EMPRO

Current Research in Pharmaceutical Sciences

Available online at www.crpsonline.com

crps online CRPS

ISSN: 2250 – 2688

Received: 25/01/2017
Revised: 16/02/2017
Accepted: 23/02/2017

Amit Kumar Tiwari, Preveena Narasimhamurthy, Gopalpur Nagendrappa
P.G. Department of Chemistry, Centre for Postgraduate Studies, Jain University, 3rd Block, Jayanagar, Bangalore - 560011, (Karnataka), India

Abhilash Thakur
Department of Applied Science, National Institute of Technical Teachers Training and Research, Bhopal - 462002 (Madhya Pradesh), India

Correspondence

Amit Kumar Tiwari
P.G. Department of Chemistry
Centre for Postgraduate Studies
Jain University, 3rd Block, Jayanagar
Bangalore - 560011, (Karnataka), India

Email: meettiwari@gmail.com

DOI: 10.24092/CRPS.2017.070103

Website: www.crpsonline.com

Quick Response Code:

QSAR Modeling for Relative Toxicity Prediction of (3-(2-chloroquinolin-3-yl)oxiran-2-yl)(phenyl) methanone Derivatives

Amit Kumar Tiwari, Preveena Narasimhamurthy, Gopalpur Nagendrappa, Abhilash Thakur

ABSTRACT

2-Chloroquinoline-3-carbaldehyde and its substituted products are extremely versatile intermediates for synthesizing a variety of compounds containing quinoline moiety, which find many pharmaceutical and other applications. Quantitative structure-activity relationship (QSAR) plays an important role in toxicity prediction. The present study deals with acute toxicity predictions LD_{50} (median lethal dose) values of (3-(2-chloroquinolin-3-yl)oxiran-2-yl)(phenyl) methanone and its derivatives in rat by oral exposure through QSAR modelling software package T.E.S.T. In the present study the toxicity (LD_{50}) is evaluated using a variety of QSAR methodologies, such as hierarchical clustering, the Food and Drug Administration (FDA) MDL, nearest neighbor and a consensus model. For compounds No. 1 to 4, 7, 10 and 11 hierarchical clustering method does not provide the LD_{50} values; however, other methods have successfully provided the toxicity estimation for the same. The said software helps to predict the exact LD_{50} values when compared to experimental data reported in the range (>2000 to >5000 mg/kg). This is a preliminary observation from screening of LD_{50} values using the said software package. Further study may be relevant using other software to compare the predicted data.

Key words: QSAR analysis, Chloroquinoline, Rat, T.E.S.T., Toxicity, median lethal dose

1. INTRODUCTION

2-Chloroquinoline-3-carbaldehyde and its substituted products are remarkably versatile intermediates for synthesizing a variety of compounds containing quinoline moiety, which find many pharmaceutical and other applications[1]. The aldehydic and the chloro functional groups in chloroquinoline carbaldehyde molecule serve as synthons to prepare a range of useful products and have been of interest to medicinal chemists in the past decades. In keeping with this trend, the present authors have aimed to design eleven new compounds, namely, (3-(2-chloroquinolin-3-yl)oxiran-2-yl) (phenyl) methanone and its derivatives (1-11)[2]. In the present study, the authors have described in silico toxicity estimation of these compounds in detail.

Toxicity study is the study of adverse effects of chemical and physical agents and the degree to which a substance can harm humans or animals. Toxicity studies can be of chronic toxicity and acute toxicity, which involves harmful effects on an organism through a single or short term exposure. Sub-chronic toxicity is defined as the ability of a toxic substance to cause effects for more than one year but less than the life time of exposed organism. Chronic toxicity, on the other hand, is the ability of the substance or mixture of substances to cause harmful effects over an extended period, usually upon repeated and continuous exposure. Quantitative structure–activity relationship (QSAR) is a mathematical model that attempts to relate the structure-derived features (molecular descriptors) of a chemical compound to its biological or physicochemical activity.

