



Original Research Article

In Silico Identification of Some Novel Ketoamides as Potential Pan-Genotypic HCV NS3/4A Protease Inhibitors with Drug-Likeness, Pharmacokinetic ADME Profiles and Synthetic Accessibility Predictions

Stephen Ejeh* , Adamu Uzairu, Gideon Adamu Shallangwa, Stephen Eyije Abechi, Muhammad Tukur Ibrahim

Department of Chemistry, Faculty of Physical Science, Ahmadu Bello University, P.M.B. 1044, Zaria, Kaduna State, Nigeria

ARTICLE INFO

Article history

Submitted: 12 February 2022

Revised: 06 April 2022

Accepted: 14 April 2022

Available online: 18 April 2022

Manuscript ID: [AJCA-2202-1302](https://doi.org/10.22034/AJCA.2022.329332.1302)

Checked for Plagiarism: Yes

DOI: [10.22034/AJCA.2022.329332.1302](https://doi.org/10.22034/AJCA.2022.329332.1302)

KEYWORDS

Structure-based design

HCV NS3/A4 protease

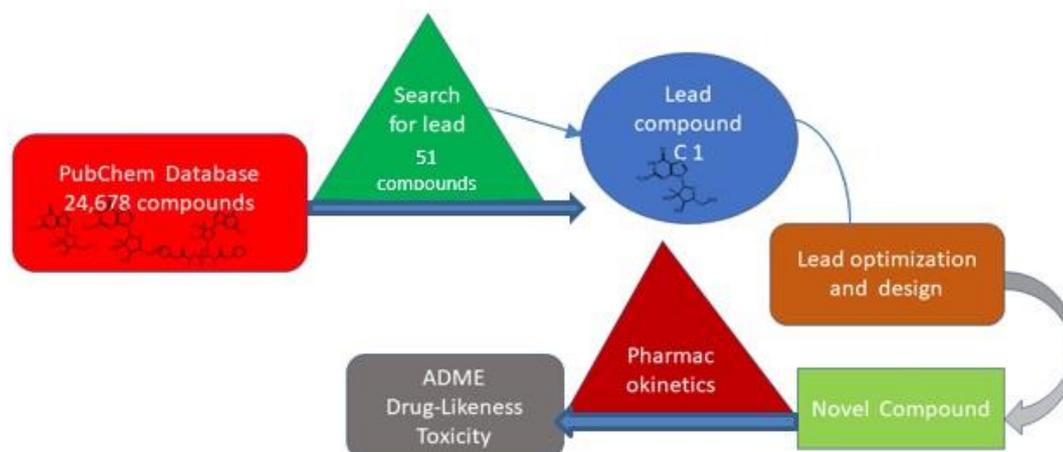
Drug-likeness

In silico approach

ABSTRACT

Hepatitis C virus (HCV) infection promotes death rates worldwide. As a result, there is a constant need to improve current HCV therapy and produce new drugs. The NS3/A4 enzyme plays a critical role in the HCV entire lifespan and proliferation. Consequently, inhibitors of the HCV NS3/A4 enzyme are a great spot to start exploring new drug candidates. In this study, the high throughput in silico screening of the Pubchem database was used to analyze a set of ketoamides as HCV NS3/A4 enzyme inhibitors to find a novel potential drug as a lead candidate. To Voxilaprevir as a reference medicine, our findings revealed that three HCV NS3/A4 protease inhibitors (Pubchem CID: 44158040, 44158107, and 11479303) were identified as the best drugs for blocking hepatitis C virus NS3/A4 protease. The QSAR was performed to study the relationships between the structural features of the targets and their binding affinity by developing statistical models. The reported compounds had a higher binding affinity for the target receptor than Voxilaprevir, the reference drug. This study could be significant in understanding the physicochemical and binding affinity of HCV NS3/A4 inhibitors to find new and improved HCV antiviral drugs.

GRAPHICAL ABSTRACT



* Corresponding author: Ejeh, Stephen

✉ E-mail: ejehstephen@gmail.com

© 2022 by SPC (Sami Publishing Company)

Introduction

Hepatitis C virus (HCV) infection promotes death rates worldwide. Persistent HCV infection is a serious health problem worldwide, impacting up to 3% of the populace and killing over 300,000 people annually [1-3]. Its gradual spread and hard-to-detection make it a hidden pandemic, and most infections proceed into a persistent condition that lasts for years. About 60%-80% of people infected with HCV suffer persistent hepatitis, with 20% developing cirrhosis and about 2%-5% of patients dying from liver cirrhosis and malignancy [4,5].

In 2019, there were approximately 58 million severe recorded cases around the globe, with 75% of them occurring in low- and middle-income countries (LMICs) [6]. New HCV infections are occurring in high-risk populaces worldwide, including drug users and men who have sex with males [7]. HCV is primarily transmitted in Sub-Saharan Africa through improper medical practices and infected blood transfusions. Needlestick injuries in medical professionals, mother-to-child spread (MTCT), and societal behaviors like piercing and tattooing are also possible transmission routes [7,8].

