Molecular phylogeny, Sequence-based drug design, Docking built virtual screening, dynamics simulations, and ADMET properties of thiazolino 2-pyridone amide derivatives as an inhibitor of Chlamydia trachomatis and SARS-CoV-2 protein

Emmanuel Israel Edache[a, b, 1], Adamu Uzairu[b], Paul Andrew Mamza[b], Gideon Adamu Shallangwa[b]

[a]Department of Pure and Applied Chemistry, University of Maiduguri, Borno State, Nigeria
[b]Department of Chemistry, Ahmadu Bello University, Zaria, Nigeria

Abstract: The propagation of emerging diseases and the expensive cost and time lost by using the classic methods, especially in the current scenario with the world being plagued by SARS-CoV-2 and Chlamydia trachomatis diseases, make finding another way to invent new medication very important. That's why we used computational approaches to predict protein-ligand interactions of thiazolino 2-pyridone amide derivatives. The high-throughput virtual screening requires extensive combing through existing datasets in the hope of finding possible matches to screen for new molecules able to inhibit SARS-CoV-2 and Chlamydia trachomatis diseases. In this study, 46 thiazolino-2-pyridone amide derivatives were chosen for planning the powerful inhibitors by utilizing various strategies: QSAR analysis, phylogenetic analysis, homology modeling, docking simulation, molecular dynamics (MD) simulation, as well as ADMET Screening. The 2D QSAR investigation uncovers that these compounds show a satisfactory connection with bioactivity. From that point onward, phylogenetic analysis and homology modeling were used to model the selected receptors, which were then evaluated using both the SAVES and PROSA servers, indicating the best correctness of the modeled protein with the experimental results. Additionally, a docking simulation investigation was carried out to comprehend the 46 thiazolino-2-pyridone amide derivatives' interactions with homologous proteins. Additionally, MD simulations coupled with MM/GBSA verified the chosen complex systems' stability over 1000 ps. Two compounds were chosen as possible inhibitors based on these findings. The expected thiazolino-2-pyridone amide's oral bioavailability and toxicity have been discovered under the ADMET. Thus, these discoveries can be leveraged to develop novel molecules with the necessary action.

Keywords: SARS-CoV-2, Chlamydia trachomatis, Thiazolino 2-pyridone amide, Docking, MD simulations, ADMET, and MM/GBSA calculation

1. Introduction

The current SARS-CoV-2 (Covid-19) emergency has demonstrated that the world is not equipped to answer with solutions for treating existing illnesses because no U.S. therapeutics have been approved by the Food and Drug Administration (FDA) or the European Medicines Agency (EMEA). This implies we ought to have been prepared for a sickness that has been recognized for nearly four decades. The ongoing SARS-CoV-2 episode is now being shown to be incredibly expensive in terms of

death and financial repercussions. As of August 7, 2022, there had been over 200 million confirmed COVID-19 cases worldwide, with over four million deaths (WHO, https://covid19.who.int/) [1]. SARS-CoV-2 has been spreading swiftly throughout the world since its discovery in December 2019. SARS-CoV-2 transmits through respiratory droplets from one individual to another, high concentrations of aerosols, and, on rare occasions, feces or urine [1, 2]. Although the number of deaths caused by SARS-CoV-2 is steadily increasing, a

---

[1] Corresponding Authors

e-mail: eiedache@unimaid.edu.ng

Emmanuel Israel Edache, Adamu Uzairu, Paul Andrew Mamza, Gideon Adamu Shallangwa

powerful pharmacological treatment other than vaccines has not been developed at this time. Although immunizations are being investigated as a possible therapeutic approach for SARS-coronavirus-2, vaccines can be ineffective when strains are constantly changing. New strains of coronaviruses seem to be emerging constantly, necessitating the pursuit of various treatment strategies in addition to vaccinations [2-4]. The SARS-CoV-2 major protease (Mpro) is commonly considered a target in studies of new drug development, drug screening, and therapeutic repurposing because of its critical role in viral replication and transcription [3, 5]. Thus, inhibiting this enzyme may be a potential treatment for COVID-19. Several investigations into SARS-coronavirus-2 Mpro and its putative antagonists have been published in the existing literature [6-8]. According to the World Health Organization (WHO), more than a million people worldwide experience sexually transmitted illnesses regularly [9]. Infections such as Chlamydia trachomatis (C. trachomatis) can occur without causing any visible symptoms. In some cases, it is possible to be infected without realizing it [10, 11]. Any symptoms that do appear can be mistaken for those of other illnesses. C. trachomatis is known as a "silent" infection because many people are unaware that they have it [12]. Chlamydia was the most frequent sexually transmitted disease in 1994, according to the Centers for Disease Control and Prevention (CDC) [13]. In the United States alone, about 1.5 million cases of C. trachomatis were reported in 2015 [14]. A rate of 478.8 instances per 100,000 has been gradually increasing since 2001 [15]. C. trachomatis is multiple times more pervasive than syphilis and multiple times more predominant than gonorrhea [16]. Chlamydia has been documented to harm the reproductive system in severe and possibly irreversible ways [17]. If left untreated, chlamydia can result in (1) men suffering from chronic prostatitis, which causes pain and erectile dysfunction (ED); (2) a greater chance of contracting HIV/AIDS and SARS-CoV-2; and (3) infertility in women for the rest of their lives, as well as a painful condition known as a pelvic inflammatory disease [18-21]. Various causes of C. trachomatis and SARS-CoV-2 have been proposed. Because of the shortfall of medication that can eradicate SARS-CoV-2 and C. trachomatis,

numerous researchers all over the planet have zeroed in their examination on finding new medications that are productive against these sicknesses. Many C. trachomatis and SARS-CoV-2 antagonists have been documented in the literature [21-24]. Antibiotics and antivirals alone won't be enough to solve this issue because the setting for the spread of STIs and SARS-CoV-2 [25] varies widely based on social, economic, and geographic factors. Computer-Aided Drug Design (CADD) has emerged as a proficient method for recognizing potential lead compounds and helping the improvement of potential medications for many illnesses [26, 27]. The identification of prospective lead compounds from vast compound libraries is currently done using a variety of computational methods. Drug discovery applications using the CADD technique are improving constantly [28]. The current trend in drug design is to logically create powerful treatments with more targets, greater efficacy, and fewer side effects, particularly toxic ones. Computational modeling can help speed up the drug discovery process through simulation and modeling. Since theoretical research is progressing in the direction of rational drug design, understanding the relationship between physicochemical properties and molecular structure enables scientists to create new active molecules to a reasonable degree [29]. With the improvement of in silico techniques in recent years, the number of novel molecular entities endorsed by the U.S. Food and Drug Administration (FDA) has grown significantly One of the key strategies for modern pre-clinical drug discovery is computer-aided drug design (CADD), in which a variety of computational methods and software tools are often combined to produce the desired results. In the modern day, medicines and other physiologically active chemicals are discovered, developed, and analyzed using CADD methods including docking, MD simulation, and pharmacophore modeling perspectives. The CADD technique, which combines a structural and ligand-based pharmacophore model to identify similar active compounds against a specific target protein, may swiftly analyze a humongous compound's binding affinity to a target glycoprotein. A compound's bioactivity can be assessed anytime it attaches to a certain protein and causes a particular response. However, a molecular docking approach made the

Emmanuel Israel Edache, Adamu Uzairu, Paul Andrew Mamza, Gideon Adamu Shallangwa

process quicker. A CADD can predict pharmacokinetics and pharmacology properties such as absorption, distribution, metabolism, excretion (ADME), and even toxicity of a drug. Quantitative Structure-Activity Relationship (QSAR) modeling is a ligand-based drug design method for both exploring and exploiting the relationship between chemical structure and its biological activities. To predict the activities of anti-chlamydia trachomatis compounds, quantum chemical descriptors like molecular orbital, dipole moment, charge, etc. and molecular property descriptors like hydrophobic, steric coefficient, etc. have been applied to develop 2D QSAR models [30]. In the present research, we executed QSAR studies on some chemical libraries using genetic function approximation multiple linear regression (GFA-MLR). The best model out of the many generated model will be systematically analyzed. The results gained from these methods were equated for validation. This research primarily focused on CADD procedures such as virtual screening (docking simulation), ADMET, and MD simulation-MM/GBSA techniques to find potential oral medicines against C. trachomatis, as well as SARS-CoV-2 Mpro.

2. Computational Method
2.1 Data set

In this research, a band of 46 thiazolino-2-pyridone amide derivatives was obtained from the PubChem database (https://pubchem.ncbi.nlm.nih.gov/AID 1293489) and was chosen to build the QSAR models. These substances were created through exploration, and they display a variety of structural characteristics and biological activity levels as shown by their half-maximal inhibitory concentration (IC_{50}). This QSAR strategy depends on the speculation that the action Y (subordinate variable, pIC_{50}) is straightly reliant upon specific autonomous factors Xn (descriptors) as indicated by the following condition:

$$Y = aX1 + bX2 + cX3 + \ldots\ldots\ldots + k \quad (1)$$

Where $a, b, c,$ and so on are the coefficients of the descriptors X and k is the intercept in the linear QSAR model. The chemical names of the 46 thiazolino 2-pyridone amide derivatives investigated, as well as their pIC50 (bioactivities), are shown in Table 1. The IC_{50} action values were recently estimated in μM/mL, and afterward, we changed them to pIC_{50} with the usage of the subsequent expression pIC_{50} = -log IC_{50} to work with the computation. Table 1 shows the pIC_{50} results that were obtained. One of the chemical structure of thiazolino-2-pyridone amide derivatives is presented in Fig. 1.
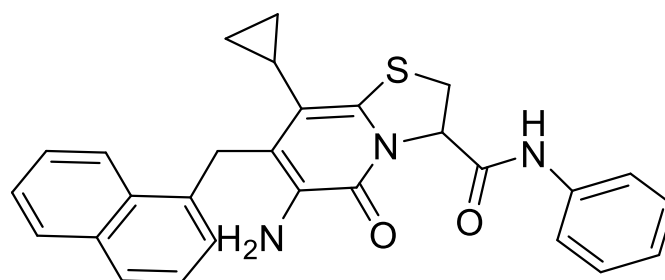

Figure 1. Molecular structure of thiazolino 2-pyridone amide.

Table 1. Compounds used in the generation of the 2D-QSAR models and docking simulation.			
S/N	Compound Name	pIC50	A*	B*
1	(S)-6-amino-7-(4-chlorobenzyl)-8-cyclopropyl-5-oxo-N-(p-tolyl)-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.6021	-7.1	-7.7
2	(S)-6-amino-8-cyclopropyl-7-(2,3-dimethylbenzyl)-5-oxo-N-(p-tolyl)-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	6.0000	-7	-7.5
3	(R)-8-cyclopropyl-6-(morpholinomethyl)-7-(naphthalen-1-ylmethyl)-5-oxo-N-phenyl-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.3010	-6.7	-8.3
4	8-cyclopropyl-7-(naphthalen-1-ylmethyl)-5-oxo-N-phenyl-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.6021	-7.5	-7.9

Emmanuel Israel Edache, Adamu Uzairu, Paul Andrew Mamza, Gideon Adamu Shallangwa

5	(S)-8-cyclopropyl-7-(naphthalen-1-ylmethyl)-5-oxo-N-phenyl-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.6021	-8.1	-8.4
6	(R)-8-cyclopropyl-7-(naphthalen-1-ylmethyl)-5-oxo-N-phenyl-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.6021	-7.1	-7.7
7	(S)-8-cyclopropyl-7-(naphthalen-1-ylmethyl)-5-oxo-N-phenyl-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.6021	-8.1	-8.5
8	(S)-8-cyclopropyl-N-(2-fluorophenyl)-7-(naphthalen-1-ylmethyl)-5-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.6021	-8.1	-7.7
9	(S)-8-cyclopropyl-7-methyl-5-oxo-N-phenyl-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.0000	-6.4	-6.9
10	(S)-3-(cyclohexylcarbamoyl)-8-cyclopropyl-7-(naphthalen-1-ylmethyl)-5-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-6-carboxylic acid	5.0000	-7.7	-7.7
11	(S)-6-amino-8-cyclopropyl-7-(naphthalen-1-ylmethyl)-5-oxo-N-phenyl-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	6.6201	-7.3	-8.2
12	(S)-6-amino-8-cyclopropyl-7-(naphthalen-1-ylmethyl)-5-oxo-N-(m-tolyl)-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	6.0000	-7.3	-8
13	(S)-6-amino-8-cyclopropyl-7-(naphthalen-1-ylmethyl)-5-oxo-N-(p-tolyl)-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	6.6201	-7.3	-7.9
14	(S)-N-cyclohexyl-8-cyclopropyl-7-(2,3-dimethylbenzyl)-5-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.6021	-7.7	-8
15	(S)-N-cyclohexyl-8-cyclopropyl-7-(2,3-dichlorobenzyl)-5-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.6021	-7.4	-7.9
16	(S)-N-cyclohexyl-8-cyclopropyl-7-(3,4-dimethylbenzyl)-5-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.6021	-7.1	-7.2
17	(S)-N-cyclohexyl-8-cyclopropyl-5-oxo-7-(quinolin-5-ylmethyl)-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.0000	-8.1	-8.2
18	(S)-8-cyclopropyl-7-(2,3-dimethylbenzyl)-5-oxo-N-(p-tolyl)-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	6.0000	-7	-8
19	(S)-7-(4-chlorobenzyl)-8-cyclopropyl-5-oxo-N-(p-tolyl)-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	6.0000	-7.1	-6.4
20	(S)-6-amino-8-cyclopropyl-7-(naphthalen-1-ylmethyl)-5-oxo-N-(pyrimidin-4-yl)-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.6021	-7.6	-7.9
21	(S)-7-benzyl-N-cyclohexyl-8-cyclopropyl-5-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.6021	-7	-7.5
22	(S)-N-cyclohexyl-8-cyclopropyl-7-(3-methylbenzyl)-5-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.6021	-7.1	-7.8
23	(S)-7-((4-chlorocyclohexyl)methyl)-8-cyclopropyl-5-oxo-N-phenyl-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.6021	-6.9	-6.3
24	(S)-N-cyclohexyl-8-cyclopropyl-7-(4-methoxybenzyl)-5-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.6021	-7	-7.4
25	(S)-8-cyclopropyl-N-(3-fluorophenyl)-7-(naphthalen-1-ylmethyl)-5-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.6021	-8	-8.6
26	(S)-N-(3-chlorophenyl)-8-cyclopropyl-7-(naphthalen-1-ylmethyl)-5-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.6021	-7.5	-8.6
27	(S)-8-cyclopropyl-7-(naphthalen-1-ylmethyl)-5-oxo-N-(m-tolyl)-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	6.0000	-7.3	-8.5
28	(S)-8-cyclopropyl-N-(3-ethylphenyl)-7-(naphthalen-1-ylmethyl)-5-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.6021	-7.1	-8.4
29	(3S)-8-cyclopropyl-7-(naphthalen-1-ylmethyl)-5-oxo-N-(3-(trifluoromethyl)cyclohexyl)-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.6021	-7.6	-8.6
30	(S)-8-cyclopropyl-N-(3-methoxyphenyl)-7-(naphthalen-1-ylmethyl)-5-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.6021	-8.2	-8.2

