

Theoretical Investigation of the Cooperation of Iminoguanidine with the Enzymes-Binding Domain of Covid-19 and Bacterial Lysozyme Inhibitors and their Pharmacokinetic Properties

Emmanuel Israel Edache^{1,2*}, Adamu Uzairu², Paul Andrew Mamza², Gideon Adamu Shallangwa²

¹Department of Pure and Applied Chemistry, University of Maiduguri, Borno State, Nigeria.

²Department of Chemistry, Ahmadu Bello University, Zaria, Nigeria.

*Corresponding author: Emmanuel Israel Edache, email: edacheson2004@gmail.com

Received January 24th, 2022; Accepted August 1st, 2022.

DOI: <http://dx.doi.org/10.29356/jmcs.v66i4.1726>

Abstract. The investigation for innovative treatments for *Pseudomonas aeruginosa* and SARS-CoV-2 is a burgeoning field. Even though scientists and pharmaceutical companies have made significant contributions to the research of multi-drug resistance infections from a variety of perspectives, these diseases remain incurable. As a result, developing novel and more effective drugs for proteins associated with *Pseudomonas aeruginosa* and SARS-CoV-2 has become a top priority in recent years. In this regard, the protein is known as a putative inhibitor of vertebrate lysozyme [*Pseudomonas aeruginosa*] and chain A, spike protein S1 [SARS-CoV-2], and it is one of the key targets for the development of new drug candidates that could be used as inhibitors in both *Pseudomonas aeruginosa* and SARS-CoV-2 chemotherapies. The structural characteristics and binding mechanism of certain inhibitors of *Pseudomonas aeruginosa* and SARS-CoV-2 receptor were investigated using Quantitative-Structure Activity Relationship (QSAR), homology modeling, molecular docking, and molecular dynamics simulation in this study. To create a 3D model of the receptors, a homology modeling approach was used. The X-ray crystal structures of chain A inhibitor of vertebrate lysozyme (PDB: 4PS6) and chain A spike protein S1 (PDB: 7MZF), respectively, were employed as templates for this technique. The final projected structure is obtained and examined by the programs ERRAT, VERIFY3D, and PROCHECK, confirming that the final model is credible. The anticipated structure was fine-tuned before being employed in docking simulations. The results of the structure-based virtual screening show that two potent new lead molecules, compounds 7 and 15, have the most noteworthy affinity to the predicted 3D receptors. The docked compound 15 was subjected to a 1 ns molecular dynamics (MD) simulation. Compound 15 produced more hydrophobic and van der Waal's contacts, according to MD simulations, and binds to SARS-CoV-2 spike protein S1 more closely. The Lipinski rule of five assessment revealed that these lead compounds had significant pharmacological characteristics. As a result, the current research will aid in the development and synthesis of another class of chain A inhibitors of vertebrate lysozyme and chain A spike protein S1 inhibitors that restore drug compound susceptibilities.

Keywords: *Pseudomonas aeruginosa*; SARS-CoV-2; QSAR; homology modelling; Ramachandran plot; docking simulations; molecular dynamic simulations; ADME.

Resumen. La investigación de tratamientos innovadores para *Pseudomonas aeruginosa* y SARS-CoV-2 es un campo floreciente. A pesar de que los científicos y las compañías farmacéuticas han hecho contribuciones significativas a la investigación de infecciones por resistencia a múltiples medicamentos desde una variedad de perspectivas, estas enfermedades siguen siendo incurables. Como resultado, el desarrollo de fármacos novedosos y más eficaces para las proteínas asociadas con *Pseudomonas aeruginosa* y SARS-CoV-2 se ha

convertido en una prioridad en los últimos años. En este sentido, la proteína es conocida como un supuesto inhibidor de la lisozima de vertebrados [*Pseudomonas aeruginosa*] y de la cadena A, proteína espiga S1 [SARS-CoV-2], y es uno de los objetivos clave para el desarrollo de nuevos fármacos candidatos que podrían utilizarse como inhibidores tanto en *Pseudomonas aeruginosa* como en quimioterapias de SARS-CoV-2. Las características estructurales y el mecanismo de unión de ciertos inhibidores de *Pseudomonas aeruginosa* y el receptor SARS-CoV-2 se investigaron utilizando la Relaciones Cuantitativas de Estructura-Actividad - (QSAR, por sus siglas en inglés), el modelado por homología, el acoplamiento molecular y la simulación de dinámica molecular. Para crear un modelo 3D de los receptores, se utilizó un enfoque de modelado por homología. Las estructuras cristalinas de rayos X del inhibidor de la cadena A de la lisozima de vertebrados (PDB: 4PS6) y la proteína de espiga de cadena A S1 (PDB: 7MZF), respectivamente, se emplearon como plantillas para esta técnica. La estructura final proyectada se obtuvo y examinó con los programas ERRAT, VERIFY3D y PROCHECK, confirmando que el modelo final es creíble. La estructura anticipada se afinó antes de ser empleada en simulaciones de acoplamiento. Los resultados del cribado virtual basado en la estructura generaron dos nuevas y potentes moléculas líderes, los compuestos 7 y 15, que tienen la afinidad más notable con los receptores 3D predichos. El compuesto acoplado 15 fue sometido a una simulación de dinámica molecular (DM) de 1 ns. El compuesto 15 produjo más contactos hidrófobos y de van der Waals, según las simulaciones de MD, y se une más de cerca a la proteína espiga S1 del SARS-CoV-2. La evaluación de la regla de cinco de Lipinski reveló que estos compuestos líderes tenían características farmacológicas significativas. Como resultado, la investigación actual ayudará en el desarrollo y la síntesis de otra clase de inhibidores de la cadena A de la lisozima de vertebrados y los inhibidores de la proteína de espiga de cadena A S1 que restauran las susceptibilidades de los compuestos farmacológicos.

Palabras clave: *Pseudomonas aeruginosa*; SARS-CoV-2; QSAR; homología; gráficas de Ramachandran; acoplamiento molecular; simulaciones de dinámica molecular; ADME.

Abbreviations

ADME: Absorption, distribution, metabolism, excretion
CCC: Concordance Correlation Coefficient
DFT: Density functional theory
MAE: Mean Absolute Error in fitting
MVD: Molegro virtual docker
NAMD: NANOScale Molecular Dynamics program
NPT: number of constant pressure and temperature

NVT: Number of constant volume and temperature
PDB: Protein data bank
PRESS: Predictive Residual Sum of Squares
RMSD: Root mean square deviation
s: Standard error of estimate
SASA: Solvent-accessible surface area
VMD: Visual Molecular DynamicS

Introduction

Pseudomonas aeruginosa (*P. aeruginosa*) is an opportunistic bacterial nosocomial pathogen-multi-drug resistance pathogen that causes severe infections to cystic fibrosis (CF) patients [1], hospitalized, and immunocompromised patients [2]. These *P. aeruginosa* oftentimes colonize similar destinations, including the urinary lot, burn and surgical wounds, bacteremia, septicemia (blood poisoning), pneumonia, bronchitis, diarrhea, keratitis, skin and wound infections, and the upper respiratory tract [3,4]. *P. aeruginosa* is the most pervasive bacterial microbe that colonizes the lung, and their co-colonization is related to helpless patient outcomes, i.e., patients with cystic fibrosis [5]. The virulence nature of this pathogen in the cystic fibrosis patient's lung and severity of cystic fibrosis lung disease is closely linked to iron homeostasis [6]. The Centers for Disease Control and Prevention in 2019, has considered *P. aeruginosa* as a serious health threat (https://www.cdc.gov/drugresistance/biggest_threats.html). *P. aeruginosa* is related to an expanded bounty of

other bacterial and viral pathogens, for example, Covid-19 (SARS-CoV-2), HIV and AIDs, STIs, diabetes, etc. Lysozymes are the significant cell divider polymer in the microorganisms [7], it provides resistance against turgor and its cleavage by hydrolases [8]. Lysozyme is an old bactericidal protein that is important for the antibacterial protection arrangement of vertebrate and invertebrate creatures [9]. Several studies support a role for these lysozyme inhibitors in bacteria–host interactions [10]. SARS-CoV-2 is the most recent coronavirus to be discovered, and it mostly affects the respiratory tract (nose and lungs), causing colds, bronchitis, and pneumonia [11]. The connection of the viral spike protein and the host angiotensin-converting enzyme 2 (ACE2) causes an underlying passage and ensuing replication in the upper respiratory tract [12]. Due to higher amounts of ACE2 expression [13] and viral RNA (vRNA) at this region [14], the nasal cavity is believed to be the main site of viral replication rather than the oral cavity [15]. Speedy and uncontrolled propagation can bring about contamination of the lower respiratory tract and serious ailments. This virus is very contagious and has a high fatality rate [16]. It takes 2-14 days to incubate SARS-CoV-2 after exposure [17]. Dry cough, fever, shortness of breath, loss of taste or smell, skin rashes, conjunctivitis, exhaustion, muscle soreness, diarrhea, and fatigue are the first symptoms of SARS-CoV-2 infection, which are joined by other biological indicators such as lymphopenia [18]. The World Health Organization (WHO) listed severe acute respiratory syndrome-related (SARS) out of the nine (9) as one of the highly pathogenic viruses likely to cause epidemics. Under 90 days later its underlying rise in Wuhan, China WHO proclaimed serious intense respiratory disorder related to Covid-19 (SARS-CoV-2) a pandemic [19] and at present no drugs are available to address this dangerous pathogen [20]. Older and an individual with prior ailments or comorbidities (cardiovascular sickness, diabetes, lungs infection, gram-positive or negative bacterial, and cancer) are bound to foster serious ailments. The spike glycoprotein creates a homotrimer that allows the virus to enter host cells. The spike glycoprotein is of tremendous relevance in the development of SARS-CoV-2 treatment techniques since it is the key component that the virus employs to bind to receptors on the host cell surface. If humans are infected with SARS-CoV-2, the spike protein could be used as a target for therapeutic and vaccine development [21,22]. *P. aeruginosa* and SARS-CoV-2 are dangerous diseases considering their negative consequences for human health and their ability to transmit from one person to another [23]. To avoid these infection relapses, new anti- *P. aeruginosa* and anti-SARS-CoV-2 drugs are needed.

Iminoguanidine is a key chemical class in drug discovery owing to its broad range of therapeutic potential, which include, antibacterial [24-26], antiviral [25,26], ability to recognize and separate anions from competitive aqueous environments [27], and biomass-based direct air capture [28]. This beneficial property of iminoguanidine sparks our interest in creating a tool for identifying novel iminoguanidine analogs that have potential as antibacterial agents.

There is an immediate need for ways to combat the disease's increased prevalence, and counteraction seems a legitimate methodology. This computer-aided drug design (CADD) approach can contribute to many stages of the drug discovery process, notably to performing a search for active compounds by virtual screening [29]. Computational methods, i.e., computer-aided drug design (CADD) has been successfully applied in pharmacy, pharmaceutical chemistry, molecular biology, and biochemical researcher for years to fight the problems related to drug design, drug discovery, and drug development [30]. The current work examines the use of grounded quantitative structure-activity relationship (QSAR) modeling in conjunction with genetic approximation-multiple linear regression (GA-MLR), docking simulations, molecular dynamics simulations (MDs), and ADME as relevant to iminoguanidine derivatives to investigate their role as a negative catalyst against *P. aeruginosa* and SARS-CoV-2. Accordingly, this research may be practicable for the design and development of persisting multi-drug resistance pathogen activity.

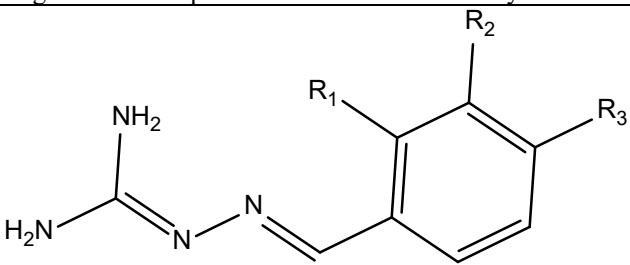
Experimental

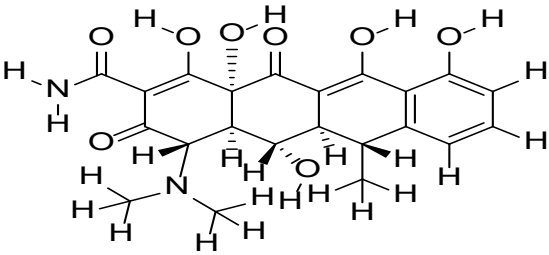
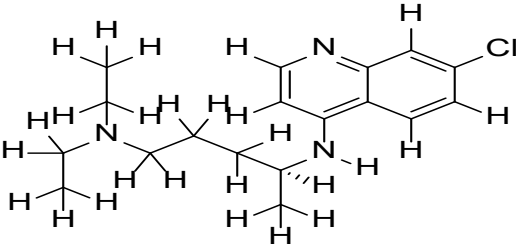
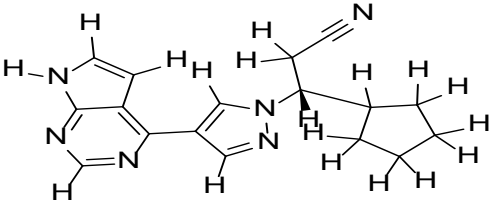
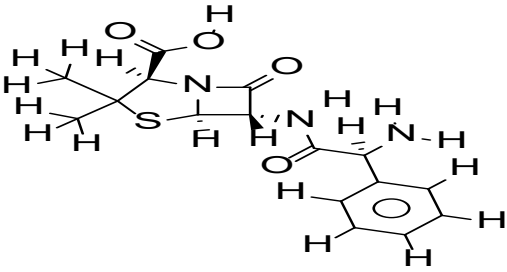
Methods

A total number of 25 iminoguanidine derivatives reported by PubChem with the accession number AID_1315713 were collected to perform this study. The MIC₅₀ values (minimum inhibitory concentrations) were converted to pMIC₅₀ (-logMIC₅₀) values that were utilized in the QSAR investigation as dependent

variables. The compound's pMIC₅₀ values are within the QSAR prerequisite range according to Shirvania and Fassihi [31]. The 3D structures of all the compounds including the reference compounds were prepared using MarvinView Eurpium.6 (<http://www.chemaxon.com>) and their energies were minimized with the Spartan'14 V1.1.4 PM6. The energy minimized molecules were subjected to optimization via DFT (density functional theory) method with B3LYP function [32,33] and 6-31G+ basic set [34] with the GAUSSIAN 09 package [35]. The optimized structures were imported into PaDEL-Descriptors v2.20 [36], which calculated the molecular descriptors for each molecule. The calculated descriptors from the PaDEL software saved in an excel spreadsheet were pre-treated using NEW_V-WSP version 1.2 (DTC) laboratory software to sequester highly correlated descriptors from the distantly related ones, thus reducing redundancy in the calculated values. The pre-treated descriptors saved in the text file were transferred into the QSARINS software for further analysis [37]. The chemical structure of the molecules and their pMIC₅₀ values are available in Table 1. The QSARIN v2.2.4 software was adopted to perform a Genetic algorithm (GA) and model development. GA approach allows the proper selection of some exceptionally good descriptors to improve a model's quality and predictive accuracy and by doing so, a multiple linear regression (MLR) is curated to obtain the additive relationship between biological endpoints and molecular descriptors [38]. Thus, the technique is referred to as GA-MLR. Before developing the 2D-QSAR model, the data set were divided into 70 % training and 30 % test by applying the random selection implemented in QSARINS software.