The hepatitis C virus gene contains a polyprotein with around 3000 amino acids that are separated into two groups: The structural and non-structural (NS) proteins are the critical targets for antiviral drug development because of their critical role in HCV virus propagation [2,9]

Current HCV treatment has a poor sustained viral chance of success, the rapid development of drug resistance, and significant side effects, leading to treatment discontinuation [10]. As a response, more effective anti-HCV drugs having minimal adverse effects are desperately required. The discovery of the most critical information to facilitate the identification of unique bioactive compounds that could be utilized to cure HCV infection and then provide medically successful

therapy is undoubtedly a critical global health priority [11].

Pharmaceuticals can be designed *in silico* using molecular docking by enhancing lead candidates targeted against specific receptors. A docking method can be used to find the best small molecule (ligand) binding mode to an active receptor domain. As a result, drug research aims to generate molecules that are bound to a receptor more strongly than the native ligand [12]. The biological reaction catalyzed by the target molecule can be modified or inhibited in this way. Cheminformatics has become extensively utilized in pharmaceutical research due to the many advantages of being quite time-consuming, cost-effective, and high reliability in the computational screening and modeling of viable drugs. High-throughput screening techniques that employ *in vitro* tests to determine the biological responses of a variety of molecules towards a specific macromolecule are often used to identify drugs by chance in a trial-and-error procedure. This is a lengthy and expensive process. If the target's 3D structure is known, a program to mimic docking can help drug discovery. This *in silico* study identifies promising therapeutic candidates more quickly and at a lower cost by virtual screening of drug repositories.

Laboratory investigations (synthesis), cytotoxic screening, clinical trials, and other procedures could be utilized to investigate the drug molecules discovered using the *in silico* approach. Most docking approaches employ an energy-based ranking system to find optimum thermodynamically stable ligand orientation when a ligand is attached to a target. Lower energy scores are assumed to indicate better protein-ligand bindings when compared to higher energy scores. Consequently, docking can be conceived as an iterative algorithm aiming to find the ligand-binding formation with the least energy [13]. The present research seeks to discover possible pan-genotypic HCV NS3/4A protease inhibitors with enhanced

pharmacokinetic profiles and lower hepatotoxicity threats. In this article, we used a docking study to screen the Pubchem database for possible lead (hit) molecules. We further build a QSAR model to find underlying molecular characteristics that can facilitate the discovery of new biomolecules as anti-HCV medications.

Materials and Methods

Dataset

53 ketoamides compounds were used as possible inhibitors of the HCV NS3/4A enzyme, which were retrieved from a website (<https://pubchem.ncbi.nlm.nih.gov/>). Table SM1 in the Supplementary Material contains the datasets and Pubchem compound identification number (CID) and Energy score (E_{score}).

Optimization

Optimizing a molecule's equilibrium or concept energy geometry is referred to as optimization. Chemdraw v 12.0 was used to sketch the structures of the molecules, which was then

transferred to the Spartan 14 program to minimize the geometry using B 3 LYP level of theories and 6-31G as the basis set [13].

Target preparation and docking procedure

The receptor utilized in this research was retrieved from a web database (PDB) uploaded on 14th February 2019 by Taylor, and a co-worker [14] with the relevant details: PDB ID (6NZT), the resolution (1.40 Å), R-value free (0.197), the R-value work (0.176), and the R-value observed (0.177). The native molecule was detached from the receptor before being introduced into the Pyrx software for docking. As discussed in our prior research, the comprehensive docking strategy was applied [15]. Figure 1 illustrates the receptor.

Theoretical background of the energy terms in the docking Algorithm

The energy terms that make up the binding energy scoring function, E_{score} , are presented in Equation 1.

$$E_{score} = E_{inter} + E_{intra} \quad (1)$$



PDB ID: 6NZT

Figure 1. Structure of the receptor (HCV NS3/A4 protease) in complex with native ligand (Voxilaprevir)

E_{inter} represents protein-ligand interaction energy, shown in Equation 2.

$$E_{inter} = \sum_{i \in \text{ligand}} \sum_{j \in \text{protein}} \left[E_{PLP}(r_{ij}) + 332.0 \frac{q_i q_j}{4r_{ij}^2} \right] \quad (2)$$

All heavy atoms in the ligand and protein, together with any coenzyme atoms, may be present. The E_{PLP} word refers to a piecewise linear potential, further explained down. The second phrase describes attractive forces between charged atoms. It is a Coulomb potential with the dielectric constant $D(r)=4r$, which varies with distance. The last term in Equation 2 gives the clash penalty of the electrostatic energy in kcal/mol and a threshold value with a distance of 2.0 Å, so that distances < 2.0 Å confirm that no energy input is greater than the clash penalty.

