	QMRF identifier (JRC Inventory):**Q13-414-0035**
	QMRF Title:**QSAR for rat chronic LOAEL**
	Printing Date:**Dec 11, 2019**

1.QSAR identifier

1.1.QSAR identifier (title):

QSAR for rat chronic LOAEL

1.2.Other related models:

Same endpoint and dataset as presented model in reference [sect.9.2;ref.1].

1.3.Software coding the model:

QSARModel 4.0.4

Molcode Ltd., Turu 2, Tartu, 51014, Estonia

http://www.molcode.com

2.General information

2.1.Date of QMRF:

18.10.2010

2.2.QMRF author(s) and contact details:

Molcode model development team Molcode Ltd. Turu 2, Tartu, 51014, Estonia

models@molcode.com http://www.molcode.com

2.3.Date of QMRF update(s):

2.4.QMRF update(s):

2.5.Model developer(s) and contact details:

Molcode model development team Molcode Ltd. Turu 2, Tartu, 51014, Estonia

models@molcode.com http://www.molcode.com

2.6.Date of model development and/or publication:

14.10.2010

2.7.Reference(s) to main scientific papers and/or software package:

[1]Karelson M, Dobchev D, Tamm T, Tulp I, Jänes J, Tämm K, Lomaka A, Savchenko D & Karelson G (2008). Correlation of blood-brain penetration and human serum albumin binding with theoretical descriptors. ARKIVOC 16, 38-60.

[2]Karelson M, Karelson G, Tamm T, Tulp I, Jänes J, Tämm K, Lomaka A, Savchenko D & Dobchev D (2009). QSAR study of pharmacological permeabilities. ARKIVOC 2, 218–238.

2.8.Availability of information about the model:

Model is proprietary, but the training and test sets are available.

2.9.Availability of another QMRF for exactly the same model:

None to date

3.Defining the endpoint - OECD Principle 1

3.1.Species:

rat

3.2.Endpoint:

4.Human Health Effects 4.14.Repeated dose toxicity

3.3.Comment on endpoint:

Rat chronic LOAEL

3.4.Endpoint units:

µmol/kg

3.5.Dependent variable:

log(LOAEL)

3.6.Experimental protocol:

3.7.Endpoint data quality and variability:

Statistics:

max value: 4.31

min value: -1.10

standard deviation: 1.16

skewness: -0.62

4.Defining the algorithm - OECD Principle 2

4.1.Type of model:

QSAR

4.2.Explicit algorithm:

Multilinear regression QSAR

log(LOAEL) = 3.244

-4.240E-003*Final heat of formation (AM1)

-0.228*Kier&Hall index (order 3)

+3.778E-002*Lowest resonance energy (AM1) for C - N bonds

-1.148*Max net atomic charge (AM1) for N atoms

-0.405*Square root of Charged (Zefirov) Surface Area of Cl atoms

4.3.Descriptors in the model:

[1]Final heat of formation (AM1) kcal semiempirically calculated heat of formation

[2]Kier&Hall index (order 3) unitless third order Kier and Hall valence connectivity index

[3]Lowest resonance energy (AM1) for C - N bonds [eV] lowest resonance energy between bonded C and N atoms

[4]Max net atomic charge (AM1) for N atoms [au] maximum (highest) atomic partial charge over nitrogens in molecule

[5]Square root of Charged (Zefirov) Surface Area of Cl atoms [Å] square root of charged curface area of Cl atoms based on Zefirovs charge distribution

4.4.Descriptor selection:

Initial pool of ~1000 descriptors. Stepwise descriptor selection based on a set of statistical selection rules:

One-parameter equations: Fisher criterion and R2 over threshold, variance and t-test value over threshold, intercorrelation with another descriptor not over threshold)

Two-parameter equations: intercorrelation coefficient below threshold, significant correlation with endpoint, in terms of correlation coefficient and t-test)

Stepwise trial of additional descriptors not significantly correlated to any already in the model.

4.5.Algorithm and descriptor generation:

1D, 2D, and 3D theoretical calculations. Quantum chemical descriptors derived from AM1 calculation. Model developed by using multilinear regression.

4.6.Software name and version for descriptor generation:

QSARModel 4.0.4

QSAR/QSPR package that will compute chemically meaningful descriptors and includes statistical tools for regression modelling

Molcode Ltd, Turu 2, Tartu, 51014, Estonia

http://www.molcode.com

4.7.Chemicals/Descriptors ratio:

15.2 (76 chemicals / 5 descriptors)

5.Defining the applicability domain - OECD Principle 3

5.1.Description of the applicability domain of the model:

Applicability domain based on training set:

a) by chemical identity: structurally heterogeneous chemical structures but not phosphorus containing compounds

b) by descriptor value range: The model is suitable for compounds that have the descriptors in the following range:

Final heat of formation (AM1): -283 - 111

Kier&Hall index (order 3): 0 - 11.5

Lowest resonance energy (AM1) for C - N bonds: -33.4 - 0

Max net atomic charge (AM1) for N atoms: -0.859 - 0.626

Square root of Charged (Zefirov) Surface Area of Cl atoms: 0 - 4.59

5.2.Method used to assess the applicability domain:

Compounds should not conatin phosphorus.

Range of descriptor values in training set with ±30% confidence.

Descriptor values must fall between maximal and minimal descriptor values of training set ±30%.