Therefore, this method has been established for the predictive and ultimately diagnostic abilities. This can be used to predict the biological activity, viz., IC_{50}, LC_{50}/LD_{50}, EC_{50} etc. or class, viz., inhibitor versus non-inhibitor type of compounds before the actual bioassay. The molecular descriptors for QSAR are used on the basis of thermodynamic, steric and electronic parameters[3,4]. These parameters include partition coefficient, molecular volume, surface area, molecular refractivity etc. Also, the structural descriptors, which provide information about the various toxicological and pharmacokinetic aspects of the synthesized molecules, include E-state functions, kappa index, chi index, Lipinski's five rules and Wiener index[5].

An in silico method is also based on quantitative structure–activity relationship (QSAR) models, which can be used to understand drug action, design new compounds or drugs and screen chemical libraries [6,7,8,9]. The experimental measurement as bioassay with animals for compounds is difficult, more expensive and time-consuming. In order to mitigate the enormous difficulties associated with the animal tests,appreciable efforts have been expended in developing computational methods to predict biological activity through QSAR along with statistical modeling[10]. Recently, the European Chemicals Legislation, Registration, Evaluation and Authorization of Chemicals (REACH) have suggested the use of in silico method as a study for reliable toxicological risk assessment[11-12]. There are several recommended toxicity prediction softwares viz. TOPKAT (Toxicity Prediction by Komputer Assisted Technology)[13], DRAGON[14], ADMET (Absorption, Distribution, Metabolism, Elimination, and Toxicity)[15], V-life MDS[16] and ADME[17], T.E.S.T. (Toxicity Estimation Software Tool), PADEL[18], MDL QSAR (Elsevier MDL, 2006), Molconn-z (Edusoft-LC, 2006) etc.

According to USEPA[5], T.E.S.T. software is a simple QSAR model to calculate the toxicity of chemicals using a simple linear function of molecular descriptors, which is as follows:

$$\text{Toxicity} = ax_1 + bx_2 + c$$

Where, x_1 and x_2 are the independent descriptor variables and a, b, and c are fitted parameters. The T.E.S.T (Toxicity Estimation Software Tool) software provides multiple prediction methodologies, which has greater confidence in the predicted toxicities (as assuming the predicted toxicities are closely similar from different methods). In addition some researchers may have more confidence in particular QSAR approaches based on value added experience.

In the present study an attempt has been made to predict acute toxicity of (3-(2-chloroquinolin-3-yl)oxiran-2-yl) (phenyl) methanone and its ten derivatives (1-11) in the rat oral exposure for LD_{50} values using QSAR modeling software package. The comparisons were made between existing LD_{50} values obtained from different experimental and predicted LD_{50} values using T.E.S.T. software for the said compounds. The compounds 1-11 were evaluated for their toxicity using an in silico QSAR model. Since these predictions may be advantageous with respect to time and cost. Therefore by using Toxicity Estimation Software Tool (T.E.S.T.) version 4.2, a highly reliable QSAR model was applied for the present study.

2. MATERIALS AND METHODS

2.1 Drug and chemicals used

In this analysis, we investigate using the lethal dose that kills fifty percent of a test population (the LD_{50}) for determining the relative toxicity of a number of substances. In general, the smaller the LD_{50} value, the more toxic the chemical, and the larger the LD_{50} value, the lower the toxicity. When systemic toxicity and other specific toxicity data are unavailable for the chemical(s) of interest, during emergency responses, computed LD_{50} values may be employed to determine the relative toxicity of a series of chemicals. In the present study, a group of 11 new chemical entities ((3-(2-chloroquinolin-3-yl)oxiran-2-yl)(phenyl) methanone and its derivatives (1-11) as tabulated in Table 1 and 2, have been evaluated using three available rat oral QSAR LD_{50} models by using the said software. The present study was conducted to evaluate the toxicity (LD_{50}). T.E.S.T. estimates toxicity using a variety of QSAR methodologies, such as hierarchical clustering, the Food and Drug Administration (FDA) MDL, nearest neighborand a consensus model. The required descriptors are calculated without requiring any external programs. The structure of a chemical can be simply entered through the use of multiple tools including a chemical sketcher window, a text file containing SMILES notations, or importing it from a database of structures. After entering the structure, a chemical's toxicity can be estimated using one of several advanced methodologies. T.E.S.T. version 4.0 contains LD_{50} values from 7,420 chemicals.

