U.S. EPA - OPPT QSAR and Expert System Tools

Tala Henry

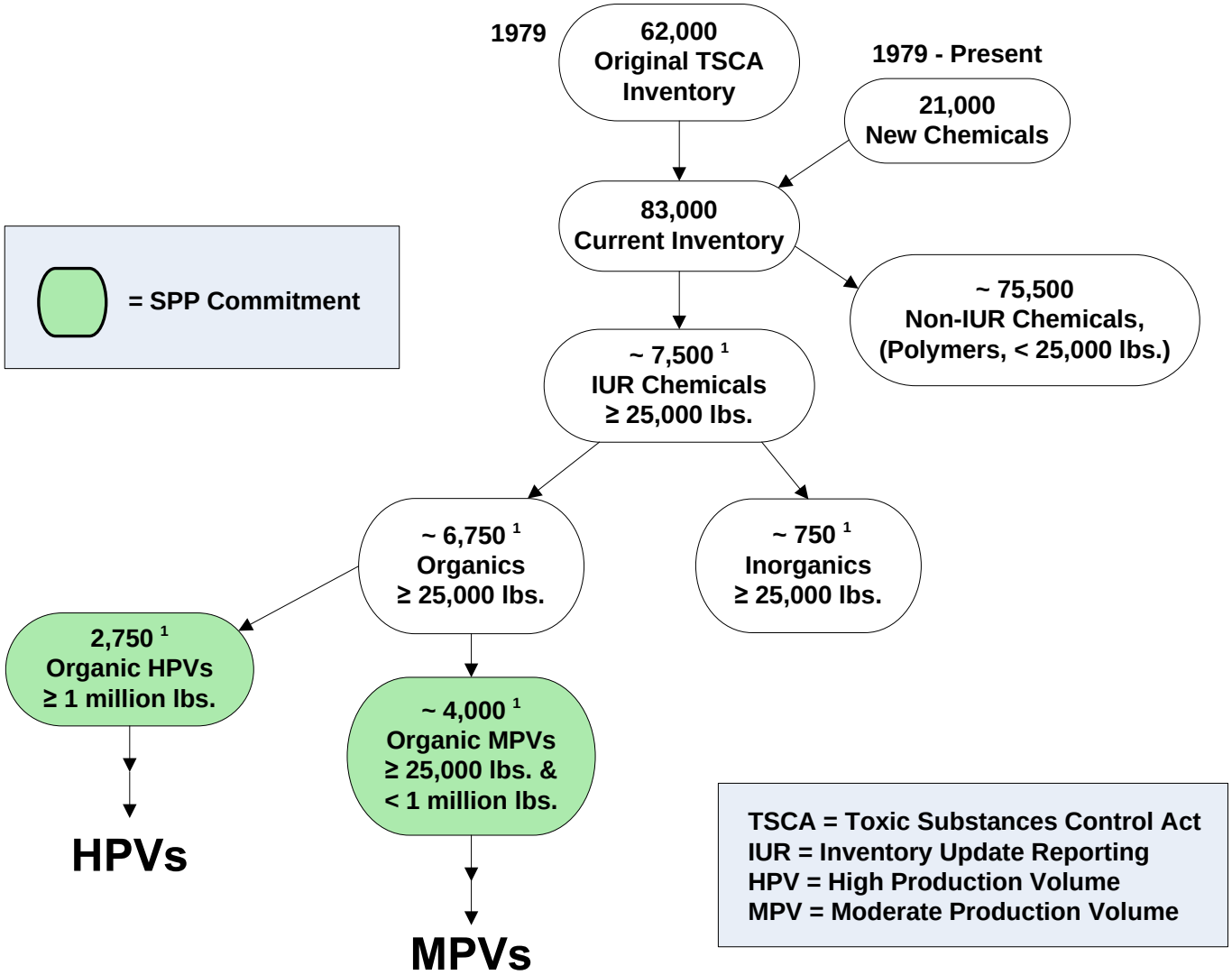
Office of Pollution Prevention and Toxics

PPDC Meeting
October 8, 2008

Office of Pollution Prevention and Toxics

- Administer the Toxic Substances Control Act (TSCA) and the Pollution Prevention Act (PPA)
- Review Pre-Manufacture Notices (PMN) for new industrial chemicals
- Request and/or perform testing, assessment, and risk reduction of existing industrial chemicals
- Manage “national chemicals” (e.g. PCBs, Hg, Lead)
- Represent USA on International chemical issues (e.g. POPs)
- Advocate & facilitate pollution prevention
- Encourage & facilitate partnership programs, e.g. DfE, HPV Challenge, Green Suppliers Network and Green Chemistry

TSCA Chemical Universe



[1] Statistics are based upon preliminary 2006 IUR data; the actual numbers may change slightly when official statistics are available.

Note: The 2006 IUR introduces new reporting thresholds.

New Chemicals Program Overview

Section 5 of TSCA

- Requires a manufacturer or importer of a new chemical substance to submit a “premanufacture notice” (PMN) to EPA 90 days before intended start of production or import
- Designed to prevent health and/or environmental risks before they occur
- Regulatory decisions must be made within 90 days and often in the absence of data
- ~ 1500 – 2000 PMN Chemicals assessed per year; over 41,000 PMNs reviewed to date

PMN Process

- Information Submitted: chemical identity, use, production/import volume, byproducts, human exposure, release to environment, *existing* test data (most often acute toxicity; little fate or environmental effects)
- TSCA allows CBI claims: only 5% of PMNs submitted to date have been completely non-confidential
- Review of PMNs to determine if need to prohibit or limit manufacturing, processing or use pending development of needed information
- Following EPA review, if manufacturing or importing commences:
  - EPA must be notified (NOC)
  - Chemical is added to the TSCA inventory

Structure Activity Team (SAT) Meeting

- Key technical meeting to consider potential hazard to human health and the environment

- Hazard profile based on:
  - Physical-chemical properties
  - Routes of absorption
  - PMN data
  - Structure activity relationship (SAR) analysis:
    - Analogs & categories
    - QSAR estimates

Categories & Analogs

OncoLogic®

Exposure Report

- ChemSTEER - Chemical Screening Tool for Exposures and Environmental Releases
  - Estimates workplace exposures and releases resulting from manufacture or use of an industrial chemical
  - Uses EPA default assumptions (worst case scenarios)
- E-FAST - Screening level assessment of exposures to the environment and non-workers
  - Aquatic species
  - Consumers
  - General population

Existing Chemicals Program Overview

Security and Prosperity Partnership (SPP)

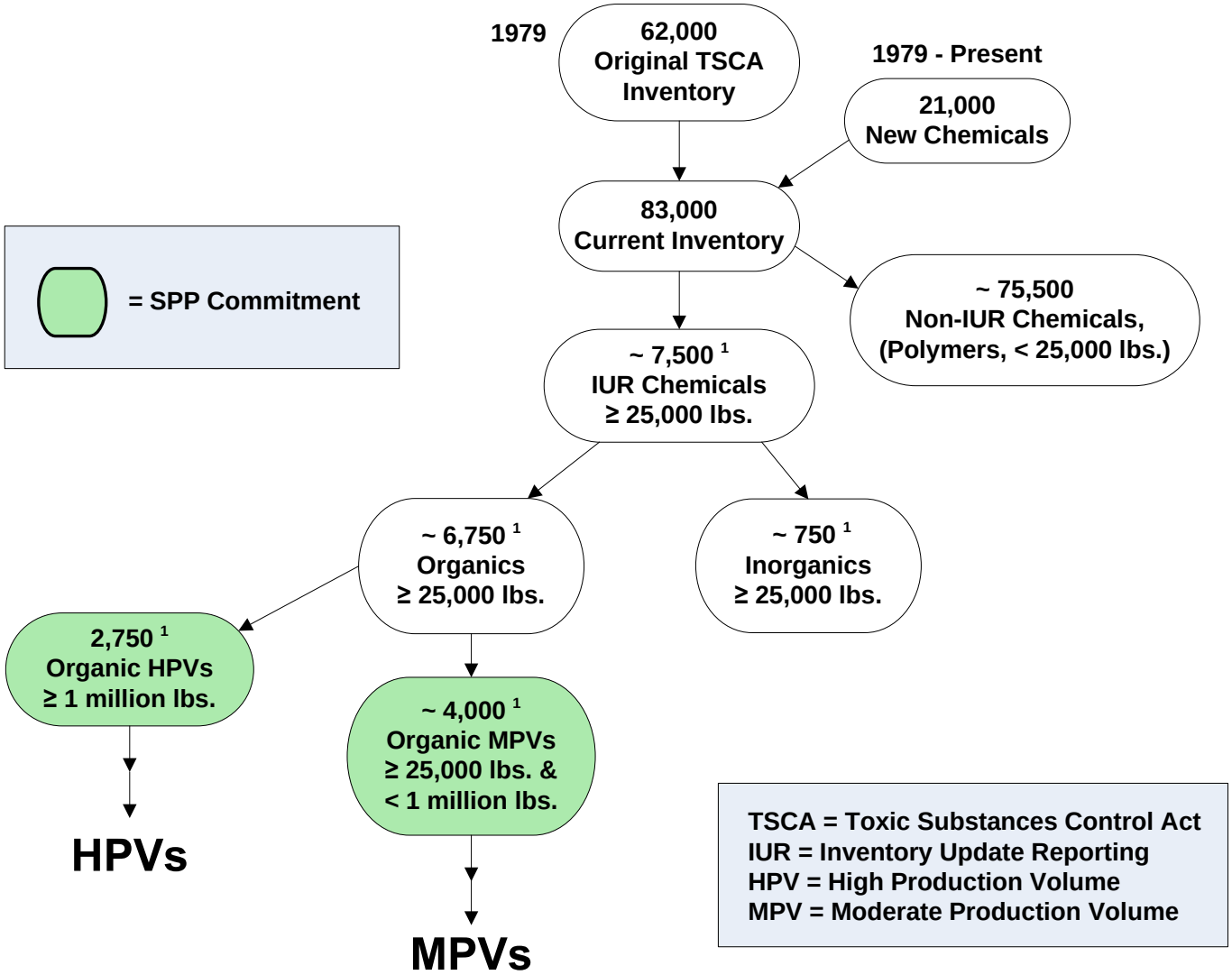
- Goal: enhance regulatory cooperation among Canada, Mexico, U.S.

- U.S. Commitment, by 2012:
  - Assess and initiate needed action on the over 6,750* existing chemicals produced above 25,000 lbs/yr in the U.S.

  - Includes work on organic High Production Volume (HPV's) and Moderate Production Volume (MPV's)

* Based on preliminary statistics from 2006 IUR data

U.S. SPP Commitments



[1] Statistics are based upon preliminary 2006 IUR data; the actual numbers may change slightly when official statistics are available.

Note: The 2006 IUR introduces new reporting thresholds.