Emmanuel Israel Edache, Adamu Uzairu, Paul Andrew Mamza, Gideon Adamu Shallangwa

31	(S)-8-cyclopropyl-N-(3-(methylcarbamoyl)phenyl)-7-(naphthalen-1-ylmethyl)-5-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.0000	-8.1	-8.5
32	(S)-8-cyclopropyl-N-(4-fluorocyclohexyl)-7-(naphthalen-1-ylmethyl)-5-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.6021	-7.4	-8.6
33	(S)-N-(4-chlorophenyl)-8-cyclopropyl-7-(naphthalen-1-ylmethyl)-5-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.6021	-7.5	-8.4
34	(S)-8-cyclopropyl-7-(naphthalen-1-ylmethyl)-5-oxo-N-(p-tolyl)-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.6021	-7.6	-8.5
35	(S)-8-cyclopropyl-N-(4-methoxyphenyl)-7-(naphthalen-1-ylmethyl)-5-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.6021	-7.3	-8.6
36	(S)-N-(4-carbamoylphenyl)-8-cyclopropyl-7-(naphthalen-1-ylmethyl)-5-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.6021	-7.4	-8.7
37	(S)-8-cyclopropyl-7-(naphthalen-1-ylmethyl)-5-oxo-N-(4-sulfamoylphenyl)-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.0000	-7.6	-8.3
38	(S)-8-cyclopropyl-7-(naphthalen-1-ylmethyl)-5-oxo-N-(pyridin-2-yl)-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.0000	-7.7	-8.4
39	(3S)-8-cyclopropyl-7-(naphthalen-1-ylmethyl)-5-oxo-N-(1l2-piperidin-3-yl)-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.0000	-7.9	-8.3
40	(S)-8-cyclopropyl-7-(naphthalen-1-ylmethyl)-5-oxo-N-(pyridin-4-yl)-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.3010	-7.7	-8.3
41	(S)-8-cyclopropyl-7-(naphthalen-1-ylmethyl)-5-oxo-N-(pyrimidin-4-yl)-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.6021	-7.1	-8.4
42	(S)-8-cyclopropyl-7-(naphthalen-1-ylmethyl)-5-oxo-N-(thiazol-2-yl)-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.3010	-7.2	-8.2
43	(3S)-8-cyclopropyl-7-((decahydronaphthalen-1-yl)methyl)-N-methyl-5-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.0000	-7.1	-7.3
44	(S)-N-cyclohexyl-8-cyclopropyl-7-(naphthalen-1-ylmethyl)-5-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.0000	-7	-8
45	(S)-N-benzyl-8-cyclopropyl-7-(naphthalen-1-ylmethyl)-5-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.6021	-7.4	-7.9
46	(S)-8-cyclopropyl-N-methyl-7-(naphthalen-1-ylmethyl)-5-oxo-N-phenyl-2,3-dihydro-5H-thiazolo[3,2-a]pyridine-3-carboxamide	5.3010	-7.3	-7.6
Reference Drug				
47	(2R,5R,6S)-6-((S)-2-amino-2-cyclohexylacetamido)-3,3-dimethyl-7-oxo-4-thia-1-azabicyclo[3.2.0]heptane-2-carboxylic acid	Ampicillin	-6	-7.2
48	(2R,3S,4R,5R,8R,10S,11R,12R,13S,14R)-11-(((2S,3R,4S,6R)-4-(dimethylamino)-3-hydroxy-6-methyltetrahydro-2H-pyran-2-yl)oxy)-2-ethyl-3,4,10-trihydroxy-13-(((2R,4R,5S,6S)-5-hydroxy-4-methoxy-4,6-dimethyltetrahydro-2H-pyran-2-yl)oxy)-3,5,6,8,10,12,14-heptamethyl-1-oxa-6-azacyclopentadecan-15-one	Azithromycin	-5.7	-6.3
49	(4S,4aR,5S,5aR,6R,12aR)-4-(dimethylamino)-1,5,10,11,12a-pentahydroxy-6-methyl-3,12-dioxo-3,4,4a,5,5a,6,12,12a-octahydrotetracene-2-carboxamide	Doxycycline (Vibramycin)	-6.7	-8
50	(R)-2-(ethyl(4-((7-methyl-1,2-dihydroquinolin-4-yl)amino)pentyl)amino)ethan-1-ol	Hydroxychloroquine	-5.5	-6.6

Emmanuel Israel Edache, Adamu Uzairu, Paul Andrew Mamza, Gideon Adamu Shallangwa

51	(R)-N-((2R,4R,5R)-6-cyclohexyl-5-(2-(2,6-dimethylphenoxy)acetamido)-4-hydroxy-1-phenylhexan-2-yl)-3-methyl-2-(2-oxotetrahydropyrimidin-1(2H)-yl)butanamide	Lopinavir	-6	-7.2
52	(2,3-dihydrothiazol-5-yl)methyl ((2S,3R,5S)-3-hydroxy-5-((2S)-2-(3-((2-isopropyl-2,3-dihydrothiazol-4-yl)methyl)-3-methylureido)-3-methylbutanamido)-1,6-diphenylhexan-2-yl)carbamate	Ritonavir	-7.7	-8.7
53	(3R)-3-cyclopentyl-3-(4-(2,3,7,7a-tetrahydro-1H-pyrrolo[2,3-d]pyrimidin-4-yl)-2,5-dihydro-1H-pyrazol-1-yl)propanenitrile	Ruxolitnib	-7.1	-8

A* C. Trachomatis (Binding affinity kcal/mol), B* SARS-CoV-2 main protease (Binding affinity kcal/mol)

2.2 Compounds optimization and selection of descriptors

The data set descriptors were obtained by optimizing the compounds in Gaussian 09 [31] with the B3LYP/6-31G (d,p) [32, 33] basic set. In the program PaDEL freeware v2.21 [34], the descriptors were created using the improved structures of all compounds. An enormous number of descriptors were determined and the factors were pre-filtered by: (1) physically erasing every one of the missing qualities; (2) the correlation coefficients (R) between every descriptor and the pIC50 values were determined to prohibit the descriptors with low connection with biological activity. (3) barring if steady qualities > 80%; and (4) excluding if correlations > 95% using the QSARINS v2.2.4 program [35]. We partitioned the data set of 46 compounds into two sets. The test set was chosen utilizing an irregular determination strategy by setting a likelihood opportunity of around 30%. The dataset was divided into 33 and 13 molecules in the training and test sets, respectively. Then, the genetic algorithm (GA) technique as executed in QSARINS v2.2.4 [35-37] was applied as a secondary screening to get the descriptors that generally added to the biological activity. Afterward, measurable boundaries like "SEC (standard error of calibration), PRESS (prediction error sum of squares), coefficient of determination (R^2), adjusted coefficient of determination (R_adj^2), Mean of Squared Errors of a model (MSE), Fischer's value (F-test), Variance Inflation Factor (VIF), coefficient of determination of Leave-One-Out cross-validation (Q_LOO^2), coefficient of determination of external test (R_test^2), and Y-randomization parameters (R_Yscr^2 and Q_Yscr^2)" for modeling, internal and external validation measures were used to assess the fitness, steadiness, and prescient force of the QSAR model as per their conditions in Table 2.

Table 2. Comparison of the statistical parameter with Golbraikh and Tropsha criteria [35-38]			
Criteria	Parameter	Equation	Threshold
Fitting criteria	R^2	$R^2 = 1 - \dfrac{\sum(Y_{obs} - Y_{cal})^2}{\sum(Y_{obs} - \bar{Y}_{cal})^2}$	➢ 0.6
	R^2_{adj}	$R^2_{adj} = \dfrac{(N-1)(R^2 - P)}{N - P - 1}$	➢ 0.6
	MSE	$MSE = \dfrac{\sum(Y_{obs} - Y_{cal})^2}{N}$	*A low value*
	F	$F = \dfrac{\sum(Y_{cal} - \bar{Y}_{cal})^2 - N - P - 1}{\sum(Y_{obs} - Y_{cal})^2 \times P}$	*A high value*
Internal validation criteria	Q^2_{LOO}	$Q^2_{Loo} = 1 - \dfrac{\sum(Y_{cal} - Y_{obs})^2}{\sum(Y_{obs} - \bar{Y}_{obs})^2}$	➢ 0.5
	R_{Yscr}	*Average of 100 R_{Yscr}*	$< R$
	R^2_{Yscr}	*Average of 100 R^2_{Yscr}*	$< R^2$
	Q^2_{Yscr}	*Average of 100 Q^2_{Yscr}*	$< Q^2_{LOO}$
	cR^2_p	$cR^2_p = R\sqrt{R - (Average\ R_{Yscr})^2}$	➢ 0.5

Emmanuel Israel Edache, Adamu Uzairu, Paul Andrew Mamza, Gideon Adamu Shallangwa

External validation criteria	R^2_{Pred}	$R^2_{Pred} = 1 - \dfrac{\sum (Y_{obs(test)} - Y_{test})^2}{\sum (Y_{obs(test)} - \bar{Y}_{obs(test)})^2}$	➢ 0.5
	$\bar{r}_{m(test)}$	$\bar{r}_{m(test)} = \dfrac{\lvert r^2_m + r'^2_m \rvert}{2}$	➢ 0.5
	Δr^2_{test}	$\Delta r^2_{test} = \lvert r^2_m + r'^2_m \rvert$	< 0.3
	$\Delta r^2_{o(test)}$	$\Delta r^2_{o(test)} = \lvert r^2_o + r'^2_o \rvert$	< 0.3
	$\dfrac{r^2 - r^2_o}{r^2}$	$\dfrac{r^2 - r^2_o}{r^2}$	< 0.1
	$\dfrac{r^2 - r'^2_o}{r^2}$	$\dfrac{r^2 - r'^2_o}{r^2}$	< 0.1
	K	$K = \dfrac{\sum Y_{obs} Y_{cal}}{\sum Y^2_{cal}}$	$0.85 \le K \le 1.15$
	K'	$K' = \dfrac{\sum Y_{obs} Y_{cal}}{\sum Y^2_{obs}}$	$0.85 \le K' \le 1.15$

$Y_{obs}, Y_{cal}, \bar{Y}_{obs}, \bar{Y}_{cal}, N, and\ P$: refers to the observed, predicted, mean of the observed, mean of the predicted, number of compounds, and number of descriptors, respectively.

2.3 Homology modeling

The inclusion membrane protein IncA (C. trachomatis) and SARS-coronavirus-2 Mpro sequence used in this study were obtained from the National Center for Biotechnology Information (https://www.ncbi.nlm.nih.gov/) database in the Fasta format.

2.3.1 Templates searching and modeled C. trachomatis and SARS-CoV-2 protein quality

The amino acid sequence of inclusion membrane protein IncA (C. trachomatis) (Query ID: WP_057222102.1) [39] and Chain A, SARS-coronavirus-2 Mpro (Query ID: 6YZ6_A) has been uploaded from further comparative modeling. The BLASTp [40] program was used to find the template protein with the highest self-image in the query sequence. These templates were downloaded from the PDB (https://www.rcsb.org/) database. Phylogenetic analysis is used to establish the evolutionary relationships between the query and template protein sequences. The Molecular Evolutionary Genetics Analysis (MEGA X) program was used for phylogenetic analysis [41]. The results of an analysis can be obtained in the form of a phylogram. The UPGMA method [42] was used to infer the phylogenetic analysis. The evolutionary distances were calculated using the Poisson correction method [43] and are measured in amino acid substitutions per site. The comparative models of C. trachomatis and SARS-coronavirus-2 were built with MODELLER v10.2 [44, 45]. The accuracy of the designed homology model was validated using various techniques on the SAVES server

(http://nihserver.mbi.ucla.edu/SAVES/). The validation techniques, such as Verify3D [46, 47], which decides the similarity of an atomic model (3D) with its amino acid arrangement (1D), and Rampage to analyze the Ramachandran plot [48, 49], ERRAT [50], and PROSA server to calculate Z-score [51], were carried out.

2.4 Molecular docking preparations

After valid 3D coordinates of the modeled receptors were obtained from homology modeling using MODELLER v10.2, and a few inherited problems to the modeled 3D receptors (for example, missing side-chain, loop, and hydrogens), The receptors shown were arranged using the Protein Planning Wizard module from the AutoDookTools v1.5.7 suite. The AutoDookTools v1.5.7 suite is programmed to remove water molecules from the active site, fill in the missing side chain, add hydrogens, and assign charges.

2.5 Active Site Prediction

After preparing the protein with AutoDockTools v1.5.7, it is minimized in 500 steps using the steepest descendent algorithm, which is used by GROMOS96 43B1 parameters installed in DeepView v4.1.0. The active site residues of the modeled receptors are predicted using compute pocket (AutoSite 1.1) implemented in AGFR v1.2 [52] to be recognized for docking research. These active site residues must be predicted, and the binding region is chosen for derivative docking based on the pocket AS score with the highest value. Here, the active/binding site of the modeled

Emmanuel Israel Edache, Adamu Uzairu, Paul Andrew Mamza, Gideon Adamu Shallangwa

receptors is anticipated and docked with 46 thiazolino-2-pyridone amide derivatives plus 7 controls/standard drugs for the modeled receptors' activity. Table 3 displays the AS score, radius of

gyration, buriedness, and the position of the grid of each protein. Analysis of specific docking results was carried out using Discovery Studio visualizer 2020.

Table 3. Grid setting for specific docking.

Protein	AS Score	RadGr	Buriedness	Grid	
				Center	Dimensions
C. trachomatis	364.85	7.53	0.75	-39.001	22.500
				39.513	32.250
				31.111	19.500
SARS-CoV-2	430	6.60	0.81	10.392	23.250
				-0.600	27.000
				22.438	18.000

2.6 Pharmacokinetics properties

The pharmacokinetics of the best-docked ligands were calculated by the webserver SwissADME server (www.swissadme.ch). The SwissADME server is a free online tool that predicts pharmacokinetics, human intestinal absorption (HIA), bioavailability, and blood-brain barrier (BBB) permeability with a quick, accurate, and user-friendly interface.

2.7 Molecular Dynamics Simulations

Following a successful docking protocol, nanoscale molecular dynamics (NAMD) v2.14 [53] was used to conduct MD simulations on the best complexes. The production run was conducted in an NPT ensemble for 1 ns with no constraints. For 10,000 steps in the NVT ensemble, MD simulations of the two complexes (i.e., the C. trachomatis and SAR-CoV-2 complexes) were run in a solvent environment. The temperature of the simulation system was set to 310 K. Energy minimization for the complexes was based on the RMSD variation of $C\alpha$, RMSF, SASA, and the RG plot. The visual molecular dynamics (VMD) software v1.9.3 [54] was used to display, analyze, and animate trajectories visually. The molecular mechanics generalized Born surface area (MM/GBSA) free binding energies of ligand-receptor complexes were calculated using MolAlCal software [55].

3. Results and discussion

3.1 Structure-Activity Relationship for Thiazolino 2-pyridone amide Derivatives

The following descriptors are chosen as the most suitable descriptors for developing the QSAR model through the multiple linear regression (MLR) techniques after performing the genetic approximation (GA) analyses: nAromBond, ATSC3v, MATS8m, and SdO. The correlation matrix found in Table S1 displays the following four descriptors along with the correlation

coefficients between them. The generated QSAR model 1 showed a strong correlation between the chemical descriptors and bioactivity. The cross-validation regression coefficient (Q LOO2 = 0.7047) indicates that this QSAR model has a 70% prediction accuracy, while the regression coefficient ($R2 = 0.7772$) indicates that there is a 78% correlation between the bioactivity and chemical descriptors in the training dataset.

$pIC_{50} = 0.0749$ **(nAromBond) – 0.0011 (ATSC3v) – 1.6109 (MATS8m) – 0.0272 (SdO) + 5.0927 Model 1**

where "nAromBond = Number of aromatic bonds, ATSC3v = Centered Broto-Moreau autocorrelation – lag 3/weighted by van der Waals volumes, MATS8m = Moran autocorrelation - lag 8/weighted by mass, and SdO = Sum of atom-type E-State: =O."

Fitting criteria:

$R^2 = 0.7772$, $R^2_{adj} = 0.7453$, R^2-$R^2_{adj} = 0.0318$, LOF = 0.0505, Kxx = 0.2779, Delta K = 0.0994, RMSE tr = 0.1702, MAE tr = 0.1191, RSS tr = 0.9556, CCC tr = 0.8746, s = 0.1847 F = 24.4147

Internal validation criteria:

$Q^2_{LOO} = 0.7047$, R^2-$Q^2_{LOO} = 0.0724$, RMSE cv = 0.1959, MAE cv = 0.1404, PRESS cv = 1.2662, CCC cv = 0.8353, $Q^2_{LMO} = 0.6487$, $R^2_{Yscr} = 0.1254$, $Q^2_{Yscr} = -0.2482$, RMSE AV $_{Yscr} = 0.3367$

External validation criteria:

RMSE ext = 0.3706, MAE ext = 0.2539, PRESS ext = 1.7854, Q^2-$F_1 = 0.1383$, Q^2-$F_2 = 0.1116$, Q^2-$F_3 = -0.0568$, CCC ext = 0.5221

The Fisher test (F-test), the mean squared error (MSE), and the coefficient of determination (R2) of the 2D-QSAR model developed by using the GA-MLR based on fitting criteria evaluation were precisely evaluated, and we obtained a lower root mean square error MSE = 0.1191 indicating that the model is more reliable. The cross-validation correlation coefficient Q LOO2, shown in Table S2,

Emmanuel Israel Edache, Adamu Uzairu, Paul Andrew Mamza, Gideon Adamu Shallangwa

indicates the model's robustness and accuracy. The P-value (lower than 0.05) shown in Table S2 indicates that the model acquired is statistically significant at the 95% level. These descriptors' VIF values in the training set were calculated and found to be less than 10 (Table S2), demonstrating the independence of the identified model's fitness from the descriptors [56]. To inspect the general significance and commitment of the descriptor with pIC50 values. We work out the mean impact (MF) as per the following condition:

$$MF_j = \frac{\beta_j \Sigma_i^n d_{ij}}{\Sigma_j^m (\beta_j \Sigma_j^n d_{ij})} \underline{\hspace{3cm}} (2)$$

Where MF_j and β_j address the mean effect of every descriptor j and the coefficient of the descriptor j, separately. d_{ij} and m allude to the value of proposed descriptors for each compound and the number of descriptors.

