Technical information on alternative methods

Andrew Worth

European Commission, Joint Research Centre, Systems Toxicology Unit, Italy

**CADASTER workshop on the use of QSAR models in REACH,
Slovenia, 1-2 September 2011**

http://ihcp.jrc.ec.europa.eu/our_labs/computational_toxicology/

Overview

1. Historical perspective on the regulatory use of QSAR
2. Reporting formats for QSAR models and their predictions
3. Ongoing development in reporting formats for MoA-based hazard assessment
4. Framework for assessing model predictions
5. JRC tools of possible interest to CADASTER

Appendix – example of QMRF

Regulatory use of (Q)SAR before REACH

- Extensive use of grouping and read-across, generally without documented rationale
- Occasional use of QSARs in risk assessment, PBT assessment, and classification & labelling (mainly Existing Substances), generally without documented rationale
- Direct replacement of experimental data for physicochemical properties and environmental fate
- Filling of data gaps for ecotoxicological endpoints, usually to supplement experimental data
- Filing of data gaps for human health endpoints very limited, and only as supporting information

Standardised (Q)SAR Reporting Formats

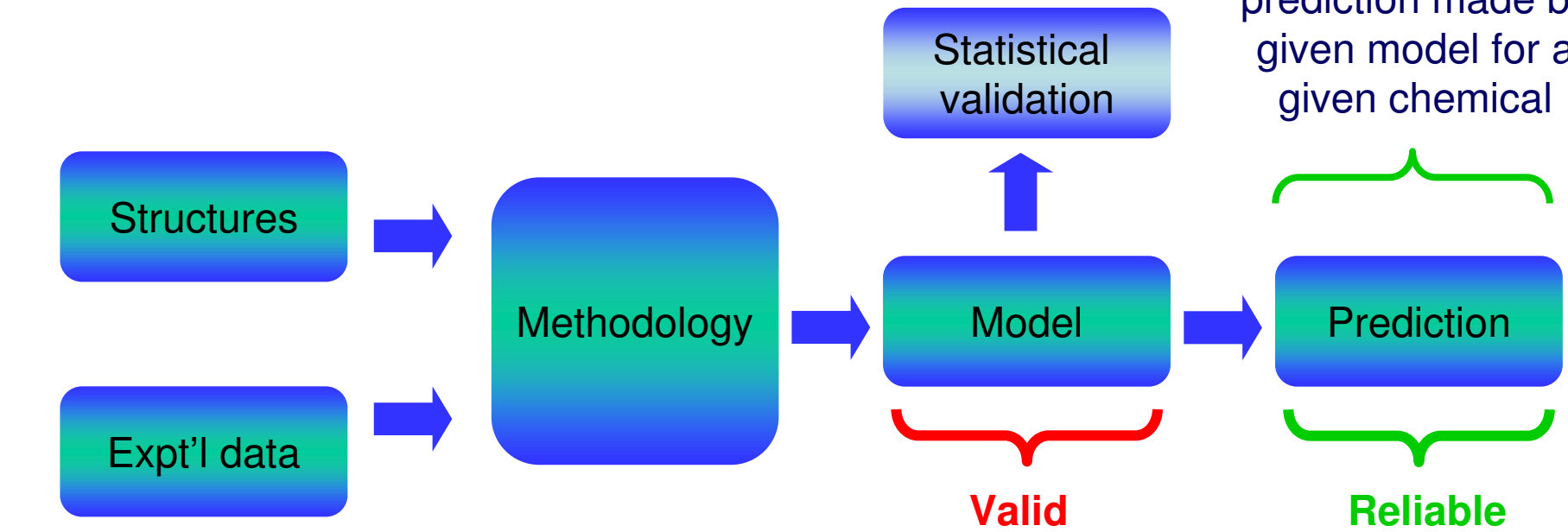
The need for "adequate and reliable" documentation is met by using standardised reporting formats:

QMRF

Robust summary of a (Q)SAR model, which reports key information on the model according to the 5 OECD validation principles.

QPRF

Description and assessment of the prediction made by given model for a given chemical



(Q)SAR Reporting Formats: QMRF

QMRF captures information on fulfilment of OECD validation principles, but no judgement or "validity statement" is included

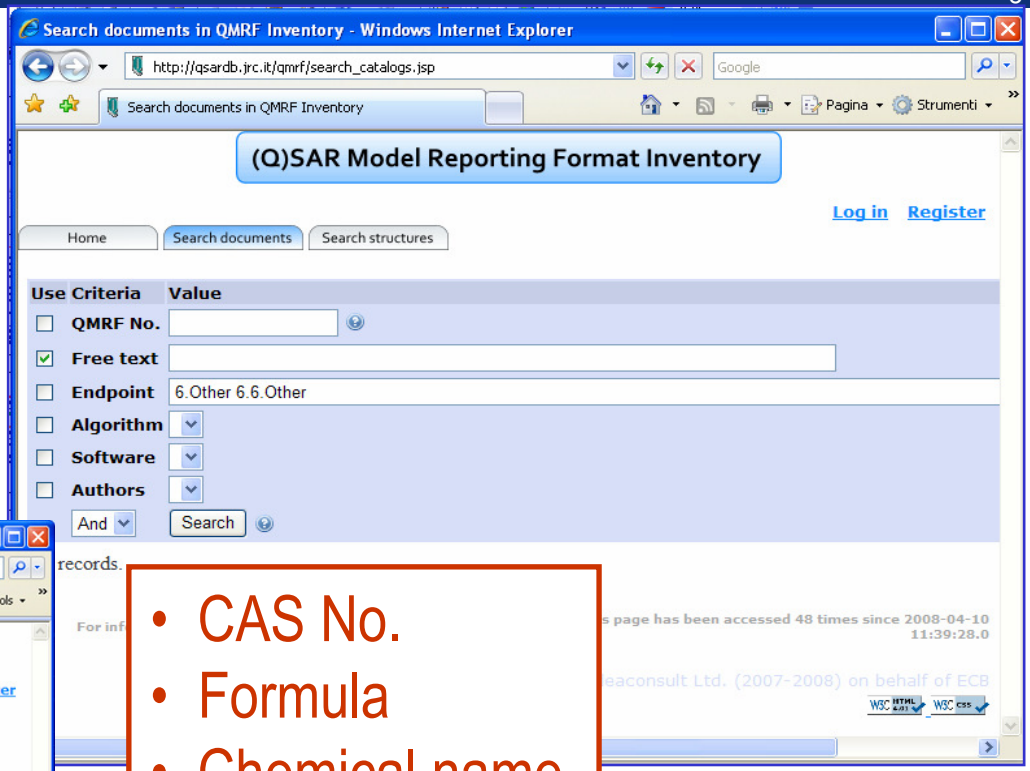
A (Q)SAR should be associated with the following information:

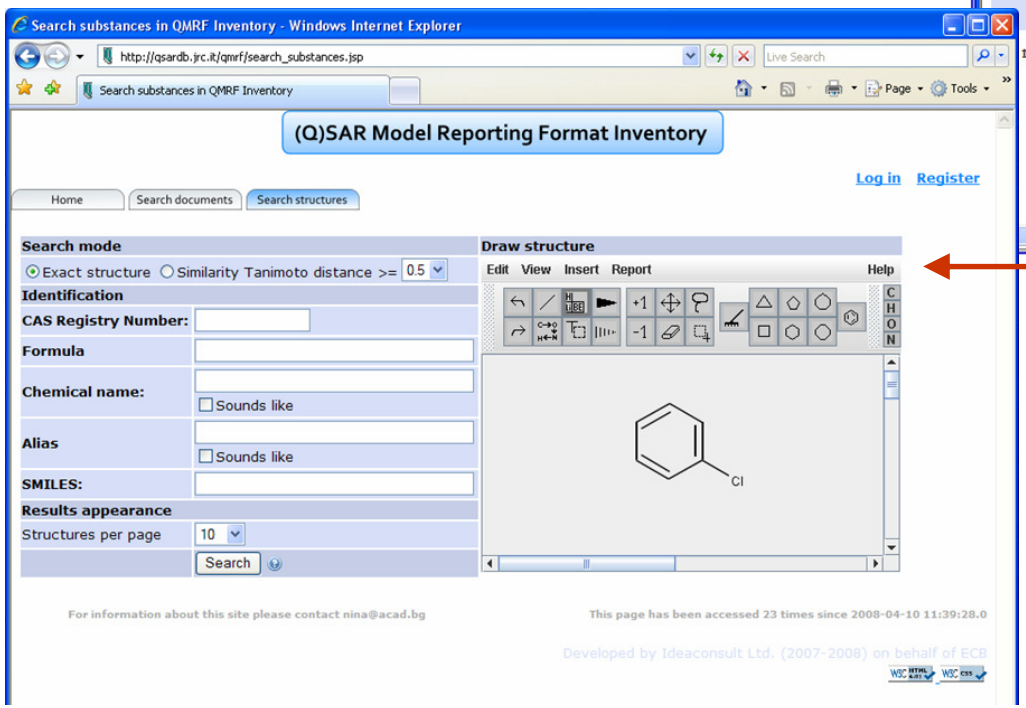
1. a defined endpoint
2. an unambiguous algorithm
3. a defined applicability domain
4. appropriate measures of goodness-of-fit, robustness and predictivity
5. a mechanistic interpretation, if possible