Chemical Assessment and Management Program (ChAMP)

ChAMP was created to implement commitments the United States made under the SPP and other related actions:

I. Risk-based Prioritizations for HPVs (≥ 1 M lbs/yr)

II. Hazard-based Prioritizations for MPVs (25K – 1M lbs/yr)

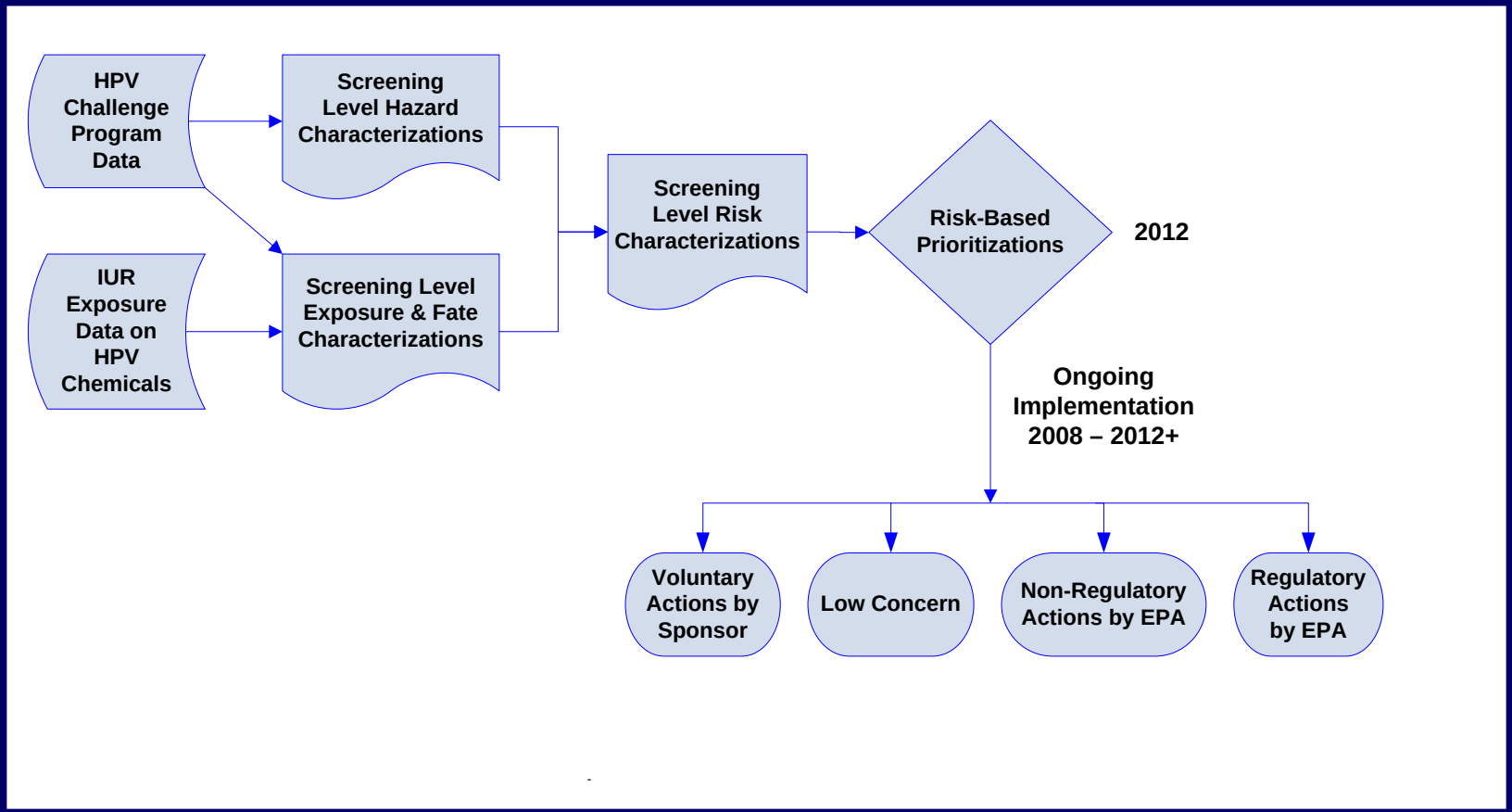
III. Inorganic HPV Chemicals

IV. Inventory Reset

HPV Chemical Risk-Based Decision Making

- 1998 – High Production Volume (HPV) “Challenge”
  - a voluntary initiative to develop and make public, screening-level health and environmental effects information on HPV chemicals
  - producers and importers identify and provide existing toxicity data/information and conduct new testing if data do not exist
  - a complete data submission contains data or information on 18 internationally agreed to “SIDS” (Screening Information Data Set) endpoints
  - Essentially same as OECD HPV Programme\
  - Both US and OECD HPV Programs encourage the use of Categories
- 2006 – Inventory Update Reporting
  - Industrial Processing and Use Information
  - Commercial and Consumer Use Information
  - Used in products intended for children

HPV Chemical Risk-Based Decision Making



MPV Chemical Hazard-Based Screening Decision Process

- Produced or imported at quantities ≥ 25,000 lbs/yr and ≤ 1 million lbs/yr.
- Apply available existing data and OPPT Category & Structure Activity Relationships (SAR) analysis to assess hazard and fate.
- Basic exposure/use data are available only for MPVs produced at ≥ 300,000 lbs at a site
- Use Hazard Characterizations (HCs) to identify MPVs that require follow-up: testing, risk management, etc.

Approach for MPV Hazard Characterization

1986–2006 IUR

New Chemicals

Identify Structural Clusters

Cluster 1 …

Analog Identification Methodology

Physical-Chemical Properties – HPV, PhysProp, EPISuite, SPARC

Environmental Fate – HPV, PhysProp, EPISuite

EPI

PBT Profiler
A Component of OPPT's P2 Framework
Assessing Chemicals in the Absence of Data

Ecotox break pt 1

Ecotoxicity

Ecotoxicity – HPV, ECOTOX, ECOSAR

ECOSAR

HH break pt 1

HH break pt 2

Human Health Endpt 1

Human Health Endpt 2

Human Health Endpt 3

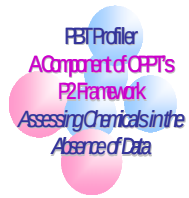
OncoLogic®

Meeting the SPP Goals

- 2007
  - Developed process for screening-level Hazard Characterizations (HCs), Exposure Characterizations (ECs), Risk Characterizations (RCs) and Risk-Based Prioritizations (RBPs)
  - Posted over 150 HCs
- 2008
  - Posted additional ~150 HCs
  - Posted RBPs for 150 chemicals
  - Post initial MPV HCs
- 2009
  - Continue posting RBPs for HPV chemicals and significantly increase development/posting MPV HCs

Filling the Gaps: OPPT Predictive Tools

OncoLogic®



- ***EPI Suite™*** **–** Suite of QSAR programs for estimating physical-chemical properties & environmental fate parameters.
- ***ECOSAR*** – a library of QSARs for predicting aquatic toxicity and an expert system for selecting the appropriate QSAR
- ***PBT Profiler*** – estimates **P**ersistence, **B**ioaccumulation and **T**oxicity and distribution in water, soil, sediment, and air using a Level III multi-media model.
- ***OncoLogic®*** Cancer Expert System – Predicts concern levels for cancer potential based on "knowledge rules"
- ***E-FAST*** – General Population, Consumer, and Environmental Exposures
- ***ChemSTEER*** **–** ***Chem***ical ***S***creening ***T***ool for ***E***xposures & ***E***nvironmental ***R***eleases

Appropriate Use of Predictive Models

- Priority setting tools allow one to compare large numbers of chemicals using standardized criteria
- Screening level tools can help to screen out or eliminate chemicals of low concern and identify where additional evaluation is needed
- Higher tier tools provide more definitive, detailed analyses, often used in combination with screening level tools and available data
- Models can be helpful in filling data gaps and to supplement and/or interpret data

OPPT QSAR and Expert System Tools

- Fate and Toxicity Tools and Models
  - Analogs & Categories
  - AIM
  - ECOSAR
  - Oncologic®
  - PBT Profiler
- Exposure Assessment Tools and Models
  - ChemSTEER
  - E-FAST
- Other detailed models are also used as needed

Analogs & Categories

- Analogs - chemicals whose physicochemical and toxicological properties are likely to be similar as a result of structural similarity
- Category - A group of analogs
- Similarities based on:
  - a common functional group (e.g., aldehyde, epoxide, ester, etc.)
  - common precursors and/or breakdown products
  - similar biological acvtivity and/or mechanisms of action
- Categories can be reduced (subcategories) or expanded
- Analogs & Categories Facilitate:
  - Evaluation of the reliability of model estimates
  - Filling data gaps
  - Evaluations based on a greater weight of evidence
  - Strategic testing
  - Identification of safer alternative or substitutes

Categories - Lessons

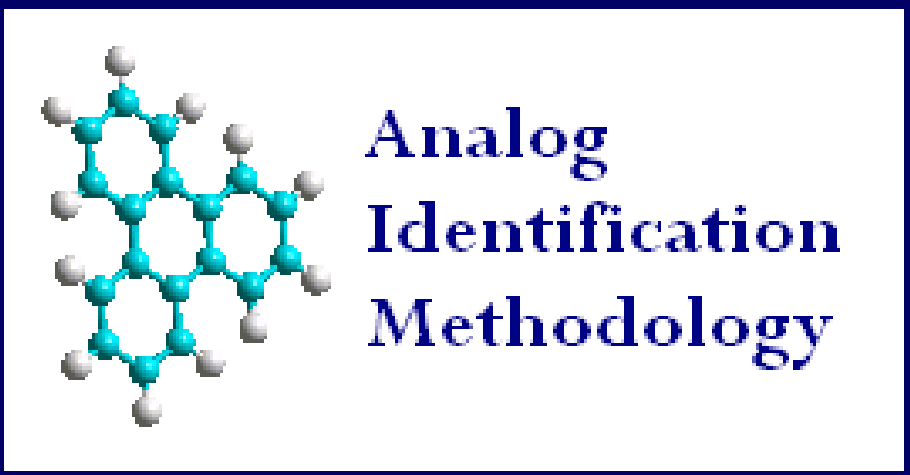
- Based on ~2 decades of using Categories to assess hazard and risk to 1000s of chemicals, U.S. EPA and others have found:

- Category evaluation
  - is based on a greater weight of evidence; increases confidence in conclusions
  - provides better basis for establishing biological plausibility

- Category analysis facilitates strategic testing
  - weight of evidence used for deciding need for additional testing
  - defines the nature and scope of any testing needs
  - testing often completed faster

