mean square fluctuations (RMSF)

Proteins undergo structural fluctuations as a result of movements of the alpha carbon of their residues around their equilibrium conformations.⁶⁸ From Figure 9 and Table 6, the total and average global RMSF is greater in the Helicase-GA51 complex than all the other holo structures. Also, the total and average global RMSF values of the Helicase-Triphenylmethane complex is greater than those of the Helicase-GA12 and Helicase-GA20 complexes.

The total and average regional RMSF values at Pocket 26 remained highest for the Helicase-GA51 complex. The values for Helicase-GA20 complex were higher than those of the Helicase-Triphenylmethane complex. The Helicase-GA12 complex has the lowest values. Similarly, the range of RMSF followed the same order with Helicase-GA51 and Helicase-GA20 complexes having the highest values.

Put together, ligand binding with GA51 induced the greatest instability as seen in the global, regional RMSF values, and the range of RMSF. Ligand binding with GA20 also induced more instability of the SARS-CoV-2 Helicase than the standard and this would negatively affect the physiological function of the enzyme.

Principal components analysis

During the MDS of a protein, new conformations are being generated during the trajectory. Principal component analysis (PCA) is used to determine the statistical significance and relationships of these conformations.⁶⁹

Of all the holo structures, the Total global motions (average of PC1, PC2, and PC3) was highest in Helicase-GA51 complex. Also, Helicase-GA20 and Helicase-GA12 complexes showed greater global motion than the Helicase-Triphenylmethane complex (Figure 10 and Table 6).

In a similar manner, the Total regional motions (average of PC1, PC2, & PC3) was highest in Helicase-GA51 complex making it the most unstable at Pocket 26 of all the holo structures.

Based on the greatest motions, the best global conformations are PC3 of the Apo protein, PC1 of Helicase-Triphenylmethane complex, PC2 of Helicase-GA12 complex, PC3 of the Helicase-GA20 complex, and PC3 of the Helicase-GA51 complex. Similarly, the best conformations that produced the greatest motions at Pocket 26 are PC3, PC3, PC1, PC1, and PC3 of the Apo protein, Helicase-Triphenylmethane, Helicase-GA12, Helicase-GA20, and Helicase-GA51 complexes respectively (Table 6).

The cosine contents of the principal components reveal the convergence of the MD simulation. Monitoring convergence is essential for sampling quality as results should be accurate and reproducible. Table 6 shows good quality cosine content data except a slight non-convergence at the PC3 of the Helicase-Triphenylmethane complex.⁷⁰

Put together, while all lead compounds induced greater motions of the SARS-CoV-2 Helicase than the standard, GA51 induced the greatest

motions. This implies that GA51 would induce the greatest structural distortion of the viral protein at global and regional (Pocket 26) levels.

Table 3: Molecular docking scores of ligands against SARS-CoV-2 Helicase

| Ligand | Binding (Kcal/mol) affinity |
|------------------|-----------------------------|
| Triphenylmethane | -7.4 |
| Gibberellin A12 | -8.0 |
| Gibberellin A20 | -8.4 |
| Gibberellin A51 | -8.6 |

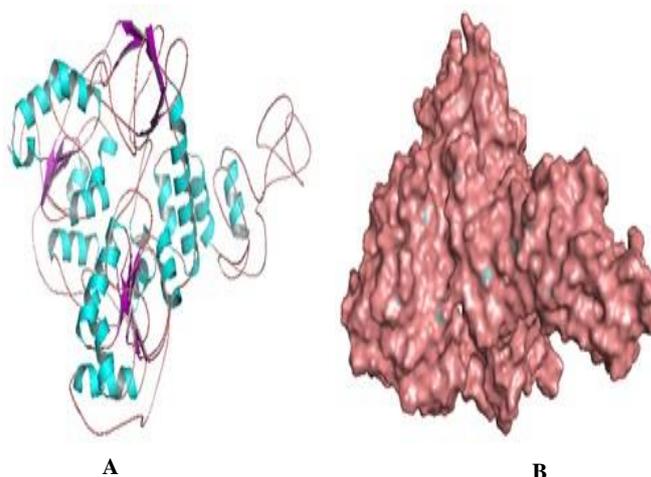


Figure 1: (A) Cartoon model of the crystal structure of SARS-CoV-2 Helicase (QHD43415_12.pdb). Beta sheets (magentas), Alpha helix (cyan) and Loops (pink). (B) Surface representations

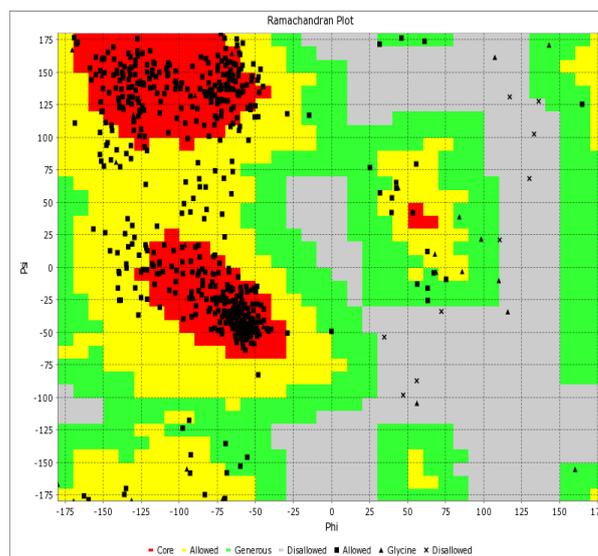


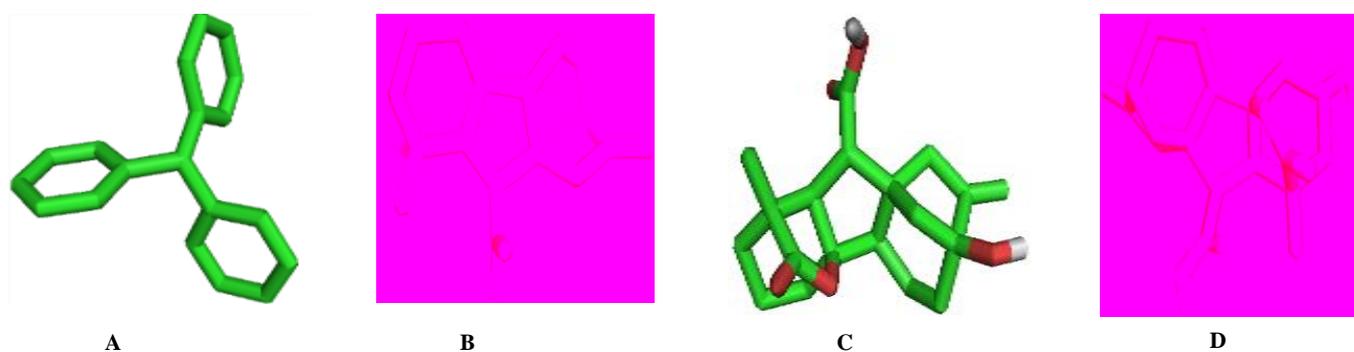
Figure 2: Ramachandran plot for SARS-CoV-2 Helicase (QHD43415_12.pdb).

