Data Article

Dataset on theoretical bio-evaluation of 1,2,4-thiadiazole-1,2,4-triazole analogues against epidermal growth factor receptor kinase down regulating human lung cancer

Abel Kolawole Oyebamiji, Sunday Adewale Akintelu, Oreoluwa P. Amao, Mary Oluwatosin Kaka, Adetoun E. Morakinyo, Folake Ayobami Amao, Banjo Semire

ABSTRACT

Data from eight 1,2,4-thiadiazole-1,2,4-triazole derivatives were used to observe the anti-epidermal growth factor receptor kinase activities of 1,2,4-thiadiazole-1,2,4-triazole analogues thereby reducing human lung cancer. The software used to achieve this work were Spartan 14, Pymol, mgltools_win32_1.5.6, Auto dock vina and biovia2019.ds2019client. Also, the developed QSAR model was developed using the screened descriptors so as to inspect the closeness between the experimental IC$_{50}$ and the predicted IC$_{50}$. More so, the binding affinity from 1,2,4-thiadiazole-1,2,4-triazole derivatives - epidermal growth factor receptor kinase complexes using molecular docking approach were reported. Also, the ADMET properties for

Keywords
1,2,4-Thiadiazole
EGFRK
1,2,4-Triazole
Lung cancer
QSAR
DFT
Docking

ARTICLE INFO

Article history:
Received 31 March 2021
Revised 21 May 2021
Accepted 14 June 2021
Available online 17 June 2021

* Corresponding author.
E-mail addresses: abeloyebamiji@gmail.com, oyebamiji.abel@adelekeuniversity.edu.ng (A.K. Oyebamiji).

https://doi.org/10.1016/j.dib.2021.107234
2352-3409/© 2021 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
selected compounds and proposed compounds with better binding affinity were reported.

© 2021 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

### Specification Table

| Subject          | Computational Chemistry       |
|------------------|--------------------------------|
| Specific subject area | Drug Design                  |
| Type of data     | Table                         |
|                  | Figure                        |
|                  | Statistical model             |
| How data were acquired | Spartan’14, Pymol 1.7.4.4, mgltools_win32_1.5.6; AutoVina, EduPyMOL-v1.7.4.4-Win32; biovia2019.ds2019client |
| Data format      | Raw figures                   |
| Parameters for data collection | B3LYP with 6-31G* as basis set was used for optimization due to its efficiency and QSAR software (Gretl and Material studio software) were carefully selected and employed to develop QSAR models. Also, the docking software (EduPyMOL-v1.7.4.4-Win32, biovia2019.ds2019client, mgltools_win32_1.5.6 and Autodock vina) were employed and validated for the purpose of reliability. |
| Description of data collection | The study began with optimizing the chosen ligands using quantum chemical method. The obtained descriptors from the optimized drug-like compounds were extracted and used to develop QSAR model using multiple linear regression method. Also, the developed QSAR model was used to predict the biological activities of new set of triazole based drug-like compounds and further subjected to docking. The results obtained were collected and interpreted. |
| Data source location | Computational Chemistry Research Laboratory, Department of Pure and Applied Chemistry, Ladoke Akintola University of Technology, P.M.B. 4000, Ogbomoso, Oyo State, Nigeria |
| Data accessibility | The experimental and predicted data can be gained access to in the data article |

### Value of the Data

- The data obtained from the optimized 3D structure of the investigated molecular compounds will assist researchers to identify the parameters that define anti-epidermal growth factor receptor kinase activities of 1,2,4-thiadiazole-1,2,4-triazole derivatives.
- The calculated data used in developing quantitative structure-activities relationship (QSAR) model will help scientists in proposing library of 1,2,4-thiadiazole-1,2,4-triazole based compounds with efficient biological activities.
- The calculated scoring will reveal the strength of individual investigated molecule to inhibit epidermal growth factor receptor kinase with PDB ID: 1m17.
- The proposed compounds will help scientists to know compound(s) with better inhibitory concentration than those mentioned in the literatures.