The mean effect recognizes important data of the descriptors model created for determining the activity of compounds and how strong is the model [57, 58]. The involvement of a descriptor in comparison to other descriptors in the model is indicated by the MF value (Table S2). The indicators suggest that the values of these descriptors will either increase or decrease, improving the anti-C. trachomatis inhibitory activities of the study molecules [59, 60]. A positive sign in the model for nAromBond (and more or less relevant descriptor). The activity (pIC50) and the descriptor (MF = 2.261) are positively correlated. It is presumptively true that an increase in aromatic bonds of 10% will increase the compound's inhibitory activity. On the other hand, a 44% reduction in ATSC3v values will result in a rise in compound (anti-C. trachomatis) activities. The inhibitory activities of the molecules are most strongly influenced by MATS8m with the highest MF value. It has been hypothesized that MATS8m, the Moran autocorrelation - lag 8/weighted by mass, which has a positive MF value, positively affects anti-C. trachomatis activity. 33% of this descriptor will lead to an increase in anti-C. trachomatis inhibitory activity. The 2D descriptor SdO represents the Sum of atom-type E-State: =O this descriptor contributes negatively to MF value (Table S2) which indicates the inhibitory activity of thiazolino 2-pyridone amide derivatives will increase if the Sum of atom-type E-State: =O is reduced by 14%. In summary of the mean effect (MF), The importance of the pIC50 values of the compounds will be outstanding because MATS8m is the most significant descriptor and has the best mean effect value. The descriptors are listed below in order of how they affect pIC50 values. ATSC3v > MATS8m > nAromBond > SdO. The comparison

and consistent distribution of the experimentally determined pIC50 activity values and the values predicted by the GA-MLR model are displayed in Fig. S1. The experimental and predicted activity values are well correlated, as shown by the clustering of points along the line of best fit, demonstrating the robustness and dependability of the built model. The final model's inter-correlation between the descriptors and the bioactivities was plotted with Kxy versus Q2 LMO, and the LMO (leave many out) parameter values were shown around the model parameters in Fig. S2, indicating that the model was stable and reliable. The observed model was not randomly obtained, as shown by the VIF provided in Table S2 and Kxy vs. Q2 LMO in Fig. S2, and the structural dependence of the training set supported the reliability of the model. As shown in Table S3, the GA-MLR model's parameters, including the compounds' status, had a significant impact on how each studied compound's chemical data were converted into numerical values. Additionally, the XY (scatter) plot of noticed pIC50 against the residuals of both the training and test sets was displayed in Fig. S3. The sporadic appearance of these residuals on one or the other side of zero on the plot shows that the model is liberated from regular error. The estimations performed were rehashed with randomized exercises of the mixtures in the training set, to assess the heartiness of the GA-MLR model. Ten (10) iterative runs were performed to recreate 10 MLR models indiscriminately. In this assessment, we produced MLR models by arbitrarily blending the dependent variable while keeping the independent variables unaltered [61]. R2 and Q LOO2 values for the new models are lower than for the original model. This shows that the original GA-MLR model is stable and that its strong statistical properties are not accidental. As indicated in the results of Table S4, all new models cannot match the original model. Furthermore, a value of CRp^2 = 0.72 criteria is superior to 0.5; indicating high confidence in the original GA-MLR model. In light of the examination of the outcomes, we can presume that the 2D-QSAR models we have created are extremely effective and well-equipped for anticipating new subordinates of thiazolino-2-pyridone amide. The Y-scramble plot (Fig. S4) of Kxy versus R2Yscr was used to test the external validity, and Q2Yscr displayed a lower value compared to the values in the model. The leveraged approach evaluated the model's structural applicability domain and provided a cut-off hat value. The diagonal elements of the HAT matrix are used to calculate HAT values:

$$H = X(X^T X)^{-1} X^T \underline{\hspace{2cm}} (3)$$

Emmanuel Israel Edache, Adamu Uzairu, Paul Andrew Mamza, Gideon Adamu Shallangwa

Where the leverage value ((h*) can be calculated according to the following equation:

$$h^* = \frac{3g'}{n} \underline{\hspace{3cm}} \quad (4)$$

where g' is the number of model variables plus one, and n is the number of objects used to calculate the model.

The standardized residuals in cross-validation greater than 2.5 standard deviations (δ) units can be used to confirm the response applicability domain. Both structural outliers (i.e., compounds with a high leverage value (H > h*)) and response outliers (i.e., compounds with cross-validated standardized residuals greater than 2.5 standard deviations (δ) units) have been defined using the fixed thresholds. The Williams plot for the developed QSAR models is presented in Fig. S5. According to the plot, it is observed that the leverage value (h*) dashed lines were found to be 0.455 for the developed QSAR models, which indicated the presence of structural (or influential) and response outliers. The William plot shows that all molecules are within the applicability domain (AD), except compound 37 from the training set is considered an outlier (structural outlier) because it exceeds the critical leverage (h*), compounds 11, 13, and 15 are considered response outliers because they have a normalized residual outside the 2.5 δ units. As a result, the QSAR model correctly predicted the 42 remaining compounds when they were tested in the applicability domain, even for these chemicals (37), whose h values are beyond h*, the predicted pIC_{50} is close to their experimental values as seen in Table S3. The Insubria graph enables the identification of the compounds for which the model interpolates or extrapolates its predictions. The recheck of the applicability domain (William's plot) of the built model is possible through the Insubria graph. The results of the Insubria graph study for the model are graphically reported in Fig. S6. The leverage for the prediction set is plotted against the predicted values in the graph. By using the Insubria graph, we were able to define the model's reliable prediction zone based on the structural similarity to the training compounds (leverage value) and the predicted biological activity value. If both of the following conditions hold— H < h* and Y_min < Y_pred < Y_max (Y_min and Y_max are the minimal and the maximal value of bioactivity in the training set)— then the predicted results are believed to be accurate. Except for compounds 15, 37, 43, and 44, every compound from the prediction set was found to be within the applicability domain of the model. This demonstrates that the model obtained in this work has high applicability to design new thiazolino 2-pyridone amide derivatives, and we can apply them to screen and prioritize them for future experiments or for filling the data gap.

3.1.1 Galbraith and Tropsha's criteria
The outcomes of model 1 were contrasted with the acceptable limit thresholds established by Golbraikh and Tropsha (Table 1) [38, 62], and they also met the requirements for standard validation as set forth by the OECD. The outcomes of model 1 demonstrated the validity and suitability of our suggested model.

3.2 Homology modeling results
A comparative search against the NCBI database yielded two highly similar protein sequences: 6E7E.pdb and 6E6A.pdb. The query cover of 6E7E.pdb and 6E6A.pdb, predicted from the deduced amino acid sequences, is 58% each for C. trachomatis. To investigate the protein sequence evolutionary relationship with the query sequence (Query ID: WP_057222102.1) using MEGA X software. The phylogenetic tree was obtained based on previously obtained multiple sequence alignment using the MEGA-X phylogeny option using the UPGMA method. By analyzing the tree (Fig. 2), we can take observations. However, 6E7E.pdb and 6E6A.pdb showed the highest sequence similarity to the Inclusion membrane protein sequence. The cluster gave bootstrap values higher than 60. Among 2 Inclusion membrane protein sequences (6E7E.pdb and 6E6A.pdb), the query WP 057222102.1 inclusion membrane protein IncA C. trachomatis was found to be closely linked with 6E7E.pdb and 6E6A.pdb sequence (Fig. 2).

3.2.1 Sequence analysis and domain prediction of C. Trachomatis
Since both templates are closely related to the query sequence, the 6e6a.pdb crystal structure E-value 9e-111 was selected as the best template. The 3D structure of the constructed PDB protein structure was generated using MODELLER v10.2. A total of 5 models were generated by the MODELLER software (http://www.salilab.org/modeler), which was further evaluated by Structural Analysis and Verification Server (SAVES) (http://nihserver.mbi.ucla.edu/SAVES/) and the two last parameters were determined using Protein Structure Analysis (ProSA)-web server (https://prosa.services.came.sbg.ac.at/prosa.php) [51]. In this study, model 2 returned the highest DOPE score (Table 4). The matchmaker 6e6a.pdb, chain A (template) with WP_057222102.1, chain A (query), sequence alignment score = 796.7. The root means square deviation (RMSD) between 162 pruned atom pairs is 0.224 angstroms (Å) (across all 163 pairs: 0.281) (Fig. 3).

Emmanuel Israel Edache, Adamu Uzairu, Paul Andrew Mamza, Gideon Adamu Shallangwa

The sequence alignment between the inclusion membrane protein IncA (WP_057222102.1) and triclinic crystal form of IncA G144A point mutant (PDB ID: 6E6A) receptor was produced using MODELLER v10.2 aligned program (Fig. 4). The constraints are typically met by presuming that aligned residues in the template and target proteins are aligned at similar distances and angles. A homology model of the inclusion membrane protein IncA is the process output.
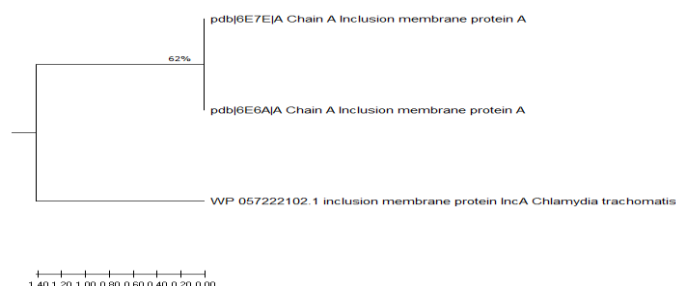
Figure 2. Phylogenetic tree of C. trachomatis protein and its selected homologs.

Table 4. The top five protein models, along with their parameter scores, were retrieved from MODELLER.

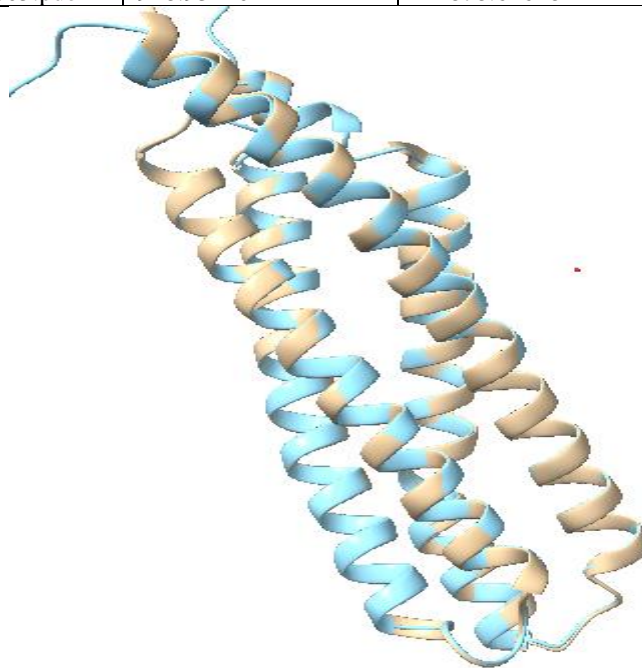| Filename	molpdf	DOPE score	GA341 score
ABCD.B99990001.pdb	917.60547	-24656.01758	1.00000
ABCD.B99990002.pdb	887.83081	-24999.32617	1.00000
ABCD.B99990003.pdb	947.85773	-24901.06055	1.00000
ABCD.B99990004.pdb	1047.82397	-24697.75391	1.00000
ABCD.B99990005.pdb	928.93420	-24878.64648	1.00000

Figure 3. Model template alignment: ChimeraX v1.3 superimposed WP 057222102.1 on the template (PDB ID: 6e6a.pdb), and the RMSD was 0.224. The target is highlighted in gray, while the template structure is displayed in green.

The PROCHECK server provided the Ramachandran plot (http://services.mbi.ucla.edu/SAVES/Ramachandran/) [63] for the comparative modeling protein. This plot's purpose is to validate the protein model that was found. The concealing on the plot addresses the various locales as displayed in Fig. S7. The haziest regions in red relate to the "center"

Emmanuel Israel Edache, Adamu Uzairu, Paul Andrew Mamza, Gideon Adamu Shallangwa

locales showing the most positive mixes of phi-psi values. Preferably, one would expect to have more than 90% of the buildups in these "center" locales. One of the better predictors of the quality of a stereo-chemical is the level of deposits in the "center" locations. According to Fig. S7, the produced model—the one with the most build-ups of residues in the permitted area and the fewest deposits in the rejected locale—was chosen as the best fit model for further investigation of the protein that was demonstrated. Since 97.2% of the deposits are placed in the ideal location on the plot, it follows that the local proteins' conformational components correspond to those of the similarly native protein. The number of residues present in the allowed region is 7 (2.8%) and the number of residues in generously allowed regions is 0 (0.0%) and residues in disallowed regions 0 (0.0%). The phi-psi torsion angles for each residue in the structure are displayed on the Ramachandran plot (except those at the chain termini). Triangles are used to distinguish glycine residues since they are not confined to the plot areas designated for the other side chain types. The ERRAT (overall quality factor) [50] program was used to calculate the non-bonded interactions between different types of atoms, and the analysis indicated an overall quality

factor of 85.052%. (Fig. S8). Yellow areas on the structure indicate areas that can be ruled out with a 95% confidence level. A good protein structure is expected to have an error value of 5% or less. Red areas represent those that can be rejected at a 99% level. An acceptable protein environment is indicated by the analysis of the ERRAT graph. The loading quality of each residue was then tested, as determined by the Verify-3D [45] method (http://services.mbi.ucla.edu/Verify3D/), which reflects the profile obtained concerning the residues. According to Sahu and Shulkla's discussion of acceptable side-chain environments (Fig. S9), the compatibility score in the Verify-3D graph above zero corresponds to these environments. The Dope per residue score vs alignment position plot generated by MODELLER in Fig. 5 shows that the inclusion membrane protein IncA (query ID: WP_057222102.1)) energy graph (red color) lies exactly on one chain of the triclinic crystal form of IncA G144A point mutant template (PDB ID; 6E6A) energy graph peaks (green color) which show primary similarity and similitude. This implies that the model has overall self-consistency regarding the relationship between sequence and structure.
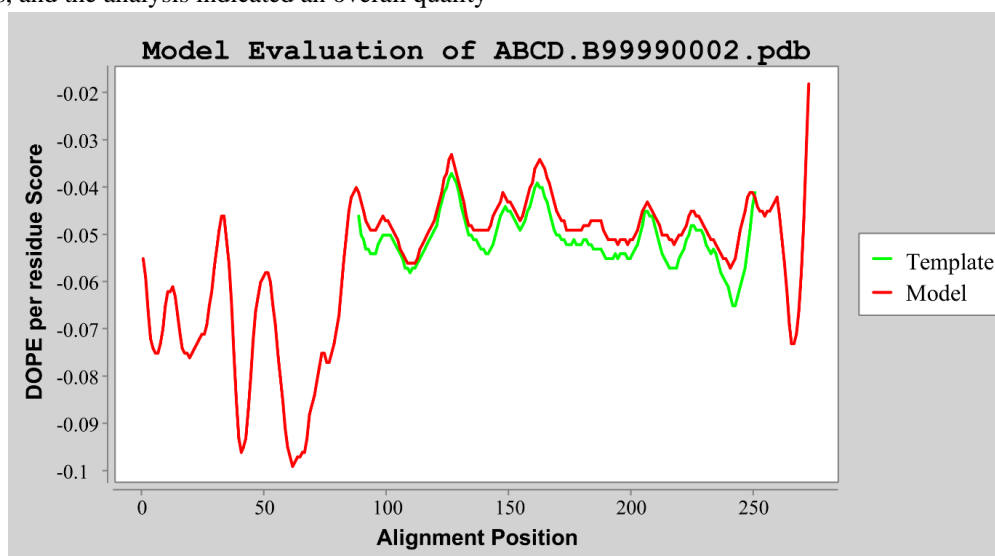

Figure 5. Dope profiles of Chlamydia trachomatis model (shown in red color) and template 6E6A.pdb (shown in green color).