Table 1. Chemical structures of iminoguanidine compounds as well as their activity levels.

Compound No.					
	R1	R2	R3	MIC ₅₀ (ug/ml)	pMIC ₅₀
1	H	H	Cl	10.7	1.0294
2	H	H	N(CH ₃) ₂	32.5	1.5119
3	OH	H	H	58.5	1.7672
4	OCH ₃	H	H	38.4	1.5843
5	H	H	NO ₂	61.4	1.7882
6	F	H	H	28.6	1.4564
7	H	Naph ^c	Naph ^c	31.7	1.5011
8	H	Cl	H	19.2	1.2833
9	Cl	H	H	16.3	1.2122
10	H	H	CH(CH ₃) ₂	20.6	1.3139
11	H	H	OCH ₃	30.5	1.4843
12	Naph ^c	Naph ^c	H	30.4	1.4829
13	H	H	F	87.4	1.9415
14	H	H	Br	19.7	1.2945
15	H	OCH ₂ C ₆ H ₆	H	25.5	1.4065
16	Br	H	H	62.9	1.7987
17	OH	Br	H	145	2.1614
18	H	H	C ₆ H ₆	34.6	1.5391
19	H	OH, OH ^a	H	338	2.5289
20	H	H	OH	443	2.6464
21	H	OH	H	391	2.5922

22	H	OCH ₃	H	48.6	1.6866
23	H	F	H	70	1.8451
24	H	Br	H	23	1.3617
25	H	Cl	Cl	24.1	1.3820
Reference compounds					
Doxycycline (Vibramycin)					
Chloroquine					
Ruxolitnib					
Ampicillin					

2D-QSAR Model validation

Internal and external validations of the QSAR model are carried out by using the model to forecast training and test data. To assess the model's validity, we calculate several parameters. To begin, the coefficients

of determination (R^2) are computed to assess the strength of the relationship between actual and predicted values. The slope value (k) and $(r^2 - r_0^2/r^2)$ are also used to assess the model's performance. The correlation coefficients between actual and expected values with and without intercept are represented by the r_0^2 and r^2 parameters, respectively. Ultimately, we compute the modified r^2 (r_m^2) to show the model's predictive performance [39,40]. The following equations can be used to calculate the validation parameters:

$$R^2 = 1 - \frac{\sum(Y - \hat{Y})^2}{\sum(Y - \bar{Y})^2} \quad \text{Eq. 1}$$

$$r^2 = \frac{[\sum(Y - \bar{Y})(\hat{Y} - \bar{\hat{Y}})]^2}{\sum(Y - \bar{Y})^2 \times \sum(\hat{Y} - \bar{\hat{Y}})^2} \quad \text{Eq. 2}$$

$$r_0^2 = 1 - \frac{\sum(Y - K \times \hat{Y})^2}{\sum(Y - \bar{Y})^2} \quad \text{Eq. 3}$$

$$K = \frac{\sum(Y \times \hat{Y})}{\sum \hat{Y}^2} \quad \text{Eq. 4}$$

$$r_m^2 = r^2 \times (1 - \sqrt{r^2 - r_0^2}) \quad \text{Eq. 5}$$

Where, Y and \hat{Y} denote the actual and predicted value of pMIC50, respectively. The threshold values for all the parameters used in these studies were detailed according to Edache et al [41].

Protein generation and preparation

Pseudomonas aeruginosa and the 2019 novel coronavirus protein sequence were downloaded from the NCBI database (<https://www.ncbi.nlm.nih.gov/>) with accession number KJJ19201 and PDB: 7E23, respectively. A BLAST exploration was carried out to find appropriate templates. The UPGMA algorithm [42] was used to infer the evolutionary history. The evolutionary distances were calculated using the Poisson correction method [43] and are measured in amino acid substitutions per site. Each descending clade's proportion of sites with at least one unambiguous base in at least one sequence is shown next to each internal node in the tree. This study looked at five and six different amino acid sequences of *P. aeruginosa* and SARS-CoV-2 spike proteins, respectively. For each sequence pair, all unclear locations were deleted (pairwise deletion option). In total, there were 215 and 239 positions in the final dataset, respectively. MEGA X was used to undertake evolutionary analyses [44]. Pair-wise alignment of the template and query sequences were constructed using MODELLER v10.1 to check the results of the phylogenetic tree analysis [45-47]. For query and template sequences, BioEdit [48] was used to build a sequence alignment file in PIR format. Python Scripts were used to construct the Multiple Template Modeling (MTM) protocol from module 5 structures. With the help of Swiss-Pdb Viewer [49], for energy minimization, the model with the most reduced DOPE score (Discrete Optimized Protein Energy) was chosen. For *P. aeruginosa* and SARS-CoV-2 spike, the template and the predicted modelled structures were superimposed with RMSD values of 0.176 and 0.261, respectively (Fig. 1).

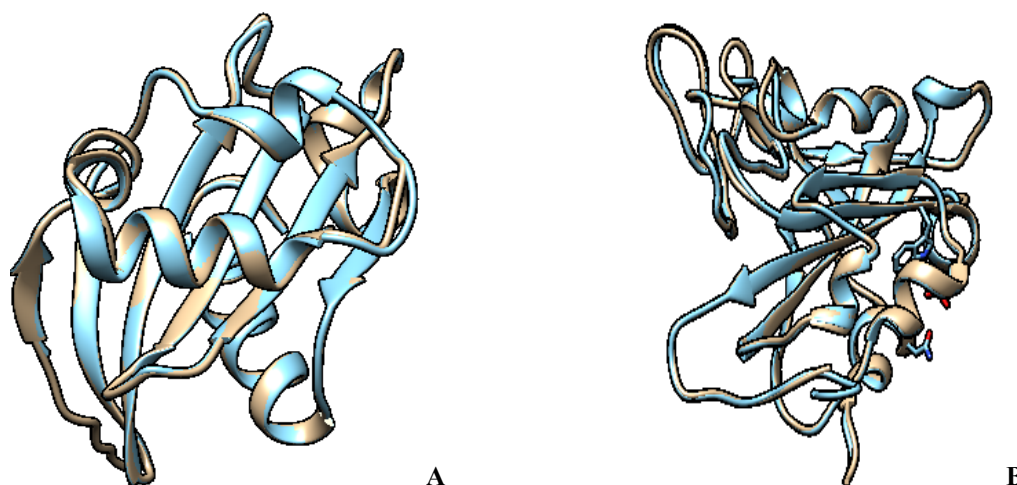


Fig. 1. The modelled protein structures of (A) *P. aeruginosa* and (B) SARS-CoV-2 spike protein from Modeller v10.1 using multiple template modeling.

Equipment and programming

Molegro Virtual Docker 6.0 2013 [50] and Autodock Vina with PyRx v0.8 [51] were used for docking simulations with core i3 windows processors. The molecular docking simulations of all compounds were performed with the predicted modelled crystal structure of *P. aeruginosa* and SARS-CoV-2. Molegro Virtual Docker (MVD) which was considered as one of the best and accurate docking packages were used to study the interactions energies between the stationary modelled proteins and the ligands. The design of the demonstrated proteins was created by pressing icons, and the amino acid flaws that presented as a warning were repaired. The protein surface and binding cavity have both been identified. After resetting the icon, the design of the demonstrated proteins is ready for docking simulation. The docking wizard was used to perform interactions of the protein-ligand and the parameters for the docking interactions were set as ten (10) runs, 50 population size, 1500 iterations, and lastly, the maximum number of poses is 4. Next was the PyRx [51] docking simulations setup. The modelled proteins were prepared using PyRx v0.8 software and the docking experiments were subsequently accomplished using the AutoDock Vina package [52]. The network place for docking of the ligands against the displayed protein of *P. aeruginosa* was set as exhaustiveness = 8, center $x = 16.8669$, center $y = 53.3921$, center $z = -4.3156$, size $x = 55.3409311867$, size $y = 155.016245728$, size $z = 141.969700775$, and against the modelled SARS-CoV-2 spike protein was set as center ($x = -1.5643$, $y = 12.818$, and $z = 12.3015$), size ($x = 49.2203297997$, $y = 44.6869613934$, and $z = 51.2521103764$), respectively. The Discovery studio 2020 client was used to visualize the outcomes of the docking simulation.

Molecular Dynamics Simulations

Utilizing the CHARMM-GUI web server [53], a cubic box of TIP3 water particles with an edge distance of 10 Å was utilized to submerge the docked complex. A 0.15M KCl counterion was utilized to neutralize the system. To deliver the inner strain energies of the whole system, the whole framework or system was exposed to 5,000 run steps of energy minimization utilizing the steepest drop integrator. The NVT equilibration was then performed at 303.15 K for 0.01 ns. 500,000 runs i.e., 1 ns of unrestrained isothermal isobar group (NPT) were performed for the production and were completed at 303.15 K, utilizing 1 atm (1.01325 bar) with 2 fs time steps. All computations were performed using the CHARMM36m force field [54], NAMD [55], and VMD [56] for representation.

Results and discussion

The QSAR study was carried out to explain the multiple linear relationships between the 2D descriptors and the bioactivity. To explain the diversity in the bioactivity of iminoguanidine derivatives, the molecular descriptors (Centered Broto-Moreau autocorrelation - lag 1 / weighted by I-state (ATSC1s) and Sum of E-State descriptors of strength for potential hydrogen bonds of path length 7 (SHBint7) were chosen. The GA-MLR equation for the prediction of the bioactivity is given in equation 6:

$$pMIC_{50} = 0.4444 - 0.147ATSC1s - 0.0332SHBint7 \quad \text{Eq. 6}$$

The following are the mathematical parameters/ values of the 2D QSAR model:

$$R^2 = 0.8445; R_{adj}^2 = 0.8238; R^2 - R_{adj}^2 = 0.0207; LOF = 0.0404; RMSE_{train} = 0.1564; MAE_{train} = 0.1268; RSS_{train} = 0.4402; CCC_{train} = 0.9157; S = 0.1713; F = 40.7393$$

$$Q_{LOO}^2 = 0.7652; R^2 - Q_{LOO}^2 = 0.0793; RMSE_{cv} = 0.1922; MAE_{cv} = 0.1558; PRESS_{cv} = 0.6647; CCC_{cv} = 0.8685; R_{Yscr}^2 = 0.1169; Q_{Yscr}^2 = -0.3234; R_{Yrnd}^2 = 0.1180; Q_{Yrnd}^2 = -0.3188$$

$$RMSE_{test} = 0.4930; MAE_{test} = 0.3954; PRESS_{test} = 1.4585; R_{test}^2 = 0.7763; Q^2 - F_1 = -0.1334; Q^2 - F_2 = -0.9302; Q^2 - F_3 = -0.5455; CCC_{test} = 0.6663; r_m^2 \text{ delta} = 0.6164.$$

The developed model represents a linear combination of two descriptors. One of them belongs to autocorrelation descriptor (ATSC1s) and electro-topological state atom type descriptor (SHBint7). The model in (equation 6) indicates that the inhibitory activity of compounds against the strain of *P. aeruginosa* depends on autocorrelation and topological state properties. Table 2 shows the p-values, variance inflation factor (VIF), and the mean effect (MF) for the variables used in the model proposed.

Table 2. The validation parameters, variance inflation factor, and mean effect.

Variable	Coeff.	Std. coeff.	Std. err.	(+/-) Co. int. 95%	p-value	VIF	MF
Intercept	0.4444		0.1373	0.2926	0.0052		
ATSC1s	-0.1470	-1.0263	0.0163	0.0347	0.0000	1.25	1.068
SHBint7	-0.0332	-0.4710	0.0080	0.0171	0.0008	1.25	-0.068

Internal and external validation was used to estimate the significance and to ensure the validity of the models [57]. The model disclosed a cross-validation coefficient (Q_{LOO}^2) value of 0.7652 and the difference between R^2 and Q_{LOO}^2 is very small (less than 0.01) that indicates the consistency and stability of the model. The mean absolute error (MAE) measurement of the model accesses the degree of the prediction error. For the validation set, an MAE_{cv} of 0.16, and PRESS of 0.66 is obtained signifying an improved prediction. The certainty of the model was asserted by a test set, which gave a determination correction coefficient (R_{test}^2) value of 0.7763. The predicted values of the bioactivity of the training set and test set of the model are introduced in Table 3. Table 3 constitutes the computed values of the 2 descriptors.

Table 3. Observed, predicted biological activity, and HAT of the training and predicted set of compounds against *P. aeruginosa*.

