Computer tools for Threshold of Toxicological Concern Elucidation

Dr. Nina Jeliazkova

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

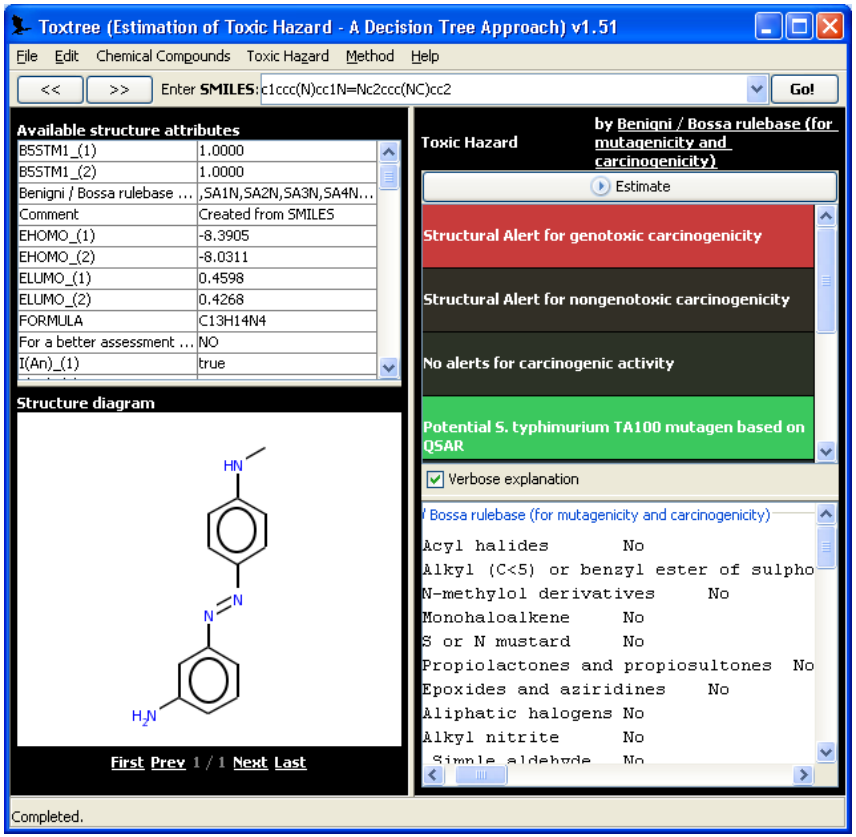
Content

- Toxtree
  - Cramer rules
  - Verhaar scheme for predicting toxicitymode 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



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

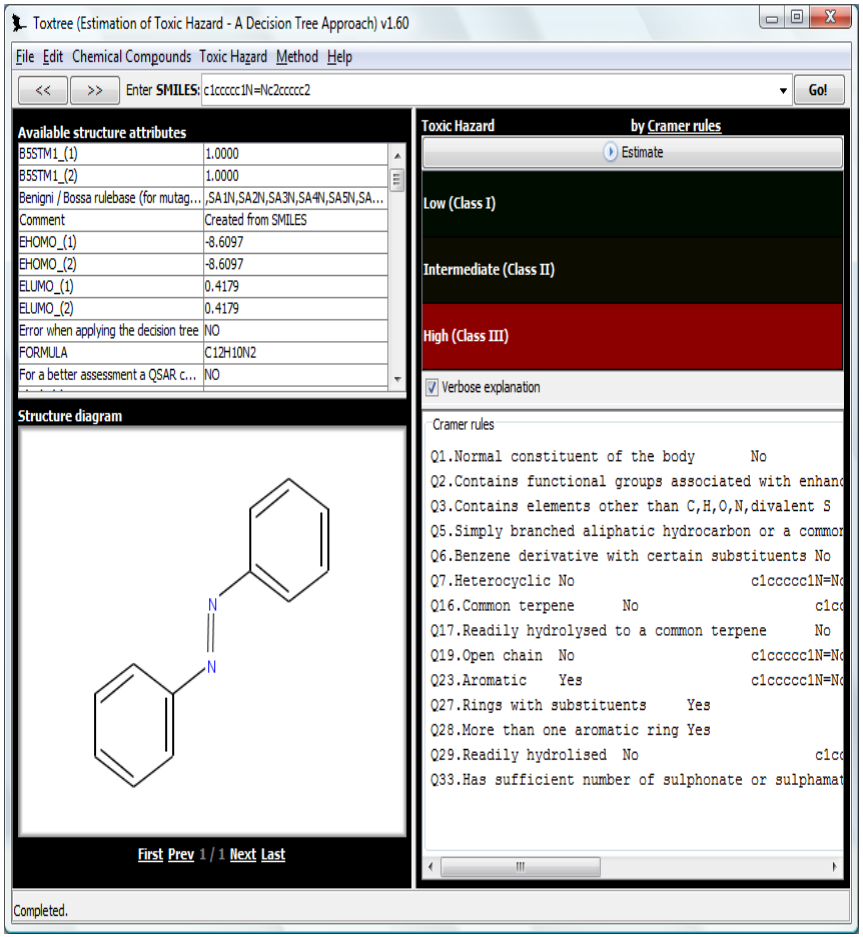
Toxtree: Cramer rules

- 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. Jeliazkova, 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

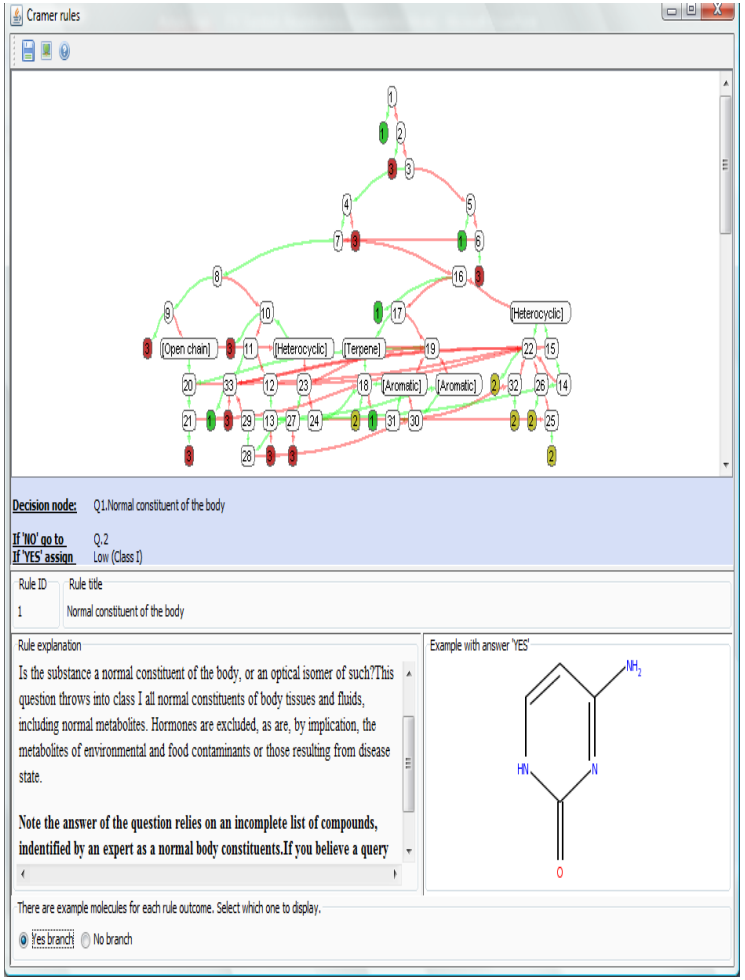


Toxtree: Cramer rules

- 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**

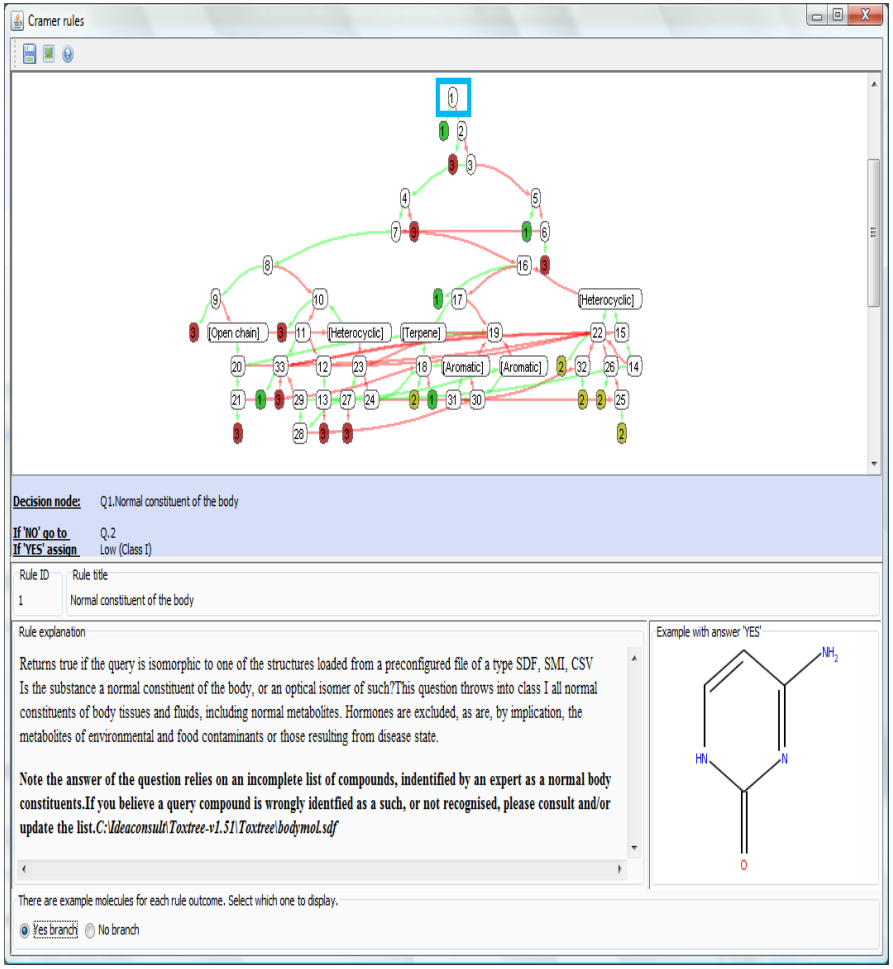
Cramer classification scheme

- 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



Toxtree internals:

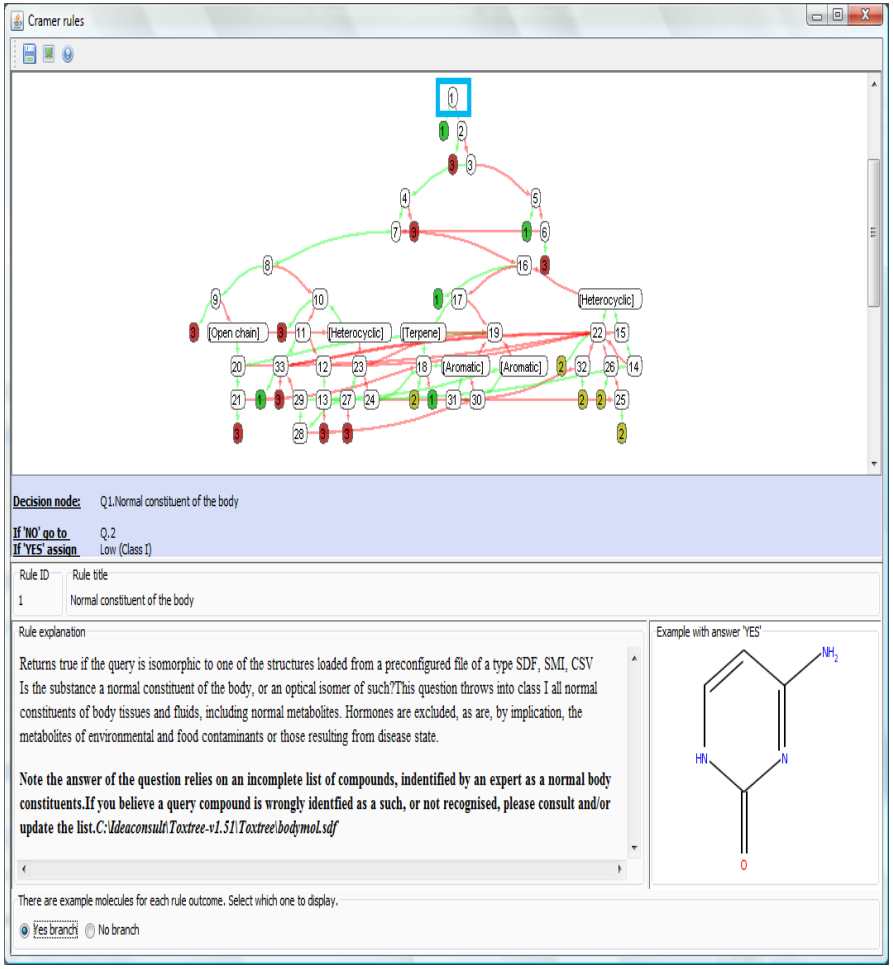
"Normal constituents of the body"



- 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"

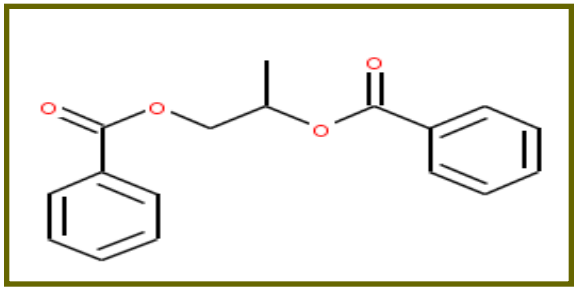


- Rule 22 :"Common component of food"
- File: foodmol.sdf
- 108 compounds

Toxtree internals

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

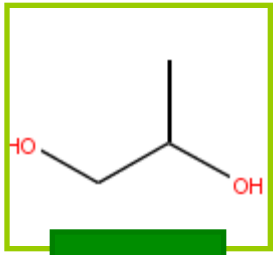
Q1.Normal constituent of the body No
Q2.Contains functional groups associated with enhanced toxicity No
Q3.Contains elements other than C,H,O,N, divalent S No
Q5.Simply branched aliphatic hydrocarbon or a common carbohydrate No
Q6.Benzene derivative with certain substituents No
Q7.Heterocyclic No
Q16.Common 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



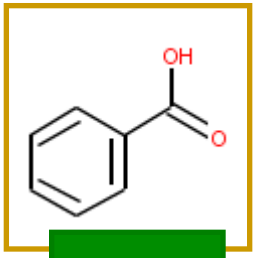
Query compound

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)**

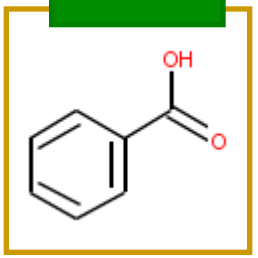
Hydrolysis products



Low (Class I)



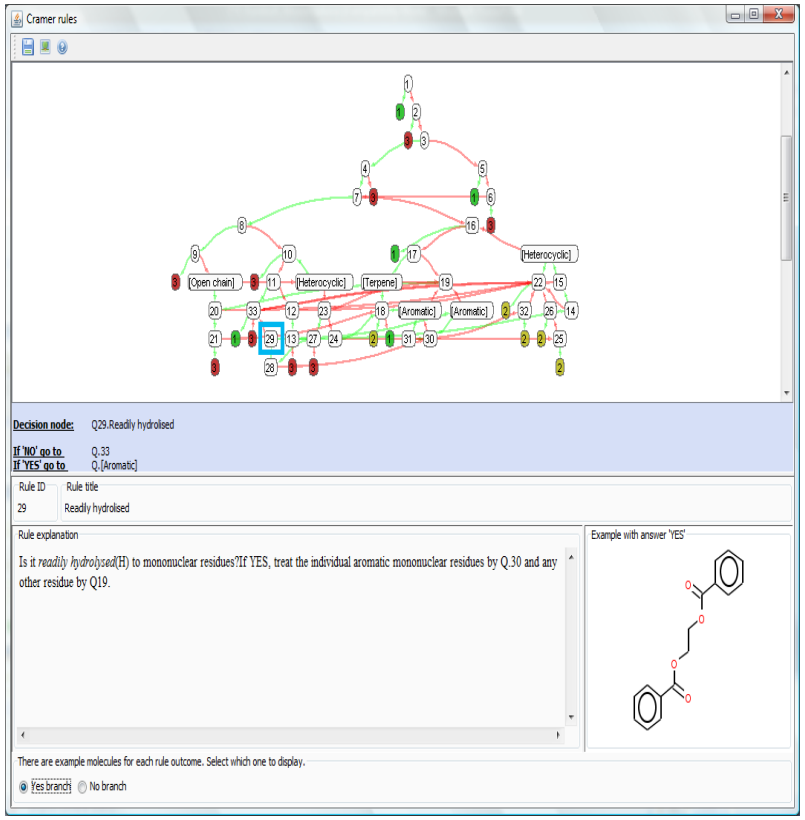
Low (Class I)

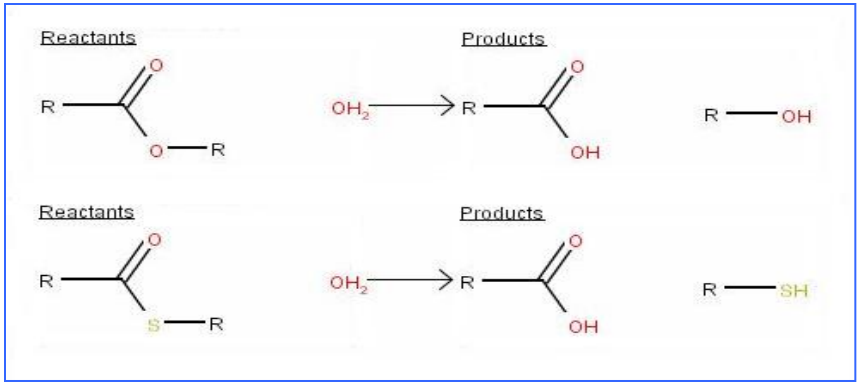


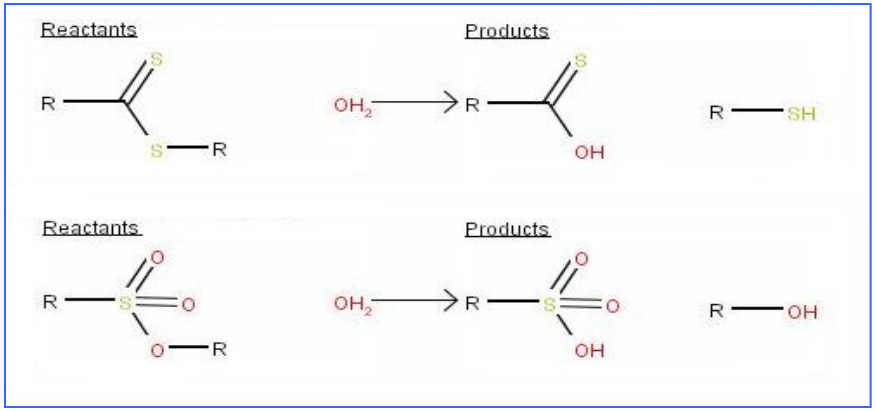
Q30.Aromatic Ring with complex substituents No
Q18.One of the list (see explanation) No **Low (Class I)**

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.



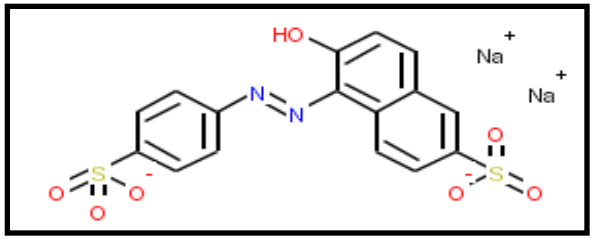




Toxtree internals

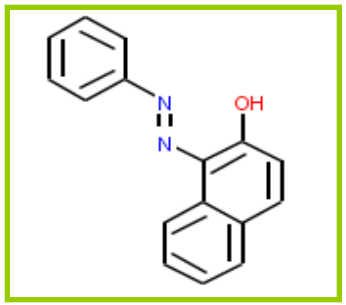
Series of questions are applied on query compound and also on reaction products

- 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



Query compound

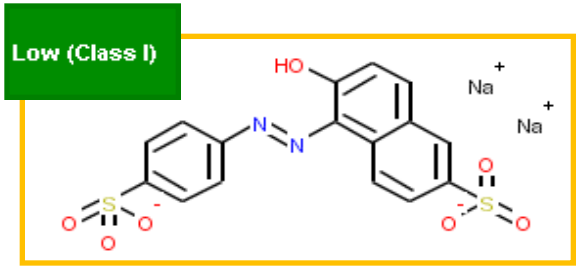
Q7.Heterocyclic No
Q16.Common terpene No
Q17.Readily hydrolysed to a common terpene No
Q19.Open chain No
Q23.Aromatic Yes
Q27.Rings with substituents Yes
Q28.More than one aromatic ring Yes
Q29.Readily hydrolised No



