JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 13(3), SEPTEMBER, 2025
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbuftejoste.com.ng

# Quantum Modelling for Anti-Human African Trypanosomiasis Activity of Substituted 2 - Phenylimidazopyridines; QSAR Approach

Hassan S., Oladunni N., Uwaiya E., Magaji I. Y., Amos A. K., Omotola M. B.
*Department of Science Laboratory Technology,*
*Nigerian Institute of Leather and Science Technology, Zaria*

**ABSTRACT**
*Human African Trypanosomiasis (HAT), commonly known as sleeping sickness, remains a critical public health concern in sub-Saharan Africa. It is caused by the protozoan parasites Trypanosoma brucei gambiense and Trypanosoma brucei rhodesiense, transmitted by the tsetse fly (Glossina). HAT is a disease driven by an underlying lack of effective treatment options, with current therapies suffering from high toxicity, poor bioavailability, and resistance issues. QSAR analysis was performed on the anti-trypanosomal activity of substituted 2-phenylimidazopyridines. A large number of molecular descriptors were calculated using semi-empirical methods, and a genetic function algorithm (GFA) approach was used to develop the best predictive model with high statistical significance ($R^2$ = 0.89228, Radj = 0.87074, $Q^2$ = 0.81896, cRp = 0.15428, $R^2$pred = 0.72734, $r^2_0$ = 0.82387, $r^2_0$ = 0.81947, $(r^2_0 - r^2_0) / r^2$ = 0.01695, $r^2_0 - r^2_0$ = 0.0044, k = 0.71507, $r^2 - r^2_0 + r^2$ = 0.02219). The study reveals that the molecular descriptors (AMR, AATSC7c, and E3u) significantly contribute to the enhanced anti-trypanosomal activity of these compounds, providing valuable insights for future drug design.*

**ARTICLE INFO**
*Article History*
Received: March, 2025
Received in revised form: June, 2025
Accepted: August, 2025
Published online: September, 2025

**KEYWORDS**
QSAR, Model, Trypanosomiasis, Activity

## INTRODUCTION

Human African Trypanosomiasis (HAT), commonly known as sleeping sickness, remains a critical health issue in rural sub-Saharan Africa, transmitted by the tsetse fly (*Glossina*). Despite near eradication efforts in the mid-1960s, HAT experienced a resurgence in the late 1990s due to factors such as inadequate sanitation and favorable habitats for the tsetse fly vector, particularly in countries like the Democratic Republic of the Congo (DRC), Angola, Central African Republic, southern Sudan, and Uganda [19; 22]. The disease is caused by *Trypanosoma brucei gambiense* and *Trypanosoma brucei rhodesiense*, with other trypanosome species such as *T. brucei brucei*, *T. congolense*, and *T. evansi* also implicated [19].

HAT progresses through two stages: the hemolymphatic stage, characterized by peripheral infection and non-specific symptoms, and the central nervous system (CNS) stage, where the parasite crosses the blood-brain barrier (BBB) [5]. Despite significant progress in reducing the number of reported cases to historically low levels, the disease remains endemic in parts of sub-Saharan Africa, with the Democratic Republic of Congo (DRC) being a particularly affected region [8].

Human African Trypanosomiasis (HAT) continues to be a significant public health challenge in sub-Saharan Africa. Current treatments are hampered by high toxicity, poor oral bioavailability, and complex administration, particularly in late-stage disease where the parasite affects the central nervous system (CNS) [23; 8]. Existing drugs are stage-specific, have severe side effects, and face challenges with blood-brain barrier (BBB) penetration [34]. The high antigenic variation of *Trypanosoma brucei* complicates vaccine development, intensifying the need for new, effective drugs [24; 23]. Despite extensive control efforts, the elimination target for

*Corresponding author: Hassan S.*
hassansamuel84@gmail.com
*Department of Science Laboratory technology, Nigerian Institute of Leather and Science Technology, Zaria.*

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 13(3), SEPTEMBER, 2025
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbuftejoste.com.ng

HAT remains elusive due to challenges such as inadequate vector control and difficulties in achieving comprehensive case detection and treatment [5].

QSAR modeling has become an essential tool in developing new anti-HAT drugs, allowing researchers to predict compound efficacy and toxicity before conducting costly experimental studies [17]. These models analyze molecular descriptors such as hydrophobicity, electronic distribution, and steric factors to optimize chemical structures for improved activity against T. brucei [2]. Advances in machine learning algorithms have further enhanced the predictive power of QSAR models, facilitating the identification of promising drug candidates for future development [35].

## METHODOLOGY

### Data Collection

60 compounds of substituted 2-Phenylimidazopyridines, along with their biological activities against Human African Trypanosomiasis (HAT), were collected from well-established literature [36].

### Molecular Optimization and Calculation of

### *Molecular Descriptors*

Molecular descriptors are numerical representations of chemical information encoded within a molecular structure. The process involved finding the lowest energy geometry of the molecules. Geometry optimization was performed using Spartan "14" V1.1.2 on an HP Pavilion with the Microsoft Windows 7 Ultimate operating system. Quantum chemical and constitutional descriptors were calculated using Spartan "14" V1.1.2 software, while topological and geometrical descriptors were computed using PaDE Descriptor 2.18 software.

### *Data Set Division*

The compounds were randomly divided into a training set and a test set. The training set was used to develop the QSAR model, while the test set was used for external validation of the model.

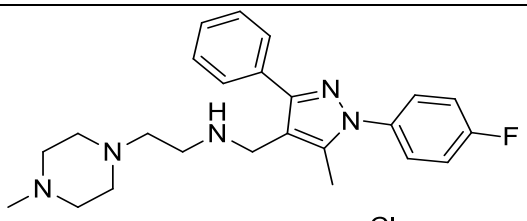
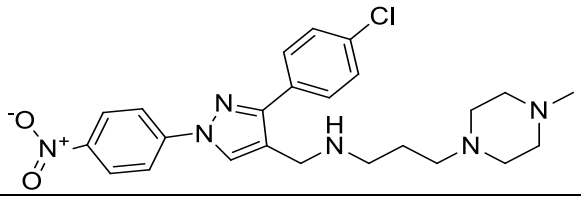
### *Development of QSAR Model*

The QSAR model was developed using the Genetic Function Approximation (GFA) technique on Padel descriptor software. GFA generated multiple models, from which the best-performing model was selected.

### *Model Validation*

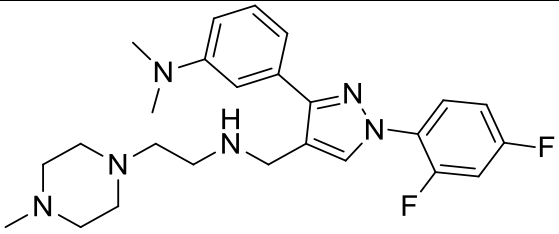
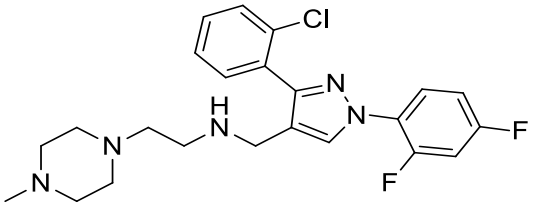
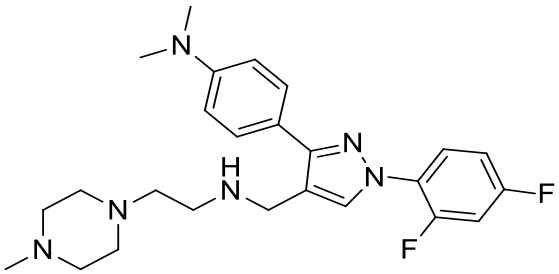
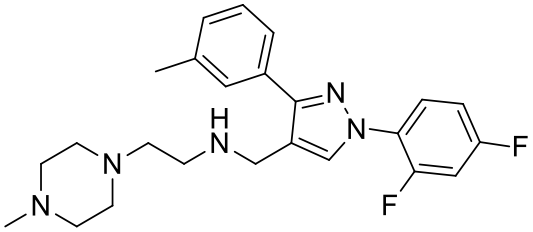
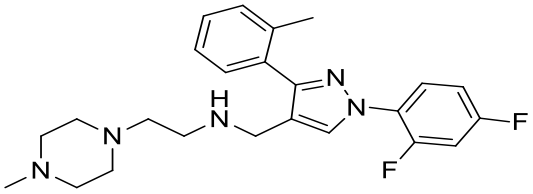
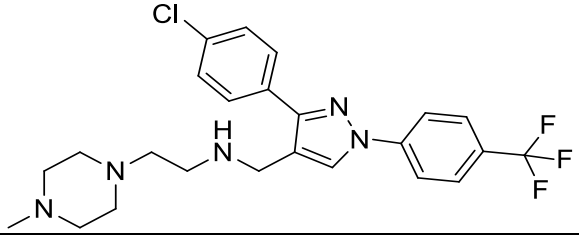
Both internal and external validation were performed. Internal validation was conducted using the training set, while external validation was carried out using the test set.