Table 4: Hydrogen bond analysis

| Complex | Number of bonds | Residues | Distance (H-A) | Distance (D-A) | Bond angle |
|-----------------------|-----------------|----------|----------------|----------------|------------|
| Hel- Triphenylmethane | 0 | | | | |
| Hel-GA12 | 3 | LEU235 | 3.22 | 4.01 | 134.88 |
| | | LEU235 | 2.24 | 3.09 | 145.5 |
| | | LEU235 | 2.76 | 3.16 | 105.78 |
| Hel-GA20 | 2 | PHE24 | 2.22 | 3.18 | 159.23 |
| | | PHE133 | 2.64 | 3.03 | 102.66 |
| Hel-GA51 | 2 | PHE24 | 2.09 | 3.06 | 159.85 |
| | | PHE133 | 2.83 | 3.31 | 109.57 |

Table 5: Other Protein-ligand interactions

| Complex | Hydrophobic Int. Residue | Distance | Salt bridge Residue | Distance | p-Stacking Residue | Distance |
|----------------------|--------------------------|----------|---------------------|----------|--------------------|----------|
| | | | | | | |
| Hel-Triphenylmethane | VAL6 | 3.45 | | | PHE24 | 4.86 |
| | PRO23 | 3.90 | | | | |
| | PHE24 | 3.98 | | | | |
| | PHE133 | 3.63 | | | | |
| Hel-GA12 | ARG129 | 3.60 | | | | |
| | PRO234 | 3.60 | | | | |
| Hel-GA20 | PRO23 | 3.50 | | | | |
| | ARG129 | 3.86 | | | | |
| | PRO234 | 3.94 | | | | |
| Hel-GA51 | PRO23 | 3.67 | | | | |
| | ARG129 | 3.99 | | | | |

**Figure 3:** The 3D chemical structures (stick model) of standard and lead compound. (A) Triphenylmethane, (B) GA12, (C) GA20 and (D) GA51*Dynamic cross correlation map (DCCM)*

DCCM is a widely accepted tool used to analyze trajectories of molecular dynamics simulation. It is used to calculate the time-correlation atomic motions of a system as it reveals the heat map of the cross correlation of residual fluctuations.^{71,72}

Figure 11 and Table 6, shows heat maps depicting a complex pattern of correlated, non-correlated, and anti-correlated motions in Apo and Holo structures.

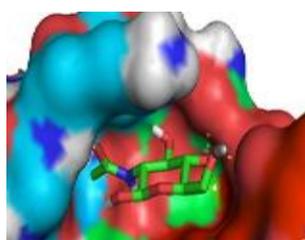
Comparative results reveal that atomic motions in the Helicase- GA51 complex showed the most intense anti-correlated motions in the holo structures. The Helicase-Triphenylmethane complex has predominantly non-correlated motions in residues 1-300. On the contrary, the Helicase-GA51 complex has predominantly correlated motions between residues 1-100 and non-correlated motions between residues 100-300. While the heat map of the Helicase-GA12 complex closely resembles that of the Helicase-Triphenylmethane, that of the

Helicase-GA20 complex is less intense than that of the Helicase-GA51 complex with respect to anti-correlation motions. Put together, the standard and all the lead compounds induced anti-correlation motions on SARS-CoV-2 Helicase. GA51 induced the greatest anti-correlation motions.

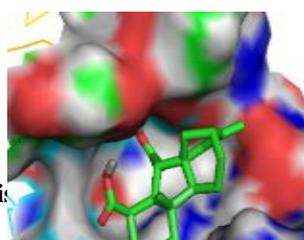
Found abundantly in *Pisum sativum* (green peas) and *Abelmoschus esculentus* (okra), Gibberellins (GAs) are plant hormones that regulate its various physiological processes.^{73,74} Notably in cucumbers, the Mosaic virus infection is associated with a reduced concentration of endogenous gibberellins which culminates in stunted roots and reduced leaf and stem growth.⁷⁵ In this study, GA29 and GA20 are predicted to be good drug candidates and also have a modulatory effect on the killer cell immunoglobulin-like receptor 2DS2 of Natural Killer Cells.⁷⁶ Similarly, GAs have been implicated as modulators of plant innate immunity.⁷⁷

Table 6: Summary of data from Molecular Dynamics Simulations of Apo and Holo Structures

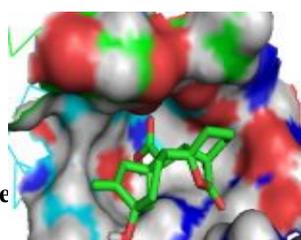
| MDS Parameters | Helicase | Helicase-Triphenylmethane | Helicase-GA12 | Helicase-GA20 | Helicase-GA51 |
|--|----------|---------------------------|---------------|---------------|---------------|
| RMSD | | | | | |
| Total RMSD | 41.05 | 42.84 | 36.06 | 44.06 | 51.06 |
| Average RMSD | 1.95 | 2.04 | 1.72 | 2.10 | 2.46 |
| Lowest RMSD | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Highest RMSD | 2.42 | 2.87 | 2.25 | 2.61 | 3.32 |
| Time Frame of Highest RMSD | 1 | 1 | 1 | 1 | 1 |
| Time Frame of Lowest RMSD | 9 | 21 | 21 | 21 | 15 |
| RMSD Peak Distribution | | | | | |
| 0.00 - 0.49A | 1 | 1 | 1 | 1 | 1 |
| 0.50 – 0.99A | 0 | 0 | 0 | 0 | 0 |
| 1.00 – 1.49A | 1 | 0 | 2 | 1 | 0 |
| 1.50 – 1.99A | 5 | 6 | 13 | 1 | 3 |
| 2.00 – 2.49A | 14 | 12 | 5 | 16 | 6 |
| 2.50 – 2.99A | 0 | 2 | 0 | 2 | 6 |
| 3.00 – 3.49A | 0 | 0 | 0 | 0 | 5 |
| RMSF | | | | | |
| Total Global RMSF | 598.93 | 639.76 | 610.30 | 622.78 | 724.06 |
| Average Global RMSF | 1.00 | 1.06 | 1.02 | 1.04 | 1.20 |
| Total Regional (Pocket 26) RMSF | 8.33 | 8.20 | 7.21 | 8.25 | 8.57 |
| Average Regional (Pocket 26) RMSF | 0.76 | 0.75 | 0.66 | 0.75 | 0.78 |
| Least Fluctuation | 0.45 | 0.44 | 0.45 | 0.40 | 0.39 |
| Highest Fluctuation | 3.48 | 3.52 | 2.83 | 4.06 | 4.37 |
| Range of RMSF | 3.03 | 3.08 | 2.38 | 3.66 | 3.98 |
| PCA | | | | | |
| Total global motions (mean of PC1, PC2 & PC3) | 20.57 | 19.20 | 20.30 | 20.7 | 20.9 |
| Average global motions (mean of PC1, PC2 & PC3) | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 |
| Total Regional (Pocket 26) Motion (mean of PC1, PC2 & PC3) | 0.27 | 0.23 | 0.24 | 0.24 | 0.25 |
| Average Regional (Pocket 26) Motion (mean of PC1, PC2 & PC3) | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 |
| Best global Conformation | PC3 | PC1 | PC2 | PC3 | PC3 |
| Best regional Conformation (Pocket 26) | PC3 | PC3 | PC1 | PC1 | PC3 |
| PC1 Eigenvalue | 19.90% | 30.70% | 31.70% | 30.70% | 50.20% |
| PC2 Eigenvalue | 15.70% | 0.16% | 14.50% | 20.50% | 10.50% |
| PC3 Eigenvalue | 13.30% | 0.13% | 9.20% | 9.30% | 8.50% |
| Total | 48.90% | 59.80% | 55.40% | 60.50% | 69.20% |
| PC1 cosine content | 0.87 | 0.86 | 0.82 | 0.87 | 0.87 |
| PC2 cosine content | 0.23 | 0.32 | 0.71 | 0.76 | 0.87 |
| PC3 cosine content | 0.13 | 0.34 | 0.70 | 0.53 | 0.68 |