ID	Name	Status	Observed act.	Pred. by model eq.	HAT i/i (h*=0.5000)
1	2	Prediction	1.5119	1.3261	0.1013
2	3	Training	1.7672	1.7977	0.5221
3	4	Training	1.5843	1.2615	0.1377

4	5	Training	1.7882	2.012	0.1304
5	6	Training	1.4564	1.5377	0.4325
6	7	Training	1.5011	1.4094	0.0839
7	8	Training	1.2833	1.4373	0.0793
8	9	Training	1.2122	1.3869	0.081
9	10	Training	1.3139	1.2556	0.1206
10	11	Training	1.4843	1.4962	0.0718
11	12	Training	1.4829	1.3931	0.0868
12	13	Prediction	1.9415	2.1189	0.1704
13	14	Training	1.2945	1.4085	0.084
14	15	Training	1.4065	1.4088	0.0839
15	16	Prediction	1.7987	1.3922	0.087
16	17	Prediction	2.1614	1.7301	0.4965
17	18	Prediction	1.5391	1.3749	0.0904
18	19	Prediction	2.5289	3.5361	1.6041
19	20	Training	2.6464	2.4252	0.3379
20	21	Training	2.5922	2.4252	0.3379
21	22	Training	1.6866	1.4962	0.0718
22	23	Training	1.8451	2.1189	0.1704
23	24	Training	1.3617	1.4085	0.084
24	25	Training	1.382	1.4102	0.0837

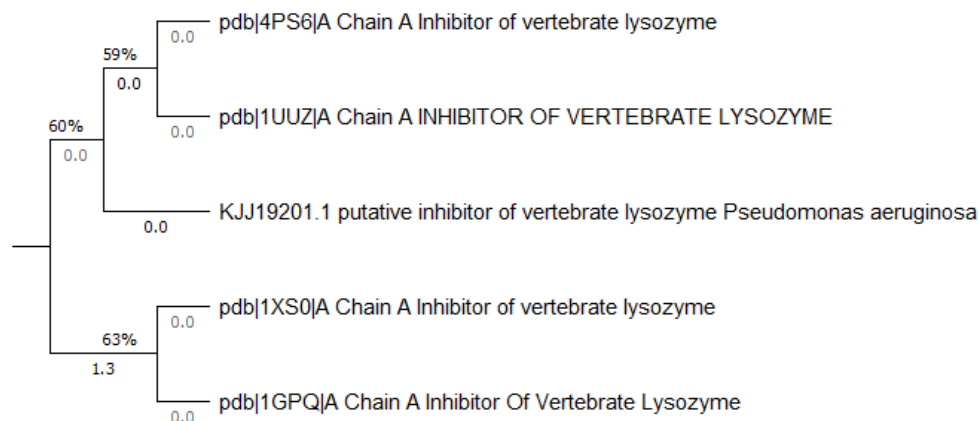
To additionally affirm whether or not there is a multicollinearity issue in the revealed model as remarked by Bolboac and Lorentz [58], the VIF values of these descriptors in the preparation (training) set were registered and gotten to be under 5 demonstrating the wellness of the announced model. The descriptors in the revealed model were measurably huge and the symmetrical idea of the descriptors was affirmed, i.e., the model is great, liberated from multicollinearity issue, and satisfactory since no descriptor has its VIF values more noteworthy than 5 (If a model has a descriptor with VIF esteem more prominent than 5, the model isn't acknowledged, thusly, should be re-evaluated) [59]. The importance and commitment of a descriptor concerning other descriptors in the model were determined by the mean effect (MF) values. The positive sign of Centered Broto-Moreau autocorrelation - lag 1 / weighted by I-state (ATSC1s), indicates that the molecule of the compound should have larger electronic properties (polarizability and charges as weights) in the molecule will increase the biological activity. This descriptor showed the highest significance than the other descriptors according to their p-value. The negative sign of the Sum of E-State descriptors of strength for potential hydrogen bonds of path length 7 (SHBint7) that account for the intrinsic electronic state of the atom as perturbed by the electronic influence of all other atoms in the molecule, shows that, the lesser values of these descriptor attributed to the compounds with greater bioactivity values.

The fitting criteria R^2 , Q_{LOO}^2 , R_{test}^2 , CCC and relatively low values of the errors represented by: LOF , $RMSE_{train}$, MAE_{train} , RSS_{train} , $RMSE_{cv}$, MAE_{cv} , $PRESS_{cv}$, $RMSE_{test}$, and MAE_{test} show that there is a strong linkage between the observed bioactivity with the predicted by the built model as shown in Fig. S1 of Supplementary material, and it also confirmed the robustness, goodness of fit and predictive ability of the model. The model predicted is correct not only for the training set, but also for the test set meanwhile the error values ($RMSE_{train}$, MAE_{train} , RSS_{train} , $RMSE_{cv}$, MAE_{cv} , $PRESS_{cv}$, $RMSE_{test}$, and MAE_{test}) are at the same level and there are no significant large residual values. The observed activity is plotted against the residuals in Fig. S2 since the scatter plot were all scattered within the baseline shows that the model is free from systematic

error. To assess the resilience of the derived models, we used Y-randomization [38, 57, 59] to confirm that the correlations between the bioactivity and descriptors are not by chance. The average value of the Y-randomization test is $R_{Yscr}^2 = 0.1169$; $Q_{Yscr}^2 = -0.3234$; $R_{Yrnd}^2 = 0.1180$; and $Q_{Yrnd}^2 = -0.3188$. The outcomes of the Y-scrambling for the repetitions are shown in Fig. S3. It was found that all randomization or scrambling values of the random models are lesser than the innovative ones and the intercepts are inside the acceptable values recommended in the literature [60]. This result shows that the obtained model is not by chance. The applicability domains (AD) are crucial because they are specifically sought in the OECD's validation processes [61]. The AD of this model was assessed by a leverage analysis which is based on the standardized residuals' variation. The standard deviation (δ) and the leverage value (h^*) are set to ± 2.5 and 0.5, respectively. We find the presence of outliers as shown in Fig. S4. The response outlier is a chemical structure that is below or above the standard deviation of ± 2.5 . While the structurally influential outlier is a chemical structure that is outside the leverage value. Compound 17 was found to have a standardized residual more than the standard deviation of ± 2.5 is a response outlier. The compounds number 3 of the training set with its leverage value reasonably higher than the threshold value ($h^* = 0.5$) and compound number 19 of the test set has its h^* far greater than the limit value of 0.5 are structurally influential outlier (Fig. S4). Insubria plot [62] was used to evaluate the position of the molecules lacking experimental response. The Insubria plot (Fig. S5) is similar to the Williams plot (Fig. S4). The Insubria plot is also confirmed that compounds 3 and 19 are a structural influential outlier. As a result of our 2D-QSAR model, we can create novel compounds with higher anti-P aeruginosa activity.

Phylogenetic tree and Homology modeling

The sequence of *P. aeruginosa* and SARS-CoV-2 spike glycoproteins were obtained from the NCBI (National Center for Biotechnology Information) database (<http://www.ncbi.nlm.nih.gov/>) with accession number KJJ19201.1 and PDB 7E23, respectively, were used to BLASTp search against the protein data bank to select the template structures. The query sequence and the BLAST template structures were first analyzed through the MEGA-X [63] software to investigate the phylogenetic relationship based on the test UPGMA (Unweighted pair-group method with arithmetic) method as shown in Fig. 2. The phylogenetic tree which is obtained by MUSCLE [64] shows the evolutionary relationship between the proteins sequences as it was a rooted tree. Phylogenetic analysis demonstrated clear sequestration of *P. aeruginosa* and SARS-CoV-2 in two different biological groups and these biological groups were also divided into subgroups. All the biological groups are grouped according to their protein sequence similarity. The *P. aeruginosa* query sequence (accession number KJJ19201.1) shows high similarity to PDB 4PS6 and 1UUZ, respectively as shown in Fig. 2(A). The phylogenetic sequence analysis of the spike protein of SARS-CoV-2 (PDB: 7E23) and the other five templates shows a close similarity (Fig. 2(B)). Concerning overall sequence similarity PDB 7MZF spike protein S1 has the highest similarity to PDB 7E23. The chain A inhibitors of vertebrate lysozyme (PDB 4PS6 and PDB 1UUZ) and spike protein S1 (PDB 7MZF) have very close relationships to the query sequence, respectively.



A

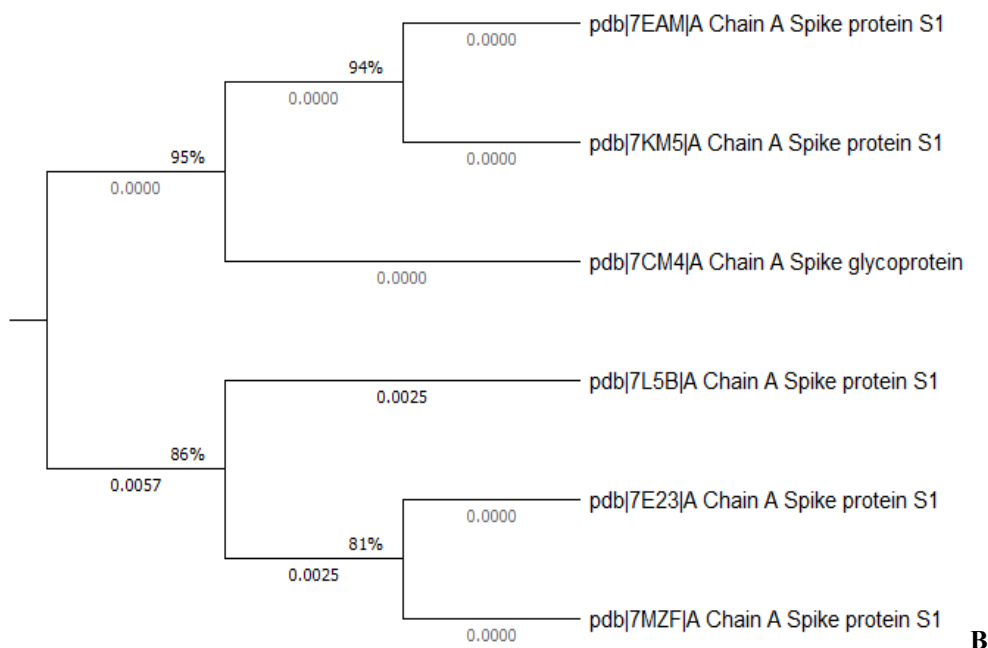
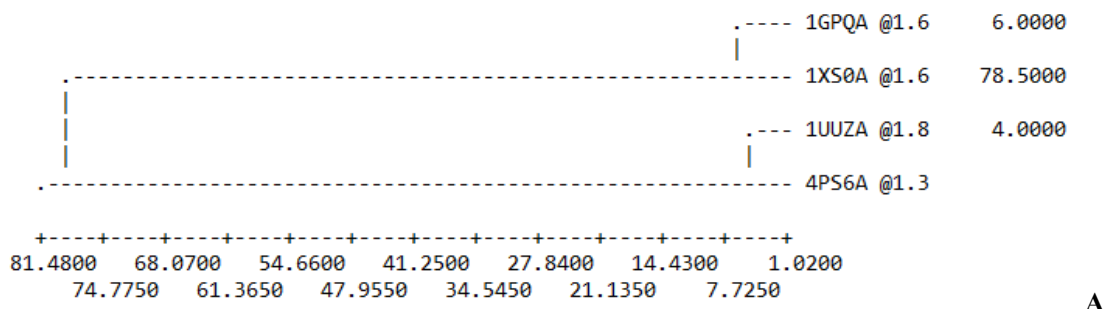


Fig. 2. Phylogenetic analysis of **(A)** *P. aeruginosa* and **(B)** SARS-CoV-2 spike protein homologs were constructed by the UPGMA method using the MEGA X program.

Further analysis using MODELLER 10.1 was performed on the query sequence and the template sequences structures to confirm the phylogenetic tree analysis and a better insight into their protein structure similarity. According to the result `malign3D` command implemented in MODELLER 10.1, the tree diagram calculates a weight pair group average clustering on a distance matrix (Fig. 3). Out of the template structure, the best homologous structure belonged to an inhibitor of vertebrate lysozyme (PDB: 4PS6) (Fig. 3(A)) and spike protein S1 (PDB: 7mzf) (Fig. 3(B)) because of its similarity and high resolution with the query structure was selected for homology modeling. The results obtained thereof confound to the phylogenetic analysis. The selected PDB was used as a template structure for the construction of the *P. aeruginosa* and SARS-CoV-2 spike protein, respectively. *Pseudomonas aeruginosa*, with the accession number of KJJ19201.1, has 199 amino acids in its protein sequence, while SARS-CoV-2 spike protein S1 with PDB code 7E23 has 196. The *Pseudomonas aeruginosa* (PDB 4PS6) and SARS-CoV-2 spike protein S1 (PDB code 7MZF) were selected as the template based on the phylogenetic tree analysis and dendrogram command using the `align2d` script implemented in MODELLER. An alignment of *Pseudomonas aeruginosa* (accession number of KJJ19201.1) and SARS-CoV-2 spike protein S1 with PDB code 7E23 with the template sequences are presented in Fig. 4. The MODELLER constructed five (5) different structure modelled proteins. The receptor generated by MODELLER was ranked and scored using discrete optimized protein energy (DOPE) score (Data presented in Table 3 and 4). The top one out of 5 models with the lowest DOPE scores were selected and assessed using ERRAT, VERIFY3D, and Ramachandran plots.

We perform energy minimization to the assessed receptors to remove side chains and later use it in the molecular docking and MD simulations.