Table 3.1: The PubChem CID, Structure, and the activities ($pEC_{50}$) of Series 2

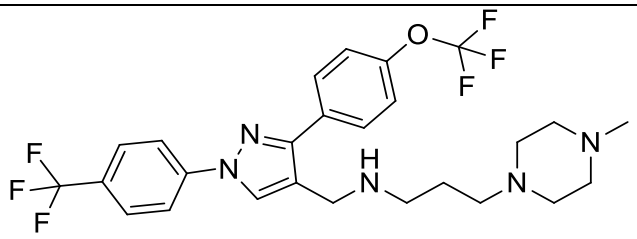
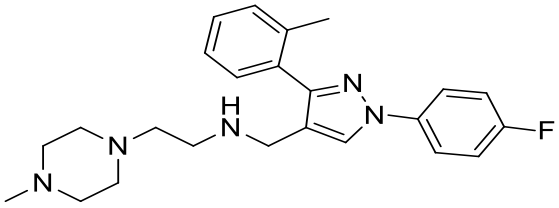
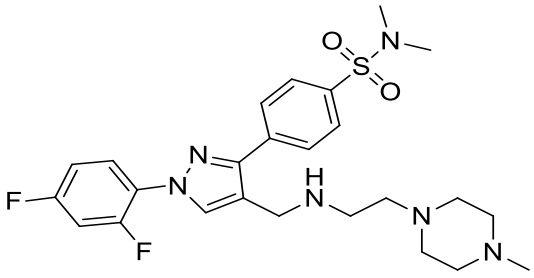
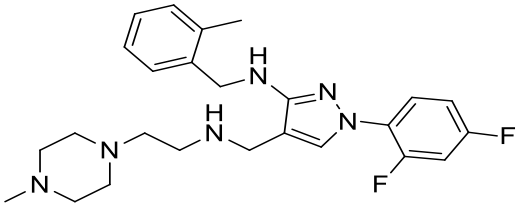
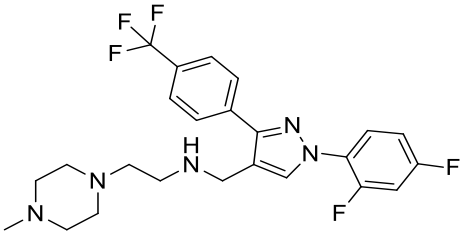
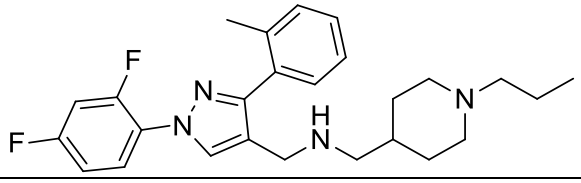
| S/No | Pubchem CID | Structure | $pEC_{50}$ |
|---|---|---|---|
| 1 | 72545083 |  | 5.0706 |
| 2 | 73346556 |  | 6.8861 |

*Corresponding author: Hassan S.*
*hassansamuel84@gmail.com*
*Department of Science Laboratory technology, Nigerian Institute of Leather and Science Technology, Zaria.*

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 13(3), SEPTEMBER, 2025
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbuftejoste.com.ng

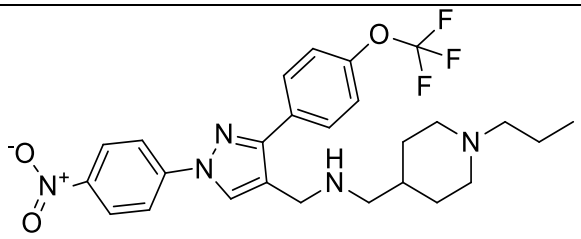
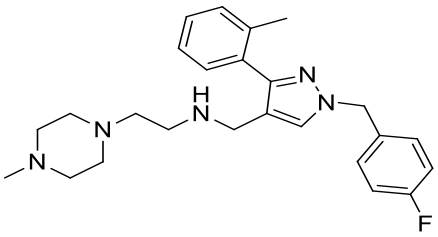
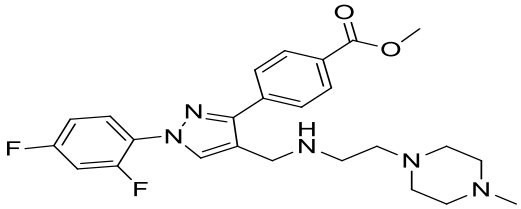
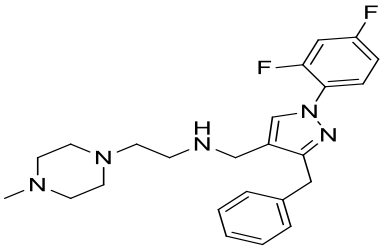
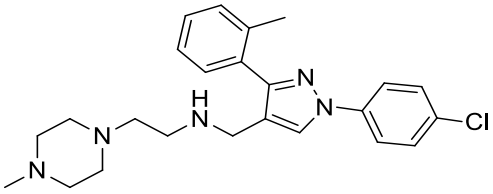
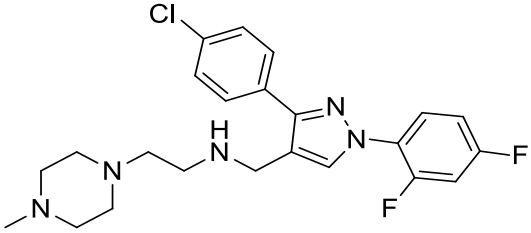
| S/No | Pubchem CID | Structure | $pEC_{50}$ |
|---|---|---|---|
| 3 | 72546770 |  | 5.4949 |
| 4 | 72546771 |  | 5.4815 |
| 5 | 72546531 |  | 5.8861 |
| 6 | 72546533 |  | 5.7696 |
| 7 | 72545819 |  | 5.5528 |
| 8 | 73348042 |  | 6.4437 |

*Corresponding author: Hassan S.*
*hassansamuel84@gmail.com*
*Department of Science Laboratory technology, Nigerian Institute of Leather and Science Technology, Zaria.*