s. Thi



e Cre



on 4.0

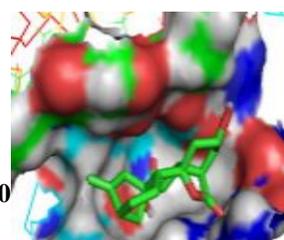


Figure 4: Binding site of SARS-CoV-2 Helicase interacting with standard and lead compounds. (A) Helicase-Triphenylmethane complex, (B) Helicase-GA12 complex, (C) Helicase-GA20 complex and (D) Helicase-GA51 complex

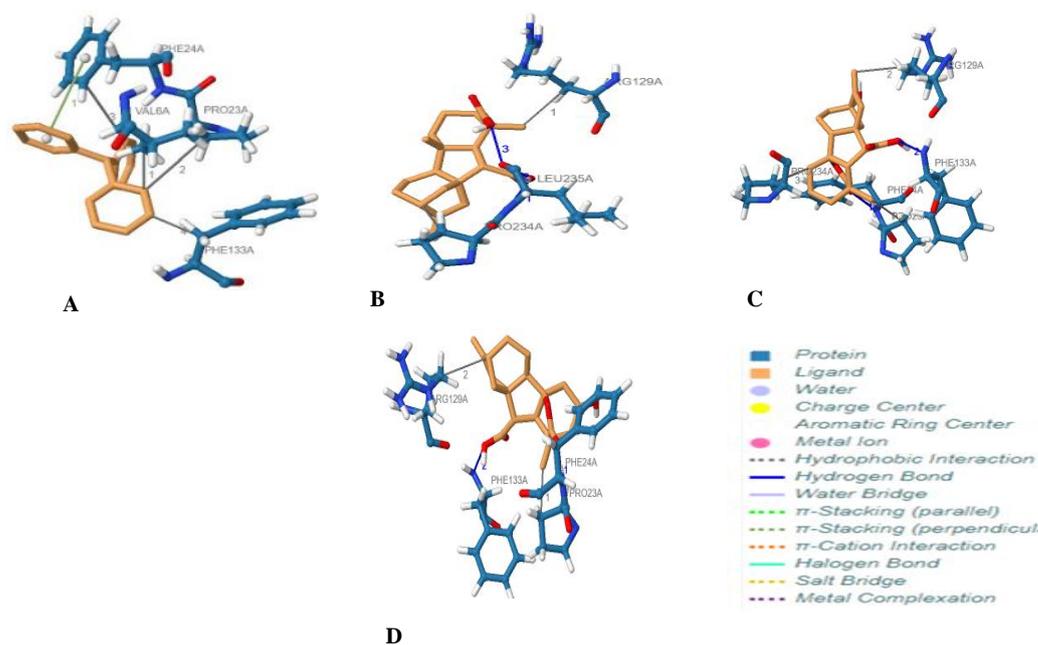


Figure 5: Protein-Ligand interactions of SARS-CoV-2 Helicase with standard and lead compound. (A) Helicase-Triphenylmethane complex, (B) Helicase-GA12 complex, (C) Helicase-GA20 complex and (D) Helicase-GA51 complex

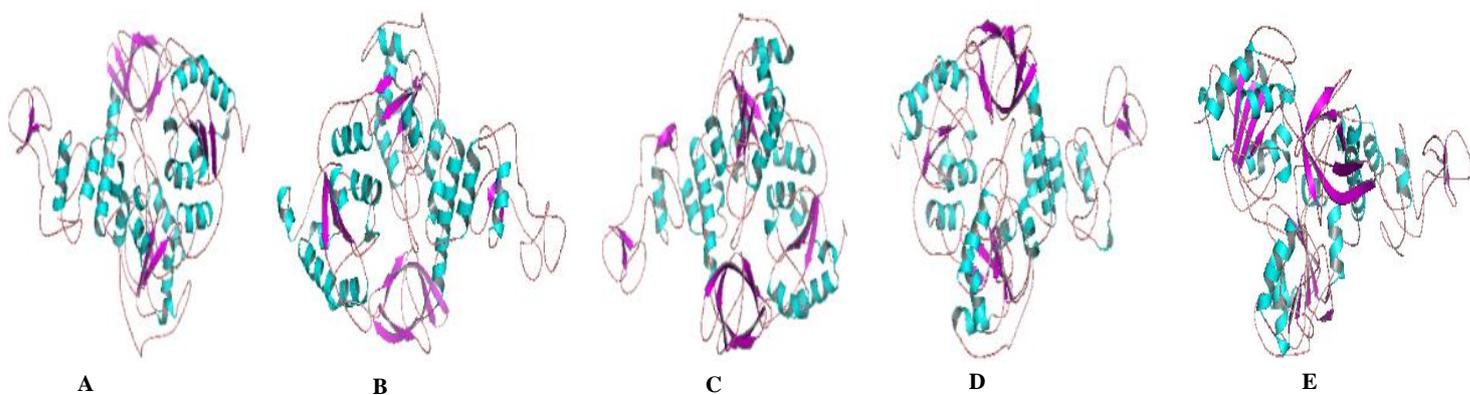


Figure 6: Cartoon model of the crystal structure of SARS-CoV-2 Helicase Apo and Holo structures (without water and ions) after molecular dynamics simulation. Beta sheets (magenta), Alpha helix (cyan) and Loops (pink). (A) Helicase, (B) Helicase-Triphenylmethane complex, (C) Helicase-GA12 complex, (D) Helicase-GA20 complex and (E) Helicase-GA51 complex

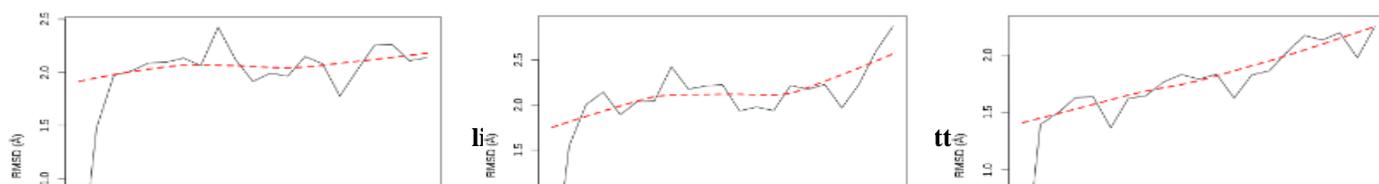


Figure 7: RMSD for Apo and Holo structures. (A) Helicase, (B) Helicase-Triphenylmethane complex, (C) Helicase-GA12 complex, (D) Helicase-GA20 complex and (E) Helicase-GA51 complex.