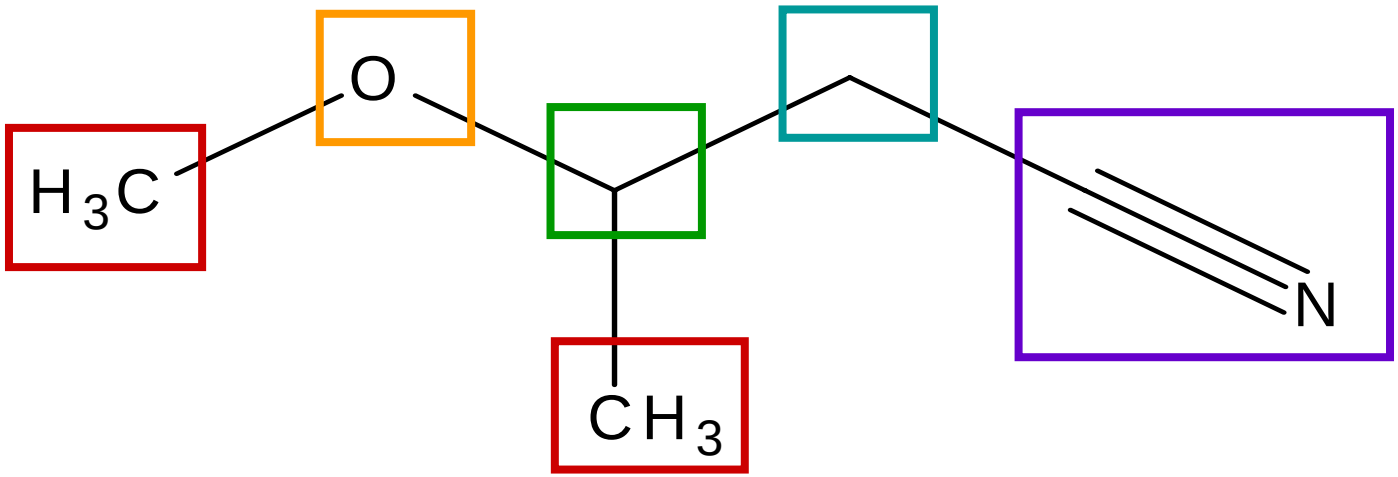
- Chemical Categories are a practical way to:
  - Extrapolate data gathered for HPV chemicals to lower volume chemicals,
  - To meet goals of assessing large number chemicals (U.S. EPA TSCA; REACH; Canadian DSL), and
  - Guide/Organize Integrated Testing Strategies

Analog Identification Methodology (AIM)



AIM Methodology

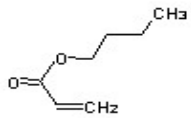
AIM identifies analogs using a chemical fragment-based approach (645 individual *fragments* indexed in the database)



- Web-based program provides user with a list of chemical analogs linked to publically available toxicity data.
- From a database of 31,031 chemicals/structures that are indexed to publicly available toxicity sources such as: TSCATS, AEGLS, IRIS, HPV Challenge, ATSDR, NLM-HSDB, NTP, RTECS, IUCLID

Found 12 Analog(s) for 2-Propenoic acid, butyl ester:

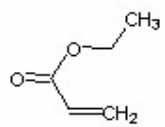
Exact Chemical Match:

# of Studies: 3	BUTYL ACRYLATE [CAS No. 141-32-2]		
	Toxicity Data Available for this Compound		
	On-Line Databases	U.S. Government Documents	Other Sources
	TSCATS HSDB		RTECS IUCLID

- Based on its structure, this chemical may belong to the acrylates/methacrylates. Members of this category may have potential human health concerns. More information and category definitions

- This compound may metabolize in the body to products that may cause concerns for human health. Analogs for metabolites should also be investigated. Metabolism classes found: terminal double bond, ester, and acrylate

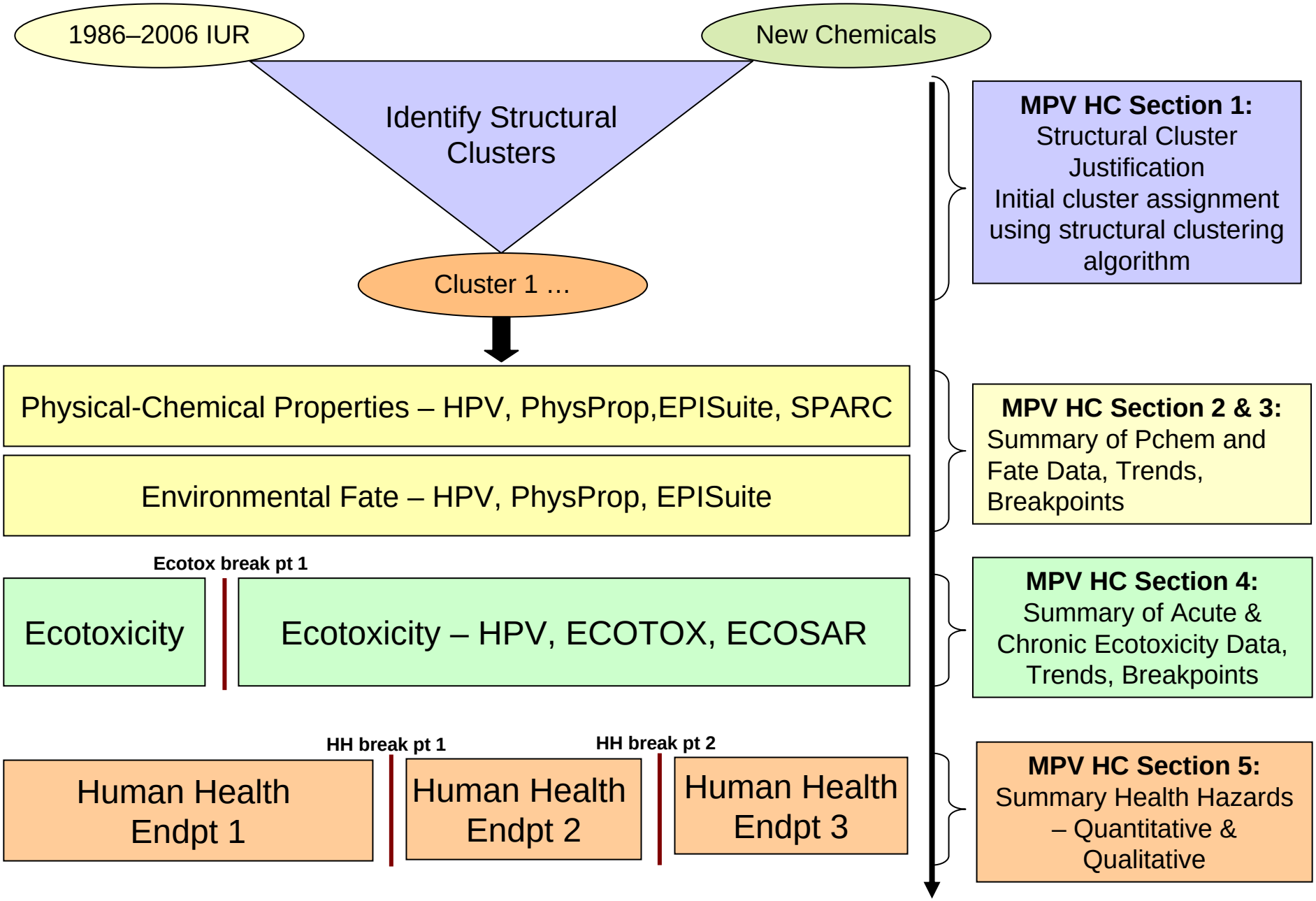
Analogs, Ordered by Number of Studies:

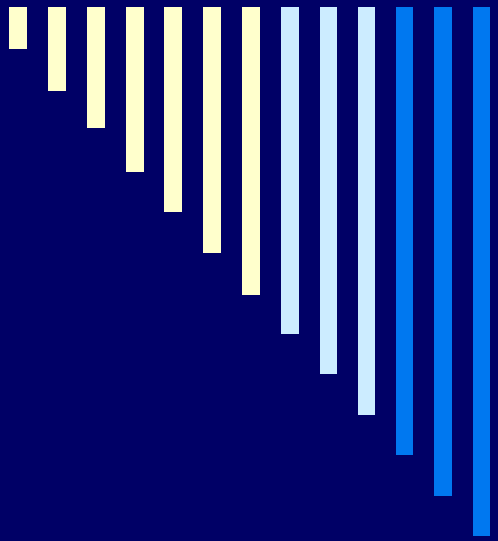
# of Studies: 5	ETHYL ACRYLATE [CAS No. 140-88-5]		
Analog # 1 	Toxicity Data Available for this Compound		
	On-Line Databases	U.S. Government Documents	Other Sources
	TSCATS HSDB	NTP	RTECS IUCLID DSSTox Cancer

AIM Clustering Tool/Category Builder

- AIM fragment matching algorithm expanded to organize data sets to identify “structural clusters” of chemicals
- Applied to multiple EPA databases (PMNs, HPV, 8(e), IUR) to formulate structure-based categories; “structural clusters”

Approach for MPV Hazard Characterization





EPI Suite™

E*stimation *P*rograms *I*nterface *Suite

EPI Suite™

- Estimates physical/chemical properties and environmental fate and transport
- Runs estimation programs sequentially with chemical structure as only input
- Can use experimental data to improve the estimations of the other properties
- Includes PHYSPROP, a database of measured p/chem and fate properties for >40,000 chemicals
- Includes an online User's Manual

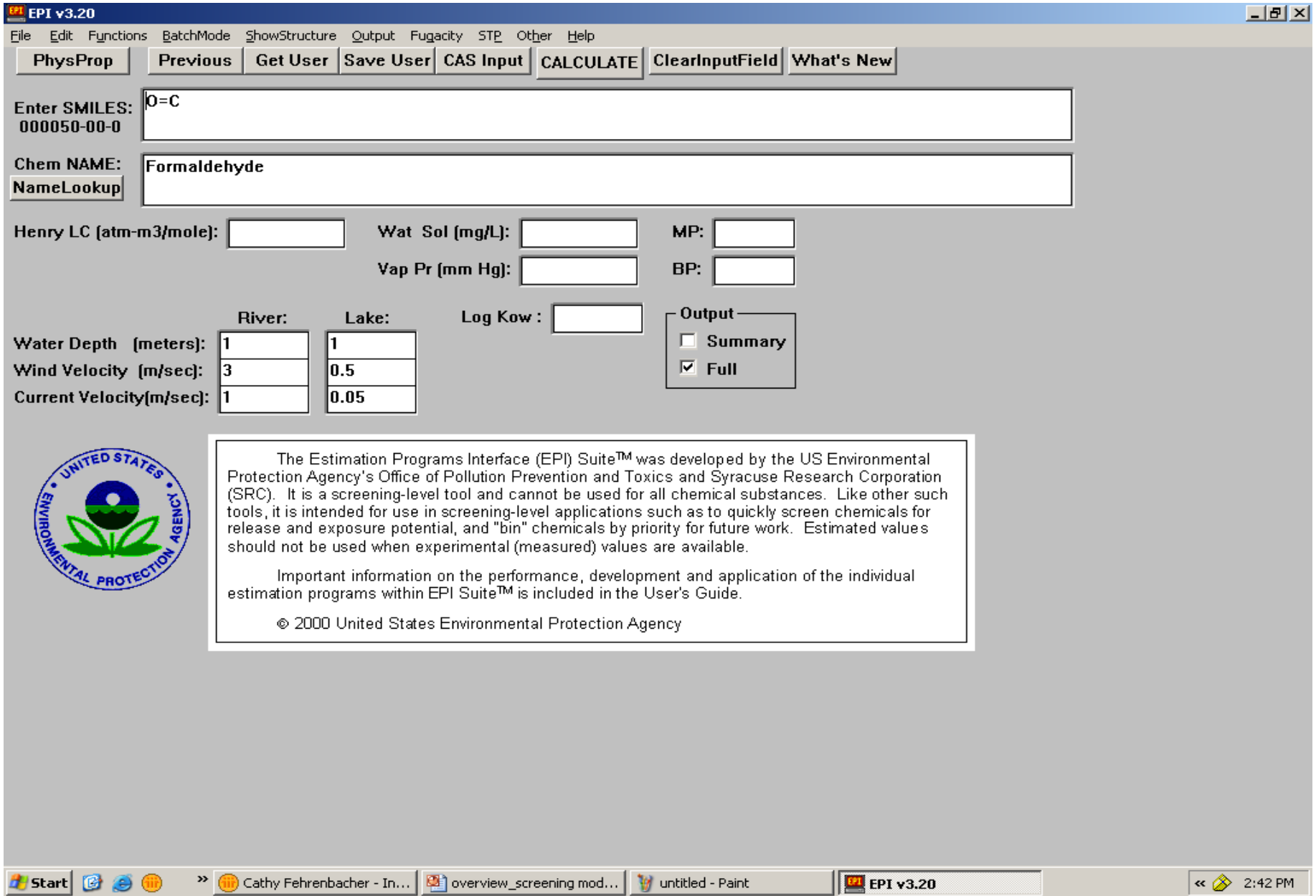
EPI Suite™ P/Chem Modules

- MPBPWIN™ – melting point, boiling point, and vapor pressure
- KOWWIN™ – octanol/water partition coefficient
- WS/KOWWIN™ – water solubility from Kow
- WATERNT™ – water solubility based on a fragment constant method
- HENRYWIN™ – air/water partition coefficient
- PCKOCWIN™ – ability to sorb to the organic portion of soil and sediment
- BCFWIN™ – ratio of a chemical's concentration in the tissue of an aquatic organism to the concentration in the ambient water
- HYDROWIN™ – acid and base-catalyzed hydrolysis
- AEROWIN™ – fraction of airborne substance sorbed to airborne particulates
- AOPWIN™ – atmospheric persistence
- BIOWIN™ – aerobic and anaerobic biodegradation
- BioHCWIN – biodegradation half-life for compounds containing only carbon and hydrogen (e.g., hydrocarbons)
- KOAWIN – octanol/air partition coefficient

