In silico evaluation and docking studies of pyrazole analogs as potential autophagy modulators against pancreatic cancer cell line MIA PaCa-2

Hiba Hashim Mahgoub Mohamed 1,2,* , Amna Bint Wahab Elrashid Mohammed Hussien 3 and Ahmed Elsadig Mohammed Saeed 1

1 Department of Chemistry, College of Science, Sudan University of Science and Technology, 2288, Khartoum, Sudan
hibahashim.m@hotmail.com (H.H.M.); amsaeed@gmail.com (A.E.M.S.)
2 Department of Chemistry, Ibn Sina University, 10995, Khartoum, Sudan
3 College of Animal Production Science and Technology, Sudan University of Science and Technology, 2288, Khartoum, Sudan
annaelrasheed@gmail.com (A.B.W.E.M.H.)

* Corresponding author at: Department of Chemistry, College of Science, Sudan University of Science and Technology, 2288, Khartoum, Sudan.
e-mail: hibahashim.m@hotmail.com (H.H.M. Mohamed).

ABSTRACT
A quantitative structure activity relationship (QSAR) model for a series of N-(1-benzyl-3,5-dimethyl-1H-pyrazole-4-yl) benzamide derivatives having autophagy inhibitory activities as potent anticancer agents was developed by the multiple linear regressions (MLR) method. In this study, previous compounds were used in the model development were divided into a set of fifteen compounds as training set and set of four compounds as test set. A model with high prediction ability and high correlation coefficients was obtained. This model showed $r = 0.968$, $r^2 = 0.937$ and $Q^2 = 0.880$, the QSAR model was also employed to predict the experimental compounds in an external test set, and to predict the activity of a new designed set of 3,5-dimethyl-4-substituted-pyrazole derivatives (1-15), result showed that compound 3 has the most promising inhibition activity ($EC_{50} = 0.869 \mu M$) against human pancreatic ductal adenocarcinoma cell MIA PaCa-2 compared to the reference chloroquine with ($EC_{50} = 14 \mu M$). Thus, the model showed good correlative and predictive ability. Docking studies was performed for designed compounds, docking analysis showed the best compound 1 with high docking affinity of -24.8616 kcal/mol.

1. Introduction
Autophagy is a conserved intracellular degradation process that delivers substrates including bulk cytoplasm, organelles, aggregate-prone proteins, and infectious agents to lysosomes [12], it is induced under various conditions of cellular stress, which prevents cell damage and promotes survival in the event of energy or nutrient shortage. There are three types of autophagy: macroautophagy [3], microautophagy [4] and Chaperone-mediated autophagy (CMA) [5]. The molecular mechanism of autophagy involves several conserved ATG (autophagy-related) proteins, various stimuli lead to the formation of the phagophore; the elongation of the phagophore results in the formation of the characteristic double-membrane autophagosome. The autophagosomes fuses with the lysosome and release its inner compartment into lysosomal lumen, after fusion, a series of acid hydrolases are involved in degradation of the sequestered cytoplasmatic cargo. The small molecules resulting from the degradation, particularly amino acids are transported back to the cytosol for protein synthesis and maintenance of cellular functions under starvation conditions [6-9].

In cancer, the role of autophagy is highly complex and dependent on cancer type and stage [10], autophagy has been shown to act as a tumor suppressing to constrains tumor initiation in normal tissue, some tumor in last stage of progression rely on autophagy for tumor promotion and maintenance [11-13]. Therefore, targeting autophagy and discovering autophagy inhibitions and its modulation has considerable potential as anticancer agents used in therapeutic approach, especially in pancreatic ductal adenocarcinoma (PDAC), which it represents a viable approach to fight pancreatic cancer.

Quantitative structure activity relationships is a mathematical equation relating chemical structure with its physical, chemical and biological effect [14], QSAR model is useful for understanding the factors controlling activity and for designing new compounds for therapeutic areas [15-17], it requires a compound set that has been tested against an identified molecular target, cell tissue, or even microorganism,
under the same experimental conditions and possesses the minimum variance in the observed responses [18]. Once a suitable dataset has been selected, the main step of modeling requires molecular/physicochemical properties, followed by variable selection, model generation from different algorithms and validation process using internal and external dataset [19-21].

