

Computer tools for Threshold of Toxicological Concern Elucidation



Dr. Nina Jeliaskova

Ideaconsult Ltd.
4 A.Kanchev str.
Sofia 1000
Bulgaria
Email: nina@acad.bg
Phone +359 886802011
Fax +359 2 9526265

Content

❑ Toxtree

- Cramer rules
- Verhaar scheme for predicting toxicity mode of action
- Benigni-Bossa rulebase for carcinogenicity and mutagenicity
- ILSI decision tree (Kroes)
- Other plug-ins

❑ Ambit

- Open source software for chemo-informatics data management, allowing storage of large number of chemical structures and toxicological data and providing flexible means for exploration of structural and similarity spaces
- TTC use cases:
 - ❑ Cumulative dose distribution per class
 - ❑ Datasets comparison

Toxtree 1.60

<http://toxtree.sourceforge.net>

- Estimates toxic hazard by applying a decision tree approach.
- Full-featured and flexible user-friendly open source software
 - New decision trees with arbitrary rules can be built with the help of graphical user interface or by developing new plug-ins in Java code
 - GPL license
- Platform independent
- Input:
 - datasets from various compatible file types
 - SMILES
 - built-in 2D structure diagram editor.
- Output
 - SDF, MOL, CSV, MS Excel, CML, TXT, PDF, HTML
- Batch mode
- 10 classification schemes (plug-ins) for various endpoints assessment available
- More plug-ins under development

The screenshot displays the Toxtree v1.51 application window. At the top, the title bar reads "Toxtree (Estimation of Toxic Hazard - A Decision Tree Approach) v1.51". The menu bar includes "File", "Edit", "Chemical Compounds", "Toxic Hazard", "Method", and "Help". Below the menu bar, there is a text input field for "Enter SMILES:" containing the string c1ccc(N)cc1N=Nc2ccc(NC)cc2, followed by a "Go!" button. The main interface is divided into several panels. On the left, the "Available structure attributes" panel lists various molecular descriptors and their values for the input structure. In the center, the "Structure diagram" panel shows a 2D chemical structure of 4-(4-aminophenyl)azobenzene. On the right, the "Toxic Hazard" panel displays the results of the assessment using the Benigni / Bossa rulebase. It includes a list of structural alerts and a detailed explanation of the classification.

Attribute	Value
BSSTM1_(1)	1.0000
BSSTM1_(2)	1.0000
Benigni / Bossa rulebase ...	SA1N,SA2N,SA3N,SA4N...
Comment	Created from SMILES
EHOMO_(1)	-8.3905
EHOMO_(2)	-8.0311
ELUMO_(1)	0.4598
ELUMO_(2)	0.4268
FORMULA	C13H14N4
For a better assessment ...	NO
I(An)_(1)	true

Toxic Hazard
by Benigni / Bossa rulebase (for mutagenicity and carcinogenicity)

Estimate

- Structural Alert for genotoxic carcinogenicity
- Structural Alert for nongenotoxic carcinogenicity
- No alerts for carcinogenic activity
- Potential *S. typhimurium* TA100 mutagen based on QSAR

☒ Verbose explanation

/ Bossa rulebase (for mutagenicity and carcinogenicity)

Category	Result
Acyl halides	No
Alkyl (C<5) or benzyl ester of sulpho	No
N-methylol derivatives	No
Monohaloalkene	No
S or N mustard	No
Propiolactones and propiosultones	No
Epoxides and aziridines	No
Aliphatic halogens	No
Alkyl nitrite	No
Simple aldehyde	No

Completed.

Toxtree 1.60 modules:

1. Cramer rules
2. Verhaar scheme for predicting toxicity mode of actions
3. Skin irritation and corrosion potential estimation
4. Eye irritation and corrosion potential estimation
5. Benigni-Bossa rulebase for carcinogenicity and mutagenicity prediction
6. Cramer rules with extensions
7. Structure Alerts for the in vivo micronucleus assay in rodents
8. START (Structural Alerts for Reactivity in Toxtree) biodegradation and persistence plug-in
9. Michael Acceptors identification

Toxtree: Cramer rules

- Cramer classification scheme is used to assess TTC based on structural information:

Cramer G. M., R. A. Ford, R. L. Hall, Estimation of Toxic Hazard - A Decision Tree Approach, J. Cosmet. Toxicol., Vol.16, pp. 255-276, Pergamon Press, 1978

G. Patlewicz, N. Jeliaskova, R.J. Safford, A.P. Worth and B. Aleksiev, An evaluation of the implementation of the Cramer classification scheme in the Toxtree software, SAR and QSAR in Environmental Research, Vol. 19, Nos. 3-4, April-June 2008, 1-30

- Software implementation
 - Since toxTree v1.00 (2005)
 - IdeaConsult Ltd. For JRC

Toxtree (Estimation of Toxic Hazard - A Decision Tree Approach) v1.60

File Edit Chemical Compounds Toxic Hazard Method Help

Enter SMILES: c1ccccc1N=Nc2ccccc2 Go!

Available structure attributes	
B5STM1_(1)	1.0000
B5STM1_(2)	1.0000
Benigni / Bossa rulebase (for mutag...	SA1N,SA2N,SA3N,SA4N,SA5N,SA...
Comment	Created from SMILES
EHOMO_(1)	-8.6097
EHOMO_(2)	-8.6097
ELUMO_(1)	0.4179
ELUMO_(2)	0.4179
Error when applying the decision tree	NO
FORMULA	C12H10N2
For a better assessment a QSAR c...	NO

Structure diagram

First Prev 1 / 1 Next Last

Completed.

Toxic Hazard by Cramer rules

Estimate

Low (Class I)

Intermediate (Class II)

High (Class III)

☒ Verbose explanation

Cramer rules

- Q1.Normal constituent of the body No
- Q2.Contains functional groups associated with enhance...
- Q3.Contains elements other than C,H,O,N,divalent S...
- Q5.Simply branched aliphatic hydrocarbon or a common...
- Q6.Benzene derivative with certain substituents No
- Q7.Heterocyclic No c1ccccc1N=Nc2ccccc2
- Q16.Common terpene No c1ccccc1N=Nc2ccccc2
- Q17.Readily hydrolysed to a common terpene No
- Q19.Open chain No c1ccccc1N=Nc2ccccc2
- Q23.Aromatic Yes c1ccccc1N=Nc2ccccc2
- Q27.Rings with substituents Yes
- Q28.More than one aromatic ring Yes
- Q29.Readily hydrolysed No c1ccccc1N=Nc2ccccc2
- Q33.Has sufficient number of sulphonate or sulphonam...

Toxtree: Cramer rules

- ❑ Chemicals are divided into three structural classes based on a decision tree. This comprises some 33 structural rules and places evaluated compounds into one of three classes:
 - **Class I** substances are simple chemical structures with efficient modes of metabolism suggesting a low order of oral toxicity; (Fifth percentile NOEL (mg/kg bw/day) 3.0 Human exposure threshold (mg/person/day) 1.8)
 - **Class III** substances are those that permit no strong initial presumption of safety, or may even suggest significant toxicity or have reactive functional groups (Fifth percentile NOEL (mg/kg bw/day) 0.15 Human exposure threshold (mg/person/day) 0.09)
 - and finally, **Class II** are intermediate (Fifth percentile NOEL (mg/kg bw/day) 0.91 Human exposure threshold (mg/person/day) 0.54)
- ❑ Conservative scheme
 - For example presence of halogen (with the exception of hydrochloride) or phosphorus leads to **Class III**

Cramer classification scheme

- Questions assess different features:
 - Structural features (functional groups, ring substituents, etc.)
 - Propensity of a reaction (e.g. hydrolysis) is questioned and if positive, the reaction residues are subject for further analysis
 - Natural occurrence in body and in traditional foods
 - Logic of the tree relies primarily on knowledge of common metabolic pathways
- Designed to be used by an expert
- Software implementation presents some challenges

The screenshot displays the 'Cramer rules' software window. At the top, a decision tree is visualized with nodes numbered 1 through 33, connected by green and red lines. Below the tree, the 'Decision node' is set to 'Q1: Normal constituent of the body'. The 'If 'NO' go to' is 'Q:2' and the 'If 'YES' assign' is 'Low (Class I)'. A table lists Rule ID 1 with the title 'Normal constituent of the body'. The 'Rule explanation' section contains the text: 'Is the substance a normal constituent of the body, or an optical isomer of such? This question throws into class I all normal constituents of body tissues and fluids, including normal metabolites. Hormones are excluded, as are, by implication, the metabolites of environmental and food contaminants or those resulting from disease state. Note the answer of the question relies on an incomplete list of compounds, identified by an expert as a normal body constituents. If you believe a query'. To the right, under 'Example with answer 'YES'', is a chemical structure of 2-aminopyridine-4-carbaldehyde (Nc1cc(C=O)cnc1). At the bottom, there are radio buttons for 'Yes branch' (selected) and 'No branch'.

Toxtree internals:

“Normal constituents of the body”

Cramer rules

Decision node: Q1.Normal constituent of the body

If 'NO' go to: Q.2

If 'YES' assign: Low (Class 1)

Rule ID	Rule title
1	Normal constituent of the body

Rule explanation

Returns true if the query is isomorphic to one of the structures loaded from a preconfigured file of a type SDF, SMI, CSV

Is the substance a normal constituent of the body, or an optical isomer of such? This question throws into class I all normal constituents of body tissues and fluids, including normal metabolites. Hormones are excluded, as are, by implication, the metabolites of environmental and food contaminants or those resulting from disease state.

Note the answer of the question relies on an incomplete list of compounds, identified by an expert as a normal body constituents. If you believe a query compound is wrongly identified as a such, or not recognised, please consult and/or update the list. C:\Ideaconsoft\Toxtree-v1.51\Toxtree\bodymol.sdf

There are example molecules for each rule outcome. Select which one to display.

