	QMRF identifier (JRC Inventory): **Q13-410-0065**
	QMRF Title: **Toxtree QSAR 6: mutagenicity of aromatic amines in Salmonella typhimurium TA100**
	Printing Date: **Dec 11, 2019**

1.QSAR identifier

1.1.QSAR identifier (title):

Toxtree QSAR 6: mutagenicity of aromatic amines in Salmonella typhimurium TA100

1.2.Other related models:

1.3.Software coding the model:

Toxtree

Standalone software application downloadable from the Joint Research Centre (JRC) website http://ecb.jrc.ec.europa.eu/qsar/qsar-tools/index.php?c=TOXTREE

2.General information

2.1.Date of QMRF:

June 2010

2.2.QMRF author(s) and contact details:

[1]Romualdo Benigni Experimental and Computational Carcinogenesis Unit, Environment and Health Department, Istituto Superiore di Sanita Viale Regina Elena 299, 00161 Rome, Italy rbenigni@iss.it

[2]Cecilia Bossa Experimental and Computational Carcinogenesis Unit, Environment and Health Department, Istituto Superiore di Sanita Viale Regina Elena 299, 00161 Rome, Italy cecilia.bossa@iss.it

2.3.Date of QMRF update(s):

2.4.QMRF update(s):

2.5.Model developer(s) and contact details:

Romualdo Benigni rbenigni@iss.it

2.6.Date of model development and/or publication:

2007

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

Benigni R, Bossa C, Netzeva T, Rodomonte A & Tsakovska I (2007). Mechanistic QSAR of aromatic amines: new models for discriminating between mutagens and nonmutagens, and validation of models for carcinogens. Environmental and Molecular Mutagenesis 48, 754-771.

2.8.Availability of information about the model:

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

3.Defining the endpoint - OECD Principle 1

3.1.Species:

Salmonella typhimurium TA100 strain

3.2.Endpoint:

4.Human Health Effects 4.10.Mutagenicity

3.3.Comment on endpoint:

3.4.Endpoint units:

Nonmutagen / Mutagen

3.5.Dependent variable:

Mutagenic activity. The yes/no call was assigned based on the analysis of induced revertants. A chemical was considered to be a mutagen if the number of revertant colonies induced was at least twice the control value.

3.6.Experimental protocol:

Training and test set data are retrieved from the literature (different laboratories, with standard protocols employing S. typhimurium TA100 strain, with S9 metabolic activation system). The compilation of the training set data, with literature sources and details on the experimental protocol, is in Debnath et al (1992), section 9.2 ref 2. Details on the test set data are in the reference quoted in point 2.7.

3.7.Endpoint data quality and variability:

Literature sources were critically reviewed by the authors. The yes/no call was established by the authors.

4.Defining the algorithm - OECD Principle 2

4.1.Type of model:

QSAR

4.2.Explicit algorithm:

QSAR

QSAR6

The algorithm was generated through Canonical Discriminant Analysis. W is a mutagenicity probability score. W(mean, Class) is the average value for each of the two classes in the training set. When calculated for an individual chemical, W indicates to which class the chemical should be assigned, based on the Threshold (given below) that separates the two classes. The model has raw coefficients. In this way it can be applied directly to the descriptors, without any preliminary transformation of the values.

w = - 2.85 HOMO + 1.84 LUMO + 0.70 MR2 + 0.69 MR3 + 1.90 MR6 + 3.36 Idist

w(mean,Class1) = 26.09
N1 = 47 (non-mutagens)

w(mean,Class2) = 23.99
N2 = 64 (mutagens)

Threshold = 25.04

4.3.Descriptors in the model:

[1]Highest Occupied Molecular Orbital (HOMO) and the Lowest Unoccupied Molecular orbital (LUMO) eV PM3 molecular orbital energies

[2]MR2 , MR3 , MR6 Molar Refractivity contributions of substituents in position 2, 3, and 6 to the amino group

[3]Idist The indicator variable Idist is a structural parameter coding for the presence (Idist = 1, otherwise Idist = 0) of bulky substituents on the positions 3'-, 4'- and 5'- of 4-aminobiphenyl.