E_{intra} is the internally potential of the ligand, given by Equation 3.

$$E_{intra} = \sum_{i \in \text{ligand}} \sum_{j \in \text{protein}} E_{PLP}(r_{ij}) + \sum_{\text{flexible bonds}} A[1 - \text{Cos}(m \cdot \theta - \theta_0)] + E_{clash} \quad (3)$$

The double summation applies to all-atom sets in the ligand, omitting those bound by two or fewer bonds. The second part is a torsional energy factor parameterized by the bonded atoms' hybridization kinds. θ represents the bond's torsional angle in degrees. If many torsions could be established, the mean of the torsional energy bond input was employed. If the difference across two heavy atoms (more than two bonds apart) < 2.0 Å, the last part, E_{clash} , imposes a penalty of 1000. As a result, the E_{clash} term penalizes impracticable ligand conformations.

QSAR method

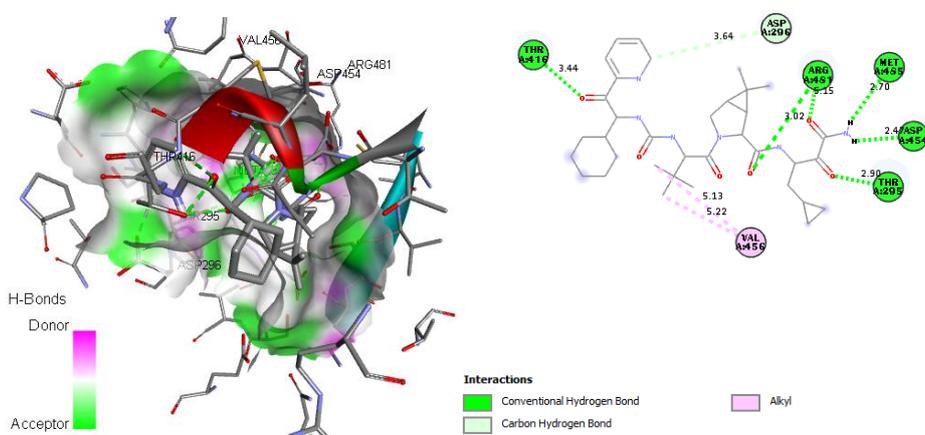
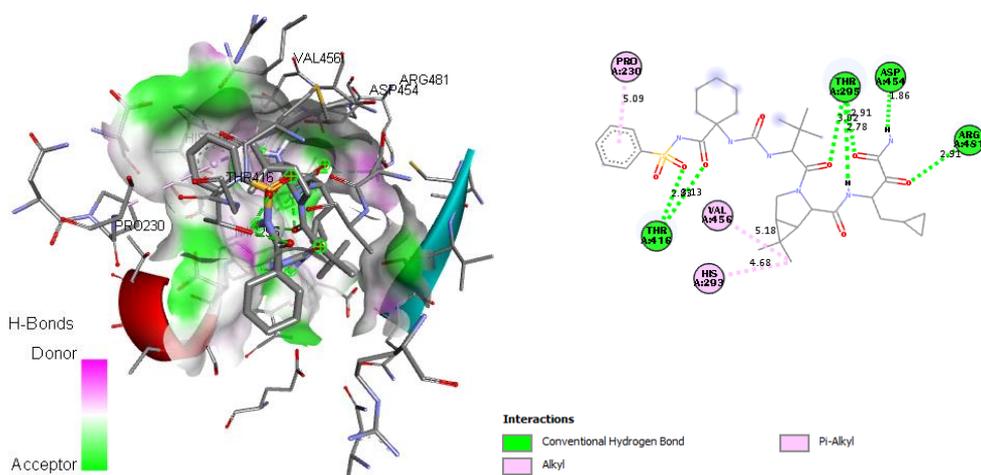
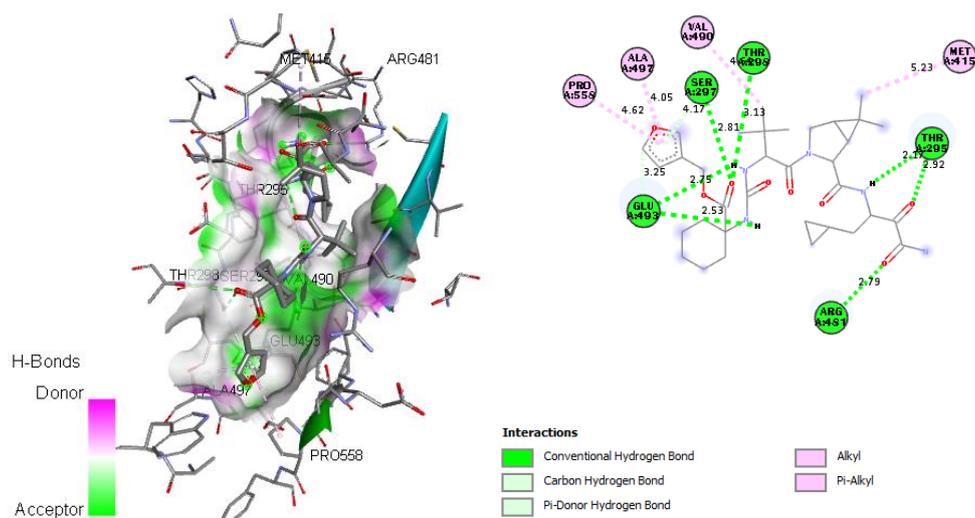
The datasets were split into model building data and model validation data of 70% and 30%, respectively, as outlined in our prior study [2].