☒ Yes branch ☐ No branch

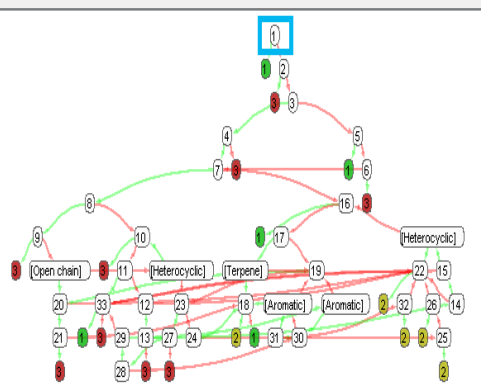
Example with answer 'YES'

- Rule 1 : “Normal constituent of the body”
- File: bodymol.sdf
- 68 compounds in ToxTree 1.0-1.60
- 440 compounds ToxTree 1.61

Toxtree internals:

“Common components of food”

Cramer rules



Decision node: Q1. Normal constituent of the body

If 'NO' go to: Q.2
If 'YES' assign: Low (Class 1)

Rule ID: 1 Rule title: Normal constituent of the body

Rule explanation

Returns true if the query is isomorphic to one of the structures loaded from a preconfigured file of a type SDF, SMI, CSV

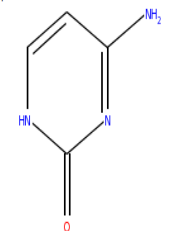
Is the substance a normal constituent of the body, or an optical isomer of such? This question throws into class I all normal constituents of body tissues and fluids, including normal metabolites. Hormones are excluded, as are, by implication, the metabolites of environmental and food contaminants or those resulting from disease state.

Note the answer of the question relies on an incomplete list of compounds, identified by an expert as a normal body constituents. If you believe a query compound is wrongly identified as a such, or not recognised, please consult and/or update the list. C:\Ideaconsoft\Toxtree-v1.51\Toxtree\bodymol.sdf

There are example molecules for each rule outcome. Select which one to display.

☒ Yes branch ☐ No branch

Example with answer 'YES'



- Rule 22 : “Common component of food”
- File: foodmol.sdf
- 108 compounds

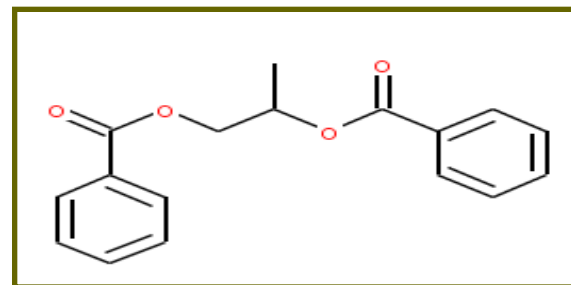
Toxtree internals

Series of questions are applied on query compound and also on reaction products

Q1.Normal constituent of the body No
Q2.Contains functional groups associated with enhanced toxicity No
Q3.Contains elements other than C,H,O,N, divalent S No
Q5.Simply branched aliphatic hydrocarbon or a common carbohydrate No
Q6.Benzene derivative with certain substituents No
Q7.Heterocyclic No
Q16.Common terpene No
Q17.Readily hydrolysed to a common terpene No
Q19.Open chain No
Q23.Aromatic Yes
Q27.Rings with substituents Yes
Q28.More than one aromatic ring Yes
Q29.Readily hydrolysed Yes

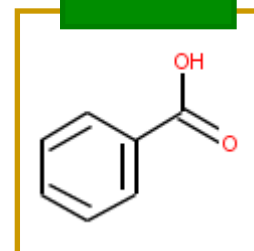
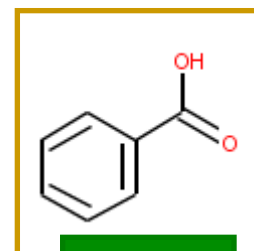
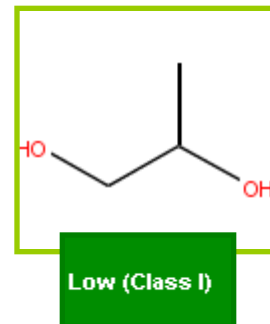
Q19.Open chain Yes
Q20.Aliphatic with some functional groups Yes
Q21.3 or more different functional groups No
Q18.One of the list (see explanation) No **Low (Class I)**

Q30.Aromatic Ring with complex substituents No
Q18.One of the list (see explanation) No **Low (Class I)**



Query compound

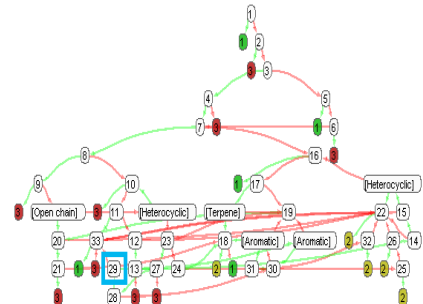
Hydrolysis
products



Toxtree internals: Hydrolysis reactions

A limited number of hydrolysis reactions (Cramer rules #15, #17, #29, #30 and #31) are implemented, based on an expert advice.

Cramer rules



Decision node: Q29: Readily hydrolysed

If 'NO' no to: Q.33

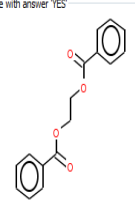
If 'YES' no to: Q. [Aromatic]

Rule ID: 29

Rule title: Readily hydrolysed

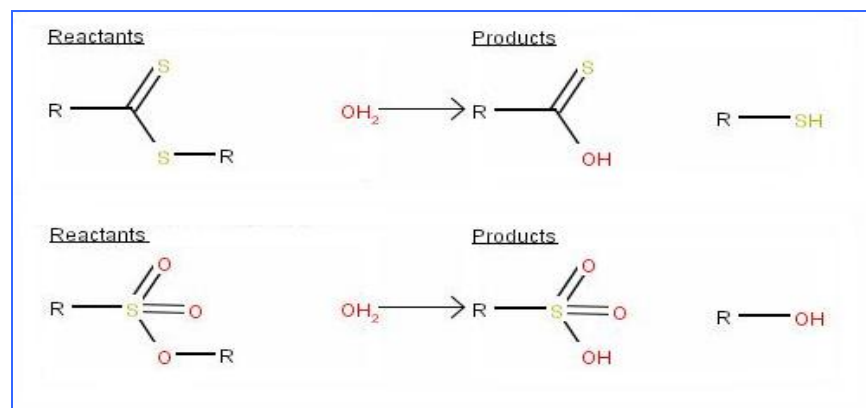
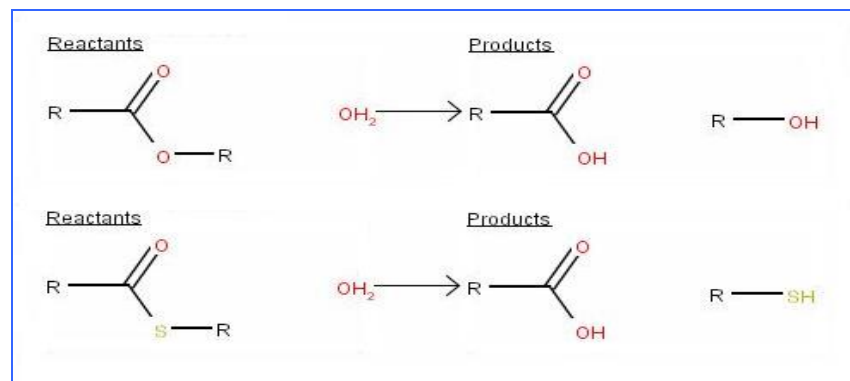
Rule explanation: Is it readily hydrolysed(H) to mononuclear residues? If YES, treat the individual aromatic mononuclear residues by Q.30 and any other residue by Q19.

Example with answer 'YES'



There are example molecules for each rule outcome. Select which one to display.

☒ Yes branch ☐ No branch



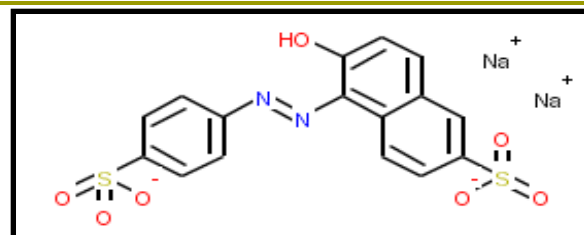
Toxtree internals

Series of questions are applied on query compound and also on reaction products

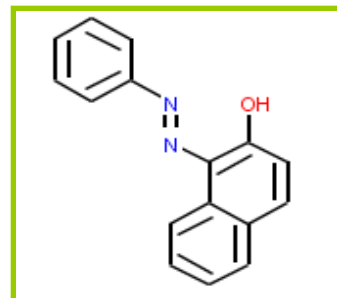
- Q1.Normal constituent of the body No
- Q2.Contains functional groups associated with enhanced toxicity: No
- Q3.Contains elements other than C, H, O, N, divalent S: Yes
- Q4.Elements not listed in Q3 occurs only as a Na, K, Ca, Mg, N salt, sulphamate, sulphonate, sulphate, hydrochloride Yes

- Q7.Heterocyclic No
- Q16.Common terpene No
- Q17.Readily hydrolysed to a common terpene No
- Q19.Open chain No
- Q23.Aromatic Yes
- Q27.Rings with substituents Yes
- Q28.More than one aromatic ring Yes
- Q29.Readily hydrolysed No

- Q33.Has sufficient number of sulphonate or sulphamate groups
Yes **Low (Class I)**

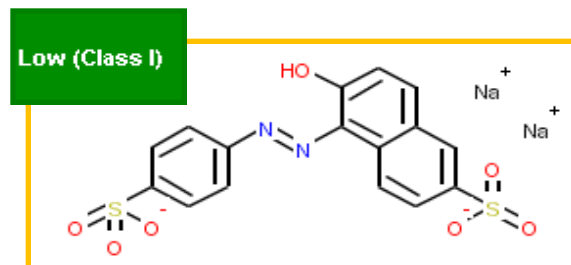


Query compound



Unsulphonated compound

Query compound (again)



Low (Class I)

Toxtree internals

Rule4: Elements not listed in Q3 occurs only as a Na, K, Ca, Mg, N salt, sulphamate, sulphonate, sulphate, hydrochloride

Decision node: Q4.Elements not listed in Q3 occurs only as a Na,K,Ca,Mg,N salt, sulphamate, sulphonate, sulphate, hydrochloride ...

If 'NO' assign: High (Class III)

If 'YES' go to: Q,7

Rule ID: 4

Rule title: Elements not listed in Q3 occurs only as a Na,K,Ca,Mg,N salt, sulphamate, sulphonate, sulphate, hydrochloride ...