Weighted pair-group average clustering based on a distance matrix:



Weighted pair-group average clustering based on a distance matrix:

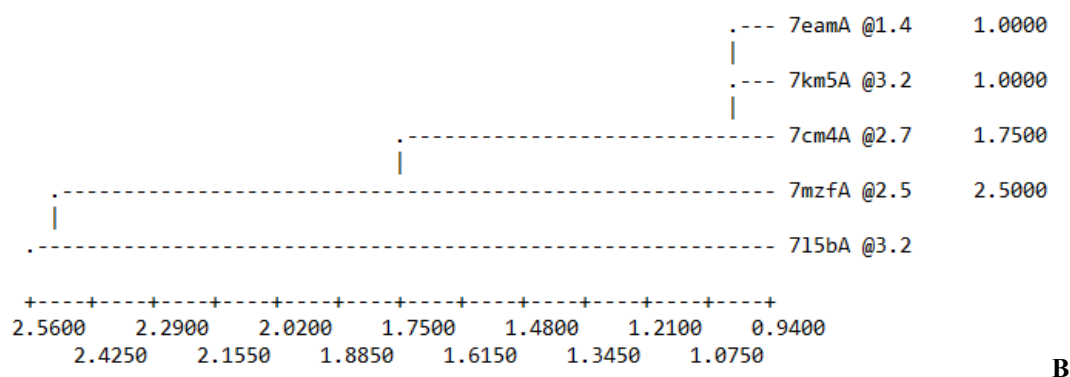


Fig. 3. The weighted pair-group matrix of **(A)** *P. aeruginosa* and **(B)** SARS-CoV-2 spike protein homologs were constructed by the modeller program.

```

aln.pos      10   20   30   40   50   60
4PS6A  -----G
qseq
MPVIRRMELPVDSSELVAVVPFQAASGGHRRIDCFHFTRSKEITMNGVSRLLSLALLGAALHWAP
AQA
_consrvd

aln.p 70   80   90   100  110  120  130
4PS6A
EEQPRLFELLGQPGYKATWHAMFKGESDVPKWVSDASGPSSPSTLSLEGQPYVLANSCKPHDCG
NNR
qseq
EEQPRLFELLGQPGYKATWHAMFKGESDVPKWVSDASGPSSPSTLSLEGQPYVLANSCKPHDCG
NNR
_consrvd *****

aln.pos 140  150  160  170  180  190
4PS6A
LLVAFRGDKSAAAYGLQVSLPDEPAEVMQTPSKYATYRWYGEPSRQVRELLMKQLES DPNWK
    
```

```

qseq
LLVAFRGDKSAAAYGLQVSLPDEPAEVMQTPSKYATYRWYGEPSRQVRELLMKQLESDPNWK
_consrvd *****

A
aln.pos    10    20    30    40    50    60
7mzfA
NLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSP TKLNDLCFTNVYAD
SF
qseq
TNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSP TKLNDLCFTNVYA
DSF
_consrvd *****

aln.p 70    80    90    100   110   120   130
7mzfA
VIRGDEV RQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYR LFRKSNLKPFE
RDI
qseq
VIRGDEV RQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYR LFRKSNLKPFE
RDI
_consrvd *****

aln.pos 140   150   160   170   180   190
7mzfA
STEIQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPGS
qseq  STEIQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCG---
_consrvd *****

```

B

Fig. 4. Pairwise sequence alignment between the template and the query (A) and putative inhibitor of vertebrate lysozyme [*Pseudomonas aeruginosa*], (B) chain A, Spike protein S1 [severe acute respiratory syndrome coronavirus 2].

Model number two (PA-model02.pdb) and model number one (CoV-model01.pdb) having lower DOPE scores as presented in Table 4 and Table 5, respectively were considered as the best model of *P. aeruginosa* and SARS-CoV-2.

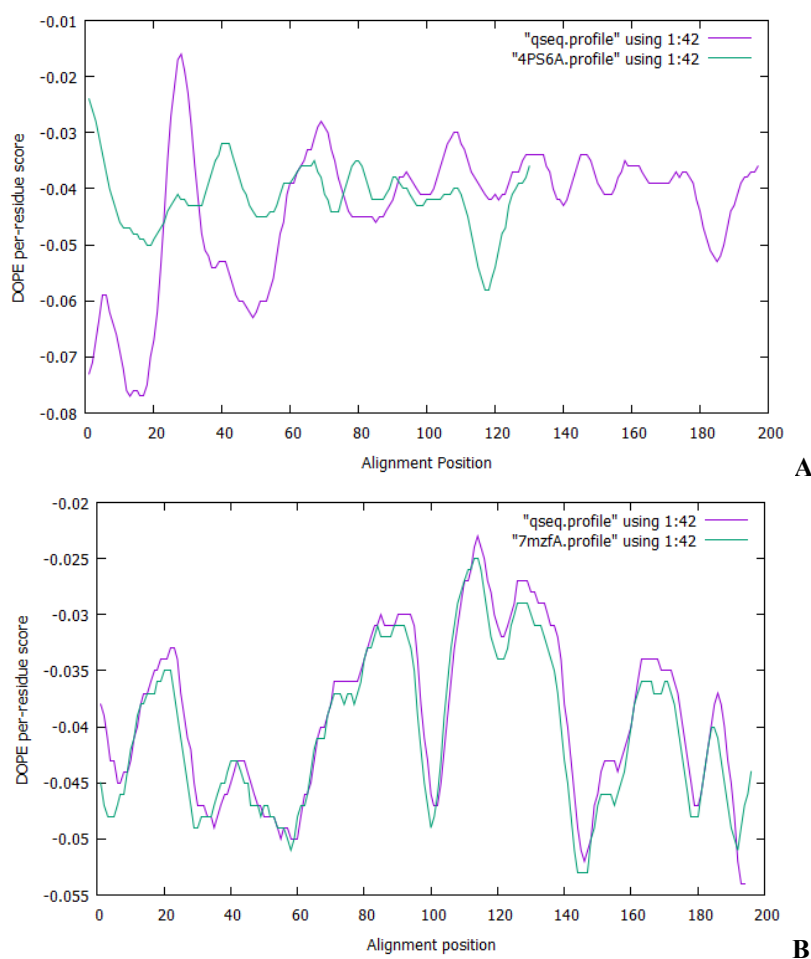
Table 4. Comparative analysis of five models' DOPE scores predicted by modeller.

Filename	molpdf	DOPE score	GA341 score
PA-model01.pdb	946.82159	-15805.88770	1.00000
PA-model02.pdb	886.59521	-15917.49902	1.00000
PA-model03.pdb	877.72742	-15809.07910	1.00000
PA-model04.pdb	931.51996	-15808.23438	1.00000
PA-model05.pdb	914.30719	-15800.19141	1.00000

Table 5. A study comparing the DOPE scores of five models predicted by the modeller.

Filename	molpdf	DOPE score	GA341 score
CoV-model01.pdb	1122.96790	22118.33008	1.00000
CoV-model 02.pdb	1048.29150	-21991.14258	1.00000
CoV-model 03.pdb	978.01141	-22080.57617	1.00000
CoV-model 04.pdb	1034.12805	-22053.22070	1.00000
CoV-model 05.pdb	1049.48303	-22097.23047	1.00000

The crystal structures of *Pseudomonas aeruginosa* and SARS-CoV-2 were predicted, polished, and validated. The final predicted structure of *P. aeruginosa* protein has a modeller objective (molpdf) score of 886.59521, genetic algorithm 342 (GA341) score of 1.0000, and the discrete optimized protein energy (DOPE) score of -15917.49902 was picked out. The modelled crystal structures SARS-CoV-2 spike protein S1 has a molpdf score of 1122.96790, GA341 score of 1.0000, and a DOPE score of -22118.33008 was also picked out. Following the sequence alignment, the DOPE score profile was generated using MODELLER's "assess dope" function to give an insight into the quality of input alignment (Fig. 5). Fig. 5 shows that the DOPE score profile of the modelled (query) and template proteins meet over a wide range of residues, and the results suggest no problematic locations.

**Fig. 5.** The DOPE score visibility of (A) *P. aeruginosa* and (B) SARS-CoV-2 spike protein incurred from Modeller software.

Validity of the structures analyzed for the *P. aeruginosa* receptor by the Ramachandran plot shows 94 % residues residing in most favored regions, 4.8 % in the additional allowed region, and 0.6 % in the generously allowed region. Only one residue was settled in the disallowed region, which constituted 0.6 % of the total protein, which constituted a good quality of the predicted model (Fig. 6). Other tests, such as Verify 3D and ERRAT also known as “overall quality factor”, were used to ensure the quality of the revised models. For non-bonded atomic interactions, the overall quality factor with larger scores indicates higher quality [65]. In general, a range of greater than 50 (> 50) indicates a high-quality model [66]. In this research, the ERRAT score for predicted modelled *P. aeruginosa* protein has an overall quality factor of 83.582 % (Fig. 7). Therefore, this confirmed that modelled putative inhibitor of vertebrate lysozyme *pseudomonas aeruginosa* protein has a high resolution and quality protein models. Verify-3D score profile access the model's quality. The Verify 3D profile of the modelled protein is shown in Fig. 8, and residues with an averaged 3D-1D score larger than zero are regarded as reliable. Some of the residues in the modelled protein have a computability score that is greater than zero. It was discovered that the structurally and functionally significant residues in the modelled potential inhibitor of vertebrate lysozyme *pseudomonas aeruginosa* protein had a score ranging from 0.30-0.60, indicating that the model's quality is comparable to that of high-resolution crystal structures.

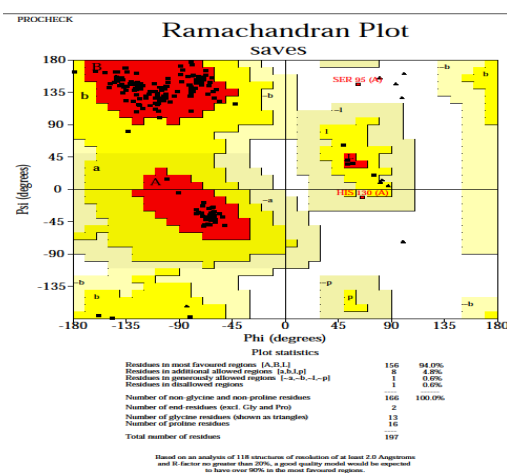


Fig. 6. Ramachandran plot calculation for 3D model of *P. aeruginosa*.

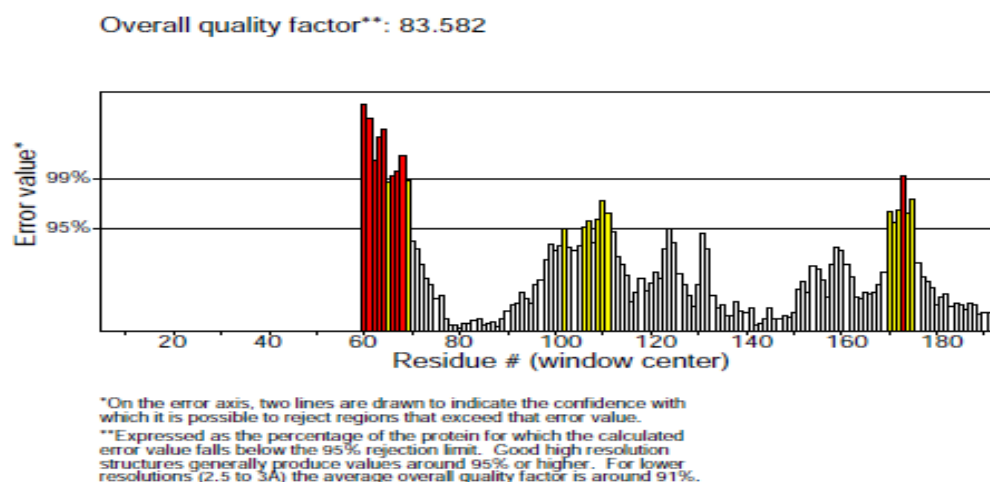


Fig. 7. ERRAT gave a quality factor of 83.582 % to the modelled *P. aeruginosa* protein structure.

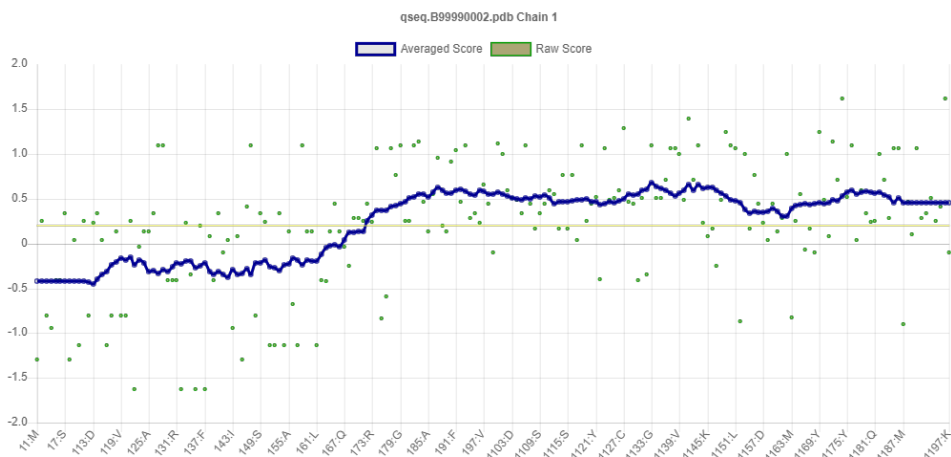


Fig. 8. Verify 3D score diagram validating the modelled *P. aeruginosa* protein.

Ramachandran plot analysis was carried out on the modelled SARS-CoV-2 spike protein showed 91.1 % residues residing in highly favored regions, 8.9 % in additional allowed regions, 0.0 % residues in generously allowed and disallowed regions, respectively (Fig. 9). This also validates that the modelled crystal structure is a good quality model. The quantity of the structural error at each amino acid residue in the 3D structural model was given by the ERRAT plot. The overall quality factor of the computed model was 79.032 % (Fig. 10). Also, the Verify3D plot of the SARS-CoV-2 modelled spike protein (Fig. 11) was obtained for the structure validation and it showed as PASS. Verify-3D ratings near 1.0 correspond to scores that would be expected for a legitimate protein of similar size. It was also discovered that more than 90 % of the residues had a score of more than 0.2, implying that more than 90 % of the residues were complemented by the 1D-3D model, implying that the predicted model's quality is adequate.