- **Principles adopted by 37th Joint Meeting of Chemicals Committee and Working Party on Chemicals, Pesticides & Biotechnology; 17-19 Nov 2004**
- **ECB preliminary Guidance Document published in Nov 2005**
- **OECD Guidance Document published in Feb 2007**
- **OECD Guidance summarised in REACH guidance (IR and CSA) 2008**

JRC QSAR Model Database

- QMRF No.
- Free text
- Endpoint
- Algorithm
- Software
- Authors



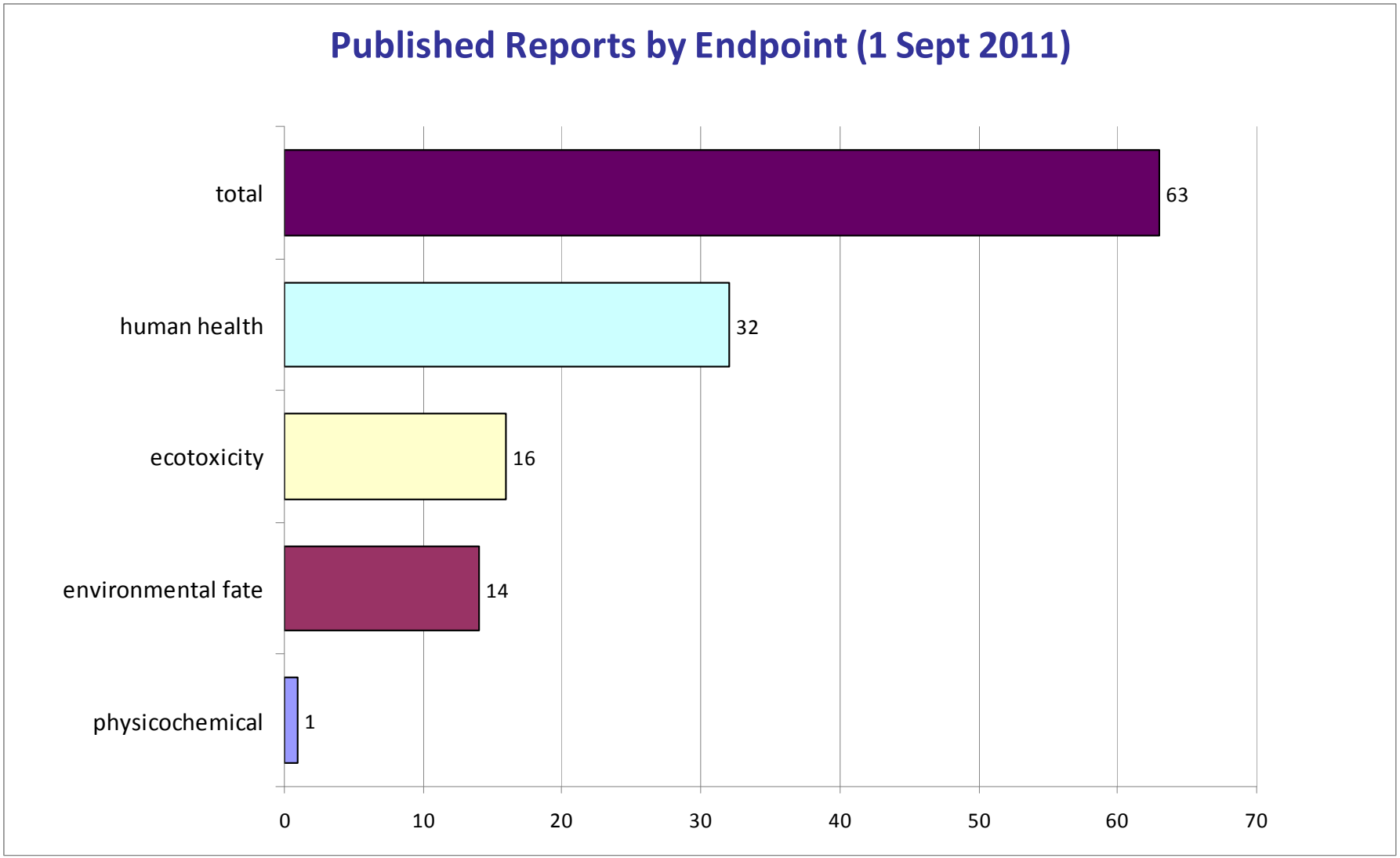


- CAS No.
- Formula
- Chemical name
- SMILES

Offline QMRF editor

http://qsardb.jrc.ec.europa.eu

QMRFs in JRC database



http://qsardb.jrc.ec.europa.eu

1 September 2011: 63 QMRFs published

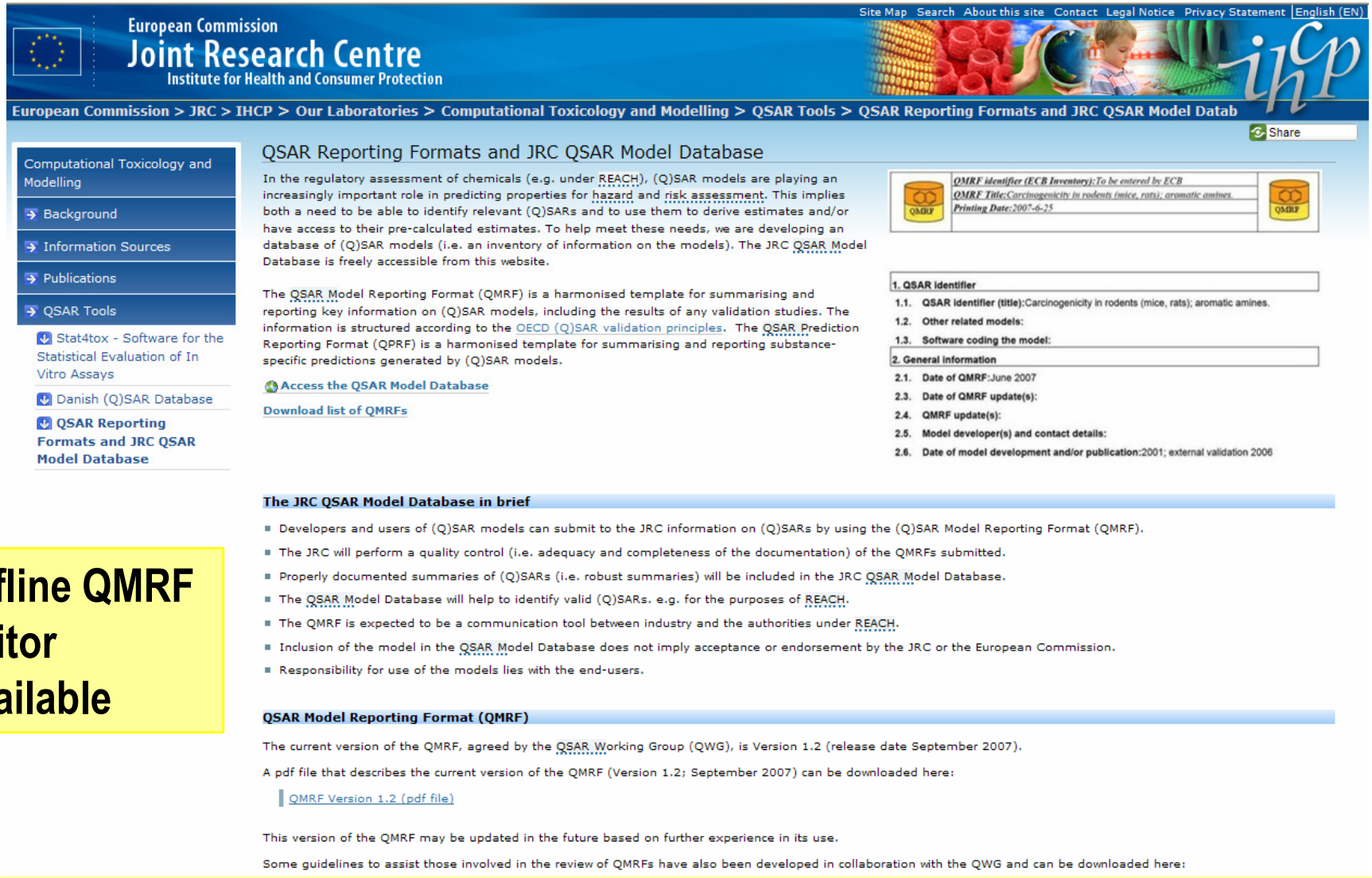
(Q)SAR Reporting Formats: QPRF

QPRF captures information on the substance and its prediction, and is intended to facilitate considerations of the adequacy of a prediction

1. Substance information
2. General (administrative) information on QPRF
3. Information on prediction (endpoint, algorithm, applicability domain, uncertainty, mechanism)
4. Adequacy (includes judgement and indicates whether additional information is needed for WoE assessment)

- Assessment of adequacy depends on reliability and relevance of prediction, but also on the availability of other information, and the consequence of being wrong
- Not just a scientific consideration, but also a policy decision