The current work aimed to obtain a QSAR model of \( N(1\text{-benzyl}-3,5\text{-dimethyl}-1H\text{-pyrazole}-4\text{-yl}) \) benzamide derivatives in order to predict biological activity against human pancreatic ductal adenocarcinoma cell MIA PaCa-2, validate the predictive ability of the developed model through validation methods, calculate the statistical parameter to prove quality of model, and use the obtained model to predict the biological activity against human pancreatic ductal adenocarcinoma cell MIA PaCa-2 on a set of designed compounds (1-15). And conducting docking studies for all designed compounds (1-15) and selected protein 6s6a.

2. Experimental

2.1. QSAR studies

2.1.1. Data set

A data set comprised of nineteen \( N(1\text{-benzyl}-3,5\text{-dimethyl}-1H\text{-pyrazole}-4\text{-yl}) \) benzamide derivatives was used in the present study. All compounds and associated data were obtained from literature [22]. The biological activity data were reported as (EC\(_{50}\)) values half maximal effective concentration in MIA PaCa-2a pancreatic cancer line. The (EC\(_{50}\)) values were converted into (pEC\(_{50}\)) using the formula: pEC\(_{50}\) = -log EC\(_{50}\), values along with the \( N(1\text{-benzyl}-3,5\text{-dimethyl}-1H\text{-pyrazole}-4\text{-yl}) \) benzamide derivatives structures can be found on Table 1. Chemical structures of the compounds were done using the ACD/ChemSketch v14.01 software (ACD, Copyright 1994-2013 Advanced Chemistry Development, Inc.), molecular modeling was performed using the Molecular Operating Environment software package (MOE, v2009.10; Chemical Computing Group Inc.). The QSAR model was derived from nineteen molecules which were randomly divided into training set of fifteen molecules and test set of four molecules was used to validate QSAR model.

2.1.2. Molecular descriptors generation

Different molecular descriptors (physicochemical properties) [23] were calculated for each molecule after the low energy conformer of structures were generated, these descriptors included electronic, spatial, and structural descriptor were calculated using MOE and ACD lab programs. In order to select the best subset of descriptors, and avoid difficulties in forming QSAR models, hence the predictivity and the generalization of the model fail under these conditions, highly correlated descriptors were excluded using correlation matrix, the nine descriptors used to generate QSAR model denoted as molecular weight (MW), molar volume (MV), molar refractivity (Mr), sum of atomic polarizabilities (S -aPol), density (D), index of refraction (InR), surface tension (ST), and Log octanol/water partition coefficient (log P (o/w)) reported in Table 2.

### Table 1. Biological activities and structures of \( N(1\text{-benzyl}-3,5\text{-dimethyl}-1H\text{-pyrazole}-4\text{-yl}) \) benzamide compounds obtained from literature [22].

| Compound | R\(^1\) | R\(^2\) | R\(^3\) | EC\(_{50}\) | pEC\(_{50}\) |
|----------|--------|--------|--------|--------|--------|
| 1        | H      | H      | H      | 10.00  | 5.000  |
| 2        | H      | C(O)NH\(_2\) | H      | 9.20   | 5.036  |
| 3        | H      | H      | C(O)NH\(_2\) | 42.00  | 4.377  |
| 4        | H      | C(O)NMe | H      | 6.20   | 5.208  |
| 5        | H      | C(O)N(Me)\(_2\) | H      | 6.40   | 5.194  |
| 6        | H      | C(O)NHEt | H      | 11.00  | 4.959  |
| 7        | H      | C(O)NHPf | H      | 8.90   | 5.051  |
| 8        | H      | C(O)NHPa | H      | 6.40   | 5.194  |
| 9        | -      | -      | -      | 14.00  | 4.854  |
| 10       | H      | Me     | H      | 6.20   | 5.208  |
| 11       | OMe    | H      | H      | 8.50   | 5.071  |
| 12       | H      | OMe    | H      | 6.20   | 5.208  |
| 13       | H      | H      | OMe    | 8.00   | 5.097  |
| 14       | Cl     | H      | H      | 6.30   | 5.201  |
| 15       | H      | Cl     | H      | 5.90   | 5.229  |
| 16       | H      | H      | Cl     | 4.20   | 5.377  |
| 17       | H      | CF\(_3\) | H      | 2.30   | 5.638  |
| 18       | H      | H      | CF\(_3\) | 0.80   | 6.097  |
| 19       | Me     | H      | CF\(_3\) | 0.62   | 6.208  |
| Chloroquine | - | -     | -      | 14.00  | 4.854  |
The QSAR model was developed from the training set compounds where the independent variables molecular descriptors and dependent response variable (pEC50) were subjected to multiple linear regressions (MLR) analysis, several QSAR models were developed. The comparison of squared correlation coefficients of the models reported in Table 3.