To identify errors in theoretical and experimental models, the ProSA server was employed. ProSA assesses the model packing by calculating the likelihood of finding residues at a given distance and assesses the degree of solvation—the

interaction between the model and the solvent-that took place. The quality and dependability of the generated model are determined and assessed based on the total of all these probabilities. It is used to look for potential errors in the 3D-modeled protein.

Emmanuel Israel Edache, Adamu Uzairu, Paul Andrew Mamza, Gideon Adamu Shallangwa

A Z-score and a plot of the input 3D model's residue energies are two features that the program shows. The Z-score assesses the general model quality and gauges how far the structure's total energy deviates from an energy distribution derived from all of its random conformations. The modeled protein's overall model quality is shown by the ProSA Z-score of -4.42 (Fig. S10). The score demonstrates an exceptionally solid construction and is well within the scope of scores normally found for proteins of comparable size. The energy plot (Fig. S11) shows the nearby model quality by plotting knowledge-based energies as a component of amino acid sequence position. This demonstrates that the model built by the MODELLER is extremely reliable and deserving of consideration for further research.

3.3 Sequence analysis and domain prediction of SARS-CoV-2 main protease

The phylogenetic analysis of the query sequence and the templates were carried out using the UPGMA method and computed using MEGA X [65]. The studies revealed the evolutionary development of both the query sequence (PDB ID: 6YZ6) amongst the various templates. Multiple sequence alignment using MUSCLE was used to create a phylogram, which showed that 7KFI and 6YZ6 were strongly linked to a conserved theoretical protein from 3C-like proteinase of SARS-CoV-2 of 7CB7 and 6XA4 (Fig. 6).

MODELLER v10.2 built an incredible homology model of SARS-CoV-2 given the grouping character between the chain of The main protease (Mpro) SARS-CoV-2 (PDB: 6YZ6) and the template protein sequences. The dendrogram tree generated by MODELLER as shown in Fig. 7 confirmed the phylogenetic analysis above. The protein (PDB ID: 6XA4) has the least values, as a result, the 3D crystal structure of the SARS-CoV-2 3C-like proteinase (PDB:6XA4) with a resolution of 1.65 angstroms was chosen as a template. It has an E-value of 0.0, and 100% query identity and percentage identity with the query SARS-CoV-2 Mpro (PDB: 6YZ6) amino acids suggesting that they both have the same 3D structure. The template's sequence alignment (PDB ID: 6XA4) revealed the highest preservation of amino acid residues between dissimilar species in PDB ID: 6YZ6 (Fig. 8). As the least disruptions may have taken place throughout their phylogenetic origin, where it discovered their essential functional role, the solid preservation of core residues in the vicinity of the PDB ID: 6YZ6 domain region explains their evolutionary significance.
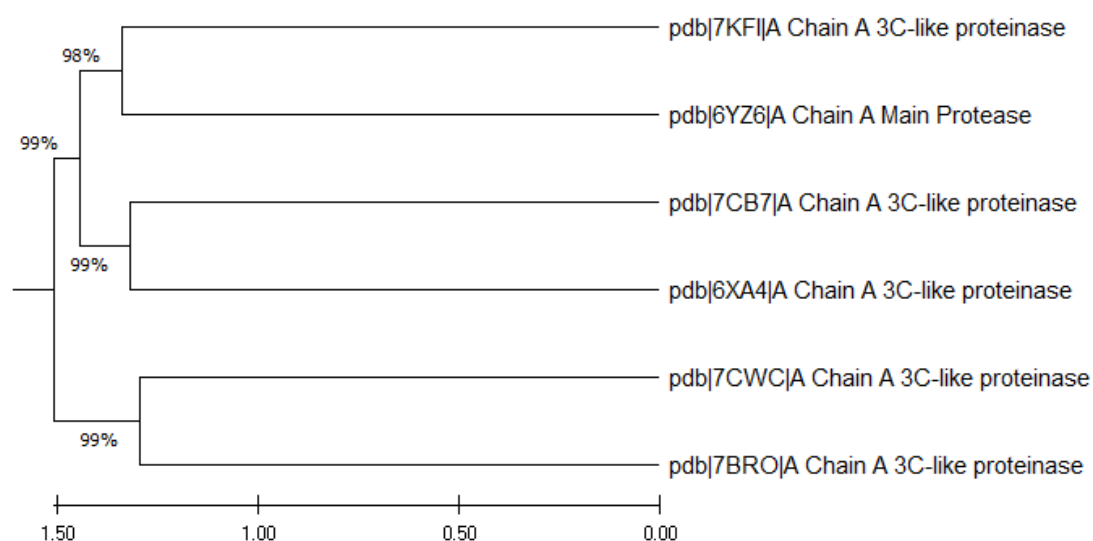

Figure 6. Tree obtained with MEGE-X using the UPGMA method.

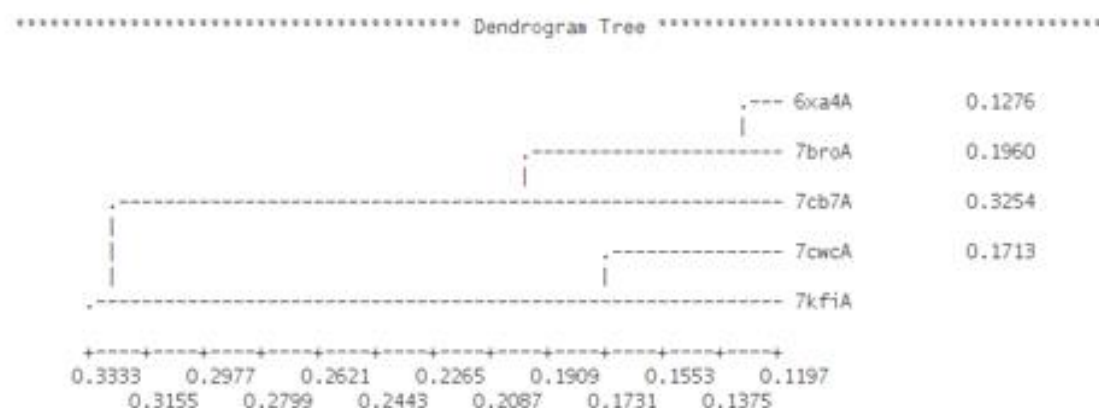**Emmanuel Israel Edache, Adamu Uzairu, Paul Andrew Mamza, Gideon Adamu Shallangwa**

```
************************************* Dendrogram Tree *************************************

                                                        .--- 6xa4A        0.1276
                                                        |
                                    .---------------------------- 7broA    0.1960
                                    |
                    .------------------------------------------------- 7cb7A   0.3254
                    |
                    |                               .--------------- 7cwcA    0.1713
                    |                               |
                    .------------------------------------------- 7kfiA

          +----+----+----+----+----+----+----+----+----+----+----+
        0.3333   0.2977   0.2621   0.2265   0.1909   0.1553   0.1197
          0.3155   0.2799   0.2443   0.2087   0.1731   0.1375
```

Figure 7. The weighted pair-group matrix of the template sequences

```
************************************* Alignments *************************************

_aln.pos        10        20        30        40        50        60
6xa4A     -GFRKMAFPSGKVEGCMVQVTCGTTTLNGLWLDDVVYCPRHVICTSEDMLNPNYEDLLIRKSNHNFLV
ABCD      SGFRKMAFPSGKVEGCMVQVTCGTTTLNGLWLDDVVYCPRHVICTSEDMLNPNYEDLLIRKSNHNFLV
_consrvd   ****************************************************************


_aln.p  70        80        90        100       110       120       130
6xa4A     QAGNVQLRVIGHSMQNCVLKLKVDTANPKTPKYKFVRIQPGQTFSVLACYNGSPSGVYQCAMRPNFTI
ABCD      QAGNVQLRVIGHSMQNCVLKLKVDTANPKTPKYKFVRIQPGQTFSVLACYNGSPSGVYQCAMRPNFTI
_consrvd   *****************************************************************


_aln.pos  140       150       160       170       180       190       200
6xa4A     KGSFLNGSCGSVGFNIDYDCVSFCYMHHMELPTGVHAGTDLEGNFYGPFVDRQTAQAAGTDTTITVNV
ABCD      KGSFLNGSCGSVGFNIDYDCVSFCYMHHMELPTGVHAGTDLEGNFYGPFVDRQTAQAAGTDTTITVNV
_consrvd   *****************************************************************


_aln.pos  210       220       230       240       250       260       270
6xa4A     LAWLYAAVINGDRWFLNRFTTTLNDFNLVAMKYNYEPLTQDHVDILGPLSAQTGIAVLDMCASLKELL
ABCD      LAWLYAAVINGDRWFLNRFTTTLNDFNLVAMKYNYEPLTQDHVDILGPLSAQTGIAVLDMCASLKELL
_consrvd   *****************************************************************


_aln.pos  280       290       300
6xa4A     QNGMNGRTILGSALLEDEFTPFDVVRQCSGVTF-
ABCD      QNGMNGRTILGSALLEDEFTPFDVVRQCSGVTFQ
_consrvd   *********************************
```

Figure 8. Sequence alignment of the crystal structure of the SARS-COV-2 main protease. The symbol (*) refers to identical amino acids.

The modeled protein structure was submitted to the SAVES website and Ramachandran plots were generated. The stereochemical quality of the modeled protein, including its main chain, bond lengths, and bond angles, is assessed using Ramachandran plots. A good model should have more than 90% of its residues in the most advantageous areas as discussed earlier. The Ramachandran plot analysis (Fig. S12) of the comparative protein structure model predicted found 94.7%, 4.2%, 0.8%, and 0.4% residues in the favorable allowed, additionally allowed, generously allowed, and disallowed regions accordingly. Likewise, the ERRAT score values obtained for the modeled protein structure were 94.966% (Fig. S13). The modeled protein coordinates that determine the stereo-chemical qualities of predicted models were found to be stable and reliable. The profile score above zero in the VERIFY-3D graph (Fig. S14), corresponds to an acceptable environment of the model. The result shows that the modeled protein possesses the verification where at least 95% of the amino acids have scored ≥ 0.2 in the 3D/1D profile. The dope

Emmanuel Israel Edache, Adamu Uzairu, Paul Andrew Mamza, Gideon Adamu Shallangwa

score profile graphs generated by MODELLER in Fig. 9 show that the SARS-CoV-2 Mpro (PDB ID: 6YZ6) energy graph (red color) lies exactly on one chain of SARS-CoV-2 3C-like proteinase template (PDB ID: 6XA4) energy graph peaks (green color) which show primary similarity and closeness. Matchmaker 6XA4.pdb, chain A with modeled protein structure (ABCD.B99990005.pdb), sequence alignment score = 1585.1. RMSD between 302 pruned atom pairs is 0.420 angstroms. Fig. 10 presented the superimposition of the query (PDB ID: 6YZ6) with the template (PDB ID: 6XA4.pdb) by ChimeraX v1.3. The template structure is displayed in blue and target with grey color. The protein structure analysis (ProSA) of the 3D model SARCoV-2 Chain A, Mpro gives a Z-score of -7.27 indicating that the model is in the variety of acceptable protein structures that have been experimentally elucidated using X-ray crystallography (Fig. S15), The plot of residue energies, where positive values correlate to problematic or incorrect elements of the input structure [51]. Fig. S16 shows that majority of the computed values were negative, as shown by the plot of residue energies. The use of both of the proposed 3D models could be considered highly dependable, according to all of these validation techniques, which was strongly advised.

3.4 Computational screening for Thiazolino 2-pyridone amide derivatives.

After achieving a comparable model of the 3D structure that is acceptable, C. trachomatis and SARS-CoV-2 Mpro, docking simulations were carried out next. These dockings were done to examine the interactions and binding affinities of each compound. The most popular method for screening the binding modes of the retained homology C. trachomatis and SARS-CoV-2 protein models is docking simulation. This study used AutoDock-vina [66], a feature of the EasyDockVina v2.2 program, to clarify the predicted binding pocket using a set of compounds of thiazolino-2-pyridone amide derivatives.

3.4.1 Recognition of C. trachomatis protein binding anti-bacterial compounds with molecular docking

To investigate the binding affinity of the synthesized compounds, we screened all 46 selected compounds and seven (7) reference (control) compounds as presented in Table S1. The synthesized and the reference compounds were docked well in the modeled protein binding site and set up many van der Waals, electrostatic interactions, as well as hydrogen bonding, and hydrophobic interactions. The more negative binding affinity for the reference drug is Ritonavir (-7.7 kcal/mol) against the C. trachomatis modeled protein. In the Autodock-vina mode, the reference drug Ritonavir maintained a conventional hydrogen bond and carbon-hydrogen bond with the following residues: ASP 177 and GLN229 which have two carbon-hydrogen bond interactions. The residues had the following bond lengths: 3.83 Å, 4.06 Å, and 5.92 Å, respectively. The amino acid residues ILE174, LEU233, and GLN229 play a crucial role in hydrophobic interactions (Fig. 11). The rest of the thiazolino 2-pyridone amide derivatives that were docked using the precision mode in Autodock-vina led to the identification of seven potential compounds exhibiting better binding than the reference drugs with a binding affinity of -8.0 to -8.2 kcal/mol. The interacting residues for these lead compounds are shown in Table S5 and the best two with conventional hydrogen docking pose are depicted in Fig. 12. Higher binding affinities, equal to -8.2 and -8.1 Kcal/mol, were demonstrated by the lead compounds 30 (pIC50 = 5.6021) and compound 31 (pIC50 = 5.000). The compounds (30 and 31) have the best binding affinities with the modeled C. trachomatis protein compared to other compounds and the standard drugs, respectively. Fig. 12A shows the interactions of compound 30 with the prevalence of hydrophobic interactions and one conventional hydrogen bonding. As we can see, the hydrophobic interactions with the important residues ILE174, THR236, and LEU233 have been exposed. Meanwhile, the residues GLN240 add value to the conventional hydrogen bond interaction. Most interacting residues with compound 31 and the modeled protein structure were observed to be GLU173, ILE174, THR236, LEU233, and GLY238 with electrostatic (pi-anion) and hydrophobic interaction (as seen in Fig. 12B). It was also indicated that compound 31 shows conventional hydrogen bonding between the modeled receptor (inclusion membrane protein IncA [Chlamydia trachomatis]) with residue GLN232 with a distance of 4.34 Å, between H-

Emmanuel Israel Edache, Adamu Uzairu, Paul Andrew Mamza, Gideon Adamu Shallangwa

donor and acceptor atom. Based on these findings, the structure of compounds (30 and 31) can be adopted as a new potential reference in the development and design of new drugs against C. trachomatis diseases.
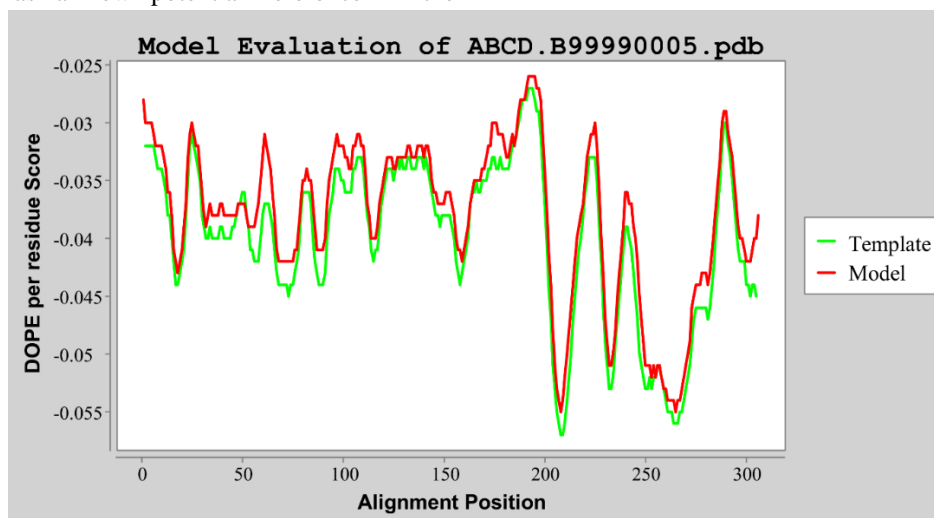

Figure 9. Dope profiles of the SARS-CoV-2 model (displayed in red color) and template 6XA4.pdb (displayed in green color).
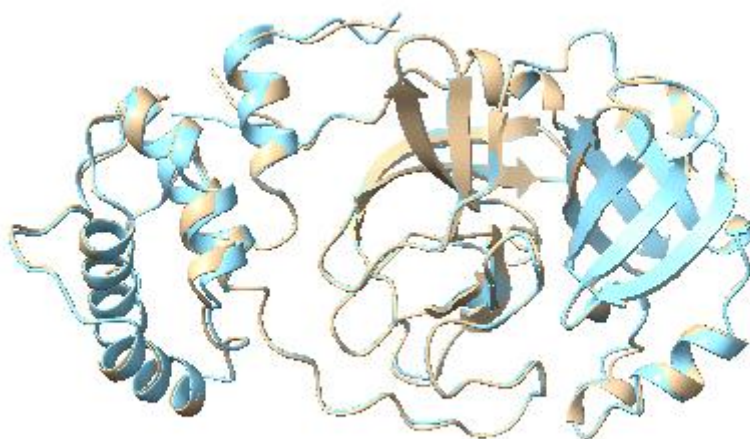

Figure 10. Superimposition of PDB ID: 6YZ6 with the template (PDB ID: 6XA4.pdb) by ChimeraX v1.3, and RMSD is 0.224. SARS-CoV-2 model template alignment.