Unsulphonated compound

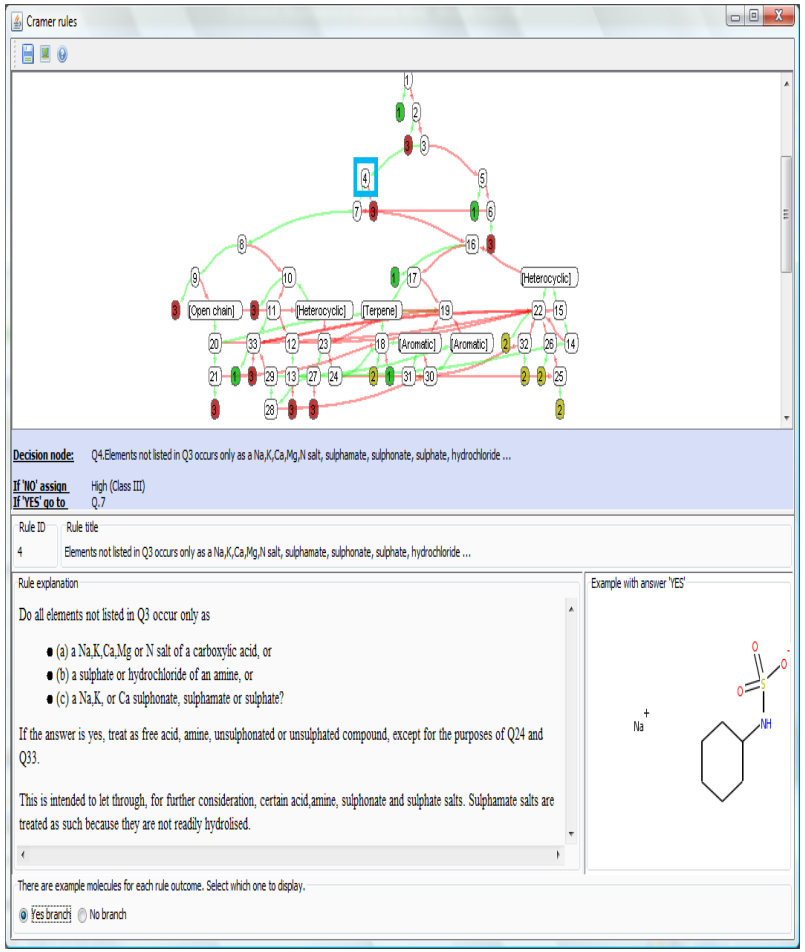
Query compound (again)

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



Toxtree internals

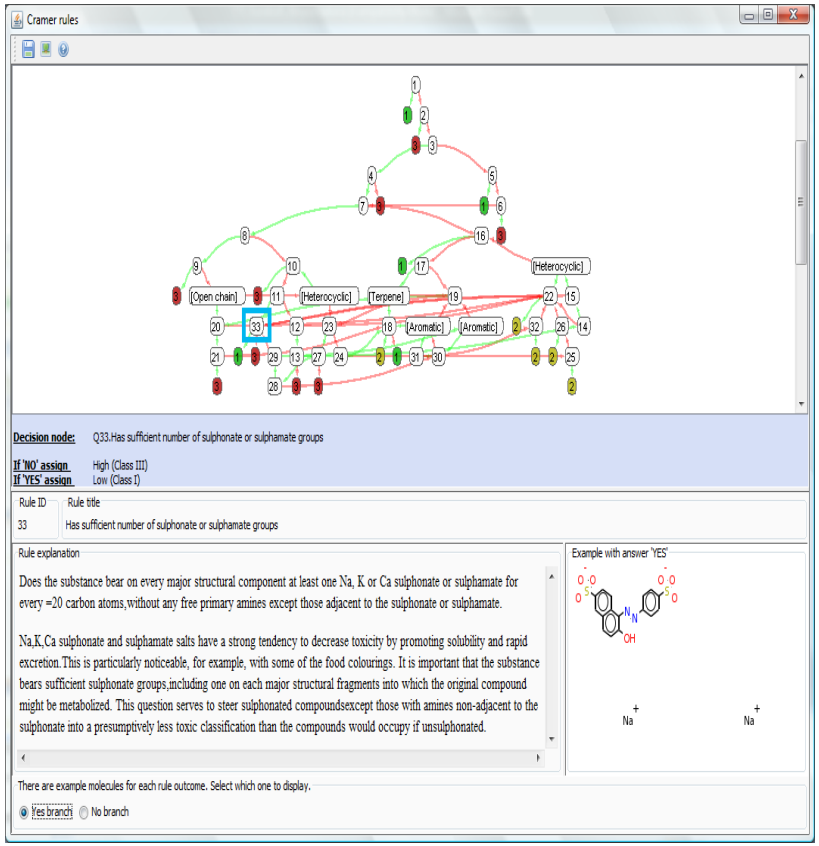
Rule4: Elements not listed in Q3 occurs only as a Na, K, Ca, Mg, N salt, sulphamate, sulphonate, sulphate, hydrochloride

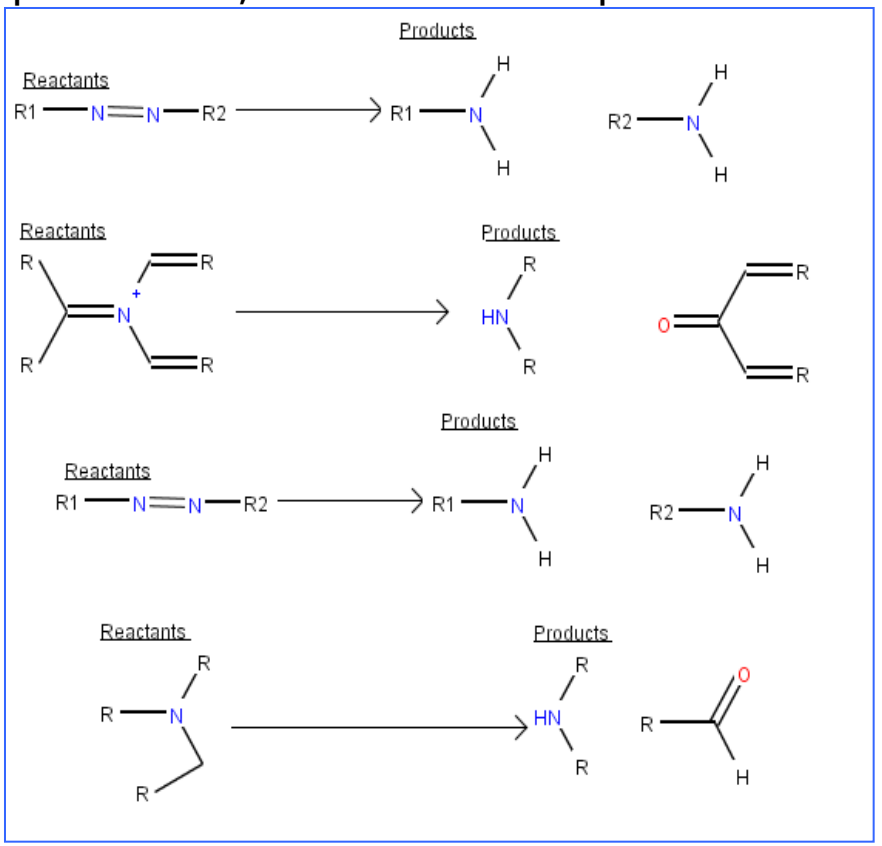


- Do all elements not listed in Q3 occur only as (a) a Na,K,Ca,Mg or N salt of a carboxylic acid, or
- (b) a sulphate or hydrochloride of an amine, or
- (c) a Na,K, or Ca sulphonate, sulphamate or sulphate?
- If the answer is yes, **treat as free acid, amine, unsulphonated or unsulphated compound, except 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 hydrolised.

Toxtree internals: Metabolism

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





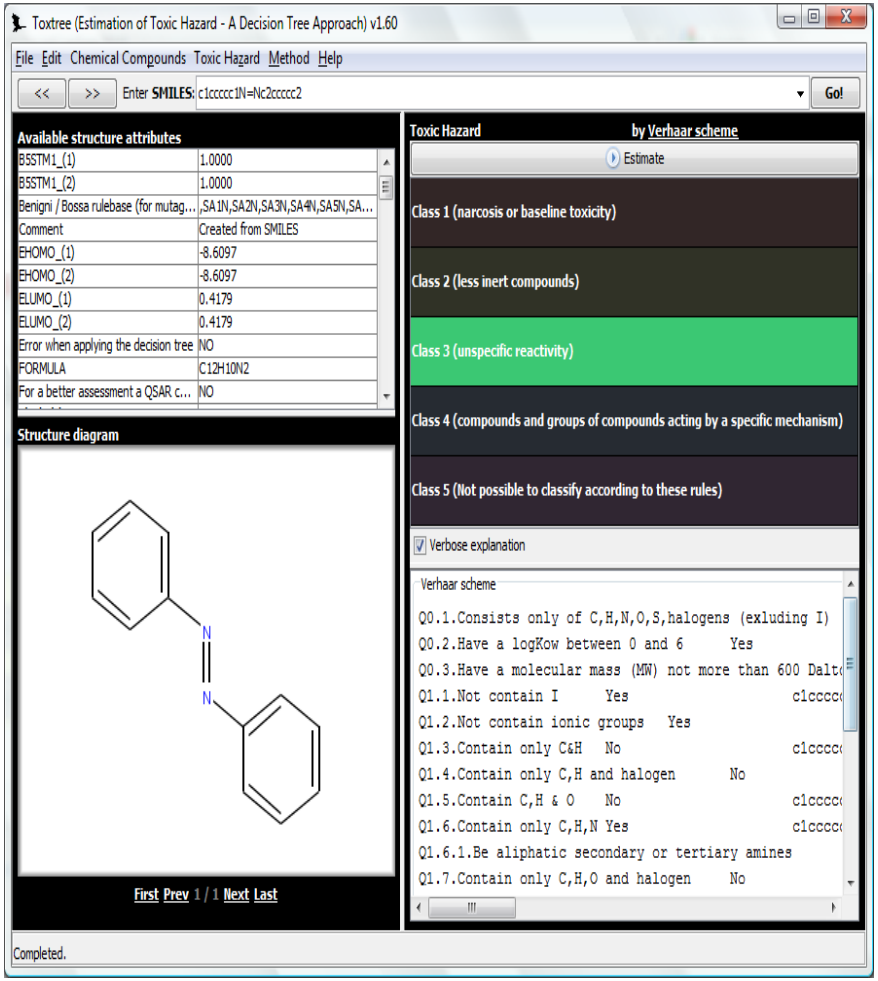
Toxtree: Verhaar scheme

- 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.

