CHEMPROP – CHEMICAL PROPERTIES ESTIMATION SOFTWARE SYSTEM

Ralph Kühne[1], Barbara Wagner[1], Ralf-Uwe Ebert[1], David Seemann[1], Gerrit Schüürmann[1,2]

[1] *Helmholtz Centre for Environmental Research – UFZ, Department Ecological Chemistry*

[2] *Technical University Bergakademie Freiberg*

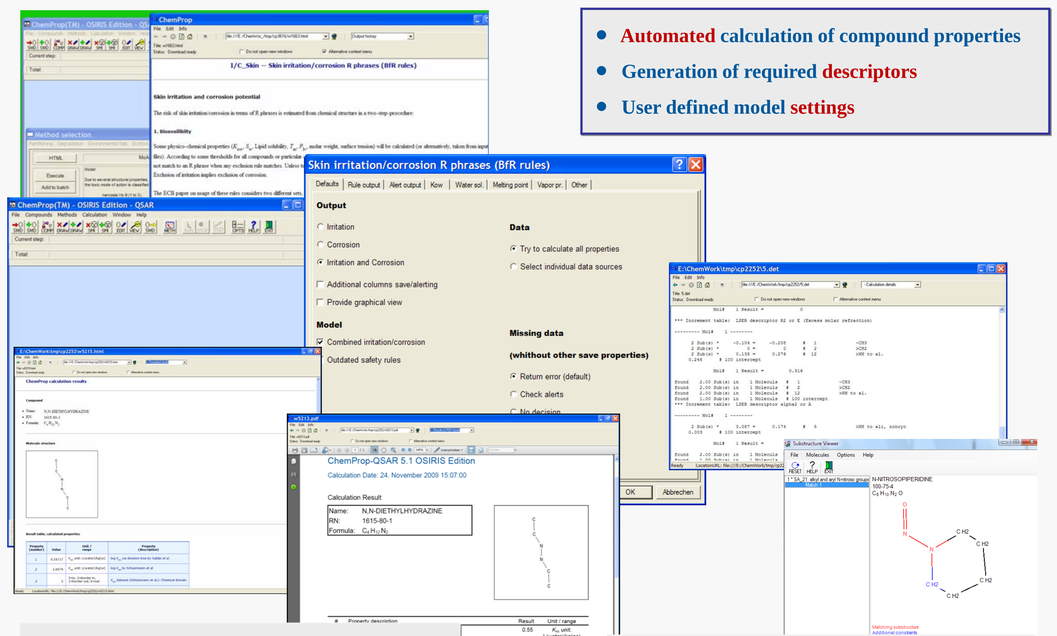
Contact:
ralph.kuehne@ufz.de

CHEMPROP OVERVIEW

- **QSARs for physical chemistry, environmental fate and ecological and human toxicology**
- **Quantitative and categorical models, decision trees, automated read-across**
- **Support for automated model selection**
- **Applicability domain and model uncertainty**
- **Database with structure and substructure searching**
- **Interface to OSIRIS webtool**

QSAR MODULE

- **Automated calculation of compound properties**
- **Generation of required descriptors**
- **User defined model settings**



Basic Output: HTML and PDF Report

- Description of **compound** with formula
- Tabulated **calculation results** (with easy export to **spreadsheet** opportunity)
- Model **documentation**

Extended Output

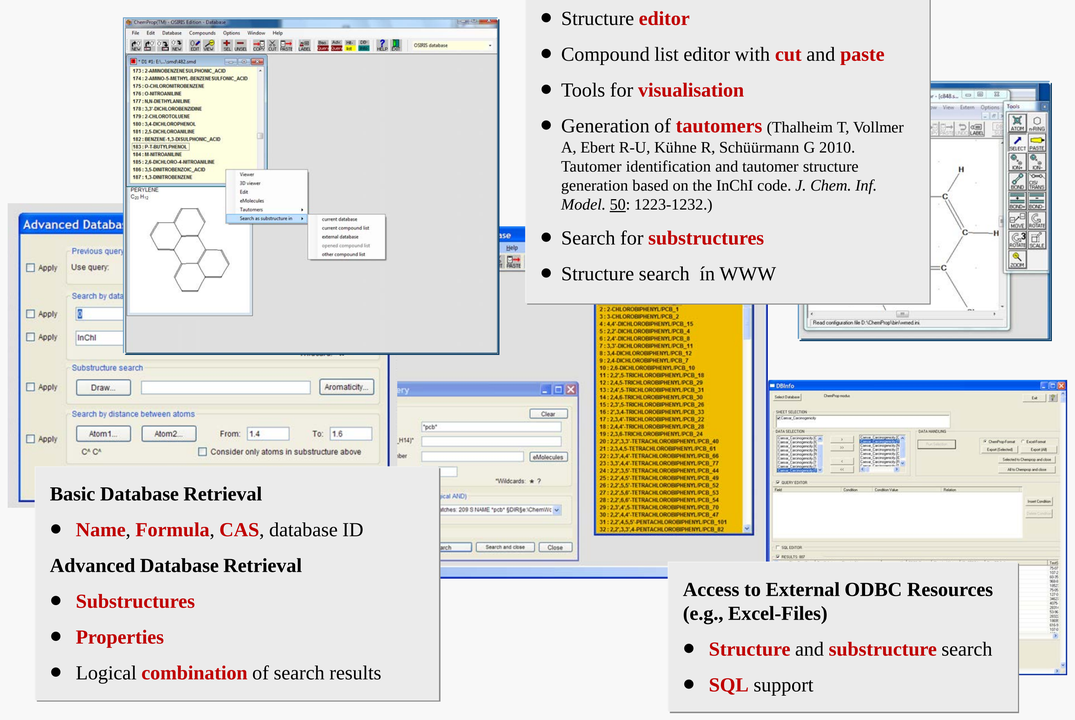
- Calculation **details**
- **Visualisation** of key **substructures**
- Session **protocol** including links to all reports