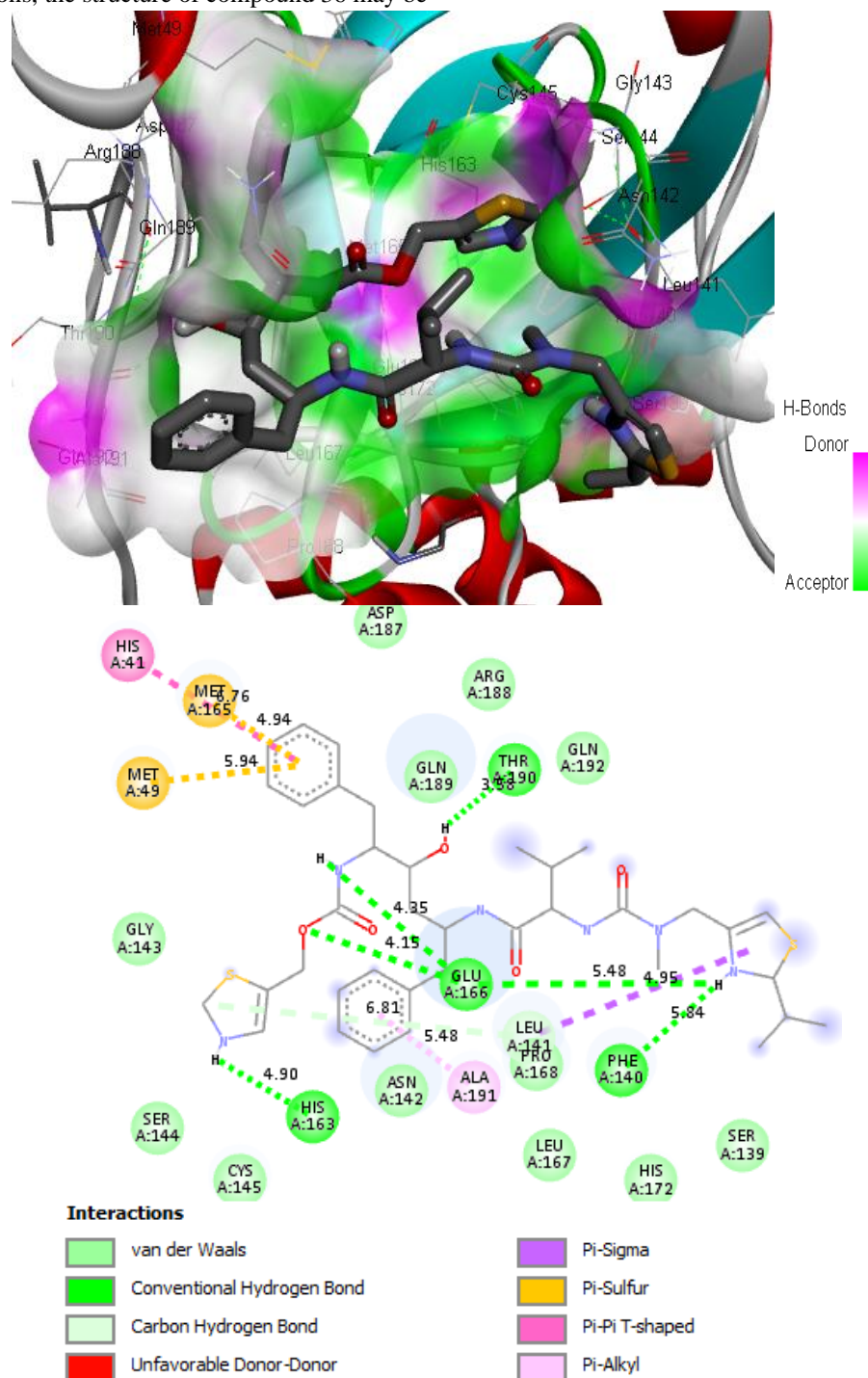**Emmanuel Israel Edache, Adamu Uzairu, Paul Andrew Mamza, Gideon Adamu Shallangwa**
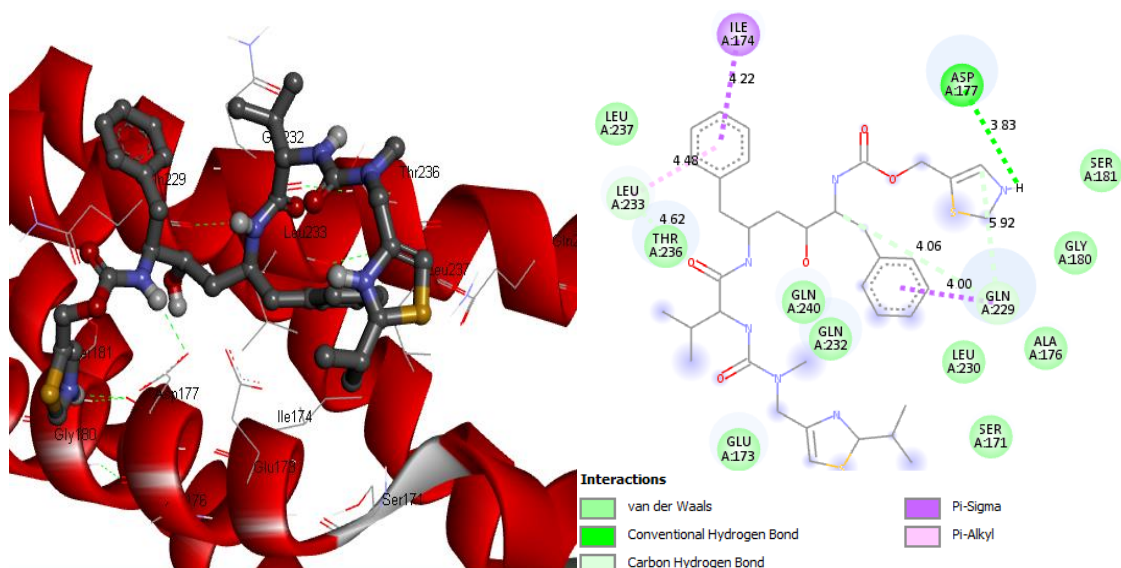
Figure 11. 3D and 2D interaction representation of the reference ligand Ritonavir in the catalytic site of the Chlamydia trachomatis protein structure.
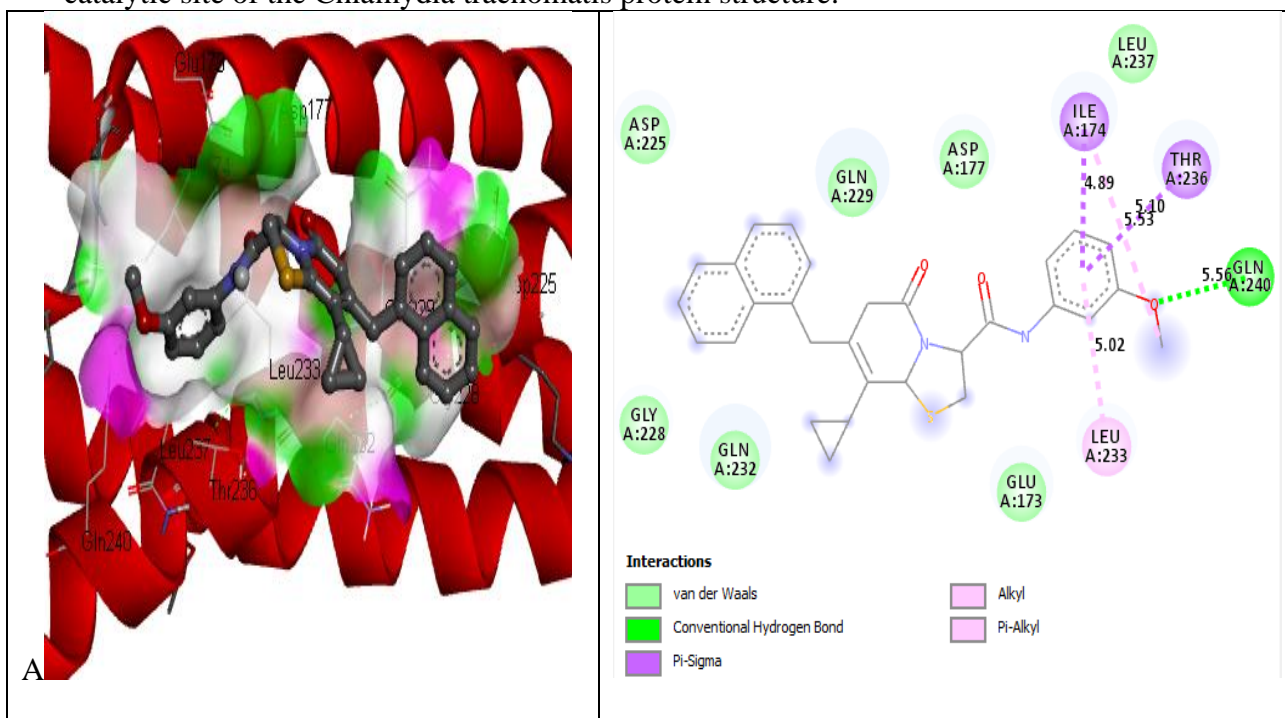
Emmanuel Israel Edache, Adamu Uzairu, Paul Andrew Mamza, Gideon Adamu Shallangwa
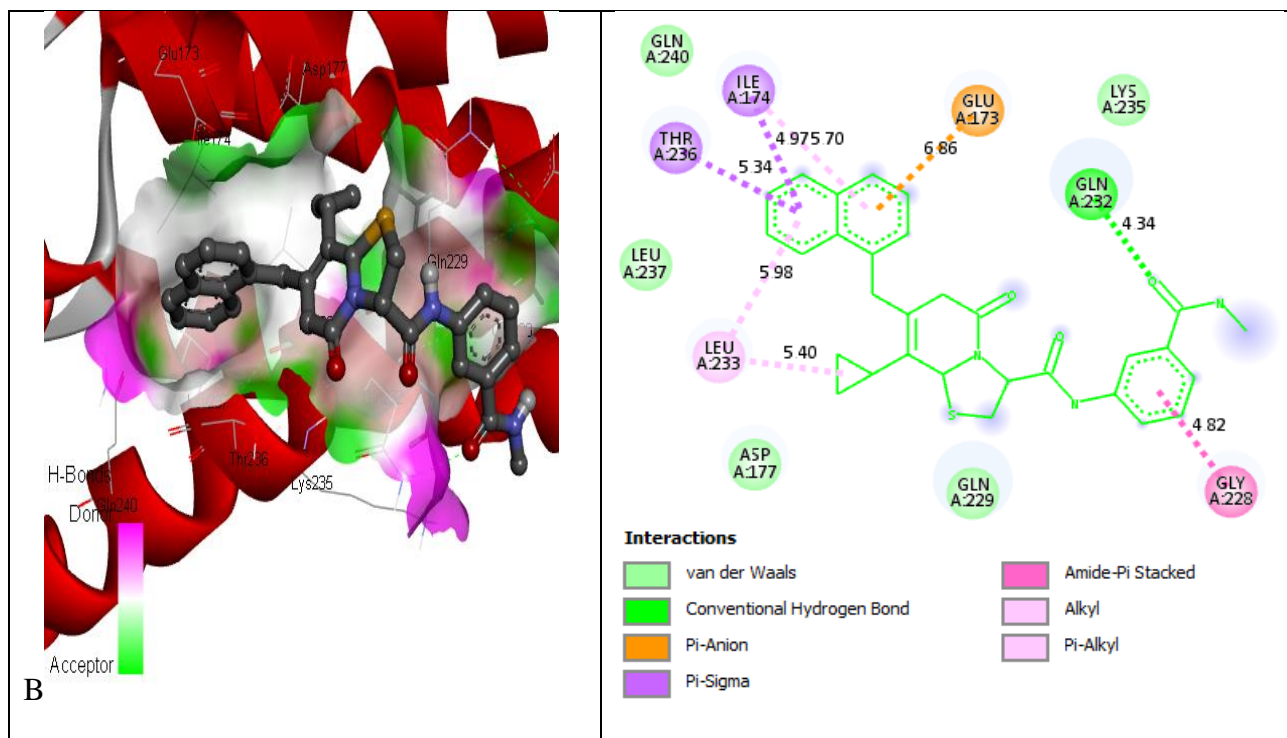
Figure 12. 2D and 3D interaction representation of the best compounds with conventional hydrogen bond interactions (A) Compound 30 and (B) compound 31 in the catalytic site of chlamydia trachomatis protein structure.

3.4.2 Recognition of SARS-CoV-2 Mpro binding antiviral compounds with molecular docking

To identify the kinds of amino acid residues held to account for the molecular interaction at the catalytic site, the inhibitors' docking simulation was examined on the SARS-CoV-2 modeled receptor. The calculated binding affinities for each compound range from -6.3 to -8.7 kcal/mol, while Ritonavir, a reference drug, was determined to have a binding affinity of -8.7 kcal/mol (Table S1). From Fig. 13, we can see that the referenced drug (Ritonavir) interacts with modeled SARS-CoV-2 Mpro by forming 6 conventional hydrogen bonds, which are THR190 (3.58 Å), GLU166 (4.15 Å, 4.35 Å, 5.48 Å), PHE140 (5.84 Å), and HIS163 (4.90 Å). The reference drugs also have one carbon-hydrogen bond (LEU141 (6.81 Å), 3 hydrophobic interactions, two miscellaneous interactions, and one unfavorable donor-donor interaction. The referenced drug (Ritonavir) interacted with the active sites MET49 (5.94 Å) and MET165 (4.94 Å) via a miscellaneous interaction and interacted with the active residue sites HIS41 (6.76 Å), ALA191 (5.48 Å), and LEU141 (4.95 Å) via hydrophobic interactions. Moreover, Ritonavir interacted with residues ASP187, ARG188, GLN189, GLN192,

SER139, HIS172, LEU167, PRO168, ASN142, CYS145, SER144, and GLY143 through van der Waals interactions with one unfavorable donor-donor interaction with THR190. Also, compound 36 interacted with SARS-CoV-2 with a-8.7 Kcal/mol binding affinity score (Table S1). The connection between the modeled Mpro of SARS-coronavirus and the ligand is shown in Fig. 14. Compound 36 interacts with modeled SARS-CoV-2 Mpro by forming 4 conventional hydrogen bonds, one carbon-hydrogen bond, 2 hydrophobic bonds, and with no record of unfavorable dumps. we can see that compound 36 interacted with four residues THR26 (3.90 Å), GLY143 (3.62 Å), CYS145 (4.06 Å), and SER144 (3.56 Å) through conventional hydrogen bonds, one carbon-hydrogen bond with MET165 (5.53 Å), one hydrophobic (Pi-Alkyl) interaction with PRO168 (6.40 Å), and one miscellaneous (Pi-Sulfur) interaction with CYS145 (5.53 Å), respectively. Whereas twelve residues of the modeled protein of SARS-CoV-2 Mpro ARG188, GLN192, LEU167, THR190, GLN189, HIS164, GLU166, HIS163, LEU141, ASN142, THR24, and THR25 participated in the compound to protein van der Waals interactions. We can also notice that the distances between the structure of

Emmanuel Israel Edache, Adamu Uzairu, Paul Andrew Mamza, Gideon Adamu Shallangwa

compound 36 and the amino acid residues with which it interacts in the complex were shorter compared to those with which referenced drug (Ritonavir) interacts in the complex, especially with interactions that occurred through conventional and carbon-hydrogen bonds. Based on these observations, the structure of compound 36 may be

more stable in complex than the referenced drug (Ritonavir). Based on these hypotheses, the structure of compound 36 can be adopted as a new potential reference in the development and design of new drug inhibitors of the SARS-coronavirus Mpro.

