

# QSAR Model Generation of Phthalazinones as Poly (ADP-Ribose) Polymerase Inhibitors by the Genetic Algorithm and Multiple Linear Regression (GA-MLR) Method: A Ligand-Based Approach for Cancer Drug Design

George Oche Ambrose<sup>1\*</sup>, Olanrewaju John Afees<sup>2,3</sup>, Uchechukwu Joseph Kalu<sup>4</sup>, Abiodun Wisdom Oshireku<sup>5</sup>, Afolayan Daniel Todimu<sup>1</sup>, Olayinka Emmanuel Oluwasegun<sup>6</sup>, Toba Olatoye<sup>1</sup>, Fagbemi Ranti-Ade Rebecca<sup>1</sup> and Adekunle Precious<sup>1</sup>

<sup>1</sup>Department of Biochemistry, University of Ilorin, Ilorin, Nigeria

<sup>2</sup>Department of Anatomy, University of Ilorin, Ilorin, Nigeria

<sup>3</sup>Ben Carson School of Medicine, Ilesha-Remo, Nigeria

<sup>4</sup>University of Calabar, Calabar, Nigeria

<sup>5</sup>University of Ibadan, Ibadan, Nigeria

<sup>6</sup>Olabisi Onabanjo University, Ogun State, Nigeria

## Abstract

Poly (ADP-ribose) polymerase-1 (PARP-1), an enzyme known for catalyzing the attachment (covalently) of polymers of ADP-ribose moieties on itself and its target proteins, has been reported in recent study to regulate gene expression in prostate cancer. BRCA mutations are associated in the sensitivity of PARP inhibitors. The present study aimed to develop a Quantitative Structure-Activity Relationship (QSAR) model with Phthalazinones, inhibitors of PARP-1. Phthalazinones were divided into training and test sets to build the QSAR model. Among the several topological, constitutional, geometrical, electronic and hybrid descriptors generated as inputs to the model, three variables were selected by adopting the genetic algorithm subset selection method (GA). The correctness of the proposed model was accounted for by using the following evaluation techniques: Y-randomization, Validation of the external data test set and cross-validation. The model was found to have a good predictive ability and could be used for designing similar group of compounds.

**Keywords:** Parp-1; GA-MLR; Phthalazinones; Y-randomization

## Introduction

PARP-1 is increasingly gaining attention as an anticancer pharmacological target in both preclinical examinations and clinical trials. In the treatment of breast cancer, ovarian cancer, prostate cancer, pancreatic cancer and unspecified solid tumors, some PARP-1 inhibitors/antagonist have received FDA approval such as olaparib (AZD2281), veliparib (ABT-888), and rucaparib (AG-014699) [1-4]. Inhibition of PARP-1 results to the accumulation of DNA damage. This occurs by impairing single-strand DNA break repair (SSBR) and trapping PARP-1 at single-strand break sites, which leads to inhibition of DNA replication [4]. The role of PARP-1 includes cell proliferation, survival and death; this is due to its effects on the regulation of multiple biological processes [5,6]. Recent studies have shown that metastasis, deterioration and angiogenesis in tumors are associated with elevated expression of PARP-1 protein [7,8]. The role of PARP-1 in cancer therapy makes it an interesting target for inhibition by small molecules. Recently, a novel series of phthalazinones as inhibitors of PARP-1 have been reported by Loh and colleagues [9]. The experimental estimations of the inhibitory activity of chemical molecules is difficult, time-consuming and expensive, therefore, a great deal of effort has been required into attempting the measurements of activity via statistical modeling. QSAR analysis is an effective approach in research which has been applied into rational drug design and the mechanism of drug actions. In QSAR studies, biological activities of compounds are expressed as a function of their various structural properties which explains how the variation in biological activity relies on changes in the chemical structures [10]. The advances of QSAR study depends largely on choosing a robust statistical methods for producing the predictive model and also the required structural properties for expressing the essential features within those chemical structures. In recent times, genetic algorithms (GA) are widely adopted methods for variable

selection [11-13]. In this study, we report the development of a QSAR model for PARP-1 inhibition which has not yet been done.