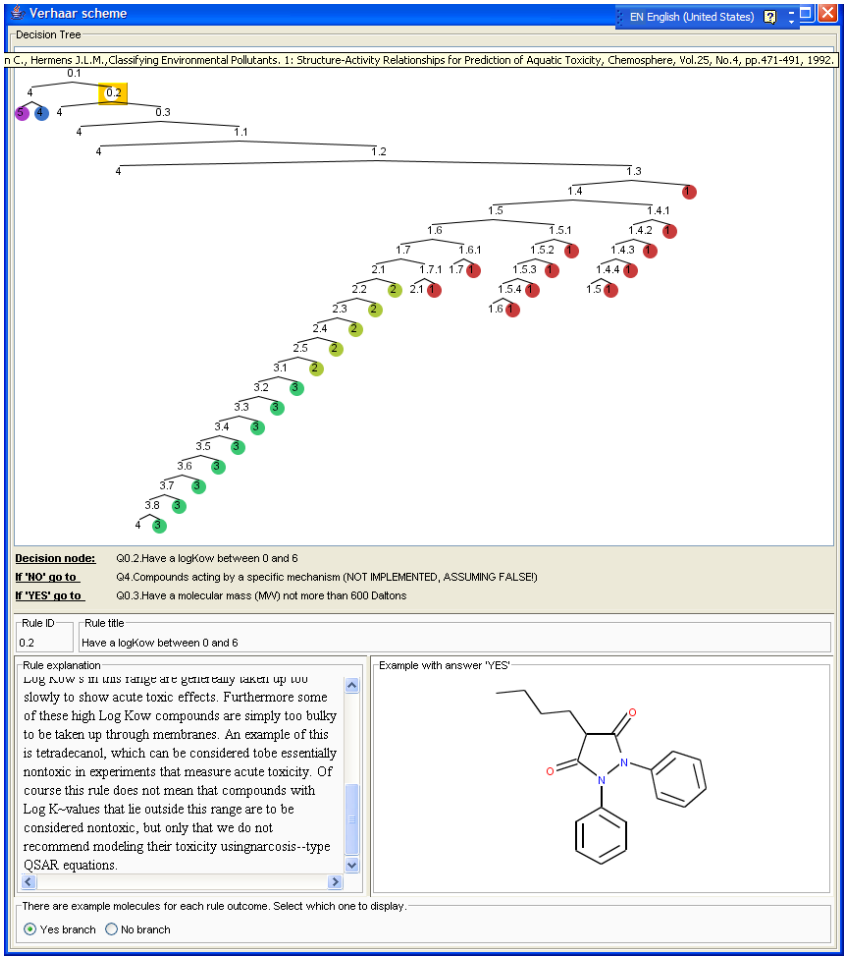
Verhaar, 2000 Verhaar HJM, Solbe J, Speksnijder J, van Leeuwen CJ and Hermens JLM (2000) Classifying environmental pollutants: Part 3. External validation of the classification system. Chemosphere 40, 875-883

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



Toxtree: 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"



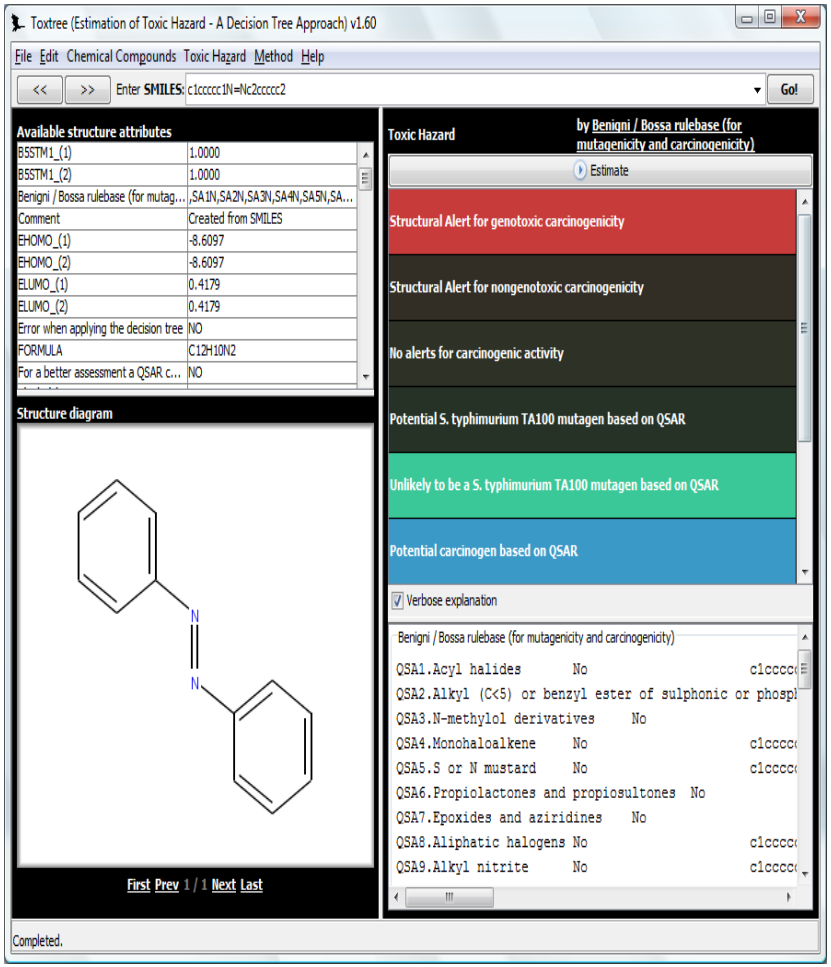
Toxtree: Benigni-Bossa rulebase for carcinogenicity and mutagenicity

- Predicting genotoxic and non-genotoxic carconogenicity 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

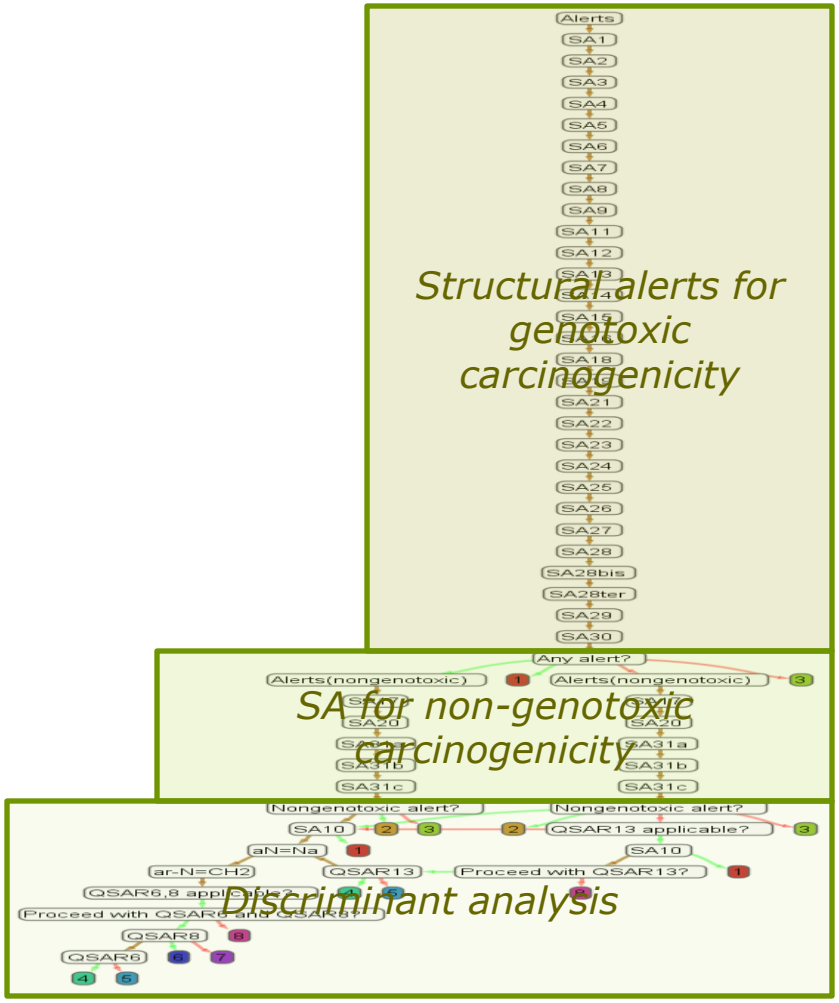
A software tool for predicting mutagenicity and carcinogenicity http://ecb.jrc.it/documents/QSAR/EUR_23241_EN.pdf