EPI Suite™ Environmental Fate Modules

- WVOLWIN™ – rate of volatilization from rivers and lakes
- STPWIN™ – removal in a simulated sewage treatment plant
- LEVEL3EPI™ Fugacity Model – partitioning of chemicals between air, soil, sediment and water under steady state conditions for a default model “environment”

EPI Suite™ Input Screen



ECOSAR

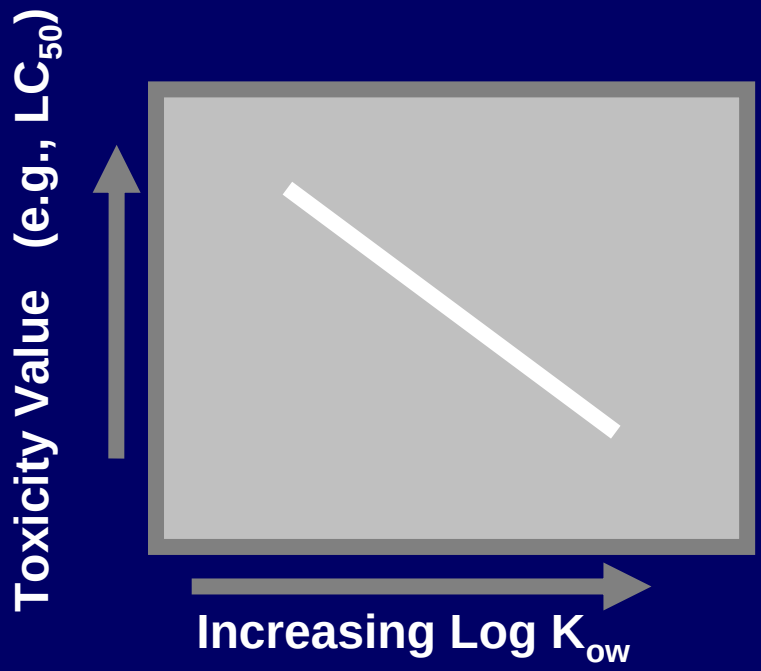
Aquatic toxicity estimates based on
Structure Activity Relationships (SAR)

http://www.epa.gov/oppt/newchems/tools/21ecosar.htm

ECOSAR

- A library of QSARs for predicting aquatic toxicity and an expert system for selecting the appropriate QSAR
  - Acute and chronic toxicity endpoints
  - Fish, aquatic invertebrates, algae, and others
- Extensive documentation and User's Manual

How ECOSAR Predicts Toxicity



- Chemicals are grouped into classes. For **most** classes, regression equations relate *predicted* log K_{ow} to toxicity
  - Example SAR, acrylates
    - log LC_{50} (mM/L) = -1.46 - 0.18 log K_{ow}
- SARs are based on measured toxicity data
  - Toxicity is predicted at pH 7, TOC < 2 mg/L, moderate water hardness (150 mg/L $CaCO_3$), 100% active ingredient

ECOSAR Chemical Classes with Excess Toxicity

40 Chemical Classes

ECOSAR will identify appropriate SAR class

- Acid chlorides
- Acrylates
- Acrylates, methacrylates
- Alcohols, propargyl
- Aldehydes
- Amines, aliphatic
- Anilines
- Anilines, amino, meta or 1,3-substituted
- Anilines, amino, ortho or 1,2-substituted
- Anilines, amino, para, or 1,4-Substituted anilines, dinitroanilines
- Diazoniums, aromatic
- Epoxides, monoepoxides
- Epoxides, diepoxides
- Esters
- Esters, monoesters, aliphatic
- Esters, diesters, aliphatic
- Esters, phosphate
- Esters, phthalate
- Hydrazines
- Ketones, diketones, aliphatic

Chemical Domains in ECOSAR

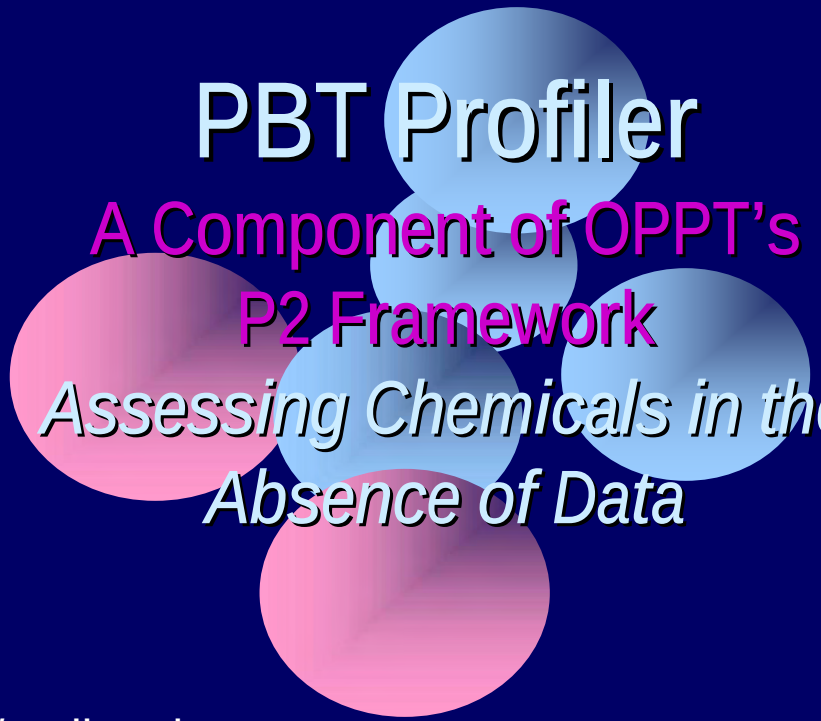
- Neutral Organics
- Classes with Excess Toxicity
- Surfactants
- Polymers*
- Dyes*
- Inorganics*
- Organometallics*

* SARs for polymers, organometallics, or inorganics are not yet available, and only a limited number of dye SARs are available in the current version of ECOSAR

ECOSAR

- “Stand-alone”, available for download at: www.epa.gov/oppt/newchems/21ecosar.htm
- Integral part of EPI Suite™, available for download at: www.epa.gov/opptintr/exposure/docs/EPISuitedl.htm

The PBT Profiler



http://www.epa.gov/pbt/tools/toolbox.htm

The PBT Profiler

- Estimates physical/chemical and fate properties
  - Persistence
  - Bioaccumulation: as fish BCF
  - Toxicity: fish chronic value (ChV) from ECOSAR
- Estimates distribution in water, soil, sediment, and air using Level III Fugacity model
- Compares P, B, and T estimates to EPA criteria
  - *New Chemicals PBT Policy – Federal Register: November 4, 1999 (Volume 64, Number 213), pages 60194-60204*
  - *TRI Reporting Criteria – Federal Register: October 29, 1999 (Volume 64, Number 209), pages 58666-58753*

P, B & T Criteria

Persistence	Not Persistent	Persistent	
Water, soil, sediment	< 60 d	≥ 60 d	> 180 d
Air	≤ 2 d		> 2 d

Bioaccumulation	Not Bioaccumulative	Bioaccumulative	
Fish BCF	< 1,000	≥ 1,000	≥ 5,000

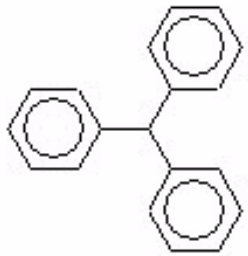
Toxicity	Low Concern	Moderate Concern	High Concern
Fish ChV (EPA New Chemical Program Criteria)	> 10 mg/L Or No Effects at Saturation	0.1-10 mg/L	< 0.1 mg/L

PBT Profiler Output

PBT Profiler Estimate = PBT

Screening estimates indicate this chemical may be a PBT - a P2 Assessment may allow further evaluation

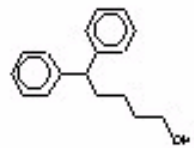
Media	Half-Life (days)	Percent in Each Medium	BCF	Fish ChV (mg/l)
Water	38	8%	2,700	0.027
Soil	75	53%		
Sediment	340	39%		
Air	1	1%		



P2 Considerations and more information

PBT Profiler Estimate = PBT

Media	Half-Life (days)	Percent in Each Medium	BCF	Fish ChV (mg/l)
Water	15	18%	170	0.12
Soil	30	73%		
Sediment	140	9%		
Air	0.75	1%		



P2 Considerations and more information

Oncologic®

Cancer Expert System

Prediction of concern levels for cancer potential based on “knowledge rules”

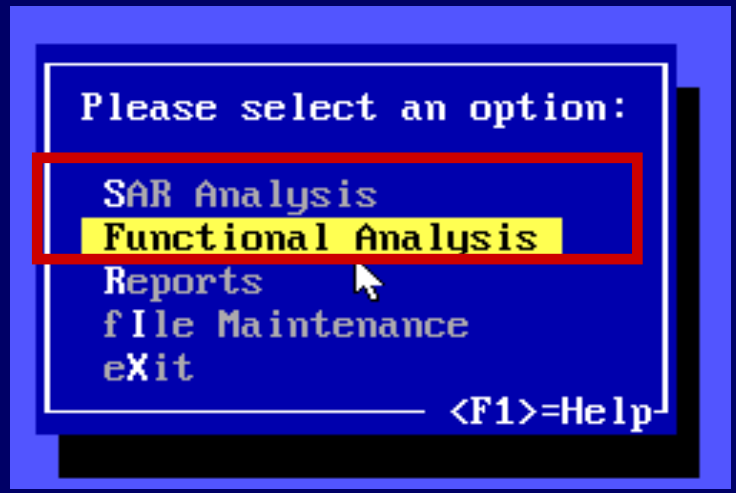
http://www.epa.gov/oppt/newchems/tools/oncologic.htm