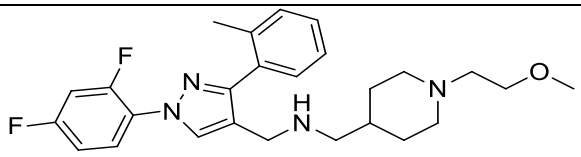
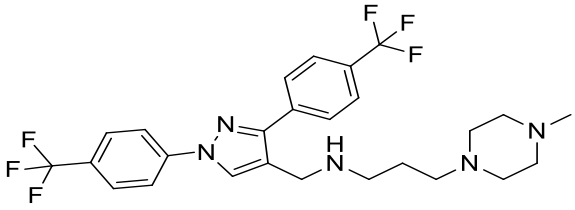
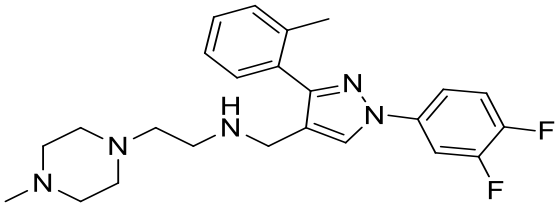
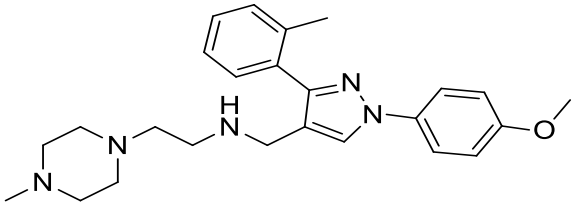
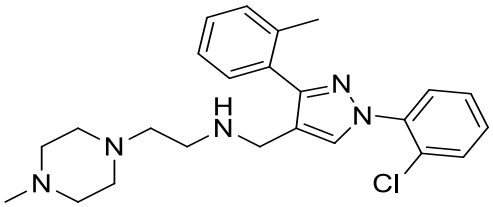
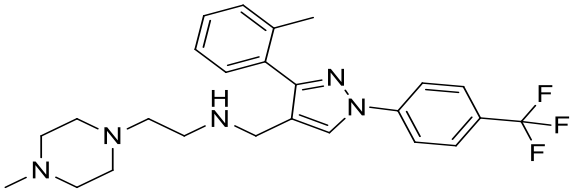
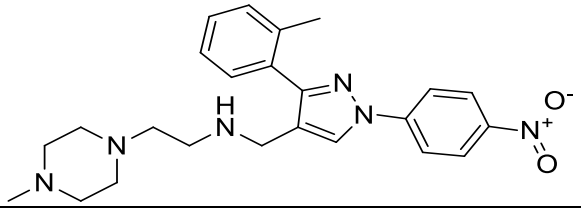
| S/No | Pubchem CID | Structure | $pEC_{50}$ |
|---|---|---|---|
| 9 | 73348043 |  | 6.1249 |
| 10 | 72547539 |  | 5.4318 |
| 11 | 72546534 |  | 5.7447 |
| 12 | 72546773 |  | 5.3665 |
| 13 | 72546287 |  | 6.2676 |
| 14 | 72545573 |  | 5.6383 |

*Corresponding author: Hassan S.*
*hassansamuel84@gmail.com*
*Department of Science Laboratory technology, Nigerian Institute of Leather and Science Technology, Zaria.*

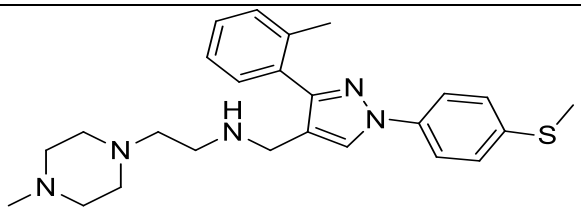
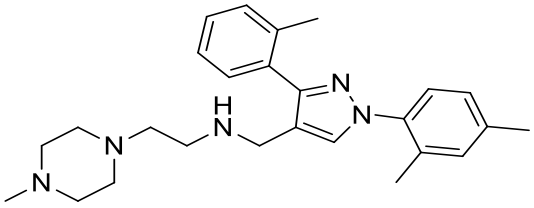
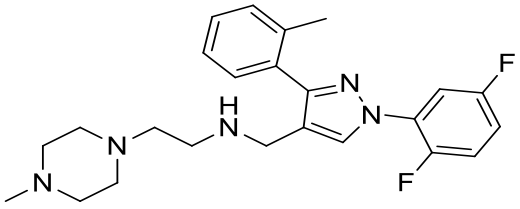
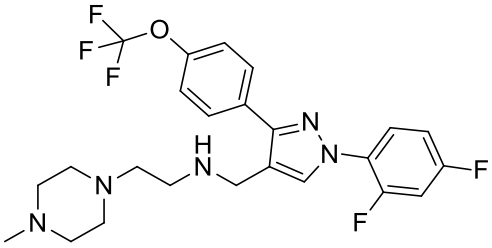
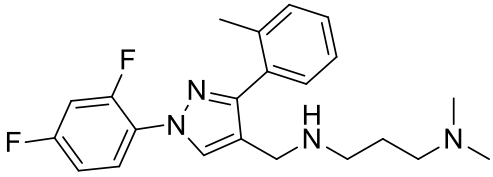
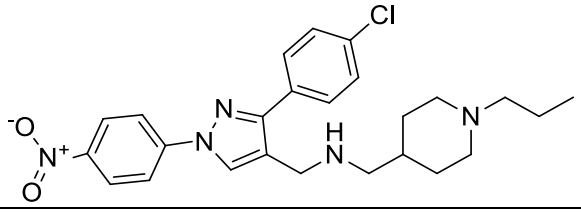
JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 13(3), SEPTEMBER, 2025
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbuftejoste.com.ng

| S/No | Pubchem CID | Structure | $pEC_{50}$ |
|---|---|---|---|
| 15 | 73349579 |  | 6.2218 |
| 16 | 72547543 |  | 5.0000 |
| 17 | 72546772 |  | 5.4202 |
| 18 | 72547034 |  | 4.7959 |
| 19 | 72547037 |  | 6.1871 |
| 20 | 72546285 |  | 6.4202 |

*Corresponding author: Hassan S.*
*hassansamuel84@gmail.com*
*Department of Science Laboratory technology, Nigerian Institute of Leather and Science Technology, Zaria.*

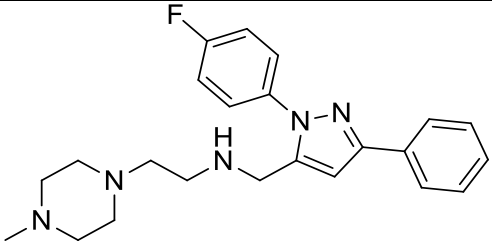
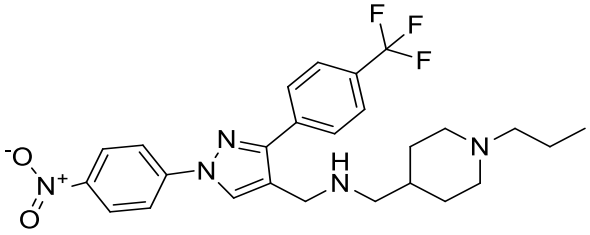
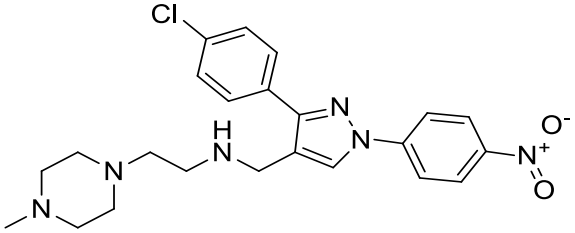
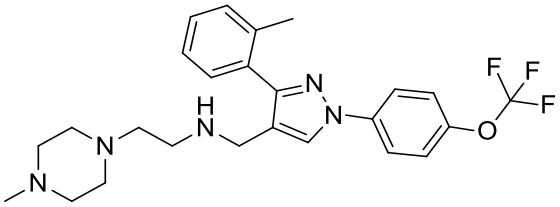
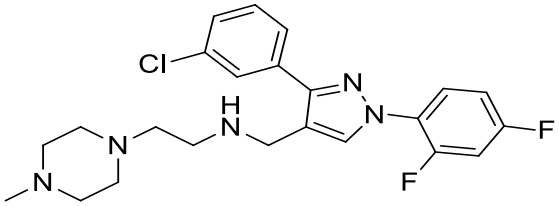
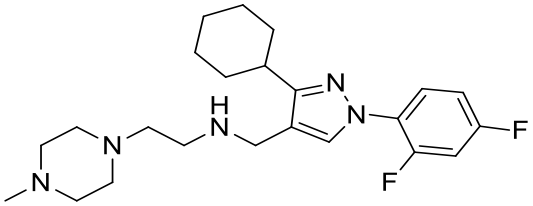
| S/No | Pubchem CID | Structure | $pEC_{50}$ |
|---|---|---|---|
| 21 | 72545820 |  | 5.5229 |
| 22 | 73351139 |  | 6.4202 |
| 23 | 72547297 |  | 5.5850 |
| 24 | 72547538 |  | 24 |
| 25 | 72547540 |  | 25 |
| 26 | 72547035 |  | 26 |
| 27 | 72547036 |  | 27 |