The data reliability is tested by plotting a graph between experimental and predicted values of similar compounds (compounds whose similarity coefficient with test compound is greater than 0.5). The confidence on the predicted value is high, if the plot between predicted and experimental values of similar compounds gives an ideal line.

It was reported that T.E.S.T. software package estimates toxicity using a variety of QSAR methodologies[19], such as hierarchical clustering, the Food and Drug Administration (FDA) MDL, nearest neighbor and a consensus model. These methods are simply the average of the predicted toxicities from other QSAR methodologies, considering the applicability domain in each

method[20]. Generally molecular descriptors are physical characteristics of the structure of chemicals viz. the molecular weight or the number of benzene rings of a chemical. The overall pool of descriptors in the software (T.E.S.T.) contains 797 two-dimensional descriptors. The descriptors include the classes of descriptors viz. E-state values and E-state counts, constitutional descriptors, topological descriptors, walk and path counts, connectivity, information content, 2Dautocorrelation, Burden eigenvalue, molecular property (such as the octanol-water partition coefficient), kappa, hydrogen bond acceptor/donor counts, molecular distance edge, and molecular fragment counts. Following important methods were described in instruction manual for the present software[5]:

2.2 Hierarchical clustering method

In T.E.S.T., the hierarchical clustering method utilizes a variation of the Ward's Minimum Variance Clustering Method to contribute a series of clusters from the initial training set as per Romesburg[21]. According to Ruiz et al. [22], the change in variance caused by combining clusters j and k is in equation given below:

$$\Delta\sigma^2 = \frac{n_j n_k}{n_j + n_k} \sum_{i=1}^{d} (C_{j,i} - C_{k,i})^2 \quad (1)$$

where nj is the number of chemicals in cluster j, Cj,i is the centroid (or average value) for descriptor i for cluster j, and d is the number of descriptors (~800)[19]. It was noted that the predicted value for a given test chemical is calculated using the equally weighted average of the model predictions from the closest cluster from each step in the hierarchical clustering.

2.3 FDA MDL QSAR Method

In T.E.S.T., the FDA MDL method is based on the work of Contrera et al.[23]. In this method, it was noted that predictions for each test chemical are made using a unique cluster (constructed at runtime). It contains structurally similar chemicals selected from the overall training set. It is different from the Hierarchical method, where the predictions are made using one or more clusters, which are constructed a priori using Ward's method. For individual test chemical, a cluster is constructed using the 30 most similar chemicals from the training set as defined by the cosine similarity coefficient, SCi,k, which is calculated using the equation given below, as per USEPA (2012):

$$SC_{i,k} = \frac{\sum_{j=1}^{\#\,descriptors} x_{ij}\, x_{kj}}{\sqrt{\sum_{j=1}^{\#\,descriptors} x_{ij}^2 * \sum_{j=1}^{\#\,descriptors} x_{kj}^2}} \quad (2)$$

where Xij is the value of the j-th normalized descriptor for chemical i (normalized with respect to all of the chemicals in the original training set) and Xkj is the value of the j-th descriptor for chemical k. The entire pool of approximately 800 descriptors is always used to calculate the similarity coefficient in equation (2). A multiple linear regression model is then built for the new cluster using a genetic algorithm-based method, and the toxicity can be easily predicted [21].

2.4 Nearest Neighbor Method

In T.E.S.T. (USEPA, 2012), the nearest neighbor method is a simplification of the variable selection of kNN approach. It was observed in the nearest neighbor method, the toxicity is simply predicted as the average of the toxicity of the three most similar chemicals from the training set. The similarity is defined in terms of the cosine similarity coefficient (Equation 2).