Accessing the identified entities

The web applications SwissADME (www.swissadme.ch/) and pkCSM (<http://structure.bioc.cam.ac.uk/pkcsm>) were used to quantify the drug-likeness and pharmacokinetics ADMET profile of these newly discovered entities to verify their reliability.

Result and Discussions

In this study, the most thermodynamically viable ligand geometry was identified, and lead (hit) molecules were discovered as HCV NS3/4A enzyme inhibitors using computational docking models, with docking scores ration as binding energy. The majority of the screened molecules are potent against the HCV NS3/4A enzyme, and the best HCV NS3/4A enzyme inhibitors were ranked according to their Lower energy docking score. The docking results revealed that molecules 6, 15, and 45 were the best hits because they established covalent and noncovalent interactions at the receptor's binding domain. Also evidenced by the energy terms and their binding sequences (Figures 2, 3, and 4, respectively) seem to be consistent with that of the standard drugs (Voxilaprevir) (Figure 5). Those molecules pose maximum interactions with the critical amino acid residues in the target binding domain as detailed in Tables SM2, SM3, SM4, SM5, respectively.



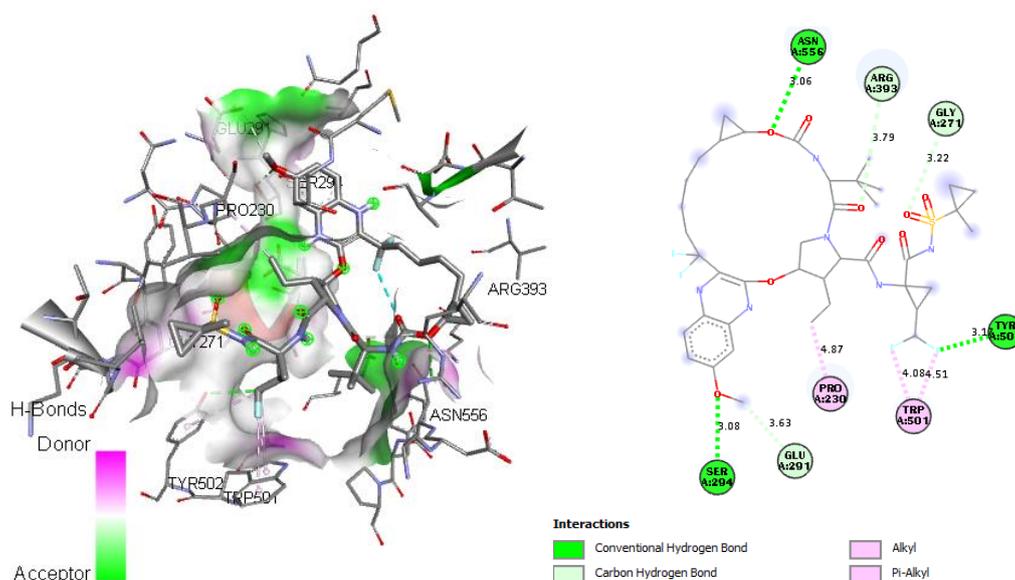


Figure 5. The 3D and 2D representation of reference compound (R) with the target receptor (PDB ID: 6NZT)

The identified molecules were then evaluated for drug-likeness using the Lipinski concept of five, which essentially asserts that if a compound infringes over two of the characteristics listed in Table 1, it will be improperly absorbed [15] and is one of the most valuable approaches in the initial stages of drug improvement. Because the discovered molecules did not break more than two of Lipinski's threshold, they are likely to be classified as drug-like molecules. Additionally, ABS standards were used to evaluate the template and all discovered compounds, including the standard molecule, with values of 0.17 (17%). The reliability coefficient of a molecule with excellent

permeability and bioavailability pattern reflects total compliance with the Lipinski idea of five, which gives a value of more than 10%, is reliable [16]. The synthetic accessibility of the compounds in consideration was also assessed on a scale of 1 to 10 (very simple to synthesize to difficult to synthesize) [16]. As shown in Table 1, the discovered compounds have synthetic accessibility ranging from 6.02 to 6.38, making them moderately easy to synthesize.