*Corresponding author: Hassan S.*
*hassansamuel84@gmail.com*
*Department of Science Laboratory technology, Nigerian Institute of Leather and Science Technology, Zaria.*

| S/No | Pubchem CID | Structure | $pEC_{50}$ |
|---|---|---|---|
| 28 | 72547038 |  | 28 |
| 29 | 72547295 |  | 29 |
| 30 | 72547296 |  | 30 |
| 31 | 72546286 |  | 6.4437 |
| 32 | 16187217 |  | 5.1739 |
| 33 | 73352610 |  | 6.5229 |

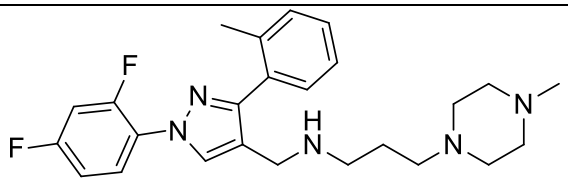
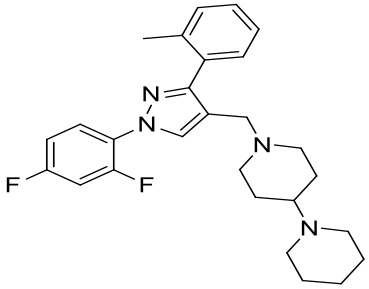
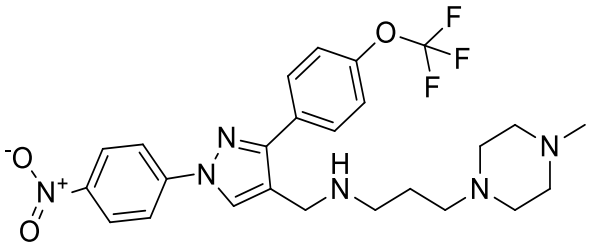
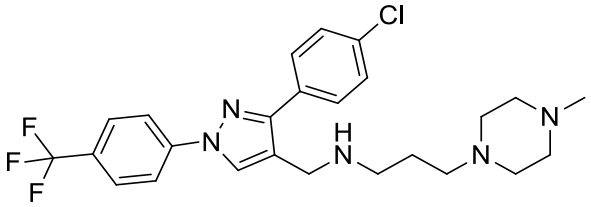
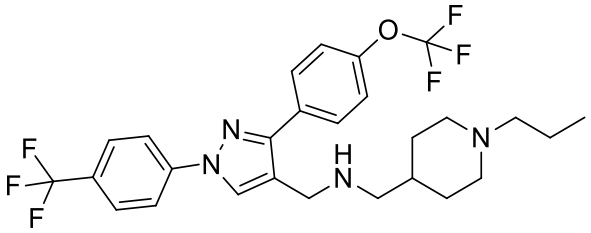
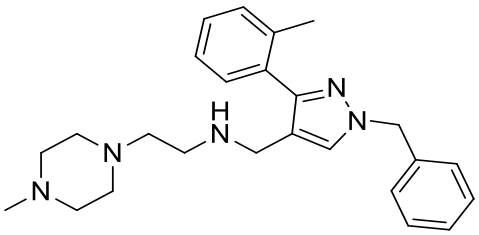
*Corresponding author: Hassan S.*
*hassansamuel84@gmail.com*
*Department of Science Laboratory technology, Nigerian Institute of Leather and Science Technology, Zaria.*

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 13(3), SEPTEMBER, 2025
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbuftejoste.com.ng

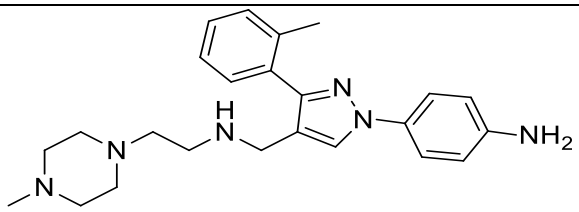
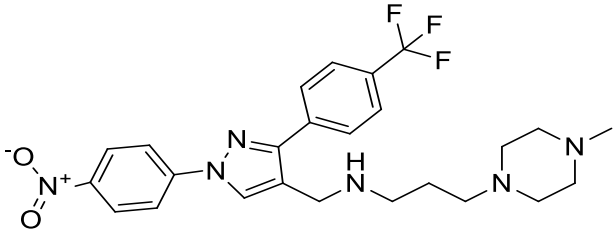
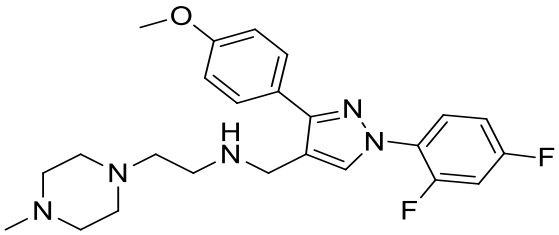
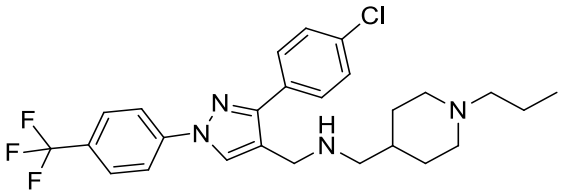
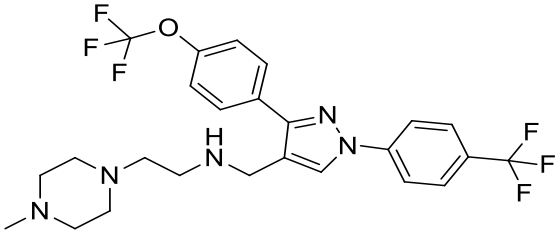
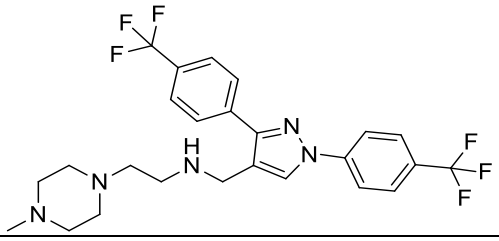
| S/No | Pubchem CID | Structure | $pEC_{50}$ |
|---|---|---|---|
| 34 | 72545084 |  | 4.8861 |
| 35 | 73352611 |  | 6.8239 |
| 36 | 73352612 |  | 6.5376 |
| 37 | 72547039 |  | 6.0458 |
| 38 | 72546288 |  | 6.2366 |
| 39 | 72546530 |  | 5.9208 |

*Corresponding author: Hassan S.*
*hassansamuel84@gmail.com*
*Department of Science Laboratory technology, Nigerian Institute of Leather and Science Technology, Zaria.*