2.5 Consensus Method

In the consensus method of T.E.S.T., the predicted toxicity is simply the average of the predicted toxicities from the above mentioned QSAR methodologies considering the applicability domain of individual method[24]. It was suggested, if only a single QSAR methodology can make a prediction, then the predicted value is unreliable and unable to use. This method typically provides the highest prediction accuracy by the predictions from the other above mentioned methods. In addition this method provides the highest prediction coverage because several methods with slightly different applicability domains are used to make a prediction[22.]

2.6 Statistical external validation

In T.E.S.T., the predictive ability of each of the QSAR methodologies was evaluated using statistical external validation as per Gramatica and Pilutti[25]. According to Golbraikh et al.[26], a QSAR model is acceptable on predictive power if the following equations are satisfied:

$$q^2 > 0.5;$$

$$R^2 > 0.6;$$

$$\frac{(R^2 - R_o^2)}{R^2} < 0.1 \text{ and } 0.85 \le k \le 1.15$$

where q2 is 'the leave one out' correlation coefficient for the training set, R2 is correlation coefficient between the observed and predicted toxicities for the test set, Ro2 is correlation coefficient between the observed and predicted toxicities for the test set with the Y-intercept set to zero (where the regression line is given by Y=kX). The prediction accuracy was evaluated in terms

of equations (4) and (5). In addition, the accuracy will be evaluated in terms of the RMSE (root mean square error), and the MAE (mean absolute error) for the test set. It has been demonstrated that q2 (the leave one out correlation coefficient for the training set) is not correlated with R2 for the test set.

2.7 Applicability domains

A concept of the applicability domain (AD) was created and used to avoid such an incorrect extrapolation of activity predictions in T.E.S.T. According to Ruiz et al. (2012), the QSAR model can predict the potential toxicity of any chemical but the predictive confidence may vary. Generally each model is processed using a training set of chemicals, which cover only a small fraction of the entire chemical world and it was observed that its prediction capability is restricted to its AD, called as its descriptor space. As a result of this, only a certain fraction of chemicals of an external data set can be reasonably predicted. So it is promising to determine the chemical of interest falls within or outside the AD of a particular model. In context, varying degrees of uncertainties could be validated with such a prediction. For model ADs, features and limitations need to be understood thoroughly for the appropriate interpretation of predictive results.

3. RESULTS & DISCUSSION

The toxicity data of ((3-(2-chloroquinolin-3-yl)oxiran-2-yl)(phenyl) methanone and its derivatives (1-11) is computed by four methods, namely, consensus method, hierarchical clustering, the Food and Drug Administration (FDA) method and nearest neighbor method and are given in Table 4. For compound no. 1 to 4, 7, 10 and 11 hierarchical clustering method does not provide the LD_{50} values; however, other methods have successfully provided the toxicity estimation of the same.

Based on these findings the order of toxicity according to consensus method was found in following manner, wherein the compound no. 4 is found to be the most toxic and compound no. 1 is found to be the safest among all and the lead compound.

4 >5>11>10>7>6>8>9>3>2>1

Hierarchical method is not found reliable for the given series as only four compounds namely 5, 6, 8 and 9 values only available. However, among these the order of toxicity was found in following manner, wherein the compound no. 6 is found to be the most toxic and compound no. 8 is found to be the safest among all and the lead compound.

6>5>9 >8

The order of toxicity according to FDA method was found to be in the following manner, wherein the compound no. 7 is found to be the most toxic and compound no. 9 is found to be the safest among all and the lead compound.

7>5>4>3>10>11>2>1>6 >8>9

According to nearest neighbor method, the order of toxicity was found to be in following manner, wherein the compound no. 8 is found to be the most toxic and compound no. 1 is found to be the safest among all and the lead compound.