## Materials and Methods

### Accession of experimental data

In this present Quantitative Structure-Activity Relationship (QSAR) study, a set of twenty-six (26) Phthalazinones was retrieved from the ChEMBL database (<http://www.ebi.ac.uk/chembl>) with accession ID of ChEMBL1141921 [14]. This dataset represent a novel series of potent inhibitors of poly (ADP-ribose) polymerase in terms of IC<sub>50</sub> (μM). The biological activity data (IC<sub>50</sub>) were then converted to PIC<sub>50</sub> values using the formula PIC<sub>50</sub> = (-Log (IC<sub>50</sub> X)) (was used as the depended variable). The structures of Phthalazinones are listed in Table 1 with their observed activities.

### Accession of chemical structures

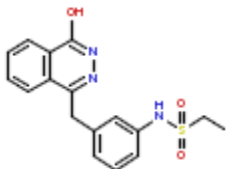
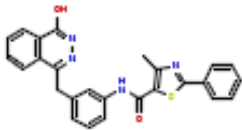
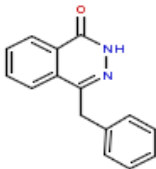
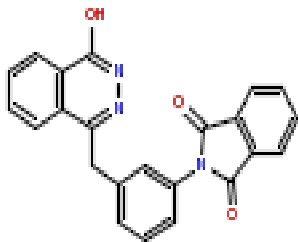
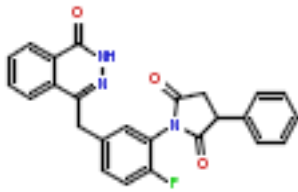
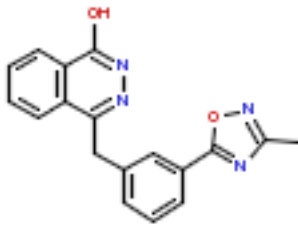
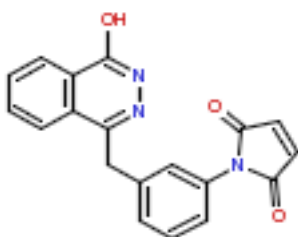
The canonical smiles of Phthalazinones obtained from the ChEMBL

**\*Corresponding author:** George Oche Ambrose, Department of Biochemistry, University of Ilorin, Ilorin, Nigeria, Tel: +031-821304; Fax: + 031-221937; E-mail: [ocheab1@gmail.com](mailto:ocheab1@gmail.com)

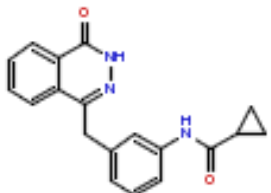
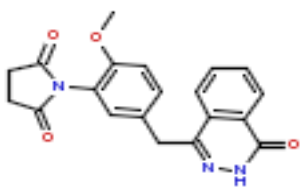
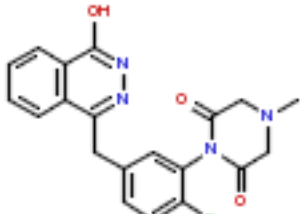
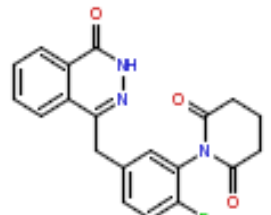
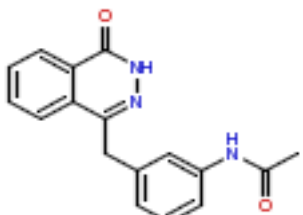
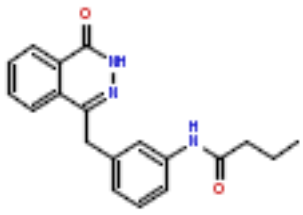
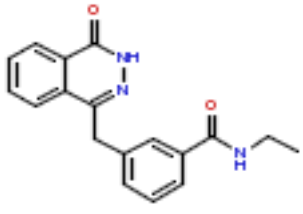
**Received** September 29, 2018; **Accepted** October 29, 2018; **Published** November 05, 2018