Information on Reporting Formats

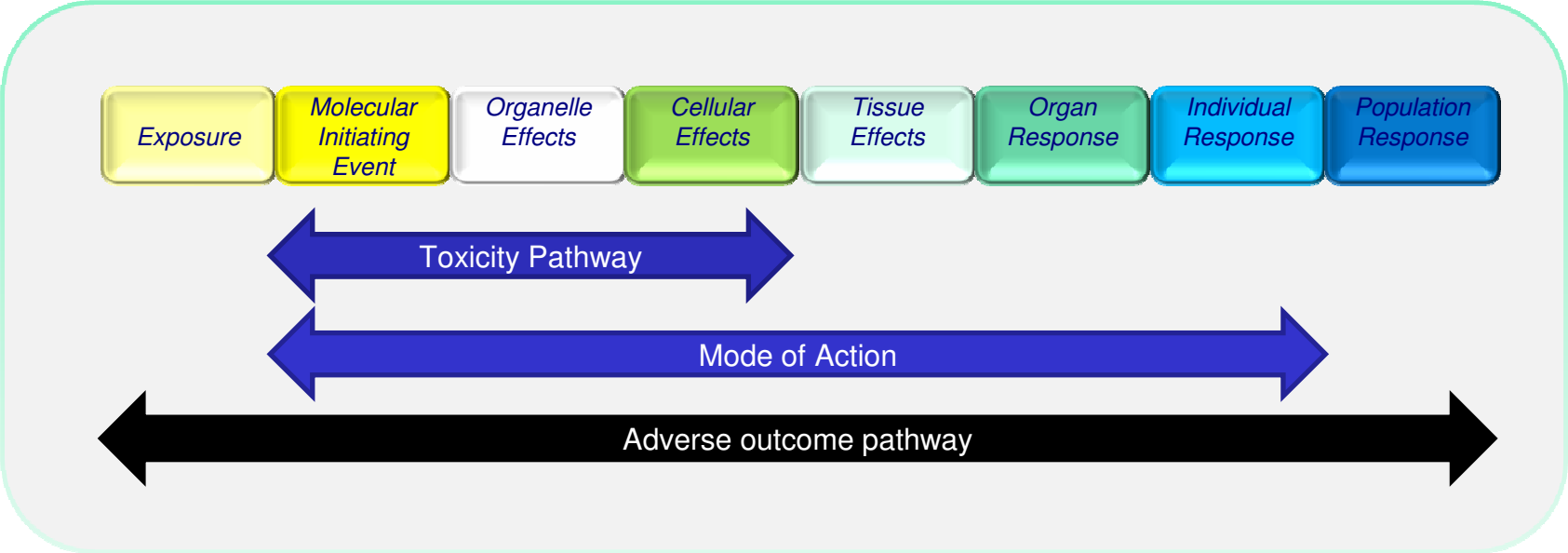


Offline QMRF editor available

http://ihcp.jrc.ec.europa.eu/our_labs/computational_toxicology/qsar_tools/QRF

Towards a template for Intermediate Effects

- Trend towards assessment based on Toxicity Pathways, Mode-of-Action (MoA) and Adverse Outcome Pathway (AOP)
- Development of OECD Harmonised Template (OHT) 201
- OECD project, led by JRC
- Compatibility with IUCLID, OECD Toolbox, Effectopedia (http://www.effectopedia.org/go/) and other tools



JRC
EUROPEAN COMMISSION

10 questions for assessing model predictions

1. Is the predicted endpoint clearly defined?
2. Is the predicted endpoint a *direct* information requirement?
3. Is the model training set fully available (for statistical models)?
4. Is the method used to develop the model well documented?
5. Is information available concerning the performance of the model?
6. In the case of a statistical model, is there evidence of overfitting?
7. Does the model training set contain the chemical of interest ?
8. Does the model make reliable predictions for analogues of the chemical of interest?
9. Is the prediction substantiated with argumentation based on the applicability domain of the model?
10. Can the prediction be easily reproduced?

Worth et al (2011). A Framework for assessing in silico Toxicity Predictions: Case Studies with selected Pesticides. JRC report EUR 24705 EN.

Example: CAESAR mutagenicity predictions

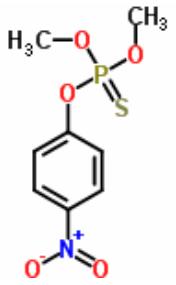
Q1	**Is the predicted endpoint clearly defined?**
A1	**Yes, the endpoint is Ames (*S. Typhimurium*) mutagenicity**
Q2	**If the predicted endpoint is clearly defined ("yes" to Q1), does it represent a direct information requirement under the legislation of interest, or is it related to one of the information requirements?**
A2	**Yes, genotoxicity test data are required under most types of chemicals legislation (e.g. industrial chemicals, pesticides, biocides)**
Q3	**If the model is statistically based (as opposed to knowledge-based), is the model training set fully available?**
A3	**Yes, the training and test set are published (http://www.caesar-project.eu)**
Q4	**Is the method used to develop the model documented or referenced (e.g. in a scientific paper or QMRF)**
A4	**Yes, a QMRF is in preparation, based on the following publications:** Ferrari T, Gini G & Benfenati E (2009). Support vector machines in the prediction of mutagenicity of chemical compounds. Proc NAFIPS 2009, June 14-17, Cincinnati, USA, p 1-6. Ferrari T & Gini G (2010). A new multistep model to predict mutagenicity from statistic analysis and relevant structural alerts. Central Chemistry 4, Suppl 1, S2.

Example: CAESAR mutagenicity predictions

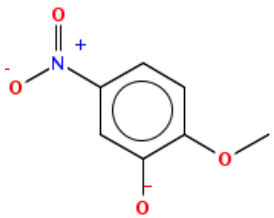
Q5	Is information available (in terms of statistical properties) concerning the performance of the model, including its goodness-of-fit, predictivity, robustness and error of prediction (uncertainty)?
A5	Yes. Information on the accuracy (82.1%), sensitivity (90.6%) and specificity (71.4%) are provided.
Q6	If the model is statistically based (as opposed to knowledge-based), does examination of the available statistics indicate that the model may have been overfitted?
A6	The model is statistically based but should not be overfitted because the ratio of chemicals (3380) to descriptors (42) is 80.5.

Q7	Does the model training set contain the chemical of interest?
A7	The model training set includes some pesticides including parathion-methyl but not sodium nitroguaiacolate.

Methyl parathion
Dimethoxy-(4-nitrophenoxy)-thioxo-phosphorane
CAS 298-00-0
S=P(Oc1ccc(cc1)[N+]([O-])=O)(OC)OC
Mutagen
Correctly predicted by CAESAR



Sodium Nitroguiacolate
2-methoxy-5-nitro-phenolate
CAS 67233-85-6
[Na+].[O-]c1cc(ccc1OC)[N+]([O-])=O
Non mutagen
Incorrectly predicted as mutagen by CAESAR



Example: CAESAR mutagenicity predictions

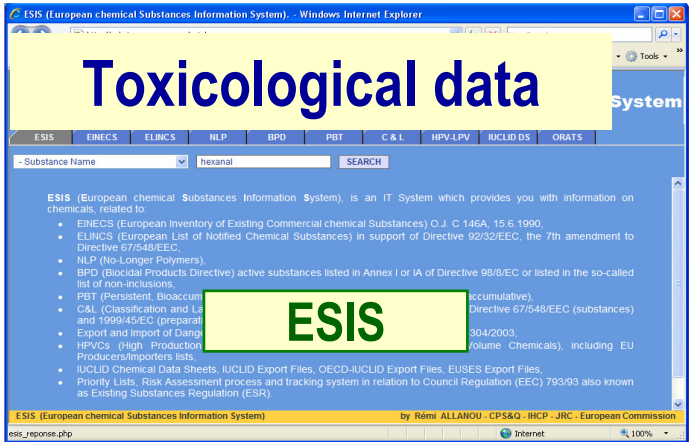
Q8	Does the model make reliable predictions for analogues of the chemical structure of interest?
A8	**Yes, the Caesar software gives the chance to examine, for each compound submitted, the six most similar compounds found in the model training set. For these compounds the experimental value for the selected endpoint is shown, together with the prediction made by the model. The similarity measure employed by the Caesar software takes into account functional group similarity, constitutional similarity, ring similarity and fingerprint similarity.** For parathion methyl (correctly predicted by the software), the similar structures obtained are: parathion methyl (input structure), aminofenitrothion, 1-ethenoxy-4-nitro-benzene, fenitrooxon, o-nitroanisole, N-hydroxy-N-(4-nitrophenyl)acetamide. All of them are predicted correctly by the software. For nitroguaiacolate (wrongly predicted by the software) the similar structures obtained are: o-nitroanisole, 1-ethoxy-3-nitro-benzene, 2,5-dinitrophenol, p-nitrosoanisole, 2-methoxy-1,3,5-trinitro-benzene, 1-ethenoxy-4-nitro-benzene. All of them are predicted correctly by the software.

Example: CAESAR mutagenicity predictions

Q9	Is the model prediction substantiated with argumentation based on the applicability domain of the model?
A9	Yes, Caesar addresses the applicability domain in several ways, namely by: a) checking whether the compound of interest falls in the descriptor space – if the compound is out of domain, this is noted in the output; b) providing a similarity score (1=identity) for the structure-based comparison with analogues; c) visual representation of the most similar compounds; d) by revealing the known and predicted toxicities for the analogues, thereby indicating the prediction error. Thus Caesar provides an assessment based on both the input (descriptor) space and the output (toxicological endpoint) space.
Q10	Can the model prediction be easily reproduced?
A10	Yes, the software is accessible in the form of a freely accessible web platform (http://www.caesar-project.eu) The software is easy to use, even for non-specialists.