The resulting QSAR model Equation (3) exhibited a high regression coefficient. The model was justified by statistical parameters such as the correlation coefficient (r), squared correlation coefficient (r²), cross-validated regression coefficient (Q²), standard error of estimate (s), F-test value (F), and the root mean squared error (RMSE), and validated using random test set compounds Table 4, and was evaluated for the robustness of its predictions via the cross-validation coefficient.

The developed model was validated internally by training set compounds using leave-one-out (LOO) cross-validation technique. In this technique, one compound is eliminated from the data set at random in each cycle and the model is built using the rest of the compounds. The model thus formed is used for predicting the activity of the eliminated compound. The process is repeated until all the compounds are eliminated once. The cross-validated regression coefficient (Q2) was calculated.

External validation was performed in order to determine the predictive capacity of the developed model as judged by its application for the prediction of test set activity values.

The observed activities and those calculated by QSAR model (Equation 3) for training set and test set were presented in Table 5 and 6.
Table 6. Predicted pEC<sub>50</sub> values of test set

| Compound | pEC<sub>50</sub><sup>exp.</sup> | pEC<sub>50</sub><sup>pred</sup> | Residuals |
|----------|----------------|----------------|-----------|
| 2        | 5.0360         | 4.5665         | 0.4695    |
| 8        | 5.1940         | 4.9549         | 0.2391    |
| 12       | 5.2080         | 5.1417         | 0.0663    |
| 17       | 5.6380         | 6.1317         | -0.4937   |

Table 7. Structures and predicted pEC<sub>50</sub> values for designed 3,5-dimethyl-4-substituted-pyrazole derivatives against human pancreatic cancer cell line (MIA PaCa-2).

| Compound | R | pEC<sub>50</sub><sup>pred</sup> |
|----------|---|----------------|-----------|
| 1        |   | 5.1410         |
| 2        |   | 4.9845         |
| 3        |   | 6.0610         |
| 4        |   | 5.3002         |
| 5        |   | 4.8785         |
| 6        |   | 5.2189         |
| 7        |   | 4.9615         |
| 8        |   | 5.4439         |
| 9        |   | 5.4630         |
| 10       |   | 4.9848         |
| 11       |   | 5.0520         |
| 12       |   | 4.8002         |
| 13       |   | 5.2539         |
| 14       |   | 5.2989         |
| 15       |   | 4.7922         |
2.1.5. Predict the activity of designed 3,5-dimethyl-4-substituted-pyrazole derivatives

Chemical structures of the designed 3,5-dimethyl-4-substituted-pyrazole derivatives (1-15) were done using the ACD/ChemSketch, the developed QSAR model (Equation (3)) was used to predict their activity against human pancreatic ductal adenocarcinoma cell line MIA PaCa-2. The predicted activity expressed as pEC50 along with the structures reported in Table 7.

2.2. Molecular docking

Docking is a molecular modelling technique that is used to predict how a protein interacts with small molecules (ligands) by predicting the most possible type of interaction, the binding affinities, and the orientations of the docked ligands at the active site of the target protein. Molecular docking study was carried out in order to elucidate which of the designed 3,5-dimethyl-4-substituted-pyrazole derivatives (1-15) has the best binding affinity against the mechanistic (or mammalian) target of rapamycin complex 1 (mTORC1). The structure of mTORC1 used in the study was obtained from Protein Data Bank with PDB code 6s6a, structures of the designed 3,5-dimethyl-4-substituted-pyrazole derivatives (1-15) were prepared and saved as mol files, the prepared compounds were docked with prepared structure of 6s6a protein using MOE program. The bindingscore (S) of the complexes and amino acid interactions are reported in Table 8.

3. Results and discussion

3.1. QSAR studies

The studied compounds which were an autophagy modulator showed a promising role as anticancer agents. In the present work, structure activity relationship model was developed that could correlate the structural features with biological activity.

The developed model showed squared correlation coefficient (r2 = 0.937) which indicates the correlation between the activity (dependent variable) the molecular descriptors (independent variable) for the training set data, and squared cross-validation (Q2 = 0.880) which indicates that the newly developed QSAR model has a good prediction. Three molecular descriptors denoted as log octanol/water partition coefficient (log P(o/w)), density (D), and surface tension (ST) were significantly correlated with anticancer activity. It is evident from the Equation (3) that amongst the molecular descriptors, log P(o/w) and D are positively correlated, that mean the biological activity increases when the values of these descriptors are positively increased. On the other hand, the descriptor ST negatively correlated with anticancer activity, that mean the biological activity decreases when the value of this descriptor is negatively correlated with anticancer activity.