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



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

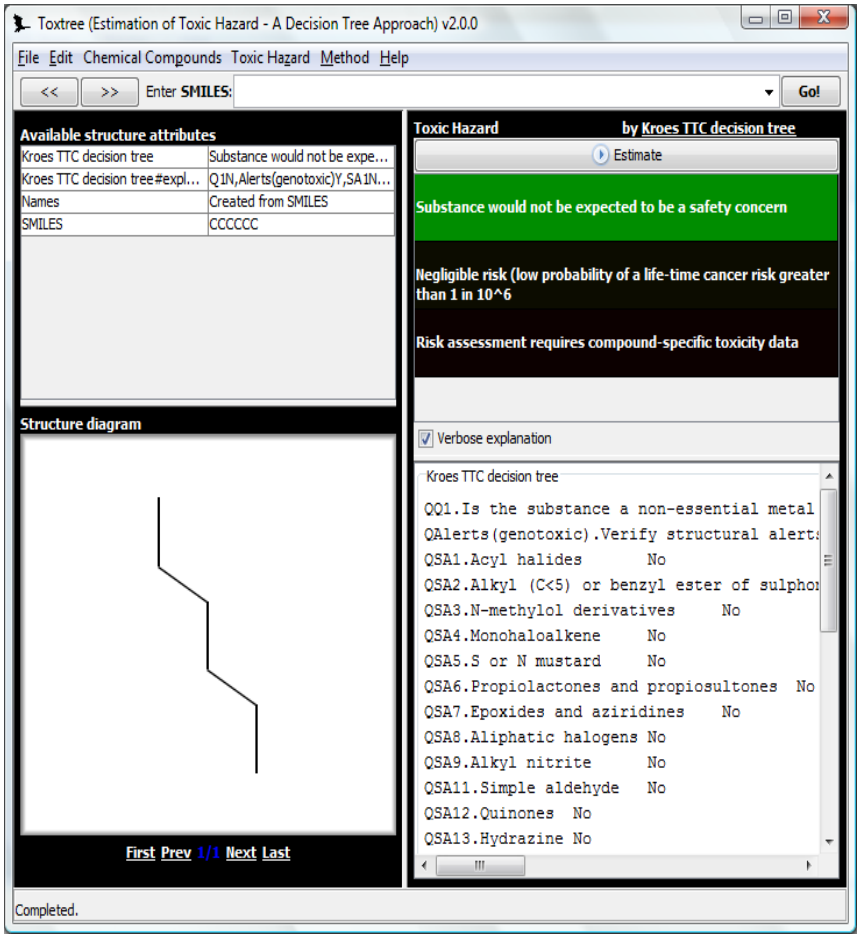


Toxtree: Descriptors calculation

- **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/

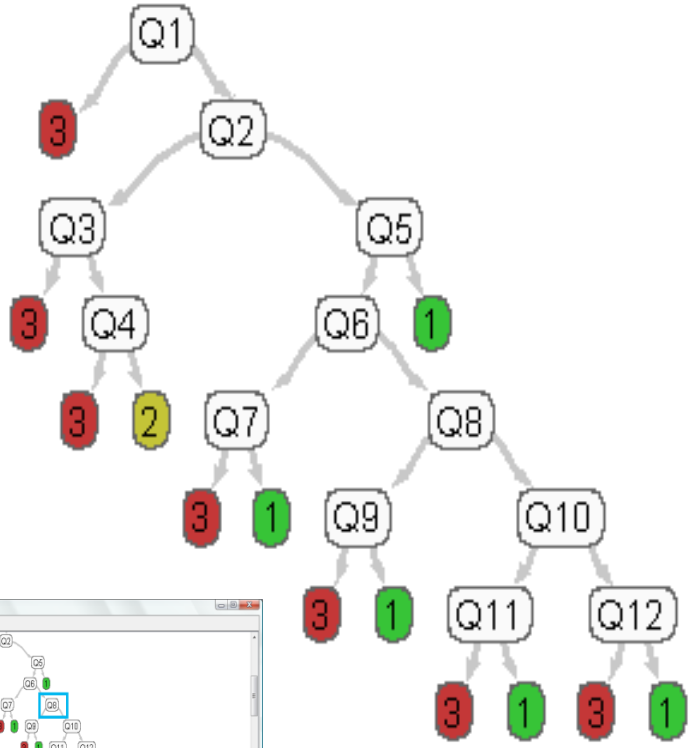
Toxtree: TTC decision tree

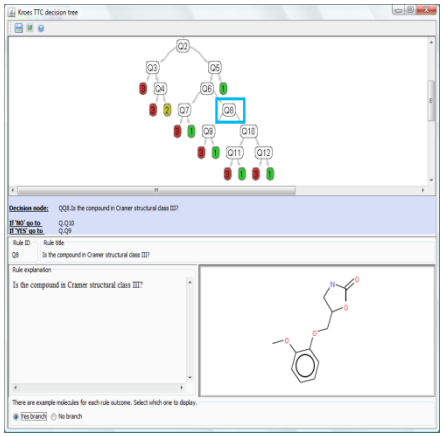
- Decision tree proposed by ILSI Europe to decide whether substances can be assessed by the TTC approach
- **Kroes et al., Food and Chemical Toxicology 42, p76, 2004**
- Software implementation
  - **Beta test in ToxTree v2.00 (2009)**
  - IdeaConsult Ltd.



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 predictiing 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)

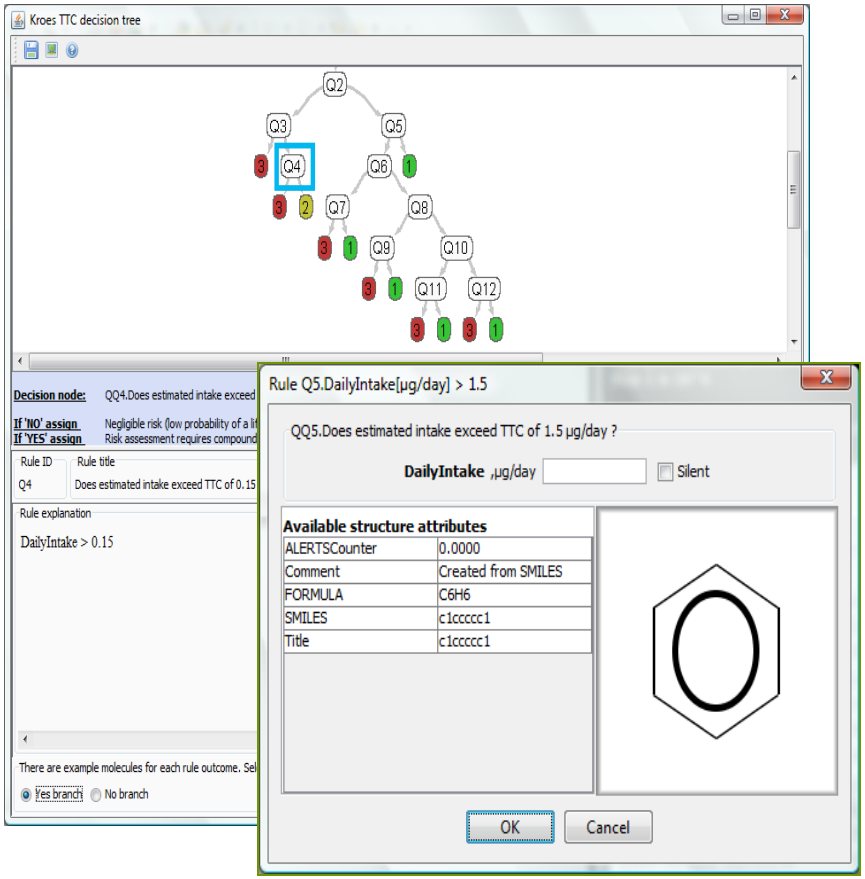




TTC decision tree: TTC thresholds:

- Q4: Does estimated intake exceed TTC of 0.15 µg/day ?
- Q5: Does estimated intake exceed TTC of 1.5 µg/day ?
- Q7: Does estimated intake exceed TTC of 18µg/day ?
- Q9: Does estimated intake exceed 90 µg/day ?
- Q11: Does estimated intake exceed 540 µg/day ?
- Q12: Does estimated intake exceed 1800 µg/day ?

- Requires user input
- DailyIntake, µg/day



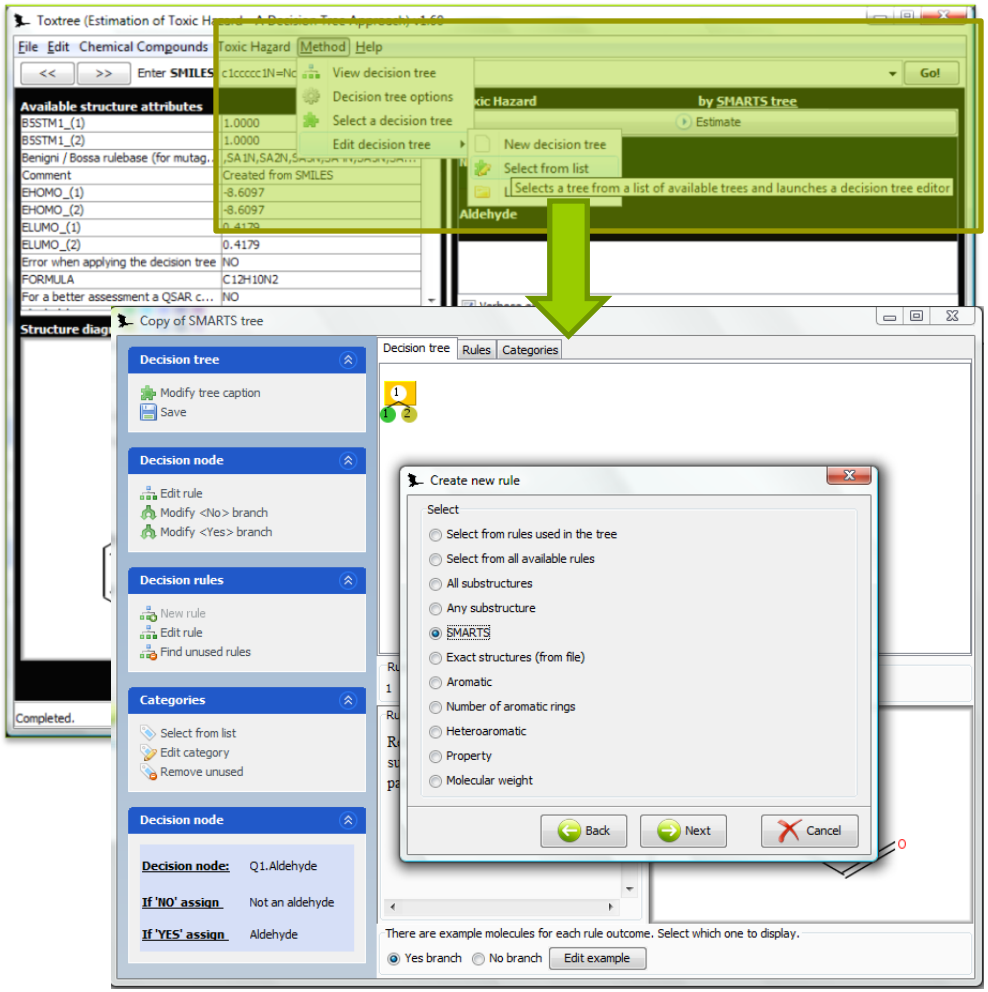
Toxtree : more plug-ins available:

- 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:

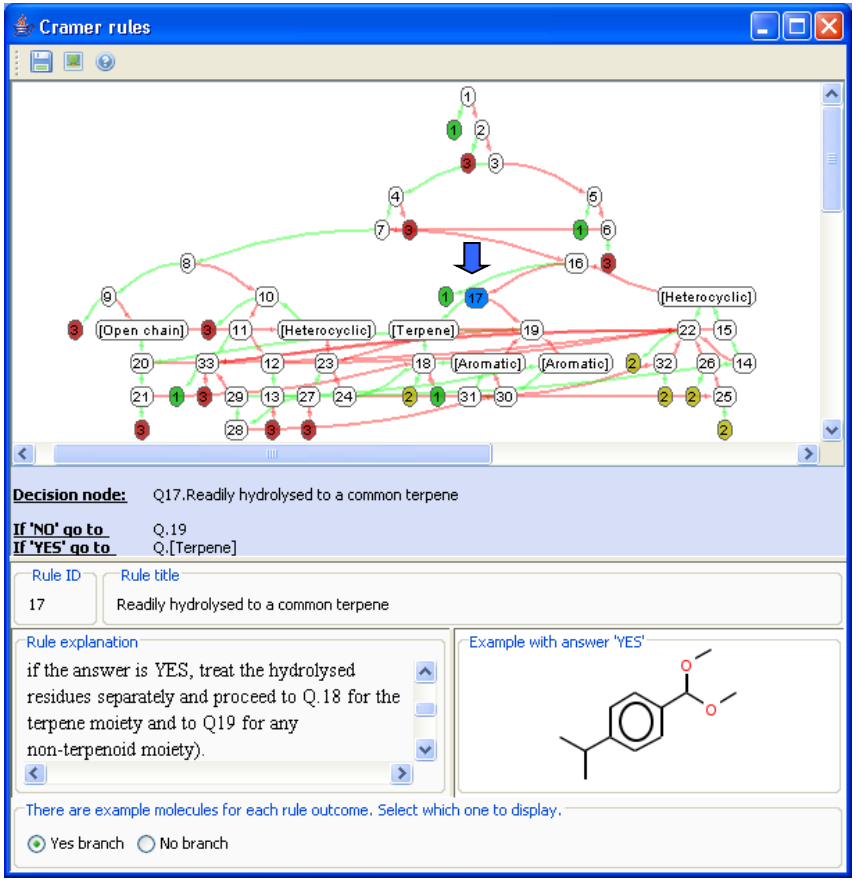
New rules easy to be added by end users

- 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

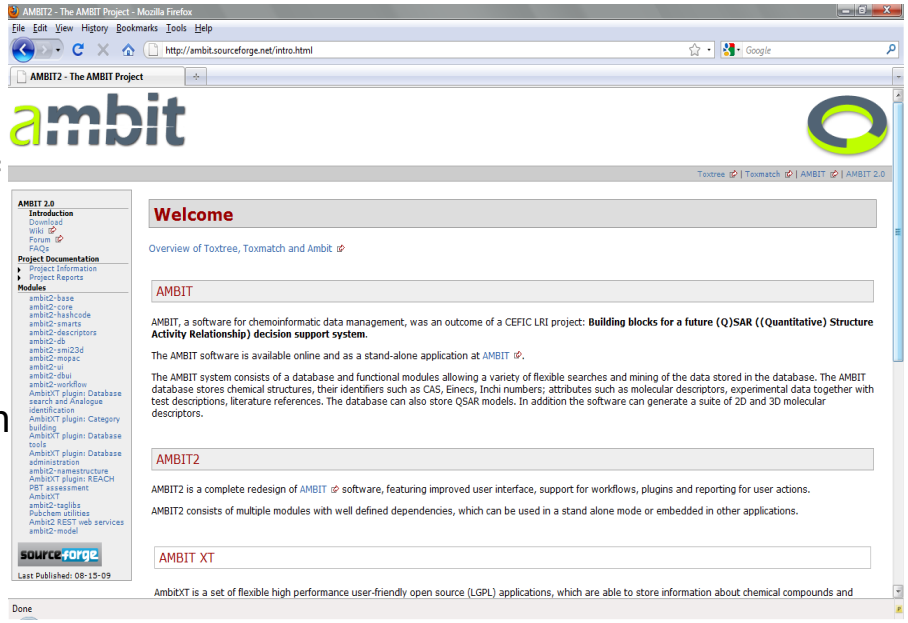
- 10 plugins
- Structure alerts defined by SMARTS patterns
- Descriptors calculation
- Explanations and examples for each rule outcome
- Improved visualization and reporting
- Can be embedded in other standalone and web applications
- New rules easy to be added by end users
- New modules can be created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



AMBIT

http://ambit.sourceforge.net

- Ambit is 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
- 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:

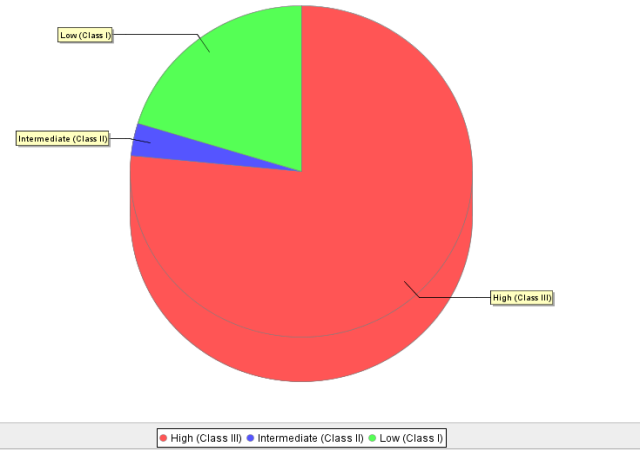
- EINECS list
- Bioconcentration factor (BCF) Gold Standard Database
- ECETOC Aquatic Toxicity (EAT) Database Supplement to ECETOC. 2003. Aquatic Hazard Assessment II. Technical Report No. 91. European Centre for Ecotoxicology and Toxicology of Chemicals, Brussels, Belgium
- Compilation of historical 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 -

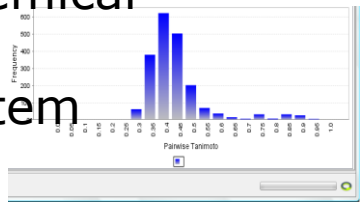
easily compare and characterize datasets

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

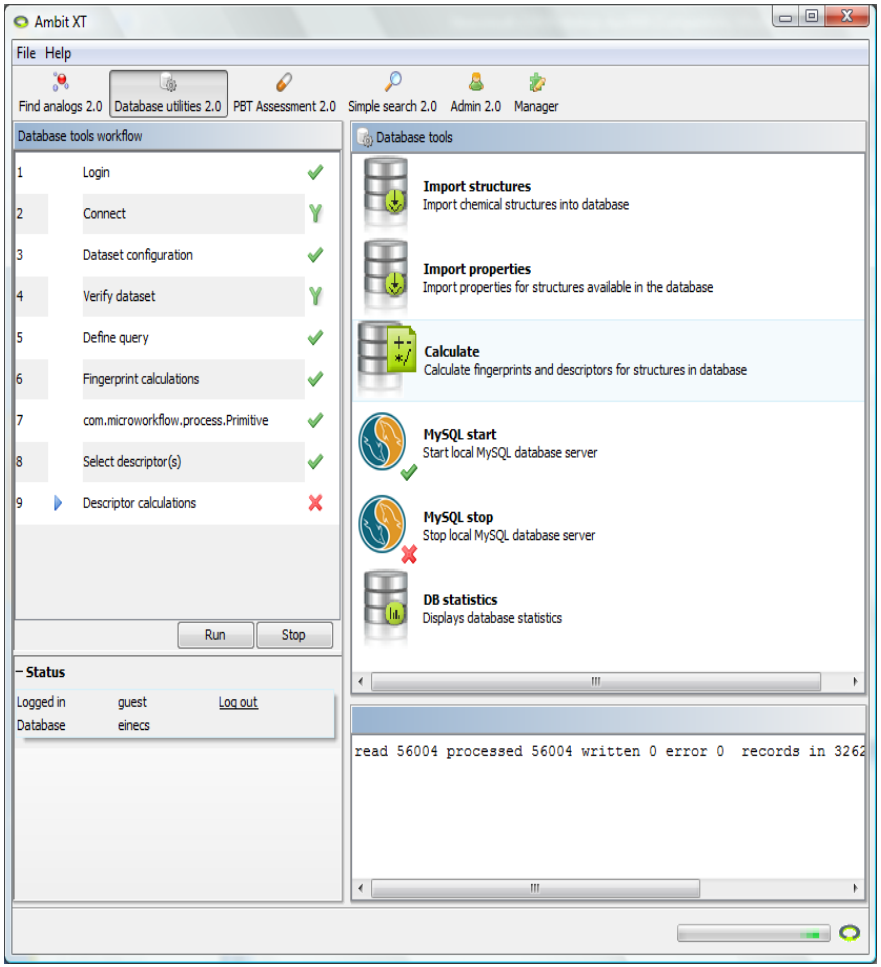


- EINECS
- DSSTOX CPDBAS: Carcinogenic Potency Database Summary Tables - All Species
- LLNA (Gerberick 2005. Dermatitis 16(4): 157-202.)
- Munro reference database (Munro 1996, Food and Chemical Toxicology 34 (1996) 829-867)
- DSSTOX IRISTR: EPA Integrated Risk Information System (IRIS) Toxicity Review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

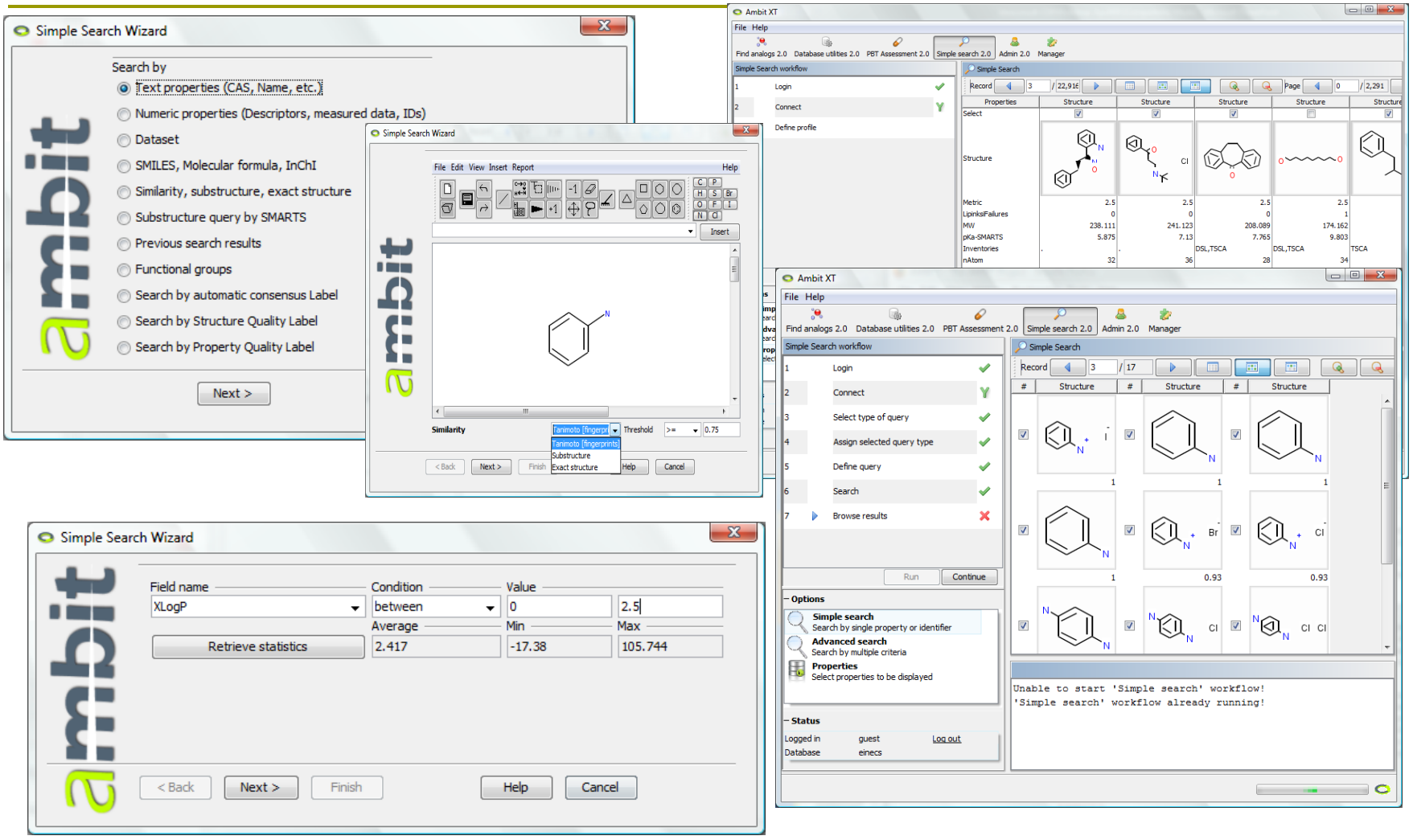


AMBIT XT – workflow support

- 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)