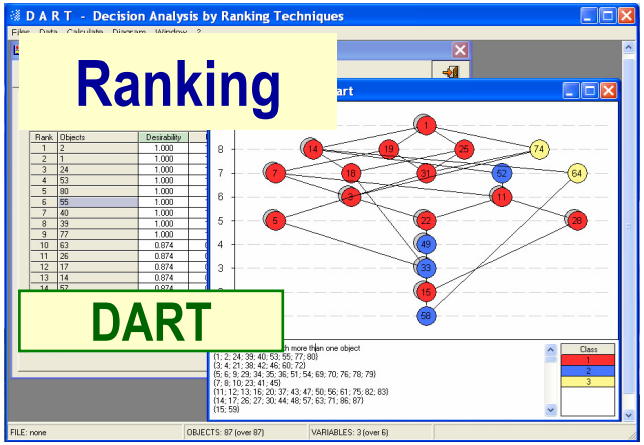
Computational tools developed by the JRC

Toxicological data



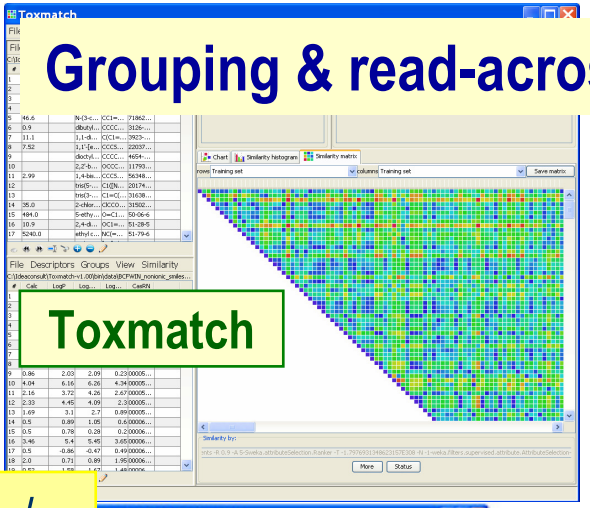
ESIS

Ranking



DART

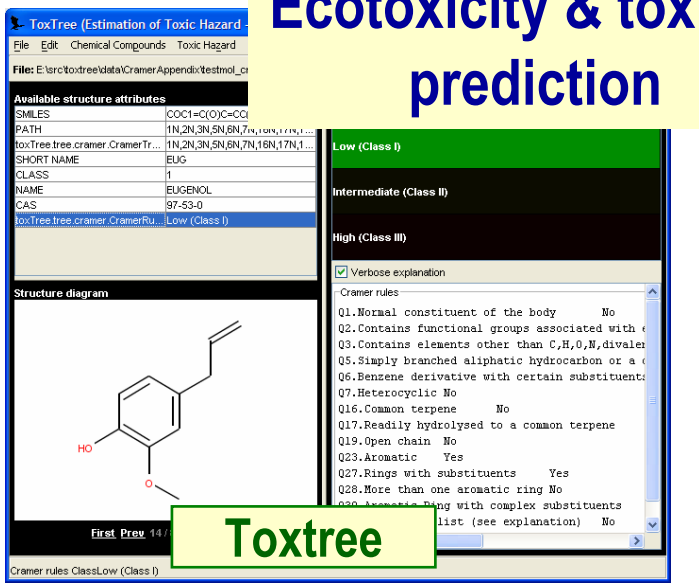
Grouping & read-across



Toxmatch

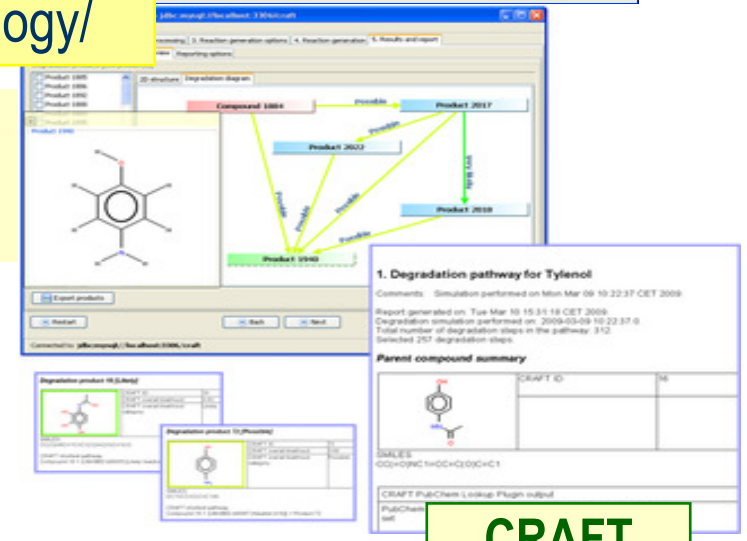
http://ihcp.jrc.ec.europa.eu/our_labs/computational_toxicology/