Figure 13. 3D and 2D interaction representation of the reference ligand (Ritonavir) in the catalytic site of the SARS-coronavirus-2 Mpro.

Emmanuel Israel Edache, Adamu Uzairu, Paul Andrew Mamza, Gideon Adamu Shallangwa
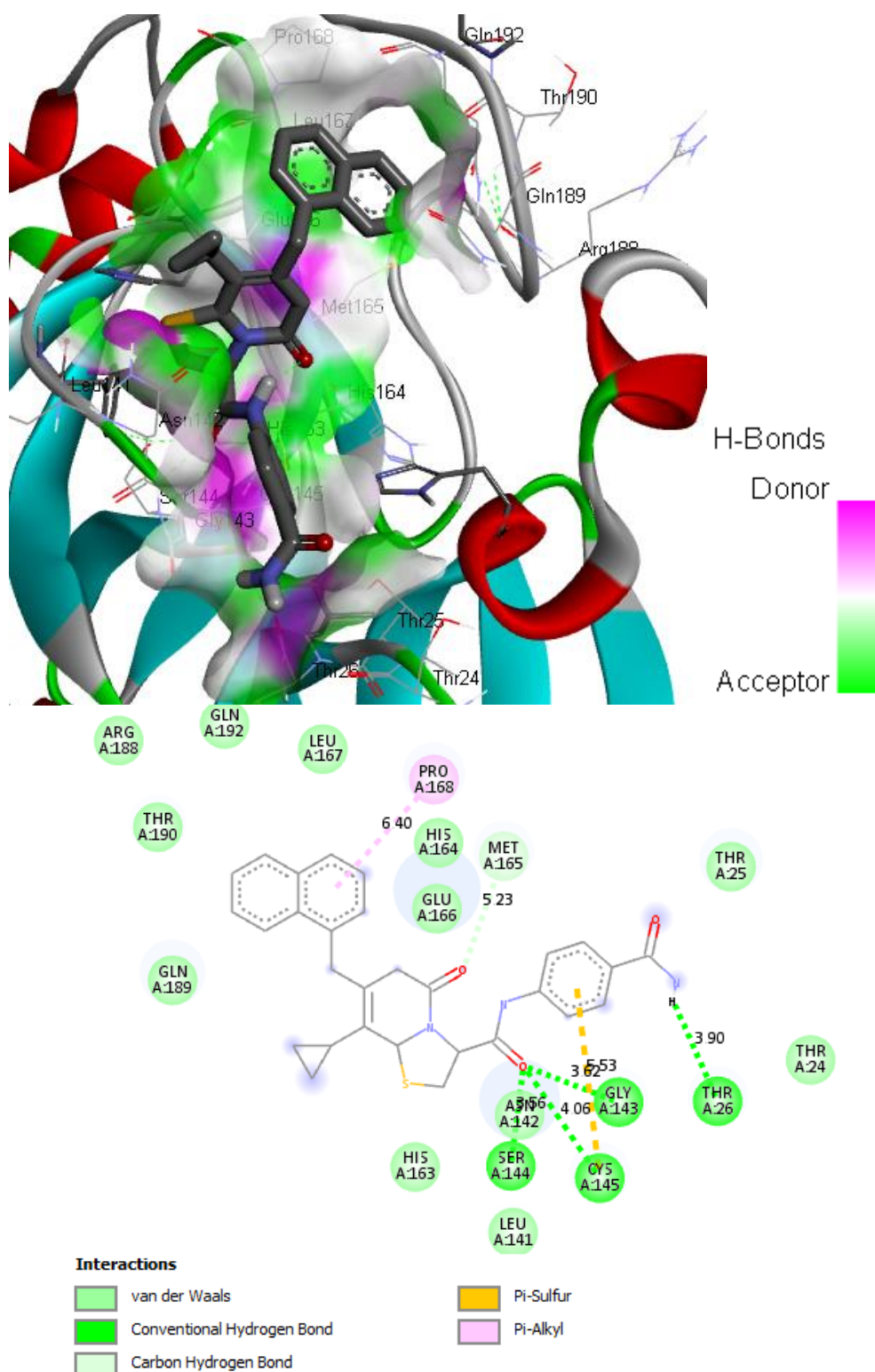
Figure 14. 3D and 2D interaction representation of compound 36 in the catalytic site of the SARS-coronavirus-2 Mpro.

3.5 Valuation of drug-like properties

Pharmacokinetic properties, bioavailability score, water-solubility, drug-likeness, pharmacokinetics, and medicinal chemistry properties of the top-ranked compounds (30, 31, 36, and reference drug (Ritonavir)) using the SwissADME server (www.swissadme.ch (accessed on 24 July 2022)) were carried out [67, 68]. The final technique to

Emmanuel Israel Edache, Adamu Uzairu, Paul Andrew Mamza, Gideon Adamu Shallangwa

identify molecules that adhere to the following rules: Lipinski's rule [69], Veber's rule [70], Ghose's rule [71], and Egan's rule [72] is ADME (absorption, distribution, metabolism, excretion, and toxicity). In order to create new anti-C. trachomatis and anti-SARS-CoV-2 medications, the top-ranked molecules will be sorted to determine which compound is the most effective. The various laws (Lipinski, Veber, Ghose, Muegge, and Egan) that we have discussed are based on how drugs behave in the body. Table S6 contains the findings of the analysis of the predicted pharmacokinetic properties. The compounds (30, 31, 36, and Ritonavir) present MLogP (MLogP ≤ 4.15) values in the range 2.01 to 4.25. The molecular weight (MW ≤ 500) of compounds (30 and 36) was in the acceptable range, while compounds 31 and Ritonavir are out of the range. The hydrogen bond acceptors and hydrogen bond donors of all the selected compounds were in the acceptable range (HA ≤ 10, HB ≤ 5) except for the reference drug (Ritonavir), this shows that compounds (30 and 36) met all criteria of the Lipinski rule. With the exception of Ritonavir, which violated Veber's rule, the results of rotatable bond (NRB) and topological polar surface area (TPSA) values indicate that the compounds possess excellent bioavailability and become more versatile for proficient interaction with a particular binding pocket. When developing active substances as therapeutic agents, oral bioavailability is frequently a crucial factor to take into account. Understanding the molecular characteristics that restrict oral bioavailability is thus a key objective in drug research. Lipophilicity, molecular size, molecular polarity, insolubility, insaturation of molecular structure, and molecular flexibility are some of the most crucial factors taken into account when evaluating oral bioavailability. According to Fig. S17, compound 30 has all six parameters in the desired range out of the six that were taken into consideration (Fig. S17A), while compounds (31 and 36) have insaturation parameters a bit out of the range (Fig. S17B and S17C). The reference drug has all the parameters except insaturation out of the desired range (Fig. S17D). Accordingly, these compounds (30, 31, and 36) are supposed to have suitable oral bioavailability. A quick screening of molecules to determine whether a compound has the potential to be used as an oral drug is called a

bioavailability score. With the exception of the drug ritonavir, all of the compounds have high bioavailability scores of 0.55 (or 55%), indicating good bioavailability. In contrast to the Ritonavir drug, which has broken the rule of Lipinski, Veber, Ghose, Muegge, and Egan and displays a bioavailability score of 17%, indicating poor bioavailability, all the chosen compounds demonstrated zero PAINS and Brenk alert and can be used as lead compounds. Synthetic accessibility (SA) is one of the most important components of computer-aided drug design (CADD) activities because it aids in choosing the most promising molecule that was created and put through biological testing. The synthetic accessibility (SA) of the selected compounds is evaluated using the synthetic accessibility index. When the synthetic accessibility (SA) values are close to 1, very easy to synthesize, and if SA values are close to 10, very difficult and complicated [73]. Through the obtained results (Table S7), the values of synthetic accessibility (SA) are limited to 4.53 to 7.02. From this assessment, compounds (30 and 36) can be easily synthesized in the laboratory without complexity. The drug Ritonavir also had the highest SA value (7.02), indicating that it would be the most challenging to synthesize in a lab. When a medication is meant to be consumed orally, the drug's gastrointestinal (GI) absorption becomes crucial. This study predicted that the three (3) predicted compounds would be quickly absorbed from the GI tract. The reference drug Ritonavir has a low GI absorption. However, compounds (30 and 36) are accounted for as high, and that implies these mixtures are profoundly caught up in the digestive tract human gastrointestinal absorption (HIA). Contrarily, the blood-brain barrier (BBB permeant) is a significant barrier to the delivery of drugs to the central nervous system (CNS). The BBB is inversely correlated with molecular weight roughly along the square route. The BBB is sufficiently permeable for drugs with higher molecular weight to affect the central nervous system (CNS) [74]. The three compounds that were chosen (Table S8) have a low probability of penetrating the BBB, according to the calculations, and may not have any effect on CNS function. It is noteworthy that after administration, most especially compounds (30 and 36) are supposed to create no neurotoxicity, essentially not in that frame of mind, since they

Emmanuel Israel Edache, Adamu Uzairu, Paul Andrew Mamza, Gideon Adamu Shallangwa

won't cross the BBB (Fig. S18 and Table S8). Cytochrome P450 (CYP), a crucial enzyme for detoxification in metabolism, is a key player. All of the body's tissues contain CYP enzymes [75]. This enzyme oxidizes unwanted microorganisms to help with their removal. Numerous drugs can be both activated and inhibited by the cytochrome CYP enzyme. Drug metabolism may be affected by this enzyme inhibitor, which could result in the drug having an adverse effect [76]. Evaluation of a compound's capacity to inhibit cytochromes (CYP) is therefore essential. The most well-known liver cytochrome P450 isoforms CYP1A2, CYP2C19, CYP2C9, CYP2D6, and CYP3A4 play a crucial function in drug metabolism and the efficacious and carcinogenic concentrations of drugs in the body [77]. Even though only CYP1, CYP2, CYP3, and CYP4 are responsible for the metabolism of drugs, only the types (1A2, 2C9, 2C19, 2D6, and 3A4) are therefore accountable for metabolization for far more than 90% of drugs that pass the initial stage of metabolism [78]. The main players in drug metabolism are the two isoforms 2D6 and 3A4 [79]. The processing of organic molecules by CYP and P-GP may work in concert to enhance tissue and organism defense [80, 81]. An efflux drug transporter known as P-glycoprotein (P-GP) protects cells from the harmful effects of prescription drugs by removing toxins and foreign substances out from cells [82-85]. All proposed C. trachomatis and SARS-CoV-2 inhibitors are an inhibitor of P-GP (Table S7). Due to the significance of this issue, we used computational methods to try and predict how compounds (30 and 36) will affect liver enzymes. The findings in Table S8 suggest that some of the cytochrome P450 isoforms may inhibit one or more of the proposed compounds. These ligands are therefore anticipated to have less hepatotoxic effects. In order to verify the human intestinal absorption (HIA) and blood-brain barrier (BBB) access as described in Table S8, the compounds' WLogP and TPSA values (Table S6) were plotted. The egg-shaped graph is divided into three areas a white area (HIA), a yolk (BBB access), and a grey area (no HIA or BBB access) (Fig. S18). The presence of the compounds (30, 31, and 36) in the white portion of the plot suggests that they most likely penetrated the gastrointestinal absorption (HIA). It is anticipated that the reference drug (Ritonavir) in the grey area will have

insufficient brain penetration and intestinal absorption. Additionally, this boiled-egg method gives whether or not compounds (30, 31, 36, and Ritonavir) were P-glycoprotein substrates (PGP). P-glycoprotein substrates (PGP+) and P-glycoprotein non-substrates (PGP-) are represented by the red and blue dots, respectively. As shown in Fig. S18, all the selected compounds represented by a blue dot are P-glycoprotein non-substrates (PGP-), which conformed with the results in Table S8. The findings indicate that compounds (30) and (36) had good oral bioavailability and could be considered anti-C. trachomatis and anti-SARS-CoV-2 drugs, respectively. These compounds met the pharmacokinetic requirements for drug-like compound behavior. The results presented in Tables S7 – S9 as well as Fig. S17 and Fig. S18 indicate that the Lipinski, Veber, Egan, and Muegge rules are all satisfied by the compounds (30 and 36), indicating that there are no issues with oral bioavailability for these substances. Additionally, these substances have demonstrated a high capacity for absorption. Based on this hypothesis, we can confirm these compounds (30 and 36) cause no oral bioavailability issues and have great medication like properties. Subsequently, the compounds (30 and 36) are the favored possibility to be utilized as new C. trachomatis and SARS-CoV-2 inhibitors, respectively among the 46 thiazolino-2-pyridone amide derivatives molecules studied.

3.6 Molecular dynamic simulation

A crucial tool for predicting potential structural conformations between ligand and receptor or enzyme is a molecular dynamics simulation (MDS) study. For all of the complexes to the initial structure, RMSD, RMSF, SASA, and Rg values of C atoms were estimated in order to track the impact of the simulation on the conformational stability of the receptors. The outcomes were plotted as an element of the reproduction's time. The overall RMSD value of compound 30 complex is 2.13 Å. The RMSD changed significantly between 851 and 997 ps before becoming more steady after 1 ns (Fig. S19A). Fig. S19B illustrates the RMSF (root-mean-square fluctuations) of the complex structure of compound 30 produced during the MD simulation. For the purpose of describing the mobility of specific residues, these fluctuations were tabulated. The loop region of residue 172 exhibits a maximum fluctuation at about 7.5 nm, as shown in Fig. S19B.

Emmanuel Israel Edache, Adamu Uzairu, Paul Andrew Mamza, Gideon Adamu Shallangwa

Even though very few fluctuations were seen during the compound 30 complex interaction, the majority of these fluctuations were smaller than 2.5 nm, demonstrating a strong binding between compound 30 and the inclusion membrane protein IncA. In the simulation process, the progression of the corresponding ligand had a greater impact on the protein. According to RMSF analysis, compound 30's binding to the target protein showed a marginally nonsignificant fluctuation, revealing that the ligand binding had no discernible effects on the protein's conformation. Solvent Accessible Surface Area (SASA) is the surface area in which ligands could indeed engage with solvent molecules. In Fig. S19C, which plots the average SASA values for the compound 30 complex against simulation time, fluctuations in the ligand-protein complex are seen between 450 and 700 ps. The transformation in the protein's formability following the ligand interaction is represented by the radius of gyration. In the first 55 ps, the Rg values varied between 20.48 nm to 21.00 nm. Nevertheless, following this point and up until the simulation's final result, the values stayed largely stable in the 21.01 to 21.75 nm range. The average Rg values of 21.2 nm are calculated from the compound 30 complex. After experiencing some fairly noticeable fluctuations, the complex's radius of gyrating tends to stabilize at about 900 ps (Fig. S19D). The inclusion membrane protein IncA C. trachomatis can interact with compound 30 without changing its structural folding in the dynamic environment, according to the Rg results, which demonstrate that the ligand-protein systems achieve a level of compactness. The final frame of the compound 30 complex forms conventional and carbon H-bonds after a successful MDS; one of them is stronger than the other. According to Fig. S19E, compound 30 forms one H-bond with THR236 by contributing an electron pair with a bond length of 4.18 and another bond with the NH of LEU233 by contributing an electron pair with a bond length of 3.39.