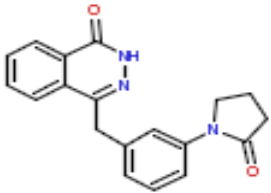
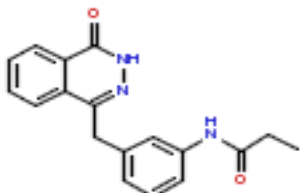
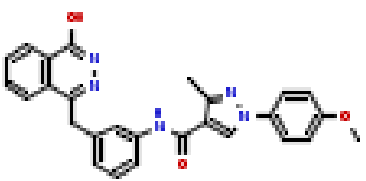
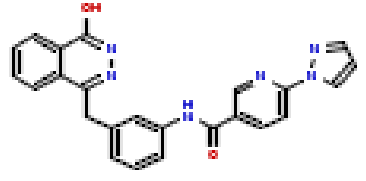
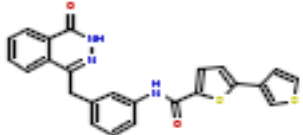
**Citation:** Ambrose GO, Afees OJ, Kalu UJ, Oshireku AW, Todimu AD, et al. (2018) QSAR Model Generation of Phthalazinones as Poly (ADP-Ribose) Polymerase Inhibitors by the Genetic Algorithm and Multiple Linear Regression (GA-MLR) Method: A Ligand-Based Approach for Cancer Drug Design. J Proteomics Bioinform 11: 183-191. doi: [10.4172/0974-276X.1000485](https://doi.org/10.4172/0974-276X.1000485)

**Copyright:** © 2018 Ambrose GO, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

S/N	CHEMBL ID	CHEMICAL STRUCTURES	IC50 (nM)	pIC50 (nM)	Normalized Data (nM)
1	CHEMBL193917		290	6.54	0.189
2	CHEMBL383578		27	7.57	0.64
3	CHEMBL66761		770	6.11	0
4	CHEMBL370692		180	6.75	0.281
5	CHEMBL196450		6.8	8.17	0.904
6	CHEMBL193918		189	6.72	0.268
7	CHEMBL363617		13	7.89	0.781

8	CHEMBL195913		19	7.72	0.706
9	CHEMBL196444		5	8.3	0.961
10	CHEMBL197192		9.8	8.01	0.833
11	CHEMBL371244		36	7.44	0.583
12	CHEMBL193512		370	6.43	0.14
13	CHEMBL195402		12	7.92	0.794
14	CHEMBL436298		5	8.3	0.961

15	CHEMBL196507		55	7.26	0.504
16	CHEMBL371205		33	7.48	0.601
17	CHEMBL195966		9.5	8.02	0.838
18	CHEMBL370217		4.1	8.39	1
19	CHEMBL381652		90	7.05	0.412
20	CHEMBL425560		90	7.05	0.412
21	CHEMBL381208		50	7.3	0.522

22	CHEMBL196559		120	6.92	0.355
23	CHEMBL371425		20	7.7	0.697
24	CHEMBL382950		10	8	0.829
25	CHEMBL193903		47	7.33	0.535
26	CHEMBL372450		56	7.25	0.5

**Table 1:** Structures and biological activity of compounds.

database were converted to SDF files with 2D and 3D coordinates using data Warrior software version 4.7.2. The 2D and 3D QSAR model generated in this study was derived from the training dataset of 20 molecules while the predictive potential of this model was evaluated by the test set of 6 molecules with uniformly distributed biological activities. Table 2 shows the observed and predicted biological activities of the training and test datasets.

### Descriptors generation

In order to develop a Quantitative Structure-Activity Relationship (QSAR) model, the biological activity of compounds must be quantitatively represented by molecular descriptors. The Chemistry Development Kit (CDK) descriptor version 1.0 was used for the calculation of different descriptors under the following categories: Topological descriptors, Geometric descriptors, Hybrid descriptors, Electronic descriptors and Constitutional descriptors. The calculated descriptors were arranged in a data matrix. The preprocessing or pretreatment of the independent variables (i.e., descriptors) was done by removing invariable (constant column) and other descriptors based on a variance cut-off of 0.0001 and correlation coefficient cut-off of 0.99 using J Frame VWSP version 1.0.