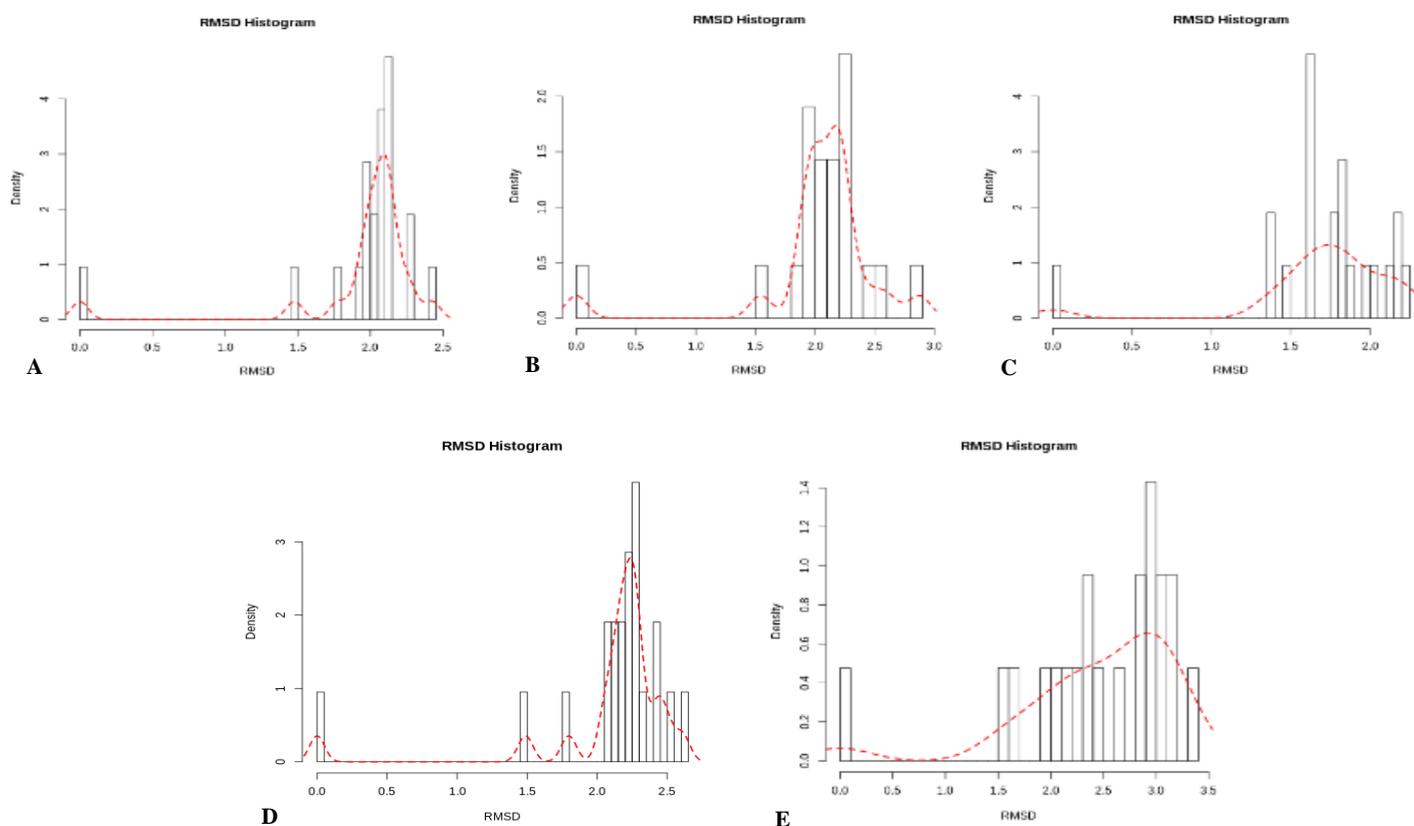


Figure 8: RMSD histogram for Apo and Holo structures. (A) Helicase, (B) Helicase-Triphenylmethane complex, (C) Helicase-GA12 complex, (D) Helicase-GA20 complex and (E) Helicase-GA51 complex