### 1. Data Description

As shown in Table 1, 2D structure of the compounds under investigation were obtained from the work done by Yazala et al., (2020) [1] for further analysis. The compounds were converted to 3D structure before optimization using density functional theory via 6–31G* as basis set.
Table 1
2D and 3D structures of 1,2,4-Thiadiazole-1,2,4-Triazole derivatives, the IUPAC name and Observed IC$_{50}$ value.

| R    | 3D Structure | IUPAC Name                                                                 | Obs IC$_{50}$ |
|------|--------------|-----------------------------------------------------------------------------|---------------|
| a    | H            | (5-(3,4,5-Trimethoxyphenyl)-3-(4-(3-(3,4,5-Trimethoxyphenyl)-1,2,4-Thiadiazol-5-yl)Phenyl)-1H-1,2,4-Triazol-1-yl)(Phenyl)Methanone | 2.98          |
| b    | 3,4,5-trimethoxy | (3,4,5-Trimethoxyphenyl)(5-(3,4,5-Trimethoxyphenyl)-3-(4-(3-(3,4,5-Trimethoxyphenyl)-1,2,4-Thiadiazol-5-yl)Phenyl)-1H-1,2,4-Triazol-1-yl)Methanone | 0.17          |
| c*   | 3,5-dimethoxy | (3,5-Dimethoxyphenyl)(5-(3,4,5-Trimethoxyphenyl)-3-(4-(3-(3,4,5-Trimethoxyphenyl)-1,2,4-Thiadiazol-5-yl)Phenyl)-1H-1,2,4-Triazol-1-yl)Methanone | 1.79          |

(continued on next page)
| Table 1 (continued) |     |                                                                                       |         |
|---------------------|-----|---------------------------------------------------------------------------------------|---------|
| d                   | 4-methoxy | (4-Methoxyphenyl)(5-(3,4,5-Trimethoxyphenyl)-3-(4-(3,4,5-Trimethoxyphenyl)-1,2,4-Thiadiazol-5-yl)Phenyl)-1H-1,2,4-Triazol-1-yl)Methanone | 2.10    |
| e                   | 4-nitro  | (5-(3,4,5-Trimethoxyphenyl)-3-(4-(3,4,5-Trimethoxyphenyl)-1,2,4-Thiadiazol-5-yl)Phenyl)-1H-1,2,4-Triazol-1-yl)(4-Nitrophenyl)Methanone | 1.64    |
| f                   | 4-chloro | (4-Chlorophenyl)(5-(3,4,5-Trimethoxyphenyl)-3-(4-(3,4,5-Trimethoxyphenyl)-1,2,4-Thiadiazol-5-yl)Phenyl)-1H-1,2,4-Triazol-1-yl)Methanone | 1.69    |
| g*                  | 4-cyano  | 4-[[5-(3,4,5-Trimethoxyphenyl)-3-(3,4,5-Trimethoxyphenyl)-1,2,4-Thiadiazol-5-yl]Phenyl]-1H-1,2,4-Triazol-1-yl]CarbonylBenzonitrile | 1.90    |
| h*                  | 4-methyl | (5-(3,4,5-Trimethoxyphenyl)-3-(4-(3,4,5-Trimethoxyphenyl)-1,2,4-Thiadiazol-5-yl)Phenyl)-1H-1,2,4-Triazol-1-yl)(p-Tolyl)Methanone | 11.50   |

* Denotes test set.
Table 2
Calculated 2D Descriptors from optimized 1,2,4-Thiadiazole-1,2,4-Triazole derivatives.