### Data normalization

Due to the existence of much variability in the range and distribution of each variable in the data set, the calculated values of the descriptors of each compound with their corresponding biological activity were subjected to a statistical technique known as min-max normalization using Normalize. The Data software version 1.0. In min-max normalization, the minimum and maximum value of each variable is adjusted to a uniform range between 0 and 1 according to the following equation:

$$x_{normalized} = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (1)$$

Where  $x_{normalized}$  represents the min-max normalized value,  $x_i$  represents the value of interest,  $x_{min}$  represents the minimum value, and  $x_{max}$  represents the maximum value.

### Selection of training and test set

The dataset of 26 Phthalazinones molecules was divided into training and test set based on Kennard-Stone method [15,16] using the J Frame Division software version 1.0. In this method, dissimilarity

Training Set		Selected Descriptors		Observed and Predicted values		Outlier Information
Compounds	khs.aasC	bpol	TPSA	Observed	Predicted	Outlier
2	1	0.793	0.219	0.64	0.523117788	-
3	0	0	0.231	0	8.01E-04	-
4	0.667	0.399	0.519	0.281	0.704651415	-
5	0.5	0.676	0.868	0.904	0.876213482	-
6	0.667	0.297	0	0.268	0.174108841	-
7	0.333	0.272	0.588	0.781	0.561471385	-
8	0.333	0.377	0.739	0.706	0.694655479	-
9	0.333	0.423	0.939	0.961	0.897221843	-
10	0.5	0.659	0.799	0.833	0.80664574	-
12	0.333	0.52	0.363	0.14	0.24932188	-
13	0.167	0.416	0.898	0.794	0.729157094	-
14	0.5	0.405	0.822	0.961	0.902829557	-
15	0.167	0.411	0.712	0.504	0.530150087	-
16	0.333	0.654	0.882	0.601	0.770868974	-
17	0.5	0.71	0.855	0.838	0.852648456	-
18	0.333	0.549	1	1	0.927528838	-
20	0.167	0.538	0.62	0.412	0.395320488	-
22	0.167	0.487	0.587	0.355	0.374099969	-
24	1	1	0.489	0.829	0.755849723	-
25	0.667	0.648	0.502	0.535	0.616338063	-
Test set		Selected Descriptors		Observed and Predicted values		AD information
Compounds	khs.aasC	bpol	TPSA	Observed	Predicted	AD
1	0.333	0.714	0.385	0.189	0.218491495	-
11	0.167	0.294	0.947	0.583	0.816249215	-
19	0.167	0.284	0.71	0.412	0.563695923	-
21	0.167	0.411	0.967	0.522	0.804909229	-
23	0.167	0.411	0.824	0.697	0.650828612	-
26	0.667	0.675	0.374	0.5	0.470829812	-

**Table 2:** Normalized values of selected descriptors and the observed/predicted Y values (Normalized values).

value gives an idea to handle training and test set size. This method is used for MLR model with pIC50 activity values as dependent variable and the various 2D and 3D descriptors calculated for the molecules as independent variables.

### QSAR Model development

In this study, QSAR model was developed from the dataset using the Multiple linear regression (MLR) method to screen potential leads against PARP-1 within a training dataset set (20 compounds). The total molecular descriptors (108) was calculated for each compound using CDK algorithm. Finally, a robust QSAR model equation was derived by MLR; Irrelevant descriptors were removed based on the Inter Correlation cut-off of 0.99 and Variance cut off of 0.001 using the Genetic Algorithm v4.1 software which leads to a selection of three (3) descriptors (one 3D and two 2D) in the final QSAR regression equation (Table 2). The model creates a relationship in the form of a straight line (linear) equation that best approximates all the individual data points. Regression equation takes the form.

$$Y = b_1x_1 + b_2x_2 + b_3x_3 \quad \text{Equation 2}$$

Where Y is dependent variable, 'b's are regression coefficients for corresponding 'x's (independent variable), 'c' is a regression constant or intercept.

