In silico identification of the potential natural inhibitors of SARS-CoV-2 Guanine-N7 methyltransferase.

Adekunle B. Rowaiye[1], Olukemi A. Onuh[1], Joy Awulika Oladimeji-Salami[1], Doofan Bur[1], Moses Njoku[2], Nma Helen Ifedilichukwu[1], Comfort Ojochenemi John[1], Olanike Binuyo[1], Faith Udo Pius[1]

[1]Department of Medical Biotechnology, National Biotechnology Development Agency, Lugbe, Abuja, Nigeria.

[2] Human Virology Division, Department of Microbiology and Biotechnology, Nigerian Institute of Pharmaceutical Research & Development, Abuja, Nigeria

*Email of corresponding author: adekunlerowaiye@gmail.com

Abstract

The outbreak of the COVID-19 pandemic caused by the SARS-CoV-2 has triggered intense scientific research into the possible therapeutic strategies that can combat the ravaging disease. One of such strategies is the inhibition of an important enzyme that affects an important physiological process of the virus. The enzyme, Guanine 7 Methyltransferase is responsible for the capping of the SARS-CoV-2 mRNA to conceal it from the host's cellular defense. The study aims at computationally identifying the potential natural inhibitors of the SARS-CoV-2 Guanine-N7 methyltransferase binding at the active site (Pocket 41). A library of small molecules was obtained from edible African plants and were molecularly docked against the SARS-CoV-2 Guanine-N7 methyltransferase (QHD43415_13. pdb) using the Pyrx software. Sinefungin, an approved antiviral drug which had a binding score of -7.6 kcal/ mol with the target was chosen as a standard. Using the molecular descriptors of the compounds, a virtual screening for oral availability was performed using the Pubchem and SWISSADME web tools. The online servers PKCSM and Molinspiration were used for further screening for pharmacokinetic properties and bioactivity respectively. The molecular dynamic simulation and analyses of the Apo and Holo proteins was performed using the GROMACS software on the Galaxy webserver. The lead compounds are Crinamidine, Marmesin and Sinensetin which are obtained from waterleaf, mango, and orange plants respectively. All the lead compounds performed better than the

standard. Crinamidine is predicted to show the greatest inhibitory activity. Further tests are required to further investigate the inhibitory activities of the lead compounds.

Introduction

Corona virus disease 2019 (COVID-19) is a novel infection which began in China resulting in a worldwide outbreak. The disease was declared a global health emergency and later recognized a pandemic by the World Health Organization in March 2020 [1]. As at the 25th of July, 2020, the global number of reported cases of the disease stood at 15,975,268 with 643,476 deaths and 9,766,873 recoveries [2]. COVID-19 is caused by Severe Acute Respiratory Syndrome Corona virus 2 (SARS-CoV-2) which causes mild to severe respiratory illness with symptoms such as fever, cough, and shortness of breath. The ailment becomes life-threatening in the presence of co-morbidities such as diabetes, hypertension, and cardiovascular diseases [3,4]. There is currently no WHO approved drug or vaccine for the cure or prevention of COVID-19. SARS-CoV-2 belongs to a large family of viruses consisting of multiple strains that are known to cause illnesses ranging from the common cold to more severe diseases such as the Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS) [4,5]. SARS-CoV-2 is a positive-sense, single-stranded RNA virus possessing the largest and most complex genome (about 30 Kb), packed inside a nucleocapsid protein and enveloped with several structural proteins [6]. The size of the viral particle is in the range of 8 0-90nm and there are bulbous surface projections that form crown-like patterns (corona) on the surface of the particles [7]. The potential therapeutic strategies for the treatment of COVID-19 include immunomodulation and viral inhibition. Several enzymes or structural proteins of SARS-CoV-2 are potential drug targets as they directly affect physiological processes such as RNA synthesis, replication, assembly and human cell receptor binding [8,9]. Guanine N-7-MethylTransferase (GNMT) is one of such targets and it the enzyme responsible for the capping of SARS-CoV-2 mRNA. For many life-sustaining processes such as replication, protein translation and metabolism, viruses require a host cell as they lack the proper cellular machinery. Viral propagation within host cell require the transcription of viral mRNA. To do this, the viral mRNA

assumes molecular anonymity to evade detection in the host cell cytoplasm. The viral mRNA undergoes structural modification by a 5' cap structure. By evading the host cell defense system, viral mRNA can be effectively translated into proteins. The addition of the guanine N-7-methylguanosine cap is necessary for the maturation, stability, nuclear export and efficient translation of viral mRNA. Eukaryotic mRNA is modified by the addition of the 5′ cap structure which is a 7-methylguanosine linked to the first transcribed nucleotide by a 5′-5′ triphosphate bridge [10]. The mRNA cap is formed on the first transcribed nucleotide of transcripts by three sequential enzymatic activities; triphosphatase, guanylyltransferase and methyltransferase [11,12]. The 5′ triphosphate of pre-mRNA is hydrolyzed to diphosphate by a 5 ′-triphosphatase, to which Guanosine monophosphate (GMP) is added by the RNA guanylyltransferase to create the cap intermediate, GpppN. Guanine-N-7-methyl transferase (GNMT) also known as mRNA cap guanine-N7 methyltransferase is the enzyme that catalyzes the chemical reaction and most importantly plays a necessary part of the RNA capping reaction. RNA guanine N7 methyltransferase creates the mature cap, m7GpppN, and a byproduct, AdoHcy (S-adenosyl homocysteine) through the methylation of the cap intermediate utilizing the methyl donor, AdoMet [13]. The GNMT in coronaviruses belongs to a large class of SAM (S-Adenosyl methionine)-dependent methyltransferases and is an exoribonuclease [14]. Additionally, they have been shown to be linked with a unique 3' to 5' exoribonuclease (ExoN) domain in non-structural protein 14 (nsp14). The diversity of the capping apparatus makes viral RNA capping an attractive target for drug design and development [14,15]. Accordingly, the inhibition of GNMT which may induce potent antiviral activity makes it an important drug target [16]. This implies that incompletely-capped mRNAs can be recognized by immune sensors which trigger innate immunity pathways that culminate in the expression of type I interferon and other cytokines that have antiviral activity in neighboring cells [17,18]. The active site of GNMT is found in Pocket 41 and it includes residues ARG 289, VAL 290, TRP 292, GLY 333, PRO 335, ASP 352, ALA 353, GLN 354, PRO 355, CYS 356, SER 357, TRP 385, ASN 386, CYS 387, ASN 388 and PHE 426 [19].

Materials and Methods

Preparation, analysis and validation of target protein structure: The 3D structure of SARS-CoV-2 GNMT in the Protein Data Dank (pdb) format (ID: QHD43415_13. pdb) was obtained

from the I-TASSER online server with an estimated Template Modelling (TM) score of 0.99 [20]. The web server, Volume, Area, Dihedral Angle Reporter (VADAR 1.8) was used to reveal the architecture of GNMT. The structure of the target was further analysed using the Ramanchandran plot obtained from the MolProbity web server [21]. **Ligand preparation:** A library of 1,048 compounds obtained from edible African plants such as fruits, spices, and vegetables were downloaded from PubChem database [22]. All the compounds had been pre-screened for Lipinski (hydrogen bond donor (HBD) ≤ 5, hydrogen bond acceptor (HBA) ≤ 10, molecular weight ≤ 500, and logP ≤ 5) and Veber (polar surface area (PSA) ≤ 140, and rotatable bonds ≤ 10) rules [23]. The 3D structures of all the compounds, and that of the standard, Sinefungin (PubChem CID 65482) were downloaded from PubChem in the structure-data file (sdf) format [22].

Molecular docking and virtual screening: In preparation for molecular docking, all the ligands were uploaded on the virtual screening software, PyRx (Python prescription) 0.8 version using the Open Babel plug-in tool [24] and converted from sdf to Protein Data Bank, Partial Charge, & Atom Type (pdbqt) format [25]. For stable conformation, the Universal Force Field (UFF) was used as the energy minimization parameter and conjugate gradient descent as the optimization algorithm. Using AutoDock Vina plug-in tool in Pyrx, all ligands and the standard were docked against the target protein, SARS-CoV-2 GNMT using the following grid parameters [26]. Centre X = 92.432, Y = 92.529, Z = 92.555 and Dimensions (Angstrom): X = 87.658, Y = 97.427, Z = 64.081 [24]. Using the Microsoft Excel software, the docked results were exported in comma-separated values (csv) format and screened using the docking score of the standard, Sinefungin (-7.6 kcal/ mol) as the cut off. The SWISSADME, pkCSM, and Molinspiration webservers were used to predict the molar refractivity, pharmacokinetic properties, and bioactivity of all the ligands respectively [27, 28 29, 30]. The SMILES for Sinefungin and the ligands were downloaded from PubChem. Binding site analyses: Using the Pymol software, the target protein was superimposed with the docked poses of all the front-runner compounds [31]. The Protein-Ligand Interaction Profiler (PLIP) webserver was used to evaluate the resultant protein-ligand complexes for hydrogen bonds, salt bridges and other protein-ligand interactions. The analyses carried out include the name and number of residues, exhaustiveness, bond distance and bond angle [32]. The binding pockets of the target protein was analysed with the Fpocket web server [19].

Molecular Dynamic Simulations (MDS) and Analyses: A 2-nanoseconds MDS of the Apo and Holo structures of SARS-CoV-2 GNMT was performed using the GROMACS software of the Galaxy (versions 2019.1 and 2019.1.4) supercomputing server [33]. For ligand parameterization, LigParGen server was used to generate GROMACS-compatible topology files for the small molecules. OPLS-AA/ 1.14*CM1A was the force field parameter used [34, 35]. After initial conversion to topology files, solvation, energy minimization and equilibration (NVT and NPT), a 1,000,000-step MDS was performed. The analyses of trajectories were done using the BIO 3D tool on the Galaxy super-computing platform [36]. These include the Principal Component Analysis (PCA), per residue Root Mean Square Fluctuation (RMSF) of the protein backbone, and Root Mean Square Deviation of atomic positions (RMSD) and Dynamical Cross-Correlation Matrix (DCCM). [37]. The radius of gyration and the B factor was also analysed using the MDWeb webserver [38].

Results and Discussion.