OncoLogic® - Expert System How it Works

- Mimic the thinking and reasoning of human experts using knowledge based rules for chemical classes to predict cancer concern
  - Assigns a baseline concern level ranging from low to high
  - Evaluates how substituents on the chemical may affect carcinogenicity

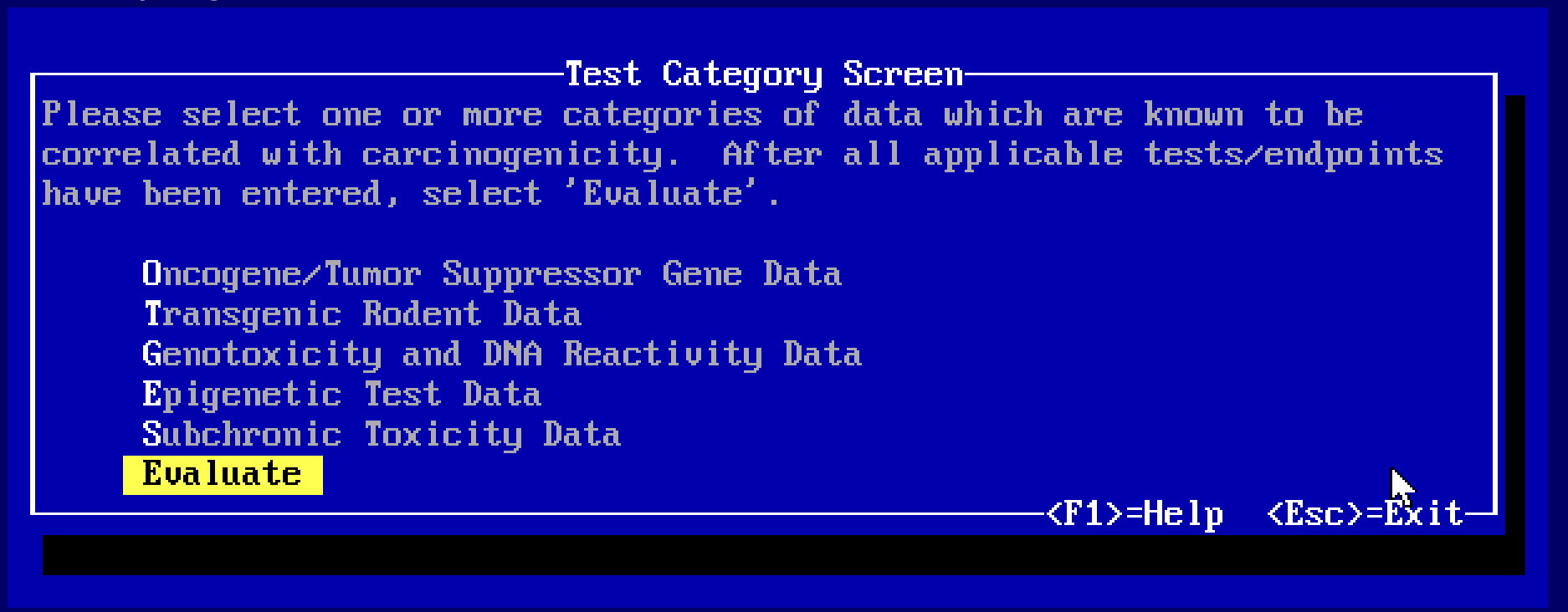
Running OncoLogic®

- Two methods to predict carcinogenicity
  - SAR Analysis
    - Knowledge rules
  - Functional Analysis
    - Uses results of specific mechanistic/non-cancer studies



The Functional Arm of OncoLogic®

- OncoLogic® can use results from some shorter-term tests to support a cancer concern.
- Results indicate whether chemical may be an initiator, promoter, or progressor



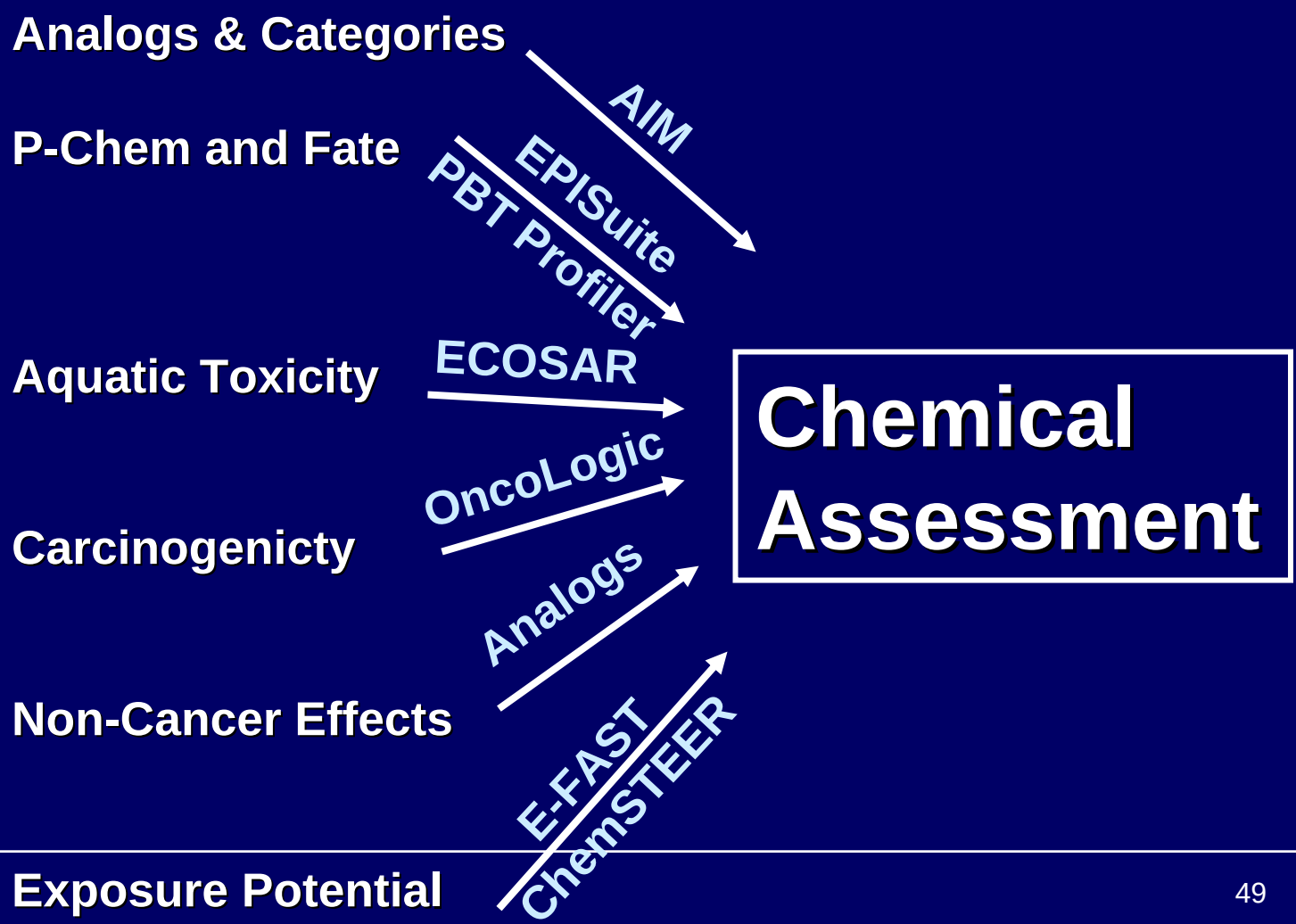
Major Data Sources Used to Develop Cancer Knowledge Rules

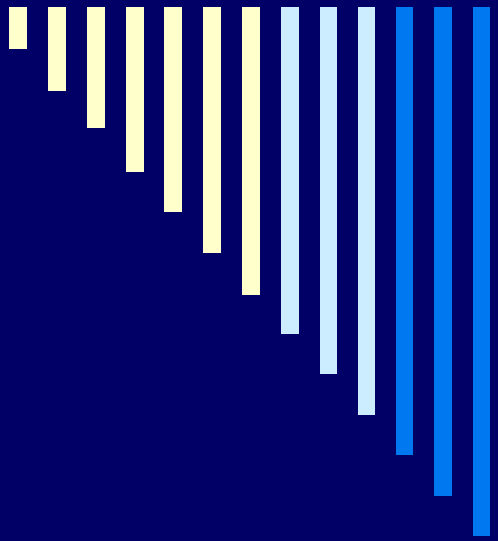
- Chemical Induction of Cancer monograph series
- IARC monograph series
- NCI/NTP technical reports
- Survey of compounds which have been tested for carcinogenic activity, PHS Publ. 149
- Non-classified EPA submission data from various EPA program offices

OncoLogic Concern Levels

OncoLogic Concern	Definition
Low	Unlikely to be carcinogenic
Marginal	Likely to have equivocal carcinogenic activity
Low – Moderate	Likely to be weakly carcinogenic
Moderate	Likely to be a moderately active carcinogen
Moderate – High	Highly likely to be a moderately active carcinogen
High	Highly likely to be a potent carcinogen

Computational Tools Are Used Together in OPPT Chemical Assessments





Peer Review, Validation, Evaluation for Regulatory Purposes

Considerations when Developing Approaches/Models in OPPT

- Developed to meet specific quality criteria
- May use TSCA CBI data
- Transparency and internal QC review is critical – we use these models every day in regulatory decision-making
- May participate in OECD or other reviews or programs
- OPPT uses established EPA policy and procedures for the development, verification, and use of tools and models:
  - EPA Peer Review Handbook
  - Information Quality Guidelines
  - Science Advisory Board Consultations & Reviews
  - EPA Council for Regulatory Environmental Modeling (CREM) Guidance
  - Risk Assessment Forum Guidance

Verification/Validation Procedures – Examples for SARs

- *U.S. EPA/European Community (EC) Joint Project on the Evaluation of (Quantitative) Structure Activity Relationships;* found our methods to be accurate 60-90% of the time depending on the endpoint assessed
- OECD QSAR Validation Principles
- OECD QSAR Validation Guidance

Peer Review of Screening Level Models

- 2005-2006 – SAB Peer Review of EPI Suite™
- 2000-2005 – Independent Peer Review of E-FAST
- 2003-2006 – Independent Peer Review of ChemSTEER evaporation from open surface models
- Ongoing – Independent Peer Review of Generic Scenarios in ChemSTEER
- 1992 and 1998-1999 – Peer Review of Oncologic
- 1989 - 2006 - Ongoing Independent Peer Reviews of SAR Equations in ECOSAR

OPPT Models and Methods Widely Used

- New Chemicals
  - PMN Reviews
  - Sustainable Futures
- Existing Chemicals
  - Design for the Environment
  - PBT Initiative
  - HPV Chemicals
- Office of Pesticide Programs – Inerts
- Office of Water – OncoLogic for CCL
- Office of Air Quality Planning and Standards - HAPS
- Other Federal Agencies
  - Customs & Border Protection
  - Dept. of Defense
  - Food & Drug Administration
  - Federal Aviation Administration
- International
  - European Union
  - Canada
  - The Netherlands

US Contributions to International Activities on the Use and Application of Predictive Modeling

EPA/OPPT Contributions to International Efforts

- 1993 – US & EC Validation Exercise & Publication: US *EPA/EC Joint Project on the Evaluation of (Quantitative) Structure Activity Relationships (QSARS)*
- 2004 – OECD Workshop & Publication: *Principles for Validation of (Q)SARs*
- 2006 – OECD Publication: *Regulatory Uses and Applications in OECD Member Countries of (Quantitative) Structure-Activity Relationships [(Q)SAR] Models in the Assessment of New and Existing Chemical*
- 2005 – OECD Exposure Policy Dialogue & Publication: *Approaches to Exposure Assessment in OECD member countries*
- 2007 – OECD Publication: *Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models*
- Ongoing – OECD (Q)SAR Steering Committee: defines strategic directions and develops guidance for developing and implementing (Q)SAR approaches for regulatory purposes in OECD member countries
- Ongoing – OECD Environmental Exposure Assessment Task Force: defines strategic directions and develops exposure methods, models and guidance

Organization for Economic Co-operation and Development

QSAR Application Toolbox

The Toolbox development is a collaborative effort to make specialized QSAR software and data from individual organisations accessible to all OECD stakeholders.