### Model validation

Model validation is necessary in QSAR modeling, it confirms the reliability of the developed QSAR model along with the acceptability of each step during model development [17]. Model validation is

done to test the internal stability and predictive ability of the QSAR models. The developed QSAR models in this study were validated by the following method:

**Internal validation:** Internal validation was carried out using leave-one-out (LOO-) method. In the leave-one-out (LOO) method of cross validation, the process of removing a molecule, and creating and validating the model against the individual molecules is performed for all the Q<sub>2</sub> (rCV<sub>2</sub>) values and reported. The rCV<sub>2</sub> (cross-validation regression coefficient) was calculated using equation (3), which describes the internal stability of a model.

$$rCV^2 = 1 - \frac{\sum (Y_{obs} - Y_{pred})^2}{\sum (Y_{obs} - \bar{Y})^2} \quad (3)$$

In the above equation, Y-means the average activity value of the training dataset, while Y<sub>obs</sub> and Y<sub>pred</sub> represent the observed and predicted activity values respectively. A high rCV (>0.5) suggests a reasonably robust model [18].

**Estimation of the predictive ability of a QSAR model:** After the internal validation process, the high predictive power of a QSAR model should be estimated from an external test set of compounds that are not used in building of the QSAR model. The external validation or predictive capacity of the obtained model was judged by predictive R<sub>2</sub> (R<sub>pred</sub><sup>2</sup>) as shown in given equation:

$$r_{pred}^2 = 1 - \frac{\sum (Y_{pred(test)} - Y_{(test)})^2}{\sum (Y_{(test)} - \bar{Y}_{(training)})^2} \quad (4)$$

Where  $Y_{pred(test)}$  and  $Y_{(test)}$  indicate the predicted and observed activity values, respectively, for test set compounds and  $Y_{(training)}$  indicates the average bioactivity of compound in the training set. An acceptable predictive power of a QSAR model ( $R_{pred}^2$ ) should be  $>0.6$  for the test set molecules [19-21].

## Results

### GA-Multiple linear regression

According to the inter-correlation coefficients of the descriptors, highly correlated descriptors were removed from the study by a genetic algorithm method using a correlation regression cut-off of 0.99. According to the rule of thumb in MLR (ratio of sample size to the number of descriptors should be greater than or equal to 5), a tetra-parametric model can be expected with the current training set of 20 compounds. This can be shown below:

$$pIC50 = -0.2481(+/-0.1155) + 0.7582(+/-0.2042) \text{ khs.aasC} - 0.2811(+/-0.2448) \text{ bpol} + 1.0775(+/-0.1539) \text{ TPSA}$$

$$n = 20, R^2 = 0.8038, R_a^2 = 0.767, F = 21.85072, q^2 = 0.6727, r_{pred}^2 = 0.61915, SEE=0.1421, SDEP= 0.1641, PRESS : 0.32297$$

The above equation indicates that the model obtained with GA-MLR showed good squared correlation coefficient ( $R^2$ ) value and good internal predictive power ( $rCV^2$ ) with an excellent external predictive power ( $r_{pred}^2$ ). The scatter plot which is plotted between observed and predicted pIC50 values for training set and test set are shown in the Figure 1a and b respectively. A plot of the residual for the predicted values for both the test and training data sets against the experimental

pIC50 values is shown in Figure 2. It can be deduced from the plot that the model did not show any proportional and systematic error. This is because the propagation of the residuals on both sides of the zero are random. The derived QSAR model fitted with GA-MLR presents a significant relationship between pIC50 values (dependent variable) and the selected descriptors (independent variables). The value of the regression coefficient ( $R^2=0.8038$ ) indicates the existence of ~80.4% correlation between the activity and the selected descriptors in the training dataset, while the value of the cross-validation regression coefficient ( $q^2 = 0.6727$ ) suggests ~67.2% prediction accuracy of this QSAR model. This QSAR model fitted with GA-MLR can be used to predict future observations.  $R_{pred}^2 = 0.61915$ , shows the predictive power of the model. To judge the overall significance of the regression coefficients, the variance ratio (F) is computed. The F value has two degrees of freedom: p, N-p-1. For overall significance of the regression coefficients, the F value should be high. Also, for a good model, the standard error of estimate (SEE) of Y should be low. Finally, model predictivity is judged using the predicted residual sum of squares (PRESS) and cross-validated  $R^2$  ( $Q^2$ ) for the model while the value of standard deviation of error of prediction (SDEP) is calculated from PRESS.