|       | ALogP | ALogP2 | AT8Sm | apol | nAtom | nHeavyAtom | nH | nC | nN | nO | Predicted IC50 |
|-------|-------|--------|-------|------|-------|------------|----|----|----|----|----------------|
| a     | -0.355| 0.126025| 14180 | 96.28458 | 79 | 48 | 31 | 45 | 6 | 7 | 2.92 |
| b     | -0.5104| 0.260508| 17313.71 | 107.9713 | 91 | 54 | 37 | 54 | 10 | 0.11 |
| c     | 0.0422| 0.001781 | 17360.75 | 110.2629 | 93 | 54 | 39 | 59 | 9 | 0.07 |
| d     | -0.4068| 0.165486 | 15153.33 | 100.1802 | 83 | 50 | 33 | 56 | 5 | 2.05 |
| e     | 0.3051| 0.093086 | 15506.86 | 98.32179 | 81 | 51 | 30 | 35 | 6 | 1.73 |
| f     | 0.3893| 0.151554 | 15489.79 | 97.79779 | 79 | 49 | 30 | 35 | 7 | 1.75 |
| g     | -0.3203| 0.102592 | 14962.87 | 98.47779 | 80 | 50 | 30 | 36 | 6 | 2.22 |
| h     | 0.287| 0.082369 | 14677.11 | 99.37817 | 82 | 49 | 33 | 36 | 5 | 2.47 |

Note: ALogP = Lipophilicity; AT8Sm = Broto-Moreau autocorrelation of a topological structure-lag 8 / weighted atomic masses; apol = atomic polarizabilities; nAtom = number of atom; nHeavyAtom = number of heavy atom; nH = number of hydrogen; nC = number of carbon; nN = number of Nitrogen; nO = number of Oxygen.

Table 3
Calculated 3D Descriptors from optimized 1,2,4-Thiadiazole-1,2,4-Triazole derivatives.

|       | EHOMO | ELUMO | BG | DM | MW | AREA | VOL | OVALITY | POL | HBD | HBA | Predicted IC50 |
|-------|-------|-------|----|----|----|------|-----|----------|-----|-----|-----|----------------|
| a     | -5.88 | -2.43 | 3.45 | 7.66 | 623.646 | 608.08 | 584.88 | 1.8 | 88.01 | 3 | 12 | 3.30 |
| b     | -5.89 | -2.44 | 3.45 | 8.56 | 713.724 | 695.76 | 666.13 | 1.89 | 94.56 | 3 | 15 | 1.77 |
| c     | -5.89 | -2.4 | 3.5 | 7.6 | 711.752 | 705.99 | 675.28 | 1.9 | 94.64 | 3 | 14 | 0.83 |
| d     | -5.88 | -2.35 | 3.53 | 8.14 | 653.672 | 636.07 | 611.25 | 1.83 | 90.13 | 3 | 13 | 2.72 |
| e     | -5.91 | -3.04 | 2.87 | 9.48 | 668.643 | 632.71 | 606.19 | 1.83 | 89.87 | 3 | 15 | 1.29 |
| f     | -5.91 | -2.4 | 3.51 | 5.76 | 658.091 | 622.4 | 597.76 | 1.81 | 89.04 | 3 | 12 | 1.50 |
| g     | -5.91 | -2.46 | 3.45 | 9.75 | 648.656 | 628.04 | 603.83 | 1.82 | 89.54 | 3 | 13 | 1.67 |
| h     | -5.66 | -3.96 | 1.7 | 8.87 | 637.673 | 609.29 | 599.56 | 1.77 | 89.61 | 3 | 12 | 1.38 |

Note: EHOMO = Highest occupied molecular orbital energy; ELUMO = Lowest unoccupied molecular orbital energy; BG = Band Gap; DM = Dipole Moment; MW = Molecular weight; VOL = Volume; POL = Polarizability; HDB = Hydrogen bond donor; HBA = Hydrogen Bond Acceptor.