Rule explanation:

Do all elements not listed in Q3 occur only as

- (a) a Na,K,Ca,Mg or N salt of a carboxylic acid, or
- (b) a sulphate or hydrochloride of an amine, or
- (c) a Na,K, or Ca sulphonate, sulphamate or sulphate?

If the answer is yes, treat as free acid, amine, unsulphonated or unsulphated compound, except for the purposes of Q24 and Q33.

This is intended to let through, for further consideration, certain acid,amine, sulphonate and sulphate salts. Sulphamate salts are treated as such because they are not readily hydrolysed.

There are example molecules for each rule outcome. Select which one to display.

☒ Yes branch ☐ No branch

Example with answer 'YES'

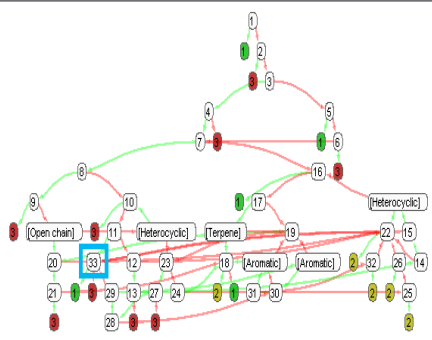
[Na+].[O-]P(=O)([O-])Nc1ccccc1

- Do all elements not listed in Q3 occur only as (a) a Na,K,Ca,Mg or N salt of a carboxylic acid, or
- (b) a sulphate or hydrochloride of an amine, or
- (c) a Na,K, or Ca sulphonate, sulphamate or sulphate?
- If the answer is yes, **treat as free acid, amine, unsulphonated or unsulphated compound, except for the purposes of Q24 and Q33.** This is intended to let through, for further consideration, certain acid,amine, sulphonate and sulphate salts. Sulphamate salts are treated as such because they are not readily hydrolysed.

Toxtree internals: Metabolism

A limited number of metabolic reactions (used by Cramer rule #33) reactions are implemented, based on an expert advice.

Cramer rules



Decision node: Q33.Has sufficient number of sulphonate or sulphamate groups

IF 'NO' assign: High (Class III)
IF 'YES' assign: Low (Class I)

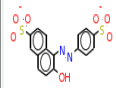
Rule ID	Rule title
33	Has sufficient number of sulphonate or sulphamate groups

Rule explanation

Does the substance bear on every major structural component at least one Na, K or Ca sulphonate or sulphamate for every >20 carbon atoms, without any free primary amines except those adjacent to the sulphonate or sulphamate.

Na,K,Ca sulphonate and sulphamate salts have a strong tendency to decrease toxicity by promoting solubility and rapid excretion. This is particularly noticeable, for example, with some of the food colourings. It is important that the substance bears sufficient sulphonate groups, including one on each major structural fragments into which the original compound might be metabolized. This question serves to steer sulphonated compounds except those with amines non-adjacent to the sulphonate into a presumably less toxic classification than the compounds would occupy if unsulphonated.

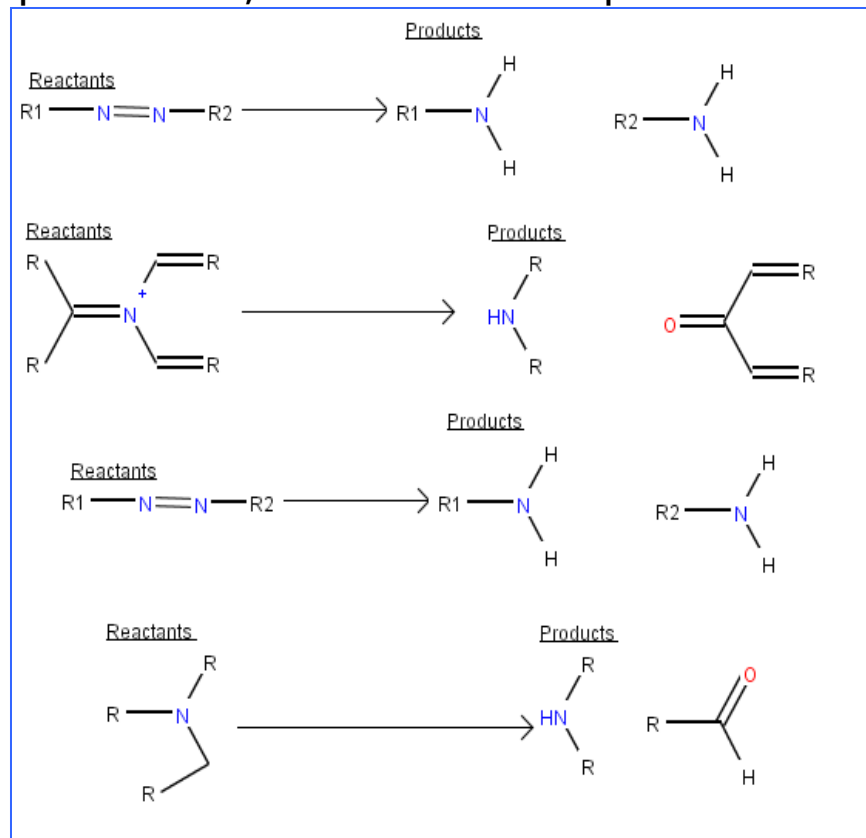
Example with answer 'YES'



Na⁺ Na⁺

There are example molecules for each rule outcome. Select which one to display.

☒ Yes branch ☐ No branch



Toxtree: Verhaar scheme

- Predicting toxicity mode of action

Verhaar, 1992 Verhaar HJM, van Leeuwen CJ and Hermens JLM (1992) Classifying environmental pollutants. 1. Structure-activity relationships for prediction of aquatic toxicity. Chemosphere 25, 471-491.

Verhaar, 2000 Verhaar HJM, Solbe J, Speksnijder J, van Leeuwen CJ and Hermens JLM (2000) Classifying environmental pollutants: Part 3. External validation of the classification system. Chemosphere 40, 875-883

- Software implementation
 - Since toxTree v1.20 (2005)
 - IdeaConsult Ltd. for AMBIT project

Toxtree (Estimation of Toxic Hazard - A Decision Tree Approach) v1.60

File Edit Chemical Compounds Toxic Hazard Method Help

Enter SMILES: c1ccccc1N=NC2=CC=CC=C2 Go!

Available structure attributes	
BSSTM1_(1)	1.0000
BSSTM1_(2)	1.0000
Benigni / Bossa rulebase (for mutag... SA1N,SA2N,SA3N,SA4N,SA5N,SA...	
Comment Created from SMILES	
EHOMO_(1)	-8.6097
EHOMO_(2)	-8.6097
ELUMO_(1)	0.4179
ELUMO_(2)	0.4179
Error when applying the decision tree	NO
FORMULA	C12H10N2
For a better assessment a QSAR c...	NO

Structure diagram

First Prev 1 / 1 Next Last

Completed.

Toxic Hazard by Verhaar scheme

Estimate

Class 1 (narcosis or baseline toxicity)

Class 2 (less inert compounds)

Class 3 (unspecific reactivity)

Class 4 (compounds and groups of compounds acting by a specific mechanism)

Class 5 (Not possible to classify according to these rules)

☒ Verbose explanation

Verhaar scheme

Q0.1.Consists only of C,H,N,O,S,halogens (excluding I) Yes

Q0.2.Have a logKow between 0 and 6 Yes

Q0.3.Have a molecular mass (MW) not more than 600 Daltons Yes

Q1.1.Not contain I Yes c1ccccc1

Q1.2.Not contain ionic groups Yes

Q1.3.Contain only C&H No c1ccccc1

Q1.4.Contain only C,H and halogen No

Q1.5.Contain C,H & O No c1ccccc1

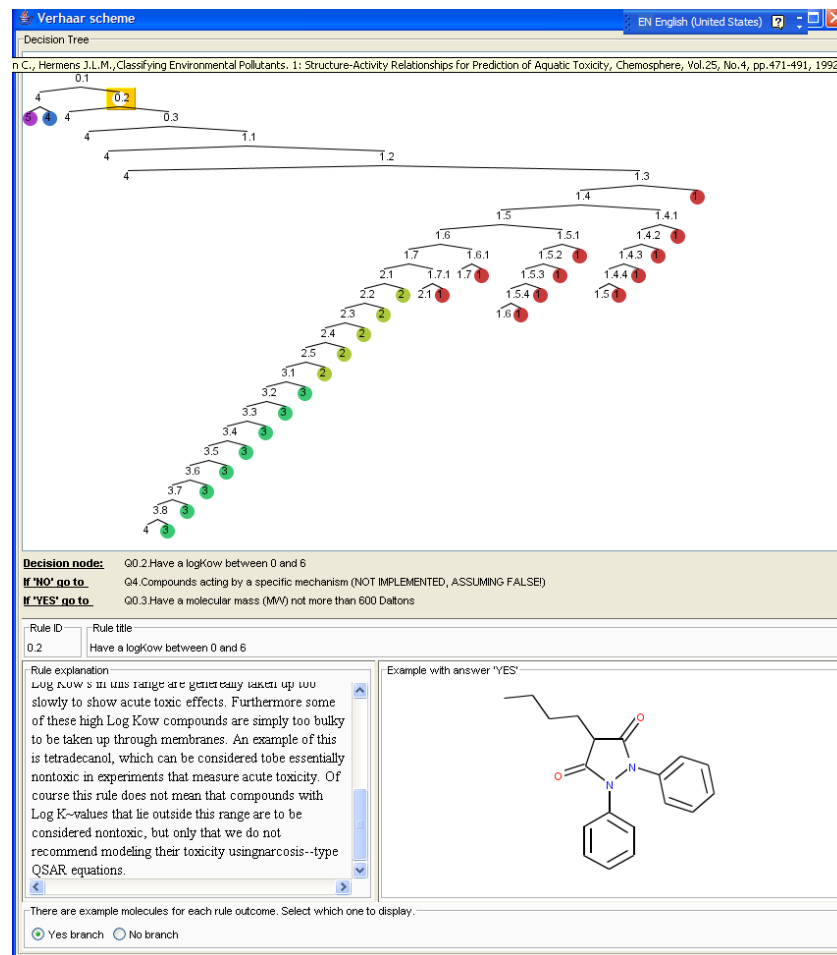
Q1.6.Contain only C,H,N Yes c1ccccc1

Q1.6.1.Be aliphatic secondary or tertiary amines

Q1.7.Contain only C,H,O and halogen No

Toxtree: Verhaar Scheme

- ❑ 34 rules
- ❑ 5 classes
 - **Class 1.** Narcosis or baseline toxicity
 - **Class 2** Less inert compounds
 - **Class 3** Unspecific reactivity
 - **Class 4** Compounds and groups of compounds acting by a specific mechanism
 - **Class 5** Not possible to classify according to these rules
- ❑ Compounds that cannot be classified as belonging to class 1,2 or 3 and that are not known to be compounds acting by a specific mechanism can only be classified as "not possible to classify according to these rules"