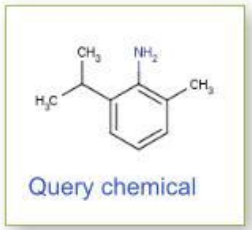
AMBT XT – Search facilities



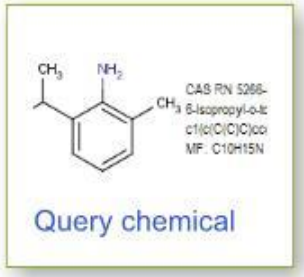
A workflow in AMBIT XT

1. Load chemical
Search by identifier or specify structure

2. Load data
Properties related to the selected endpoint

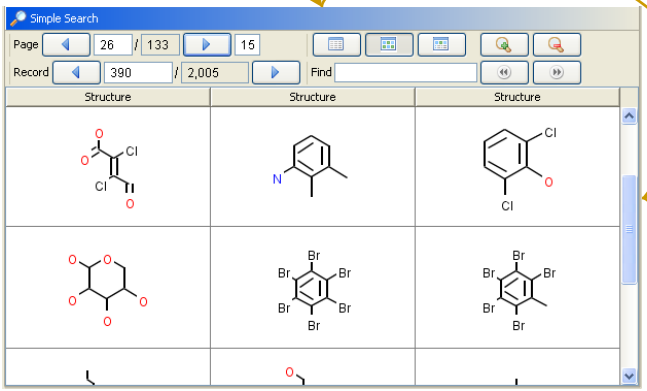


3. Predict properties
Calculate descriptors and properties not available in the database



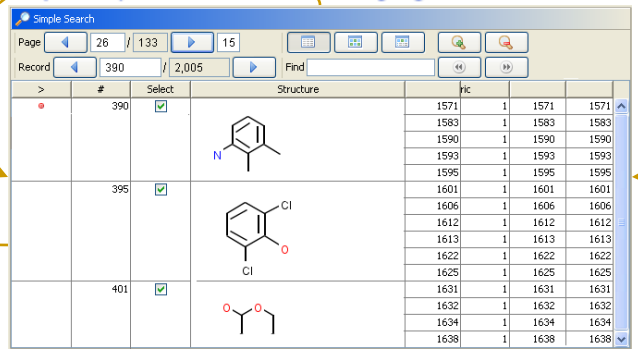
4. Find similar substances
Substances with similar structure and descriptor values

5. Add more substances
Specify similar substances by hand

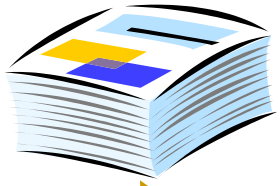


6. Load data
Retrieve properties related to the selected endpoint

7. Analyse data
Calculate probability that the datapoint is questionable and further testing might be valuable

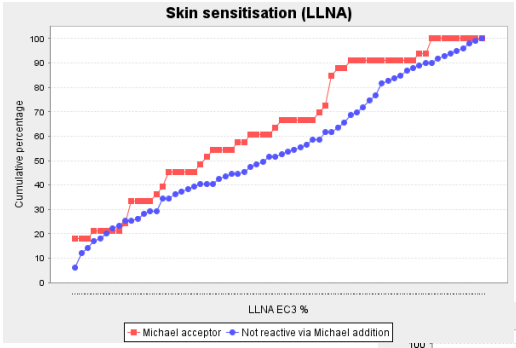


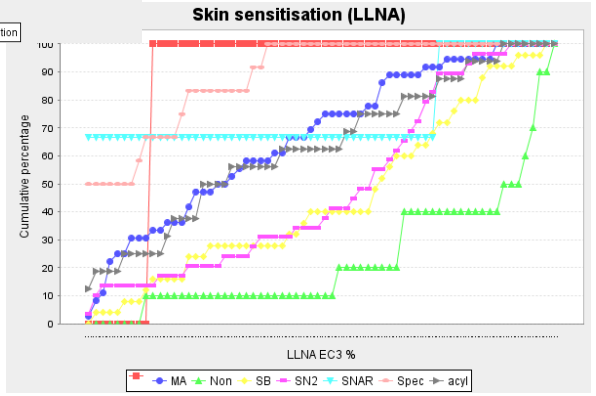
8. Generate report
Data matrix and probabilities