Table 4
Calculated QSAR model Using 3D descriptors.

| Equation | F | P-value | R² | Adjusted R² | MSE | CVR² |
|----------|---|---------|----|-------------|-----|------|
| IC50 = 238.668 + 37.9589(EHOMO) - 0.0195043(MW) | 0.004 | ≤0.001 | 0.971 | 0.951 | 0.4138 | 0.972 |

Table 2 shows series of descriptors from 2D structures were obtained and screened for anti-epidermal growth factor receptor kinase activities. The selected 2D descriptors for QSAR model were AT8Sm, ALogP, Apol, nATOM, nH, nC, nN, nO, nHeavyAtom. Also, the descriptors (dipole moment, highest occupied molecular orbital, band gap, lowest unoccupied molecular orbital, molecular weight, area, volume, polar surface area, ovality, log P, polarisability, hydrogen bond donor and hydrogen bond acceptor) from the 3D structure of the 1,2,4-thiadiazole-1,2,4-triazole analogues which describe anti-epidermal growth factor receptor kinase activity of 1,2-thiadiazole-1,2,4-triazole analogues is shown in Table 3.

Table 4 shows the quantitative structure-activity relationship model developed from the descriptors obtained from 3D-structures of the investigated 1,2,4-thiadiazole-1,2,4-triazole derivatives via Dataset Division GUI 1.2 software [2,3] and Gretl software [4]. The statistical factors calculated for the developed QSAR model were squared correlation coefficient (R²) (0.971), adjusted squared coefficient (Adj.R²) (0.951), P-Value (≤0.001), F-Value (0.004), cross validation correlation coefficient (CVR²) (0.972) and mean square error (0.4138). The screened and the chosen calculated descriptors which described anti-epidermal growth factor receptor kinase activity of 1,2,4-thiadiazole-1,2,4-triazole analogues used in the developed model were highest occupied
Fig. 1. Graph showing correlation between the predicted and observed IC$_{50}$.

Table 5
Calculated QSAR model Using 2D descriptors.

| Equation                      | F     | P-value | R$^2$ | Adjusted R$^2$ | CVR$^2$ |
|-------------------------------|-------|---------|-------|----------------|---------|
| IC$_{50}$ = -0.000894663(ATS8m) + 15.608990873 | 587.69 | $\leq 0.001$ | 0.994 | 0.993 | 0.971 |

molecular orbital and molecular weight. The correlation between the experimental IC$_{50}$ and the predicted IC$_{50}$ was displayed in Fig. 1.

Also, Table 5 showed the developed QSAR model using 2D descriptors using material studio, DTC-QSAR tool v1.0.4 and DataPreTreatmentGUI 1.2 software. The statistical factors considered were squared correlation coefficient (R$^2$) (0.994), adjusted squared coefficient (Adj.R$^2$) (0.993), P-Value ($\leq 0.001$) and F-Value (587.69). The selected calculated parameter obtained from 2D structure of 1,2,4-thiadiazole-1,2,4-triazole derivatives that defined anti-epidermal growth factor receptor kinase used in the developed model was ATS8m.

Fig. 2 showed five (5) proposed molecular compounds (AB1-AB5) which were designed by adding new derivatives to the parent compound used in this work and their biological activities were calculated using the developed QSAR model with 3D descriptors. The predicted inhibition concentration (IC$_{50}$) for the proposed compounds was 3.483, 0.682, 1.782, 0.806 and 2.296 for compound AB1-AB5.

The 3D model of the validated docking calculation overlying the native drug is shown in Fig. 3. The observed similarity and the root mean square deviation (RMSD) between the re-docked native molecule and the native ligand were near to 1. The calculated binding affinity between 1,2,4-thiadiazole-1,2,4-triazole derivatives and human lung cancer cell line (PDB ID: 1m17) [5] is shown in Table 6. The residues and the type of interaction involved in the interaction between 1,2,4-thiadiazole-1,2,4-triazole derivatives and human lung cancer cell line (PDB ID: 1m17) is also displayed in Table 6 and Fig. 4.
Fig. 2. Proposed 1,2,4-thiadiazole-1,2,4-triazole based compounds.
**Fig. 3.** Overlay of native drug-like compounds over re-docked drug compound.