### Y-Randomization

The Y-randomization test was carried out in order to ensure that there is no random correlation. By this, we could validate the established QSAR model and confirm that the selected descriptors are not random, and consequently, the result model should have low statistical quality. Random MLR models are generated in this test. This is done by randomly shuffling the dependent variable while keeping the independent variables as it is. The newly established QSAR models are expected to give significantly low values of  $R^2$  and  $Q^2$  for several trials; which confirm that the developed QSAR models are robust [22]. In this study, five trials of Y-randomization was carried out and the five random models generated gave lower values of  $R^2$  and  $Q^2$  thereby validating the original model (the established GA-MLR model) (Table 3). Another parameter;  $cRp^2$  is also estimated, which should be greater than 0.5 for passing the test [22] following the equation below:

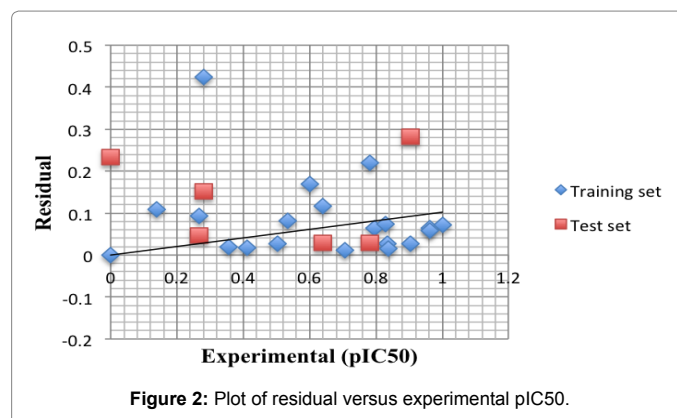
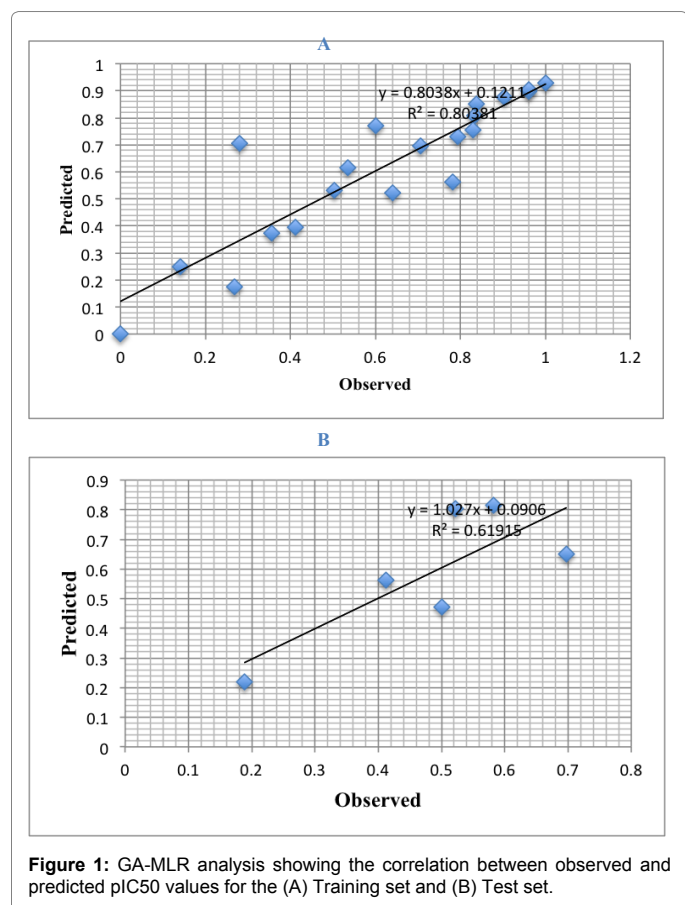
$$cRp^2 = R'(R^2 - (\text{Average } Rr)^2)^{1/2}$$

Where  $Rr$  = Average 'R' of random models

Table 4 shows that the  $cRp^2$  calculated in this study is 0.688462594 which is greater than 0.5 and thus confirmed that the test is passed.

