	QMRF identifier (JRC Inventory): **Q13-31-0037**
	QMRF Title: **Catalogic QSAR for aquatic toxicity**
	Printing Date: **Dec 11, 2019**

1.QSAR identifier

1.1.QSAR identifier (title):

Catalogic QSAR for aquatic toxicity

1.2.Other related models:

1.3.Software coding the model:

Catalogic v5.10.7

Professor Ovanes Mekenyan

http://oasis-lmc.org/

2.General information

2.1.Date of QMRF:

23 July 2009

2.2.QMRF author(s) and contact details:

Grace Patlewicz DuPont Haskell Global Centers for Health & Environmental Science 1070 Elkton Rd, Newark, DE 19711, USA grace.y.tier@usa.dupont.com

2.3.Date of QMRF update(s):

2.4.QMRF update(s):

2.5.Model developer(s) and contact details:

Professor Ovanes Mekenyan Laboratory of Mathematical Chemistry, University As Zlatarov, Bourgas, Bulgaria omekenya@btu.bg http://oasis-lmc.org/

2.6.Date of model development and/or publication:

November 2005 is the release date referenced in the software.

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

Catalogic v5.10.7 http://oasis-lmc.org/

2.8.Availability of information about the model:

The model is proprietary since it is housed within the Catalogic expert system. Structural rules are used to determine which model for acute toxicity is relevant. These rules characterise the training sets and whether they comprise amines/phenols, unspecific reactivity or basesurface narcotics. The training set for the basesurface narcotics model is taken from published sources of data as far as possible (see literature references in section 9).

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

3.Defining the endpoint - OECD Principle 1

3.1.Species:

Daphnia Magnia

3.2.Endpoint:

3.Ecotoxic effects 3.1.Short-term toxicity to Daphnia (immobilisation)

3.3.Comment on endpoint:

EC50, the median effective concentration. Acute toxicity is expressed in this test as the median effective concentration (EC50) for immobilisation.This is the concentration, in terms of initial values, which immobilises 50% of the Daphnia in a test batch within a continuous period of exposure (in this case 48 hours) which must be stated.

3.4.Endpoint units:

mg/L

3.5.Dependent variable:

Log(1/EC50) was the dependent variable used. An EC50 value reported in mg/L was first divided by 1000 to convert to a g/L unit. This was then divided by the molecular weight to achieve a mol/L. Finally the EC50 mol/L was inverted and the log to the base 10 taken. In the dependent variable reported - the EC50 is in mol/L

3.6.Experimental protocol:

This is described in more detail in EU Test Method C.2 or OECD Test Guideline 202

3.7.Endpoint data quality and variability:

Reasonably high quality data (in terms of the data arising from the same protocol). In the majority of cases the data will have been taken from the EPA Acquire database which is well established and carefully curated. However which specific data points come from which references can not be determined from Catalogic. For that reason, there is expected to be some variability in the data as the training set would have been likely been compiled secondhand from training sets previously collected by other researchers such as Zhao et al (1998) and Ohe et al (2005).

4.Defining the algorithm - OECD Principle 2

4.1.Type of model:

QSAR

4.2.Explicit algorithm:

Multilinear regression QSAR

Regression equation which relates LogBCFtox and E_{LUMO} to the Log(1/EC50) value

4.3.Descriptors in the model:

[1]LogBCFtox none provided Used to model partition

[2]ELUMO eV Energy of the Lowest Unoccupied Molecular Orbital

4.4.Descriptor selection:

These descriptors were chosen to account for the hydrophobicity and electrophilicity.

4.5.Algorithm and descriptor generation:

Expert system - Catalogic is an hybrid expert system comprising structure-activity, structure-metabolism rules together with 3D QSARs. The basesurface narcotics EC50 Daphnia model is one of the 3D QSARs within Catalogic. As a first step, within Catalogic, structure-activity rules determine the mode of action - reactive, basesurface narcotic, ester, phenols/anilines. This is to ensure that only substances which

are presumed to act by a certain mode of action are progressed to the specific 3D QSAR model for a refined estimate of EC50. A multiple linear regression is used to compute the two descriptors (LogBCFtox and ELUMO) to provide a quantitative output for the EC50 value.

4.6.Software name and version for descriptor generation:

Catalogic v5.10.7

4.7.Chemicals/Descriptors ratio:

15.0 (30 chemicals / 2 descriptors)

5.Defining the applicability domain - OECD Principle 3

5.1.Description of the applicability domain of the model:

A three-step applicability domain assessment was carried out to account for the structural fragments and the parameter ranges of the training set materials

5.2.Method used to assess the applicability domain:

Parameter ranges was the first layer defined by the range of preliminary 3D parameters included in the model. The LogBCFtox and E_{LUMO}values were computed for all the 32 training set chemicals and their respective conformers. A parametric coverage domain was then extracted to demonstrate how much an input chemical belonged to the parameter(s) probability distribution. A third structural domain was extracted which utilises atom centred fragments as a means of describing the structures within the training set.