Ecotoxicity & toxicity prediction

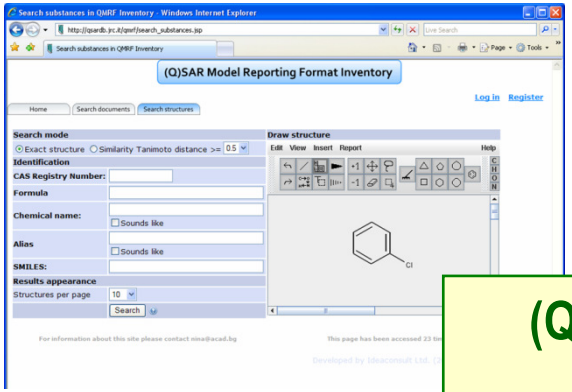


Toxtree

Metabolism & fate prediction



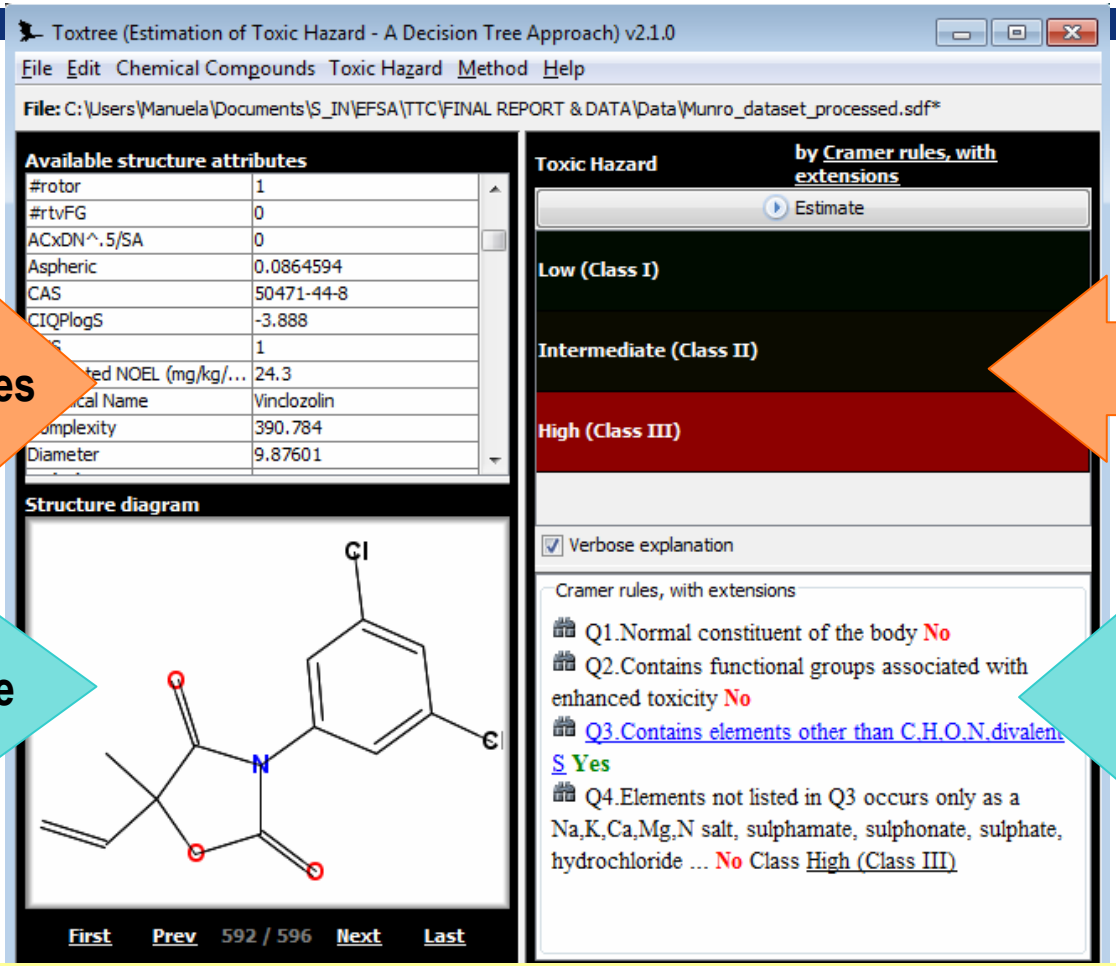
CRAFT



(Q)SAR model database

METIS

Toxicity Estimation Tool: Toxtree



Compound properties

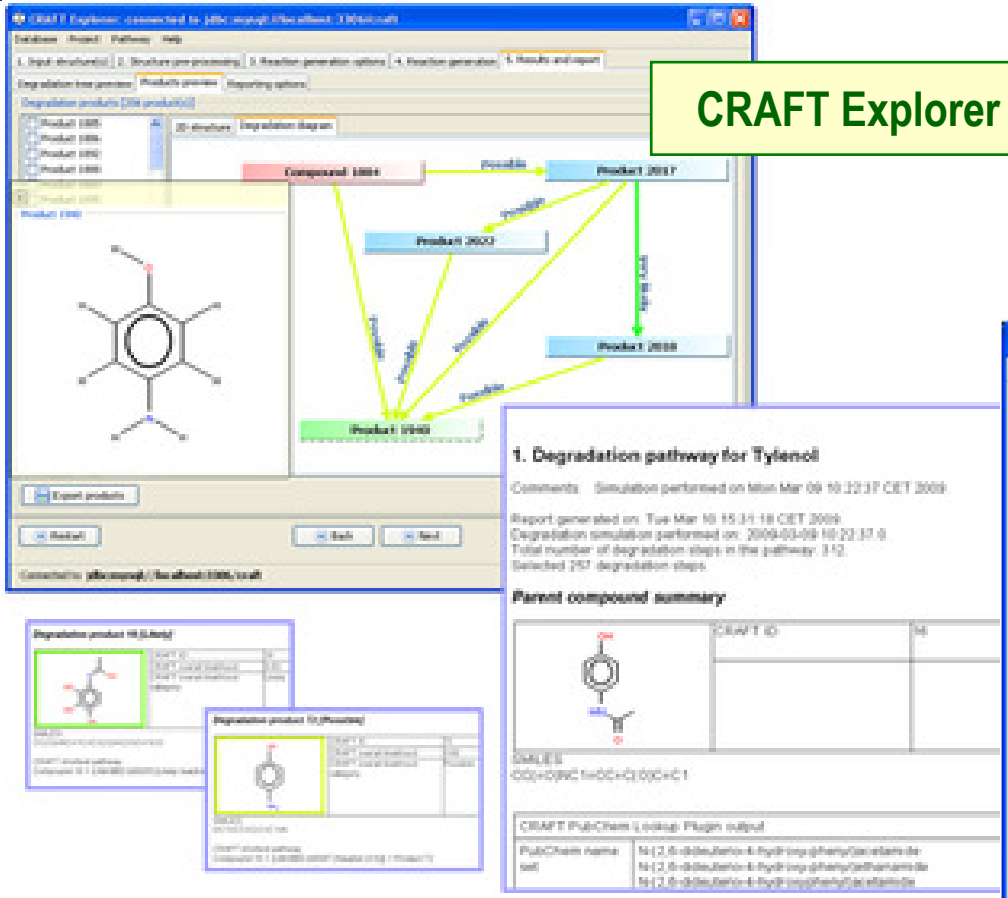
Compound structure

Prediction

Reasoning

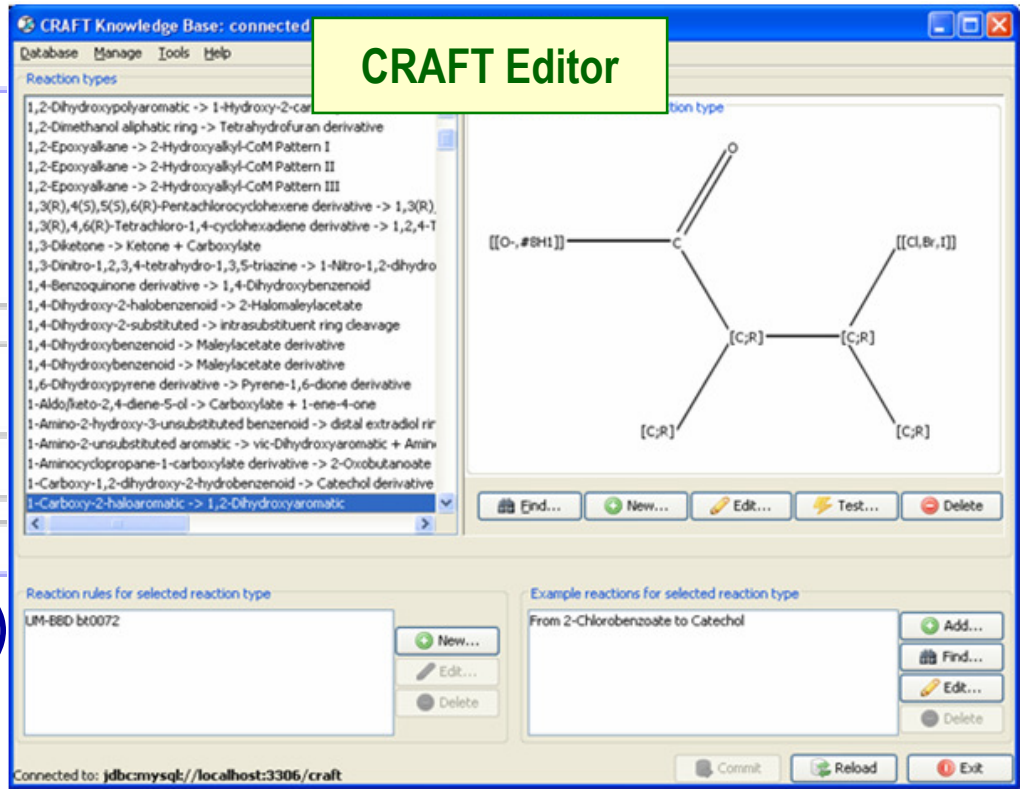
- Downloadable versions from JRC and Sourceforge (http://toxtree.sourceforge.net)
- Online version: OpenTox - ToxPredict (http://www.opentox.org/)
- Version 2.5.0 (August 2011) includes Verhaar, Extended Verhaar (Enoch 2008), START biodegradation, ISSMIC organic functional groups

CRAFT: Chemical Reactivity and Fate Tool



CRAFT Explorer

- Generates & visualises reactions, ranks transformation products
- User can modify rulebase

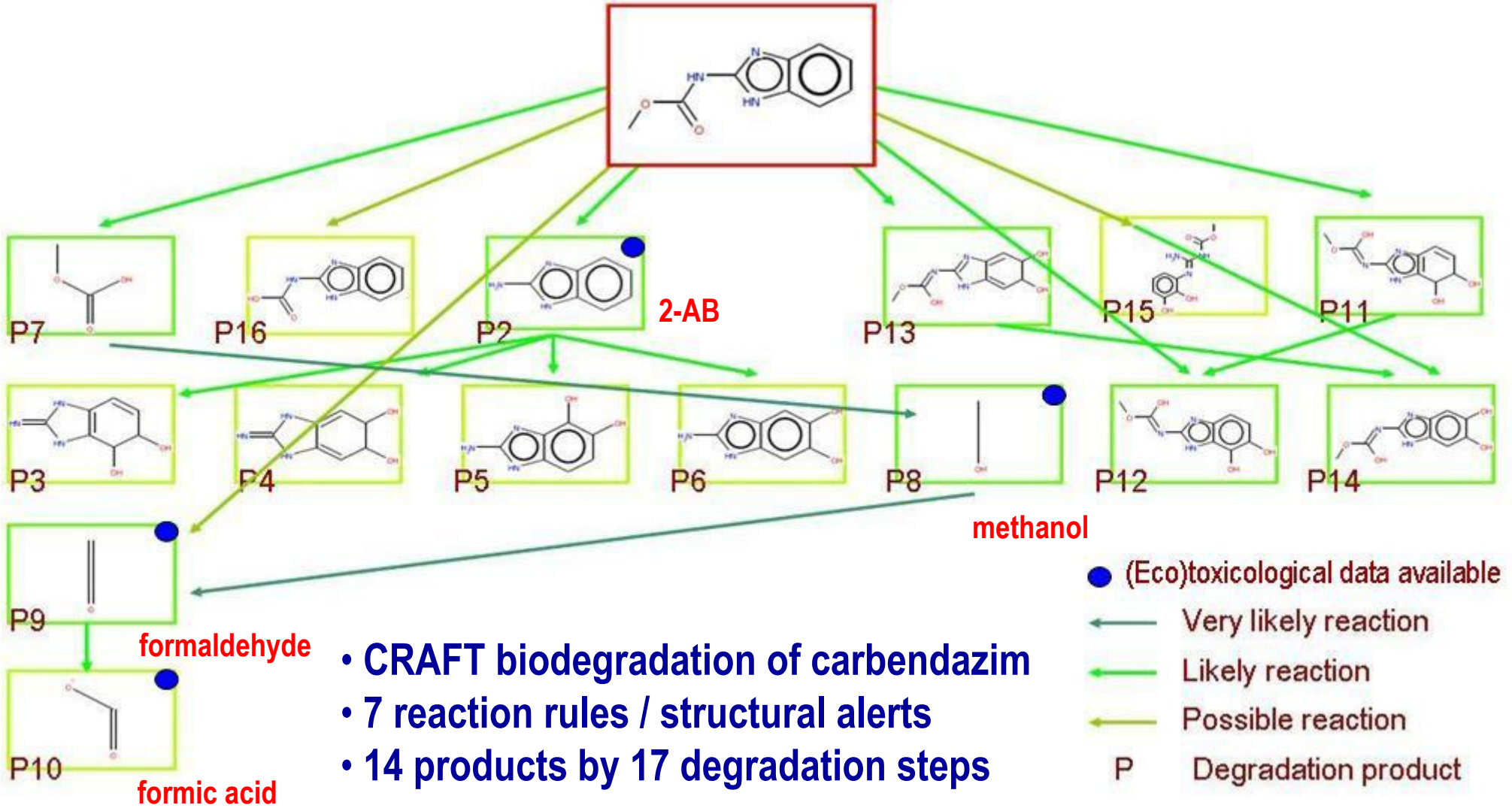


CRAFT Editor

- Developed by Molecular Networks (Germany) on behalf of JRC
- Data model based on AMBIT technology
- Compatible with METIS reaction editor