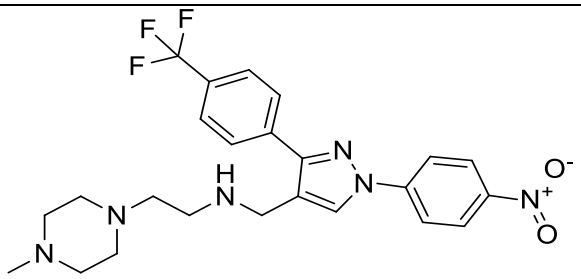
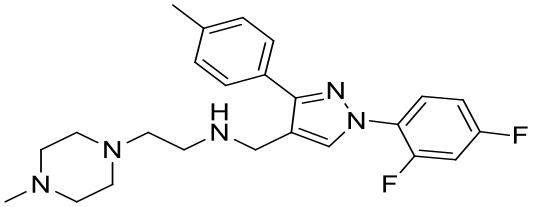
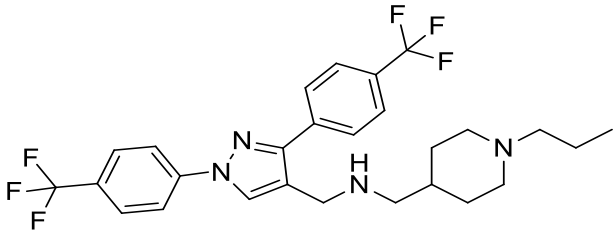
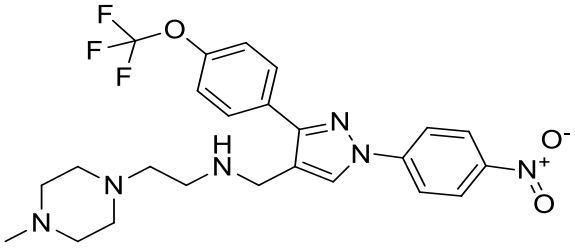
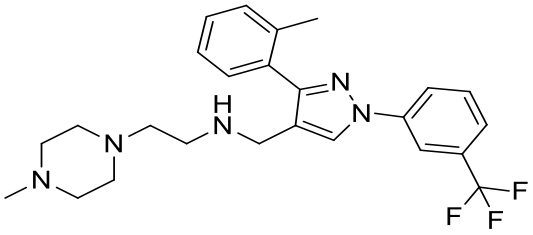
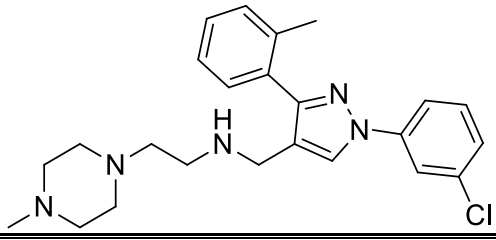
JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 13(3), SEPTEMBER, 2025
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbuftejoste.com.ng

| S/No | Pubchem CID | Structure | $pEC_{50}$ |
|---|---|---|---|
| 40 | 72545572 |  | 5.6576 |
| 41 | 72545822 |  | 5.1249 |
| 42 | 73354107 |  | 6.4685 |
| 43 | 73354108 |  | 6.4437 |
| 44 | 73354109 |  | 6.3872 |
| 45 | 72547541 |  | 5.1249 |

*Corresponding author: Hassan S.*
*hassansamuel84@gmail.com*
*Department of Science Laboratory technology, Nigerian Institute of Leather and Science Technology, Zaria.*

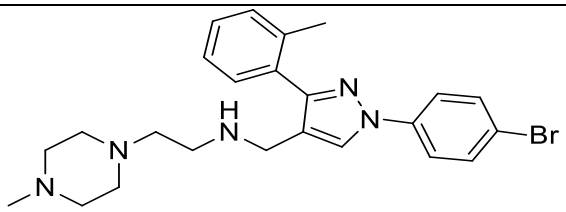
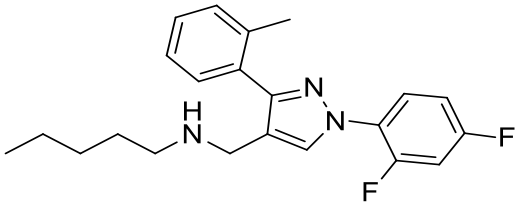
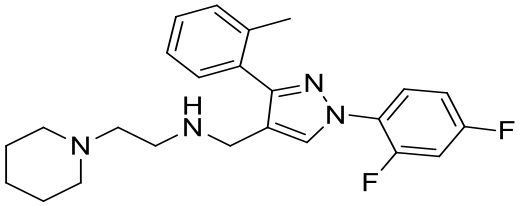
| S/No | Pubchem CID | Structure | $pEC_{50}$ |
|---|---|---|---|
| 46 | 72547542 |  | 5.0000 |
| 47 | 73354110 |  | 6.9208 |
| 48 | 72546769 |  | 5.6778 |
| 49 | 73355627 |  | 6.5376 |
| 50 | 73355628 |  | 6.3565 |
| 51 | 73355629 |  | 6.3098 |

*Corresponding author: Hassan S.*
hassansamuel84@gmail.com
*Department of Science Laboratory technology, Nigerian Institute of Leather and Science Technology, Zaria.*

| S/No | Pubchem CID | Structure | $pEC_{50}$ |
|---|---|---|---|
| 52 | 73355631 |  | 6.9586 |
| 53 | 72546529 |  | 6.0915 |
| 54 | 73357174 |  | 6.4815 |
| 55 | 73357175 |  | 6.3979 |
| 56 | 72547292 |  | 6.0410 |
| 57 | 72547293 |  | 6.0000 |

*Corresponding author: Hassan S.*
*hassansamuel84@gmail.com*
*Department of Science Laboratory technology, Nigerian Institute of Leather and Science Technology, Zaria.*

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 13(3), SEPTEMBER, 2025
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbuftejoste.com.ng

| S/No | Pubchem CID | Structure | $pEC_{50}$ |
|---|---|---|---|
| 58 | 72547294 |  | 5.9208 |
| 59 | 72545823 |  | 4.7447 |
| 60 | 72546284 |  | 4.6990 |

## RESULTS

Table 4.1: Model 1

| MODEL |
|---|
| $\boldsymbol{pEC_{50}} = 7.55735 + 0.29138 * \text{GGI4} + 10.90034 * JGI1 - 0.17042 * XLogP - 0.64795 * TDB5p - 2.9115 * FPSA - 2$ (4.1) |

Table 4.2: The statistical parameters for the model and minimum accepted values for QSAR models

| Symbol | Name | Threshold value | Model Value |
|---|---|---|---|
| $R^2$ | Co-efficient of determination | $\geq 0.6$ | 0.7770 |
| $R^2_{adj}$ | Adjusted Square correlation Co-efficient | $\geq 0.6$ | 0.7302 |
| $Q^2$ | Cross-validation co-efficient | $\geq 0.5$ | 0.7310 |
| $R^2 - Q^2$ | Difference between $R^2$ and $Q^2$ | $\leq 0.3$ | 0.0460 |
| $R^2_{ext}$ | Coefficient of determination for external test set | $\geq 0.5$ | 0.7741 |
| $cR^2_p$ | y-randomization parameter | $\geq 0.5$ | 0.7057 |
| $N_{test\ set}$ | Minimum number of an external test set | $\geq 5$ | 12 |

Table 4.3: The descriptor, descriptions, and class of all descriptors in Model 1

| S/N | Descriptor | Description | Class |
|---|---|---|---|
| 1 | GGI4 | Topological charge index of order 4 | 2D |
| 2 | JGI1 | Mean topological charge index of order 1 | 2D |
| 3 | XLogP | XLogP | 2D |

*Corresponding author: Hassan S.*
hassansamuel84@gmail.com
*Department of Science Laboratory technology, Nigerian Institute of Leather and Science Technology, Zaria.*

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 13(3), SEPTEMBER, 2025
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbuftejoste.com.ng

| S/N | Descriptor | Description | Class |
|---|---|---|---|
| 4 | TDB5p | 3D topological distance-based autocorrelation - lag 5 / weighted by polarizabilities | 3D |
| 5 | FPSA-2 | FPSA-2 / total molecular surface area | 3D |

Table 4.4: Y-Randomization test for model 1

| Iteration | R | $R^2$ | $Q^2$ |
|---|---|---|---|
| Random 1 | 0.402125275 | 0.161704737 | -0.143494795 |
| Random 2 | 0.295413285 | 0.087269009 | -0.199552574 |
| Random 3 | 0.363753985 | 0.132316962 | -0.096005134 |
| Random 4 | 0.171332013 | 0.029354659 | -0.338449885 |
| Random 5 | 0.421955796 | 0.178046693 | -0.09374659 |
| Random 6 | 0.512175741 | 0.26232399 | 0.054663457 |
| Random 7 | 0.529031812 | 0.279874658 | 0.066134148 |
| Random 8 | 0.407913314 | 0.166393272 | -0.096194226 |
| Random 9 | 0.276997432 | 0.076727577 | -0.256093395 |
| Random 10 | 0.306755027 | 0.094098647 | -0.142063448 |
| Random Models Parameters | | | |
| Average r: | 0.415355604 | | |
| Average $r^2$: | 0.204098032 | | |
| Average $Q^2$: | -0.04670533 | | |
| $cR_p^2$ | 0.705714316 | | |