3.2. Docking study

Molecular docking study was carried out between the target (mTORC1) and designed 3,5-dimethyl-4-substituted-pyrazole derivatives (1-15). All compounds were found to inhibit the receptor by occupying the active sites of the target (mTORC1). The binding affinity values for designed compounds range from -24.8616 to -18.0398 kcal/mol as reported in Table 8.

Table 8. Binding scores and interactions of the docked designed 3,5-dimethyl-4-substituted-pyrazole derivatives (1-15) on the active site of 6s6a.

| Compound | S (kcal/mol) | Amino acid interaction | Type of interaction | Length (Å) |
|----------|--------------|------------------------|--------------------|------------|
| 1        | -24.8616     | ArgA37                 | π-cation interaction| -          |
|          |              | LysA128                | π-cation interaction| -          |
|          |              | ThrA21                 | Metal complexation (Mg) | 2.16       |
|          |              | ThrA42                 | Metal complexation (Mg) | 2.22       |
| 2        | -22.0658     | ArgA37                 | π-cation interaction| -          |
|          |              | LysA128                | π-cation interaction| -          |
|          |              | ThrA21                 | Metal complexation (Mg) | 2.16       |
|          |              | ThrA42                 | Metal complexation (Mg) | 2.22       |
| 3        | -20.1774     | ArgA37                 | π-cation interaction| -          |
|          |              | LysA128                | π-cation interaction| -          |
| 4        | -24.8373     | ArgA37                 | π-cation interaction| -          |
|          |              | ThrA21                 | Metal complexation (Mg) | 2.16       |
|          |              | ThrA42                 | Metal complexation (Mg) | 2.22       |
| 5        | -18.9748     | ArgA37                 | π-cation interaction| -          |
|          |              | AspA130                | Hydrogen bond      | 2.03       |
| 6        | -22.5543     | ArgA37                 | π-cation interaction| -          |
|          |              | LysA128                | π-cation interaction| -          |
| 7        | -18.0398     | ArgA37                 | π-cation interaction| -          |
|          |              | ArgA37                 | Hydrogen bond      | 3.04       |
| 8        | -19.4697     | ArgA37                 | π-cation interaction| -          |
|          |              | ArgA37                 | π-cation interaction| -          |
|          |              | LysA128                | π-cation interaction| -          |
| 9        | -20.3244     | ArgA37                 | π-cation interaction| -          |
| 10       | -22.4642     | LysC179                | π-cation interaction| -          |
|          |              | SerC76                 | Hydrogen bond      | 3.04       |
| 11       | -21.7914     | ArgA37                 | π-cation interaction| -          |
|          |              | ArgA37                 | Hydrogen bond      | 2.97       |
|          |              | LysA128                | π-cation interaction| -          |
| 12       | -21.0138     | ArgA37                 | π-cation interaction| -          |
|          |              | LysA128                | π-cation interaction| -          |
| 13       | -24.0266     | ArgA37                 | π-cation interaction| -          |
|          |              | LysA128                | π-cation interaction| -          |
| 14       | -22.5484     | ArgA37                 | π-cation interaction| -          |
|          |              | ArgA37                 | π-cation interaction| -          |
| 15       | -18.7338     | ArgA37                 | π-cation interaction| -          |
|          |              | LysA128                | π-cation interaction| -          |
Figure 1. Predicted versus experimental pEC₅₀ values of (a) training set, (b) cross validation set, and (c) test set against human pancreatic cancer MIA PaCa-2.

Figure 2. 2D molecular docking model of compounds 1, 4 and 13 with 6s6a.

Figure 3. 3D model of the interaction between compounds 1, 4 and 13 with 6s6a.
However, three ligands (1, 4, and 13) have higher binding score, which ranges from -24.8616 to -24.0026 kcal/mol, ligand 1 formed two π-cation interaction with ArgA37 and LysA128, and two metal complexations with Mg ion with ThrA21 and ThrA42. Ligand 4 formed three interactions, a π-cation interaction with ArgA37 and two metal complexations with Mg ion with ThrA21 and ThrA42. Ligand 13 formed two π-cation interactions with ArgA37 and LysA128 (Figure 2 and 3).