8>11>4>5>6>10>9>2>7>3>1

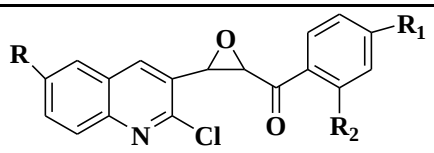
Based on these findings, compound no. 1 and 9 are found to be the lead compounds, subject to suitability of other parameters, while compounds no. 4, 6, 7 and 8 are found to be the most toxic and further study is recommended for these compounds. The prediction value for Mean Absolute Error (MAE) for different compounds studied for their toxicity and the other similar compounds are provided in Fig. 1-4 respectively. Though the mean absolute error value falls within the acceptable corridor for the all seven graphs, it exceeds the acceptable range for compound no. 1, 8 and 9.

Researchers nowadays rely a lot on QSAR models for Toxicity predictions, since they reduce the time consumed, the cost incurred and also eliminate the problem of animal testing. These models are highly reliable and are used widely. Rat oral QSAR LD_{50} models were made for ((3-(2-chloroquinolin-3-yl)oxiran-2-yl)(phenyl) methadone and its derivatives (1-11). The toxicity levels were found in the range of 246.05 mg/kg to 4956.14 mg/kg having the highest similarity coefficient of 0.88. These LD_{50} values of the compounds are assigned to various toxic levels according to standard toxicity scale. The reliability of the data is solely based on user confidence. The current findings serve as a base for researchers to further investigate the properties of these new compounds

4. CONCLUSION

The study concludes that some of ((3-(2-chloroquinolin-3-yl)oxiran-2-yl)(phenyl) methanone and its derivatives (1-11) are found to be toxic, some are moderately toxic and while few others are relatively less toxic. However, further studies on the same are recommended. The study also helped to found the lead compound among the test compounds based on their relative toxicity level.

Table 1. General structure for compounds 1–11



Compound	R	R_1	R_2
1	H	H	H
2	CH_3	H	H
3	OCH_3	H	H
4	H	Cl	Cl
5	CH_3	Cl	Cl
6	OCH_3	Cl	Cl
7	H	CH_3	H
8	CH_3	CH_3	H
9	OCH_3	CH_3	H
10	Cl	Cl	Cl
11	Cl	CH_3	Cl

Table 2. IUPAC names of test compounds 1–11.

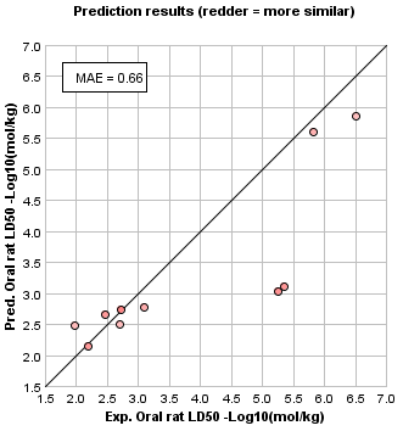
Compound	Chemical Name
1	(3-(2-Chloroquinolin-3-yl)oxiran-2-yl)(phenyl)methanone
2	(3-(2-Chloro-6-methylquinolin-3-yl)oxiran-2-yl)(phenyl)methanone
3	(3-(2-Chloro-6-methoxyquinolin-3-yl)oxiran-2-yl)(phenyl)methanone
4	(3-(2-Chloroquinolin-3-yl)oxirane-2-yl)(2,4-dichlorophenyl)methanone
5	(3-(2-Chloro-6-methylquinolin-3-yl)oxirane-2-yl)(2,4-dichlorophenyl)methanone
6	(3-(2-Chloro-6-methoxyquinolin-3-yl)oxirane-2-yl)(2,4-dichlorophenyl)methanone
7	(3-(2-Chloroquinolin-3-yl)oxirane-2-yl)(p-tolyl)methanone
8	(3-(2-Chloro-6-methylquinolin-3-yl)oxirane-2-yl)(p-tolyl)methanone
9	(3-(2-Chloro-6-methoxyquinolin-3-yl)oxirane-2-yl)(p-tolyl)methanone
10	(3-(2,6-Dichloroquinolin-3-yl)(2,4-dichlorophenyl)oxirane-2-yl)methanone
11	(3-(2,6-Dichloroquinolin-3-yl)oxirane-2-yl)(p-tolyl)methanone

Table 3. Structures of test compounds 1–11 by SMILES notation & Molecular formula.