As can be seen from the data for the SARS-CoV-2 Mpro complex in Fig. S20, all systems achieved equilibrium within the first few picoseconds of the simulation and remained stable thereafter. The SARS-CoV-2 Mpro complex (compound 36 complex) has a total RMSD value of 4.01. The detailed representation of the trajectories clearly demonstrated that the systems had reached equilibrium and presented initial proof that systems are stable under physiological conditions (Fig. S20A). Figure S20B displays the RMSF of the C atoms in each residue of both proteins and complexes. The total RMSF value for compound 36 complex was discovered to be less than 0.81, denoting the amino acids' good stability in an aqueous environment. Few residues, particularly terminal ones, demonstrated greater fluctuations, which is attributable to their hanging position and propensity for fluctuations. Fig. S20C shows the SASA of compound 36 complex. The data clearly show that SASA fluctuated during the simulation period. For example, they fluctuate from 100 to 200 ps and 700 to 800 ps. The system becomes stable from 950 ps to 1 ns. The Rg of compound 36 complex is presented in Fig. S20D. The data unambiguously demonstrates that Rg fluctuated over the course of the simulation before stabilizing at around 950 ps. The average Rg of compound 36 complex was found to be 22.32 nm. The data a little farther confirms that all systems are stable in a physiological environment Fig. S20E shows new amino acids (GLU166 and MET165) that helped the compound 36 complex remain stable during the simulation, in contrast to some amino acids that were present in the docking complex (THR190, GLN189, HIS164, THR24, HIS163, LEU141, SER144, and THR26). However, contacts with GLY143, PRO168, CYS145, GLU166, MET165, and PRO168 stay all through the direction. Conventional and carbon-hydrogen bonding profiles involving specific amino acids in the binding site support the ligand's stability. These results from MDS demonstrated that compounds 30 and 36 were properly docked in the active site and suggested that they formed stable complexes, which can be interpreted as satisfactory results.

3.6.1 MM/GBSA binding energy

We use the molecular mechanics-generalized Born surface area (MM-GBSA) method created by Miller et al. and Kollman et al. [86, 87] to determine the binding energy of complexes. The trajectories of MDS are used to calculate the binding energy. Equations 4 to 8 were used to determine the complex system's binding free energy for each snapshot.

Emmanuel Israel Edache, Adamu Uzairu, Paul Andrew Mamza, Gideon Adamu Shallangwa

$\delta G_{bind} = \delta G_{complex} - (\delta G_{protein} + \delta G_{ligand})= -RTlnK_i$ 5

$\Delta G_{bind} = \Delta H - T\Delta S \approx \Delta E_{MM} + \Delta G_{sol} - T\Delta S$ 6

$\Delta E_{MM} = \Delta E_{internal} + \Delta E_{ele} + \Delta E_{vdw}$ 7

$\Delta G_{sol} = \Delta G_{SA} + \Delta G_{GB}$ 8

Where $\delta G_{complex}$, $\delta G_{protein}$, and δG_{ligand} are the free energies of the complex, protein, and ligand, respectively. ΔE_{MM} and $-T\Delta S$ is the gas phase MM energy and conformational entropy, respectively. ΔE_{MM} contains electrostatic ΔE_{ele}, van der Waals energy ΔE_{vdw} and $\Delta E_{internal}$ of bond, angle, and dihedral energies. ΔG_{sol} is the solvation-free energy which is the sum of the nonelectrostatic solvation component ΔG_{SA} and electrostatic solvation energy ΔG_{GB}. The binding energy is determined using 1000 frames. The binding free energies, including the energy contribution for each model, are summarized in Table 4 based on our evaluation.

Table 4. Calculating MM/GBSA binding energy of compound (30 and 36) complexes

complex	protein	ligand	MM/GBSA
IncA C. trachomatis structure with Compound 30			
Av. BOND = 275.1218	Av. BOND = 267.0135	Av. BOND = 10.2889	ΔE(internal) = -2.8355
Av. ANGLE = 879.9713	Av. ANGLE = 660.0253	Av. ANGLE = 225.0678	ΔE(electrostatic) +
Av. DIHED = 1,573.7779	Av. DIHED = 1,549.7024	Av. DIHED = 18.944	ΔG(sol) = 4.8043
Av. IMPRP = 46.6093	Av. IMPRP = 47.1906	Av. IMPRP = 0.0832	ΔE(VDW) = -42.2051
Av. ELECT = -5,947.4077	Av. ELECT = -5,876.0459	Av ELECT = -76.1661	ΔG binding = -40.2363
Av. VDW = -540.1239	Av. VDW = -538.7162	Av. VDW = 40.7975	+/- 0.3449 (kcal/mol)
Mpro of SARS-CoV-2 with Compound 36			
Av. BOND: 431.6435	Av. BOND: 421.6344	Av. BOND: 10.9402	ΔE(internal) = 2.3445
Av. ANGLE: 1,325.3474	Av. ANGLE: 1,098.4201	Av. ANGLE: 227.9729	ΔE(electrostatic) +
Av. DIHED: 2,732.4421	Av. DIHED: 2,687.3247	Av. DIHED: 41.5015	ΔG(sol) = 1.7456
Av. IMPRP: 55.9942	Av. IMPRP: 55.1711	Av. IMPRP: 0.1177	ΔE(VDW) = -44.9076
Av. ELECT: -7,082.2802	Av. ELECT: -7,033.5957	Av. ELECT: -50.43	ΔG binding = -40.8176
Av. VDW: -1,260.549	Av. VDW: -1,250.4808	Av. VDW: 34.8394	+/- 0.5498 (kcal/mol)

Av =Average; ELECT = Electrostatic; VDW =van der Waals; sol = solvation; Δ =Delta

The calculated binding free energies of compounds 30 and 36 were -40.2363 and -40.8176, respectively (Table 4). For the two systems, the van der Waals interactions interactions (ΔE_{vdw}) played a major role in the total binding free energy (ΔG_{bind}). The contribution of electrostatic interaction (ΔE_{ele}) was nearly counteracted by the electrostatic solvation energy (ΔG_{GB}).

4. Conclusions

In this work, we performed computational-Aided drug design and development on 46 thiazolino 2-pyridone amide derivatives as candidates against Chlamydia trachomatis and SARS-CoV-2 Mpro. Various modeling methods were used for this study, such as QSAR, phylogenetic analysis/ homology modeling, docking simulation, ADMET, and molecular dynamics simulations. This study aims to identify the potential of thiazolino 2-pyridone amide derivatives to inhibit C. trachomatis and SARS-CoV-2 Mpro by acting on the protein binding domain of the C. trachomatis and SARS-CoV-2 Mpro. The findings of this research study demonstrate that compounds 30 and 36 possess strong inhibitors of C.s trachomatis and SARS-CoV-2, respectively. After series in silico screening compounds (30 and 36) are predicted to have one of the strongest binding affinities with the receptors and the ADME and molecular dynamics simulations results exhibit their drugability. The precise binding mechanism, however, is still a work in progress because it necessitated additional molecular dynamic simulation research. Moreover, the in vivo and in vitro effects of compound 30 against the inclusion membrane protein IncA (Chlamydia trachomatis) and compound 36 against SARS-CoV-2 Mpro would be expected to affirm the system. The work can be extended further to predict its functions and other expectations.

References

[1] M. Lyu, G. Fan, G. Xiao, T. Wang, D. Xu, J. Gao, S. Ge, Q. Li, Y. Ma, H. Zhang, J. Wang, Y Cui, J. Zhang, Y. Zhu, B. Zhang, Traditional Chinese medicine in COVID-19. Acta Pharm. Sin. B 11(11) (2021) 3337e3363. DOI: 10.1016/j.apsb.2021.09.008.

[2] S.V. Stoddard, S.D. Stoddard, B.K. Oelkers, K. Fitts, K. Whalum, K. Whalum, A.D. Hemphill, J. Manikonda, L.M. Martinez,

Emmanuel Israel Edache, Adamu Uzairu, Paul Andrew Mamza, Gideon Adamu Shallangwa

E.G. Riley, C.M. Roof, N. Sarwar, D.M. Thomas, E. Ulmer, F.E. Wallace, P. Pandey, S. Roy, Optimization Rules for SARS-CoV-2 Mpro Antivirals: Ensemble Docking and Exploration of the Coronavirus Protease Active Site. Viruses. *12* (2020) 942. DOI: 10.3390/v12090942.

[3] H., Karataş, B., Tüzün, Z., Kökbudak, Could pyrimidine derivative be effective against Omicron of SARS-CoV-2?. BRATISLAVA MEDICAL JOURNAL-BRATISLAVSKE LEKARSKE LISTY 123(7) (2022) 505-513.

[4] B., Tüzün, K., Sayin, H., Ataseven, Could Momordica Charantia Be Effective In The Treatment of COVID19?. Cumhuriyet Science Journal, 43(2) (2022) 211-220.

[5] E.I. Edache, A. Uzairu, P.A. Mamza, G.A. Shallengwa, QSAR, homology modeling, and docking simulation on SARS-CoV-2 and pseudomonas aeruginosa inhibitors, ADMET, and molecular dynamic simulations to find a possible oral lead candidate. J. Genet. Eng. Biotechnol. *20* (2022) 88. DOI: 10.1186/S43141-022-00362-Z.

[6] Md.M. Rahman, T. Saha, K.J. Islam, R.H. Suman, S. Biswas, E.U. Rahat, Md.R. Hossen, R. Islam, Md.N. Hossain, A. Al Mamun, M. Khan, Md.A. Ali Md. M.A. Halim, Virtual screening, molecular dynamics and structure–activity relationship studies to identify potent approved drugs for Covid-19 treatment, J. Biomol. Struct. Dyn. (2020). DOI: 10.1080/07391102.2020.1794974.

[7] K. Mou, M. Abdalla, D.Q. Wei, M.T. Khan, M.S. Lodhi, D.B. Darwish, M. Sharaf, X. Tu, Emerging mutations in envelope protein of SARS-CoV-2 and their effect on thermodynamic properties. Inform. Med. Unlocked. *25* (2021) 100675. https://doi.org/10.1016/j.imu.2021.100675.

[8] A. Ghosh, M. Chakraborty, A. Chandra, M.P. Alam, Structure-activity relationship (SAR) and molecular dynamics study of withaferin-A fragment derivatives as potential therapeutic lead against main protease (Mpro) of SARS-CoV-2. J Mol Model. *27* (**2021**) 97. DOI: 10.1007/S00894-021-04703-6.

[9] World Health Organization. Report on global sexually transmitted infection surveillance, 2018, p. 1-62.

[10] P.D.Tamma. S.E. Cosgrove, L.L. Maragakis, Combination Therapy for Treatment of Infections with Gram-Negative Bacteria. Clin. Microbiol. Rev. *25 (2012)* 450–470. DOI: 10.1128/CMR.05041-11.

[11] E.I. Edache, A. Uzairu, P.A. Mamza, G.A. Shallengwa, Docking Simulations and Virtual Screening to find Novel Ligands for T3S in Yersinia pseudotuberculosis YPIII, A drug target for type III secretion (T3S) in the Gram-negative pathogen Yersinia pseudotuberculosis. Chem Rev Lett. 4 (2021) 130-144. DOI: 10.22034/CRL.2021.254804.1088.

[12] E. Lanjouw, S. Ouburg, H.J. De Vries, A. Stary, K. Radcliffe, M. Unemo, 2015 European guideline on the management of Chlamydia trachomatis infections. Int. J. STD AIDS. 27(5) (2016) 333-348. DOI: 10.1177/0956462415618837.

[13] K.A. Workowski, S.M. Berman, Centers for Disease Control and Prevention sexually transmitted disease treatment guidelines. Clin. Infect. Dis. *53*(suppl_3) (2011) S59-S63. https://doi.org/10.1093/cid/cir694.

[14] H.C. Wiesenfeld, Screening for Chlamydia trachomatis infections in women. *N. Engl. J. Med. NEW*. 376(8) (2017) 765-773, DOI: 10.1056/NEJMcp1412935.

[15] H.C. Wiesenfeld, S.L. Hillier, M.A. Krohn, D.V. Landers, R.L. Sweet, Bacterial vaginosis is a strong predictor of Neisseria gonorrhoeae and Chlamydia trachomatis infection. Clin Infect Dis. 36(5) (2003) 663-668. DOI: 10.1086/367658.

[16] K.A. Redgrove, E.A. McLaughlin, The Role of the Immune Response in Chlamydia trachomatis Infection of the Male Genital Tract: A Double-Edged Sword. Front Immunol. 5 (2014) 534, DOI: 10.3389/fimmu.2014.00534.

[17] E.I. Edache, A. Uzairu, P.A. Mamza, G.A. Shallengwa, Theoretical Investigation of the Cooperation of Iminoguanidine with the Enzymes-Binding Domain of Covid-19 and Bacterial Lysozyme Inhibitors and their Pharmacokinetic Properties. J. Mex. Chem. Soc. 66(4) (2022) 513-542. DOI: http://dx.doi.org/10.29356/jmcs.v66i4.1726.

[18] H., Sadeghi-Nejad, M., Wasserman, W., Weidner, D., Richardson, D., Goldmeier, Sexually transmitted diseases and sexual

Emmanuel Israel Edache, Adamu Uzairu, Paul Andrew Mamza, Gideon Adamu Shallangwa

function. J. Sex. Med. 7(1) (2010) 389-413, DOI:10.1111/j.1743-6109.2009.01622.x.

[19] C., Brookings, D., Goldmeier, H., Sadeghi-Nejad, Sexually transmitted infections and sexual function in relation to male fertility. Korean J. Urol. 54(3) (2013) 149-56, DOI: 10.4111/kju.2013.54.3.149.

[20] A., González-Sánchez, J. J., Reyes-Lagos, Peña-Castillo, M.A., Nirmalkar, K., García-Mena, J. and Pacheco-López, G., Vaginal Microbiota Is Stable and Mainly Dominated by Lactobacillus at Third Trimester of Pregnancy and Active Childbirth: A Longitudinal Study of Ten Mexican Women. Current Microbiology, 79(8) (2022) 1-8.

[21] A. S., Roy, M. I. Q., Tonmoy, A., Fariha, I., Hami, I. K., Afif, M. A., Munim, M. et al. Multi-epitope Based Peptide Vaccine Design Using Three Structural Proteins (S, E, and M) of SARS-CoV-2: An In Silico Approach. J. Appl. Biotechnol. Rep. 8(2) (2021) 146-154, DOI:10.30491/JABR.2020.253820.1303.

[22] S. A., Kohlhoff, M. R., Hammerschlag, M. Treatment of chlamydial infections: 2014 update. Expert Opin. Pharmacother. 16(2) (2015) 205-212, DOI: 10.1517/14656566.2015.999041.

[23] E. I., Edache, A. Uzairu, G. A., Shallengwa, P. A., Mamza, Virtual screening, pharmacokinetics, and molecular dynamics simulations studies to identify potent approved drugs for Chlamydia trachomatis treatment. Futur J Pharm Sci. 7 (2021) 220, DOI: 10.1186/S43094-021-00367-4.

[24] E. S. Okeke, C. V. Olovo, N. E. Nkwoemeka, C. O. Okoye, C. E. I. Nwankwo, C. J. Onu, Microbial ecology and evolution is key to pandemics: using the coronavirus model to mitigate future public health challenges. Heliyon. 8(5) (2022) e09449. DOI: 10.1016/j.heliyon.2022.e09449.

[25] D. Kang, A. V. Revtovich, A. E. Deyanov, N. V. Kirienko, Pyoverdine inhibitors and gallium nitrate synergistically affect Pseudomonas aeruginosa. MSphere. 6 (2021) e00401-21. DOI: 10.1128/MSPHERE.00401-21.

[26] L. Scotti, M. T. Scotti, Computer aided drug design studies in the discovery of secondary metabolites targeted against age-related neurodegenerative diseases. Curr. Top.

Med. Chem. 15(21) (2015) 2239-2252. DOI: 10.2174/1568026615666150610143510.

[27] M.H. Baig, K. Ahmad, S. Roy, J. M. Ashraf, M. Adil, M.H. Siddiqui, S. Khan, M. A. Kamal, I. Provazník, I. Choi, Computer Aided Drug Design: Success and Limitations. Curr. Pharm. Des. 22(5) (2016) 572-581. DOI: 10.2174/1381612822666151125000550.