Toxtree: Benigni-Bossa rulebase for carcinogenicity and mutagenicity

- Predicting genotoxic and non-genotoxic carcinogenicity and mutagenicity by discriminant analysis and structural rules

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

A software tool for predicting mutagenicity and carcinogenicity
http://ecb.jrc.it/documents/QSAR/EUR_23241_EN.pdf

- Software implementation
 - Since toxTree v1.50 (2008)
 - IdeaConsult Ltd. for JRC

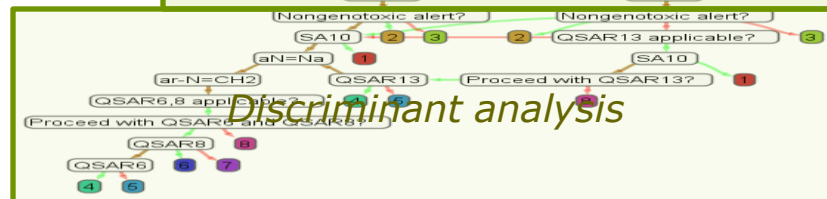
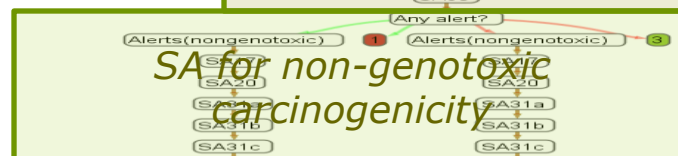
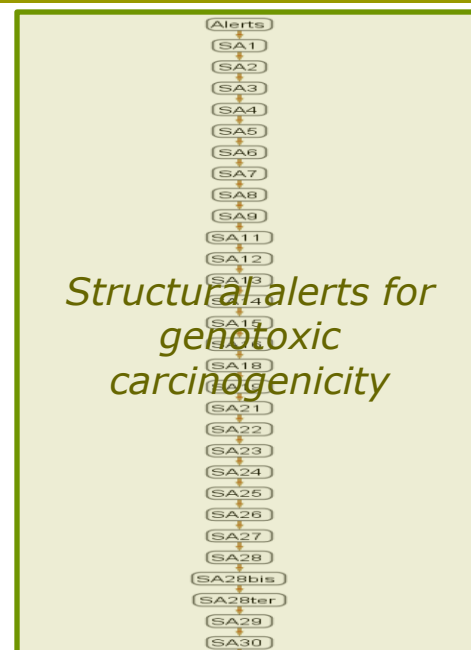
The screenshot displays the Toxtree (Estimation of Toxic Hazard - A Decision Tree Approach) v1.60 interface. The main window is divided into several sections:

- Available structure attributes:** A table listing various attributes and their values for the input compound.
- Structure diagram:** A chemical structure diagram of the input compound, which is a diphenylhydrazine derivative.
- Toxic Hazard:** A section titled "by Benigni / Bossa rulebase (for mutagenicity and carcinogenicity)" showing the estimated hazard. It includes a red bar for "Structural Alert for genotoxic carcinogenicity" and a green bar for "Unlikely to be a S. typhimurium TA100 mutagen based on QSAR".
- Verbose explanation:** A section providing a detailed explanation of the prediction, listing various QSAR rules and their outcomes.

The input SMILES string is c1ccccc1N=Nc2ccccc2. The structure diagram shows a diphenylhydrazine derivative. The prediction results indicate a structural alert for genotoxic carcinogenicity and a potential carcinogen based on QSAR.

Toxtree: Benigni-Bossa rulebase for carcinogenicity and mutagenicity

- ❑ Set of structural alerts for genotoxic and non-genotoxic carcinogenicity
- ❑ Classes:
 - a) no structural alerts for carcinogenicity are recognised;
 - b) one or more structural alerts (SAs) are recognised for genotoxic or non-genotoxic carcinogenicity;
 - c) SAs relative to aromatic amines or a,b unsaturated aldehydes are recognised, and the chemical goes through Quantitative Structure-Activity Relationship (QSAR) analysis, which may result in a negative or positive outcome.
 - If the query chemical belongs to the classes of aromatic amines or a,b unsaturated aldehydes, the appropriate QSAR is applied and provides a more refined assessment than the SAs



Toxtree: Descriptors calculation

□ **LogP**

- Used by Verhaar scheme, Skin irritation scheme and QSARs in Benigni-Bossa rulebase
- Based on XLogP implementation (atom-additive method)
- R. Wang, Y. Fu and L. Lai, A new atom-additive method for calculating partition coefficients, *J. Chem. Inf. Comput. Sci.* 37 (1997) 615-621.

□ **Molar Refractivity**

- Ghose-Crippen molar refractivity
- Viswanadhan, V.N., Ghose, A.K., Revankar, G.R. & Robins, R.K. *J. Chem. Inf. Comput. Sci.*, (1989), 29,163-172.

□ **Partial MR, Sterimol**

- Used by QSARs in Benigni-Bossa rulebase
- Based on (limited) lookup table for substituents or aromatic rings as in Corwin Hansch, Albert Leo, David Hoekman, *Exploring QSAR: Volume 2: Hydrophobic, Electronic, and Steric Constants*

□ **eHOMO, eLUMO**

- Used by QSARs in Benigni-Bossa rulebase
- Calculated by OpenMOPAC <http://openmopac.net/>

Toxtree: TTC decision tree

- Decision tree proposed by ILSI Europe to decide whether substances can be assessed by the TTC approach
- Kroes et al., Food and Chemical Toxicology 42, p76, 2004**
- Software implementation
 - Beta test in ToxTree v2.00 (2009)**
 - IdeaConsult Ltd.

Toxtree (Estimation of Toxic Hazard - A Decision Tree Approach) v2.0.0

File Edit Chemical Compounds Toxic Hazard Method Help

Enter SMILES: Go!

Available structure attributes	
Kroes TTC decision tree	Substance would not be expected to be a safety concern
Kroes TTC decision tree #expl...	Q1N,Alerts(genotoxic)Y,SA1N...
Names	Created from SMILES
SMILES	CCCCC

Structure diagram

Structure diagram area showing a skeletal structure of a straight chain alkane (pentane).

First Prev 1/1 Next Last

Completed.

Toxic Hazard by Kroes TTC decision tree

Estimate

Substance would not be expected to be a safety concern

Negligible risk (low probability of a life-time cancer risk greater than 1 in 10⁶)

Risk assessment requires compound-specific toxicity data

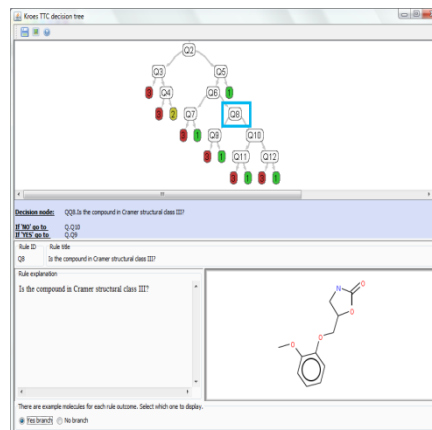
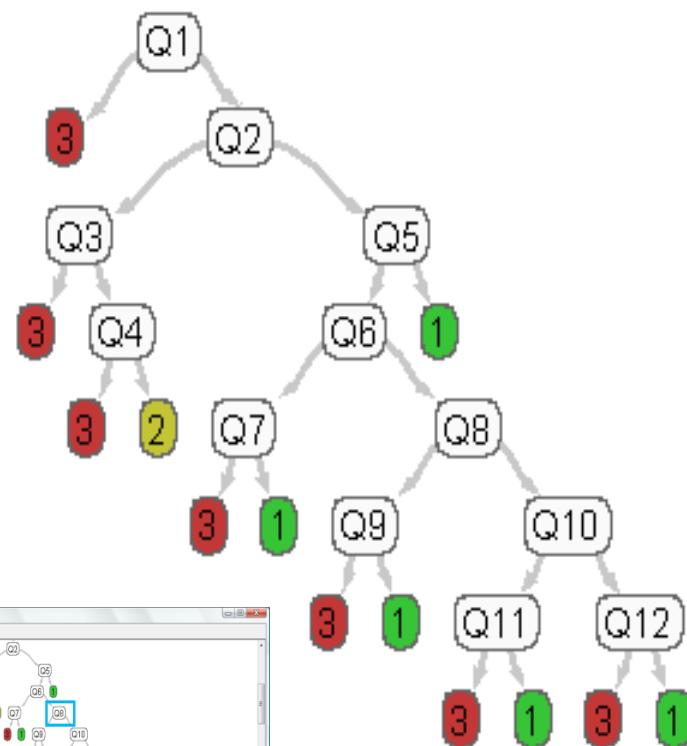
☒ Verbose explanation

Kroes TTC decision tree

Q01.Is the substance a non-essential metal
QAlerts(genotoxic).Verify structural alert:
QSA1.Acyl halides No
QSA2.Alkyl (C<5) or benzyl ester of sulpho
QSA3.N-methylol derivatives No
QSA4.Monohaloalkene No
QSA5.S or N mustard No
QSA6.Propiolactones and propiosultones No
QSA7.Epoxides and aziridines No
QSA8.Aliphatic halogens No
QSA9.Alkyl nitrite No
QSA11.Simple aldehyde No
QSA12.Quinones No
QSA13.Hydrazine No