Structural analysis, validation and preparation of SARS-CoV-2 GNMT (QHD43415_13. pdb): The Apo structure of SARS-CoV-2 GNMT (QHD43415_13. pdb) has 527 amino acids with the following constituent secondary structures: α helix 21%; beta sheets 30%; Coil 48%; and Turns 16% (Figure 1). The Total Accessible Solvent Area (ASA) is 260780 (Å) ^{2}. The geometry of SARS-CoV-2 GNMT (QHD43415_13. pdb) reveals 8.01% poor rotamers, 83.98% favored rotamers, 4.00% Ramachandran outliers, 82.29% Ramachandran favored, 3.22% Carbon Beta deviations >0.25Å, 0.00% bad bonds and 1.04% bad angles (Figure 2). The Peptide omegas of SARS-CoV-2 GNMT (QHD43415_13. pdb) include 0.00% Cis Prolines and 3.04% Twisted Peptides. The low-resolution criteria include 8.2% CaBLAM outliers and 0.96% CA Geometry outliers.
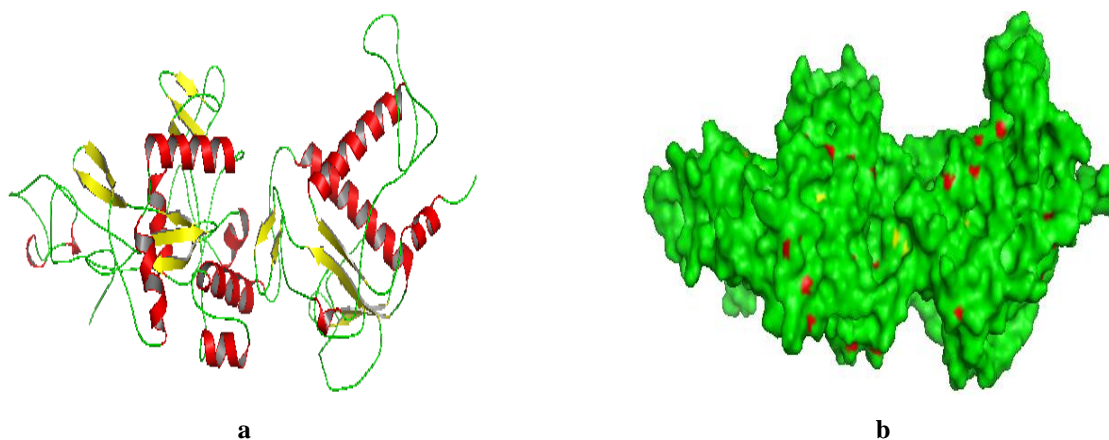
a   **b**

Figure 1: a**:** Cartoon model of the crystal structure of SARS-CoV-2 GNMT (QHD43415_13.pdb). Beta sheets (yellow), Alpha helix (red) and Loops (green) b: Surface representation.
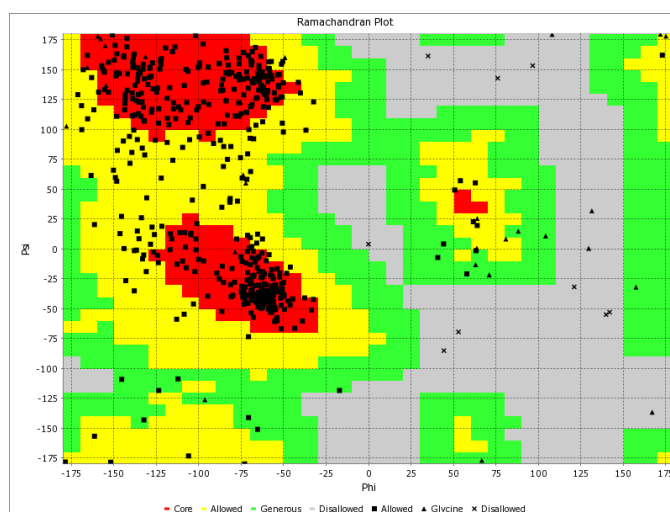

Figure 2: Ramachandran plot for SARS-CoV-2 GNMT (QHD43415_13.pdb).

Chemoinformatic profile of ligands (Figure 3, Table 1): A combination of Ghose, Lipinski and Veber rules define the molecular descriptors necessary for good oral bioavailability of drugs and their penetration through biological membranes. The molecular descriptors include a molecular weight ≤ 500 g/mol, log P ≤ 5, hydrogen bond donors ≤ 5, hydrogen bond acceptors ≤ 10, molar refractivity between 40 to 130, the number of rotatable bonds ≤ 10 and polar surface area (PSA) ≤ 140 [39, 40, 41, 42].

Results from Table 1, reveal that none of the lead compounds violated the Ghose, Lipinski and Veber rules. This suggests that they have good oral bioavailability and permeability. Therefore, we predict that these compounds are good drug candidates having met the criteria for drug-likeness assessment [43]. However, the Standard (Sinefungin) violates the Veber rule with a high TPSA value (208.65 Aª). This suggests that it would have a considerably lower intestinal absorption, blood-brain barrier permeation, and cellular potency than the lead compounds [44].

The molecular complexity of a compound measured by the ratio of sp3 hybridized carbons over the total carbon count of the molecule (Fraction Csp3). It is an important property in determining the success of drug development. A value of at least 0.25 indicates saturation [45]. From (Table 1), all lead compounds and standard are saturated suggesting molecular stability. Crinamidine has a higher saturation than the standard while Sinensetin has the lowest.

Due to problematic structural moieties, promiscuous bioactive compounds interact with multiple biological targets, and aggregate under assay conditions giving false positive results. While this might be good for polypharmacology, unintended interactions might likely lead to many undesired side effects [46]. From (Table 1), all lead compounds and standard are predicted to be non-promiscuous.

Beyond ligand binding to the appropriate target, it should elicit a pharmacological effect. Drug candidates are classified based on their bioactivity which includes GPCR ligands, ion channel modulators, kinase inhibitors, nuclear receptor ligands, protease inhibitors and other enzyme inhibitors [47]. In this study, the results showed that only the standard and Crinamidine had poor bioactivity scores as Nuclear Receptor Ligand and Kinase inhibitors respectively. All other scores for standard and lead compounds revealed moderate to good bioactivity against the targets. Furthermore, all lead compounds showed good activity as enzyme inhibitors. While the standard showed the highest enzyme inhibition, Marmesin showed the least activity (Table 1.). [29, 48].
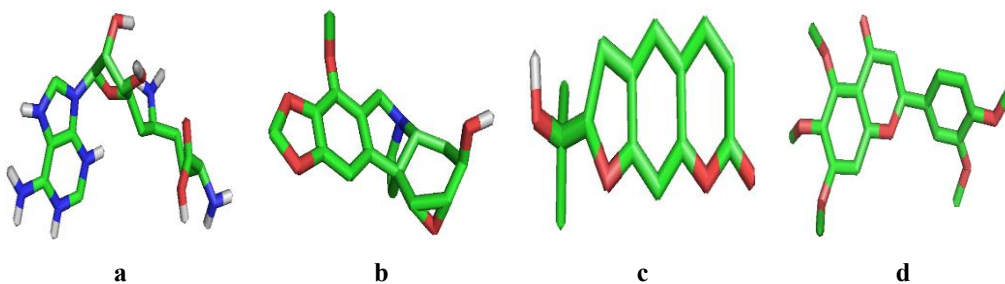
Figure 3: The 3D chemical structures (stick model) of standard and lead compound a: Sinefungin

b: Crinamidine c: Marmesin d: Sinensetin

Table 1: Chemo-informatic properties of standard and lead compounds

	Sinefungin (Standard)	Crinamidine	Marmesin	Sinensetin
Molecular Weight (g/mol)	381.39	346.37	354.31	320.29
XLogP3	-4.31	-0.43	-1.05	0.33
Hydrogen Bond Donors	6	3	4	5
Hydrogen bond acceptors	10	6	9	7
# heavy atoms	27	25	25	23
# rotatable bonds	7	2	4	2
TPSA (Aª)	208.65	111.90	138.82	119.61
Molar Refractivity	92.73	88.47	83.12	80.83
Saturation (fraction csp^3)	0.60	0.63	0.44	0.25
PAIN Alert	0	0	0	0
GCPR ligand	1.15	0.22	-0.16	0.37
Ion channel modulator	0.66	0.02	-0.26	0.07
Kinase Inhibitor	0.74	-0.62	-0.26	0.11
Nuclear Receptor Ligand	-1.03	0.67	-0.14	0.48
Protease Inhibitor	0.57	0.14	-0.16	0.23
Enzyme Inhibitor	1.14	0.36	0.29	0.39

Pharmacokinetic properties of ligands: Pharmacokinetic properties play an important role in drug discovery and development. The primary goal of drug to discovery or design projects are to identify potential drug candidates that have the greatest efficacy and least toxicity. To avoid failures in the drug development process, it is proper to identify good Absorption, Distribution, Metabolism, Excretion, and Toxicity (ADMET) properties of the front-runner compounds through *in silico* methods [28]. An excellent drug candidate should have good ADMET properties at therapeutic doses [28, 49].

The penetration of a target molecule by a drug candidate is a good marker of its therapeutic potential and its influenced by absorption parameters such as human intestinal absorption (poor: <30%), caco2 permeability (high:> 0.9), water solubility (insoluble: less than -4.0 Log mol/L), and skin permeability (low: LogKp > −2.5). From Table 2, data suggests that the standard and all lead compounds to have good human intestinal absorption property, and skin permeability. The ability to penetrate human epithelial colorectal adenocarcinoma cells is lowest in the standard, and highest in Crinamidine.

The pharmacological markers for distribution include CNS permeability (permeable Log PS > -2; poor Log PS < -3), BBB permeability (permeable: Log BBB > 0.3; poor <: Log BBB <-1), Volume of distribution steady state (Low: Log VDss <- 0.15; High: Log VDss > 0.45), and Fraction unbound. From Table 2, Sinensetin has a high VDSS, while the values for Marmesin and Crinamidine are below the pharmacological range. This can be corrected by dosage.

The standard and Marmesin have a poor ability to permeate into the brain tissue, while other lead compounds can permeate. The standard and all the lead compounds have a poor CNS permeability. The fraction unbound values for standard and all lead compounds are within acceptable range.

P-glycoprotein is a transmembrane efflux pump which pumps its substrates from inside to outside the cell [50]. All the lead compounds except Marmesin were shown to be P-glycoprotein substrates which implies that they should be co-administered with a P-glycoprotein inhibitor to prevent a potential reduction in absorption and oral bioavailability resulting in decreased retention time of the drug [51]. However, all lead compounds, and the standard showed no inhibition to P-glycoprotein I and II indicating less likelihood of its substrates inducing cellular toxicity, and drug interactions [52,53].

The predicted metabolic behavior of bioactive compounds is a determinant of their inclusion or elimination in the drug discovery process. The inhibition or non-inhibition of the isomers of the Cytochrome P450 enzyme determines whether the drug candidates would undergo biotransformation or accumulate in the cellular spaces with toxic tendencies. If drug candidates are Cytochrome P450 enzyme substrates they would be administered with inhibitors to facilitate their metabolism [54]. From Table 2, all lead compounds are neither inhibitors nor substrates of CYP1A2, CYP2C19, CYP2C9, CYP2D6 and CYP3A4 enzymes.

The predicted excretion values for Total Clearance for the standard, and the lead compounds are within pharmacological range [23]. Similarly, they all are predicted to be non-substrates of Renal Organic Cation Transporter 2 (OCT2). This implies that they will all be eliminated from the blood into the proximal tubular cell by the Renal OCT2 [24].

The toxicity profile for the standard and all lead compounds suggests that are non-mutagenic, non- cardiotoxic, non-hepatotoxic, and non-dermatotoxic as revealed in their AMES toxicity, hERG I & II toxicity, hepatotoxicity and skin Sensitization predictions respectively [28].

The dose administered at clinical trials is determined by the maximum recommended tolerated dose. Values less than 0.477 log mg/kg/day are considered low while values higher than 0.477 log mg/kg/day are considered as high. From Table 2 the predicted values suggest that Marmesin and Crinamidine are the most and least potent compounds respectively [55]. The predicted values for Oral Rat Acute Toxicity and Oral Rat Chronic Toxicity should be considered alongside factors such as concentration of drug, dose and the length of time it is administered [55].

Inhibition of 50% of the growth of *T.pyriformis, a* protozoan bacterium (IGC50) is a toxicity marker in drug discovery. When the pIGC50 value is greater than -0.5 log Ug/L, the drug candidate is considered as toxic. Results from Table 2, all lead compounds, and the standard are predicted to be toxic against *T.pyriformis* suggesting antibacterial effect properties (that might be unharmful to human cells) [55]. Similarly, in flathead Minnows, the log LC50 is the log of a compound to cause death of 50% of the population. High acute toxicity is indicated by values less than 0.3 log mM. The results from Table 2 shows that all lead compounds, and the standard are not toxic to Minnows [55].