Using web-based pkCSM programs, the ADMET profiles of discovered molecules and the standard molecule have also been analyzed and listed in Table 2. The BBB penetration ratings are used to

Table 1. Derived drug-likeness focusing on Lipinski's rule

Molecule CID	MWt	HBA	HBD	Log p	TPSA	Violation	BA	SA
Lipinski's Rule	<500	≤10	≤5	≤5	<140			
44158040	655.78	8	4	3.29	190.14	2	0.17	6.28
44158107	714.87	8	5	3.17	222.32	2	0.17	6.38
11479303	650.81	7	4	3.52	180.66	2	0.17	6.02
R	868.93	15	3	4.17	203.60	3	0.17	8.83

CID Compound identification number, MWt molecular weight, HBA hydrogen bond acceptor, HBD hydrogen bond donor, TPSA topological polar surface area, BA bioavailability, SA synthetic accessibility, T Template molecule, R Reference Molecule (Voxilaprevir)

establish if a substance will pass through the blood-brain barrier (BBB). A Log BB > 0.3 suggests that a molecule can readily flow through the blood-brain barrier, whereas a Log BB < -1 suggests that such a molecule was poorly dispersed [15]. In this study, all of the molecules investigated efficiently penetrated the blood-brain barrier (Table 2).

A compound with a Papp > $8 \times 10^{-6} \text{ cm/s}$ is reported to have the highest Caco-2 permeability, which can be used to estimate orally administered permeability in vitro [15]. According to the findings of this investigation, all of the molecules investigated have only a moderate cell uptake in Caco-2 cells.

Table 2. ADMET profiles of discovered molecules, and reference molecule

Molecule CID	BBB	CaCo2	HIA	MRTD
44158040	4.462	0.705	63.111	-0.484
44158107	4.411	0.852	57.160	-0.216
11479303	2.259	0.842	64.097	-0.522
R	3.027	1.200	89.420	-0.392

^{CID}Compound identification number, ^{BBB}blood-brain barrier (Log BB), ^{CaCo2}cell absorption (Log Papp in 10^{-6} cm/s), ^{HIA}Human Intestinal Absorption (% absorbed), ^{MRTD}maximum tolerated dose (Log mg/kg/day), ^TTemplate molecule, ^RReference Molecule (Voxilaprevir)

Equating the binding affinities of the molecules with their structural features and constructing comprehensive statistical models using both linear (multiple linear regression and partial least squares regression) and non-linear (support

The HIA estimates how much of a chemical will be consumed via the human gut. Compounds with an absorption rate < 30% are considered poorly consumed [17]. The investigation results demonstrated that all of the substances investigated have high HIA ratings.

The maximum tolerated dose (MRTD) is a metric for determining a chemical's toxic dosage limit in humans. A low MRTD for a molecule is less than or equal to 0.477 log(mg/kg/day), whereas a high MRTD is greater than 0.477 log(mg/kg/day) [15,16]. Table 2 reveals that the lethality of all of the chemicals studied in this study was minimal.

vector machine, and neural network regression) simulation techniques which give an insight into the nature of the interaction and the model has been proposed using the equation below:

$$\text{binding energy} = -13.15902 - 0.31679 * \text{n6Ring} - 0.6008 * \text{SpMax6_Bhs} + 0.00021 * \text{VR1_Dzs} - 0.72426 * \text{GGI6} + 2.79433 * \text{SpMax8_Bhv}$$

The model fulfilled the requirements specified in Table 3 as a baseline, indicating that it can make accurate predictions. The model also fulfilled Tropsha and the OECD criteria [12,18] as it explains above 90% and predicts above 70% of the variations of the HCV NS3/4A protease inhibitors with their binding affinity using both

linear and non-linear simulation techniques as described in Table 3. It shows that the model accurately extrapolated the data and can estimate its appropriate building data. It estimated more than 70% of the data and thus fulfilled the requirement of 60% [12].

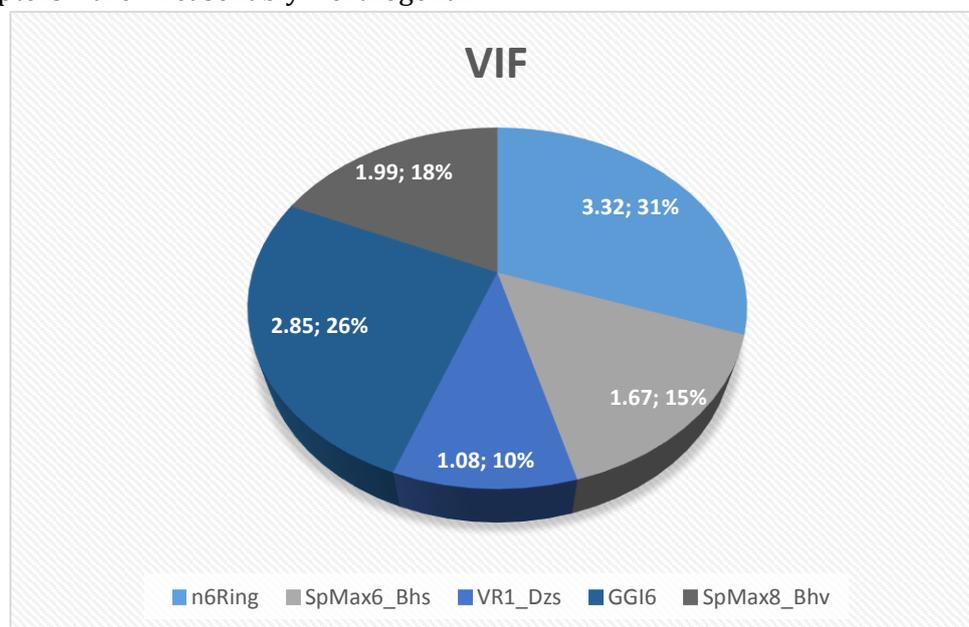
Table 3. Model external validation parameters and their threshold values using linear (MLR and PLS) and non-linear (SVM and ANN) methods