5.3.Software name and version for applicability domain assessment:

QSARModel 4.0.4

QSAR/QSPR package that will compute chemically meaningful descriptors and includes statistical tools for regression modeling

Molcode Ltd, Turu 2, Tartu, 51014, Estonia

http://www.molcode.com

5.4.Limits of applicability:

See 5.1

6.Internal validation - OECD Principle 4

6.1.Availability of the training set:

Yes

6.2.Available information for the training set:

CAS RN: Yes

Chemical Name: Yes

Smiles: No

Formula: Yes

INChI: No

MOL file: Yes

6.3.Data for each descriptor variable for the training set:

All

6.4.Data for the dependent variable for the training set:

All

6.5.Other information about the training set:

76 data points: 4 negative values; 72 positive values

6.6.Pre-processing of data before modelling:

6.7.Statistics for goodness-of-fit:

R^2= 0.790 (Correlation coefficient)

s_2= 0.554 (Standard error of the estimate)

F = 52.8 (Fisher function)

6.8.Robustness - Statistics obtained by leave-one-out cross-validation:

R^2cv = 0.755 (Cross-validated correlation coefficient)

6.9.Robustness - Statistics obtained by leave-many-out cross-validation:

R^2_{CVMO}= 0.744

6.10.Robustness - Statistics obtained by Y-scrambling:

6.11.Robustness - Statistics obtained by bootstrap:

6.12.Robustness - Statistics obtained by other methods:

ABC analysis (2:1 training : prediction) on sorted (in increasing order of endpoint value) data divided into 3 subsets (A;B;C). Training set formed with 2/3 of the compounds (set A+B, A+C, B+C) and validation set consisted of 1/3 of the compounds (C, B, A).

average R^2(fitting) = 0.798

average R^2(prediction) = 0.765

7.External validation - OECD Principle 4

7.1.Availability of the external validation set:

Yes

7.2.Available information for the external validation set:

CAS RN: Yes

Chemical Name: Yes

Smiles: No

Formula: Yes

INChI: No

MOL file: Yes

7.3.Data for each descriptor variable for the external validation set:

All

7.4.Data for the dependent variable for the external validation set:

All

7.5.Other information about the external validation set:

18 data points: 1 negative values; 17 positive values

7.6.Experimental design of test set:

From sorted data each 5th was subjected to the test set

7.7.Predictivity - Statistics obtained by external validation:

R^2= 0.725 (Correlation coefficient)

7.8.Predictivity - Assessment of the external validation set:

Descriptor value range (all in range of applicability domain):

Final heat of formation (AM1): -230 - 27.0

Kier&Hall index (order 3): 0.254 - 0.61

Lowest resonance energy (AM1) for C - N bonds: -19.4 - -17.4

Max net atomic charge (AM1) for N atoms: -0.386 - 0.566

Square root of Charged (Zefirov) Surface Area of Cl atoms: 0.820 - 3.92

7.9.Comments on the external validation of the model:

The validation correlation coefficient (R^2) for the test set is close to that of the training set.

8.Providing a mechanistic interpretation - OECD Principle 5

8.1.Mechanistic basis of the model:

"Final heat of formation (AM1)" shows stability of the compounds. Less stable molecules are more reactive and thus more toxic. "Kier&Hall index of (order 3)" is a valence connectivity index which accounts for the presence of hetero atoms and for the hybridization of atoms in the molecule and reflects the size and shape of the compound. These two descriptors are whole molecular descriptors and they probably define toxicity baseline. The remaining descriptors, "Lowest resonance energy (AM1) for C - N bonds", "Max net atomic charge (AM1) for N atoms" and "Square root of Charged (Zefirov) Surface Area of Cl atoms" are functional group dependent descriptors and they count the presence of particular groups and their electrostatic interaction potentials.

8.2.A priori or a posteriori mechanistic interpretation:

A posteriori mechanistic interpretation, consistent with published scientific interpretations of experiments.

8.3.Other information about the mechanistic interpretation:

Interpretation in general agreement with literature [sect.9.2; ref.1].

9.Miscellaneous information

9.1.Comments:

The modeling of toxicological properties is an extremely important problem. No empirical toxicological data are available for most chemicals, and the growing new ones must be evaluated or, at least estimated. Thus, reliable methods to predict environmental toxicity are required. Furthermore, the presence of toxic substances in high trophic levels can affect humans. Particular interest in the estimation of chronic lowest observed adverse effect level (LOAEL) has been raised recently due to its environmental implications. LOAEL was defined by the IUPAC as the lowest concentration or amount of a substance, found by experiment or observation, which causes an adverse alteration of morphology, functional capacity, growth, development, or life span of a

target organism distinguishable from normal (control) organisms of the same species and strain under defined conditions of exposure.

9.2.Bibliography:

García-Domenech R, de Julián-Ortiz JV & Besalú E (2006). True prediction of lowest observed adverse effect levels. Molecular Diversity 10,159-168. http://dx.doi.org/10.1007/s11030-005-9007-z

9.3.Supporting information:

Training set(s)Test set(s)Supporting information

10.Summary (JRC QSAR Model Database)

10.1.QMRF number:

Q13-414-0035

10.2.Publication date:

2013-06-27

10.3.Keywords:

Molcode;rat;chronic LOAEL;

10.4.Comments:

former Q8-10-30-288