4.4.Descriptor selection:

The descriptors were selected by screening those present in other mechanistically similar QSAR models for the aromatic amines (mutagenic potency; carcinogenic potency and activity).

4.5.Algorithm and descriptor generation:

4.6.Software name and version for descriptor generation:

Daylight

Partial MR calculated by Daylight software freely available on-line in 2001- 2006

http://www.daylight.com/

Sybyl, Tripos

HOMO and LUMO generated with PM3 Hamiltonian

http://www.tripos.com/

4.7.Chemicals/Descriptors ratio:

18.5 = 111 chemicals / 6 descriptors

(11 originally screened)

5.Defining the applicability domain - OECD Principle 3

5.1.Description of the applicability domain of the model:

The applicability domain of the model is defined in terms of:

a) structures to which it applies;

b) range of values of the descriptors in the model.

In the comparison of the test to the training set, mathematical chemical structural similarity was also applied.

5.2.Method used to assess the applicability domain:

Structures were checked by human experts. Ranges of descriptors values were calculated. Mathematical chemical similarity between test and training set was assessed as follows. The training and test sets were combined, and the overall Tanimoto similarity matrix was calculated with the Leadscope software. A Euclidian distance matrix was calculated from the similarity matrix and then subjected to Principal Component (PC) Analysis. Finally, the ranges of PC scores for the training and test sets were compared.

5.3.Software name and version for applicability domain assessment:

Leadscope v. 2.4

Used for the calculation of the Tanimoto chemical similarity index

http://www.leadscope.com/

5.4.Limits of applicability:

The range of structures to which the model applies is very large and it is detailed in Debnath et al.(1992), ref.2 section 9.2.

Ranges of descriptors for the training set:

HOMO between -9.032 and -7.528

LUMO between -1.330 and 0.722

MR2 between 0.090 and 2.980

MR3 between 0.100 and 2.980

MR6 between 0.056 and 1.500

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: Yes

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:

111 data points: 64 positive values; 47 negative values

As training set, the chemicals compiled in the publication by Debnath et al.(1992), ref 1, section 9.2, were selected.

6.6.Pre-processing of data before modelling:

6.7.Statistics for goodness-of-fit:

Squared Canonical Correlation = 0.52. The equation correctly reclassified 87.4 % (Accuracy) of the compounds (Class1, nonmutagens , 95.7 % (Specificity); Class2, mutagens, 81.3 % (Sensitivity)

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:

Yes

7.2.Available information for the external validation set:

CAS RN: Yes

Chemical Name: Yes

Smiles: Yes

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:

External validation set with 47 compounds appended: 18 nonmutagens, 29 mutagens

7.6.Experimental design of test set:

The test set consisted of literature data not included in the publication by Debnath et al. (1992), ref 1, section 9.2.
The applicability domain was taken into account by considering the factors of chemical structure, range of descriptors values, chemical similarity.

Chemicals with different functional groups (except nitro groups), or basic structures different from those detailed in Debnath et al.(1992).ref 1, section 9.2., were excluded from the test set.

Ranges of descriptors for the test set:
HOMO between -10.212 and -7.861
LUMO between -2.028 and 0.709
MR2 between 0 and 1.960
MR3 between 0.100 and 0.800
MR6 between 0.100 and 1.500

7.7.Predictivity - Statistics obtained by external validation:

Accuracy = 0.81; Sensitivity = 0.86; Specificity = 0.72.

7.8.Predictivity - Assessment of the external validation set:

The test set is representative of the applicability domain; no further chemicals were available from the literature.

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:

Whereas the principal factor that affects the relative mutagenicity (potency) of the aminoarenes is their hydrophobicity (logP), followed by electronic factors (HOMO and LUMO) and then steric factors, the model for the yes/no activity shows no influence of logP. This indicates that the potential to be active depends on a threshold of reactivity (HOMO and LUMO), and on the steric hindrance at substitution positions 2, 3, and 6 of the ring, together with steric effects on the distal ring (Idist variable). The parameters in the model are mechanistically linked to requirements for metabolic activation.