Table 2: Pharmacokinetic properties of ligands

	Sinefungin (standard)	Crinamidine	Marmesin	Sinensetin
Water solubility (log mol/L)	-2.892	-2.487	-2.21	-3.085
Caco2 permeability (log Papp in 10-6 cm/s)	-0.933	0.54	0.377	-0.119
Human Intestinal absorption (% Absorbed)	32.936	51.799	48.119	60.725
Skin Permeability (log Kp)	-2.735	-2.735	-2.822	-2.735
P-glycoprotein substrate (Yes/No)	Yes	Yes	No	Yes
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)	0.012	-1.386	-0.611	1.635
Fraction unbound (human) (Fu)	0.383	0.488	0.397	0.263
BBB permeability (log BB)	-1.582	-0.665	-1.286	-0.927
CNS permeability (log PS)	-3.928	-3.101	-3.954	-3.265
CYP2D6 substrate (Yes/No)	No	No	No	No
CYP3A4 substrate (Yes/No)	No	No	No	No
CYP1A2 inhibitor (Yes/No)	No	No	No	No
CYP2C19 inhibitor (Yes/No)	No	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)	0.564	0.744	0.716	0.347
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)	0.44	0.777	0.393	0.368
hERG I inhibitor (Yes/No)	No	No	No	No
hERG II inhibitor (Yes/No)	No	No	No	No
Oral Rat Acute Toxicity (LD_{50})(mol/kg)	2.482	1.996	2.391	2.289
Oral Rat Chronic Toxicity (log mg/kg_bw/day)	3.081	2.278	3.756	2.929
Hepatotoxicity (Yes/No)	No	No	No	No
Skin Sensitization (Yes/No)	No	No	No	No
T.Pyriformis toxicity (log ug/L)	0.285	0.285	0.286	0.296
Minnow toxicity (log mM)	4.001	3.177	4.198	3.747

Molecular docking analyses of ligands against SARS-CoV-2 GNMT: In molecular docking, the binding affinity score is a measure of the ability of the small molecule to find the optimal conformation in the protein binding pocket. Hence, the ligand with the lower binding energy suggests the greatest binding affinity making it a possible drug candidate [56].

All lead compounds have shown greater potency as drug candidate because they have stronger binding affinity than the standard. Crinamidine has the strongest binding affinity of -8.5 Kcal/mol (Table 3)

Table 3. Molecular docking scores of ligands against SARS-CoV-2 GNMT

Ligand	Binding (Kcal/mol) affinity
Sinefungin	-7.6
Crinamidine	-8.5
Marmesin	-7.9
Sinensetin	-7.7

Binding Site analyses: Hydrogen bonding plays an important role in many biochemical processes such as protein-ligand interactions. By displacing water molecules, it enhances ligand binding [57]. Also, the orientation and length of an intermolecular hydrogen bond determine the direction and specificity ligand binding [58].

Hydrogen bonds (H-bonds) are abundant in nature and are vital in protein folding, protein-ligand interactions as well as catalytic reactions. In biological systems, they are generally considered as facilitators of protein-ligand binding [59,60]. An increasing number of H-bonds between protein and drug molecule in molecular simulations is indicative of a stronger binding affinity [61].

Figure 4 & 5 and Table 4 reveals that while the standard has the highest number of intermolecular hydrogen bonds (eight) while Marmesin form the least (one). Of all the lead compounds, Crinamidine has the highest number of hydrogen bonds (four).

All hydrogen bonds of the lead compounds and standard fall within Pocket 41. Regarding the angles formed by hydrogen bonds, the standard forms four strong (greater than 130°) and four (less than 130°) hydrogen bonds with the target protein. Crinamidine forms two weak and two strong hydrogen bonds. Other lead compounds form only weak hydrogen bonds [62].

Regarding the donor to acceptor distance, the standard forms six moderate (2.5-3.2 Å) and two weak (3.2-4.0 Å) hydrogen bonds with the target protein. Crinamidine forms two moderate and two weak hydrogen bonds. Marmesin and Sinensetin forms only weak bonds [62].

The identification of potential protein-ligand interactions is an integral aspect of drug discovery as it aids the discovery of possible new drug leads, thus contributing to the advancement from hits to leads and prediction of likely explanations for side effects of approved drug candidates [63]. The most frequently observed interactions in ligand design are hydrophobic bonds, hydrogen bonds and π-stacking, followed by weak hydrogen bonds, salt bridges, amide stacking, and cation–π interactions [64]. The presence of hydrophobic interactions and salt bridges further strengthens and stabilizes the protein-ligand complexes [65].

The salt bridge is the strongest non-covalent bond, and it gives greater stability of the Protein-ligand complex [66]. From Table 5, GNMT-Crinamidine and GNMT-Marmesin complexes form salt bridges at residues ASP352 and HIS424 respectively. GNMT-Marmesin also has the highest number of hydrophobic interactions. This suggests a slightly more atom-efficient binding than other complexes. GNMT-Crinamidine has also p-stacking contributing to the small molecule interaction.
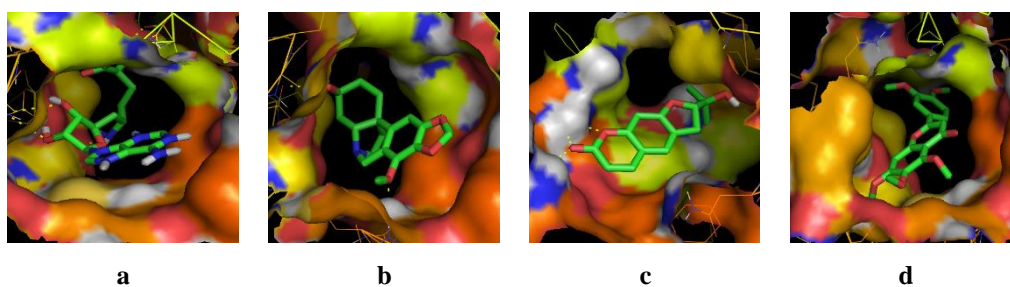

Figure 4: Binding site of SARS-CoV-2 GNMT interacting with standard and lead compounds a: GNMT-Sinefungin complex b: GNMT-Crinamidine complex c: GNMT- Marmesin complex d: GNMT-Sinensetin complex
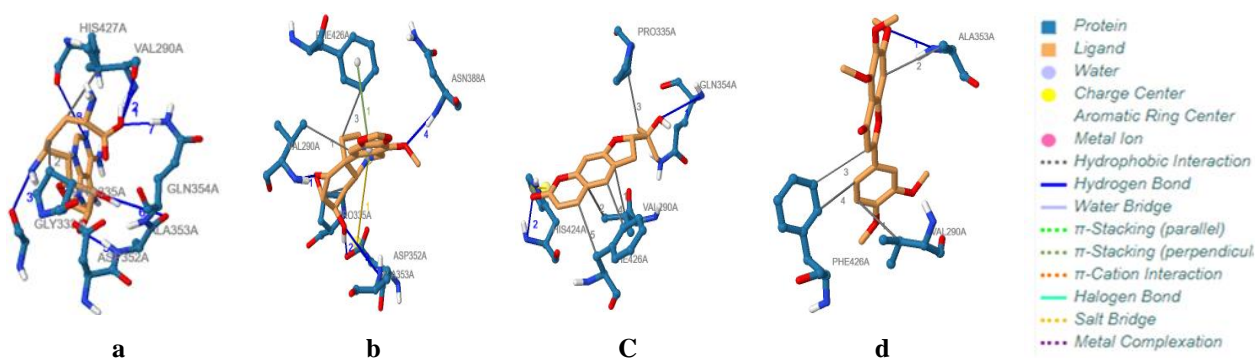

Figure 5: Protein-Ligand interactions of SARS-CoV-2 GNMT with standard and lead compound. a: GNMT-Sinefungin complex b: GNMT-Crinamidine complex c: GNMT- Marmesin complex d: GNMT-Sinensetin complex

Table 4. Hydrogen bond analysis

Complex	Number of bonds	Residues	Distance (H-A)	Distance (D-A)	Bond angle
GNMT-Sinefungin	8	VAL290	2.65	3.19	114.05
		VAL290	2.05	2.9	145.62
		GLY333	2.3	3.02	124.56
		ASP352	1.96	2.87	151.24
		ALA353	2.08	3.02	155.19
		ALA353	2.99	3.92	157.5
		GLN354	2.4	2.96	114.36
		HIS427	3.18	3.86	125.1
GNMT-Crinamidine	4	VAL290	3.34	4.04	129.2
		ASP352	2.13	3.09	159.62
		ALA353	3.38	3.95	118.21
		ASN388	2.17	3.16	167.33
GNMT-Marmesin	2	GLN354	3.46	3.78	101.13
		HIS424	2.67	3.36	126.18
GNMT-Sinensetin	1	ALA353	3.07	3.64	117.17

Table 5. Other Protein-ligand interactions

Complex	Hydrophobic Int.		Salt bridge		p-Stacking	
	Residue	Distance	Residue	Distance	Residue	Distance
GNMT-Sinefungin	VAL290	3.85				
	PRO335	3.72				
GNMT-Crinamidine	VAL290	3.82	ASP352	5.28	PHE426	4.78
	PRO335	3.8				
	PHE426	3.74				
GNMT-Marmesin	VAL290	3.79	HIS424	4.1		
	VAL290	3.85				
	PRO335	3.8				
	PHE426	3.61				
	PHE426	3.49				
GNMT-Sinensetin	VAL290	3.9				
	ALA353	3.87				
	PHE426	3.68				
	PHE426	3.72				

Analysis of MDS

Root Mean Square Deviation of Atomic Positions (RMSD): Through a computational approach, the RMSD is used assess the quality of a reproduced binding pose. The new structures induced by simulation and/or ligand binding is compared to a reference structure where the RMSD is at zero. The structural distance between the Cα atoms of the protein backbone is used as a means of evaluation. Lower RMSD values show greater stability of the biological configuration. Higher values suggest greater structural instability [56, 67,68].

In a 2-nanosecond trajectory, the RMSD of Apo and Holo structures were measured over consistent time frames (Figure 7 and Table 6). Of all the holo structures, the GNMT-Sinefungin complex has the least total and average RMSD values. The other lead compounds produced greater total and average RMSD values than the standard (Sinefungin). Crinamidine followed closely by Sinensetin induced the greatest total and average RMSD values.

There is a steep increase in RMSD of the simulated Apo protein relative to the crystal structure as the production time increased. The slope suggests that the RMSD values would increase with more simulation time. In a similar manner, the holo structures formed by the Crinamidine and Sinensetin also showed a steep increase of RMSD values all through the trajectory showing instability. This is also shown in the time frame in which their respective highest RMSD values were attained (20 and 19 respectively). The GNMT-Marmesin complex shows a gentle slope which flattens towards the end of the trajectory. The GNMT-Sinefungin complex shows the greatest stability with the least gradient of the slope.

The distribution of RMSD values of the Apo and holo structures (Figure 8 and Table 6) suggests that the greatest deviation to the right from the respective reference structures comes from the GNMT-Crinimadine complex. A total of 17 peaks were found between RMSD values 3.0 to 5.0 Å for the GNMT-Crinimadine complex while 17, 17, and 16 peaks were found in the same positions for the GNMT-Sinefungin, GNMT-Sinensetin, and GNMT-Marmesin complexes respectively. The GNMT-Crinimadine complex shows a wider RMSD range than the GNMT-Sinefungin and GNMT-Sinensetin complexes. This is because the GNMT-Crinimadine complex has a peak between the 5.00 -5.49 Å range while the GNMT-Sinefungin complex has no peak beyond 4.0 Å and GNMT-Sinensetin complex has no peak beyond 5.0 Å.

Put together, during the course of the simulation, the ligand-induced protein conformations have changed between different time points in the trajectory. The RMSD data suggests that

Crinamdine, Marmesin and Sinensetin in this order induced more structural distortion to GNMT than the standard. Crinamdine followed closely by Sinensetin showed the greatest ligand induced instability of the viral protein.

RMSF: The function of a protein is largely dependent on its function and dynamics. Protein motions are global, regional (domain or active site) and local (residue). Protein dynamics can be evaluated through the measure of the root mean square fluctuations (RMSF) of aligned residues. [69].

From Figure 9 and Table 6, the total and average global RMSF is greater in the GNMT-Crinamidine complex than all the other holo structures and least in the GNMT-Sinefungin complex (Standard). In this regard, the GNMT-Crinamidine complex is followed by the GNMT-Sinestein complex. The total and average regional (pocket 41) RMSF remained highest in the GNMT-Crinamidine complex followed by the GNMT-Sinestein complex. The lowest values are seen in the GNMT-Sinefungin complex for Pocket 41. In a similar vein, the highest fluctuation and highest range of RMSF was found in the GNMT-Crinamidine complex is followed by the GNMT-Sinestein complex. GNMT-Sinefungin complex (Standard) had the lowest values.