Table 4.5: The experimental $pEC_{50}$, predicted $pEC_{50}$ and the residual of the model 1

| Name | Experimental $pEC_{50}$ | Predicted $pEC_{50}$ | Residual |
|---|---|---|---|
| 1 | 6.8861 | 6.6613 | 0.2247 |
| 2 | 5.4949 | 5.4406 | 0.0543 |
| 3 | 5.4815 | 5.6028 | -0.1213 |
| 4 | 5.8861 | 5.5779 | 0.3082 |
| 5 | 5.5528 | 5.4756 | 0.0773 |
| 6 | 6.4437 | 6.6097 | -0.1660 |
| 7 | 6.1249 | 6.4279 | -0.3030 |
| 8 | 5.4318 | 5.6451 | -0.2133 |
| 9 | 5.7447 | 5.7448 | -1.7E-05 |
| 10 | 6.2676 | 6.1109 | 0.1567 |
| 11 | 5.6383 | 5.4719 | 0.1664 |
| 12 | 6.2219 | 6.4121 | -0.19023 |
| 13 | 5.0000 | 5.8686 | -0.8686 |
| 14 | 5.4202 | 5.3280 | 0.0923 |
| 15 | 4.7959 | 5.5957 | -0.7998 |
| 16 | 6.1871 | 5.7678 | 0.4193 |
| 17 | 6.4202 | 6.1420 | 0.2782 |
| 18 | 5.5229 | 5.3730 | 0.1499 |
| 19 | 6.4202 | 6.5710 | -0.1508 |
| 20 | 5.5850 | 5.6695 | -0.0845 |
| 21 | 5.5528 | 5.3598 | 0.1931 |
| 22 | 5.4202 | 5.2189 | 0.2012 |

*Corresponding author: Hassan S.*
*hassansamuel84@gmail.com*
*Department of Science Laboratory technology, Nigerian Institute of Leather and Science Technology, Zaria.*

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 13(3), SEPTEMBER, 2025
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbuftejoste.com.ng

| Name | Experimental $pEC_{50}$ | Predicted $pEC_{50}$ | Residual |
|---|---|---|---|
| 23 | 6.3372 | 6.1863 | 0.1510 |
| 24 | 6.1739 | 5.7855 | 0.3884 |
| 25 | 5.6198 | 5.6750 | -0.0552 |
| 26 | 6.4437 | 6.1288 | 0.3149 |
| 27 | 5.1739 | 5.3139 | -0.1400 |
| 28 | 6.5229 | 6.6648 | -0.1420 |
| 29 | 4.8861 | 5.5634 | -0.6774 |
| 30 | 6.5376 | 6.6506 | -0.1130 |
| 31 | 6.0458 | 6.1553 | -0.1096 |
| 32 | 6.2366 | 5.9515 | 0.2851 |
| 33 | 5.9208 | 5.5024 | 0.4185 |
| 34 | 5.6576 | 5.5395 | 0.1181 |
| 35 | 5.1250 | 5.1611 | -0.0361 |
| 36 | 6.4685 | 6.4827 | -0.0142 |
| 37 | 6.4437 | 6.5567 | -0.1123 |
| 38 | 6.3872 | 6.3457 | 0.0414 |
| 39 | 5.1249 | 5.2554 | -0.1305 |
| 40 | 5.0000 | 5.4651 | -0.4651 |
| 41 | 6.9208 | 6.6563 | 0.2646 |
| 42 | 6.5376 | 6.5980 | -0.0603 |
| 43 | 6.9586 | 6.5830 | 0.3756 |
| 44 | 6.4815 | 6.5573 | -0.0758 |
| 45 | 6.3979 | 6.5167 | -0.1187 |
| 46 | 6.0000 | 5.5260 | 0.4740 |
| 47 | 6.9208 | 6.8029 | 0.1179 |
| 48 | 4.7447 | 4.7638 | -0.0190 |
| 49 | 4.6990 | 4.8475 | -0.1485 |
| 50 | 5.3665 | 5.3215 | 0.0451 |
| 51 | 5.0706 | 5.5102 | -0.4396 |
| 52 | 5.7696 | 5.5504 | 0.2191 |
| 53 | 5.3665 | 5.2688 | 0.0978 |
| 54 | 6.3098 | 6.1669 | 0.1429 |
| 55 | 5.6021 | 5.4931 | 0.1090 |
| 56 | 6.8239 | 6.6335 | 0.1904 |
| 57 | 5.6778 | 5.4352 | 0.2426 |
| 58 | 6.3566 | 6.4913 | -0.1348 |
| 59 | 6.3098 | 6.6005 | -0.2907 |
| 60 | 6.0915 | 5.7110 | 0.3805 |

*Corresponding author: Hassan S.*
*hassansamuel84@gmail.com*
*Department of Science Laboratory technology, Nigerian Institute of Leather and Science Technology, Zaria.*

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 13(3), SEPTEMBER, 2025
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbuftejoste.com.ng

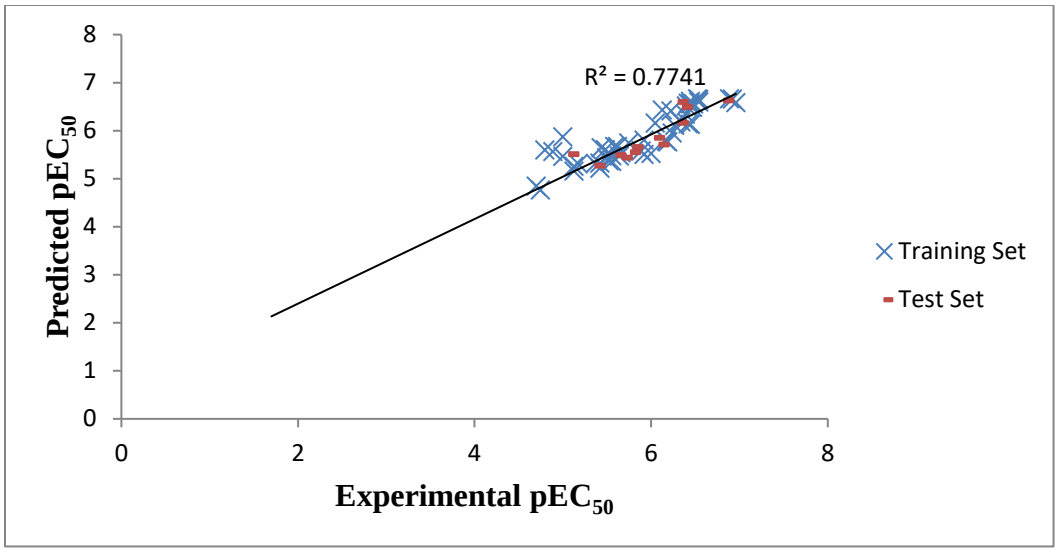



Figure 1: XY scatter plot of the model experimental versus predicted $pEC_{50}$ of both the training and test set of model 1

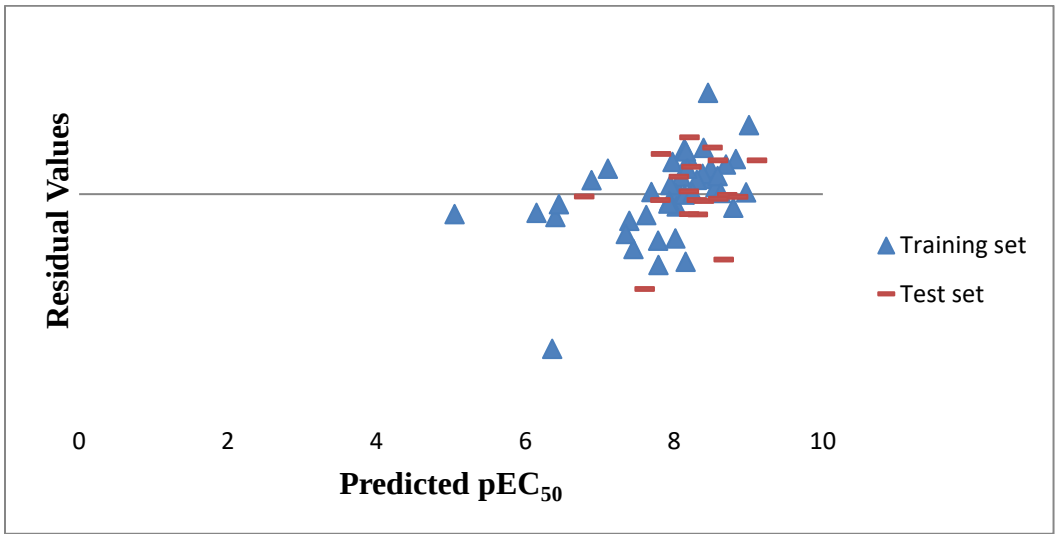



Figure 2: XY scatter plot of predicted $pEC_{50}$ versus residual values of both the training and test sets of model 1