Compound No	SMILES Notation	Molecular Formula
1	O=C(C2OC2C4=C(Cl)N=C3C=CC=CC3=C4)C1=CC=CC=C1	$C_{18}H_{12}ClNO_2$
2	O=C(C2OC2C4=C(Cl)N=C3C=CC(C)=CC3=C4)C1=CC=CC=C1	$C_{19}H_{14}ClNO_2$
3	O=C(C2OC2C4=C(Cl)N=C3C=CC(OC)=CC3=C4)C1=CC=CC=C1	$C_{19}H_{14}ClNO_3$
4	O=C(C2OC2C4=C(Cl)N=C3C=CC=CC3=C4)C1=CC=C(Cl)C=C1Cl	$C_{18}H_{10}Cl_3NO_2$
5	O=C(C2OC2C4=C(Cl)N=C3C=CC(C)=CC3=C4)C1=CC=C(Cl)C=C1Cl	$C_{19}H_{12}Cl_3NO_2$
6	O=C(C2OC2C4=C(Cl)N=C3C=CC(OC)=CC3=C4)C1=CC=C(Cl)C=C1Cl	$C_{19}H_{12}Cl_3NO_3$
7	O=C(C2OC2C4=C(Cl)N=C3C=CC=CC3=C4)C1=CC=C(C)C=C1	$C_{19}H_{14}ClNO_2$
8	O=C(C2OC2C4=C(Cl)N=C3C=CC(C)=CC3=C4)C1=CC=C(C)C=C1	$C_{20}H_{16}ClNO_2$
9	O=C(C2OC2C4=C(Cl)N=C3C=CC(OC)=CC3=C4)C1=CC=C(C)C=C1	$C_{20}H_{16}ClNO_3$
10	O=C(C4=CC=C(Cl)C=C4Cl)C1OC1C3=C(Cl)N=C2C=CC(Cl)=CC2=C3	$C_{18}H_9Cl_4NO_2$
11	O=C(C2OC2C4=C(Cl)N=C3C=CC(Cl)=CC3=C4)C1=CC=C(C)C=C1	$C_{18}H_{13}Cl_2NO_2$

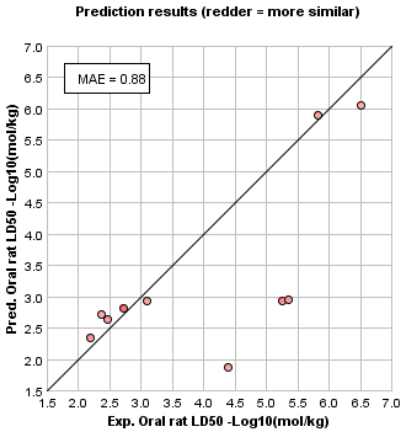
Table 4. Batch predictions of oral rat LD_{50} values (mg/kg) by four different methods.

Compound No	Conseusus method	Hierarchical method	FDA method	Nearest neighbor method
1	2971.37	N/A	3033.12	2910.89
2	2702.85	N/A	2644.73	2762.25
3	2276.49	N/A	1787.80	2898.75
4	655.29	N/A	1477.90	290.55
5	731.40	1199.28	1082.74	301.32
6	1009.28	1046.79	3131.86	313.60
7	970.49	N/A	340.98	2762.25
8	1073.89	1346.43	3738.36	246.05
9	1698.44	1137.34	4956.14	869.19
10	782.37	N/A	1931.02	316.98
11	739.24	N/A	1987.99	274.89