DATABASE MODULE

- **Structure import, editing, export**
- **Database search (structures, data)**
- **Link to external databases (ODBC, SQL)**

Structure (Molecule) Lists

- Structure **import/export**: SMILES, InChI, MDL, CML, IUCLID
- Structure **retrieval** from database
- Structure **editor**
- Compound list editor with **cut** and **paste**
- Tools for **visualisation**
- Generation of **tautomers** (Thalheim T, Vollmer A, Ebert R-U, Kühne R, Schüürmann G 2010. Tautomer identification and tautomer structure generation based on the InChI code. *J. Chem. Inf. Model.* 50: 1223-1232.)
- Search for **substructures**
- Structure search ín WWW



Basic Database Retrieval

- **Name**, **Formula**, **CAS**, database ID

Advanced Database Retrieval

- **Substructures**
- **Properties**
- Logical **combination** of search results

Access to External ODBC Resources (e.g., Excel-Files)

- **Structure** and **substructure** search
- **SQL** support

WEBTOOL INTERFACE

OSIRIS Webtool (http://osiris.simpple.com/OSIRIS-ITS/welcome.do)
WWW Based Tool to Apply Integrated Testing Strategies (ITS)
ChemProp provides calculated properties including domain information to the webtool

FURTHER INFORMATION

System Requirements

- **Windows** XP SP3 or Vista or 7
- **.NET Framework** 2.0 or higher

Web Resources

- **http://www.ufz.de/index.php?en=6738**
- Alternatively, from www.ufz.de select Divisions, Ecological Chemistry, Methods, there click on the ChemProp link

Availability

OSIRIS edition:

Publicly available due to **charge-free license agreement**

(see web page for details)

Contact

- gerrit.schuurmann@ufz.de
- ralph.kuehne@ufz.de

MODELS

Quantitative and Categorical Models
- Computerized read-across
- Structural alerts
- Regression-based relationships
  - Increment and conventional multilinear models
  - Abraham-type linear solvation-energy relationships (LSERs)
- Decision-support schemes for automated method selection

Profiles
- Model list for batch run
- Load and save model collections with specific settings

Model Sources
- UFZ developments
- Implementations of literature models

Model Selection
- Typically several models for single property
- User defined model settings and adaptions

Validation
- Applicability domain
- Uncertainty and prediction capability

Automated read-across: Schüürmann G, Ebert R-U, Kühne R 2011. Quantitative read-across for predicting the acute fish toxicity of organic compounds. *Environ. Sci. Technol.* 45: 4616-4622.

Structural alerts: Von der Ohe PC, Kühne R, Ebert R-U, Altenburger R, Liess M, Schüürmann G 2005. Structural alerts - a new classification model to discriminate excess toxicity from narcotic effect levels of organic compounds in the acute daphnid assay. *Chem. Res. Toxicol.* 18: 536-555

Regression based relationships: Schüürmann G, Ebert R-U, Nendza M, Dearden JC, Paschke A, Kühne R 2007. 9. Predicting fate-related physicochemical properties. In: van Leeuwen K, Vermeire T (eds) Risk Assessment of Chemicals. An Introduction. Springer Science, Dordrecht (NL), pp. 375-426.

Automated model selection: Kühne R, Ebert R-U, Schüürmann G 2006. Model selection based on structural similarity – method description and application to water solubility prediction. *J. Chem. Inf. Model.* 46: 636-641.

Chemical domain: Kühne R, Ebert R-U, Schüürmann G 2009. Chemical domain of QSAR models from atom-centered fragments. *J. Chem. Inf. Model.* 49: 2660-2669.

Model validation: Schüürmann G, Ebert R-U, Chen J, Wang B, Kühne R 2008. External validation and prediction employing the predictive squared correlation coefficient - test set activity mean vs training set activity mean. *J. Chem. Inf. Model.* 48: 2140-2145.