## DISCUSSION OF FINDINGS

The QSAR model developed in this study helps explain why some 1,3,4-trisubstituted pyrazoles work better than others in fighting human African trypanosomiasis. The model, built using Genetic Function Approximation (GFA), has been tested and proven to be reliable [37]. As shown in Table 4.2, the statistical results confirm that the model can make accurate predictions about the biological activity of these compounds. One of the most important values to consider is the squared correlation coefficient ($R^2$), which is 0.7770. This means that the model can explain about 78% of the differences in biological activity among the compounds.

*Corresponding author: Hassan S.*
*hassansamuel84@gmail.com*
*Department of Science Laboratory technology, Nigerian Institute of Leather and Science Technology, Zaria.*

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 13(3), SEPTEMBER, 2025
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbuftejoste.com.ng

In simple terms, most of the variations in how effective the compounds are can be predicted using the molecular descriptors chosen for this model [33]. Another key value is the cross-validation coefficient ($Q_{CV}^2$), which is 0.7310. Since this number is very close to $R^2$, it proves that the model is not just memorizing the training data—it can also make accurate predictions for new compounds [38]. The small gap between $R^2$ and $Q_{CV}^2$ (only 0.046) shows that the model is stable and not overfitted. This is important because an overfitted model may work well on training data but fail when tested on new data [11]. Additionally, the external validation coefficient ($R_{ext}^2$) is 0.7741, confirming that the model remains accurate when applied to entirely new data sets [33]. To understand why certain compounds, work better than others, we need to examine the molecular descriptors included in Model 1 (Table 4.3). These descriptors represent specific chemical properties that influence biological activity [32].

One key descriptor is GGI4 (Global Charge Transfer Index), which has a positive effect on biological activity. This means that when a molecule has a well-distributed charge, it tends to be more effective. Charge distribution plays a big role in determining how well a molecule binds to its target. A balanced charge makes it easier for the molecule to interact with the biological system, improving its overall effectiveness [14].

Another important descriptor is JGI1 (Mean Topological Charge Index of Order 1), which also has a positive impact. This descriptor measures how the charge is spread across the molecule. A molecule with a favorable charge distribution is more likely to fit well with its target, leading to better activity [9]. This suggests that electronic properties are crucial for determining how well a compound works.

However, not all descriptors have a positive impact. XLogP, which represents lipophilicity (how much a molecule likes fat), has a negative effect on biological activity. This means that if a compound is too lipophilic, it becomes less effective. Lipophilicity affects how well a drug dissolves in fats and oils, which can impact how it moves through the body [15]. While some level of lipophilicity is necessary for a drug to enter cells, too much can be harmful. Highly lipophilic compounds may get stuck in fat tissues, making them less available to reach their target. They may also be broken down too quickly or have trouble dissolving in water-based environments like blood. This result suggests that a good drug should have a balance between being water-loving (hydrophilic) and fat-loving (lipophilic) [39].

Another descriptor with a negative effect is TDB5p, which is related to the 3D structure of the molecule and how easily its electrons shift. This property, known as polarizability, affects how the molecule interacts with proteins and enzymes. If a molecule has too much charge shifting in certain areas, it may not bind properly to its target, reducing its biological activity [32].

The last important descriptor is FPSA-2 (Fractional Polar Surface Area of oxygen and nitrogen atoms), which also negatively affects activity. This means that molecules with a higher polar surface area tend to be less effective. Polar surface area measures how much of a molecule's surface is covered by atoms that like water, such as oxygen and nitrogen. While some polarity helps drugs dissolve in the bloodstream, too much can make it difficult for them to pass through cell membranes. If a drug cannot easily enter cells, it may struggle to reach its target effectively [40]. This suggests that reducing excessive polar surface area may help improve a compound's ability to work as a drug.

To check if the model's predictions are accurate, an XY scatter plot of experimental versus predicted $pEC_{50}$ values was created (Figure 1). The points in the plot are closely aligned with the regression line, showing that the model's predictions match well with real experimental values. If the points were scattered far from the line, it would mean the model's predictions were unreliable. Since most of the points are close to the line, we can be confident that the model successfully captures the relationship between molecular properties and biological activity [11].

To further confirm the reliability of the model, a residual plot was created (Figure 2). A residual plot shows the difference between

*Corresponding author: Hassan S.*
*hassansamuel84@gmail.com*
*Department of Science Laboratory technology, Nigerian Institute of Leather and Science Technology, Zaria.*

predicted and actual values. If the model were biased, we would expect to see a clear pattern in the residuals. However, in Figure 2, the residuals appear to be randomly scattered. This means the model does not systematically overestimate or underestimate certain types of molecules, proving that it is statistically sound and free from major errors [33].

To ensure the model's results were not just due to random chance, a Y-randomization test was performed (Table 4.4). This test involves randomly shuffling the biological activity values and rebuilding the model. If the original model is meaningful, the randomized versions should perform much worse. As expected, the shuffled models had much lower $R^2$ and $Q^2$ values compared to the real model. This confirms that the descriptors in the original model are truly responsible for predicting activity and that the model's accuracy is not just a coincidence (Tropsha *et al*., 2003). Additionally, the calculated $cR_p^2$ value of 0.7057 is well above the acceptable limit of 0.50, further proving the model is statistically valid (Golbraikh & Tropsha, 2002).

Lastly, Table 4.5 compares the experimental $pEC_{50}$ values, predicted values, and residuals for each compound. The small differences between the predicted and actual values show that the model is highly accurate. More importantly, the model performs well on both the training and test sets, meaning it is not just memorizing the data it was trained on but can also predict the activity of new compounds [4].

## CONCLUSION

This study has demonstrated the potential of substituted 2-phenylimidazopyridines as promising candidates for anti-HAT therapy. Through a QSAR approach, key molecular descriptors influencing their activity have been identified, providing a foundation for further optimization. Given the persistent challenges in HAT treatment—ranging from drug toxicity to blood-brain barrier penetration—rational drug design guided by QSAR modeling offers a viable pathway for developing safer, more effective therapeutics. Continued research in this field is crucial for advancing treatment options and supporting global efforts toward the eradication of HAT.

## REFERENCES

1. Baker, N., de Koning, H. P., Mäser, P., & Horn, D. (2013). Drug resistance in African trypanosomiasis: The melarsoprol and pentamidine story. Trends in Parasitology, 29(3), 110-118.
2. Barrett, M. P., & Croft, S. L. (2012). Management of trypanosomiasis and leishmaniasis. British Medical Bulletin, 104(1), 175-196.
3. Barrett, M. P., Vincent, I. M., Burchmore, R. J., Kazibwe, A. J., & Matovu, E. (2011). Drug resistance in human African trypanosomiasis. Future Microbiology, 6(9), 1037-1047.
4. Walters, W. P., & Barzilay, R. (2020). Applications of deep learning in molecule generation and molecular property prediction. *Accounts of chemical research*, *54*(2), 263-270.
5. Büscher, P., Cecchi, G., Jamonneau, V., & Priotto, G. (2017). Human African trypanosomiasis. The Lancet, 390(10110), 2397-2409.
6. Büscher, P., Mertens, P., Leclipteux, T., Pyana, P. P., & Boelaert, M. (2014). Sensitivity and specificity of HAT Sero-K-SeT, a rapid diagnostic test for sleeping sickness. PLOS Neglected Tropical Diseases, 8(10), e3183.
7. Chappuis, F., Loutan, L., Simarro, P., Lejon, V., & Büscher, P. (2005). Options for field diagnosis of human African trypanosomiasis. Clinical Microbiology Reviews, 18(1), 133-146.
8. Chtita, S., El-Abd, A. E., & El-Sayed, N. N. (2019). Predictive modeling and structure-activity relationship of anti-HAT activity using QSAR techniques. Journal of Molecular Modeling, 25(6), 134-145.
9. Consonni, V., & Todeschini, R. (2009). Molecular Descriptors. In Handbook of Chemoinformatics Algorithms (pp. 29–122). CRC Press.