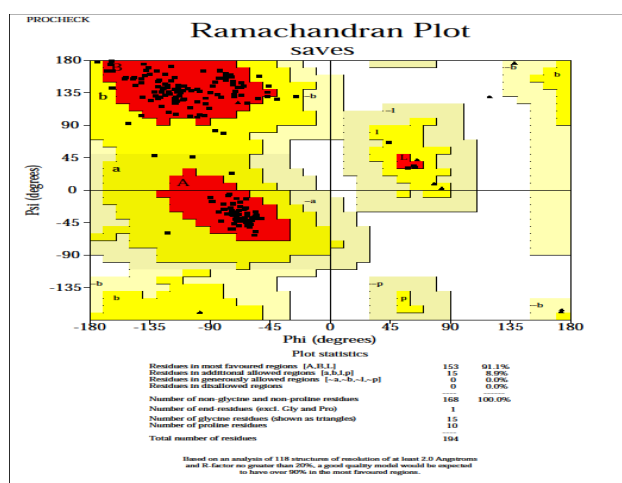


Fig. 9. Ramachandran plot computation for a 3D model of SARS-CoV-2 spike protein S1.

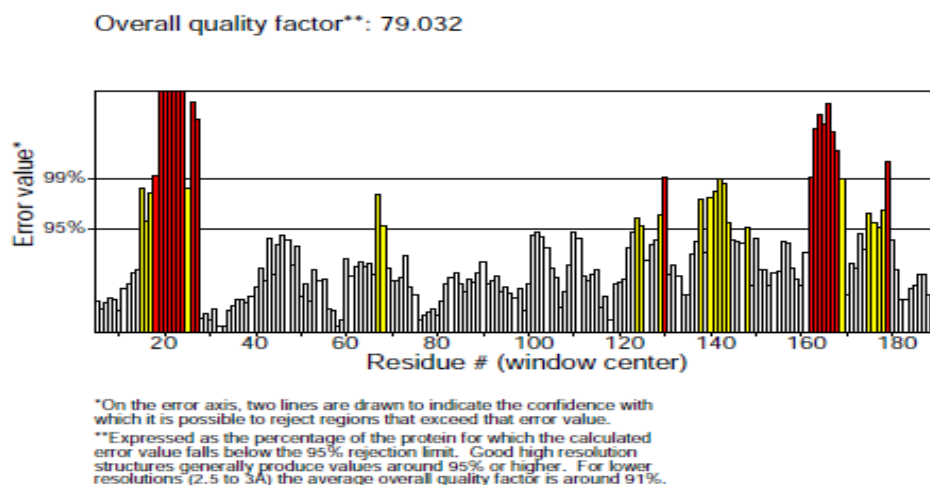


Fig. 10. ERRAT gave a quality factor of 79.032 % to the modelled SARS-CoV-2 spike protein structure.

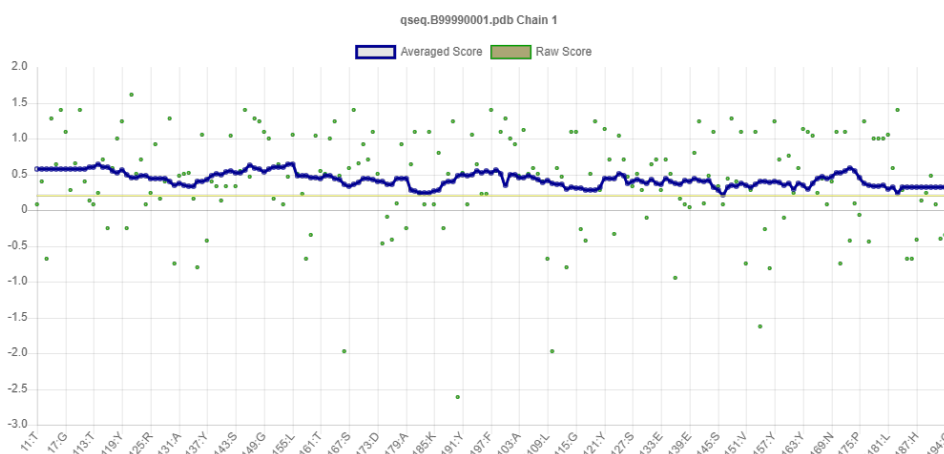


Fig. 11. Verify 3D score diagram validating the modelled SARS-CoV-2 spike protein.

Molecular Docking Results

Rating of docking simulation methodological analysis

At first, the Molegro virtual docker (MVD) and AutoDock Vina programs estimated with PyRx were selected to carry out a docking-based virtual screening approach. Both Scoring Functions were supplied to select the best ligand against *P. aeruginosa* and SARS-CoV-2.

Molecular docking simulation is of two types; namely the oriented docking simulation and the blind docking simulation [67,68]. Blind molecular docking simulation is a docking procedure in which the position of the enzyme's active site is unknown, and the grid center of attention used during grid box determination is not unique to a specific area but is applied throughout the entire enzyme area. The grid box was specifically resolved after orienting docking by establishing the position of the enzyme active site. Because the active site region has been established, oriented docking was used in this study (using Melogro Virtual Docker). In this case, an oriental docking simulation was carried out. The molecular docking simulation of 25 iminoguanidine derivatives, against modelled proteins *Pseudomonas aeruginosa* and SARS-CoV-2 Spike protein, were

performed using MVD software. The modelled protein was taken as the substrate for molecular docking. The substrate had only three cavities. From the three (3) predicted cavities the one with the highest volume was picked for molecular docking simulations. The cavity with volumes 48.64 \AA^3 (Fig. 12) and 11.78 \AA^3 (Fig. 13) are used for docking against *P. aeruginosa* and SARS-CoV-2, respectively. The MolDock score [69] was picked out for rating the inhibitor poses, and for all the modelled proteins docking executed here, the poses picked out as the best were those with the most prominent MolDock score. The MolDock score of all the ligands including the reference drugs ranges from $-125.553 \text{ kcal/mol}$ to 7320.88 kcal/mol as presented in Table 6. The reference drug (Ruxolitinib) has the highest negative MolDock score of $-125.553 \text{ kcal/mol}$ and $-117.237 \text{ kcal/mol}$, followed by compound 15 with -117.489 and -102.85 kcal/mol against *P. aeruginosa* and SARS-CoV-2 Spike protein, respectively.

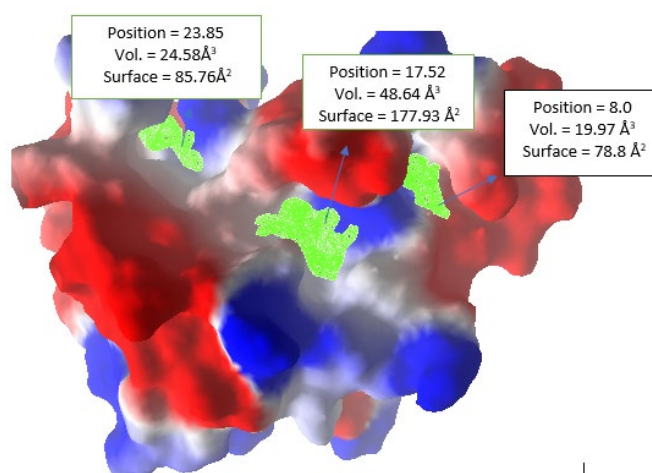


Fig. 12. Modelled protein of potential inhibitor of vertebrate lysozyme [*Pseudomonas aeruginosa*] with electrostatic surface and active site cavity (green grid).

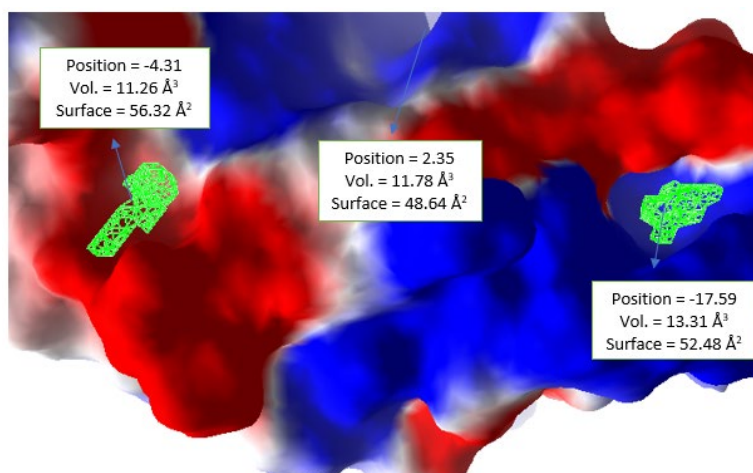


Fig. 13. The electrostatic surface and active site cavity (green grid) of modelled protein of SARS-CoV-2 Spike protein S1.

Table 6. The docking energies of *Pseudomonas aeruginosa* and SARS-CoV-2.

Compound Number	Putative inhibitor of vertebrate lysozyme [<i>Pseudomonas aeruginosa</i>]			SARS-CoV-2 Spike protein S1		
	MolDock Score	Rerank Score	HBond	MolDock Score	Rerank Score	HBond
1	-70.1638	-61.8064	-2.063	-68.012	-52.3464	-6.938
2	-78.6419	-65.4992	-0.1895	-74.8674	-63.0005	-4.2532
3	-78.0857	-62.122	-4.7344	-70.0071	-57.925	-7.4733
4	-74.7886	-62.1973	-1.0691	-75.2023	-63.9325	-5.7075
5	-75.7968	-69.6859	-2.2375	-78.1734	-71.7748	-10.972
6	-76.4574	-65.8232	-0.8352	-74.2074	-65.5021	-6.5918
7	-86.7011	-76.9983	-1.4985	-78.6897	-68.0336	-0.4485
8	-77.9807	-66.1488	-0.2075	-63.489	-42.3824	-4.0915
9	-77.39	-63.7478	-2.3958	-76.7554	-65.5159	-7.1415
10	-81.445	-62.382	-2.1289	-78.2765	-68.2709	-1.1466
11	-77.9612	-65.2889	0	-76.3297	-66.2211	-4.8914
12	-87.3695	-72.0002	-1.0544	-85.7218	-74.148	-6.5194
13	-73.5418	-63.5228	-0.6091	-65.7669	-42.6207	-4.1196
14	-70.2995	-62.085	-1.7337	-65.2352	-41.5776	-4.1817
15	-117.489	-94.1851	-2.9057	-102.85	-89.8918	-7.9081
16	-74.9906	-63.4911	-3.4383	-76.7682	-65.6931	-7.1996
17	-76.9665	-63.8531	-1.178	-67.7912	-61.0345	-8.9551
18	-86.185	-71.2461	-0.1287	-81.353	-61.7442	-2.5
19	-82.6353	-69.5908	-6.5194	-71.7263	-49.8586	-6.2159
20	-73.5011	-60.3293	-4.3336	-71.293	-61.5842	-6.1265
21	-75.3018	-62.8724	-5.2574	-77.3631	-66.4571	-7.1931
22	-84.3353	-71.885	-0.5039	-84.2272	-71.6465	-8.0651
23	-77.2196	-63.3773	-1.2512	-80.6143	-68.8821	-7.3401
24	-75.1741	-62.151	-0.0435	-80.6	-68.6933	-6.7389
25	-74.903	-64.3062	-1.2884	-68.427	-43.4749	-4.1837
Chloroquine	-111.536	-82.6976	-1.1128	-101.402	-72.2266	-4.5758
Ampicillin	-114.864	-95.187	-8.8215	-102.865	-81.7504	-4.5601
Vibramycin	3239.53	565.867	-0.6755	3342.46	660.603	-2.3881
Ruxolitnib	-125.553	-97.9409	-2.3851	-117.237	-88.9727	-6.1061

The reference drug (Ruxolitnib) against the protein of putative inhibitor of vertebrate lysozyme (*Pseudomonas aeruginosa*) formed two hydrogen bond interactions with Cys132 and Asp193 (Fig. 14(A)). Next is compound 15 which also formed two hydrogen bond interactions with the modelled protein first with Asn125, secondly with Gln189 (Fig. 14(B)). Against the modelled protein of SARS-CoV-2 Spike protein S1, the reference drug Ruxolitnib with the highest negative MolDock score of -117.237 kcal/mol formed hydrogen bond interactions with Lys126, Asn149, and Cys148, respectively (Fig. 15(A)). In Fig. 15(B), compound 15

formed five (5) hydrogen bond interactions with the same modelled proteins. The five hydrogen bond interactions of compound 15 are with Arg125, Lys126, and Ser127, respectively.



Fig. 14. Hydrogen bond interactions between (A) Ruxolitinib and (B) compound 15 with modelled protein of putative inhibitor of vertebrate lysozyme [*Pseudomonas aeruginosa*].

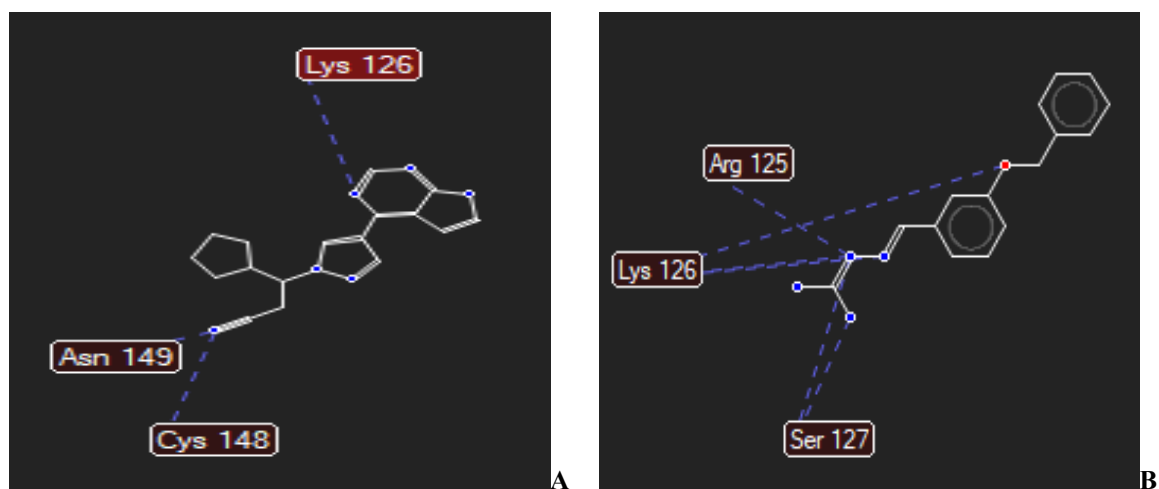


Fig. 15. Hydrogen bond interactions between (A) Ruxolitinib and (B) compound 15 with modelled protein of SARS-CoV-2 Spike protein S1.