http://www.molecular-networks.com/products/craft

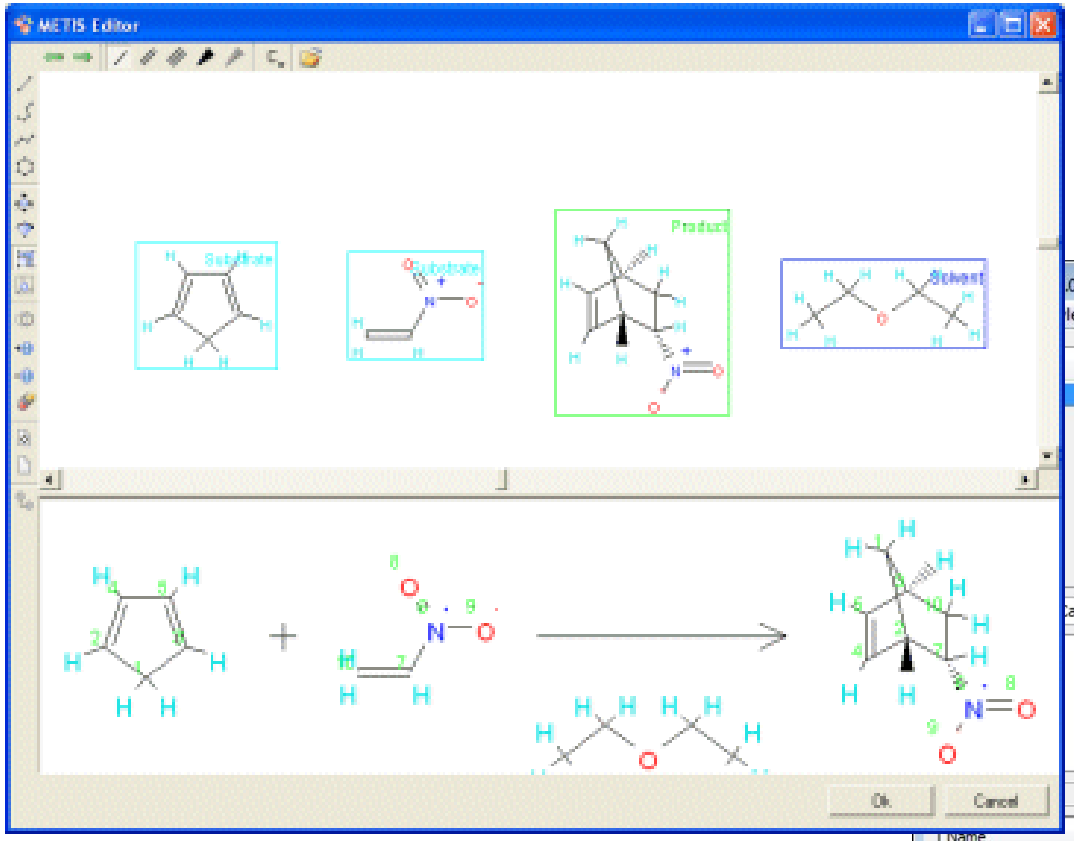
CRAFT prediction of carbendazim biodegradation



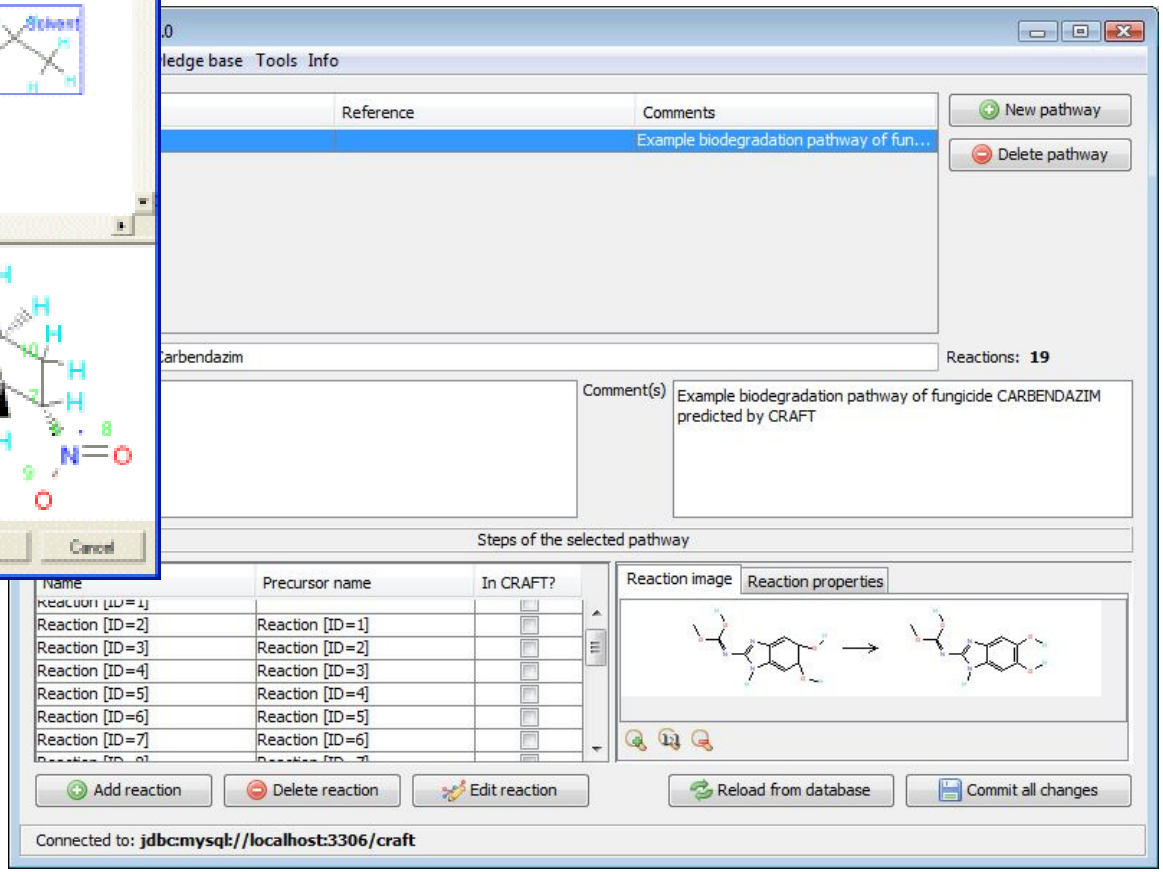
- CRAFT biodegradation of carbendazim
- 7 reaction rules / structural alerts
- 14 products by 17 degradation steps

Mostrag-Szlichtyng & Worth (2010). In silico modelling of microbial and human metabolism: a case study with the fungicide carbendazim. JRC report EUR 24523 EN.

METIS: Metabolic Information Input System



Reaction editor for storage, manipulation, and exchange of information on metabolic and degradation reactions



- **Developed by Molecular Networks (Germany) on behalf of JRC**
- **Import / export reactions from / to CRAFT & other applications**

http://www.molecular-networks.com/products/metis

Summary & conclusions

- In principle, (Q)SAR estimates can be used as direct replacements for test data, but in practice, use in weight-of-evidence assessments is more likely
- No reporting format for Integrated Testing Strategy, but tempate for intermediate effects under development
- To harmonise the use of QSARs, standardised templates for reporting the validity of QSAR models, and the adequacy of QSAR estimates, are provided in the REACH guidance documentation
- No formal validation and adoption procedures for (Q)SAR models
- Criteria for assessing the adequacy of (Q)SAR predictions?
- Examples needed to illustrate how to demonstrate adequacy

Key references

- OECD Guidance on QSAR validation (2007)

  http://www.oecd.org
- REACH Guidance on ITS and use of QSARs (2008)

  http://guidance.echa.europa.eu/docs/guidance_document/information_requirements_en.htm
- QSAR reporting formats (QMRF and QPRF) and QMRF Editor

  http://ihcp.jrc.ec.europa.eu/our_labs/computational_toxicology/qsar_tools/QRF
- QSAR Model Database

  http://qsardb.jrc.ec.europa.eu/qmrf/
- Mostrag-Szlichtyng A & Worth AP (2010). In silico modelling of microbial and human metabolism: a case study with the fungicide carbendazim. JRC report EUR 24523 EN.
- Worth A, Lapenna S, Lo Piparo E, Mostrag-Szlichtyng A & Serafimova R (2011). A Framework for assessing in silico Toxicity Predictions: Case Studies with selected Pesticides. JRC report EUR 24705 EN.