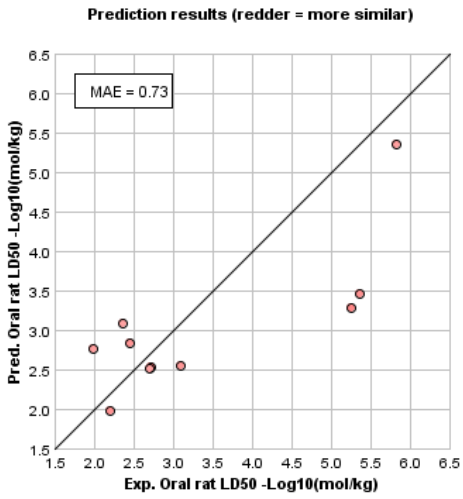
Test set chemicals	MAE *
Entire set	0.43
Similarity coefficient ≥ 0.5	0.66

Fig 1. Prediction of MAE for the test chemical ($C_{18}H_{12}ClNO_2$) and the most similar chemicals.



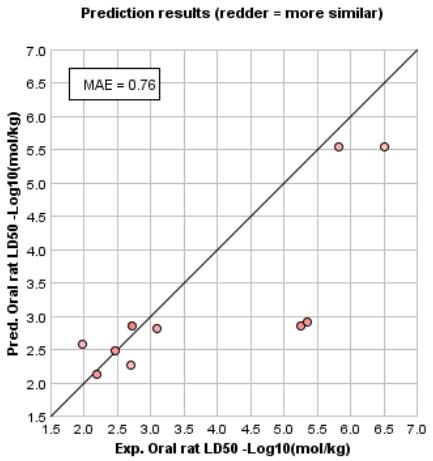
Test set chemicals	MAE*
Entire set	0.46
Similarity coefficient ≥ 0.5	0.88

Fig 2. Prediction of MAE for the test chemical ($C_{20}H_{16}ClNO_2$) and the most similar chemicals.



Test set chemi als	MAE*
Entire set	0. 7
Similarity coefficient ≥ 0.5	0.73

Fig 3. Prediction of MAE for the test chemical ($C_{20}H_{16}ClNO_3$) and the most similar chemicals.



Test set chemicals	MAE*
Entire set	0.48
Similarity coefficient ≥ 0.5	0.7

Fig 4. Prediction of MAE for the test chemical ($C_{18}H_{12}ClNO_2$) and the most similar chemicals

REFERENCES

1. Abdel-Wahab BF, Khidre RE, Farahat AA and El-Ahl AA. 2-Chloroquinoline-3-carbaldehyde: Synthesis, reactions and applications. Arkivoc, 2012; 211-276.
2. Preveena N, Nagendrappa Gopalpur, Suresha Kumara TH, Tiwari Amit Kumar, Chaithanya MS, Nagananda GS, Sujan Ganapathy PS, Tayur N Row Guru, Hosamani Amar A, Chethana PR. Synthesis of (3-(2-Chloroquinolin-3-yl)oxiran-2-yl)(phenyl)methanone derivatives and in vitro and in silico study of their various biological activities. International Journal of Pharmaceutical Science Invention. 2015; 4:6: 53-76.
3. Choplin F. Comprehensive medicinal chemistry. Elsevier Pergamon, Oxford, 2005; 4: 33-57.
4. Valentina P, Lango K and Engels M. Rationalization of physico chemical characters of 2-Phenyl-3-hydroxy-4(1H)-quinolinone-7-carboxylic acid analogs as topoisomerase inhibitors: A QSAR approach. Indian J. Pharm. Educ. Res. 2009; 43:3: 284-289.
5. T. E. S. T. Software Version 4.1 [A Program to Estimate Toxicity from Molecular Structure, developed for United States Environmental Protection Agency, Cincinatti, OH, USA]
6. Yap CW, Xue Y, Li ZR and Chen YZ. Application of support vector machines to in silico prediction of cytochrome $P4_{50}$ enzyme substrates and inhibitors. Curr. Topics Med. Chem. 2006; 6:15: 1593-1607.
7. Guido RV, Oliva G and Andricopulo AD. Virtual screening and its integration with modern drug design technologies. Curr. Med. Chem. 2008; 15:1: 37-46.
8. Schwaighofer A, Schroeter T, Mika S and Blanchard G. How wrong can we get? A review of machine learning approaches and error bars. Comb. Chem. High Throughput Screen. 2009; 12:5: 453-468.
9. Valerio LG Jr. In silico toxicology for the pharmaceutical sciences. Toxicol. Appl. Pharmacol. 2009; 241: 356-370.
10. Kovalishyn V, Kopernyk I, Chumachenko S, Shablykin O, Kondratyuk K, Pil'o S, Prokopenko V, Brovarets V and Metelytsia L. QSAR studies, design, synthesis and antimicrobial evaluation of azole derivatives. Comput. Biol. Bioinfor. 2014; 2:2: 25-32.
11. Worth AP, Bassan A, DeBruijn J, Gallegos Saliner A, Netzeva T, Patlewicz G, Pavan M, Tsakovska I and Eisenreich S. The role of the European Chemicals Bureau in promoting the regulatory use of (Q)SAR methods. SAR and QSAR Env. Res. 2007;18:1-2: 111-125.
12. Lilienblum W, Dekant W, Foth H, Gebel T, Hengstler J, Kahl R, Kramer PJ, Schweinfurth H and Wollin KM. Alternative methods to safety studies in experimental animals: role in the risk assessment of chemicals under the new European Chemicals Legislation (REACH). Arch. Toxicol., 2008; 82:4: 211-236.
13. TOPKAT Software User Guide Version 6.2 [A Program developed by Accelrys: San Diego, CA, USA]
14. Dragon Software User Guide 7.0 [A software for the calculation of 4885 molecular descriptors, developed by TALETE srl, Milano, Italy]