Put together, Crinamidine showed the most instability with the greatest fluctuations at both global and regional sites followed by Sinensetin. Globally, Sinefungin showed the least fluctuation at the regional (Pocket 41) site.

Radius of Gyration: The RoG analysis run to indicate or ascertain the compactness of the secondary structures within the 3D structure of the protein. It is measured from the center of mass of the molecule with a high RoG suggesting loose packing while a low RoG suggests a tight packing of the protein [70].

Graphical representation of the RoG reveals that the GNMT-Crinamindine complex has a steep slope in the upward direction showing the least compactness. The GNMT-Sinefungin complex also progressed upwardly howbeit with a gentle slope (Figure 10). The GNMT-Marmesin complex shows a gentle slope with a downward trend as the trajectory progressed. The GNMT-Sinensetin appear to be flat with a slight downward trend.

Trajectory data for RoG reveals that the GNMT-Crinamindine complex had the highest values of average gyration, range of gyration and percentage gyration over the trajectory. This made it the least compact of all the holo structures. The GNMT- Sinefungin complex is the most compact and only marginally different from the GNMT Marmesin complex. (Table 6). Put together, the Crinamdine followed closely by Sinensetin induced the greatest conformational changes on the target protein as shown by the least compactness. This suggests that they are better GNMT inhibitors than the standard.

B-Factor: The B-Factor or Temperature factor is an evaluation of the thermostability of the protein molecule as it measures the internal atomic motions as reflected in their flexibility or rigidity [71]. The B-factor also directly impacts on the residual factor (R factor) which is a determinant of the stereochemical quality of protein structure coordinates [72].

From Figure 11 and Table 6, the graphical plots of the B factor values show high values at the termini of the protein molecules suggesting molecular flexibility at these ends, and that the GNMT-Sinefungin complex is the most thermally stable of all the holo structures. The global average B-Factor value of the GNMT–Crinamidine complex is the highest of all the holo structures while the GNMT-Sinefungin has the lowest value. This suggests that at the global level the lead compounds-induced conformations are more thermally unstable than the conformation induced by the standard. In a similar vein, data of the regional average B factor obtained from Pocket 41 suggests that the GNMT–Crinamidine complex has the highest values of all the holo structures and this was followed by the GNMT–Sinensetin complex. In only the GNMT-Crinamideine complex, the average B-factor value for the Pocket 41 is higher than that of the global average. The GNMT-Marmesin complex has the least B factor value at the regional level. Put together, the greatest temperature-dependent atomic vibrations were induced by Crinamidine binding causing the greatest dynamic disorder of the GNMT stereochemistry.

Principal components Analysis: New conformations are generated during the molecular dynamic simulation of a protein. The statistical significance of these conformations is determined by the use of principal component analysis (PCA) [73]. Of all the holo structures, the total global motions (mean of PC1, PC2 and PC3) was highest in the GNMT-Sinensetin complex and least in the GNMT-Crinamidine complex. However, the, total regional motions (mean of

PC1, PC2 & PC3) was highest in GNMT-Sinefungin complex followed closely by the GNMT-Sinensetin and GNMT-Crinamidine complexes (Figure 12 and Table 6).

Specifically based on the greatest motions, the best global conformations are PC2 of the Apo protein, PC1 of GNMT-Sinefungin complex, PC1 of GNMT-Crinamidine complex, PC1 of the GNMT-Marmesin complex and PC3 of the GNMT-Sinensetin complex. Of all these holo structures, the GNMT-Sinensetin complex has the greatest motion. Similarly, the best conformations that produced the greatest motions at Pocket 41 are PC3, PC3, PC3, PC1 and PC2 of the Apo protein, GNMT-Sinefungin complex, GNMT-Crinamidine complex, GNMT-Marmesin complex and the GNMT-Sinensetin complexes respectively. Of all these holo structures, the GNMT-Crinamidine complex has the greatest motion at the Pocket 41.

The convergence of the MD simulation is revealed by the cosine contents of the principal components. Convergence shows sampling quality, accuracy and reproducibility. Table 6 shows Results of cosine content show good quality except for a slight non-convergence at the PC3 of the GNMT-Sinefungin complex [74].

The dynamic cross-correlation (DCC) analysis: This is a standard method for analyzing significant intermolecular contacts are rapidly substituted by side-chain flipping in molecular dynamic simulations [75]. The dynamic cross-correlation map captures the multimodal characteristics of atoms, especially at the interface of macromolecules by quantifying the correlation coefficients of motions between atoms depicting data as positive and negative correlation effect of amino acids [75,76].

From Figure 13, the strongest overall anti-correlated motion of residues occurred in the GNMT-Crinamidine complex. The active site of GNMT fall within the range of residues 289-426. The GNMT-Sinefungin complex showed non-correlated between residues 300-400 while the other residues in the active site showed moderate anti-correlation motions. The GNMT-Crinamidine complex showed strong anticorrelation motions between residues 250-450 which covers the whole area of the active site. The GNMT-Marmesin complex shows moderate anticorrelated motions at approximately residues 280-300, predominantly non-correlated motions between residues 300-350 and predominantly moderate anticorrelated motions from residues 350-400. The GNMT-Sinensetin complex showed non-correlation, moderate correlation and moderate anticorrelation motions between residues 250-300. However, greater portion consisting of residues 300-450 show moderate anticorrelation motions.

Put together, the greatest anticorrelation motions both globally and regionally (at active site) was found in the GNMT-Crinamidine complex suggesting the greatest inhibitory activity. The heat map of the GNMT-Sinensetin complex also suggests a greater inhibitory activity than the standard at the active site.
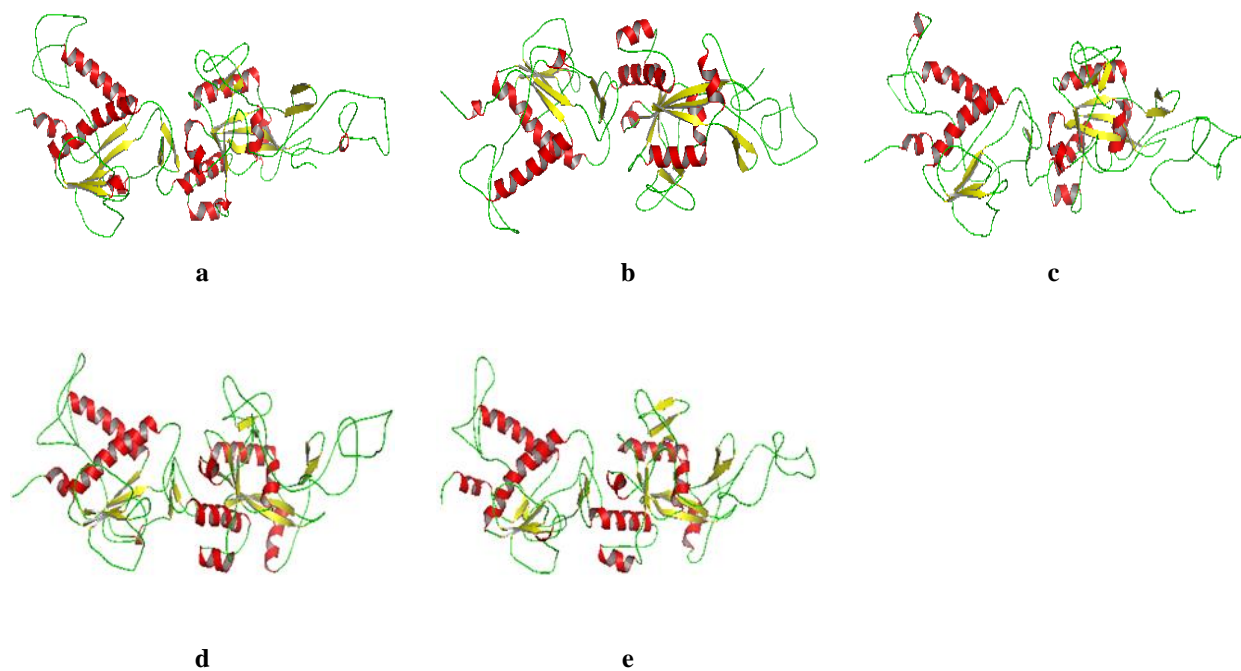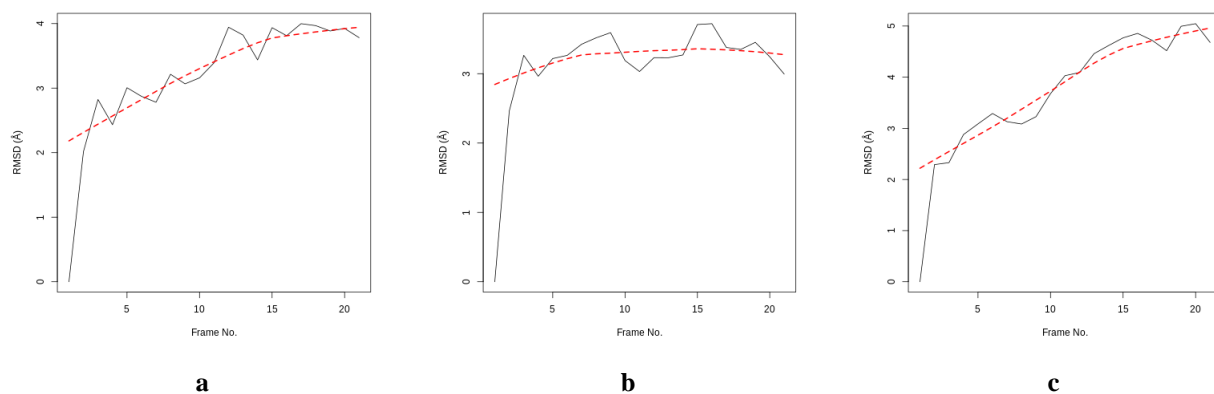

Figure 6: Cartoon model of the crystal structure of SARS-CoV-2 GNMT Apo and Holo structures (without water and ions) after molecular dynamics simulation. Beta sheets (yellow), Alpha helix (red) and Loops (green) a: GNMT b: GNMT-Sinefungin complex c: GNMT-Crinamidine complex d: GNMT- Marmesin complex e: GNMT-Sinensetin complex