5.3.Software name and version for applicability domain assessment:

Domain Manager v1.01
Professor Ovanes Mekenyan
http://oasis-lmc.org/

5.4.Limits of applicability:

This three-step process is conducted within the Domain Manager software. The parameter ranges are as follows: E_{LUMO}min 3.39 and E_{LUMO}max 3.58. The logBCFtox ranges were min -1.773 and max 3.8. The threshold for the parameter coverage was set to the default of 0.01. The default setting for the structural domain were used within the Domain manager software.

6.Internal validation - OECD Principle 4

6.1.Availability of the training set:

No

6.2.Available information for the training set:

CAS RN: Yes
Chemical Name: Yes
Smiles: Yes
Formula: No
INChI: No
MOL file: No

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:

30 data points. The references for the data are provided in section 9.

6.6.Pre-processing of data before modelling:

The endpoint data was converted into its log(1/EC50) equivalent for the purposes of modelling

6.7.Statistics for goodness-of-fit:

Minimal statistics are provided within the software are: R^2= 0.895

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

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

6.10.Robustness - Statistics obtained by Y-scrambling:

6.11.Robustness - Statistics obtained by bootstrap:

6.12.Robustness - Statistics obtained by other methods:

7.External validation - OECD Principle 4

7.1.Availability of the external validation set:

No

7.2.Available information for the external validation set:

CAS RN: No

Chemical Name: No

Smiles: No

Formula: No

INChI: No

MOL file: No

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

No

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

No

7.5.Other information about the external validation set:

7.6.Experimental design of test set:

7.7.Predictivity - Statistics obtained by external validation:

7.8.Predictivity - Assessment of the external validation set:

7.9.Comments on the external validation of the model:

8.Providing a mechanistic interpretation - OECD Principle 5

8.1.Mechanistic basis of the model:

The mechanistic basis is that neutral organic substances (i.e. not possessing any overt electrophilic centres) disrupt membranes by a non-specific narcosis mechanism. Narcosis is a non-specific reversible state of arrested activity of protoplasmic structures caused by a wide variety of organic chemicals. Within Catalogic, structural alerts and rules are in place to filter substances on the basis of their presumed mode of action. Those which fit the general narcosis equation can be

modelled by using descriptors which encompass the partitioning and electronegativity (modelled by logBCFtox and E_{LUMO}). These in turn can be related to the EC50 in Daphnia Magnia

8.2.A priori or a posteriori mechanistic interpretation:

A priori mechanistic interpretation.

8.3.Other information about the mechanistic interpretation:

9.Miscellaneous information

9.1.Comments:

This model is transparent within the Catalogic software. The training set is available and the domain assessment is automatically carried out to establish whether a chemical is likely to be within scope or not. The model has a strong mechanistic basis that is mirrored in other baseline QSARs that have been published.

9.2.Bibliography:

[1]Zhao YH, Ji GD, Cronin MTD & Dearden JC (1998). QSAR study of the toxicity of benzoic acids to Vibrio fischeri, Daphnia magna and carp. Science of the Total Environment 216, 205-215.

[2]Schuurmann G (1998). Ecotoxic Modes of Action of Chemical Substances. In Ecotoxicology: Ecological Fundamentals, Chemical Exposure, and Biological Effects (Schuurmann G, Markert B, eds). Environmental Science and Technology: A Wiley-Interscience Series of Texts and Monographs. Wiley-Interscience.

[3]Robert D & Carbo-Dorca R (1999). Aromatic compounds aquatic toxicity QSAR using molecular quantum similarity measures. SAR and QSAR in Environmental Research 10, 401-422.

[4]Zhao YH, Cronin MTD & Dearden JC (1998). Quantitative Structure-Activity Relationships of Chemicals Acting by Non-polar Narcosis - Theoretical Considerations. Quantitative Structure-Activity Relationship 17, 131-138.

[5]Ramos EU (1998). Aquatic Toxicity of Polar Narcotic Pollutants. Mechanism, Modeling and Environmental Effect Assessment. Ph.D Thesis, Utrecht University, NL.

[6]Furusjo E, Andersson M, Rahmberg M & Svenson A (2003). Estimating environmentally important properties of chemicals from the chemical structure. Swedish Environmental Research Institute, Report B1517

[7]Ohe PC, Kuhne R, Ebert R-U, Altenburger R, Liess M & Schuurmann G (2005). Structural Alerts-A New Classification Model to Discriminate Excess Toxicity from Narcotic Effect Levels of Organic Compounds in the Acute Daphnid Assay. Chemical Research in Toxicology 18(3), 536-555.

9.3.Supporting information:

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

10.Summary (JRC QSAR Model Database)

10.1.QMRF number:

Q13-31-0037

10.2.Publication date:

2013-06-27

10.3.Keywords:

Catalogic;basesurface narcosis;aquatic toxicity;Daphnia Magna;

10.4.Comments:

former Q8-29-23-53