15. ADMET Software User Guide, Version 5.5 [A software for Simulation developed by Simulation Plus Inc, SP: Lancaster, CA, USA]
16. Molecular Design Suit (MDS) TM 3.5 [A software developed by V-life Technologies, India]
17. Pallas 3.1.1.2 [An ADME Tox software developed by Computer Drug International Inc, U.S.A.]
18. Yap CW. PaDEL-Descriptor: An open source software to calculate molecular descriptors and fingerprints. J. Comp. Chem. 2011; 32:7: 1466-1474.
19. Martin TM, Harten P, Venkatapathy R, Das, S and Young DM. A hierarchical clustering methodology for the estimation of toxicity. Toxicol. Mech. Methods, 2008; 18:251-266.
20. Zhu H, Martin TM, Ye L, Sedykh A, Young DM and Tropsha A, Quantitative structure-activity relationship modeling of rat acute toxicity by oral exposure. Chem. Res. Toxicol. 2009; 22: 1913-1921.
21. Romesburg HC. Cluster Analysis for Researchers LULU Press, USA (1984).
22. Ruiz P, Begluitti G, Tincher T, Wheeler J and Mumtaz M. Prediction of acute mammalian toxicity using QSAR methods: A case study of sulfur mustard and its breakdown products. Molecules, 2012; 17: 8982-9001.
23. Contrera JF, Matthews EJ and Daniel Benz R. Predicting the carcinogenic potential of pharmaceuticals in rodents using molecular structural similarity and E-state indices. Regul. Toxicol. Pharmacol. 2003; 38: 243-259.
24. Tetko IV, Sushko I, Pandey AK, Zhu H, Tropsha A, Papa E, Oberg T, Todeschini R, Fourches D and Varnek A. Critical assessment of QSAR models of environmental toxicity against Tetrahymena pyriformis: Focusing on applicability domain and over fitting by variable selection. J. Chem. Inf. Model., 2008; 48:1733-1746.
25. Gramatica P and Pilutti P. Evaluation of different statistical approaches for the validation of quantitative structure-activity relationships. Ispra, Italy, The European Commission - Joint Research Centre, Institute for Health and Consumer Protection – ECVAM (2004).
26. Golbraikh A and ropsha A. Beware of q2!. J. Mol. Graph. Model. 2002; 20: 269-276.