Danish Ministry of the Environment
Environment Canada
European Chemicals Bureau
Japanese MITI
US EPA
…

OECD (Q)SAR Project

- OECD Principles for the Validation, for Regulatory Purposes, of (Q)SAR Models (2004)
- Guidance Document on the Validation of (Q)SAR Models (2007)
- (Q)SAR Application Toolbox phase 1 (2006-2008)

Goals for the Pilot Version of the (Q)SAR Application Toolbox - Proof-of-Concept-

- Demonstrate that the Toolbox concept will make many QSAR methods readily accessible
- Apply computational methods to the formation of chemical categories and filling data gaps
  - Facilitate selection of chemical analogues and categories
  - Integrate metabolism/mechanisms with categories/(Q)SAR
- Illustrate the importance of the domain of application in making reliable estimates
- Integrate existing data, expert knowledge and computational methods to facilitate hazard assessments

Functionalities included in the (Q)SAR Application Toolbox

- PROFILING
  - Describe the chemical(s) of Interest
  - Is the chemical included in regulatory inventories or existing chemical categories?
  - Has the chemical already been assessed by other agencies/organisations?
- ENDPOINTS
  - Would you like to search for available data on assessment endpoints for each chemical?

Functionalities included in the (Q)SAR Application Toolbox

- CATEGORY DEFINITION
  - Explore a chemical list for possible analogues for each chemical?
  - Group chemicals based on molecular similarity and reactivity analysis?
  - Identify chemicals with anomalous metabolic pathways or toxicity mechanisms?
  - Group chemicals based on common metabolite?

Typical Queries included in the (Q)SAR Application Toolbox

- FILLING DATA GAPS
  - Fill data gaps in a chemical category using read-across, trend analysis or QSAR models
- REPORT
  - Design a data matrix of a chemical category for printing/exporting results

Contractor / Developer

Laboratory of Mathematical Chemistry

Bourgas, Bulgaria

http://www.oasis-lmc.org/

A collaborative effort of all OECD member countries and stakeholders

Special thanks:

- European Commission
- CEFIC
- Danish Ministry of the Environment
- US EPA
- Environment Canada
- NITE Japan

Organization for Economic Co-operation and Development

QSAR Application Toolbox

U.S. Contributions Phase 1:

- EPISuite – OPPT
- ECOSAR – OPPT
- New Chemicals Categories (Alerts) - OPPT
- ECOTOX – ORD

U.S. Contributions Phase 2:

- Ecotoxicity Structural Alerts – OPPT (ECOSAR) & ORD (ASTER)
- OncoLogic or Structural Alerts – OPPT
- ER Binding Structural Alerts – ORD
- ECOTOX – ORD (updates)
- Mammalian Effects DB – ORD
- ToxRef DB – ORD & OPP

Logical sequence of components usage

Chemical input	Profiling	Endpoints	Category Definition	Filling data gap	Report

Logical sequence of components usage

Chemical input	Profiling	Endpoints	Category Definition	Filling data gap	Report

User Alternatives for Chemical ID:

A. Single target chemical

- **Name**
- **CAS#**
- **SMILES/InChi**
- **Draw Chemical Structure**
- **Select from User List/Inventory**

B. Group of chemicals

- **User List/Inventory**
- **Specialized Databases**

➢ **Toolbox Inventories:**

- ✓US EPA TSCA
- ✓Canadian DSL
- ✓OECD HPVCs,
- ✓US EPA HPVCs
- ✓EU EINECS
- ✓Japanese MITI
- ✓DANISH EPA

OECD Toolbox v 0.5

QSAR Application Toolbox

Organization for Economic Co-operation and Development

Options | Tracks

Chemical input | Profiling | Endpoints | Category definition | Filling data gap | Report

Single chemical

- Chemical Name
- CAS #
- SMILES / InChi
- Drawing
- Select from an existing lists
- Select from an inventory

Chemical list

- User Lists
- Regulatory Inventories
- Database

Reset

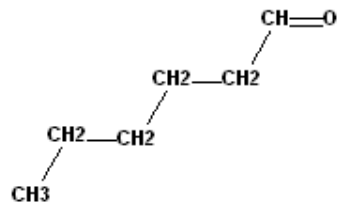
Structure drawing

C | H | 1,2 | T | C | Single | H

Smiles/InChi: CCCCCC=O — Draw

N/A — Edit names

Templates | Work



drag the mouse to move structure, alt + drag to zoom, shift + drag to rotate

OK | Cancel

Single chemical

Start | Inbox - Micros... | Impots.gouv.f... | interim report | QSAR_Toolbo... | OECD Toolbo... | EN | 11:09 AM

Logical sequence of components usage

Chemical input	Profiling	Endpoints	Category Definition	Filling data gap	Report

General characterization of chemical

- **Inventory Affiliation: HPVC, LPVC, TSCA, DSL, etc.**
- **Chemical Type: polymers, mixtures, discrete, inorganic/organic**
- **Chemical Class: aldehyde, phenol, esters, acids etc.**
- **In Existing Category? (OECD, US EPA etc)**
- **Previous Hazard/Risk Assessment Report?**
- **Hazard Indentification:**
  - **Protein/DNA binding**
  - **Structural Alerts (BioByte parent analogue)**
  - **Cramer Classification (ToxTree)**

OECD Toolbox v 0.5

QSAR Application Toolbox

Organization for Economic Co-operation and Development

Options | Tracks | Chemical input | Profiling | Endpoints | Category definition | Filling data gap | Report | STOP

Apply

Profiling methods

- **Predefined**
  - [x] OECD categorization
  - [x] US EPA Categorization
  - [x] Database affiliation
  - [x] Inventory affiliation
  - [x] Substance type
- **Mechanistic**
  - [x] OASIS Acute Toxicity MOA
  - [x] DNA Binding
  - [x] Protein Binding
  - [x] Superfragment profiling
  - [ ] Cramer Classification Tree (Tc
  - [ ] Veerhar/Hermens reactivity ru
- **Empiric**
  - [ ] Chemical elements
  - [ ] Groups of elements
  - [x] Natural functional groups

Metabolism

- **Documented**
  - [ ] Observed Microbial metabolisr
  - [ ] Observed Liver metabolism
- **Simulated**
  - [ ] Microbial metabolism simulator
  - [ ] GI tract simulator
  - [ ] Liver metabolism simulator

Show Category Boundaries

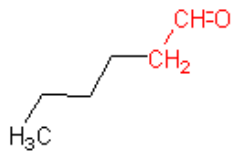
Create a new profiler

Delete profiler

Endpoints	1 (Target)
Structure	
Substance Information	
CAS Number	66-25-1
OECD Global portal	eChemPortal
Name (OECD name)	hexaldehyde
Structural Formula	C(=O)CCCCC
Profile	
OECD categorization	(N/A)
US EPA Categorization	Aldehydes
Database affiliation	Danish EPA ECOTOX OASIS Aquatic
Inventory affiliation	Danish EPA Canadian DSL US EPA HPVC EU EINECS MITI Japan US EPA TSCA
Substance type	Discrete chemical
OASIS Acute Toxici...	Aldehydes
DNA Binding	No Binding
Protein Binding	Schiff base formation
Superfragment pro...	No superfragment
Natural functional ...	Aldehyde Alkane, primary Alkane, secondary

Profile Explainer

Target chemical



Classified by:
Protein Binding

Classified as:
Schiff base formation

Met requirements

The target chemical has the fragment CC{H}=O in its structure.
The target chemical does not have the fragment c1ccccc1C{H}=O in its structure.

Single chemical

Start | Inbox - Micros... | Impots.gouv.f... | interim report | QSAR_Toolbo... | OECD Toolbo... | EN | 11:11 AM

Protein Binding (mechanistic) - Profiling Scheme Browser

Advanced

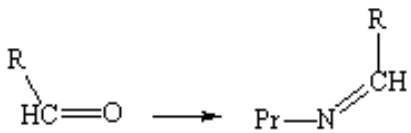
Protein Binding - Category definitions

- Arylcarboxylate aminolysis
- Disulfide exchange
- Disulfide formation
- Electrophilic substitution of arenesulfinic ...
- Electrostatic interaction with proteins
- Free radical protein adduct formation
- Haloalkane free radical protein adduct fo...
- Mercury thiolate formation
- Michael-type nucleophilic addition
- Michael-type nucleophilic ring opening
- N - Hydroxylamine condensation with pr...
- Nucleophilic addition to azomethynes
- Nucleophilic addition to carbodiimides
- Nucleophilic addition to dithiane-type az...
- Nucleophilic addition to isocyanates
- Nucleophilic addition to isothiocyanates
- Nucleophilic addition to ketones
- Nucleophilic addition to lactones
- Nucleophilic addition to thiocyanates
- Nucleophilic cycloaddition to diketones
- Nucleophilic heterocycle ring opening
- Nucleophilic substitution of acyl halides
- Nucleophilic substitution of alkyl phospho...
- Nucleophilic substitution of alkyl sulfates
- Nucleophilic substitution of alkyl sulfonates
- Nucleophilic substitution of cyclic dicarbo...
- Nucleophilic substitution of dithiocarbami...
- Nucleophilic substitution of dithiocarboxy...
- Nucleophilic substitution of haloaliphatics
- Nucleophilic substitution of haloaromatics
- Nucleophilic substitution of N - Trihalome...
- Nucleophilic substitution of organic acid ...
- Nucleophilic substitution of phosphonic a...
- Nucleophilic substitution of sulfonyl halides
- Nucleophilic substitution of thiocyanates
- Peroxy acid free radical decomposition
- Protein cyanide adduct formation
- **Schiff base formation**

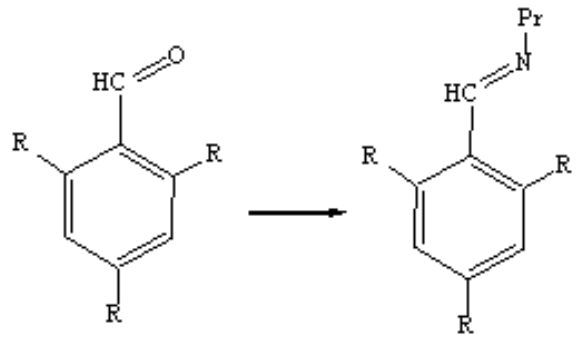
Profile Description

This category includes chemicals that potentially can cause skin sensitization effect as a result of protein conjugation via **Schiff base formation**.