TTC decision tree: Structural alerts

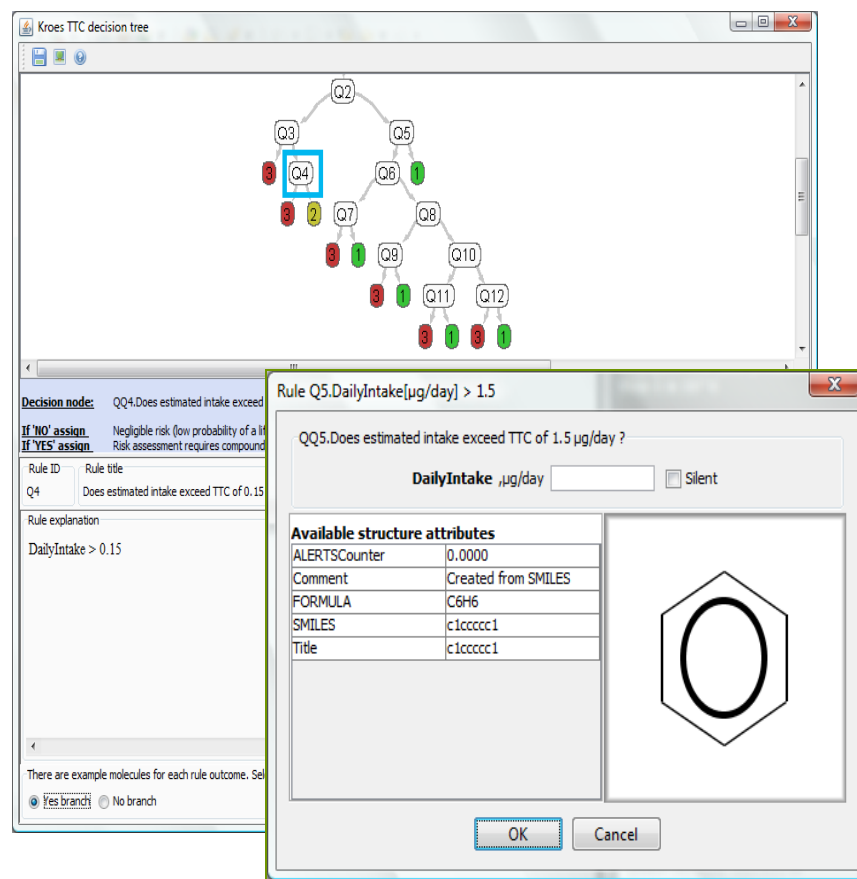
- Q1: Is the substance a non-essential metal or metal containing compound, or is it a polyhalogenated dibenzodioxin, -dibenzofuran, or -biphenyl?
- Q2: Genotoxic alerts – SA1-SA30 from Benigni-Bossa rule base for predicting cancerogenicity and mutagenicity
- Q3: Is the chemical an aflatoxin-like, azoxy-, or N-nitroso compound
- Q6: Is the compound an organophosphate?
- Q8: Is the compound in Cramer structural class III? (Uses Cramer decision tree)
- Q10: Is the compound in Cramer structural class II? (Uses Cramer decision tree)



TTC decision tree: TTC thresholds:

- Q4: Does estimated intake exceed TTC of 0.15 $\mu\text{g/day}$?
- Q5: Does estimated intake exceed TTC of 1.5 $\mu\text{g/day}$?
- Q7: Does estimated intake exceed TTC of 18 $\mu\text{g/day}$?
- Q9: Does estimated intake exceed 90 $\mu\text{g/day}$?
- Q11: Does estimated intake exceed 540 $\mu\text{g/day}$?
- Q12: Does estimated intake exceed 1800 $\mu\text{g/day}$?

- Requires user input
- DailyIntake, $\mu\text{g/day}$



Toxtree : more plug-ins available:

- ❑ A decision tree for estimating **skin irritation and corrosion potential**, based on rules published in "The Skin Irritation Corrosion Rules Estimation Tool (SICRET), John D. Walker, Ingrid Gerner, Etje Hulzebos, Kerstin Schlegel, QSAR Comb. Sci. 2005, 24, pp378-384"; Developed by IdeaConsult (Sofia,Bulgaria) for JRC
- ❑ A decision tree for estimating **eye irritation and corrosion potential**, based on rules published in "Assessment of the eye irritating properties of chemicals by applying alternatives to the Draize rabbit eye test: the use of QSARs and in vitro tests for the classification of eye irritation, Ingrid Gerner, Manfred Liebsch & Horst Spielmann, Alternatives to Laboratory Animals, 2005, 33, pp. 215-237"; Developed by Ivanka Tsakovska for JRC.
- ❑ **Cramer rules with extensions**: This plug-in is a copy of the original plug-in, plus minor extensions. Like the Cramer plug-in, this plug-in works by assigning compounds to Class I, II, or III, according to the rules from Cramer, and some extra ones. Several compounds were classified by Munro in 1996 as Class I or Class II compounds according to the Cramer rules, even though Munro reported low NOEL values upon oral administration (indicating relatively high toxicity). To overcome such misclassifications, five rules have been introduced to capture the possible toxicity of these compounds;
- ❑ **Structure Alerts for the in vivo micronucleus assay in rodents**, based on the rules, published in the document Development of structural alerts for the in vivo micronucleus assay in rodents by Romualdo Benigni, Cecilia Bossa, Olga Tcheremenskaia and Andrew Worth, European Commission report EUR 23844 EN. Developed by ISS , (Rome, Italy)
- ❑ **START (Structural Alerts for Reactivity in Toxtree) biodegradation and persistence plug-in** is based on a compilation of structural alerts for environmental persistence and biodegradability. These structural alerts are molecular functional groups or substructures that are known to be linked to the environmental persistence or biodegradability of chemicals. The rulebase utilizes the structural alerts in logical decision trees. If one or more the structural alerts embedded in the molecular structure of the chemical are recognized, the system flags the potential persistence or biodegradability of the chemical. Developed by Molecular Networks GmbH (Erlangen, Germany) for JRC
- ❑ **Michael Acceptors** Identifies Michael Acceptors by Structural Alerts as in Schultz 2007 T. Wayne Schultz, Jason W. Yarbrough, Robert S. Hunter, Aynur O. Aptula (2007) Verification of the Structural Alerts for Michael Acceptors. Chem. Res. Toxicol. 20, 1359-1363. Developed by IdeaConsult (Sofia,Bulgaria) for AMBIT project

Creating new decision schemes:

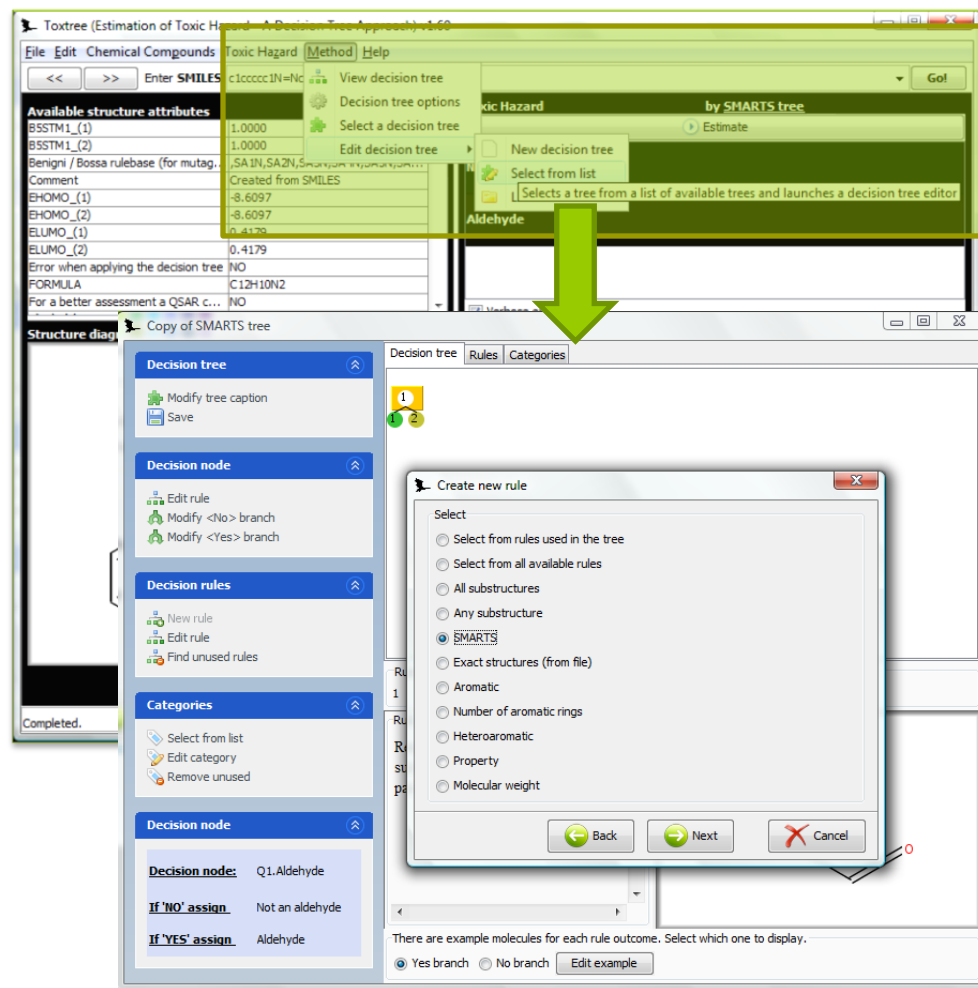
New rules easy to be added by end users

□ End-users

- Graphical Decision Tree editor
- Predefined rules
- Structure alerts defined by SMARTS patterns

□ Advanced

- Create new rules or new plug-ins by Java code
- Easy extension: drop the new jar file into extension folder
- Embed Toxtree module in other standalone or web application



Toxtree : summary

- ❑ 10 plugins
- ❑ Structure alerts defined by SMARTS patterns
- ❑ Descriptors calculation
- ❑ Explanations and examples for each rule outcome
- ❑ Improved visualization and reporting
- ❑ Can be embedded in other standalone and web applications
- ❑ New rules easy to be added by end users
- ❑ New modules can be created by end users or programmers

The screenshot displays the 'Cramer rules' software interface. At the top, a decision tree is visualized with nodes numbered 1 through 33. A blue arrow points to node 17, which is highlighted. Below the tree, the 'Decision node' is identified as 'Q17. Readily hydrolysed to a common terpene'. The interface also shows the logic for 'If 'NO' go to' (Q.19) and 'If 'YES' go to' (Q.[Terpene]). A table lists the rule ID (17) and rule title ('Readily hydrolysed to a common terpene'). The 'Rule explanation' section states: 'if the answer is YES, treat the hydrolysed residues separately and proceed to Q.18 for the terpene moiety and to Q.19 for any non-terpenoid moiety.'. To the right, an 'Example with answer 'YES'' shows a chemical structure of a terpene derivative. At the bottom, there is a section for selecting which example molecule to display, with radio buttons for 'Yes branch' (selected) and 'No branch'.