4. Conclusion

In this work, a QSAR study was performed based on theoretical molecular descriptors, the built model serves as a guide for providing the structural requirements affecting the anticancer activity against pancreatic cancer cell line MIA PaCa-2, through the identification of the most relevant selected molecular descriptors in the models, comprehensive assessment (internal and external validation) indicate that the built QSAR model was robust and satisfactory, and that the selected descriptors could account for the structural features responsible for anticancer drugs activity of the compounds. The QSAR model developed and molecular docking in this study can provide a useful tool to predict the activity of new compounds and also to design new compounds with high activity.

Disclosure statement

Conflict of interests: The authors declare that they have no conflict of interest.

Author contributions: All authors contributed equally to this work.

Ethical approval: All ethical guidelines have been adhered.

ORCID

Hiba Hashim Mahgoub Mohamed
http://orcid.org/0000-0002-1294-6130

Anna Bint Wahab Elrashid Mohammed Hussien
http://orcid.org/0000-0001-7588-0231

Ahmed Elsadig Mohammed Saeed
http://orcid.org/0000-0002-7317-8040

References

[1]. Levine, B.; Kroemer, G. Cell 2008, 132(1), 27-42.
[2]. Glick, D.; Barth, S.; Macleod, K. F. J. Pathol. 2010, 221(1), 3-12.
[3]. Shintani, T.; Klionsky, D. J. Science 2004, 306(5689), 996-995.
[4]. Li, W. W.; Li, J.; Bao, J. K. Cell. Mol. Life Sci. 2012, 69(7), 1125-1136.
[5]. Arias, E.; Cuervo, A. M. Curr. Opin. Cell Biol. 2011, 23(2), 184-189.
[6]. Mizushima, N.; Komatsu, M. Cell 2011, 147, 728-741.
[7]. Badadani, M. ISRN Cell Biology 2012, 1, 1-11.
[8]. Kaur, J.; Debnath, J. Nat. Rev. Mol. Cell Biol. 2015, 16(8), 461-472.
[9]. Yu, L.; Chen, Y.; Tootoo, S. A. Autophagy 2018, 14(2), 207-215.
[10]. Kimmelman, A. C. Gene Dev. 2011, 25(19), 1999-2010.
[11]. Yang, X.; Yu, D. D.; Yan, F.; Jing, Y. Y.; Han, Z. P.; Sun, K.; Liang, L.; Hou, J.; Wei, L. X. Cell Biosci. 2015, 5(1), 14.
[12]. Gracio, D.; Magro, F.; Lima, R. T.; Maximo, V. Hematol. Med. Oncol. 2017, 2(1), 1-4.
[13]. Marinovic, M.; Sprung, M.; Buljubasic, M.; Novak, I. Oxid. Med. Cell Longev. 2018, 2018, 1-18.
[14]. Hernandez, N.; Kiralj, R.; Ferreira, M. M.; Talavera, I. Chemometr. Intell. Lab. 2009, 98(1), 65-77.
[15]. Hasegawa, K.; Yokos, N.; Watanabe, K.; Hirata, M.; Miyashita, Y.; Sasaki, S. Chemom. Intell. Lab. 1996, 33, 63-69.
[16]. Soni, H. M.; Patel, P. K.; Chhabria, M. T.; Rana, D. N.; Mahajan, B. M.; Brahmkshatriya, P. S. Comp. Chem. 2015, 3, 45-53.
[17]. Duchowicz, P. R.; Fioretti, S. E.; Castro, E.; Wróbel, K.; Ibezm, N. E.; Baccel, D. E. ChemistrySelective 2017, 2(13), 3725-3731.
[18]. Roy, K.; Kar, S.; Das, R. N. Statistical methods in QSAR/QSPR. In A primer on QSAR/QSPR modeling; Springer, Cham, 2015.
[19]. Tong, W.; Xie, Q.; Hong, H.; Shi, L.; Fang, H.; Perkins, R. Environ. Health Persp. 2004, 112(12), 1249-1254.
[20]. Veerasamy, R.; Rajak, H.; Jain, A.; Sivadasan, S.; Varghese, C. P.; Agrawal, R. K. Int. J. Drug Des. Discov. 2011, 5, 511-519.
[21]. Golbraikh, A.; Trophash, A. ACS Sym. Ser. 2002, 16, 357-396.
[22]. Al, T.; Willett, R.; Williams, J.; Ding, R.; Wilson, D. J.; Xie, J.; Kim, D. H.; Puertollano, R.; Chen, L. ACS Med. Chem. Lett 2016, 8(1), 90-95.
[23]. Mauri, A.; Consonni, V.; Pavan, M.; Todescini, R. MATCH Commun. Math. Comput. Chem. 2006, 56(2), 237-248.