d

e

Figure 7: RMSD for Apo and Holo structures a: GNMT b: GNMT-Sinefungin complex c: GNMT- Crinamidine complex d: GNMT- Marmesin complex e: GNMT-Sinensetin complex
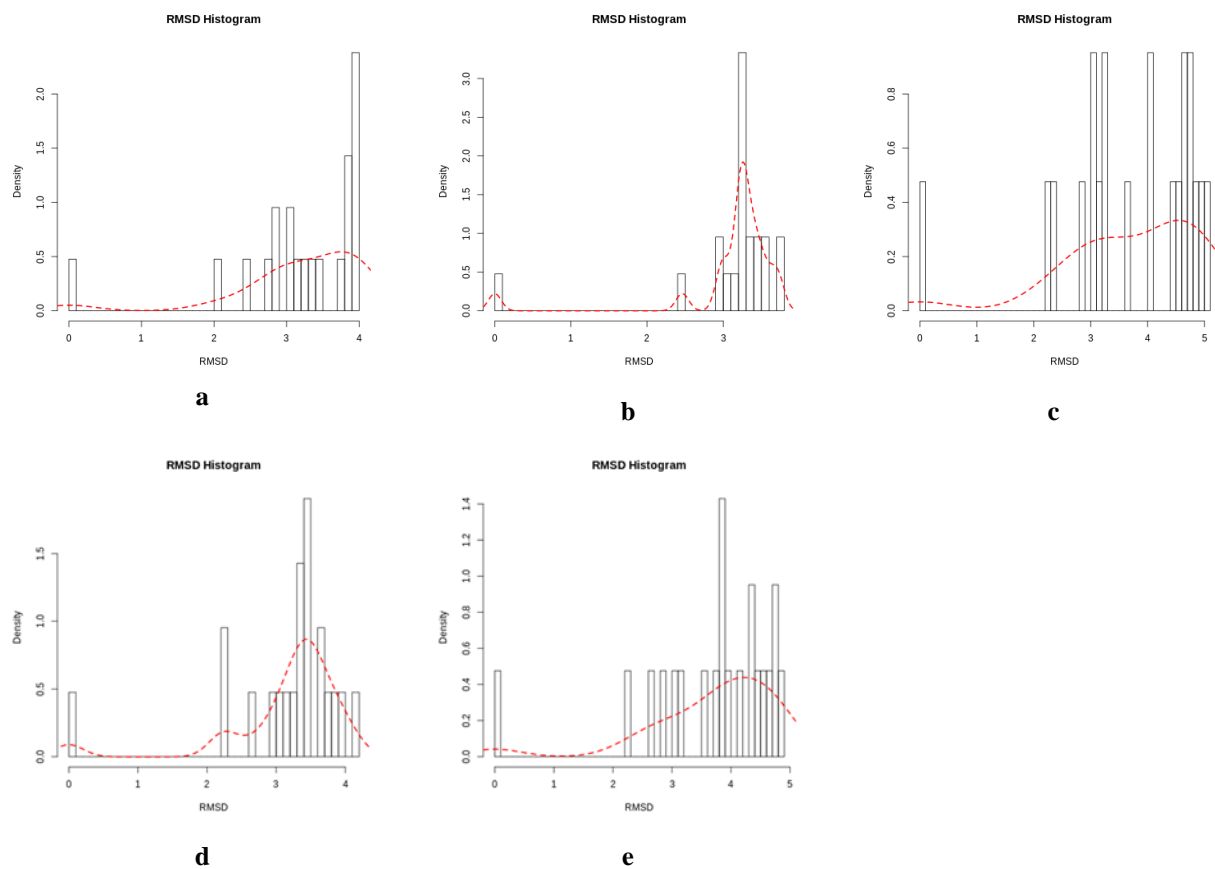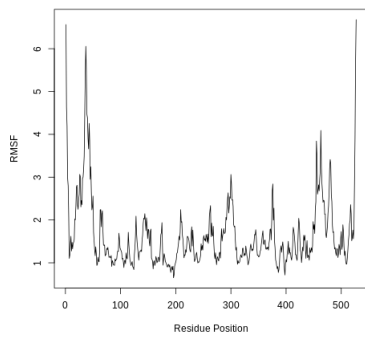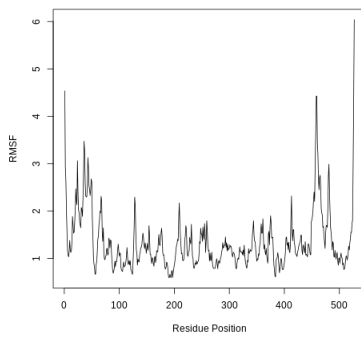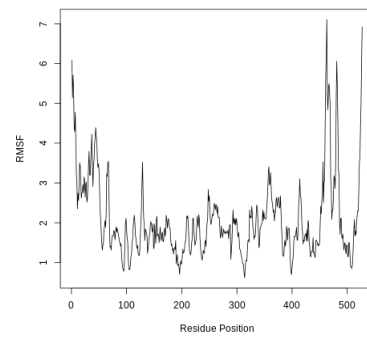

a

b

c

d

e

Fig 8: RMSD histogram for Apo and Holo structures a: GNMT b: GNMT-Sinefungin complex c: GNMT- Crinamidine complex d: GNMT- Marmesin complex e: GNMT-Sinensetin complex

Figure 9: Per-residue RMSF for Apo and Holo structures a: GNMT b: GNMT-Sinefungin complex c: GNMT- Crinamidine complex d: GNMT-Marmesin complex e: GNMT-Sinensetin complex

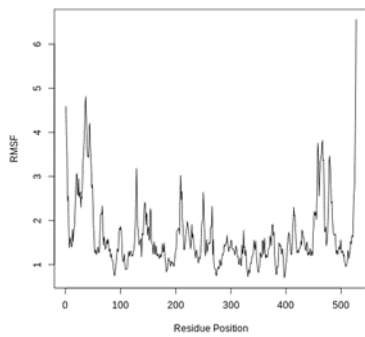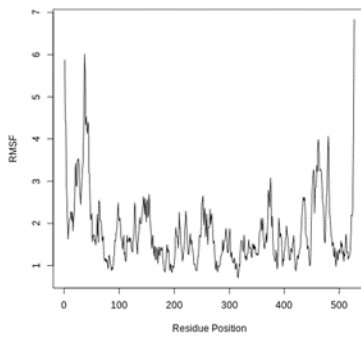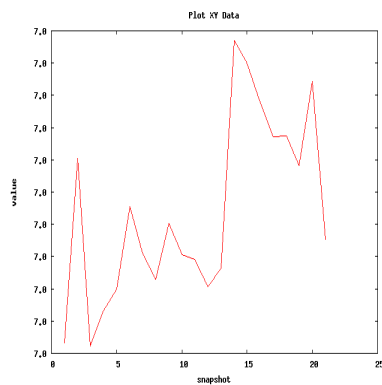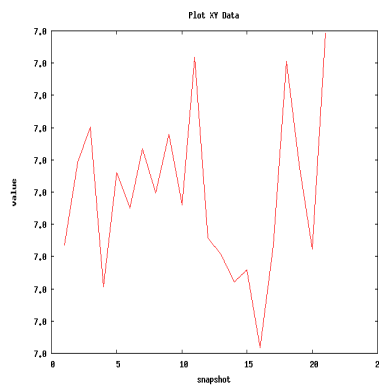**d**

e

Figure 10: Radius of gyration for Apo and Holo structures a: GNMT b: GNMT-Sinefungin complex c: GNMT- Crinamidine complex d: GNMT- Marmesin complex e: GNMT-Sinensetin complex

a

b

c

d

e
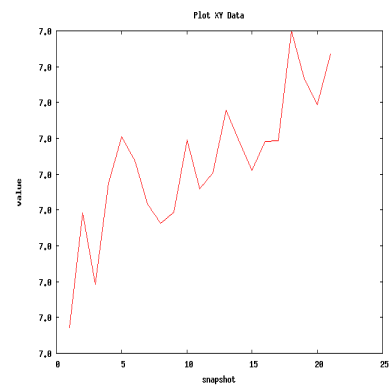
Figure 11: B Factor for Apo and Holo structures a: GNMT b: GNMT-Sinefungin complex c: GNMT- Crinamidine complex d: GNMT-Marmesin complex e: GNMT-Sinensetin complex