Appendix

An example of a QMRF

1. QSAR identifier

1.1. QSAR identifier (title): **Artificial neural network for acute fish toxicity (fathead minnow)**
1.2. Other related models: **No other related models.**
1.3. Software coding the model: **None**

2. General information

2.1. Date of QMRF: **18/09/2009.**
2.2. QMRF author(s) and contact details:
Aleksandra Mostrag-Szlichtyng; EC Joint Research Centre, Institute for Health and Consumer Protection, Via E. Fermi 2749, 21027 Ispra (VA), Italy;
aleksandra.mostrag-szlichtyng@ec.europa.eu
2.3. Date of QMRF update(s): **No QMRF update(s).**
2.4. QMRF update(s): **No QMRF update(s).**
2.5. Model developer(s) and contact details:
JRC Computational Toxicology Group
2.6. Date of model development and/or publication: **21/08/2009 (model development).**
2.7. Reference(s) to main scientific papers and/or software package:
Software package: ADMET Predictor™ 3.0; Simulations Plus, Inc. 42505 10th Street West Lancaster, CA 93534-7059 USA; http://www.simulations-plus.com/Products.aspx?grpID=1&cID=11&pID=13;
2.8. Availability of information about the model: **All information is available.**
2.9. Availability of another QMRF for exactly the same model: **No other QMRF available for the same model.**

3. Defining the endpoint – OECD Principle 1

3.1. Species:
Fathead Minnow (*Pimephales promelas*)

3.2. Endpoint:
3. Ecotoxic effects; 3.3. Acute toxicity to fish (lethality)

3.3. Comment on endpoint:
Experimental data on 96-h LC50 **(mmol/L) in fathead minnow for 577 studied chemicals were obtained from the Distributed Structure Searchable Toxicity (DSSTox) US-EPA Fathead Minnow Acute Toxicity (EPAFHM) Database. The subject of the experiments were juvenile fathead minnows (28 to 36 days-old) exposed into test substances via ninety-six-hour flow-through system (2).**

3.4. Endpoint units:
Molar 96-hours lethal concentration (LC50) in fathead minnow was expressed in (mmol/L) and inversed into decimal logarithmic scale: Log (96-h LC50) (mmol/L).

3.5. Dependent variable:
Log (96-h LC50) (mmol/L).

3.6. Experimental protocol:
The experimental protocols of biological/chemical investigations were described by Brooke et al. (3) and Geiger et al. (4). Organometallics, inorganic substances and chemicals for which the data were unavailable were excluded.

3.7. Endpoint data quality and variability:
The quality of data from DSSTox/EPAFAHM Database was verified by Russom et al. (2).

3 4. Defining the algorithm – OECD Principle 2

4.1. Type of model:

Artificial Neural Network model

4.2. Explicit algorithm:

Log (96-h LC50) model**;**

MLP-ANNE - Multilayer Perceptron Artificial Neural Network Ensembles Regression Model**;**

MLP-ANNE model was calculated with ADMET Predictor™ 3.0 software. After the procedures of (i) selecting model descriptors (i. e. removing invariant or highly correlated ones and performing sensitivity analysis to find the most relevant combination of them); (ii) splitting the input data into training pool (303 training set compounds + 173 verification test compounds) and test set (101 compounds) using Kohonen self-organising map (SOM) method; and (iii) training MLP-ANNE for different network architectures, the final model could be selected. It was characterized by the following architecture: 11-3-1 (i. e. 11 inputs [selected molecular descriptors], 3 neurons and 1 output [Log (96-h LC50), mmol/L]).

4.3. Descriptors in the model:

[1] S+logP; octanol-water partition coefficient
[2] SdCH2; atom-type electropological-state index for =CH2 groups
[3] Pi_Q4; derived from electronic properties, 4th component of the autocorrelation vector of Hückel pi atomic charges
[4] F_TpleB; constitutional descriptor, triple bonds as fraction of total bonds
[5] PolarizG ; [Å^3]; derived from electronic properties, polarizability calculated by Glen's method
[6] EEM_XFpl; derived from electronic properties, maximum sigma Fukui index on polar atoms
[7] N_Bonds; constitutional descriptor, number of bonds
[8] SsO-; atom-type electropological-state index for coordinated O- groups
[9] SHdsCH; atom-type electropological-state index for aCHa groups (aromatic carbons)
[10] StsC; atom-type electropological-state index for #C- groups
[11] Sscl; atom-type electropological-state index for -Cl groups

4. Defining the algorithm – OECD Principle 2

4.4. Descriptor selection:

ADMET Predictor™ 3.0 software calculated hundreds of descriptors **for each studied compound. Thus, the** pre-selection of "candidate" inputs **had to be performed. This procedure aimed to exclude (based on the statistical selection rules) from the initial set of available inputs those which were: (i) identical or of low variance (i. e. coefficient of variation, CV, lower than 1%); (ii) underrepresented (i. e. had non-zero values for less than 4 compounds); (iii) highly correlated (i. e. the correlation between raw descriptors was greater than 0.99999). Removing the latter resulted in the selection of 149 "candidate" inputs. In the next step, in order to find the optimal model complexity, the input gradient sensitivity analysis (SA) over all "candidates" was performed.** Finally, the set of 11 descriptors was selected**.**

4.5. Algorithm and descriptor generation:

All the descriptors were calculated with ADMET Predictor™ 3.0 software.

4.6. Software name and version for descriptor generation:

ADMET Predictor™ 3.0;
http://www.simulations-plus.com/Products.aspx?grpID=1&cID=11&pID=13;
Software for estimating certain ADMET (Absorption, Distribution, Metabolism, Elimination, and Toxicity) properties of a drug-like chemical from its molecular structure; 1998-2008; Simulations Plus Inc;
Simulations Plus, Inc. 42505 10th Street West Lancaster, CA 93534-7059 USA, Phone: +1.661.723.7723 (international), Toll free: 888.266.9294 (in the U.S. & Canada), Fax: +1.661.723.5524.

4.7. Descriptors/Chemicals ratio: **11/476 = 0.023** (43 chemicals / descriptor)

5. Defining the applicability domain – OECD Principle 3

5.1. Description of the applicability domain of the model:

Applicability domain based on the training pool, including 476 defined organic chemicals: 471 single compounds and 5 mixtures of formulation (for details please refer to the supporting files):

(i) AD by chemical classes**: the training pool compounds covered all standard chemical classes from EPAFHM Database (e.g. aliphatic and aromatic hydrocarbons, ethers, alcohols, aldehydes, ketones, amides, aliphatic and aromatic amines, sulfides, pyridines, barbitals); these compounds covered different modes of toxic action - the majority of them (200) was associated with baseline narcosis or electrophile/proelectrophile reactivity (82).**

(ii) AD by descriptor value ranges**: the model predictions were suitable for compounds characterized by the following descriptor values:**

[1] S+logP: min. -4.31; max. 6.77;
[2] SdCH2: min. 0.00; max. 5.42;
[3] Pi_Q4: min. -0.17; max. 0.45;
[4] F_TpleB: min. 0.00; max. 0.50;
[5] PolarizG: min. 3.47; max. 48.81;
[6] EEM_XFpl: min. -0.08; max. 0.45;
[7] N_Bonds: min. 1; max. 35;
[8] SsO-: min. 0.00; max. 30.90;
[9] SHdsCH: min. 0.00; max. 5.42;
[10] StsC: min. 0.00; max. 7.42;
[11] SsCl: min. 0.00; max. 35.69.

Experimental (observed) Log (96-h LC50) values for the training pool compounds varied from min. -6.38 to max. 2.96 mmol/L; for test set compounds from min. -3.25 to max. 2.85 mmol/L.

5. Defining the applicability domain – OECD Principle 3

5.2. Method used to assess the applicability domain:

Applicability Domain (AD) assessment based on the training pool compounds: (i) their chemical identity **(i.e. the presence of certain functional groups and their membership in particular chemical classes, e.g.** organometalics and inorganic substances were excluded**); (ii) the** ranges of descriptor values **describing the intrinsic properties of studied chemicals - the descriptor values of "predicted" compounds should fall between maximal and minimal descriptor values of the training pool compounds.**

5.3. Software name and version for applicability domain assessment:

ADMET Predictor™ 3.0**;**
http://www.simulations-plus.com/Products.aspx?grpID=1&cID=11&pID=13;
Software for estimating certain ADMET (Absorption, Distribution, Metabolism, Elimination, and Toxicity) properties of a drug-like chemical from its molecular structure; 1998-2008; Simulations Plus Inc;
Simulations Plus, Inc. 42505 10th Street West Lancaster, CA 93534-7059 USA, Phone: +1.661.723.7723 (international), Toll free: 888.266.9294 (in the U.S. & Canada), Fax: +1.661.723.5524.