Again, blind molecular docking simulations were carried out for all 25 compounds including the reference compounds with PyRx v0.8 software. In the molecular docking simulations with PyRx, we make use of the modelled proteins as the target. The binding affinity values obtained as a result of the molecular docking analyses are shown in Table 7. The negative value of binding affinity as presented in Table 6 suggests the stability of the protein-ligand complex. From the table, the reference compound “Vibramycin” has the highest binding affinity value of -7.4 kcal/mol for *P. aeruginosa* and SARS-CoV-2, respectively. The reference compound (Vibramycin) interacted with *P. aeruginosa* modelled protein structure (Fig. 16) showed the corresponding amino acids such as aspartic acid, phenylalanine, glycine, glutamine, tyrosine, and so on. The

docking of Vibramycin forms some hydrogen bonds and electrostatic interactions with the modelled protein. Vibramycin is observed to show three (3) hydrogen bond interactions with Ser192 (3.87 Å), Gln189 (5.33 Å), and Gln189 (4.27 Å) against *P. aeruginosa* modelled protein (Fig. 16(A)). Vibramycin exhibits three electrostatic interactions, one unfavorable acceptor-acceptor interaction, one carbon-hydrogen bond with Pro131 (6.23 Å), and two conventional hydrogen bond interactions with Ser182 (3.87 Å) and Glu184 (4.79 Å), respectively against SARS-CoV-2 protein (Fig. 16(B)).

Table 7. Autodock-Vina predicted binding affinity (kcal/mol) of *P. aeruginosa* and SARS-CoV-2 spike protein.

Ligand	<i>P. aeruginosa</i>	SARS-CoV-2
	Binding Affinity	Binding Affinity
1	-5.4	-5.8
2	-5.7	-6.3
3	-6.2	-6.2
4	-6.4	-6.2
5	-5.6	-6.2
6	-6.2	-6.1
7	-6.9	-6.9
8	-5.5	-6
9	-6.3	-6.1
10	-5.7	-6.2
11	-5.4	-6.4
12	-6.3	-6.6
13	-5.6	-6.3
14	-5.2	-5.9
15	-5.9	-7.2
16	-6.1	-6
17	-5.9	-6
18	-6.6	-7.1
19	-5.8	-6.3
20	-5.6	-6.3
21	-6.8	-6.2
22	-6	-6.3
23	-5.6	-5.9
24	-5.3	-6.4
25	-5.7	-6.1
Chloroquine	-4.1	-4.8
Ruxolitnib	-5.9	-7.0
Vibramycin	-7.4	-7.4
Ampicillin	-5.9	-6.7

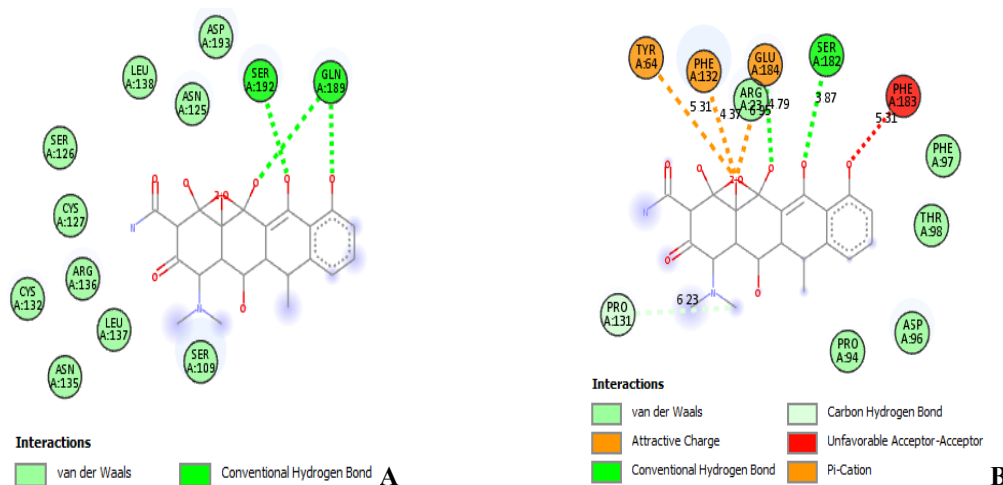


Fig. 16. The 2D illustrations of interactions between Vibramycin and the modelled (A) *P. aeruginosa* and (B) SARS-CoV-2 spike protein receptors.

Compounds 7 and 15 were selected based on their remarkable binding affinity of -6.9 and -7.2 kcal/mol for *P. aeruginosa* and SARS-CoV-2 spike protein receptors respectively compared to the reference compound's -7.4 kcal/mol. The molecular docking contact modes for compounds 7 and 15 are shown in Figures 17 and 18, respectively. The molecular docking interaction of compound 7 with the active site of modelled *P. aeruginosa* protein reveals two conventional hydrogen bonds between the hydrogen atoms of compound 7 and the amino acid Ser109 with distances of 3.45 and 4.04, respectively, and one carbon-hydrogen bond between the carbon atom and the amino acid Pro107 with a distance of 5.47. The hydrophobic interaction was performed between one of the naphthalene ring and amino acids Arg136 (5.84 Å) and Cys127 (5.04 Å). Moreover, the electrostatic interaction is observed between one of the naphthalene rings and the amino acids Asp193 (5.94 Å) and Arg136 (5.48 Å).

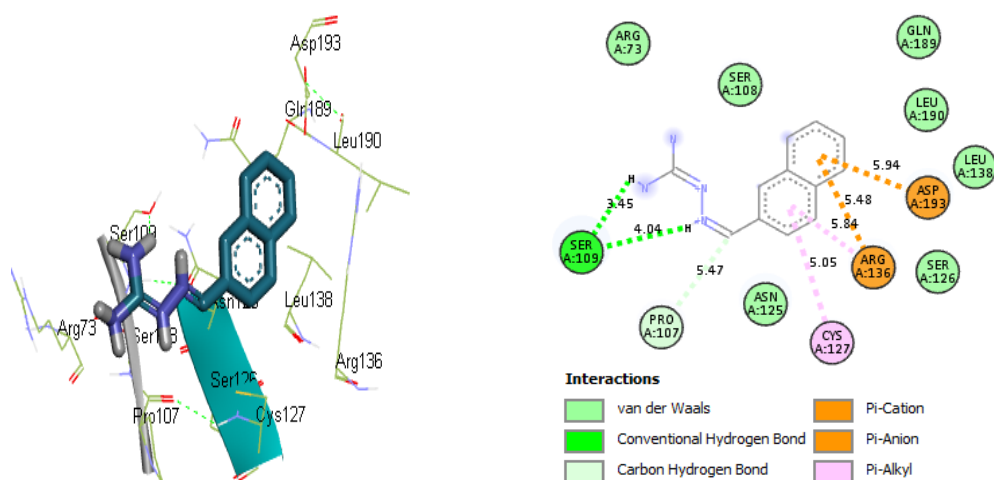


Fig. 17. The 3D and 2D representations of interactions between compound 7 and modelled *P. aeruginosa* receptor.

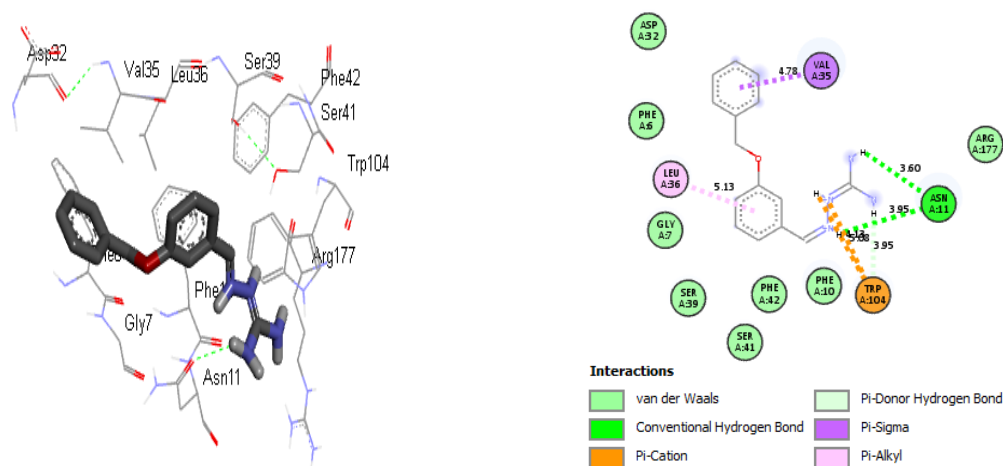


Fig. 18. The 3D and 2D representations of interactions between compound 15 and modelled SARS-CoV-2 spike protein receptor.

Nevertheless, compound 15 forms two conventional hydrogen bonds with amino acid Asn11 (3.60 and 3.95 Å), one pi-donor hydrogen bond with Trp104 (3.95 Å) with the active site of modelled SARS-CoV-2 spike protein. The hydrophobic interaction was performed between one of the benzyloxy and toluene rings with amino acids Leu36 (5.13 Å) and Val35 (4.78 Å), respectively. Furthermore, the electrostatic interaction is observed between hydrogen and the amino acid Trp104 (4.13 Å). The hydrogen bonds, the hydrophobic interactions, and the electrostatic interactions in the complexes showed that compounds 7 and 15 of iminoguanidine derivatives are most active against modelled *P. aeruginosa* receptor and modelled SARS-CoV-2 spike protein, respectively.

Lipinski's rule of five aids [70] in the differentiation of drug-like and non-drug-like compounds. It forecasts a high likelihood of success or failure due to drug-likeness for compounds that meet two or more of the rules stated in Table 7. The chemical structure of compounds 7 and 15 was submitted to the MolSoft server in SMILE format, to guess the *in silico* pharmacokinetic parameters. Based on these properties, we investigated how our lead ligands met the Lipinski rule of five (5). The table (Table 8) shows the properties of compounds 7 and 15 in which both of the compounds fulfilled the Lipinski rule of 5. The MollogS value is less than or equal to -4.0. These compounds have a low MolLogS value, indicating that they are difficult to dissolve in water [71]. The aromatic and non-polar fragments around the chemical structures are most likely to blame. Despite this, their intestinal absorption is high, exceeding 90 %.

Table 8. The pharmacokinetic parameter of compounds 7 and 15 using MolSoft server.

Compound 7	values	Compound 15	values
Molecular formula	C12 H14 N4	Molecular formula	C15 H18 N4 O
Mol. weight (≤ 500 D)	214.12	Mol. weight (≤ 500 D)	270.15
Number of HBA (≤ 10)	2	Number of HBA (≤ 10)	3
Number of HBD (≤ 5)	5	Number of HBD (≤ 5)	5
MolLogP (≤ 5)	1.24	MolLogP (≤ 5)	1.42
MolLogS	-1.96	MolLogS	-1.48
MolPSA (≤ 140)	64.39 A2	MolPSA (≤ 140)	71.89 A2
MolVol	199.97 A3	MolVol	255.68 A3
PKA of most Basic/Acidic group	11.82 / 16.27	PKA of most Basic/Acidic group	11.82/ 16.29
BBB score	3.06	BBB score	3.00

The leads in Fig. 19(A) and 19(B) showed that these molecules display ADME profile within the desired range as presented in Table 7, and therefore the values of these parameters shed light on their drug-likeness features, which could set up an essential starting point to encourage experimental procedures.

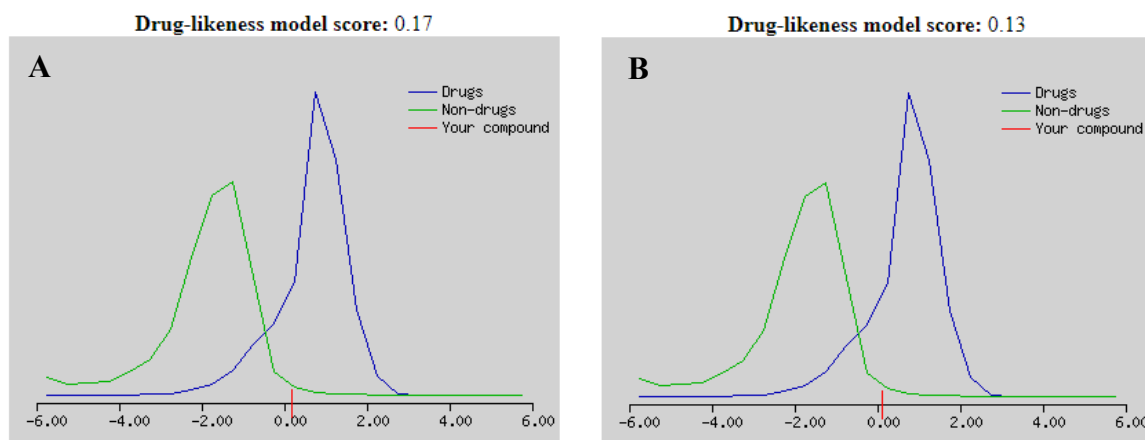


Fig. 19. The drug-likeness profile of (A) compound 7 (B) compound 15.