Name	Threshold value	MLR	PLS	SVM	ANN
Pearson Correlation (r)	$r > 0.6$	0.923	0.923	0.955	0.931
Pearson Correlation Squared (r^2)	$r^2 > 0.6$	0.777	0.777	0.830	0.790
Spearman Rank Correlation (ρ)	$\rho > 0.6$	0.954	0.954	0.982	0.952
Mean Squared Deviation (MSD)	Smaller the better	0.052	0.052	0.047	0.051
Root Mean Squared Deviation (RMSD)	Smaller the better	0.228	0.228	0.217	0.227
Cross Validated Squared CC (q^2)	$q^2 > 0.5$	0.776	0.776	0.802	0.778
External Pearson Correlation Squared (r_{ext}^2)	$r_{ext}^2 > 0.6$	0.760	0.760	0.814	0.768

^{MLR}Multiple Linear Regression, ^{PLS}Partial Least Squares Regression, ^{SVM}Support Vector Machine, ^{ANN}Neural Network Regression

The meaning and nature of the descriptors were presented in supplementary material Table SM6 with the descriptor's variance inflation factor (VIF) for better understanding. Also, the variance inflation factor (VIF) is represented pictorially in Figure 6, demonstrating that all the descriptors have VIF coefficients lesser than 5. Therefore, the developed model is statistically significant, and the descriptors are reasonably orthogonal.

Figures 7, 8, 9, and 10 reveal the estimated vs. actual binding affinity throughout the original data for Multiple Linear Regression, Partial Least Squares Regression, Support Vector Machine, and Neural Network Regression, respectively. The Figures revealed a close relationship between actual and predicted binding affinity based on the linearity of these graphs.