Table 6
Calculated binding affinity and residues involved in the interaction.

| Binding Affinity (kcal/mol) | Residues involved in the interactions | Types of Non-bonding interaction involved |
|-----------------------------|---------------------------------------|------------------------------------------|
| a -10.5                     | GLY-833, ALA-731, GLU-738, LYS-721, VAL-702, GLY-772, LEU-820, ALA-719, CYS-773, PHE-699, ASP-831 | Conventional hydrogen bond, Carbon hydrogen bond, Pi-Anion, Pi-Sigma, Pi-Sulfur, Pi-Pi Stacked, Pi-Alkyl |
| b -9.6                      | GLY-833, ALA-719, THR-766, LEU-820, PHE-699, ASP-831, VAL-702, MET-769 | Conventional hydrogen bond, Carbon hydrogen bond, Pi-Anion, Pi-Sigma, Pi-Pi Stacked, Pi-Alkyl |
| c -9.5                      | LYS-851, PRO-853, GLU-739, GLU-734, ASP-831, PHE-699, VAL-702, LEU-820, MET-769, ALA-719, THR-766, THR-830 | Conventional hydrogen bond, Carbon hydrogen bond, Pi-Anion, Pi-Sigma, Pi-Pi Stacked, Pi-Alkyl |
| d -9.7                      | PRO-853, PHE-699, VAL-702, LEU-820, ALA-719, THR-766, GLN-767 | Conventional hydrogen bond, Carbon hydrogen bond, Pi-Anion, Pi-Pi Stacked, Pi-Alkyl |
| e -10.0                     | GLY-772, LEU-820, LEU-694, VAL-702, PHE-699, LYS-721, ASP-831, GLU-738, GLU-734, ARG-817 | Conventional hydrogen bond, Carbon hydrogen bond, Pi-Anion, Pi-Alkyl |
| f -10.0                     | PRO-853, LYS-851, GLU-734, ASP-831, PHE-699, VAL-702, LEU-820, ALA-719, MET-769, THR-766, GLN-767 | Conventional hydrogen bond, Carbon hydrogen bond, Pi-Anion, Pi-Pi Stacked, Alkyl, Pi-Alkyl |
| g -10.2                     | PRO-853, GLU-734, LYS-851, GLY-833, ASP-831, PHE-699, LEU-820, VAL-702, ALA-719, THR-766, GLN-767 | Conventional hydrogen bond, Carbon hydrogen bond, Pi-Anion, Pi-Pi Stacked, Pi-Alkyl |
| h -9.3                      | ILE-735, ALA-719, LYS-721, VAL-702, ASP-831 | Pi-Anion, Pi-Sigma, Alkyl, Pi-Alkyl |
| Etoposide                   | THR-766, LEU-820, VAL-702, THR-830, ARG-817, LEU-694, GLY-772, PRO-770 | Conventional hydrogen bond, Carbon hydrogen bond, Alkyl, Pi-Alkyl |

2. Experimental Design, Materials, and Methods

The optimised compounds were divided into two divisions “training set (80%) (5 compounds) and test set (20%) (3 compounds)” by using Kennard stone algorithm approach via Dataset Division GUI 1.2 software [6,7]. The selected 3D descriptors for training set were used for deve-
Fig. 4. Residual interactions between compound a and human lung cancer cell line (PDB ID: 1m17).

Fig. 5. (A) 3D structure of the target protein (PDB ID: 1m17) and (B) 3D structure of the prepared ligand.