Molecular dynamics (MD) simulations

The study of molecular dynamics simulation in the modelled SARS-CoV-2 spike protein was being used to figure out how flexible is docked compound 15, and this evaluation was lasted up to 1 ns. To compute these RMSD, RMSF, and SASA values we used the VMD plugin with a probe radius set to 1.4 Å. Other thermodynamic parameters investigated are the total energy, potential energy, and kinetic energy of the docked protein-ligand complex [71]. To understand the structural aberrations across the simulation track, the RMSD of the c-alpha atoms in the protein-ligand complex was analyzed. The RMSD value of the complex increased at the start of the simulation, as seen in Fig. 20(A). The protein-ligand complex then stabilized at 400 and 500 ps, before rapidly increasing from 600 to 800 ps. The protein-ligand complex remained stable until the simulation's final phase, 1 ns. To estimate the amino acid residue flexibility, the RMSF of the protein-ligand complex was calculated. The majority of residues had RMSF values < 2.0 Å, as seen in Fig. 20(B). The presence of peaks indicates the areas fluctuate the most during the simulation. The fluctuations observed are around 1.9 Å (600 ps) and less around 0.4 Å (1 ns) which signifies the stability of the system. This RMSF profile indicates a complex that is less flexible and inflexible. The surface area of a group that is accessible to a solvent probe is known as the solvent-accessible surface area (SASA) [or temporarily accessible surface area (ASA)]. Protein folding and stability have always been determined by the solvent-accessible surface area (SASA). The SASA trend in the predicted protein-ligand complex was higher, indicating an increase in surface volume. Nonetheless, this complex had a high number of SASA deviations, indicating that the protein's surface area was changing significantly. The SASA value for the un-simulated protein-ligand complex has a lower value in comparison with the simulated protein-ligand (Table 9) which indicates that the simulated protein-ligand complex is more exposed to water and also increases in cavities. The SASA profile in Fig. 20(C) confirmed the claim.

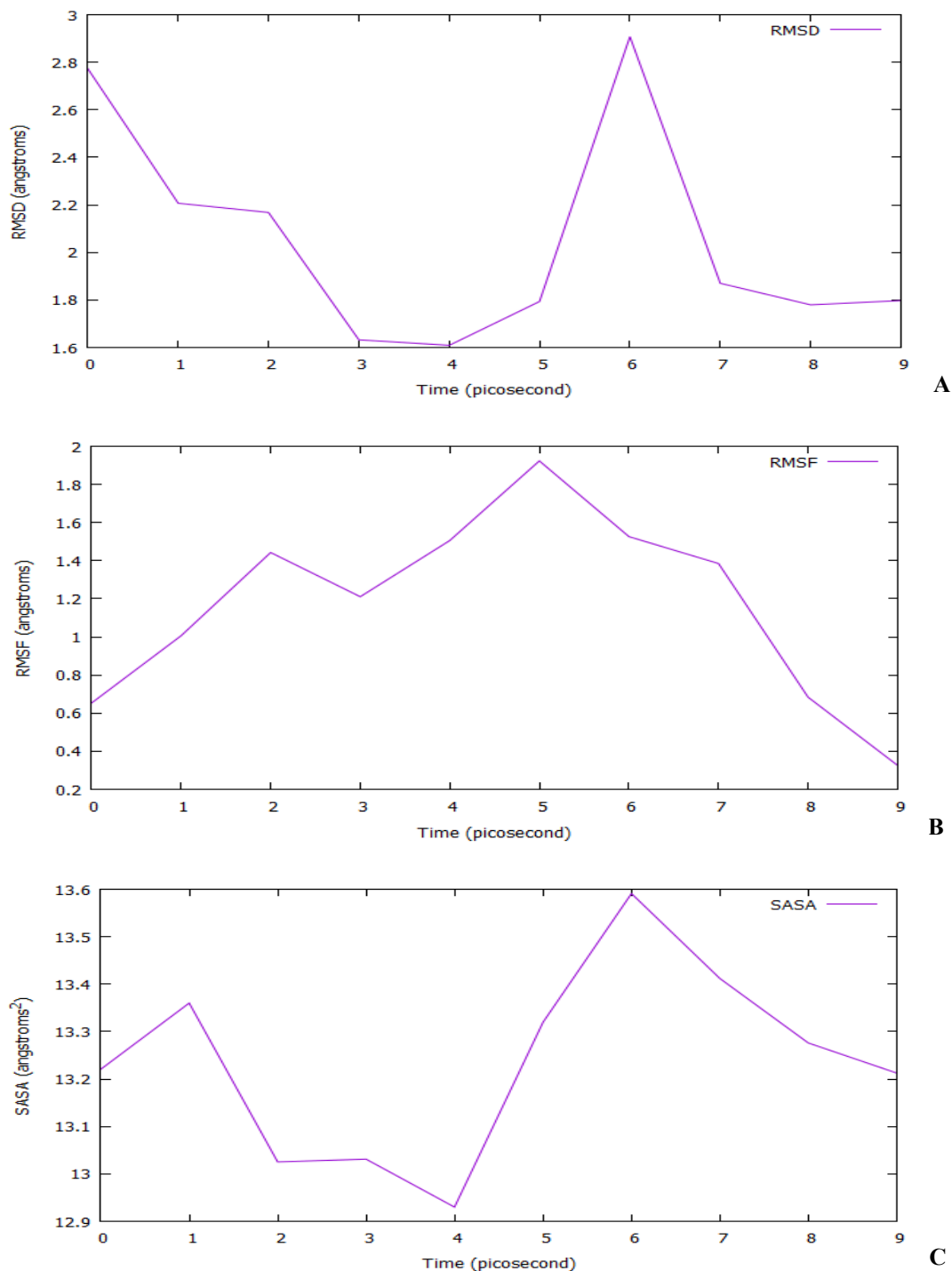


Fig. 20. The molecular dynamics simulation of the complexes. (A) Root mean square deviation (RMSD), (B) Root mean square fluctuations of the amino acid residues (RMSF), and (C) Solvent accessible surface area (SASA).

Table 9. The solvent-accessible surface area of un-simulated and simulated complexes.

The surface area of SARS-CoV-2 spike protein	
Number of non-HOH	non-H atoms=1537
Probe radius=1.40	
TOTAL, ASA=10006.94	TOTAL, MSA=8706.50
Polar ASA=4481.68	Non-polar ASA=5525.26
Polar MSA=3401.84	Non-polar MSA=5304.67
Total backbone ASA=2134.02	Total backbone MSA=2404.09
Polar backbone ASA=1414.88	Non-polar backbone ASA=719.14
Polar backbone MSA=1449.43	Non-polar backbone MSA=954.66
Polar side chain ASA=3066.80	Non-polar side chain ASA=4806.12
Polar side chain MSA=1952.40	Non-polar side chain MSA=4350.01
+charge ASA=690.04	-charge ASA=450.70
+charge MSA=484.54	-charge MSA=263.31
Structure contains 10 cavities	A
The surface area of SARS-CoV-2 spike protein after 1 nanosecond	
Number of non-HOH	non-H atoms=1733
Probe radius=1.40	
TOTAL, ASA=11687.42	TOTAL, MSA=0.00
Polar ASA=6785.50	Non-polar ASA=4901.92
Polar MSA=0.00	Non-polar MSA=0.00
Total backbone ASA=2325.91	Total backbone MSA=0.00
Polar backbone ASA=1612.36	Non-polar backbone ASA=713.55
Polar backbone MSA=0.00	Non-polar backbone MSA=0.00
Polar side chain ASA=5173.14	Non-polar side chain ASA=4188.37
Polar side chain MSA=0.00	Non-polar side chain MSA=0.00
+charge ASA=831.32	-charge ASA=575.06
+charge MSA=0.00	-charge MSA=0.00
Structure contains 13 cavities	B

A = the un-simulated protein-ligand complex; **B** = the simulated protein-ligand complex.

The MD simulations of the 10 frames and the hydrogen bond information of these 10 confirmations are listed in Fig. S6. From the figures, conformation 3, 5, 8, and 9 has 4, 2, 2, and 2 hydrogen bonds, respectively. The amino acids that appear more as conventional hydrogen bond and carbon-hydrogen bond interactions are Ser41 and Asp 32, respectively. From conformation 1 to 10, amino acid residue Ser41 appears in all the interactions either as hydrogen bond or van der Waals interactions to the ligand. After 1 ns MD simulation (the last confirmation "10"), the residue interactions and protein-ligand 2D contact contour map show that van der Waals and hydrophobic interactions were crucial contributions for the stability of the hit drug.

Fig. S7 depicts a change in total energy versus time in 1 ns of molecular dynamics simulation, which displays a slight fluctuation throughout the 1 ns. It also suggested that when the protein was simulated up to or beyond 1 ns, total energy (at or above -137713.02 KJ/mol), and kinetic energy (at or beyond 31498.99 kJ/mol), compound 15 produced a stable combination with expected protein structure. Moreover, the examined system's potential energy in 1 ns of simulation at 303.15 K was -169212.01 KJ/mol (Fig. S7(C)). As a result, under stable temperature settings, the derived equilibrium structure of the expected spike protein S1 structure was obtained at 303.15 K. These findings reveal that in the MD simulation, energy conservation was met. At the end of the MD simulation, the coordinates of the ligand in the inserted binding site were changed (Fig. S6), demonstrating the utility of MD simulations after ligand docking in the binding site. Except for Asp32, Leu36, Phe42, Phe10,

Phe6, Asn11, and Val35, the active site residues determined by docking were changed, and new residues such as Cys29, Val192, Tyr33, Val30, Ala31, Ile26, Leu181, and Val63 are positioned near ligands and could involve in the interaction, according to exploratory molecular dynamics simulations on the complex between the ligand.

Conclusion

In this paper, we outline how QSAR, homology modelling, docking, and molecular dynamic simulation were used to investigate numerous proteins, including a suspected inhibitor of vertebrate lysozyme [*Pseudomonas aeruginosa*] and chain A, spike protein S1 [SARS-CoV-2]. The 2D-QSAR was utilized to investigate iminoguanidine derivatives as a negative catalyst against *P. aeruginosa*. The GA-MLR results revealed that ATSC1s and SHBint7 are the most critical descriptors for predicting inhibitor activity. As a result, ATSC1s have a considerable beneficial effect on pMIC50 levels, but SHBint7 has a negative effect. The data gathered so far will aid in predicting the activity of novel *P. aeruginosa* inhibitor candidates. A protein structure was created with the use of various templates utilizing homology modelling techniques to explore the hit molecule's inhibitions, and validation investigations (ERRAT, VERIFY3D, and PROCHECK) indicated that the predicted protein structure was reliable and of high quality. The binding mechanism of these ligands is revealed by docking simulation studies of anticipated protein structures, as well as the amino acid involved in ligand recognition. After a successful docking simulation, exploratory molecular dynamics simulations on the ligand-protein complex disclosed that, except for Asp32, Leu36, Phe42, Phe10, Phe6, Asn11, and Val35, the rest of the residues in the active site ascertained by docking were altered, and some new residues such as Cys29, Val192, Tyr33, Val30, Ala31, Ile26, Leu181, and Val63 were positioned in the interaction. It appears reasonable to conclude that selected compound 15 could be a good option for treating *P. aeruginosa* and the SARS-CoV-2 virus and that it will hopefully have a long and interesting future in the identification and optimization of promising leads with high potential for developing new therapeutic agents. We expect that our findings should contribute to the development of effective antimicrobials for *P. aeruginosa* and the SARS-CoV-2 virus.

Acknowledgments

The authors are thankful to MarvinView, and PaDEL v2.20 developers for providing the free versions of their software. The authors are thankful to Dr. Paola Gramatica, Italy, and her team for providing QSARINS-2.2.4 (www.qsar.it).

References

1. Edache, E. I.; Uzairu, A.; Mamza, P. A. P.; Shallengwa, G. A. *J. Drug. Design. Discov. Res.* **2020**, 1, 36-52.
2. Yuan, M.; Guan, H.; Sha, D.; Cao, W.; Song, X.; Che, J.; Kan, B.; Li, J. *Antibiotics.* **2021**, 10, 1234. DOI: <https://doi.org/10.3390/antibiotics10101234>.
3. Azam, M. W.; Khan, A. U. *Drug Discov. Today.* **2019**, 24, 350-359. DOI: <https://doi.org/10.1016/j.drudis.2018.07.003>.
4. Sahu, M. K.; Kaushik, K.; Das, A.; Jha, H. *Bioresour. Bioprocess.* **2020**, 7, 14. DOI: <https://doi.org/10.1186/s40643-020-00303-z>.
5. Ahlgren, H. G.; Benedetti, A.; Landry, J. S.; Bernier, J.; Matouk, E.; Radzioch, D.; Lands, L. C.; Rousseau, S.; Nguyen, D. *BMC Pulm. Med.* **2015**, 15, 67. <https://doi.org/10.1186/s12890-015-0062-7>.
6. Gifford, A. H.; Miller, S. D.; Jackson, B. P.; Hampton, T. H.; O'Toole, G. A.; Stanton, B. A.; Parker, H. W. *Pediatr. Pulmonol.* **2011**, 46, 160-165. DOI: <https://doi.org/10.1002/ppul.21335>.