### Applicability domain

Applicability Domain (AD) refers to the response and chemical



structure space in which the QSAR model makes predictions with a given reliability [22]. We carried out the AD using standardization approach in order to find out the test set compounds that falls outside the applicability domain and also to detect training set compounds that are outliers. The software adopted for this analysis is called “AD using standardization approach”. This software is developed in java language at Drug Theoretics and Chemoinformatics laboratory. Table 4 reveals that there is no outliers among the training set which conforms with the normal distribution pattern of about 99.7% of the population remaining with the range mean of  $\pm 3$  standard deviation (SD). Thus, mean  $\pm 3$  describes the region where most of the training data set compounds belong to. Any compound found outside this region is dissimilar to the rest and majority of the compounds. Table 5 also show that no test compound is found outside the AD. Therefore, this suggests that the QSAR model developed in this study can make predictions with a given reliability. Another required aspect is how to evaluate the performance of AD. The rule that is universally accepted is

Model	R	R <sup>2</sup>	Q <sup>2</sup>
Original	0.896552594	0.803806554	0.672688449
Random 1	0.624637279	0.39017173	-0.208879566
Random 2	0.439219719	0.192913962	-0.180816154
Random 3	0.232298582	0.053962631	-0.556768219
Random 4	0.528915084	0.279751166	-0.057761159
Random 5	0.48867299	0.238801291	-0.474684813

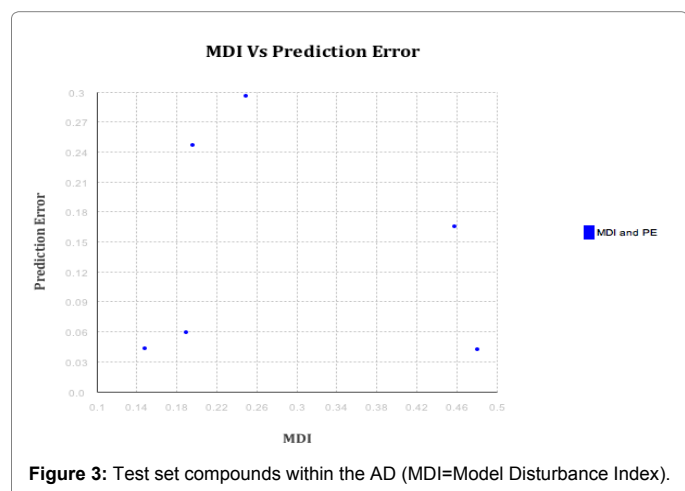
**Table 3:** Five trials Y-randomization outcome.

Random Models Parameters	
Average r :	0.462748731
Average r <sup>2</sup> :	0.231120156
Average Q <sup>2</sup> :	-0.295781982
cRp <sup>2</sup> :	0.688462594

**Table 4:** Y-randomization model's quality check parameters.

S/N	Descriptors	Description
1	khs.aasC (2D)	A fragment count descriptor that uses e-state fragments. Traditionally the e-state descriptors identify the relevant fragments and then evaluate the actual e-state value
2	bpol (2D)	Sum of the absolute value of the difference between atomic polarizabilities of all bonded atoms in the molecule (including implicit hydrogens)
3	TPSA (3D)	Sum of solvent accessible surface areas of atoms with absolute value of partial charges greater than or equal 0.2

**Table 5:** Selected descriptors with their respective description.



**Figure 3:** Test set compounds within the AD (MDI=Model Disturbance Index).

that the prediction error (PE) of the compound inside the AD should be lesser than the compound that are outside the AD [22]. Because all the test set compounds appear in the true positive quadrant, they are said to be inside the applicability domain Figure 3.

## Conclusion

In this study, GA-MLR was used in the construction of a robust QSAR model for Parp-1 inhibitors. Several validation techniques were used to validate the derived model. The model show good predictive potential for Parp-1 inhibitors which can be use to predict new Parp-1 inhibitors. These QSAR model could provide a reliable tool for the design of Parp-1 inhibitors.