**Figure 6.** Chart describing the variance inflation factor (VIF) of the model

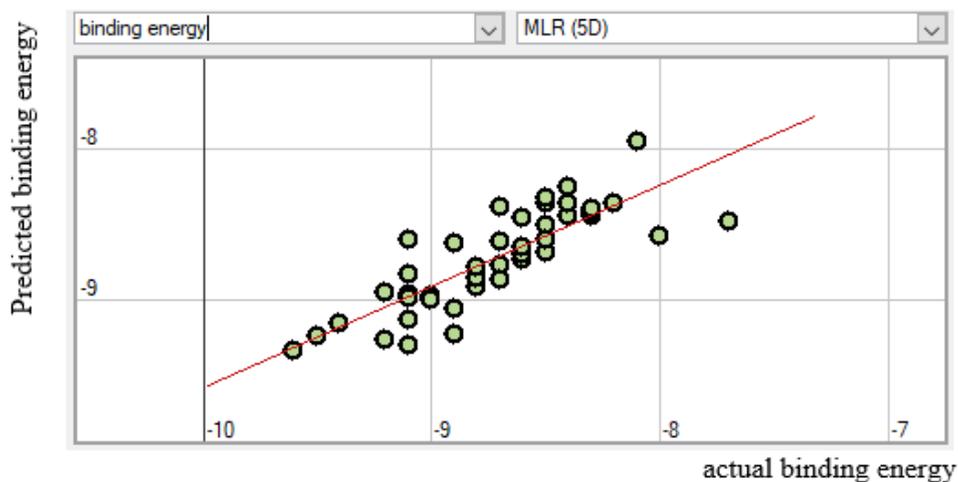


Figure 7. MLR Predicted binding energy against actual binding energy

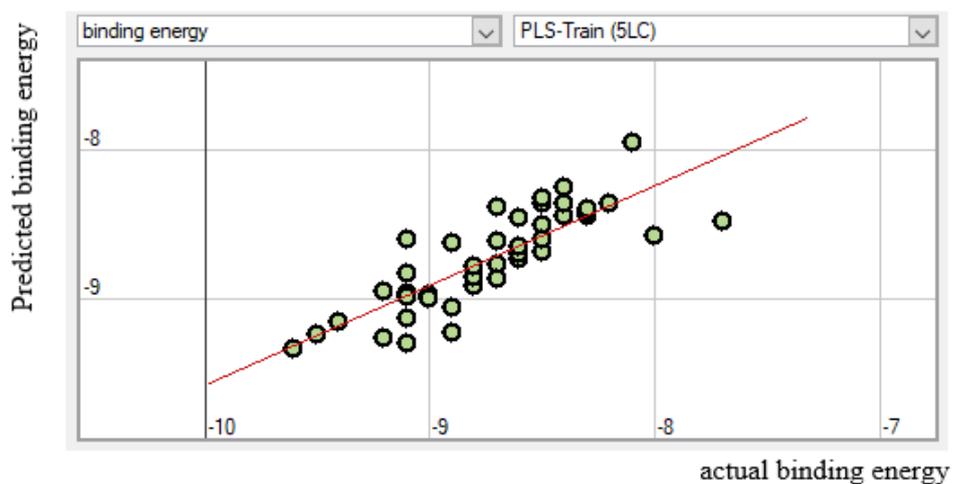


Figure 8. PLS Predicted binding energy against actual binding energy

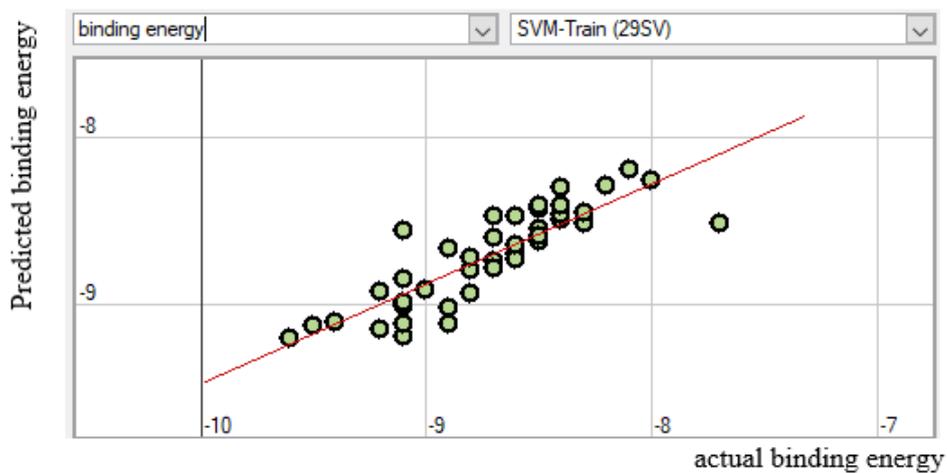


Figure 9. SVM predicted binding energy against actual binding energy

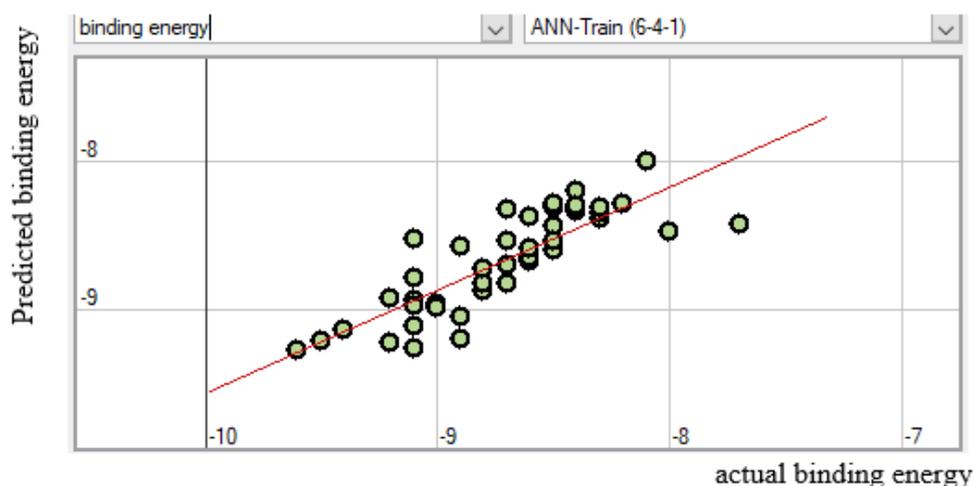


Figure 10. ANN Predicted binding energy against actual binding energy

Conclusions

Our findings revealed three HCV NS3/A4 protease inhibitors (Pubchem CID: 44158040, 44158107, and 11479303) were the best drugs for blocking the HCV NS3/A4 enzyme in comparison to Voxilaprevir as a reference medicine. By establishing statistical models, the QSAR was determined to investigate the correlations between the molecular structures of the targets and their binding affinity. The identified molecules exhibited better binding affinity than Voxilaprevir, the reference drug, with the target receptor. ADMET estimate revealed improved druggability with lower MRTD of those identified molecules. Based on the molecular docking, QSAR model, drug-likeness, and ADMET analysis, this investigation could be useful to understand the structural profile and binding affinity of HCV NS3/A4 protease inhibitors to search for novel and better HCV antiviral drugs.