[28] A. B. Gurung, M. A. Ali, J. Lee, M. A. Farah, K. M. Al-Anazi, An Updated Review of Computer-Aided Drug Design and Its Application to COVID-19. Biomed Res. Int. 2021 (2021) 1-18. DOI: 10.1155/2021/8853056.

[29] S. Belaidi, O. Youcef, T. Salah, T. Lanez, In Silico Approach for Conformational Analysis, Drug-Likeness Properties and Structure Activity Relationships of 12-Membered Macrolides. J. Comput. Theor. Nanosci., 12 (2015) 1–7. DOI:10.1166/jctn.2015.4451.

[30] E. I. Edache, H. U. Hambali, D. E. Arthur, A. Oluwaseye, O. C. Chinweuba, In-silico Discovery and Simulated Selection of Multi-target Anti-HIV-1 Inhibitors. Int Res J Pure Appl Chem. 11 (2016) 1-15. DOI: 10.9734/IRJPAC/2016/22863.

[31] M. J.; Frisch, Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H.P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J.L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; aramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, J. B.; Foresman, J. V.; Ortiz, J.; Cioslowski, D. J., Fox: Wallingford CT, 2013. Gaussian 09, Revision A.02.

[32] A. D. Becke, Density-functional thermochemistry 3. The role of exact

Emmanuel Israel Edache, Adamu Uzairu, Paul Andrew Mamza, Gideon Adamu Shallangwa

exchange. J. Chem. Phys. 1993, 98 (1993) 5648. DOI: 10.1063/1.464913.

[33] C. Lee, W. Yang, R. G. Parr, Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. Phys Rev B Condens Matter. 37 (1988) 785–789, DOI: 10.1103/physrevb.37.785.

[34] Yap, C. W., PaDEL-Descriptor: An open-source software to calculate molecular descriptors and fingerprints. J Comput Chem. 2011, 32, 1466–1474, DOI:10.1002/jcc.21707.

[35] Gramatica, P.; Chirico, N.; Papa, E.; Cassani, S.; Kovarich. S., QSARINS: A New Software for the Development, Analysis, and Validation of QSAR MLR Models. J. Comput. Chem. 2013, 34, 2121–2132, DOI: 10.1002/jcc.23361.

[36] Gramatica, P.; Cassani, S.; Chirico, N., QSARINS-Chem: Insubria Datasets and New QSAR/QSPR Models for Environmental Pollutants in QSARINS. J.Comput.Chem. 2014, 35, 1036–1044, DOI:10.1002/jcc.23576.

[37] Chirico, N.; Sangion, A.; Gramatica, P.; Bertato, L.; Casartelli, I.; Papa, E., QSARINS-Chem standalone version: A new platform-independent software to profile chemicals for physico-chemical properties, fate, and toxicity. J Comput Chem. 2021, 42 (20), 1–9, DOI: 10.1002/JCC.26551.

[38] Golbraikh, A.; Tropsha, A. Predictive QSAR modeling based on diversity sampling of experimental datasets for the training and test set selection. Mol. Divers. 2002, 5 (4), 231-243, DOI: 10.1023/A:1021372108686.

[39] Delevoye, C., Nilges, M., Dautry-Varsat, A. and Subtil, A. Conservation of the biochemical properties of IncA from Chlamydia trachomatis and Chlamydia caviae: oligomerization of IncA mediates interaction between facing membranes J. Biol. Chem. 2004, 279 (45), 46896-46906, DOI: 10.1074/jbc.M407227200.

[40] Altschul, S. F.; Gish, W.; Miller, W.; Myers E.W.; and Lipman, D. J., Basic local alignment search tool. J. Mol. Biol. 1990, 215 (3), 403-410, DOI: 10.1016/S0022-2836(05)80360-2.

[41] Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; and Tamura, K., MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Mol Biol Evol. 2018, 35, 1547-1549, DOI: 10.1093/molbev/msy096.

[42] Sneath, P. H. A.; and Sokal, R. R., Numerical Taxonomy. Freeman, San Francisco: 1973, p. 1.

[43] Zuckerkandl, E.; and Pauling, L., Evolutionary Divergence and Convergence in Proteins. In: Bryson, V. and Vogel, H.J., Eds., Evolving Genes and Proteins, Academic Press, New York: 1965, p. 97-166, DOI: 10.1016/B978-1-4832-2734-4.50017-6.

[44] Webb, B.; Sali, A., Comparative protein structure modeling using MODELLER. Curr. Protoc. Bioinforma. 2016, 20 (54), 5.6.1–5.6.37, DOI: 10.1002/cpbi.3.

[45] Webb, B.; Sali, A., Protein structure modeling with MODELLER. Protein Structure Prediction. 2014, 1137, 1-15, DOI 10.1007/978-1-4939-0366-5_1.

[46] Bowie, J. U.; Luthy, R.; and Eisenberg, D., "A method to identify protein sequences that fold into a known three-dimensional structure," Science. 1991, 253 (5016), 164–170, DOI: 10.1126/science.1853201.

[47] Eisenberg, D.; Luthy, R.; Bowie, J. U., VERIFY3D: assessment of protein models with three-dimensional profiles. Methods Enzymol. 1997, 277, 396–404, DOI: 10.1016/S0076-6879(97)77022-8.

[48] Laskoswki, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M., PROCHECK: a program to check the stereo chemical quality of protein structures. J Appl Cryst. 1993, 26, 283–291, DOI: 10.1107/S0021889892009944.

[49] Lovell, S. C.; Davis, I. W.; Arendall III, W. B.; Bakker, P. I. W.; de Word, J. M.; Prisant, M. G.; Richardson, J. S.; Richardson, D. C., Structure validation by Cα geometry: φ,ψ and Cβ deviation. Proteins Struct. Funct. Genet. 2003, 50, 437–450, DOI: 10.1002/prot.10286.

[50] Colovos, C.; Yeates, T. O., Verification of protein structures: Patterns of nonbonded atomic interactions. Protein Sci. 1993, 2 (9), 1511–1519, DOI: 10.1002/pro.5560020916.

[51] Wiederstein, M.; Sippl, M. J., ProSA-web: interactive web service for the recognition of errors in three dimensional structures of proteins. Nucleic Acids Res. 2007, 35 (suppl 2), W407-410, DOI:10.1093/nar/gkm290.

Emmanuel Israel Edache, Adamu Uzairu, Paul Andrew Mamza, Gideon Adamu Shallangwa

[52] Ravindranath, P. A; Forli, S.; Goodsell, D. S.; Olson, A. J.; Sanner, M. F., AutoDockFR: Advances in Protein-Ligand Docking with Explicitly Specified Binding Site Flexibility. PLoS Comput Biol. 2015, 11 (12), e1004586, DOI:10.1371/journal.pcbi.1004586.

[53] Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K., Scalable molecular dynamics with NAMD, J. Comput. Chem. 2005, 26, 1781–1802, DOI: 10.1002/jcc.20289 .

[54] Humphrey, W.; Dalke, A.; & Schulten, K., VMD: Visual molecular dynamics. J. Mol. Graph. 1996, 14, 33–38, DOI: 10.1016/0263- 7855(96)00018 -5.

[55] Bai, Q.; Tan, S.; Xu, T.; Liu, H.; Huang, J.; and Yao, X., MolAICal: a soft tool for 3D drug design of protein targets by artificial intelligence and classical algorithm. Briefings in Bioinformatics. 2020, 00 (00), 1–12, DOI: 10.1093/bib/bbaa161.

[56] Zhong-Yu, W., QSAR Study of Arylpyridone Oxime Based on the SVM and Elman Algorithms. Phys. Chem. Res. 2020, 8 (2), 215-223, DOI: 10.22036/pcr.2020.205099.1686.

[57] Oyeneyin, O. E.; Obadawo, B. S.; Metibemu, D. S.; Owolabi, T. O.; Olanrewaju, A. A.; Orimoloye, S. M.; Ipinloju, N.; and Olubosede, O., An Exploration of the Antiproliferative Potential of Chalcones and Dihydropyrazole Derivatives in Prostate Cancer via Androgen Receptor: Combined QSAR, Machine Learning, and Molecular Docking Techniques. Phys. Chem. Res. 2022, 10 (2), 211-223, DOI: 10.22036/pcr.2021.293051.1932.

[58] Albratty, M., Quantitative structure-activity relationship modeling and docking of some synthesized bioactive oxopyrolidines against Staphylococcus aureus, J. Saudi Chem. Soc. 2022, 26 (4), 1-19, DOI: 10.1016/J.JSCS.2022.101509.

[59] Bitam, S.; Hamadache, M.; Hanini, S., 2D-QSAR, docking, molecular dynamics, studies of PF-07321332 analogues to identify alternative inhibitors against 3CLpro enzyme in SARS-CoV disease, J. Biomol. Struct. Dyn. 2022, DOI: 10.1080/07391102.2022.2113822.

[60] Kasmia, R.; Bouachrineb, M.; and Ouammoua, A., Combined 3D-QSAR and Molecular Docking Analysis of Styrylquinoline Derivatives as Potent anti-cancer Agents. Phys. Chem. Res. 2022, 10 (3), 345-362, DOI: 10.22036/PCR.2021.304969.1967.

[61] Adeniji, E. S.; Uba, S.; and Uzairu, A., A Novel QSAR Model for the Evaluation and Prediction of (E)-N'-Benzylideneisonicotinohydrazide Derivatives as the Potent Anti-mycobacterium Tuberculosis Antibodies Using Genetic Function Approach. Phys. Chem. Res. 2018, 6 (3), 479-492, DOI: 10.22036/pcr.2018.115878.1457.

[62] Tropsha, A.; & Golbraikh, A., Predictive QSAR Modeling Workflow, Model Applicability Domains, and Virtual Screening. Curr. Pharm. Des. 2007, 13 (34), 3494–3504, DOI: 10.2174/138161207782794257.

[63] Ho B. K.; and Brasseur, R., "The Ramachandran plots of glycine and pre-proline,". BMC Struct Biol. 2005, 16 (5), 14, DOI: 10.1186/1472-6807-5-14.

[64] Sahu, R.; and Shukla, N. S., In-silico analysis of ndifferent plant protein and their essential compound with sulfonylurea binding protein of B-cell of homo sapiens for curing diabetes mellitus type II disease. Eur. Chem. Bull. 2014, 3, 568-576, DOI: 10.1016/j.sjbs.2022.103302.

[65] Tamura, K.; Stecher, G.; Peterson, D.; Filipski, A.; and Kumar, S., MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 2013, 30, 2725–2729, DOI: 10.1093/molbev/mst197.

[66] Trott, O.; Olson, A. J., AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. J. Comput. Chem. 2010, 31,455-461, DOI: 10.1002/jcc.21334.

[67] Daina, A.; Michielin, O.; Zoete, V., SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules, Sci. Rep. 2017, 7, 42717, DOI: 10.1038/srep42717.

[68] Daina, A.; Zoete, V., A boiled-egg to predict gastrointestinal absorption and brain penetration of small molecules. ChemMedChem. 2016, 11, 1117–1121, DOI: 10.1002/cmdc.201600182.

[69] Lipinski C. A., Lead-and drug-like compounds: the rule-of-five revolution. Drug Discov Today Technol. 2004, 1, 337-341, DOI: 10.1016/J.DDTEC.2004.11.007.

[70] Veber, D. F.; Johnson, S. R.; Cheng, H. Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D., Molecular properties that influence the oral bioavailability of drug candidates. J Med Chem. 2002, 45, 2615-2623, DOI: 10.1021/JM020017N.

[71] Ghose, A. K.; Viswanadhan, V. N.; Wendoloski, J. J., A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. A qualitative and quantitative characterization of known drug databases. J Comb Chem. 1999, 1, 55-68, DOI: 10.1021/cc9800071.

[72] Egan, W.J.; Merz, K.M., Jr; Baldwin, J.J. Prediction of drug absorption using multivariate statistics. J. Med. Chem. 2000, 43 (21), 3867-3877, DOI: 10.1021/JM000292E.

[73] Ertl, P.; and Schuffenhauer, A., Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. J. Cheminform. 2009, 1 (1), 1-11, DOI: 10.1186/1758-2946-1-8

[74] Banks, W. A., Characteristics of compounds that cross the blood brain barrier. BMC Neurol. 2009, 9 (Suppl. 1), 5–9. DOI: 10.1186/1471-2377-9-S1-S3.

[75] Ouassaf, M.; Belaidi, S.; Khamouli, S.; Belaidi, H.; Chtita, S., Combined 3D-QSAR and molecular docking analysis of thienopyrimidine derivatives as Staphylococcus aureus inhibitors, Acta Chim. Slov. 2021, 68, 289–303, DOI: 10.17344/acsi.2020.5985.

[76] Domínguez-Villa, F. X.; Duran-Iturbide, N. A.; Avila-Zarraga, J. G., Synthesis, molecular docking, and in silico ADME/Tox profiling studies of new 1-aryl-5-(3-azidopropyl) indol-4-ones: potential inhibitors of SARS CoV-2 main protease, Bioorg. Chem. 2021, 106, 104497, DOI: 10.1016/j.bioorg.2020.104497.

[77] De Groot, M. J., Designing better drugs: predicting cytochrome P450 metabolism. Drug Discov Today. 2006, 11 (13), 601–606, DOI:10.1016/j.drudis.2006.05.001.

[79] Zanger, U. M.; Schwab, M., Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. Pharmacol. Ther. 2013, 138 (1), 103–141, DOI: 10.1016/j.pharmthera.2012.12.007.

[80] Daoui, O.; Elkhattabi, S.; Chtita, S.; Elkhalabi, R.; Zgou, H.; Benjelloun, A. T., QSAR, molecular docking and ADMET properties in silico studies of novel 4,5,6,7-tetrahydrobenzo[D]-thiazol-2-Yl derivatives derived from dimedone as potent anti-tumor agents through inhibition of C-Met receptor tyrosine kinase. Heliyon. 2021, 7, e07463, DOI: 10.1016/j.heliyon.2021.e07463.

[81] Bouamrane, S.; Khaldan, A.; Hajji, H.; El-mernissi, R.; Maghat, H.; Ajana, M. A.; Sbai, A.; Bouachrine, M.; Lakhlifi, T., 3D-QSAR, molecular docking, molecular dynamic simulation, and ADMET study of bioactive compounds against candida albicans. Mor. J. Chem. 2022, 10, 523-541, DOI: 10.48317/IMIST.PRSM/MORJCHEM-V10I3.33141.

[82] Yan, Z.; & Caldwell, G. W., Metabolism profiling, and cytochrome P450 inhibition & induction in drug discovery. Curr. Top. Med. Chem. 2001, 1, 403-425, DOI: 10.2174/1568026013395001.

[83] Shahzadi, I., Zahoor, A.F., Tüzün, B., Mansha, A., Anjum, M.N., Rasul, A., Irfan, A., Kotwica-Mojzych, K. and Mojzych, M., 2022. Repositioning of acefylline as anti-cancer drug: Synthesis, anticancer and computational studies of azomethines derived from acefylline tethered 4-amino-3-mercapto-1, 2, 4-triazole. Plos one, 17(12), p.e0278027.

[84] Sarkı, G., Tüzün, B., Ünlüer, D. and Kantekin, H., 2023. Synthesis, characterization, chemical and biological activities of 4-(4-methoxyphenethyl)-5-benzyl-2-hydroxy-2H-1, 2, 4-triazole-3 (4H)-one phthalocyanine derivatives. Inorganica Chimica Acta, 545, p.121113.

[85] Choong, E.; Dobrinas, M.; Carrupt, P.-. A.; Eap, C. B., The permeability P-glycoprotein: a focus on enantioselectivity and brain distribution. Expert Opin. Drug Metab. Toxicol. 2010, 6, 953–965, DOI: 10.1517/17425251003789394.

Emmanuel Israel Edache, Adamu Uzairu, Paul Andrew Mamza, Gideon Adamu Shallangwa

[86] Miller, B. R., 3rd.; McGee, T. D., Jr.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E., MMPBSA.Py: An Efficient Program for End-State Free Energy Calculations. J. Chem. Theory Comput. 2012, 8 (9), 3314−3321, DOI: 10.1021/ct300418h.

[87] Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E. 3rd., Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. Acc. Chem. Res. 2000, 33, 889−897, DOI: 10.1021/ar000033j.