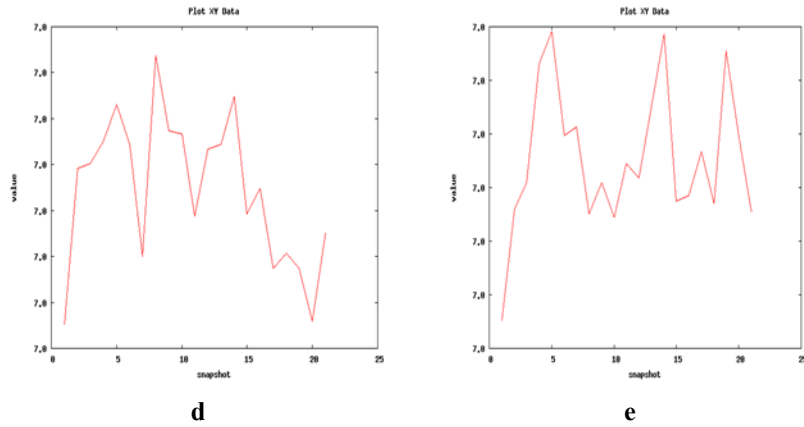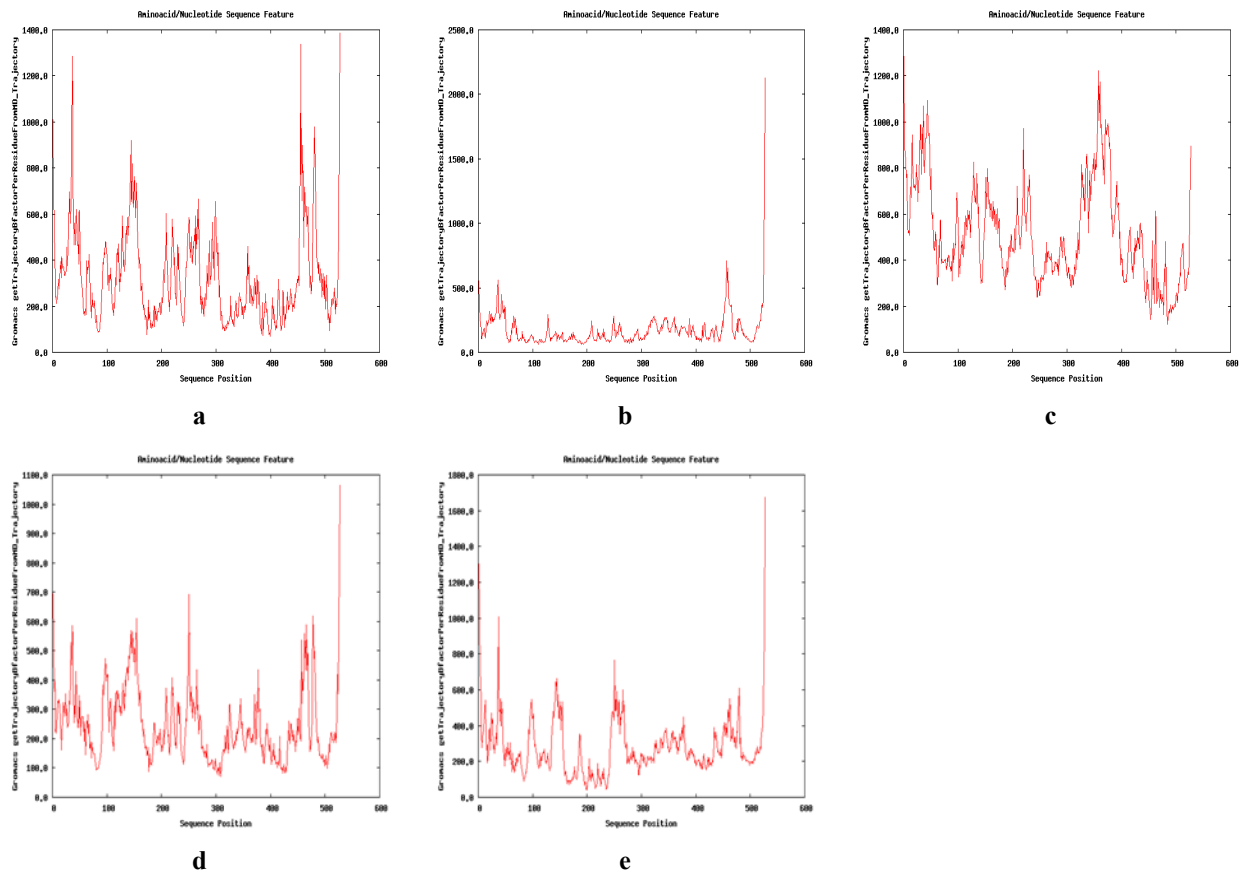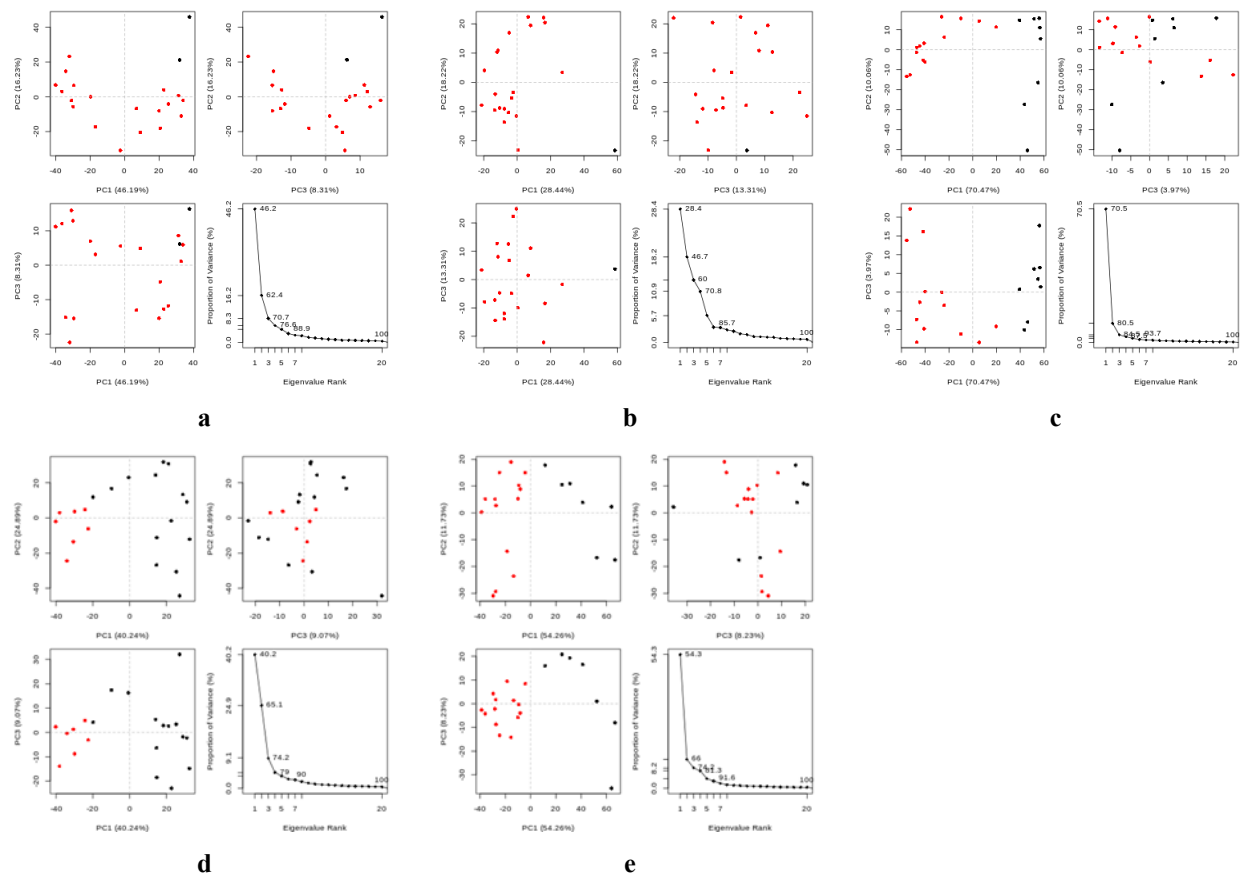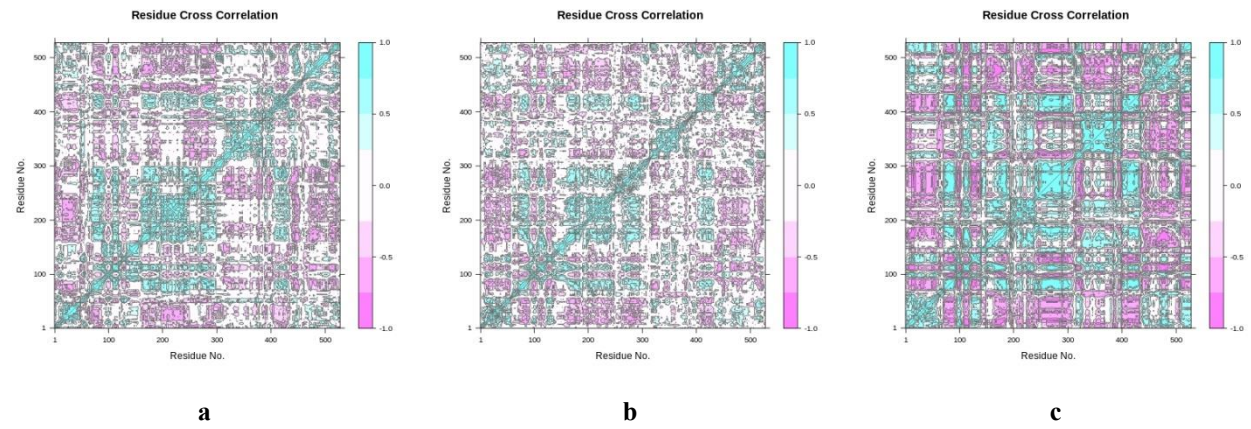Figure 12: Principle component analysis cluster plot of Apo and Holo structures. The projection of trajectory onto 1st few eigenvectors for: a: GNMT b: GNMT-Sinefungin complex c: GNMT-Crinamidine complex d: GNMT- Marmesin complex e: GNMT-Sinensetin complex

d   e

Fig 13: Dynamic cross correlation map Apo and Holo structures of 1m4k 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 -1.0 is anti-correlated. a: GNMT b: GNMT-Sinefungin complex c: GNMT- Crinamidine complex d: GNMT- Marmesin complex e: GNMT-Sinensetin complex

Table 6: Summary of data from Molecular Dynamics Simulations of Apo and Holo structures of SARS-CoV-2 GNMT

MDS Parameters	GNMT	GNMT-Sinefungin	GNMT-Crinamidine	GNMT-Marmesin	GNMT-Sinensetin
RMSD					
Total RMSD	67.269	65.504	77.78	66.383	77.69
Average RMSD	3.203	3.119	3.704	3.16	3.67
Lowest RMSD	0	0	0	0	0
Highest RMSD	3.998	3.721	5.046	4.108	4.895
Time Frame of Highest RMSD	17	16	20	17	19
Time Frame of Lowest RMSD	1	1	1	1	1
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	0	0	0	0	0
1.50 – 1.99A	0	0	0	0	0
2.00 – 2.49A	2	1	2	2	1
2.50 – 2.99A	3	2	1	2	2
3.00 – 3.49A	6	13	5	10	2

3.50 - 3.99	9	4	1	5	5
4.00 - 4.49	0	0	3	1	5
4.50 - 4.99	0	0	7	0	5
5.00 -5.49	0	0	1	0	0
RMSF					
Total Global RMSF	861.45	708.39	1100.42	864.69	946.28
Average Global RMSF	1.63	1.34	2.09	1.64	1.8
Total Regional (Pocket 41) RMSF	20.98	18.46	30.61	19.9	22.76
Average Regional (Pocket 41) RMSF	1.31	1.16	1.91	1.24	1.42
Least Fluctuation	0.65	0.59	0.62	0.7	0.71
Highest Fluctuation	6.68	6.04	7.1	6.55	6.83
Range of RMSF	6.03	5.45	6.48	5.85	6.12
PCA					
Total global motions (PC1, PC2 & PC3)	19.06382	19.42237	19.36949	19.6954	20.13226
Average global motions (PC1, PC2 & PC3)	0.03611	0.03685	0.03675	0.03737	0.0382
Total Regional (Pocket 41) Motion (PC1, PC2 & PC3)	0.44046	0.56208	0.55081	0.43385	0.55802
Average Regional (Pocket 41) Motion (PC1, PC2 & PC3)	0.02892	0.03513	0.03442	0.02712	0.03488
PC1 Eigenvalue	46.19%	28.44%	70.47%	40.24%	54.26%
PC2 Eigenvalue	16.23%	18.22%	10.06%	24.89%	11.73%
PC3 Eigenvalue	8.31%	13.31%	3.97%	9.07%	8.23%
Total	70.73%	59.97%	84.50%	74.20%	74.22%
PC1 cosine content	0.694	0.725	0.885	0.802	0.726
PC2 cosine content	0.592	0.003	0.452	0.635	0.553
PC3 cosine content	0.000	0.021	0.259	0.582	0.273
Radius of Gyration					
Average Gyration	6.9955	6.9929	6.9986	6.99292	6.9951
Maximim Gyration	7.00071	6.99792	7.00598	6.99873	6.99983
Minimum Gyration	6.99132	6.98818	6.98944	6.98703	6.98903
Range of Gyration	0.0094	0.0097	0.01654	0.117	0.0108
% Gyration	0.13	0.14	0.24	0.17	0.15
Time Frame of Max. Gyr.	14	21	18	8	5
Time Frame of Min.Gyr.	1	16	1	1	1

B Factor					
Global Average B Factor	324.51	168.72	513.95	247.71	276.95
Regional (Pocket 41) Average B Factor	220.53	167.69	696.14	154.53	260.06

Conclusion

After the virtual screening of a library of 1,048 natural compounds against the SARS-CoV-2 GNMT, three lead compounds namely Crinamidine, Sinensetin and Marmesin were identified. All the compounds showed good oral bioavailability properties except for the Standard which has a high TPSA value. The standard, and the lead compounds all showed favorable absorption, metabolism, excretion, and toxicity properties. The distribution pharmacokinetics are generally favorable except that all the compounds have a poor CNS permeability, poor BBB (except standard and Marmesin) and they are P-glycoprotein substrates (except Marmesin). The standard has the highest number of hydrogen bonds formed within the active site followed by Crinamidine. The trajectory data such as RMSD, RMSF, B-Factor, DCCM, and RoG, suggests that Crinamidine proved to cause the greatest distortion to the target protein while the standard caused the least at the global and regional levels (Pocket 41). Specifically of all the compounds, the PC3 of Crinamidine is the conformation caused the greatest distortion at the active site. Overall, the lead compounds proved to be better drug candidates than the standard in the following order: Crinamidine, Sinensetin and Marmesin.

Isolated for the *Streptomyces* species, Sinefungin is a natural nucleoside that is a derivative of S-adenosylmethionine (SAM) [77]. It has shown a wide range of biological effects which include amoebicidal, antifungal, antibacterial (*Streptococcus pneumoniae*) and antiparasitic (Plasmodium, malarial, trypanosomal, and leishmanial species) activities [77, 78,79,80]. The antiviral activity of Sinefungin has also been established as it has been shown to be an inhibitor of mRNA(guanine-7-)-methyltransferase, mRNA(nucleoside-2'-)-methyltransferase, and DNA methyltransferases [81, 82]. Sinefungin is has been shown to inhibit the multiplication of feline herpesvirus type I, Newcastle disease and vaccinia virus [82, 83].

Crinamidine is an alkaloid obtained from *Crinum latifolium* and *Talinum triangulare*. In Chinese ethnomedicine, the antiviral and antitumor properties of the extract of *Crinum latifolium* has been reported [84, 85]. Sinensetin **c**an be found in orange (*Citrus sinensis*) peel, and it has a wide range of biological activity such as antiviral, anticancer, antitumor, anti-inflammatory. Sinensetin is an important ingredient of the aqueous extract of *Orthosiphon stamineus* extract which has shown inhibitory properties against Herpes Simplex Virus type 1 [86, 87]. Marmesin can be found in mango and wheat [88]. Its inhibitory activity against the Epstein-Barr virus (EBV) has been reported [89]

It is recommended that the inhibitory effect of Crinamidine, Sinensetin and Marmesin on the active site of SARS-CoV-2 GNMT should be further investigated.