Cramer rules

Decision node: Q17. Readily hydrolysed to a common terpene

If 'NO' go to: Q.19

If 'YES' go to: Q.[Terpene]

Rule ID	Rule title
17	Readily hydrolysed to a common terpene

Rule explanation

if the answer is YES, treat the hydrolysed residues separately and proceed to Q.18 for the terpene moiety and to Q.19 for any non-terpenoid moiety).

Example with answer 'YES'

CC(C)C1=CC=C(C=C1)C(OC)OC

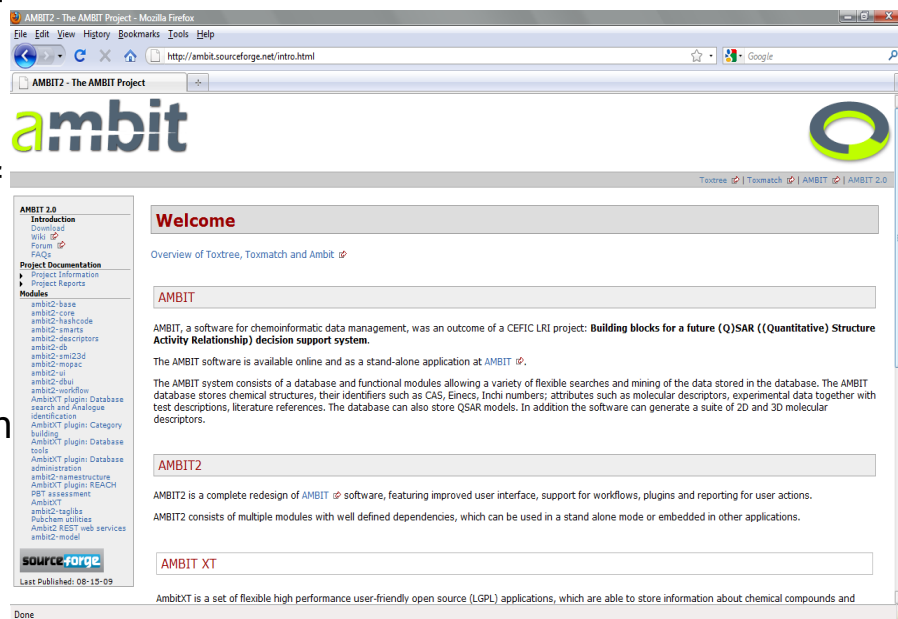
There are example molecules for each rule outcome. Select which one to display.

☒ Yes branch ☐ No branch

AMBIT

<http://ambit.sourceforge.net>

- ❑ Ambit is open source software for chemoinformatics data management, allowing storage of large number of chemical structures and toxicological data and providing flexible means for exploration of structural and similarity spaces
- ❑ Objectives:
 - Develop an open source user friendly software, providing a set of functionalities to facilitate registration of the chemicals for REACH.
 - Improve the user friendliness by introduction of workflow capabilities
 - Develop a set of defined workflows for analogue identification, category building, read across and PBT assessment.
- ❑ Close collaboration with industry
- ❑ JAVA implementation
- ❑ LGPL license
- ❑ Composed of several modules



- ❑ Applications:
 - **Standalone: AmbitXT**
- ❑ **WWW**
 - **AMBIT REST Web services**

AMBIT

Database design and content

- ❑ Generic structure, allowing to store chemical structures in arbitrary format and with arbitrary number and type of properties and descriptors
 - Properties are stored as name-value pairs
 - Support for tuples (set of related values – e.g. test study conditions and results)
 - User defined templates – the user can set a special meaning to any set of properties (e.g. properties X,Y,Z characterize skin irritation experiments)
- ❑ Data provenance – where the data came from, who imported it, Literature reference for each data item
- ❑ Fast (sub)structure and similarity searching
- ❑ Calculation of descriptors
 - By CDK, AMBIT, OpenMOPAC , ToxTree

Data: substances from following sources:

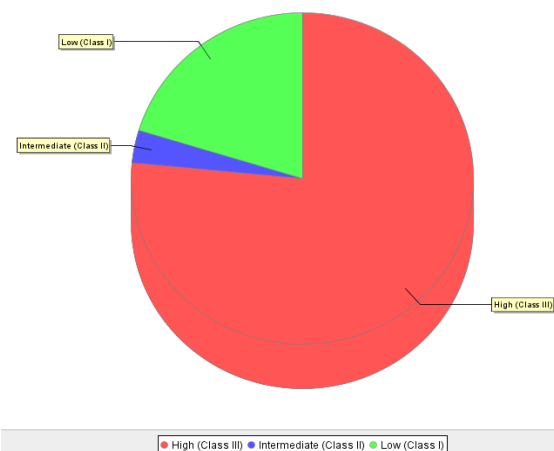
- ❑ [EINECS list](#)
- ❑ [Bioconcentration factor \(BCF\) Gold Standard Database](#)
- ❑ ECETOC Aquatic Toxicity (EAT) Database Supplement to ECETOC. 2003. Aquatic Hazard Assessment II. Technical Report No. 91. European Centre for Ecotoxicology and Toxicology of Chemicals, Brussels, Belgium
- ❑ Compilation of historical local lymph node assay data for the evaluation of skin sensitization alternatives. Gerberick GF, Ryan CA, Kern PS, Schlatter H, Dearman RJ, Kimber I, Patlewicz G, Basketter DA. (2005). Dermatitis 16(4): 157-202.
- ❑ Skin irritation and corrosion Reference Chemicals data base (1995) ECETOC Technical Report No. 66

AMBIT -

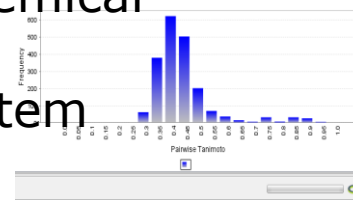
easily compare and characterize datasets

Number of compounds	EINECS	CPDBAS	LLNA	Munro	IRISTR
EINECS	100204	1071	522	542	491
CPDBAS		1547	134	300	222
LLNA			522	108	65
Munro				612	266
IRISTR					544

Distribution of Cramer classes in EINECS list compounds

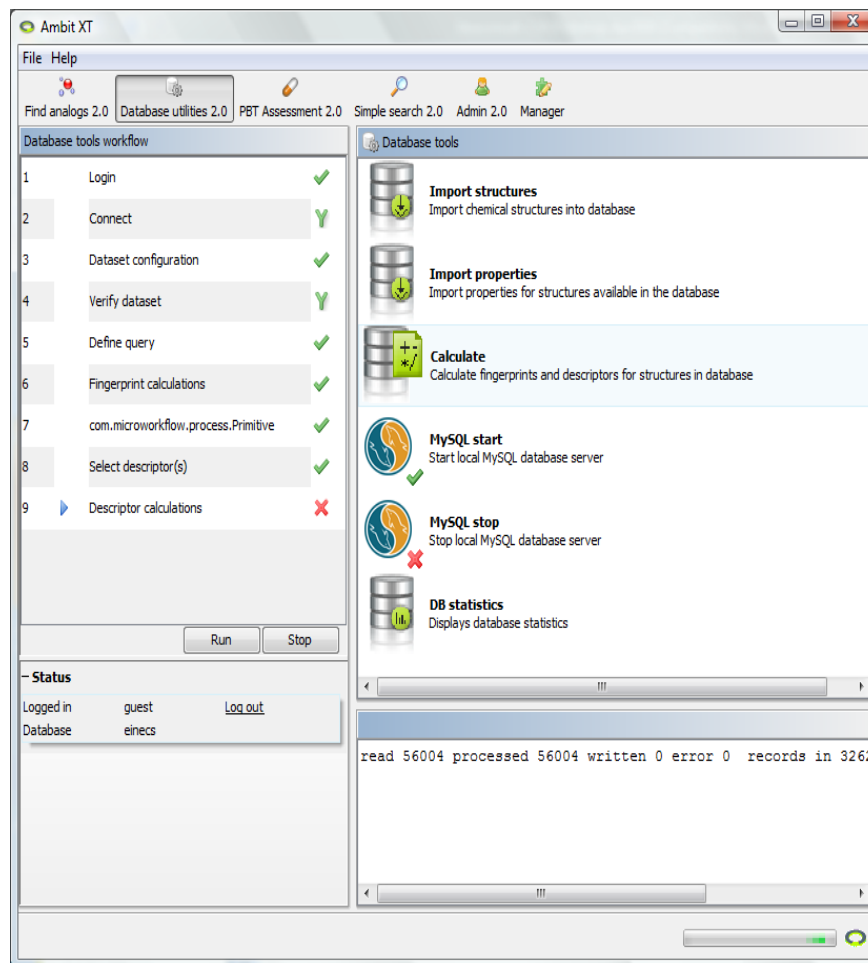


- EINECS
- DSSTOX CPDBAS: Carcinogenic Potency Database Summary Tables - All Species
- LLNA (Gerberick 2005. Dermatitis 16(4): 157-202.)
- Munro reference database (Munro 1996, Food and Chemical Toxicology 34 (1996) 829-867)
- DSSTOX IRISTR: EPA Integrated Risk Information System (IRIS) Toxicity Review Data