7. Callewaert, L.; Aertsen, A.; Deckers, D.; Vanoirbeek, K. G. A.; Vanderkelen L, et al. *PLoS Pathog.* **2008**, *4*, e1000019. DOI: <https://doi.org/10.1371/journal.ppat.1000019>.
8. Callewaert, L.; Van Herreweghe, J. M.; Vanderkelen, L.; Leysen, S.; Voet, A.; Michiels, C. W. *Trends Microbiol.* **2012**, *20*, 501-510. DOI: <https://doi.org/10.1016/j.tim.2012.06.005>.
9. Yum, S.; Kim, M. J.; Xu, Y.; Jin, X. L.; Yoo, H. Y.; Park, J. W.; Gong, J. H.; Choe, K. M.; Lee, B. L.; Ha, N. C. *Biochem Biophys Res Commun.* **2009**, *378*, 244-248. DOI: <https://doi.org/10.1016/j.bbrc.2008>.
10. Leysen, S.; Van Herreweghe J. M.; Callewaert, L.; Heirbaut, M.; Buntinx, P.; Michiels, C. W.; Strelkov, S. V. *J. Mol. Bio.* **2011**, *405*, 1233-1245. DOI: 10.1016/j.jmb.2010.12.007.
11. Li, Y.; Li, H.; Fan, R.; Wen, B.; Zhang, J.; Cao, X.; Wang, C.; Song, Z.; Li, S.; Li, X.; Lv, X. *Intervirology.* **2016**, *59*, 163-169. DOI: <https://doi.org/10.1016/j.jmb.2010.12.007>.
12. Noman, A.; Aqeel, M.; Khalid, N.; Hashem, M.; Alamari, S.; Zafar, S.; Qasim, M.; Irshad, M. K.; Qari, S. H. *Microb. Pathog.* **2021**, *150*, 104719. DOI: <https://doi.org/10.1016/j.micpath.2020.104719>.
13. Chan, J. F. W.; Yuan, S.; Kok, K. H.; To, K. K. W.; Chu, H.; Yang, J.; Xing, F.; Liu, J.; Yip, C. C. Y.; Poon, R. W. S.; Tsoi, H. W. *Lancet.* **2020**, *395*, 514-523. DOI: [https://doi.org/10.1016/S0140-6736\(20\)30154-9](https://doi.org/10.1016/S0140-6736(20)30154-9).
14. Zou, L.; Ruan, F.; Huang, M.; Liang, L.; Huang, H.; Hong, Z.; Yu, J.; Kang, M.; Song, Y.; Xia, J.; Guo, Q. *New Eng. J. Med.* **2020**, *382*, 1177-1179. DOI: <https://doi.org/10.1056/NEJMc2001737>.
15. Rhoades, N. S.; Pinski, A. N.; Monsibais, A. N.; Jankeel, A.; Doratt, B. M.; Cinco, I. R. *Cell Rep.* **2021**, *36*, 109637. DOI: <https://doi.org/10.1016/j.celrep.2021.109637>.
16. Simmons, G.; Reeves, J. D.; Rennekamp, A. J.; Amberg, S. M.; Piefer, A. J.; Bates, P. *Proc. Natl. Acad. Sci.* **2004**, *101*, 4240-4245. DOI: <https://doi.org/10.1073/pnas.0306446101>.
17. Guo, Y. R.; Cao, Q. D.; Hong, Z. S.; Tan, Y. Y.; Chen, S. D.; Jin, H. J.; Tan, K. S.; Wang, D. Y.; Yan, Y. *Military Med. Res.* **2020**, *7*, 1-10. DOI: <https://doi.org/10.1186/s40779-020-00240-0>.
18. Condé, K.; Atakla, H. G.; Garba, M. S.; Garba, I. *Afr. Med. J.* **2020**, *35*, 96. DOI: <https://doi.org/10.11604/pamj.supp.2020.35.2.24616>.
19. Kim, S. Y.; Jin, W.; Sood, A.; Montgomery, D. W.; Grant, O. C.; Fuster, M. M.; Fu, L.; Dordick, J. S.; Woods, R. J.; Zhang, F.; Linhardt, R. J. *Antiviral Res.* **2020**, *181*, 104873. DOI: <https://DOI.org/10.1016/j.antiviral.2020.104873>.
20. Muratov, E. N.; Amaro, R.; Andrade, C. H.; Brown, N.; Ekins, S.; Fourches, D.; Isayev, O.; Kozakov, D.; et al. *Chem. Soc. Rev.* **2021**, *50*, 9121-9151. DOI: <https://doi.org/10.1039/D0CS01065K>.
21. Tortorici, M. A.; Walls, A. C.; Lang, Y.; Wang, C.; Li, Z.; Koerhuis, D.; Boons, G. J.; Bosch, B. J.; Rey, F. A.; de Groot, R. J.; Veesler, D. *Nat. Struct. Mol. Biol.* **2019**, *26*, 481-489. <https://DOI.org/10.1038/s41594-019-0233-y>.
22. Vashi, Y.; Jagrit, V.; Kumar, S. *Infect., Gen. Evol.* **2020**, *84*, 104382. DOI: <https://DOI.org/10.1016/j.meegid.2020.104382>.
23. Kanaan, M. H. G. *Vet. World.* **2018**, *11*, 1445-1453. DOI: <https://doi.org/10.14202/vetworld.2018.1445-1453>.
24. Heinzl, G. A.; Huang, W.; Yu, W.; Giardina, B. J.; Zhou, Y.; MacKerell, Jr., A. D.; Wilks, A.; and Xue, F. *J Med Chem.* **2016**, *59*, 6929-6942. DOI: <https://doi.org/10.1021/acs.jmedchem.6b00757>.
25. Edache, E. I.; Uzairu, A.; Mamza, P. A.; Shallangwa, G. A. *J. Genet. Eng. Biotechnol.* **2022**, *20*, 88. DOI: <https://doi.org/10.1186/s43141-022-00362-z>.
26. Edache, E. I.; Uzairu, A.; Mamza, P. A.; Shallangwa, G. A. *J. Virol. Viral Dis.* **2022**, *2*, 1-28. DOI: <https://doi.org/10.54289/JVVD2200106>.
27. Custelcean, R. *Chem. Commun.* **2020**, *56*, 10272-10280. DOI: <https://doi.org/10.1039/D0CC04332J>.
28. Zhang Q.; Jiang, Y.; Li, Y.; Song, X.; Luo, X.; Ke, Z.; and Zou, Y. *iScience.* **2021**, *24*, 102263. DOI: <https://doi.org/10.1016/j.isci.2021.102263>.
29. Maia, E. H. B.; Assis, L. C.; Oliveira, T. A.; Silva, A. M.; Taranto, A. G. *Front. Chem.* **2020**, *8*, 343. DOI: <https://doi.org/10.3389/fchem.2020.00343>.
30. Liu, Y.; Zhang, L.; Guo, M.; Wu, H.; Xie, J.; Wei, D. *Bioresour. Bioprocess.* **2014**, *1*, 1-9. DOI: <https://doi.org/10.1186/s40643-014-0017-5>.

31. Shirvani, P.; Fassihi, A. *Mol. Simul.* **2020**, 46, 16, 1265-1280. DOI: <https://doi.org/10.1080/08927022.2020.1810853>.
32. Becke, A. D. *J. Chem. Phys.* **1993**, 98, 5648-5652. DOI: <https://doi.org/10.1063/1.464913>.
33. Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B.* **1988**, 37, 785-789. DOI: <https://doi.org/10.1103/physrevb.37.785>.
34. Petersson, G. A.; Bennett, A.; Tensfeldt, T. G.; Al-Laham, M. A.; Shirley, W. A.; Mantzaris, J. J. *J. Chem. Phys.* **1988**, 89, 2193-2218. DOI: <https://doi.org/10.1063/1.455064>.
35. Frisch, M. J.; Trucks, G.W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; et al. Revision A.02. **2016**, Gaussian Inc, Wallingford CT
36. Yap, C. W. *J. Comput. Chem.* **2011**, 32, 1466-1474. DOI: <https://doi.org/10.1002/jcc.21707>.
37. Gramatica, P.; Chirico, N.; Papa, E.; Cassani, S.; Kovarich, S. *J. Comput. Chem.* **2013**. DOI: <https://doi.org/10.1002/jcc.23361>.
38. Edache, E. I.; Uzairu, A.; Mamza, P. A.; Shallangwa, G. A. *Sci. Afr.* **2021**, 15, e01088. DOI: <https://doi.org/10.1016/j.sciaf.2021.e01088>.
39. Roy, K.; Mitra, I. *Comb. Chem. High Throughput Screen.* **2011**, 14, 450-474. DOI: <https://doi.org/10.2174/138620711795767893>.
40. Roy, K.; Chakraborty, P.; Mitra, I.; Ojha, P. K.; Kar, S.; Das, R. N. *J. Comput. Chem.* **2013**, 34, 1071-1082. DOI: <https://doi.org/10.1002/jcc.23231>.
41. Edache, E. I.; Uzairu, A.; Abechi, S. E. *J. Comput. Methods Mol. Des.* **2015**, 5, 135-149
42. Sneath, P. H.; Sokal, R. R. Numerical Taxonomy San Francisco. A statistical method for evaluating systematic relationships, **1973**, 38, 1409-1438.
43. Zuckerkandl, E.; Pauling, L. in: *Evolutionary divergence and convergence in proteins. In Evolving genes and proteins*, Academic Press., **1965**, 97-166.
44. Kumar, S.; Stecher, G.; Li, M.; Nknyaz, C.; Tamura, K. *J. Mol. Biol. Evol.* **2018**, 35, 1547-1549. DOI: <https://doi.org/10.1093/molbev/msy096>.
45. Li, Y.; Qiu, Y.; Zhang, X.; Zhu, M.; Tan, W. *Bioresour. Bioprocess.* **2019**, 6, 15. DOI: <https://doi.org/10.1186/s40643-019-0250-z>.
46. Sali, A.; Blundell, T. L. *J. Mol. Biol.* **1993**, 234, 779-815. DOI: <https://doi.org/10.1006/jmbi.1993.1626>.
47. Feng, J.; Xu, S.; Feng, R.; Kovalevsky, A.; Zhang, X.; Liu, D.; Wan, Q. *Bioresour. Bioprocess.* **2021**, 8, 102. DOI: <https://doi.org/10.1186/s40643-021-00449-4>.
48. Hall, T. A. *Nucl. Acids. Symp. Ser.* **1999**, 41, 95-98.
49. Guex, N.; Peitsch, M. C. *Electrophoresis* **1997**, 18, 2714-2723. DOI: <https://doi.org/10.1002/elps.1150181505>.
50. Thomsen, R.; Christensen, M. H. Molegro Virtual Docker 4.0 User Manual; Molegro ApS: Aarhus, Denmark, **2009**, 125-133.
51. Dallakyan, S.; Olson, A. J. *Methods Mol. Biol.* **2015**, 1263, 243-250. DOI: https://doi.org/10.1007/978-1-4939-2269-7_19.
52. Trott, O.; Olson, A. J. *J. Comput. Chem.* **2010**, 31, 455-461. DOI: <https://doi.org/10.1002/jcc.21334>.
53. Jo, S.; Kim, T.; Iyer, V. G.; Im, W. *J. Comput. Chem.* **2008**, 29, 1859-1865. DOI: <https://doi.org/10.1002/jcc.20945>.
54. Lee, J.; Cheng, X.; Swails, J. M.; Yeom, M. S.; Eastman, P. K.; Lemkul, J. A.; Wei, S.; Buckner, J.; Jeong, J. C.; Qi, Y.; Jo, S. *J. Chem. Theory Comput.* **2016**, 12, 405-413. DOI: <https://doi.org/10.1021/acs.jctc.5b00935>.
55. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, 26, 1781-1802. DOI: <https://doi.org/10.1002/jcc.20289>.
56. Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics.* **1996**, 14, 33-38.51. DOI: [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5).
57. El-Mernissi, R.; El Khatabi K.; Khaldan, A.; ElMchichi, L.; Shahinozzaman, M.; Mohammed Aziz Ajana, M. A.; Lakhilifi, T.; Bouachrine, M. *J. Mex. Chem. Soc.* **2022**, 66, 79-94. DOI: <http://dx.doi.org/10.29356/jmcs.v66i1.1578>.

58. Bolboac, S. D.; Lorentz, J. *Biomath.* **2013**, 2, 1-11. DOI: <http://dx.doi.org/10.11145/j.biomath.2013.09.089>.
59. Edache, E. I.; Uzairu, A.; Mamza, P. A.; Shallangwa, G. A. *Biomed. J. Sci. Tech. Res.* **2021**, 34, 26472- 26489. DOI: <https://doi.org/10.26717/BJSTR.2021.34.005509>.
60. Tropsha, A.; Gramatica, P.; Gombar, V. K. *QSAR Comb. Sc.* **2003**, 22, 69-77. DOI: <https://doi.org/10.1002/qsar.200390007>.
61. Ruiz, I. L.; Gomez-Nieto, M. A. *Molecules.* **2018**, 23, 2756. DOI: <https://doi.org/10.3390/molecules23112756>.
62. Gramatica, P.; Cassani, S.; Chirico, N. *J. Comput. Chem.* **2014**, 35, 1036–1044. DOI: <https://doi.org/10.1002/jcc.23576>.
63. Hall, B. G. *Mol. Biol. Evol.* **2013**, 30, 1229-1235. <https://doi.org/10.1093/molbev/mst012>.
64. Edgar, R. C. *BMC Bioinf.* **2004**, 5, 113. DOI: <https://doi.org/10.1186/1471-2105-5-113>.
65. Elengoe, A.; Naser, M. A.; Hamdan, S. *Biologia.* **2015**, 70, 1655-1671. DOI: <https://doi.org/10.1515/biolog-2015-0194>.
66. Mugabo, B.; Iradukunda, R.; Gatanganwa, J. M.; Ufitinema, J. C.; Mutangana, D. *Rwanda J. Engr. Sci. Tech. Environ.* **2021**, 4. DOI: <https://dx.DOI.org/10.4314/rjeste.v4i1.3>.
67. Tambunan, U. S. F.; Bramantya, N.; Parikesit, A. A. *BMC Bioinf.* **2011**, 12, S23. DOI: <http://www.biomedcentral.com/1471-2105/12/S13/S23>.
68. Chen, J.; Luo, X. J.; Chen, Q.; Pan, J.; Zhou, J.; Xu, J. H. *Bioresour. Bioprocess.* **2015**, 2, 39. DOI: <https://doi.org/10.1186/s40643-015-0067-3>.
69. Thomsen, R.; Christensen, M. H. *J. Med. Chem.* **2006**, 49, 3315-3321. DOI: <https://doi.org/10.1021/jm051197e>.
70. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug. Deliv. Rev.* **2001**, 46, 3–26. DOI: [https://doi.org/10.1016/s0169-409x\(00\)00129-0](https://doi.org/10.1016/s0169-409x(00)00129-0).
71. Edache, E. I.; Uzairu, A.; Shallangwa, G. A.; Mamza, P. A. *Future J. Pharm. Sci.* **2021**, 7, 1-22. DOI: <https://doi.org/10.1186/s43094-021-00367-4>.