8.2.A priori or a posteriori mechanistic interpretation:

The descriptors to be screened were selected based on the knowledge on the mechanisms of action, and on the evidence provided by similar QSARs for the aromatic amines. The descriptors actually accepted in the model

confirm the a priori hypotheses.

8.3.Other information about the mechanistic interpretation:

9.Miscellaneous information

9.1.Comments:

9.2.Bibliography:

[1]Benigni R, Bossa C, Netzeva T, Rodomonte A & Tsakovska I (2007). Mechanistic QSAR of aromatic amines: new models for discriminating between mutagens and nonmutagens, and validation of models for carcinogens. Environmental and Molecular Mutagenesis 48, 754-771.

[2]Debnath AK, Debnath G, Shusterman AJ & Hansch C (1992). A QSAR investigation of the role of hydrophobicity in regulating mutagenicity in the Ames test: 1. Mutagenicity of aromatic and heteroaromatic amines in Salmonella typhimurium TA98 and TA100. Environmental and Molecular Mutagenesis 19,37-52.

[3]Benigni R, Bossa C, Netzeva TI & Worth A P (2007). Collection and evaluation of (Q)SAR models for mutagenicity and carcinogenicity. JRC report EUR 22772 EN. Luxembourg, Office for the Official Publications of the European Communities.
http://ecb.jrc.ec.europa.eu/documents/QSAR/EUR_22772_EN.pdf

[4]Beland FA & Kadlubar FF (1990). Metabolic activation and DNA adducts of aromatic amines and nitroaromatic hydrocarbons. In Chemical carcinogenesis and mutagenesis I. Eds. Cooper CS, Grover PL. pp 265-325. Berlin: Springer Verlag

[5]Benigni R (2005). Structure-activity relationship studies of chemical mutagens and carcinogens: mechanistic investigations and prediction approaches. Chemical Reviews 105, 1767-1800.

[6]Benigni R, Giuliani A, Gruska A & Franke R (2003). QSARs for the mutagenicity and carcinogenicity of the aromatic amines. In Quantitative Structure-Activivity Relationship (QSAR) models of mutagens and carcinogens. Ed. Benigni R. pp 125-144. Boca Raton: CRC Press.

[7]Kadlubar FF & Beland FA (1985) Chemical properties of ultimate carcinogenic metabolites of arylamines and arylamides. In Polycyclic hydrocarbons and carcinogenesis. Ed. Harvey RG. pp 341-370. Washington DC: American Chemical Society.

[8]Kadlubar FF & Hammons GJ (1987). The role of cytochrome P-450 in the metabolism of chemical carcinogens. In Mammalian cytochromes 450, vol II. Ed. Guengerich FP. pp 81-130. Boca Raton: CRC Press.

[9]Woo YT & Lai DY (2001). Aromatic amino and nitro-amino compounds and their halogenated derivatives. In Patty's Toxicology. Ed. Bingham E, Cohrssen B, Powell CH. pp 969-1105. New York: John Wiley and Sons, Inc

9.3.Supporting information:

Toxtree QSAR6_training 110.sdf	http://qsardb.jrc.ec.europa.eu/qmrf/protocol/Q13-410-0065/attachment/A762
Toxtree QSAR6_test 47.sdf	http://qsardb.jrc.ec.europa.eu/qmrf/protocol/Q13-410-0065/attachment/A763

Test set(s)

10.Summary (JRC QSAR Model Database)

Supporting information

10.1.QMRF number:

Q13-410-0065

10.2.Publication date:

2013-07-02

10.3.Keywords:

Toxtree;mutagenicity;Salmonella typhimurium;TA100;S9 metabolic activation;aromatic amine;

10.4.Comments:

former Q19-35-35-290