References

1. World Health Organization (2020). Infection prevention and control during health care when covid-19 is suspected https:// www.who,int/publications-detail.
2. Worldometer, 2020 Available at https://www.worldometers.info/coronavirus/?utm_campaign=homeAdvegas1?
3. deWit, E., vanDoremalen, N., Falzarano, D. &Munster, V. J. (2016). SARS and MERS: recent 345 insights into emerging coronaviruses. *Nat. Rev. Microbiol*. 14, 523–534.
4. Gorbalenya, A. E., Baker, S. C., Baric, R. S., de Groot, R. J., Drosten, C. & Gulyaeva A. A. (2020). "The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2". *Nature Microbiology*. 5 (4): 536–544.
5. Yu, C and Deyin, G. (2016). Molecular mechanisms of coronavirus RNA capping and methylation. *Virologica Sinica*. 31(1): 3-11.
6. Li, F. Structure, Function, and Evolution of Coronavirus Spike Proteins. *Annu. Rev. Virol*. 341 (2016).
7. Wong, A. C., Li, X., Lau, S. K., & Woo, P. C. (February 2019). "Global Epidemiology of Bat *Viruses* 2019, 11, 174.
8. A.S. Omrani, M.M. Saad, K. Baig, A. Bahloul, M. Abdul-Matin, A.Y. Alaidaroos, et al. (2014). Ribavirin and interferon alfa-2a for severe Middle East respiratory syndrome

coronavirus infection: a retrospective cohort study. *Lancet Infect Dis,* 14 (2014), pp. 1090-1095.

9. Wu, C., Liu, Y., Yang, Y., Zhang, P., Zhong, W., Wang, Y., Wang, Q., Xu, Y., Li, M., Li, X. and Zheng, M., (2020). Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. *Acta Pharmaceutica Sinica B*. 10(5):766-788

10. Shuman, S., 2015. RNA capping: progress and prospects. *RNA,* 21(4), pp.735-737.

11. Shuman, S., 2002. What messenger RNA capping tells us about eukaryotic evolution. *Nature reviews Molecular cell biology,* 3(8), pp.619-625.

12. Topisirovic, I., Svitkin, Y.V., Sonenberg, N. and Shatkin, A.J., 2011. Cap and cap-binding proteins in the control of gene expression. *Wiley Interdisciplinary Reviews: RNA*, 2(2), pp.277-298.

13. Dhaval Varshney, Alain-Pierre Petit, Juan A. Bueren-Calabuig, Chimed Jansen, Dan A. Fletcher, Mark Peggie, Simone Weidlich, Paul Scullion, Andrei V. Pisliakov, and Victoria H. Cowling, (2016). Molecular basis of RNA guanine-7 methyltransferase (RNMT) activation by RAM. *Nucleic Acids Res*. 44(21): 10423–10436.

14. Chen, Y., Cai, H., Xiang, N., Tien, P., Ahola, T. and Guo, D., 2009. Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase. *Proceedings of the National Academy of Sciences,* 106(9), pp.3484-3489.

15. Chen, Y., Tao, J., Sun, Y., Wu, A., Su, C., Gao, G., Cai, H., Qiu, S., Wu, Y., Ahola, T. and Guo, D., 2013. Structure-function analysis of severe acute respiratory syndrome coronavirus RNA cap guanine-N7-methyltransferase. *Journal of virology*, 87(11): 6296-6305.

16. Sun, Y., Wang, Z., Tao, J., Wang, Y., Wu, A., Yang, Z., Wang, K., Shi, I., Chen, Y. & Guo, D. (2014). Yeast-based assays for the high-through-put screening of inhibitors of coronavirus RNA cap guanine-N7-methyltranferase. *Antiviral Res.* 104: 156-164.

17. Aouadi, W., Eydoux, C., Coutard, B., Martin, B., Debart, F., Vasseur, J.J., Contreras, J.M., Morice, C., Quérat, G., Jung, M.L. and Canard, B., (2017). Toward the identification of viral cap-methyltransferase inhibitors by fluorescence screening assay. *Antiviral research,* 144, pp.330-339.

18. Decroly, E. and Canard, B., 2017. Biochemical principles and inhibitors to interfere with viral capping pathways. *Current opinion in virology,* 24, pp.87-96.

19. Guilloux, L.V., Peter, S., & Pierre T. (2009) "Fpocket: An open source platform for ligand pocket detection", *BMC Bioinformatics,* 10:168

20. Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc* 5:725–738.

21. Ramachandran, S., Kota, P., Ding, F. & Dokholyan, N. V. (2011). PROTEINS: Structure, Function and Bioinformatics, 79: 261-270

22. Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., Li, Q., Shoemaker, B. A., Thiessen, P. A., Yu, B., Zaslavsky, L., Zhang, J., & Bolton, E. E. (2019). PubChem 2019 update: improved access to chemical data. *Nucleic acids research,* 47(D1), D1102–D1109.

23. Atatreh N, Al Rawashdah S, Al Neyadi SS, Abuhamdah SM, Ghattas MA (2019) Discovery of new butyrylcholinesterase inhibitors via structure-based virtual screening. *J Enzyme Inhib Med Chem* 34:1373–1379.

24. Dallakyan S, Olson AJ (2015) Small-molecule library screening by docking with PyRx. *Methods Mol Biol* 1263:243–250.

25. Gacche, R.N., Meshram, R.J., Dawane, A.A., Kamble, S.S., Shinde, G.P., Dhabadge, V.N. and Patil, K.K., (2016). Modeling studies of arginase from *Helicobacter pylori* divulge novel inhibitor-protein interactions. *Chemistry* 1(3): 54-61.

26. Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* 31:455–461.

27. Daina, A., Olivier, M., & Vincent, Z. (2017). SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci Rep.* 7: 42717

28. Pires, D.E.V., Tom, L., Blundell, D. B., Ascher. (2015). pkCSM: predicting small-molecule pharmacokinetic properties using graph-based signatures. *Journal of Medicinal Chemistry,* 58 (9): 4066–4072.

29. Molinspiration (2015) Calculation of Molecular Properties and Bioactivity Score. http://www.molinspiration.com/cgi-bin/properties.

30. Mishra, Shashank Shekhar, Chandra Shekhar Sharma, Hemendra Pratap Singh, Harshda Pandiya, and Neeraj Kumar (2016) "In silico ADME, Bioactivity and Toxicity Parameters Calculation of Some Selected Anti-Tubercular Drugs." *International Journal of Pharmaceutical and Phytopharmacological Research* 6, no. 6 (2016): 77-79.

31. DeLano WL (2002) Pymol: An open-source molecular graphics tool. CCP4 Newsl *Protein Crystallogr* 40:82–92.

32. Salentin S, Schreiber S, Haupt VJ, Adasme M.F, Schroeder M (2015) PLIP: fully automated protein-ligand interaction profiler. *Nucleic Acids Res* 43: W443-7.

33. Lionta, E., Spyrou, G., K Vassilatis, D. and Cournia, Z., 2014. Structure-based virtual screening for drug discovery: principles, applications and recent advances. *Current topics in medicinal chemistry,* 14(16): 1923-1938.

34. Afghan, E., Dannon, B., Marius, V.B., Daniel, B., Dave, B., Martin, Č., John, C., Dave, C,. Jeremy, G. (2016). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: update, *Nucleic Acids Research* 44(W1): W3-W10

35. Franco-Ulloa, S., Riccardi, L., Rimembrana, F., Pini, M. and De Vivo, M., 2019. Nanomodeler: A webserver for molecular simulations and engineering of nanoparticles. *Journal of chemical theory and computation,* 15(3): 2022-2032.

36. Dodda, L.S., Cabeza, V. I., Tirado-Rives, J., Jorgensen, W.L. (2017). LigParGen web server: an automatic OPLSAA parameter generator for organic ligands. *Nucleic Acids Res.* 3; 45(W1): W331-W336.

37. Abraham, M.J., R, S., SzilárdPáll, J. C., SmithBerk, H.E. (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *Software X* Volumes 1–2, Pages 19-25

38. Adam Hospital, Pau Andrio, Carles Fenollosa, Damjan Cicin-Sain, Modesto Orozco, Josep Lluís Gelpí (2012). MDWeb and MDMoby: an integrated web-based platform for molecular dynamics simulations. *Bioinformatics,* 28(9):1278-1279.

39. 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; 46:3–26

40. Okoli, P. T., Nzute, V. C., Durojaye, O. A., Chielo, O. H., Ajibo, Q. C., Udo, S. I., & Ezeh, S. M. (2019). An in-silico pharmacokinetics study on Cis-heptadeca-1, 9-Diene-4,

6-Diyne-3, 8-Diol: A nutraceutical compound with anticancer properties. *Asian Food Science Journal,* 1-7.

41. Ghose AK, Viswanadhan VN, Wendoloski JJ (1999) 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* 1: 55-68.

42. Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, et al. (2002) Molecular properties that influence the oral bioavailability of drug candidates. *J Med Chem* 45: 2615-2623.

43. Al Wasidi AS, Hassan AS, Naglah AM. In vitro cytotoxicity and druglikeness of pyrazolines and pyridines bearing benzofuran moiety. *J Appl Pharm Sci,* 2020; 10(04): 142–148.

44. David E Clark (2011). What Has Polar Surface Area Ever Done for Drug Discovery? *Future Med Chem.* 2011 Mar;3(4):469-84.

45. Clemons PA, Bodycombe NE, Carrinski HA, Wilson JA, Shamji AF, Wagner BK et al (2010) Small molecules of different origins have distinct distributions of structural complexity that correlate with protein-binding profiles. *Proc Natl Acad Sci* U.S.A 107:18787–18792.

46. Christian Feldmann, Filip Miljković, Dimitar Yonchev, and Jürgen Bajorath (2019). Identifying Promiscuous Compounds with Activity against Different Target Classes. *Molecules.* 24(22): 4185.

47. Imran, M., Choudhary, M. A., Kumar, N., Kousar, Z., Shahida, S., & Nohri, F. (2016). Synthesis, spectroscopic characterization and Petra Osiris Molinspiration (POM) analyses of dicarboxylic acid amides. *International Journal of Pharmaceutical Sciences and Research,* 7(5): 1915-1927.

48. Khan, T.; Dixit, S.; Ahmad, R.; Raza, S.; Azad, I.; Joshi, S.; Khan, A (2017). 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.* 10, 91–104.

49. Lin, Jing, Diana C. Sahakian, S. M. De Morais, Jinghai J. Xu, Robert J. Polzer, and Steven M. Winter (2003). "The role of absorption, distribution, metabolism, excretion

and toxicity in drug discovery." *Current topics in medicinal chemistry* 3, No. 10 (2003): 1125-1154

50. Bachmakov, I., Werner, U., Endress, B., Auge, D., & Fromm, M. F. (2006). Characterization of β-adrenoceptor antagonists as substrates and inhibitors of the drug transporter P-glycoprotein 1. *Fundamental & clinical pharmacology,* 20(3), 273-282.

51. Amin M. L. (2013). P-glycoprotein Inhibition for Optimal Drug Delivery. *Drug target insights,* 7, 27–34.

52. Varma MV, Ashokraj Y, Dey CS, Panchagnula R. (2003). P-glycoprotein inhibitors and their screening: a perspective from bioavailability enhancement. *Pharmacol Res.* 48(4):347-59.

53. Ondieki, G., Nyagblordzro, M., Kikete, S., Liang, R., Wang, L., & He, X. (2017). Cytochrome P450 and P-Glycoprotein-Mediated Interactions Involving African Herbs Indicated for Common Noncommunicable Diseases. Evidence-based complementary and alternative medicine: *eCAM,* 2017, 2582463.

54. Zhou SF (2008) Drugs behave as substrates, inhibitors and inducers of human cytochrome P450 3A4. *Curr Drug Metab* 9:310–322.

55. Sliwoski G, Kothiwale S, Meiler J, Lowe EW Jr (2013) Computational methods in drug discovery. *Pharmacol Rev* 66:334–395.

56. Eric W. Bell and Yang Zhang (2019). DockRMSD: an open source tool for atom mapping and RMSD calculation of symmetric molecules through graph isomorphism. *J Cheminform* 11:40-49.

57. Sarkar, A.; Kellogg, G. E. (2010). Hydrophobicity-shake flasks, protein folding and drug discovery. *Curr. Top. Med. Chem.* 2010, 10, 67–83.