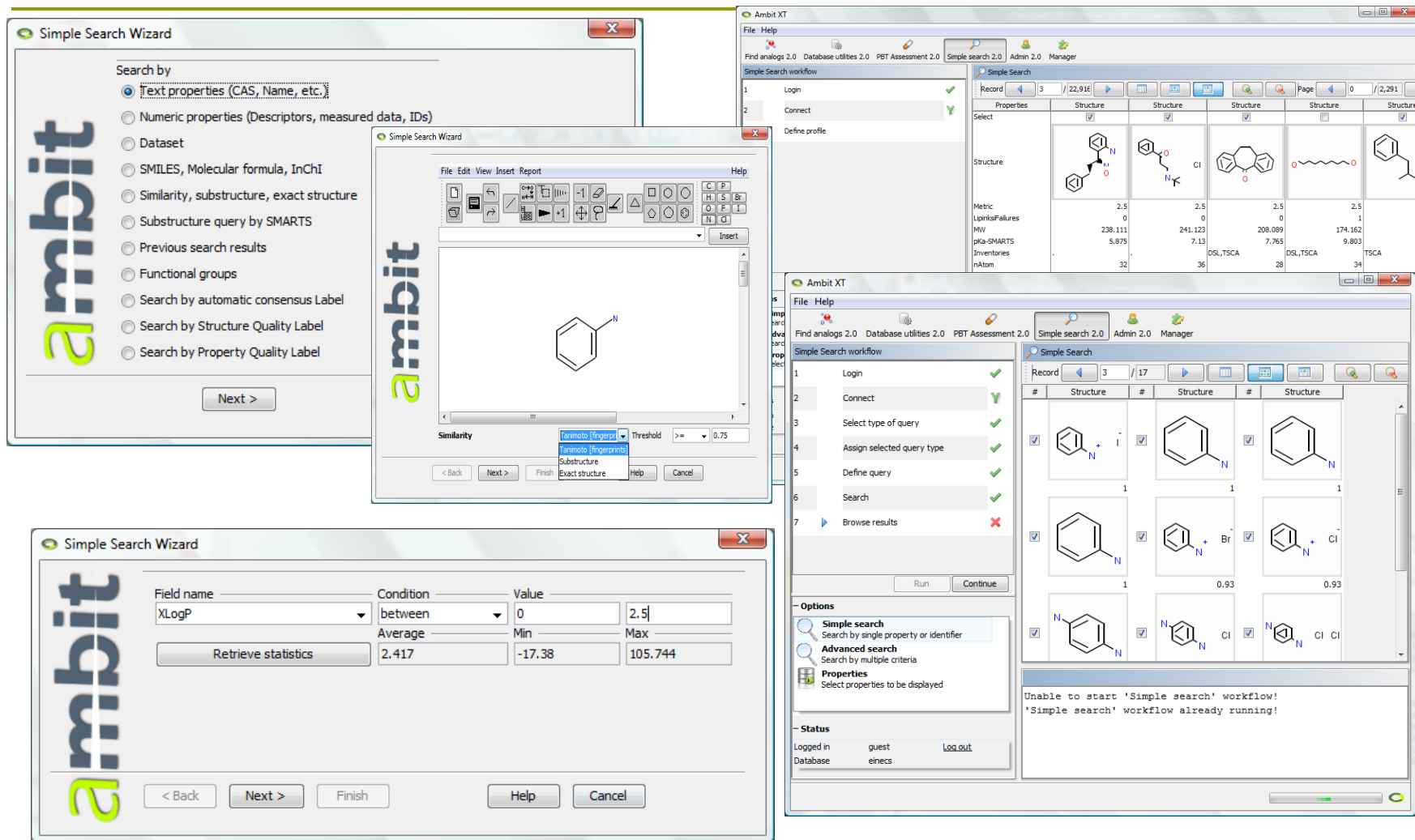


AMBIT XT – workflow support

- ❑ A standalone application (GUI for AMBIT 2.0)
- ❑ Data provenance
 - history of the updates of the chemicals information.
- ❑ Data quality
 - Easy way for comparison between different sources
- ❑ Flexible storage for measured data for different endpoints
 - Easy way to extract all relevant information for a chemical; many formats available for toxicological data
- ❑ Recording of user actions
- ❑ Easy entry of complex structural alerts to facilitate grouping
- ❑ Molecular descriptors
- ❑ Improved data entrance and visualization
- ❑ Embedded workflow engine
- ❑ Modular application (flexible plug-in support)



AMBT XT – Search facilities



The image displays four screenshots of the AMBT XT software interface, illustrating its search capabilities.

Top Left: Simple Search Wizard

Search by:

- ☒ Text properties (CAS, Name, etc.)
- ☐ Numeric properties (Descriptors, measured data, IDs)
- ☐ Dataset
- ☐ SMILES, Molecular formula, InChI
- ☐ Similarity, substructure, exact structure
- ☐ Substructure query by SMARTS
- ☐ Previous search results
- ☐ Functional groups
- ☐ Search by automatic consensus Label
- ☐ Search by Structure Quality Label
- ☐ Search by Property Quality Label

Next >

Top Right: Simple Search Wizard (Detailed)

File Edit View Insert Report Help

Simple Search workflow

1 Login ☒

2 Connect ☒

Define profile

Simple Search

Record 3 / 22,916 Page 0 / 2,291

Select

Structure

Metric

LipinskiFailures

MW

pKa-SMARTS

Inventories

nAtom

2.5

0

5.875

32

2.5

241.123

7.13

36

2.5

208.089

7.765

28

2.5

174.162

9.803

TSCA

34

Bottom Left: Simple Search Wizard (Field Selection)

Field name

XLogP

Condition

between

Value

0

2.5

Average

Min

Max

2.417

-17.38

105.744

Retrieve statistics

< Back Next > Finish Help Cancel

Bottom Right: Simple Search Wizard (Results)

File Edit View Insert Report Help

Simple Search workflow

1 Login ☒

2 Connect ☒

3 Select type of query ☒

4 Assign selected query type ☒

5 Define query ☒

6 Search ☒

7 Browse results ☒

Run Continue

Options

Simple search

Search by single property or identifier

Advanced search

Search by multiple criteria

Properties

Select properties to be displayed

Status

Logged in

guest

einecs

Log out

Unable to start 'Simple search' workflow!

'Simple search' workflow already running!

A workflow in AMBIT XT

2. Load data

Properties related to the selected endpoint

4. Find similar substances

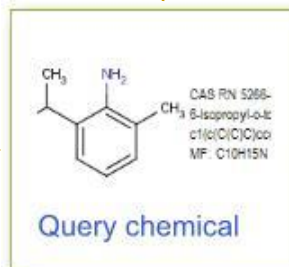
Substances with similar structure and descriptor values

5. Add more substances

Specify similar substances by hand

1. Load chemical

Search by identifier or specify structure



3. Predict properties

Calculate descriptors and properties not available in the database

Simple Search interface showing a grid of chemical structures.

Structure	Structure	Structure

6. Load data

Retrieve properties related to the selected endpoint

7. Analyse data

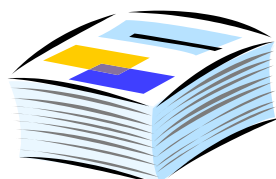
Calculate probability that the datapoint is questionable and further testing might be valuable

Simple Search interface showing a table of results.

>	#	Select	Structure	fc		
	390	<input checked="" type="checkbox"/>		1571	1	1571
				1583	1	1583
				1590	1	1590
				1593	1	1593
				1595	1	1595
	395	<input checked="" type="checkbox"/>		1601	1	1601
				1606	1	1606
				1612	1	1612
				1613	1	1613
				1622	1	1622
				1625	1	1625
	401	<input checked="" type="checkbox"/>		1631	1	1631
				1632	1	1632
				1634	1	1634
				1638	1	1638

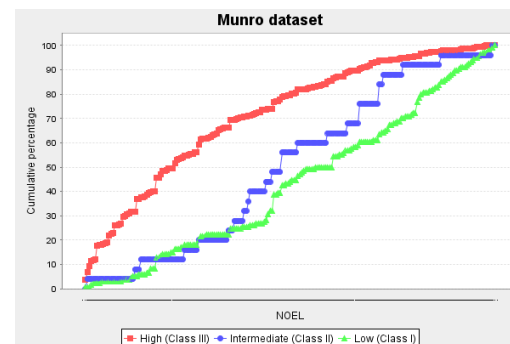
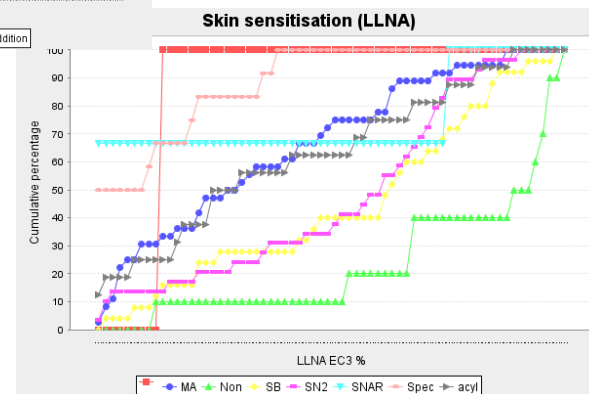
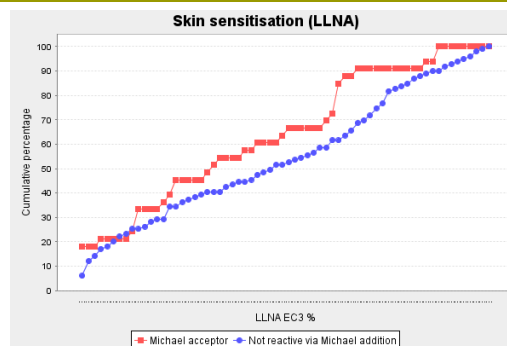
8. Generate report

Data matrix and probabilities



AMBIT – TTC use case

- Use existing data or import a dataset (multiple file formats supported)
 - E.g. Munro Reference Database
- Run descriptor calculations or classification schemes or select existing classification
 - E.g. Toxtree Cramer rules; Toxtree Michael Acceptors module
- Generate cumulative dose distribution per class



Ambit XT

File Options Help

Simple search Profile & Category REACH/PBT Assessment Database utilities Admin

Simple Search workflow

1. Login
2. Connect
3. Define profile

Run Stop

Search

Search by single property or identifier

Advanced search

Search by multiple criteria

Refine search results

Search within last results

Browse previous search results

Browse results from previous queries

Properties

Select properties to be displayed

Export results

Export results as PDF/RTF/HTML/SDF files

Statistics

Generate search results statistics (see Log)

Record 50 / 612

Page 4 / 51

12

Find

Structure

Structure

Structure

Structure

Structure

Structure

Metric

EC

CAS

Cramer ...

NAME

NOEL

EC

Consens...

Source

Quality I...