*Corresponding author: Hassan S.*
*hassansamuel84@gmail.com*
*Department of Science Laboratory technology, Nigerian Institute of Leather and Science Technology, Zaria.*

10. Elrufaie, H. A., Mohamed, L. M., Hamd, A. Y., Bala, N. A., Elbadawi, F. A., Ghaboosh, H., & Alzain, A. A. (2023). Discovery of novel natural products as rhodesain inhibitors for human African trypanosomiasis using in silico techniques. *Journal of Biomolecular Structure and Dynamics, 41*(12), 5672-5684.
11. Eriksson, L., Jaworska, J., Worth, A. P., Cronin, M. T. D., McDowell, R. M., & Gramatica, P. (2003). Methods for reliability and uncertainty assessment and for applicability evaluations of classification- and regression-based QSARs. *Environmental Health Perspectives, 111*(10), 1361–1375.
12. Fairlamb, A. H. (2003). Chemotherapy of human African trypanosomiasis: Current and future prospects. *Trends in Parasitology, 19*(11), 488-494.
13. Franco, J. R., Simarro, P. P., Diarra, A., & Jannin, J. G. (2014). Epidemiology of human African trypanosomiasis. Clinical Epidemiology, 6, 257-275.
14. Galvez, J., Garcia-Domenech, R., de Julian-Ortiz, J. V., & Pogliani, L. (1994). Topological charge indices for predicting the biological activity of drugs. *Journal of Chemical Information and Computer Sciences, 34*(3), 520–525.
15. Ghose, A. K., & Crippen, G. M. (1986). Atomic physicochemical parameters for three-dimensional QSAR. *Journal of Computational Chemistry, 7*(4), 565–577.
16. Henao-Martínez, A. F. (2023). Human African trypanosomiasis (Sleeping Sickness)—Epidemiology and control. Current Tropical Medicine Reports, 10(2), 81-89.
17. Idrissi Taourati, A. (2015). QSAR study of anti-Human African Trypanosomiasis activity for 2-phenylimidazopyridines derivatives using DFT and Lipinski's descriptors. *Journal of Molecular Modeling, 21*(6), 134.
18. Kennedy, P. G. (2013). Clinical features, diagnosis, and treatment of human African trypanosomiasis (sleeping sickness). The Lancet Neurology, 12(2), 186-194.
19. Kennedy, P. G. E. (2004). Human African trypanosomiasis of the CNS: Current issues and challenges. *The Journal of Clinical Investigation, 113*(4), 496-504.
20. Requirements of imidazopyridine analogues as anti-infective agents against trypanosomiases. *Journal of Computational Biophysics and Chemistry, 21*(01), 83-114.
21. Masand, V. H., El-Sayed, N. N., Bambole, M. U., Patil, V. R., & Thakur, S. D. (2019). Multiple quantitative structure-activity relationships (QSARs) analysis for orally active trypanocidal N-myristoyltransferase inhibitors. Journal of Molecular Structure, 1175, 481-487.
22. Masand, V. H., El-Sayed, N. N., Mahajan, D. T., Mercader, A. G., Alafeefy, A. M., & Shibi, I. G. (2017). QSAR modeling for anti-human African trypanosomiasis activity of substituted 2-Phenylimidazopyridines. *Journal of Molecular Structure,* 1130, 711-718.
23. Nnadi, C. O., Althaus, J. B., Nwodo, N. J., & Schmidt, T. J. (2018). A 3D-QSAR study on the antitrypanosomal and cytotoxic activities of steroid alkaloids by comparative molecular field analysis. Molecules, 23(5), 1113.
24. Nnadi, E. O., Olaifa, A. K., & Fadaka, A. O. (2018). Advances in the search for new anti-trypanosomal agents: A comprehensive review. *Drug Development Research, 79*(4), 450-465.
25. Papagni, R., Novara, R., Minardi, M. L., et al. (2023). Human African trypanosomiasis (sleeping sickness): Current knowledge and future

*Corresponding author: Hassan S.*
✉ *hassansamuel84@gmail.com*
*Department of Science Laboratory technology, Nigerian Institute of Leather and Science Technology, Zaria.*

challenges. Frontiers in Tropical Diseases, 4, 1087003.
26. Pépin, J., & Méda, H. A. (2001). The epidemiology and control of human African trypanosomiasis. Advances in Parasitology, 49, 71-132.
27. , P. D., Østergaard, S., Kristensen, T. N., Sørensen, P., Loeschcke, V., Mackay, T. F., & Sarup, P. (2018). Functional validation of candidate genes detected by genomic feature models. *G3: Genes, Genomes, Genetics*, *8*(5), 1659-1668.
28. Simarro, P. P., Jannin, J., & Cattand, P. (2010). Eliminating human African trypanosomiasis: Where do we stand and what comes next? PLoS Medicine, 7(8), e1000325.
29. Steverding, D. (2008). The history of African trypanosomiasis. Parasites & Vectors, 1(1), 3.
30. The Sun. (2024). Deadly 'sleeping sickness' that causes fever to psychosis 'risking worldwide spread'. The Sun.
31. Todeschini, R., & Consonni, V. (2000). Handbook of Molecular Descriptors. Wiley-VCH.
32. Tin, N., Chauhan, M., Agwamba, K., Sun, Y., Parsons, A., Payne, P., & Osan, R. (2024). Evaluating Molecular Complexity with Open-Source Machine Learning Approaches to Predict Process Mass Intensity. *ACS omega*, *9*(26), 28476-28484.
33. World Health Organization. (2023). Guidelines for the treatment of human African trypanosomiasis.
34. Hossain, M. S., & Hussain, M. H. (2025). Multi-Target Drug Design in Alzheimer's Disease Treatment: Emerging Technologies, Advantages, Challenges, and Limitations. *Pharmacology Research & Perspectives*, *13*(4), e70131.
35. Kiriiri, G. K., Njogu, P. M., & Mwangi, A. N. (2020). Exploring different approaches to improve the success of drug discovery and development projects: a review. *Future Journal of Pharmaceutical Sciences*, *6*(1), 27
36. Buckner, F. S., Buchynskyy, A., Nagendar, P., Patrick, D. A., Gillespie, J. R., Herbst, Z. & Gelb, M. H. (2020). Phenotypic drug discovery for human african trypanosomiasis: a powerful approach. *Tropical Medicine and Infectious Disease*, *5*(1), 23
37. Zhang, Y., Niu, J., & Na, S. (2019). A novel nonlinear function fitting model based on FOA and GRNN. *Mathematical Problems in Engineering*, *2019*(1), 2697317.
38. Jabeen, F. (2024). *Application of Machine Learning and Deep Learning Approaches Cheminformatic for Drug Discovery* (Doctoral dissertation, Carleton University).
39. Maghsoudi, S., Hosseini, S. A., & Ravandi, S. (2022). Liposome Carriers: Synthetic Methods and Their Applications in Drug Delivery. *Journal of Chemical Reviews*, *4*(4).
40. Manzari, M. T., Shamay, Y., Kiguchi, H., Rosen, N., Scaltriti, M., & Heller, D. A. (2021). Targeted drug delivery strategies for precision medicines. *Nature Reviews Materials*, *6*(4), 351-370.

*Corresponding author: Hassan S.*
*hassansamuel84@gmail.com*
*Department of Science Laboratory technology, Nigerian Institute of Leather and Science Technology, Zaria.*