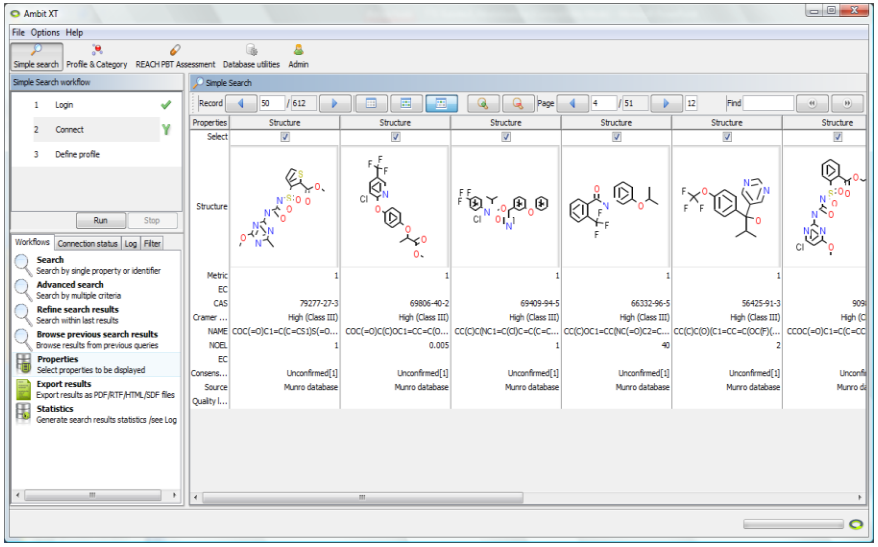


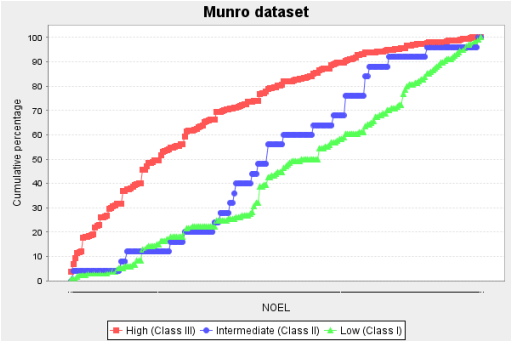
AMBIT – TTC use case

- *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*



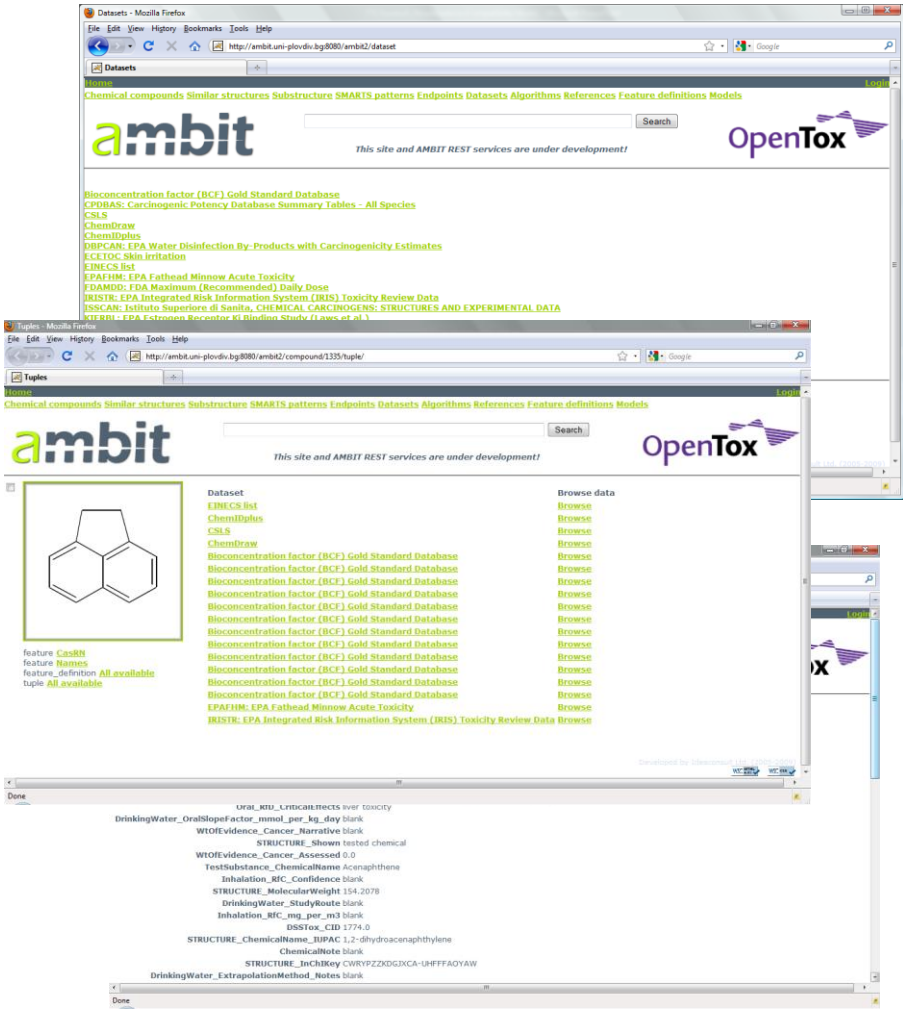




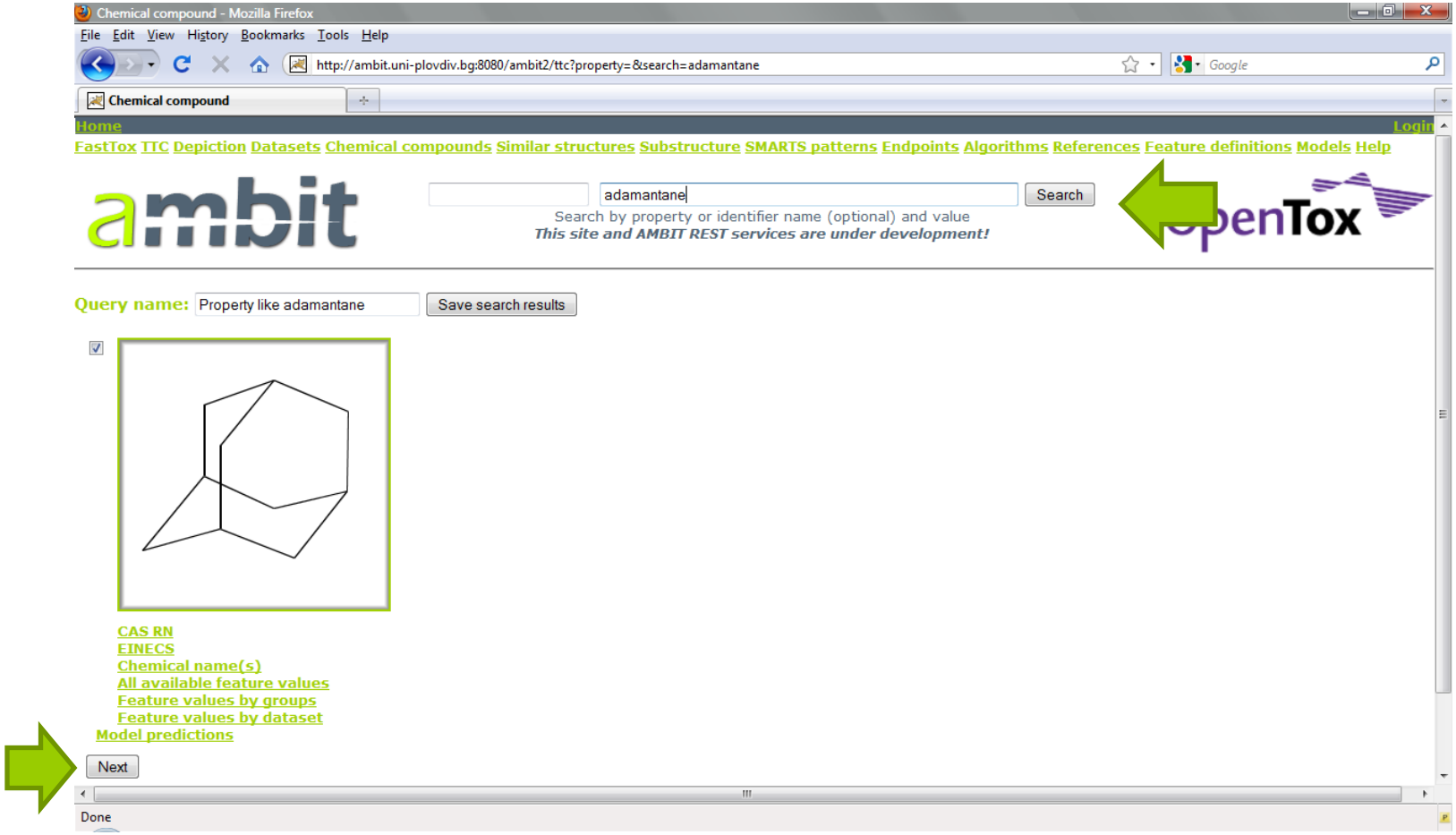


AMBIT Web services
under development

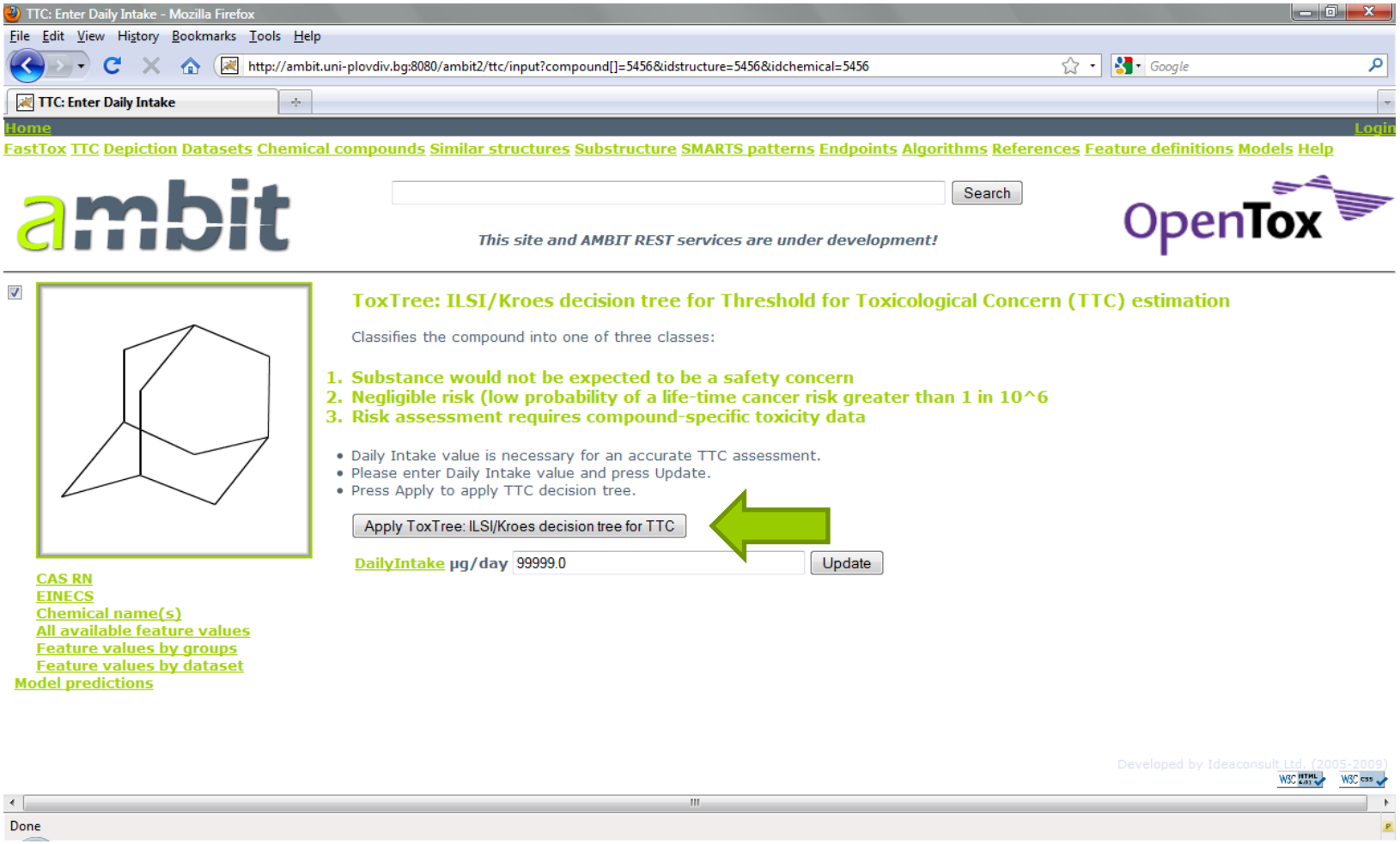
- *Ambit Web services*
  - *Data aggregation*
  - *Descriptor calculation*
  - *Model development*
  - *Model validation and comparison*
  - *Easy access by human experts (web applications) and software (web services)*
- *Based on Web services API, developed by 11 partners from EC FP7 OpenTox project*
- *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



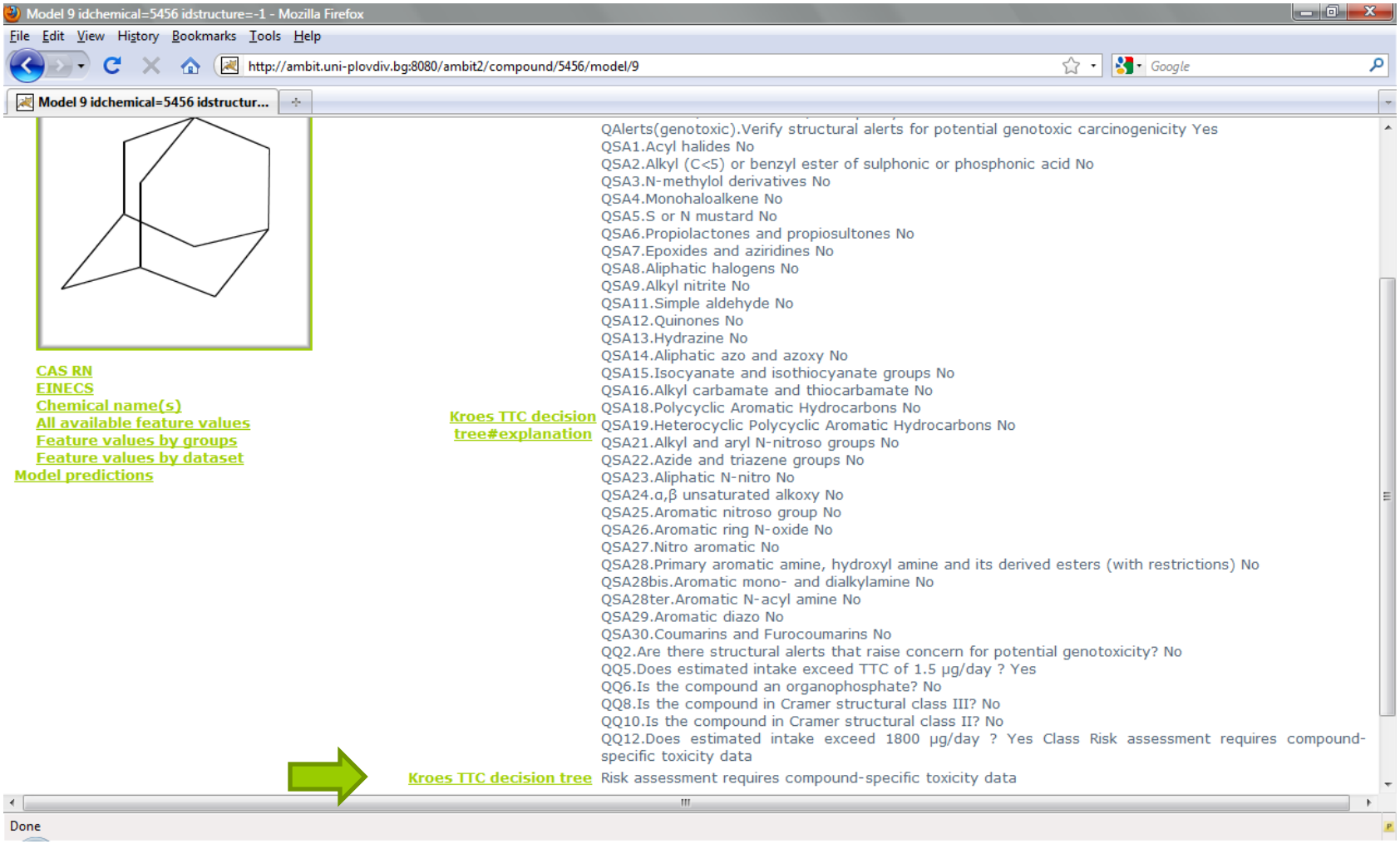
TTC estimation online



TTC estimation online



TTC estimation online



Thank you!

Questions?