79277-27-3

High (Class III)

COC(=O)C1=C(C=CS(S(=O)(=O)N1C1=CC=CC=C1)C1=CC=CC=C1)C1=CC=CC=C1

1

Unconfirmed[1]

Munro database

69806-40-2

High (Class III)

COC(=O)C(C)OC1=C(C=CC(=O)N1C1=CC=CC=C1)C1=CC=CC=C1

1

Unconfirmed[1]

Munro database

69409-94-5

High (Class III)

CC(C)C(NC1=C(C)C=C(C=C1)C1=CC=CC=C1)C1=CC=CC=C1

1

Unconfirmed[1]

Munro database

66332-96-5

High (Class III)

CC(C)OC1=C(C)C(NC1=CC(=O)C2=CC=CC=C2)C1=CC=CC=C1

1

Unconfirmed[1]

Munro database

56425-91-3

High (Class III)

CC(C)C(C)C1=C(C)C(NC1=CC(=O)C2=CC=CC=C2)C1=CC=CC=C1

2

Unconfirmed[1]

Munro database

909

High (C)

CCOC(=O)C1=C(C)C(NC1=CC(=O)C2=CC=CC=C2)C1=CC=CC=C1

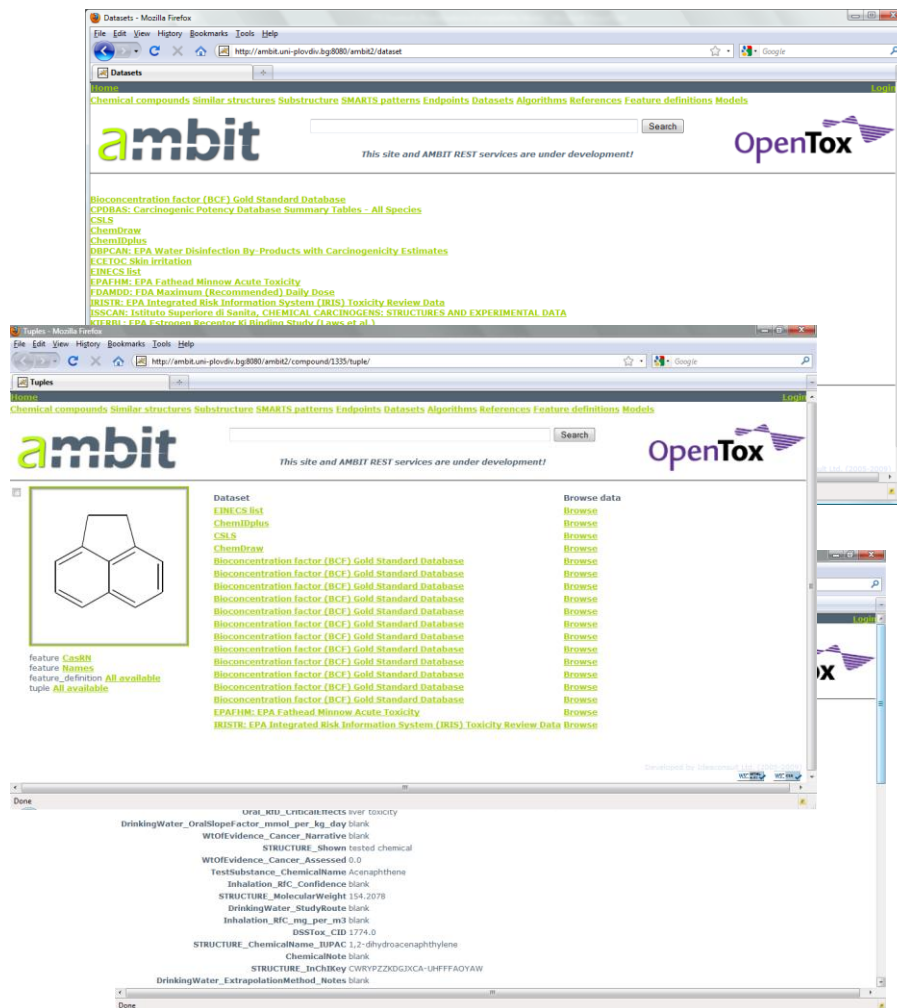
1

Unconfi...

Munro di...

AMBIT Web services under development

- ❑ **Ambit Web services**
 - Data aggregation
 - Descriptor calculation
 - Model development
 - Model validation and comparison
 - Easy access by human experts (web applications) and software (web services)
- ❑ Based on Web services API, developed by 11 partners from EC FP7 OpenTox project
- ❑ OpenTox <http://opentox.org> aims at an open source predictive toxicology framework with a unified access to toxicological data, (Q)SAR models and supporting information
 - A simple and intuitive interface for toxicological experts, that provides unified access to (Q)SAR predictions, toxicological data, (Q)SAR models and supporting information
 - An expert interface for the streamlined development and validation of new (Q)SAR models
 - An application programming interface (API) for the development, integration and validation of new (Q)SAR algorithms



TTC estimation online

Chemical compound - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://ambit.uni-plovdiv.bg:8080/ambit2/ttc?property=&search=adamantane

Chemical compound

Home Login

FastTox TTC Depiction Datasets Chemical compounds Similar structures Substructure SMARTS patterns Endpoints Algorithms References Feature definitions Models Help

ambit

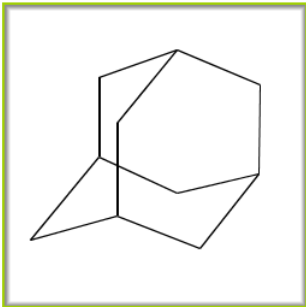
adamantane Search

Search by property or identifier name (optional) and value
This site and AMBIT REST services are under development!

openTox

Query name: Property like adamantane Save search results

☒



CAS RN
EINECS
Chemical name(s)
All available feature values
Feature values by groups
Feature values by dataset
Model predictions

Next

Done

TTC estimation online

TTC: Enter Daily Intake - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://ambit.uni-plovdiv.bg:8080/ambit2/ttc/input?compound[]=5456&idstructure=5456&idchemical=5456

TTC: Enter Daily Intake

Home Login

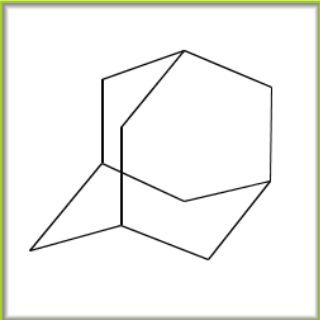
[FastTox](#) [TTC](#) [Depiction](#) [Datasets](#) [Chemical compounds](#) [Similar structures](#) [Substructure](#) [SMARTS patterns](#) [Endpoints](#) [Algorithms](#) [References](#) [Feature definitions](#) [Models](#) [Help](#)

ambit

This site and AMBIT REST services are under development!

OpenTox

☒



[CAS RN](#)
[EINECS](#)
[Chemical name\(s\)](#)
[All available feature values](#)
[Feature values by groups](#)
[Feature values by dataset](#)
[Model predictions](#)

ToxTree: ILSI/Kroes decision tree for Threshold for Toxicological Concern (TTC) estimation

Classifies the compound into one of three classes:

1. Substance would not be expected to be a safety concern
2. Negligible risk (low probability of a life-time cancer risk greater than 1 in 10^6)
3. Risk assessment requires compound-specific toxicity data

- Daily Intake value is necessary for an accurate TTC assessment.
- Please enter Daily Intake value and press Update.
- Press Apply to apply TTC decision tree.

DailyIntake $\mu\text{g/day}$

Developed by Ideaconsult Ltd. (2005-2009)

WSC HTML WSC CSS

Done

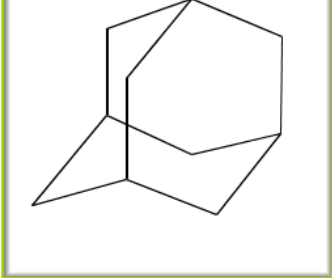
TTC estimation online

Model 9 idchemical=5456 idstructure=-1 - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://ambit.uni-plovdiv.bg:8080/ambit2/compound/5456/model/9

Model 9 idchemical=5456 idstructur...



[CAS RN](#)
[EINECS](#)
[Chemical name\(s\)](#)
[All available feature values](#)
[Feature values by groups](#)
[Feature values by dataset](#)
[Model predictions](#)

[Kroes TTC decision tree#explanation](#)

QAlerts(genotoxic).Verify structural alerts for potential genotoxic carcinogenicity Yes
QSA1.Acyl halides No
QSA2.Alkyl (C<5) or benzyl ester of sulphonic or phosphonic acid No
QSA3.N-methylol derivatives No
QSA4.Monohaloalkene No
QSA5.S or N mustard No
QSA6.Propiolactones and propionsultones No
QSA7.Epoxides and aziridines No
QSA8.Aliphatic halogens No
QSA9.Alkyl nitrite No
QSA11.Simple aldehyde No
QSA12.Quinones No
QSA13.Hydrazine No
QSA14.Aliphatic azo and azoxy No
QSA15.Isocyanate and isothiocyanate groups No
QSA16.Alkyl carbamate and thiocarbamate No
QSA18.Polycyclic Aromatic Hydrocarbons No
QSA19.Heterocyclic Polycyclic Aromatic Hydrocarbons No
QSA21.Alkyl and aryl N-nitroso groups No
QSA22.Azide and triazene groups No
QSA23.Aliphatic N-nitro No
QSA24.α,β unsaturated alkoxy No
QSA25.Aromatic nitroso group No
QSA26.Aromatic ring N-oxide No
QSA27.Nitro aromatic No
QSA28.Primary aromatic amine, hydroxyl amine and its derived esters (with restrictions) No
QSA28bis.Aromatic mono- and dialkylamine No
QSA28ter.Aromatic N-acyl amine No
QSA29.Aromatic diazo No
QSA30.Coumarins and Furocoumarins No
QQ2.Are there structural alerts that raise concern for potential genotoxicity? No
QQ5.Does estimated intake exceed TTC of 1.5 µg/day ? Yes
QQ6.Is the compound an organophosphate? No
QQ8.Is the compound in Cramer structural class III? No
QQ10.Is the compound in Cramer structural class II? No
QQ12.Does estimated intake exceed 1800 µg/day ? Yes Class Risk assessment requires compound-specific toxicity data
Risk assessment requires compound-specific toxicity data

[Kroes TTC decision tree](#)

Done

Thank you!



Questions?