Acknowledgment

Authors are grateful to the Physical chemistry research team head by Prof. Uzairu of department of the chemistry Ahmadu Bello University Zaria-Nigeria for their meaningful contributions.

Disclosure statement

The authors reported no potential conflict of interest.

ORCID

Stephen Ejeh : 0000-0002-3065-6575

Reference

- [1] H.A. Abuelizz, M. Marzouk, A.H. Bakheit, R. Al-Salahi, *RSC Adv.*, **2020**, *10*, 35820–35830. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [2] S. Ejeh, A. Uzairu, G.A. Shallangwa, S.E. Abechi, *Futur. J. Pharm. Sci.*, **2021a**, *7*, 219. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [3] M. Zając, I. Muszalska, A. Sobczak, A. Dadej, S. Tomczak, A. Jelińska, *Eur. J. Med. Chem.*, **2019**, *165*, 225–249. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [4] B. Hajarizadeh, J. Grebely, G.J. Dore, *Nat. Rev. Gastroenterol. Hepatol.*, **2013**, *10*, 553–562. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [5] W. Li, H. Si, Y. Li, C. Ge, F. Song, X. Ma, Y. Duan, H. Zhai, *Journal of Molecular Structure*, **2016**, *1117*, 227–239. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [6] World Health Organization, *Accelerating access to hepatitis C diagnostics and treatment: overcoming barriers in low-and*

- middle-income countries: global progress report 2020, **2021a** [[Google Scholar](#)], [[Publisher](#)]
- [7] World Health Organization, *Global progress report on HIV, viral hepatitis and sexually transmitted infections, 2021: accountability for the global health sector strategies 2016–2021: actions for impact: web annex 2: data methods*, **2021b**. [[Google Scholar](#)], [[Publisher](#)]
- [8] World Health Organization, *Progress report on HIV, viral hepatitis and sexually transmitted infections 2019: accountability for the global health sector strategies, 2016–2021*, **2019**. [[Google Scholar](#)], [[Publisher](#)]
- [9] Z. Qin, A. Yan, *SAR QSAR Environ. Res.*, **2020**, *31*, 281–311. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [10] P.J. Therese, D. Manvar, S. Kondepudi, M.B. Battu, D. Sriram, A. Basu, P. Yogeeswari, N. Kaushik-Basu *J. Chem. Inf. Model.*, **2014**, *54*, 539–552. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [11] Z.Y. Algamal, M.H. Lee, A.M. Al-Fakih, M. Aziz, *J. Chemom.*, **2017**, *31*, e2889. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [12] S. Ejeh, A. Uzairu, G.A. Shallangwa, S.E. Abechi, *Bull. Nat. Res. Cent.*, **2021b**, *45*, 3. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [13] Y. Shao, L.F. Molnar, Y. Jung, J. Kussmann, C. Ochsenfeld, S.T. Brown, A.T. Gilbert, L.V. Slipchenko, S.V. Levchenko, D.P. O'Neill, *Phys. Chem. Chem. Phys.*, **2006**, *8*, 3172–3191. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [14] J.G. Taylor, S. Zipfel, K. Ramey, R. Vivian, A. Schrier, K.K. Karki, A. Katana, D. Kato, T. Kobayashi, R. Martinez, *Bioorg. Med. Chem. Let.*, **2019**, *29*, 2428–2436. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [15] D.E. Pires, T.L. Blundell, D.B. Ascher, *J. Med. Chem.*, **2015**, *58*, 4066–4072. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [16] A.B. Umar, A. Uzairu, G.A. Shallangwa, S. Uba, *Bull. Natl. Res. Cent.*, **2020**, *44*, 1–11. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [17] S. Ejeh, A. Uzairu, G.A. Shallangwa, S.E. Abechi, *Chemistry Africa*, **2021c**, *4*, 563–574. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [18] F. Yan, T. Liu, Q. Jia, Q. Wang, *Sci. Total Environ.*, **2019**, *663*, 560–567. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]

HOW TO CITE THIS ARTICLE

Stephen Ejeh*, Adamu Uzairu, Gideon Adamu Shallangwa, Stephen Eyije Abechi, Muhammad Tukur Ibrahim. In Silico Identification of Some Novel Ketoamides as Potential Pan-Genotypic HCV NS3/4A Protease Inhibitors with Drug-Likeness, Pharmacokinetic ADME Profiles, and Synthetic Accessibility Predictions. *Adv. J. Chem. A*, **2022**, 5(3), 197-207.

DOI: [10.22034/AJCA.2022.329332.1302](https://doi.org/10.22034/AJCA.2022.329332.1302)

URL: http://www.ajchem-a.com/article_148357.html