Developing reliable QSAR model while the compounds for test set were used to validate the predictability of the developed QSAR model. Also, 2D descriptors were used to develop another set of QSAR model using Material studio software and the predicted IC_{50} were reported. The use of correlation coefficient in ascertaining the efficiency of develop model in QSAR study is not enough; this need validation of QSAR model where dependability and predicting ability of developed QSAR model can be confirmed [8]; therefore, internal and external validation via training set and the test set were calculated respectively. The docking study was investigated on 1,2,4-Thiadiazole-1,2,4-Triazole and epidermal growth factor receptor kinase with PDB ID: 1m17 downloaded from protein data bank (www.rcsb.org). The receptor was treated using EduPyMOL-v1.7.4.4-Win32 and the treated receptor was subjected to Autodock tool so as to locate binding site and then converted to .pdbqt format in preparation for docking calculation using autodock vina 1.1.2. The observed grid box was as follows: center (X = 23.568, Y = 9.824, Z = 59.396) and size (X = 40, Y = 40, Z = 40) as well as the spacing was set to be 1.00Å (Fig. 5).
Ethics Statement

Not Applicable.

CRediT Author Statement

Abel Kolawole Oyebamiji: Conceptualization, Methodology, Writing- Original draft preparation; Sunday Adewale Akintelu: Software and Visualization; Folake Ayobami Amao and Oreoluwa P. Amao: Data curation; Mary Oluwatosin KAKA: Writing- Reviewing and Editing; Adetoun E. Morakinyo: Writing- Reviewing and Editing; Banjo Semire: Supervision, Software, Validation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships which have or could be perceived to have influenced the work reported in this article.

Acknowledgement

We are grateful to the computational chemistry research laboratory, Department of Pure and Applied Chemistry, Ladoke Akintola University of Technology, Nigeria for the computational resources and Mrs E.T. Oyebamiji and Miss Priscilla F. Oyebamiji for their assistance in the course of this work. Also, this research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

[1] J.P. Yazala, S. Reddymasu, V. Deekala, R.R. Rudraraju, Design, Synthesis, and Biological Evaluation of 1,2,4-Thiadiazole-1,2,4-Triazole Derivatives Bearing Amide Functionality as Anticancer Agents, Arab. J. Sci. Eng. (2020), doi: 10.1007/s13369-020-04626-z.
[2] S.B. Olasupo, A. Uzairu, G. Shallangwa, S. Uba, Quantitative structure-activity relationship (QSAR) studies and molecular docking simulation of norepinephrine transporter (NET) inhibitors as anti-psychotic therapeutic agents, J. Turkish Chem. Soc. Sect. A Chem. 7 (1) (2019) 179–196.
[3] B.O. Sabitu, U. Adamu, G.A. Shallangwa, S. Uba, Computer-aided drug design and in silico pharmacokinetics predictions of some potential antipsychotic agents, Scientific African 12 (2021) e00734.
[4] A. Ousaa, B. Eldrissi1, M. Chamali, S. Chitta, A. Aouidate, M. Bouachrine, T. Lakhlifi, Quantitative structure-toxicity relationship studies of aromatic aldehydes to Tetrahymena pyriformis based on electronic and topological descriptors, J. Mater. Environ. Sci. 9 (1) (2018) 256–266.
[5] S. Jennifer, X.S. Mark, E. Charles, Structure of the Epidermal Growth Factor Receptor Kinase Domain Alone and in Complex with a 4-Anilinoquinazoline Inhibitor, The J. Biol. Chem. 277 (48) (2002) 46265–46272.
[6] A.K. Oyebamiji, I.O. Abdulsalami, B. Semire, Dataset on Insilico approaches for 3,4-dihydropyrimidin-2(1H)-one urea derivatives as efficient Staphylococcus aureus inhibitor, Data in Brief 32 (2020) 106195.
[7] A.K. Oyebamiji, O.A. Mutiu, F.A. Amao, O.M. Oyawoye, T.A. Oyedepo, B.B. Adeleke, B. Semire, Dataset on in-silico investigation on triazole derivatives via molecular modelling approach: A potential glioblastoma inhibitors, Data in Brief 34 (2021) 106703.
[8] R.O. Oyewole, A.K. Oyebamiji, B. Semire, Theoretical calculations of molecular descriptors for anticancer activities of 1, 2, 3-triazole-pyrimidine derivatives against gastric cancer cell line (MGC-803): DFT, QSAR and docking approaches, Heliyon 6 (2020) e03926.