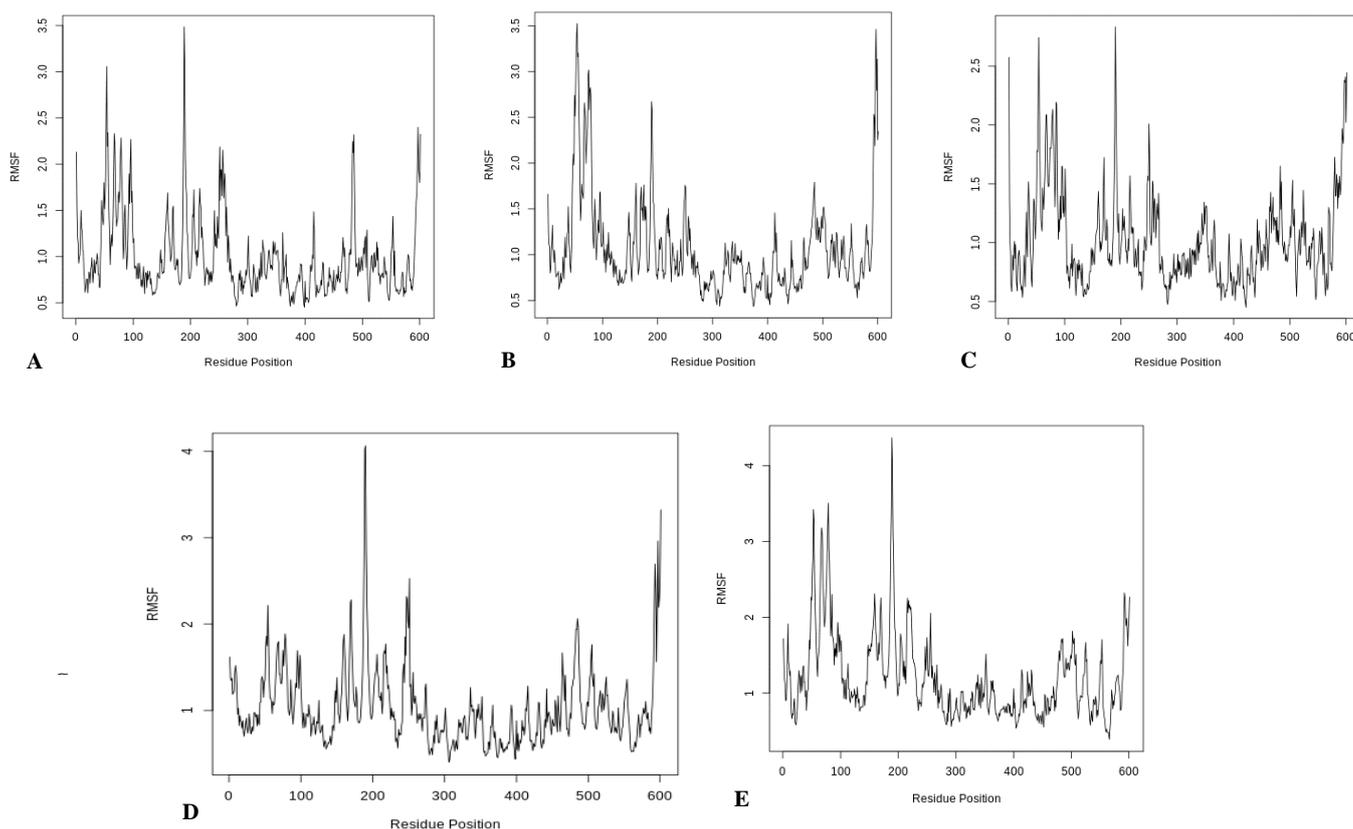


Figure 9: Per-residue RMSF for Apo and Holo structures. (A) Helicase, (B) Helicase-Triphenylmethane complex, (C) Helicase-GA12 complex, (D) Helicase-GA20 complex and (E) Helicase-GA51 complex

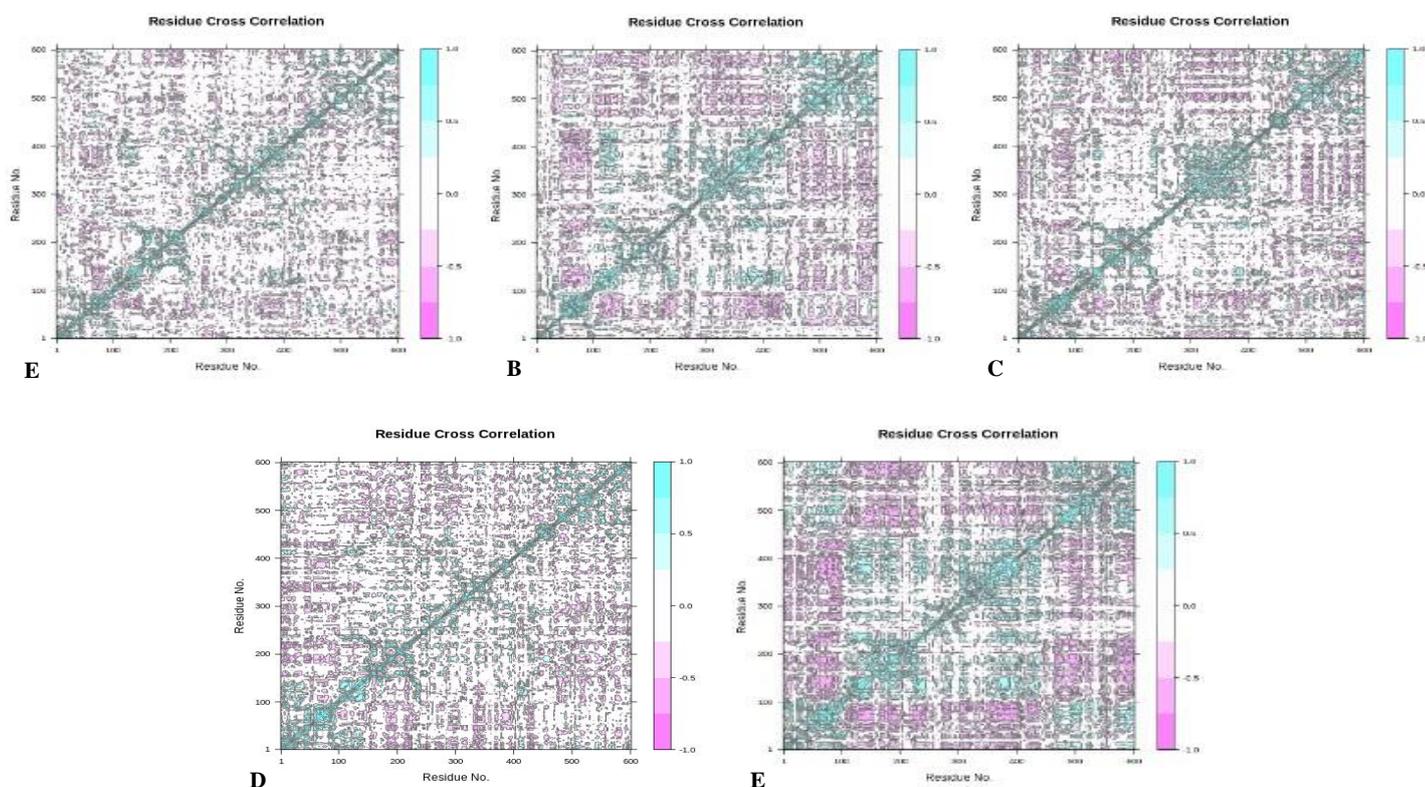


Figure 10: Dynamic cross correlation map Apo and Holo structures of SARS-CoV-2 Helicase.

Purple represents anti-correlated, dark cyan represents fully correlated while white and cyan represents moderately and uncorrelated respectively. 1.0 = correlated; 0 is non-correlated; and -1 is anti-correlated. (A) Helicase, (B) Helicase-Triphenylmethane complex, (C) Helicase-GA12 complex, (D) Helicase-GA20 complex and (E) Helicase-GA51 complex