## References

- Kaufman B, Shapira Frommer R, Schmutzler RK, Audeh MW, Friedlander M, et al. (2015) Olaparib monotherapy in patients with advanced cancer and a germline BRCA1/2 mutation. *J Clin Oncol* 33: 244-250.
- Solimando DA, Waddell JA (2017) Drug Monographs: Olaratumab and Rucaparib. *Hosp Pharm* 52: 258-263.
- Kummar S, Chen A, Ji J, Zhang Y, Reid JM, et al. (2011) Phase I study of PARP inhibitor ABT-888 in combination with topotecan in adults with refractory solid tumors and lymphomas. *Cancer Res* 71: 5626-5634.
- Lin KY, Kraus WL (2017) PARP Inhibitors for Cancer Therapy. *Cell* 169: 183.
- Jagtap P, Szabó C (2005) Poly (ADP-ribose) polymerase and the therapeutic effects of its inhibitors. *Nat Rev Drug Discov* 4: 421-440.
- Weaver AN, Yang ES (2013) Beyond DNA repair: additional functions of PARP-1 in cancer. *Front Oncol* 3: 290.
- Espinoza LA (2013) The Role of PARP Activation in Prostate Cancer. In *Advances in Prostate Cancer*. Intech.
- Wielgos ME, Rajbhandari R, Cooper TS, Wei S, Nozell S, et al. (2016) Let-7 Status Is Crucial for PARP1 Expression in HER2-Overexpressing Breast Tumors. *Mol Cancer Res* 15: 340-347.
- Loh VM, Cockcroft XL, Dillon KJ, Dixon L, Drzewiecki J, et al. (2005) Phthalazinones Part 1: The design and synthesis of a novel series of potent inhibitors of poly (ADP-ribose) polymerase. *Bioorganic & Medicinal Chemistry Letters* 15: 2235-2238.
- Bayat Z, Abad MFY (2011) Quantitative Structure-Property Relationship (QSPR) Study Of Kovats Retention Indices Of Some Of Adamantane Derivatives By The Genetic Algorithm And Multiple Linear Regression (GA-MLR) Method. *Petroleum & Coal* 53.
- Riahi S, Pourbasheer E, Ganjali MR, Norouzi P (2009) Investigation of different linear and nonlinear chemometric methods for modeling of retention index of essential oil components: Concerns to support vector machine. *J Hazard Mater* 166: 853-859.
- Depczynski U, Frost VJ, Molt K (2000) Genetic algorithms applied to the selection of factors in principal component regression. *Analytica Chimica Acta* 420: 217-227.
- Alsberg BK, Marchand Geneste N, King RD (2000) A new 3D molecular structure representation using quantum topology with application to structure-property relationships. *Chemom Intell Lab Syst* 54 75-91.
- <https://www.ebi.ac.uk/chembl/compound/inspect/ CHEMBL1141921>
- Andreasen PA, Petersen HH (2000) The plasminogen activation system in tumor growth, invasion, and metastasis *Cell Mol Life Sci* 57: 25-40.
- Agus S, Tadé MO, Vuthaluru H (2012) A Modified Kennard Stone Algorithm for Optimal Division of Data for Developing Artificial Neural Network Models. *Chemical Product and Process Modeling* 7.
- Roy K, Das RN, Ambure P, Aher RB (2016) Be aware of error measures. Further studies on validation of predictive QSAR models. *Chemom Intell Lab Sys* 15: 18-33.
- Yadav DK, Kumar S, Saloni S, Singh H, Sharma P, et al. (2017) Molecular docking, QSAR and ADMET studies of withanolide analogs against breast cancer. *Drug Des Dev Ther* 11: 1859-1870.



19. Yadav DK, Khan F (2013) QSAR docking and ADMET studies of camptothecin derivatives as inhibitors of DNA topoisomerase-I. *J Chemom* 27: 21-33.
20. Yadav DK, Mudgal V, Agrawal J, Maurya A, Bawankule D, et al. (2013) Molecular docking and ADME studies of natural compounds of Agar wood oil for topical anti-inflammatory activity. *Curr Comput Aided Drug Des* 9: 360-370.
21. Yadav DK, Kalani K, Khan F, Srivastava SK (2013) QSAR and docking based semi-synthesis and in vitro evaluation of 18 $\beta$ -glycyrrhetic acid derivatives against human lung cancer cell line A-549. *Med Chem* 9: 1073-1084.
22. Kunal R, Kar S, Ambure P (2015) On a simple approach for determining applicability domain of QSAR models. *Chemom Intell Lab Syst* 145: 22-29.