PROPERTIES (OSIRIS EDITION)

Partition coefficients

Quantitative Models

- **Solubility in Water**
  Individual fragment, property, read-across, and other models
  ACF based automated model selection
- **Octanol/Water Partition Coefficient**
  Individual fragment, LSER, and read-across models
  Compound class based automated model selection
- **Air/Water Partition Coefficient**
  Individual fragment, read-across, and LSER models
  General performance based automated model selection
  Temperature dependency
- **Octanol/Air Partition Coefficient**
  Property and LSER models
- **Melting Point**
  Individual fragment and read-across models
  General performance based automated model selection
- **Boiling Point**
  Individual fragment models
  General performance based automated model selection
- **Vapour Pressure**
  Property and neural network models (with temperature dependency)
  Read-across model (25ºC)
- **Soil Sorption Coefficient**
  Fragments, topological indexes, LSER, and read-across models
  Decision tree and equation models from other properties
  pH Dependency

Human Toxicology

Quantitative Models for PBPK related properties (mammals and human)

- **Partition Into and Between Tissues**
  LSER or property models for air/blood, air/fat, blood/fat, and blood/brain partitioning
  LSER model for human serum albumin (HSA) binding
  Fragment model for partition coefficients fat, liver, muscle, blood to air
  LSER model for skin permeability

Categorical Models

- **General Rule Sets**
  Cramer classes for toxicity (decision tree from substructures)
  Lipinski's rule of five:
  Property criteria for poor absorption or permeation
- **Mutagenicity and Carcinogenicity**
  Structural alert and read-across models
  Consensus model for carcinogenicity
- **Endocrine Disruption**
  Decision tree (structural rules and properties) for estrogenicity
  Structural alert model for androgen receptor antagonism
- **Skin and Eye**
  BfR rules for irritation and corrosion:
  Property models for bioavailability
  Structural alerts for adverse effects
  Structural alerts for skin sensitization

Environmental Fate

Quantitative Models

- **Bioconcentration in Fish (BCF)**
  Simple and complex property (K_{ow}) and read-across models
  BCF_{max} models
- **Further Fish Models**
  Complex model for bioaccumulation
  Fragment model for biotransformation
- **Plants**
  LSER models

Categorical Model

- **Fish**
  Decision tree from properties for BCF test waiving

Degradation

Categorical Models

- **Biodegradation**
  Fragment models for probability (aerobic, anaerobic)
- **Half-life in Environmental Compartments**
  Semi-quantitative read-across model for air, water, soil, sediment (9 classes)
  Simple temperature dependency

Ecotoxicology

Quantitative Models

- **Baseline and Polar Narcosis**
  Property (K_{ow}) models for several aquatic organisms
  LSER models for several aquatic organisms
- **Specific Toxicity Models**
  Read-across models for daphnids and fathead minnow
  Fragment model for fathead minnow
- **ECOSAR Type Equations**

Categorical Models

- **Mode of Action**
  Structural rule models for fish (Verhaar and Russom classes)
- **Excess Toxicity**
  Structural alerts for daphnids and fish

Other Properties and Tools

Quantitative Models and Calculations

- **Polar Surface Area**
  Fragment model
- **Molar Mass**

Other Tools

- **Chemical Domain**
  ACF based chemical domain for user defined data sets
- **Database**
  Structural database lookup

RECENT DEVELOPMENTS

UFZ Models and Models with UFZ Participation
- New and improved UFZ models for the **air/water** partition coefficient
- New and improved models for **air/blood**, **air/fat**, and **fat/blood** partitioning
- New **read-across** model for **BCF**
- Automated **waiving** scheme for **BCF**

Literature Models
- **Blood/brain** partition coefficient
- Structural alerts for **micronucleus**

Other Improvements
- **Rule examples** for structural alert models
- Extended model and program **documentation**

ACKNOWLEDGMENT

The ChemProp development has been supported significantly by the EU integrated project **OSIRIS** (Optimized Strategies for Risk Assessment of Industrial Chemicals through Integration of Non-Test and Test Information; GOCE-CT-2007-037017, 2007-2011.)
www.osiris-reach.eu

Additional financial support is gratefully acknowledged from the former EU Projects **2-Fun** (Full-chain and Uncertainty Approaches for Assessing Health Risks in Future Environmental Scenarios) contract No. 036976. and **CAESAR** (Computer Assisted Evaluation of industrial chemical Substances According to Regulations) contract No. 022674, the German Federal Environmental Agency (**UBA**), Électricité de France (**EDF**), the **RIVM**, The Netherlands, **Lhasa Limited**, UK, the **Water Research Commission of South Africa**, Environmental Risk Assessment and Management (**ERASM**), Brussels, and the Long Range Research Initiative (**LRI**) of the European Chemical Industry Council (**CEFIC**), Brussels.

Currently further developments are supported by the IGF project FORMULA (17176 BG / 2) of the IUTA Duisburg through the **AIF** within the framework of the **IGF**