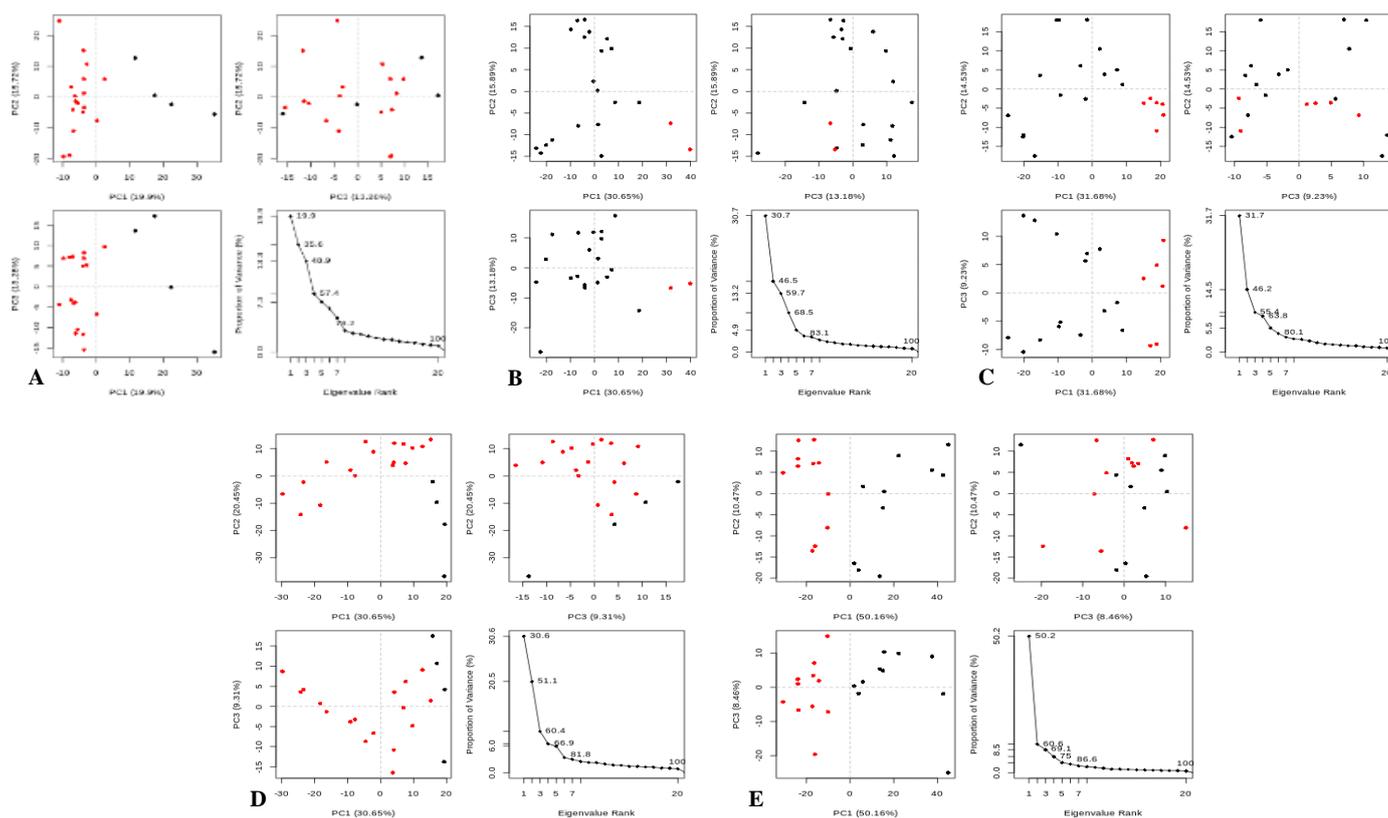


Figure 11: Principle component analysis cluster plot of Apo and Holo structures

The projection of trajectory onto 1st three eigenvectors for: (A) Helicase, (B) Helicase-Triphenylmethane complex, (C) Helicase-GA12 complex, (D) Helicase-GA20 complex and (E) Helicase-GA51 complex

Conclusion

Though Triphenylmethane has been proven to be a viral Helicase inhibitor, it is a poor drug candidate. GA51 is predicted to have the highest enzyme inhibiting activity and also the strongest binding affinity (-8.6 Kcal/mol) for SARS-CoV-2 Helicase. The other lead compounds showed a greater binding affinity than the standard which had no hydrogen bond with the target protein. Studying the time-resolved motions of Apo and Holo macromolecules, GA51 and GA20 are predicted to have better pharmacodynamics than the standard. The standard proved to be better than GA12.

Specifically, as a viral protein inhibitor, GA51 and GA20 show greater structural distortion to the SARS-CoV-2 Helicase as seen in the RMSD values, distribution of RMSD peaks, RMSF, and PCA (global and local motions) than the standard. Overall, GA51 has been predicted to show the greatest SARS-CoV-2 Helicase inhibitory activity as further confirmed by the DCCM map. It is recommended that further investigation be carried out to evaluate the inhibitory activity of GA51 and GA20 against SARS-CoV-2 Helicase using *in vivo* and *in vitro* experiments.

Conflict of interest

The authors declare no conflict of interest.

Authors' Declaration

The authors hereby declare that the work presented in this article is original and that any liability for claims relating to the content of this article will be borne by them.

Acknowledgement

We wish to express our gratitude to the Management of Medical Biotechnology Department, National Biotechnology Development

Agency, Abuja, Nigeria for providing a suitable environment for this research.

References

1. World Health Organization. Coronavirus disease (COVID-19) situation report – 86. [Online]. 2020 [cited 2020 April 15]. Available from: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200415-sitrep-86-covid-19.pdf?sfvrsn=c615ea20_6_2020.
2. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu, Y, Zhang L, Fan G, Xu J, Gu X, Cheng Z, Yu T, Xia J, Wei Y, Wu W, Xie X, Yin W, Li H, Liu M, Xiao Y, Gao H, Guo L, Xie J, Wang G, Jiang R, Gao Z, Jin Q, Wang J, Cao B. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020; 395(10223):497-506.
3. Hazafa A, ur-Rahman K, Haq I, Jahan N, Mumtaz M, Farman M, Naem H, Abbas F, Naem M, Sadiqa S, Banob S. The broad-spectrum antiviral recommendations for drug discovery against COVID-19. *Drug Metab Rev.* 2020; 52(3):408-424.
4. Bajwah S, Wilcock A, Towers R, Costantini M, Bausewein C, Simon ST, Bendstrup E, Prentice W, Johnson MJ, Currow DC, Kreuter M, Wells AU, Birring SS, Edmonds P, Higginson IJ. Managing the supportive care needs of those affected by COVID-19. *Eur Respir J.* 2020; 55(4):2000815.
5. Liu S, Shen J, Fang S, Li K, Liu J, Yang L, Hu CD, Wan J. Genetic Spectrum and Distinct Evolution Patterns of SARS-CoV-2. *Front Microbiol.* 2020; 11:593548.
6. Murphy K. SARS CoV-2 Detection from Upper and Lower Respiratory Tract Specimens Diagnostic and Infection Control Implications. *Chest.* 2020; 158(5):1804-1805.
7. Chen Y, Liu Q, Guo D. Emerging coronaviruses: Genome structure, replication, and pathogenesis. *J Med Virol.* 2020; 92(4):418-423.