58. Jeffrey, G.A. An Introduction to Hydrogen Bonding [Online]; Oxford University Press, 1997.

59. J. Gao, D. A. Bosco, E. T. Powers, J. W. Kelly (2009), Localized thermodynamic coupling between hydrogen bonding and microenvironment polarity substantially stabilizes proteins. *Nat. Struct. Mol. Biol.* 16, 684–690

60. S. Salentin, V. J. Haupt, S. Daminelli, M. Schroeder, (2014) Polypharmacology rescored: Protein–ligand interaction profiles for remote binding site similarity assessment. *Prog. Biophys. Mol.Biol.* 116, 174–186

61. Chen Deliang, Numan Oezguen, Petri Urvil, Colin Ferguson, Sara M. Dann & Tor C. Savidge (2016). Regulation of protein-ligand binding affinity by hydrogen bond pairing. *Sci Adv* 2 (3); e1501240.

62. Desiraju, G.; Steiner, T. (2001). The Weak Hydrogen Bond. In *Structural Chemistry and Biology*; Oxford University Press: Oxford,

63. Li Li, Ching Chiek Koh, Daniel Reker, J. B. Brown, Haishuai Wang, Nicholas Keone Lee, Hien-haw Liow, Hao Dai, Huai-Meng Fan, Luonan Chen & Dong-Qing Wei.(2019) Predicting protein-ligand interactions based on bow-pharmacological space and Bayesian additive regression trees. *Sci Rep* 9, 7703

64. Ferreira de Freitas, R., & Schapira, M. (2017). A systematic analysis of atomic protein-ligand interactions in the PDB. *Med Chem Comm*, 8(10), 1970–1981.

65. Kumar, S.; Nussinov, R. (2002). Close-range electrostatic interactions in proteins. *Chembiochem.* 2002, 3, 604–617.

66. Rafał Kurczab, Paweł Śliwa, Krzysztof Rataj, Rafał Kafel and Andrzej J. Bojarski (2018). Salt Bridge in Ligand–Protein Complexes—Systematic Theoretical and Statistical Investigations. *J. Chem. Inf. Model.* 58, 11, 2224–2238.

67. Maiorov VN, Crippen GM (1994). Significance of root-mean-square deviation in comparing three-dimensional structures of globular proteins. *J Mol Biol.* 1994;235(2):625-634.

68. Kirchmair, J., Markt, P., Distinto, S., Wolber, G. and Langer, T., 2008. Evaluation of the performance of 3D virtual screening protocols: RMSD comparisons, enrichment assessments, and decoy selection—what can we learn from earlier mistakes? *Journal of computer-aided molecular design,* 22(3-4): 213-228.

69. Edvin Fuglebakk, Julián Echave, Nathalie Reuter (2012). Measuring and comparing structural fluctuation patterns in large protein datasets. *Bioinformatics,* Volume 28, Issue 19, 1 October 2012, Pages 2431–2440.

70. Sneha, P. and Doss, C.G.P., (2016). Molecular dynamics: new frontier in personalized medicine. In *Advances in protein chemistry and structural biology* (Vol. 102, pp. 181-224). Academic Press.

71. Zhoutong Sun, Qian Liu, Ge Qu, Yan Feng, and Manfred T. Reetz (2019). Utility of B-Factors in Protein Science: Interpreting Rigidity, Flexibility, and Internal Motion and Engineering Thermostability. *Chem. Rev.* 2019, 119, 3, 1626–1665.

72. M. Carson, C. Bugg, L. DeLucas, S. Narayana (1994). Comparison of homology model to the experimental structure of a novel serine protease. *Acta Crystallogr. D Biol. Crystallogr.,* 50 (1994), pp. 889-899

73. David CC, Jacobs DJ (2014) Principal component analysis: a method for determining the essential dynamics of proteins. *Methods Mol Biol* 1084:193–226.

74. Sawle L, Ghosh K (2016) Convergence of Molecular Dynamics Simulation of Protein Native States: Feasibility vs Self-Consistency Dilemma. *J Chem Theory Comput* 12:861–869.

75. Hasahara K, Fukuda I, Nakamura H (2014) A Novel Approach of Dynamic Cross Correlation Analysis on Molecular Dynamics Simulations and Its Application to Ets1 Dimer-DNA Complex. Available from: https://www.researchgate.net/publication/267930440_A_Novel_Approach_of_Dynamic_Cross_Correlation_Analysis_on_Molecular_Dynamics_Simulations_and_Its_Application_to_Ets1_Dimer-DNA_Complex [accessed May 20 2020].

76. Hassan Mubashir, Shahzadi Saba, Seo Sung Y., Alashwal Hany, Zaki Nazar, Moustafa Ahmed A. (2018). Molecular Docking and Dynamic Simulation of AZD3293 and Solanezumab Effects Against BACE1 to Treat Alzheimer's Disease. *Frontiers in Computational Neuroscience,* vol 12, 2018, pp 1-34.

77. Nolan L L (1987). Molecular target of the antileishmanial action of sinefungin. *Antimicrob Agents Chemother.* 1987 Oct;31(10):1542-8.

78. Yadav Mukesh Kumar, Seok-Won Park, Sung-Won Chae, Jae-Jun Song (2014). Sinefungin, a natural nucleoside analogue of S-adenosylmethionine, inhibits Streptococcus pneumoniae biofilm growth *Biomed Res Int.* 2014;2014:156987.

79. W Trager, M Tershakovec, P K Chiang, G L Cantoni (1980). Plasmodium falciparum: antimalarial activity in culture of sinefungin and other methylation inhibitors. *Exp Parasitol* 1980 Aug;50(1):83-9.

80. D K Dube, G Mpimbaza, A C Allison, E Lederer, L Rovis (1983). Antitrypanosomal activity of sinefungin. *Am J Trop Med Hyg* 32(1):31-3.

81. Copeland Robert A, Michael E Solomon, Victoria M Richon (2009). Protein methyltransferases as a target class for drug discovery. *Nat Rev Drug Discov.* 2009 Sep;8(9):724-32.

82. Pugh C S Borchardt R T and Stone H O (1978). Sinefungin, a potent inhibitor of virion mRNA(guanine-7-)-methyltransferase, mRNA(nucleoside-2'-)-methyltransferase, and viral multiplication. *The Journal of Biological Chemistry.* 253, 4075-4077.

83. Kuroda, Y., Yamagata, H., Nemoto, M. et al. (2019). Antiviral effect of sinefungin on in vitro growth of feline herpesvirus type 1. *J Antibiot* 72, 981–985 (2019).

84. Marcel Jenny, Angela Wondrak, Elissaveta Zvetkova, Nguyen Thi Ngoc Tram, Phan Thi Phi Phi, Harald Schennach, Zoran Culig, Florian Ueberall, and Dietmar Fuchs, (2011). Crinum Latifolium Leave Extracts Suppress Immune Activation Cascades in Peripheral Blood Mononuclear Cells and Proliferation of Prostate Tumor Cells. *Sci Pharm.* 2011 Apr-Jun; 79(2): 323–335.

85. Catherine C. Ikewuchi, Jude C. Ikewuchi, and Mercy O. Ifeanacho (2017). Bioactive phytochemicals in an aqueous extract of the leaves of Talinum triangulare. *Food Sci Nutr.* 5(3): 696–701.

86. Nur Suhana Mohamad Ripim, Nuraini Fazil, Saidatul Nadrah Kholid Ibrahim, Adibah Ahamad Bahtiar, Chee Wai Yip, Nazlina Ibrahim, Norefrina Shafinaz Md Nor (2018). Antiviral Properties of Orthosiphon stamineus Aqueous Extract in Herpes Simplex Virus Type 1 Infected Cells. *Sains Malaysiana* 47(8):1725-1730.

87. M.Amzad Hossaina and Zhari Ismail (2016). Quantification and enrichment of sinensetin in the leaves of *Orthosiphon stamineus*. *Arabian Journal of Chemistry.* Volume 9, Supplement 2, November 2016, Pages S1338-S1341

88. Scalbert, A.; Andres-Lacueva, C.; Arita, M.; Kroon, P.; Manach, C.; Urpi-Sarda, M.; Wishart, D.S. (2011). "Databases on Food Phytochemicals and Their Health-Promoting Effects". *J. Agric. Food Chem.* 59 (9): 4331–4348.

89. Gwendoline Cheng Lian Ee and Ahmad Faizal Abdull Razis (2018). Phytochemical Constituents and Biological Activities of Melicope lunu-ankenda. *Molecules* 2018, 23, 2708; doi:10.3390/molecules23102708

Figure Legends

Figure 1: a**:** Cartoon model of the crystal structure of SARS-CoV-2 GNMT (QHD43415_13.pdb). Beta sheets (yellow), Alpha helix (red) and Loops (green) b: Surface representations.

Figure 2: Ramanchandran plot for SARS-CoV-2 GNMT (QHD43415_13.pdb).

Figure 3: The 3D chemical structures (stick model) of standard and lead compound a: Sinefungin b: Crinamidine c: Marmesin   d: Sinensetin

Figure 4: Binding site of SARS-CoV-2 GNMT interacting with standard and lead compounds a: GNMT-Sinefungin complex b: GNMT- Crinamidine complex c: GNMT- Marmesin complex d: GNMT-Sinensetin complex

Figure 5: Protein-Ligand interactions of SARS-CoV-2 GNMT with standard and lead compounds a: GNMT-Sinefungin complex b: GNMT-Crinamidine complex c: GNMT-Marmesin complex d: GNMT-Sinensetin complex

 Figure 6: Cartoon model of the crystal structure of SARS-CoV-2 GNMT Apo and Holo structures (without water and ions) after molecular dynamics simulation. Beta sheets (yellow), Alpha helix (red) and Loops (green) a: GNMT b: GNMT-Sinefungin complex c: GNMT-Crinamidine complex d: GNMT- Marmesin complex e: GNMT-Sinensetin complex

Figure 7: RMSD for Apo and Holo structures a: GNMT b: GNMT-Sinefungin complex c: GNMT- Crinamidine complex d: GNMT- Marmesin complex e: GNMT-Sinensetin complex

Fig 8: RMSD histogram for Apo and Holo structures a: GNMT b: GNMT-Sinefungin complex c: GNMT- Crinamidine complex d: GNMT- Marmesin complex e: GNMT-Sinensetin complex

Figure 9: Per-residue RMSF for Apo and Holo structures a: GNMT b: GNMT-Sinefungin complex c: GNMT- Crinamidine complex d: GNMT- Marmesin complex e: GNMT-Sinensetin complex

Figure 10: Radius of gyration for Apo and Holo structures a: GNMT b: GNMT-Sinefungin complex c: GNMT- Crinamidine complex d: GNMT- Marmesin complex e: GNMT-Sinensetin complex

Figure 11: B Factor for Apo and Holo structures a: GNMT b: GNMT-Sinefungin complex c: GNMT- Crinamidine complex d: GNMT- Marmesin complex e: GNMT-Sinensetin complex

Figure 12: Principle component analysis cluster plot of Apo and Holo structures. The projection of trajectory onto 1st few eigenvectors for: a: GNMT b: GNMT-Sinefungin complex c: GNMT-Crinamidine complex d: GNMT- Marmesin complex e: GMT-Sinensetin complex

Fig 13: Dynamic cross correlation map Apo and Holo structures of 1m4k 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: GNMT b: GNMT-Sinefungin complex c: GNMT- Crinamidine complex d: GNMT- Marmesin complex e: GNMT-Sinensetin complex

List of Tables

Table 1: Chemo-informatic properties of standard and lead compounds

Table 2: Pharmacokinetic properties of standard and lead compounds

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

Table 4: Hydrogen bond analysis of SARS-CoV-2 GNMT with standard and lead compounds

Table 5: Other Protein-ligand interactions

Table 6: Summary of data from Molecular Dynamics Simulations of Apo and Holo structures of SARS-CoV-2 GNMT