5.4. Limits of applicability:

The model is suitable for specified chemical classes of compounds that have particular molecular descriptors in specified ranges (p. 5.1). The most sensitive descriptor was octanol-water partition coefficient (S+logP). The values of S+logP for training pool compounds varied from -4.31 to 6.77 as the applicability domain of the model covers chemicals characterized by different modes of toxic action. Compounds characterized by S+logP values lower than 0 as well as those with S+logP higher than 6 should not be modelled as narcotics – S+logP<0 indicates unrealistically high toxic effects, while S+logP>6 indicates that the uptake of compound from water is too slow to be connected with acute toxicity. The predictions performed by narcosis-type model can be associated with high uncertainty for such compounds.

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: Yes; 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:

The MLP-ANNE model was developed and internally validated based on the "training pool" including 476 compounds (303 "training set" compounds for neural networks training + 173 "verification set" compounds for internal validation). The algorithm used for training pool selection based on Kohonen self-organizing map (SOM) method.

6.6. Pre-processing of data before modelling:

Transformation of data from 96-h LC50 to logarithmic scale: Log (96-h LC50).

6.7. Statistics for goodness-of-fit:

The MLP-ANNE model's goodness-of-fit was tested according to 303 training set compounds:
Coefficient of Multiple Determination: $R^2 = 0.755$
Root Mean Squared Error of Calibration: RMSE = 0.699
Mean Absolute Error: MAE = 0.508

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

The MLP-ANNE model was internally validated according to 173 verification set compounds. In order to find the best complexity of the model (i. e. determine the moment of stopping the training procedure and avoid overtraining) the verification set errors were monitored (early stopping technique). The finally chosen model was characterized by the following, verification-set based, statistics:
Explained variance in prediction: $Q^2 = 0.809$

6.9. Robustness – Statistics obtained by leave-many-out cross validation: **No other information available.**

6.10. Robustness – Statistics obtained by Y-scrambling: **No other information available.**

6.11. Robustness – Statistics obtained by bootstrap: **No other information available.**

6.12. Robustness – Statistics obtained by other methods: **No other information available.**

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: Yes; 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 101 compounds appended.**

7.6. Experimental design of test set:

The external validation set (i. e. test set) consisted of 101 compounds from the entire data set, selected according to Kohonen Self-Organizing Map (SOM) mathematical method. The composition of test set was determined before the beginning of neural networks training procedure. The mapping process based on 11 previously selected descriptors, gathering the structural information on the studied compounds. The size of Kohonen map was 24x24 and all chemicals were clustered into 576 2-dimensional cells of similar structure, indicated by the values of the descriptors.

7.7. Predictivity – Statistics obtained by external validation:

External validation coefficient (based on test set compounds): $Q_{EXT}^2 = 0.715$

Root Mean Squared Error of Prediction (based on test set compounds): RMSE = 0.705

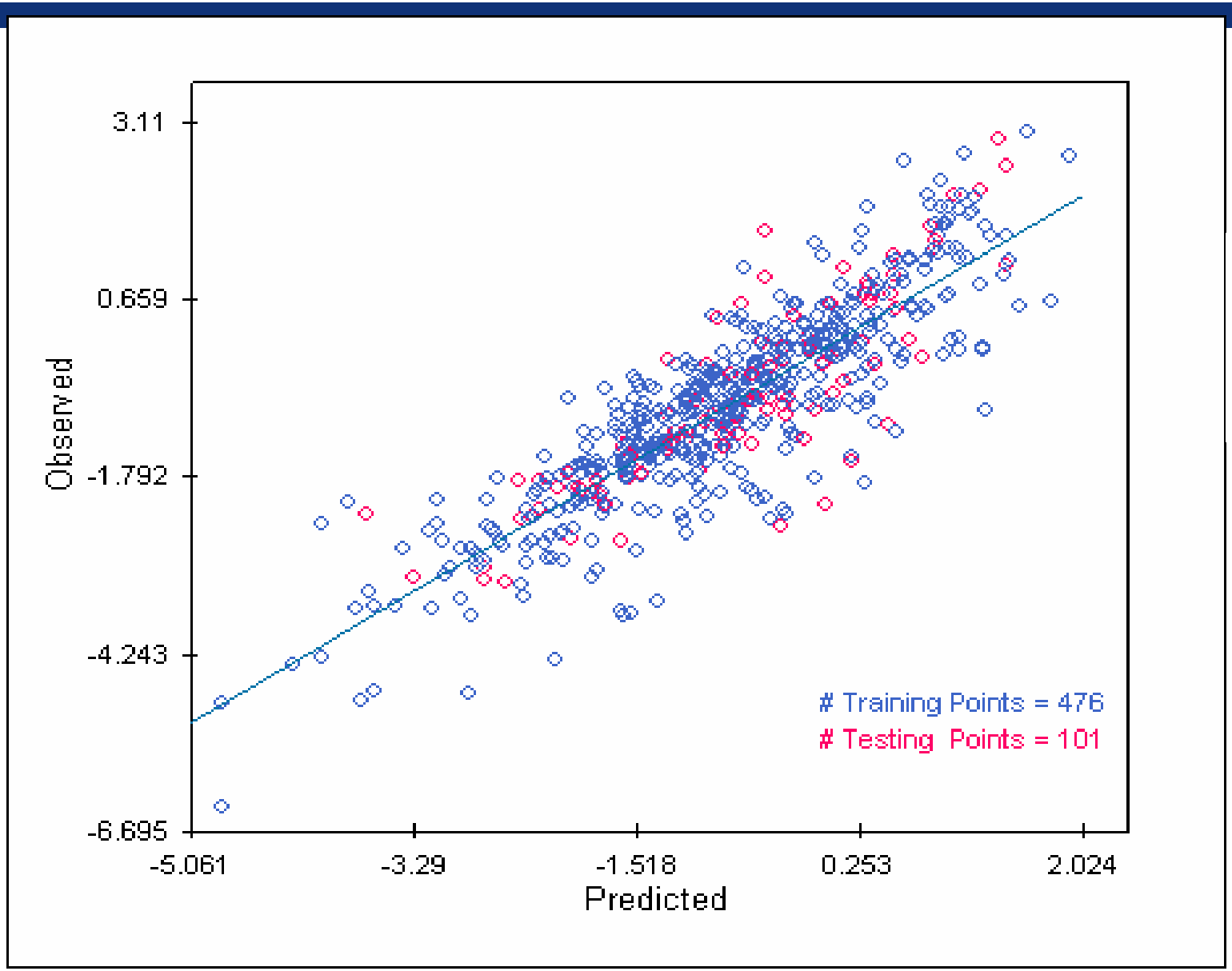
Mean Average Error (based on test set compounds): MAE = 0.515

7.8. Predictivity – Assessment of the external validation set:

The application of Kohonen SOM method allowed for determining the external validation (test) set, consisting of compounds representing the structural features and toxicological classes of the entire data set.

7.9. Comments on the external validation of the model: **No other information available.**

Validation of model – OECD Principles 4 & 5



Internal validation

$R^2 = 0.755$

RMSE = 0.699

MAE = 0.508

External validation

$Q_{EXT}^2 = 0.715$

RMSE = 0.705

MAE = 0.515

8. Mechanistic interpretation – OECD Principle 5

8.1. Mechanistic basis of the model:

As the MLP-ANNE model was developed statistically, no *a priori* assumptions have been made.

8.2. A priori or a posteriori mechanistic interpretation:

***A posteriori* mechanistic interpretation;**

The sensitivity analysis allowed to select molecular descriptors giving as much relevant information on the endpoint as possible. The most sensitive one was octanol-water partition coefficient (S+logP), which is the main mechanistically "interpretable" descriptor as far as acute aquatic toxicity is concerned. S+logP describes the kinetics of the process of uptaking chemicals from water via lipid membranes and thus indicates a baseline toxicity.

Other descriptors represent the structural features of chemicals as well as their electronic properties (e.g. polarizability, presence of polar/certain functional groups or bonds) – they reflect the polarity and the surface areas of compounds that are available for solvent (water) molecules as well as for lipid membranes of aquatic biota.

8.3. Other information about the mechanistic interpretation: **No other information available.**

9. Miscellaneous information

9.1. Comments:

The presented MLP-ANNE model is an example of the result of non-linear modelling based on the application of sophisticated mathematical and statistical approaches. As far as no equation describing the correlations between descriptors and the endpoint can be specified, the only way to transparently present the modelling procedure and its results is to describe it step-by-step in words – for this reason the model is transparent but not readily reproducible**.**

9.2. Bibliography:

(1) http://www.epa.gov/NCCT/dsstox/

(2) C.L. Russom, S.P. Bradbury, S.J. Broderius, D.E. Hammermeister and R.A. Drummond. Predicting Modes of Toxic Action from Chemical Structure: Acute Toxicity in the Fathead Minnow (*Pimephales Promelas*). Environ. Toxicol. Chem. 1997, 16 (5), 948-967.

(3) L.T. Brooke, D.J. Call, D.L. Geiger and C.E. Northcott, eds. 1984. Acute Toxicities of Organic Chemicals to Fathead Minnows (Pimephales promelas), Vol. 1. Center for Lake Superior Environmental Studies, University of Wisconsin, Superior, WI, USA

(4) D.L. Geiger, C.E. Northcott, D.J. Call and L.T. Brooke, eds. 1985. Acute Toxicities of Organic Chemicals to Fathead Minnows (Pimephales promelas), Vol. 2. Center for Lake Superior Environmental Studies, University of Wisconsin, Superior, WI, USA.

9.3. Supporting information:

Supporting Information on training and test sets appended.