8. Prajapat M, Sarma P, Shekhar N, Avti P, Sinha S, Kaur H, Kumar S, Bhattacharyya A, Kumar H, Bansal S, Medhi B. Drug targets for corona virus: A systematic review. *Indian J Pharmacol.* 2020; 52(1):56-65.
9. Singhal TA. Review of Coronavirus Disease-2019 (COVID-19). *Indian J Pediatr* 2020; 87:281-286.
10. Burrell CJ, Howard CR, Murphy FA. Coronaviruses. *Fenner and White's Medical Virology.* 2017; 437-446.
11. Satarker S and Nampoohiri M. Structural Proteins in Severe Acute Respiratory Syndrome Coronavirus-2. *Arch Med Res.* 2020; 51(6):482491.
12. Bosch BJ, van der Zee R, de Haan, CAM, Rottier PJM. The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. *J Virol.* 2003; 77(16):8801-8811.
13. Fung TS and Liu DX. Human Coronavirus: Host-Pathogen Interaction. *Annu Rev Microbiol* 2019; 73:529-557.
14. Brosh Jr RM and Matson SW. History of DNA Helicases. *Genes* 2020; 11(3):255.
15. Ren X, Yu Y, Li H, Huang J, Zhou A, Liu S, Hu P, Li B, Qi W, Liao M. Avian Influenza A Virus Polymerase Recruits Cellular RNA Helicase eIF4A3 to Promote Viral mRNA Splicing and Spliced mRNA Nuclear Export. *Front Microbiol.* 2019; 10:1625.
16. Bourgeois C, Mortreux F, Auboeuf D. The multiple functions of RNA helicases as drivers and regulators of gene expression. *Nat Rev Mol Cell Biol* 2016; 17:426-438.
17. Crute JJ, Grygon CA, Hargrave KD, Simoneau B, Faucher AM, Bolger G, Kibler P, Liuzzi M, Cordingley MG. Herpes simplex virus helicase-primase inhibitors are active in animal models of human disease. *Nat Med.* 2002; 8(4):386-391.
18. Mirza MU and Froeyen M. Structural elucidation of SARS-CoV-2 vital proteins: Computational methods reveal potential drug candidates against main protease, Nsp12 polymerase and Nsp13 helicase. *J Pharm Anal.* 2020; 10(4):320-328.
19. Jia Z, Liming Y, Ren Z, Wu L, Wang J, Guo J, Zheng L, Ming Z, Zhang L, Lou Z, Rao Z. Delicate structural coordination of the Severe Acute Respiratory Syndrome coronavirus Nsp13 upon ATP hydrolysis. *Nucleic Acids Res* 2019; 47(12):6538-6550.
20. Samatanga B, Andreou AZ, Klostermeier D. Allosteric regulation of helicase core activities of the DEAD-box helicase YxiN by RNA binding to its RNA recognition motif. *Nucleic Acids Res.* 2017; 45(4):1994-2006.
21. Habtemariam S, Nabavi SF, Banach M, Berindan-Neagoe I, Sarkar K, Sil PC, Nabavi SM. Should We Try SARS-CoV-2 Helicase Inhibitors for COVID-19 Therapy? *Arch Med Res.* 2020; 51(7):733-735.
22. Frediansyah A, Tiwari R, Sharun K, Dhama K, Harapan H. Antivirals for COVID-19: A critical review. *Clinical Epidemiology and Global Health* 2020.
23. Sayed AM, Khatta AR, AboulMag AM, Hassan HM, Rate ME, Zaidg H, Abdelmohsen UR. Nature as a treasure trove of potential anti-SARS-CoV drug leads: a structural/mechanistic rationale. *RSC Adv.* 2020; 10:19790-19802.
24. Lee J, Cho J, Ahn H, Jung W, Jeong Y. A Novel Chemical Compound for Inhibition of SARS Coronavirus Helicase. *J Microbiol Biotechnol.* 2017; 27(11):2070-2073.
25. Wu C, Liu Y, Yang Y, Zhang P, Zhong W, Wang Y, Wang Q, Xu Y, Li M, Li X, Zheng M, Chen L, Li H. Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. *Acta Pharm Sin B.* 2020; 10(5):766-788.
26. Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc* 2010; 5(4):725-738.
27. Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB 3rd, Snoeyink J, Richardson JS, Richardson DC. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* 2007; 35(Web Server Issue): W375-83.
28. Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, Li Q, Shoemaker BA, Thiessen PA, Yu B, Zaslavsky L, Zhang J, Bolton EE. PubChem 2019 update: improved access to chemical data. *Nucleic Acids Re.* 2019; 47(D1):D1102-D1109.
29. Atatreh N, Al Rawashdah S, Al Neyadi SS, Abuhamdah SM, Ghattas MA. Discovery of new butyrylcholinesterase inhibitors via structure-based virtual screening. *J Enzyme Inhib Med Chem.* 2019; 34(1):1373-1379.
30. Dallakyan S and Olson AJ. Small-molecule library screening by docking with PyRx. *Methods Mol Biol.* 2015; 1263:243-250.
31. Trott O and Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem.* 2010; 31(2):455-461.
32. Pires DEV, Tom L, Blundell DB, Ascher. pkCSM: Predicting Small-Molecule Pharmacokinetic and Toxicity Properties Using Graph-Based Signatures. *J Med Chem* 2015; 58(9):4066-4072.
33. Daina A, Olivier M, Vincent Z. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci Rep.* 2017; 7:42717.
34. Molinspiration. Calculation of molecular properties and bioactivity score. [Online]. 2020 [Cited 2020 May]. Available from: <https://www.molinspiration.com/cgi-bin/properties>.
35. DeLano WL. Pymol: An open-source molecular graphics tool. *CCP4 NewsI Protein Crystallogr.* 2002; 40:82-92.
36. Salentin S, Schreiber S, Haupt VJ, Adasme M.F, Schroeder M. PLIP: fully automated protein-ligand interaction profiler. *Nucleic Acids Res.* 2015; 43(W1):W443-447.
37. Le Guilloux V, Schmidtke P, Tuffery P. Fpocket: An open-source platform for ligand pocket detection. *BMC Bioinformatics.* 2009; 10:168.
38. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, Lindahl E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 2015; 1-2:19-25.
39. Dodda LS, Cabeza de Vaca I, Tirado-Rives J, Jorgensen WL. LigParGen web server: an automatic OPLS-AA parameter generator for organic ligands. *Nucleic Acids Res.* 2017; 45(W1):W331-W336.
40. Afgan E, Baker D, van den Beek M, Blankenberg D, Bouvier D, Čech M, Chilton J, Clements D, Coraor N, Eberhard C, Gruning B, Guerler A, Hillman-Jackson J, Von Kuster G, Rasche E, Soranzo N, Turaga N, Taylor J, Nekrutenko A, Goecks J. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res.* 2016; 44(W1):W3-W10.
41. Macalino SJY, Billones JB, Organo VG, Carrillo MCO. In Silico Strategies in Tuberculosis Drug Discovery. *Molecules* 2020; 25(3):665.
42. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev.* 2001; 46(1-3):3-26.
43. Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, Kopple KD. Molecular properties that influence the oral bioavailability of drug candidates. *J Med Chem.* 2002; 45(12):2615-2623.
44. Ghose AK, Viswanadhan VN, Wendoloski JJ. A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. *A*

- qualitative and quantitative characterization of known drug databases. *J Comb Chem* 1999; 1(1):55-68.
45. Cambridge MedChem Consulting. Brain Penetration (a work in progress) [Online] 2019. [cited 2020 Dec 02]. Available from: https://www.cambridgemedchemconsulting.com/resources/ADME/brian_penetration.html
 46. Saddala SM, Lennikov A, Huang H. Discovery of Small-Molecule Activators for Glucose-6-Phosphate Dehydrogenase (G6PD) Using Machine Learning Approaches. *Int J Mol Sci.* 2020; 21:1523.
 47. Capuzzi SJ, Muratov EN, Tropsha A. Phantom PAINS: Problems with the Utility of Alerts for Pan-Assay Interference Compounds. *J Chem Inf Model.* 2017; 57(3):417-427.
 48. Khan T, Dixit S, Ahmad R, Raza S, Azad I, Joshi S, Khan AR. Molecular docking, PASS analysis, bioactivity score prediction, synthesis, characterization and biological activity evaluation of a functionalized 2-butanone thiosemicarbazone ligand and its complexes. *J Chem Biol.* 2017; 10(3):91-104.
 49. Kesharwani RK, Vishwakarma VK, Keservani RK, Singh P, Katiyar N, Tripathi S. Role of ADMET Tools in Current Scenario: Application and Limitations. In: Singh DB (Ed.). *Computer-Aided Drug Design.* Singapore: Springer. 2020. 71-87 p.
 50. Savjani KT, Gajjar AJ, Savjani KJ. Drug solubility: importance and enhancement techniques. *ISRN Pharm* 2012; 2012:195727.
 51. Zhou SF. Drugs behave as substrates, inhibitors and inducers of human cytochrome P450 3A4. *Curr Drug Metab.* 2008; 4(9):310-322.
 52. Hacker K, Maas R, Kornhuber J, Fromm M.F, Zolk O. Substrate-Dependent Inhibition of the Human Organic Cation Transporter OCT2: A Comparison of Metformin with Experimental Substrates. *PLoS One* 2015; 10(9):e0136451.
 53. Freire E. A thermodynamic approach to the affinity optimization of drug candidates. *Chem Biol Drug Des.* 2009; 74(5):468-472.
 54. Jensen JH. Predicting accurate absolute binding energies in aqueous solution: thermodynamic considerations for electronic structure methods. *Phys Chem Chem Phys.* 2015; 17(19):12441-12451.
 55. Natarajan A, Schwans JP, Herschlag D. Using unnatural amino acids to probe the energetics of oxyanion hole hydrogen bonds in the ketosteroid isomerase active site. *J Am Chem Soc.* 2014; 136(21):7643-7654.
 56. Taylor MS and Jacobsen EN. Asymmetric Catalysis by Chiral Hydrogen-Bond Donors. *Angew Chemie Int.* 2016; 45(10):1520-1543.
 57. Salentin S, Haupt VJ, Daminelli S, Schroeder M. Polypharmacology rescored: protein-ligand interaction profiles for remote binding site similarity assessment. *Prog Biophys Mol Biol.* 2014; 116(2-3):174-186.
 58. Sarkar A and Kellogg GE. Hydrophobicity-shake flasks, protein folding and drug discovery. *Curr Top Med Chem.* 2010; 10(1):67-83.
 59. Jeffrey GA. *An Introduction to Hydrogen Bonding.* Oxford: Oxford University Press; 1997.
 60. Desiraju G and Steiner T. *The Weak Hydrogen Bond. In Structural Chemistry and Biology.* Oxford: Oxford University Press; 2001.
 61. Kumar S and Nussinov R. Close-range electrostatic interactions in proteins. *ChemBiochem.* 2002; 3(7):604-617.
 62. Kenny PW. The nature of ligand efficiency. *J Cheminform.* 2019; 11(1):8.
 63. Kufareva I and Abagyan R. Methods of protein structure comparison. *Methods Mol Biol.* 2012; 857:231-257.
 64. Stockner T, Vogel HJ, Peter Tieleman D. A salt-bridge motif involved in ligand binding and large-scale domain motions of the maltose-binding protein. *Biophys J.* 2005; 89(5):3362-3371.
 65. Krzysztof K, Gouri SJ, Ron E. Kinetics of helix unfolding: molecular dynamics simulations with milestoning. *J Phys Chem A.* 2009; 113(26):7461-7473.
 66. Rajesh D. Re: What is the importance of the RMSD value in molecular docking? [Online]. 2015 [Cited 2020 May 28]. Available from: https://www.researchgate.net/post/What_is_the_impirtance_of_the_RMSD_value_in_molecular_docking/556301ab614325252e8b45b2/citation/download.
 67. Gürsoy O and Smieško M. Searching for bioactive conformations of drug-like ligands with current force fields: how good are we? *J Cheminform.* 2017; 9(1):29.
 68. Fuglebakk E, Echave J, Reuter N. Measuring and comparing structural fluctuation patterns in large protein datasets. *Bioinformatics* 2012; 28(19):2431-2440.
 69. David CC and Jacobs DJ. Principal component analysis: a method for determining the essential dynamics of proteins. *Methods Mol Biol.* 2014; 1084:193-226.
 70. Sawle L, Ghosh K. Convergence of Molecular Dynamics Simulation of Protein Native States: Feasibility vs Self-Consistency Dilemma. *J Chem Theory Comput.* 2016; 12(2):861-869.
 71. Upadhyay SK. Dynamics of Gal80p in the Gal80p-Gal3p complex differ significantly from the dynamics in the Gal80p-Gal1p complex: implications for the higher specificity of Gal3p. *Mol Biosyst.* 2014; 10(12):3120-3129.
 72. Kasahara K, Fukuda I, Nakamura H. A Novel Approach of Dynamic Cross Correlation Analysis on Molecular Dynamics Simulations and Its Application to Ets1 Dimer-DNA Complex. *PLoS One* 2014; 9(11):e112419.
 73. Vidhyasekaran P. Gibberellin Signaling in Plant Innate Immunity. In *Plant Hormone Signalling. Systems in Plant Innate Immunity.* Netherlands: Springer; 2015. 383-401 p.
 74. Smith VA. Gibberellin A1 Biosynthesis in *Pisum sativum* L.: II. Biological and Biochemical Consequences of the le Mutation. *Plant Physiol.* 1992; 99(2):372-377.
 75. Bailiss KW. The relationship of gibberellin content to cucumber mosaic virus infection of cucumber. *Physiol Plant Pathol.* 1974; 4:73-79.
 76. Rowaiye AB, Olubiyi J, Bur D, Uzochukwu IC, Akpa A, Esimone CO. In silico screening and molecular dynamic simulation studies of potential small molecule immunomodulators of the KIR2DS2 receptor. *BioRxiv.* 2020; 2020(05-10)087148.
 77. De Bruyne L, Höfte M, De Vleeschauwer D. Connecting growth and defense: the emerging roles of brassinosteroids and gibberellins in plant innate immunity. *Mol Plant.* 2014; 7(6):943-959.