The possible electrophilic groups acting by this mechanism are illustrated below:



R = H, alkyl



R = CN, NO_2, CF_3

All aldehydes (with the exception of aryl aldehydes which possess a deactivated aldehyde group) can potentially undergo Schiff base formation with a primary amine, which is a reversible reaction (optimal at pH 3-4) and proceeds in two stages via a tetrahedral intermediate.

References:

Lipnick, R. L. Outliers: their origin and use in the classification of molecular mechanisms of toxicity, QSAR in enviromental toxicology - IV, J. L. M. Hermens A.Opperhuizer, (Eds), Elsevier, Amsterdam, 1991, pp.131-153

Patlewicz, G., Basketter, D. A., Smith, C. K., Hotchkiss, S. A. M., Roberts, D. W: Skin-sensitisation structure-activity relationships for aldehydes. Contact Dermatitis, 2001, 44, 331-336

Logical sequence of components usage

Chemical input	Profiling	Endpoints	Category Definition	Filling data gap	Report

Finding Data for SIDS and Other Endpoints

- Selecting Data Base(s):
  - ✓ Toolbox databases
    - Publicly available
    - Proprietary databases

- Selecting type of extracting data:
  - ✓ Measured Data
  - ✓ Estimated Data
  - ✓ Both

OECD Toolbox v 0.5

QSAR Application Toolbox

Organization for Economic Co-operation and Development

Options | Tracks

Chemical input | Profiling | Endpoints | Category definition | Filling data gap | Report

Data Summaries
- Tested
- Estimated
- Both

Gather data

Excel import | Excel export

Databases
- Danish EPA
- ECETOC
- ECOTOX
- OASIS Aquatic
- OASIS Bioaccumulation
- OASIS Biodegradation
- OASIS ERBA
- OASIS Genotox
- OASIS Skin sensitization
- OECD HPVC Inventory DB

Endpoints	1 (Target)
Structure	
Substance Information	
Profile	
Ecotoxicological Information	
Aquatic Toxicity (1/4)	T: 2.20E-004 mol/L, 1.80E-...

Data points

#	Endpoint	Value	DIMENSIONS	CONDITI	BIOASSAY	ORGAN/TISSUE	ADMINISTRATIC	COMMENT	AUTHOR	NAME	SOURCE	DATE	DA QU
1	LC50	2.20E-0C	mol/L	96 h	Pimephales promelas	Sp.Com. Name:Fathead minnow; Org. Age :33; Age Units:d; Org.Comment:0.0 MM, 0.000 G;	Chemical Comment:99%	Effect :MOR; Effect Measurement:MORT; Exposure Type :F; Chemical Analysis:M; LC50 Concentration Range :21000 to 23000ug/L	Geiger, D.L., L.T. Brooke, and D.J. Call	Acute Toxicities of Organic Chemicals to Fathead Minnows (Pimephales promelas)	Ctr.for Lake Superior Environ.Stud., Univ.of Wisconsin-Superi Superior, WI 5:332 p.	01-Jan-9	
2	LC50	1.80E-0C	mol/L	96 h	Pimephales promelas				Russom, C.L., Bradbury, S.P., Broderius, S.J., Drummond, R.A. and Hammermeister,	PREDICTING MODES OF TOXIC ACTION FROM CHEMICAL STRUCTURE:	Environmental Toxicology and Chemistry, Vol. 16, No. 5, pp. 948Ö967.	01-Jan-9	

Transpose

Single chemical

Start | Inbox - Micros... | Impots.gouv.f... | interim report | QSAR_Toolbo... | OECD Toolbo... | EN | 11:17 AM

Logical sequence of components usage

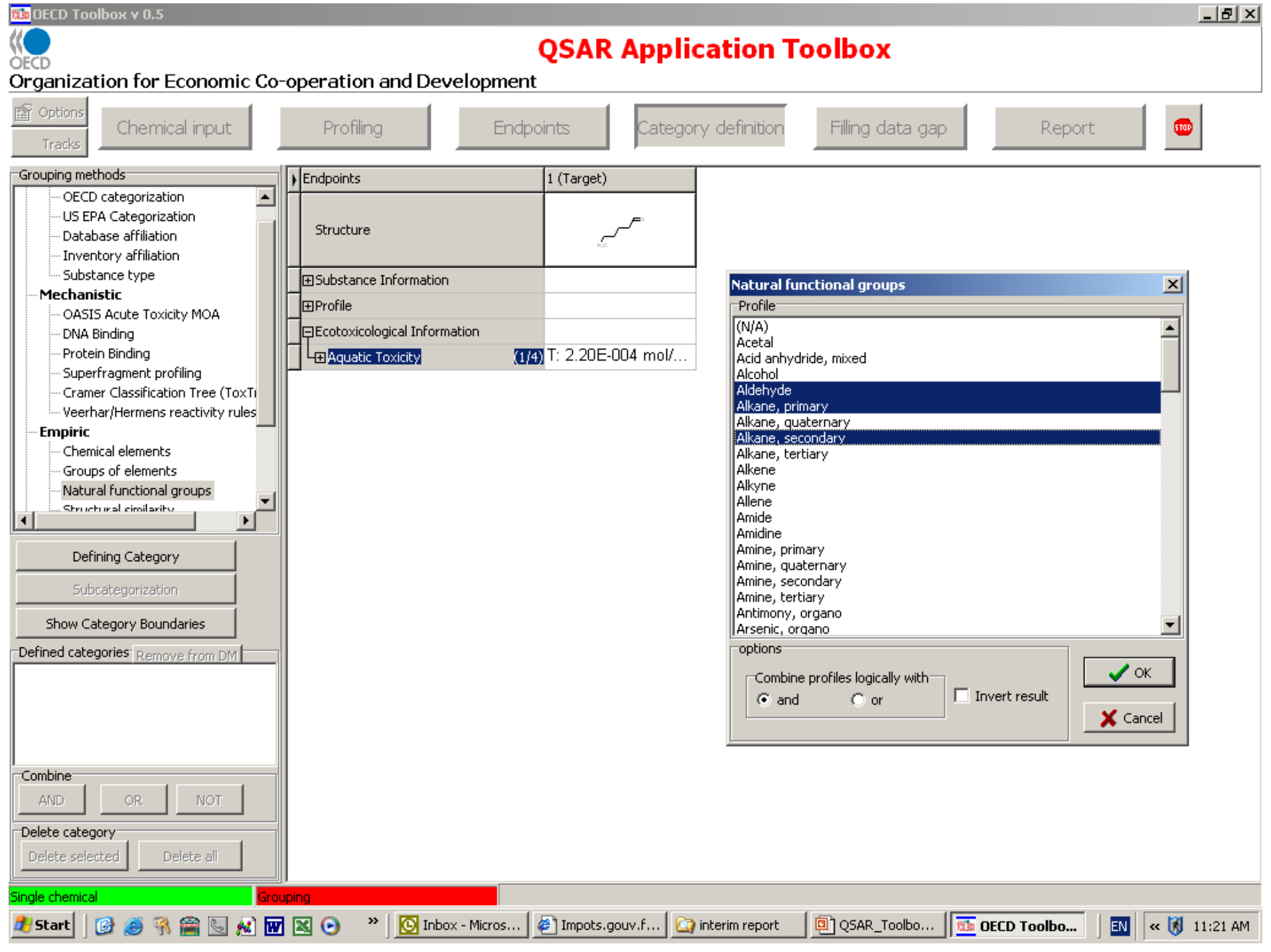
Chemical input	Profiling	Endpoints	Category Definition	Filling data gap	Report

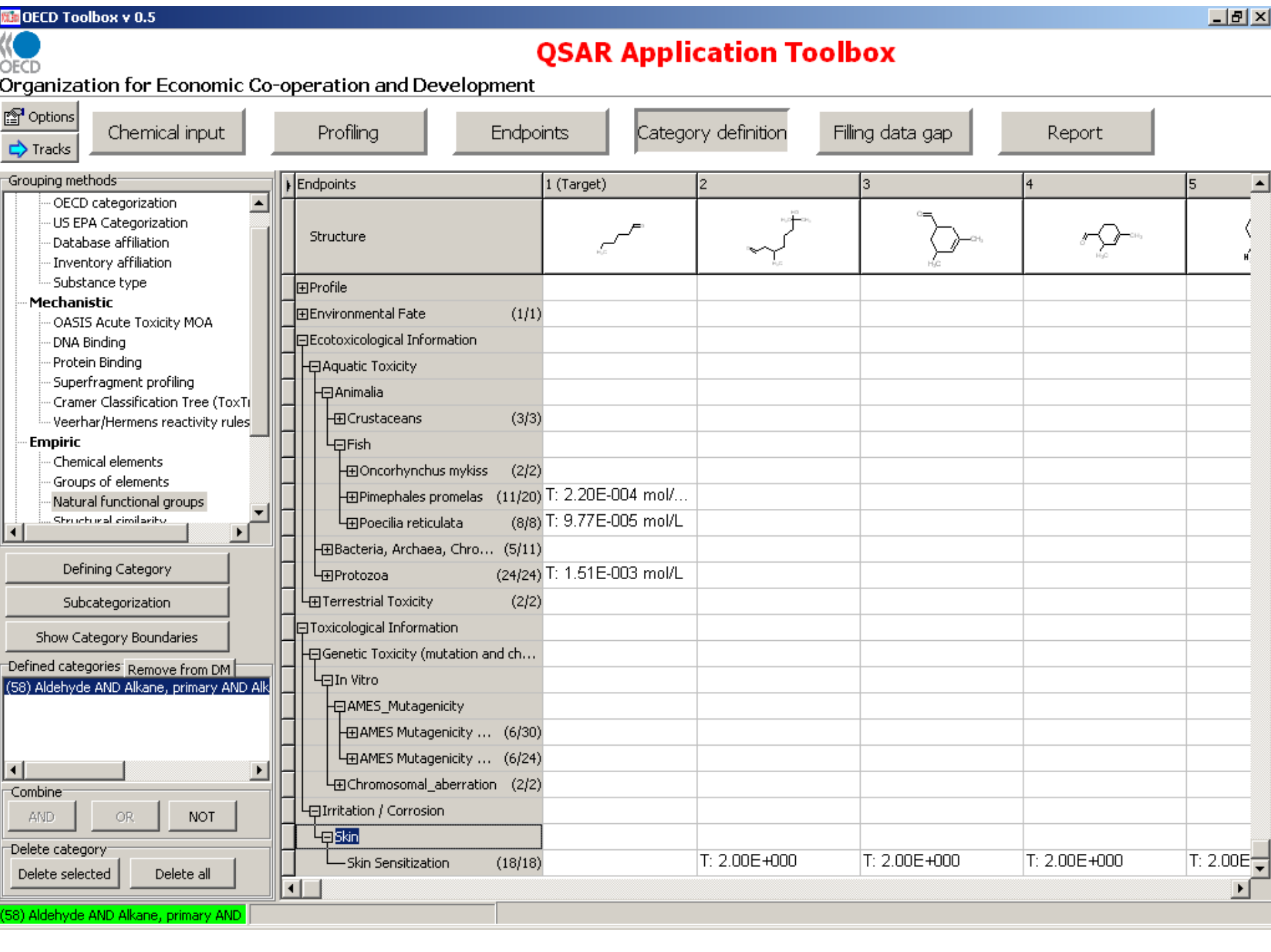
Forming Categories:

- **OECD Categories**
- **Other Established Categories (EPA)**
- **Searches for Analogues (AIM-EPA)**
- **Fragments:**
  - ✓ **DNA binding alerts**
  - ✓ **Protein binding alerts**
  - ✓ **BioByte Superfragments**
  - ✓ **Atom-centered fragments**

Pruning Categories-Subcategory Formation

- **Mechanisms boundaries**
- **Metabolism boundaries**


OECD Toolbox v 0.5
QSAR Application Toolbox
Organization for Economic Co-operation and Development
Options
Tracks
Chemical input
Profiling
Endpoints
Category definition
Filling data gap
Report
Grouping methods
OECD categorization
US EPA Categorization
Database affiliation
Inventory affiliation
Substance type
Mechanistic
OASIS Acute Toxicity MOA
DNA Binding
Protein Binding
Superfragment profiling
Cramer Classification Tree (ToxTr
Veerhar/Hermens reactivity rules
Empiric
Chemical elements
Groups of elements
Natural functional groups
Structural similarity
Endpoints
1 (Target)
Structure
Substance Information
Profile
Ecotoxicological Information
Aquatic Toxicity
(1/4)
T: 2.20E-004 mol/...
Natural functional groups
Profile
(N/A)
Acetal
Acid anhydride, mixed
Alcohol
Aldehyde
Alkane, primary
Alkane, quaternary
Alkane, secondary
Alkane, tertiary
Alkene
Alkyne
Allene
Amide
Amidine
Amine, primary
Amine, quaternary
Amine, secondary
Amine, tertiary
Antimony, organo
Arsenic, organo
options
Combine profiles logically with
and
or
Invert result
OK
Cancel
Defining Category
Subcategorization
Show Category Boundaries
Defined categories
Remove from DM
Combine
AND
OR
NOT
Delete category
Delete selected
Delete all
Single chemical
Grouping
Start
Inbox - Micros...
Impots.gouv.f...
interim report
QSAR_Toolbo...
OECD Toolbo...
EN
11:21 AM


OECD Toolbox v 0.5
QSAR Application Toolbox
Organization for Economic Co-operation and Development
Options
Tracks
Chemical input
Profiling
Endpoints
Category definition
Filling data gap
Report
Grouping methods
OECD categorization
US EPA Categorization
Database affiliation
Inventory affiliation
Substance type
Mechanistic
OASIS Acute Toxicity MOA
DNA Binding
Protein Binding
Superfragment profiling
Cramer Classification Tree (ToxTr
Veerhar/Hermens reactivity rules
Empiric
Chemical elements
Groups of elements
Natural functional groups
Structural similarity
Defining Category
Subcategorization
Show Category Boundaries
Defined categories
Remove from DM
(58) Aldehyde AND Alkane, primary AND Alk
Combine
AND
OR
NOT
Delete category
Delete selected
Delete all
Endpoints
1 (Target)
2
3
4
5
Structure
Profile
Environmental Fate (1/1)
Ecotoxicological Information
Aquatic Toxicity
Animalia
Crustaceans (3/3)
Fish
Oncorhynchus mykiss (2/2)
Pimephales promelas (11/20) T: 2.20E-004 mol/...
Poecilia reticulata (8/8) T: 9.77E-005 mol/L
Bacteria, Archaea, Chro... (5/11)
Protozoa (24/24) T: 1.51E-003 mol/L
Terrestrial Toxicity (2/2)
Toxicological Information
Genetic Toxicity (mutation and ch...
In Vitro
AMES_Mutagenicity
AMES Mutagenicity ... (6/30)
AMES Mutagenicity ... (6/24)
Chromosomal_aberration (2/2)
Irritation / Corrosion
Skin
Skin Sensitization (18/18) T: 2.00E+000 T: 2.00E+000 T: 2.00E+000 T: 2.00E
(58) Aldehyde AND Alkane, primary AND

Logical sequence of components usage

Chemical input	Profiling	Endpoints	Category Definition	Filling data gap	Report

Data gaps filling approaches

- **Read-across**
- **Trend analysis**
- **QSAR models**

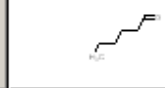
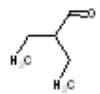
OECD Toolbox v 0.5

QSAR Application Toolbox

Organization for Economic Co-operation and Development

Options | Tracks | Chemical input | Profiling | Endpoints | Category definition | Filling data gap | Report

Read-across | Trend analysis | (Q)SAR models | Apply

Endpoints	1 (Target)	7	9	16	17
Structure					
Tetrahymena pyriformis (24/24)	T: 1.51E-003 mol/L	T: 2.24E-003 mol/L	T: 1.13E-003 mol/L	T: 1.19E-003 mol/L	T: 1.00E

Target endpoint:
Ecotoxicological Information Aquatic Toxicity
Protozoa Tetrahymena pyriformis

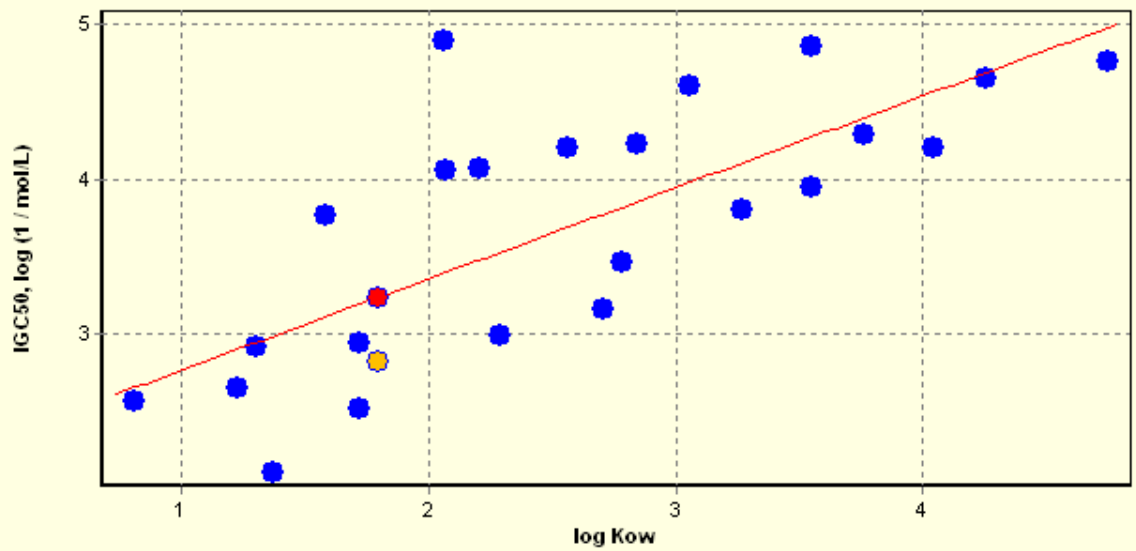
Select data points...

Endpoint
- As it is
- log(Endpoint)
- log(1/Endpoint)

Data
- All values
- Average value
- Min value
- Max value

Descriptors | Endpoint | Adequacy | Cumul. frequency | Statistic

Trend analysis evaluation, making a linear approximation,
based on 'Current subcategory', containing 24 data points in 24 analogue structure(s)
Observed target value(s): 2.82; Predicted target value: 3.24



Descriptors: log Kow

show descriptors in use only

Prediction

Accept | Cancel

Data points

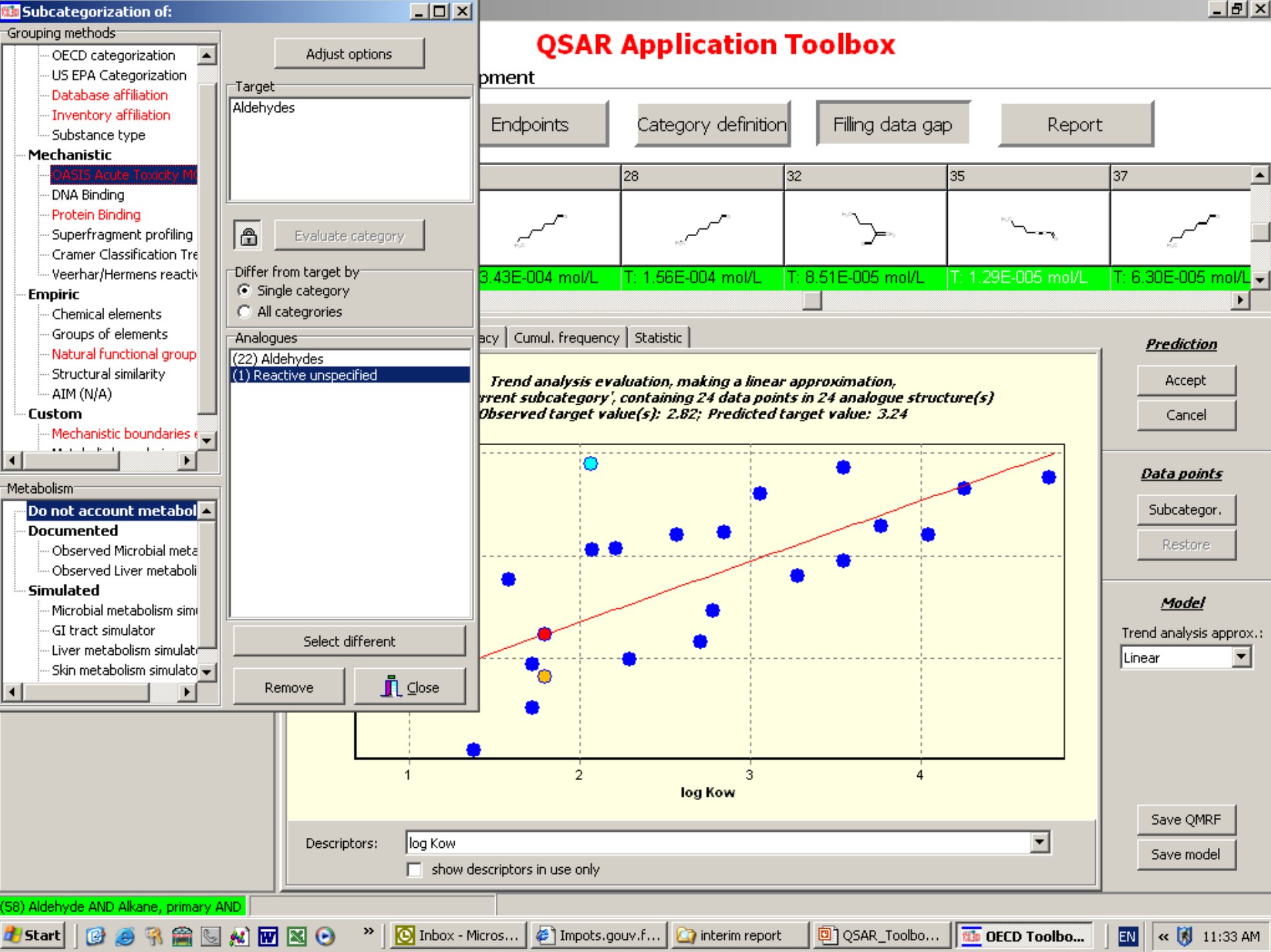
Subcategor. | Restore

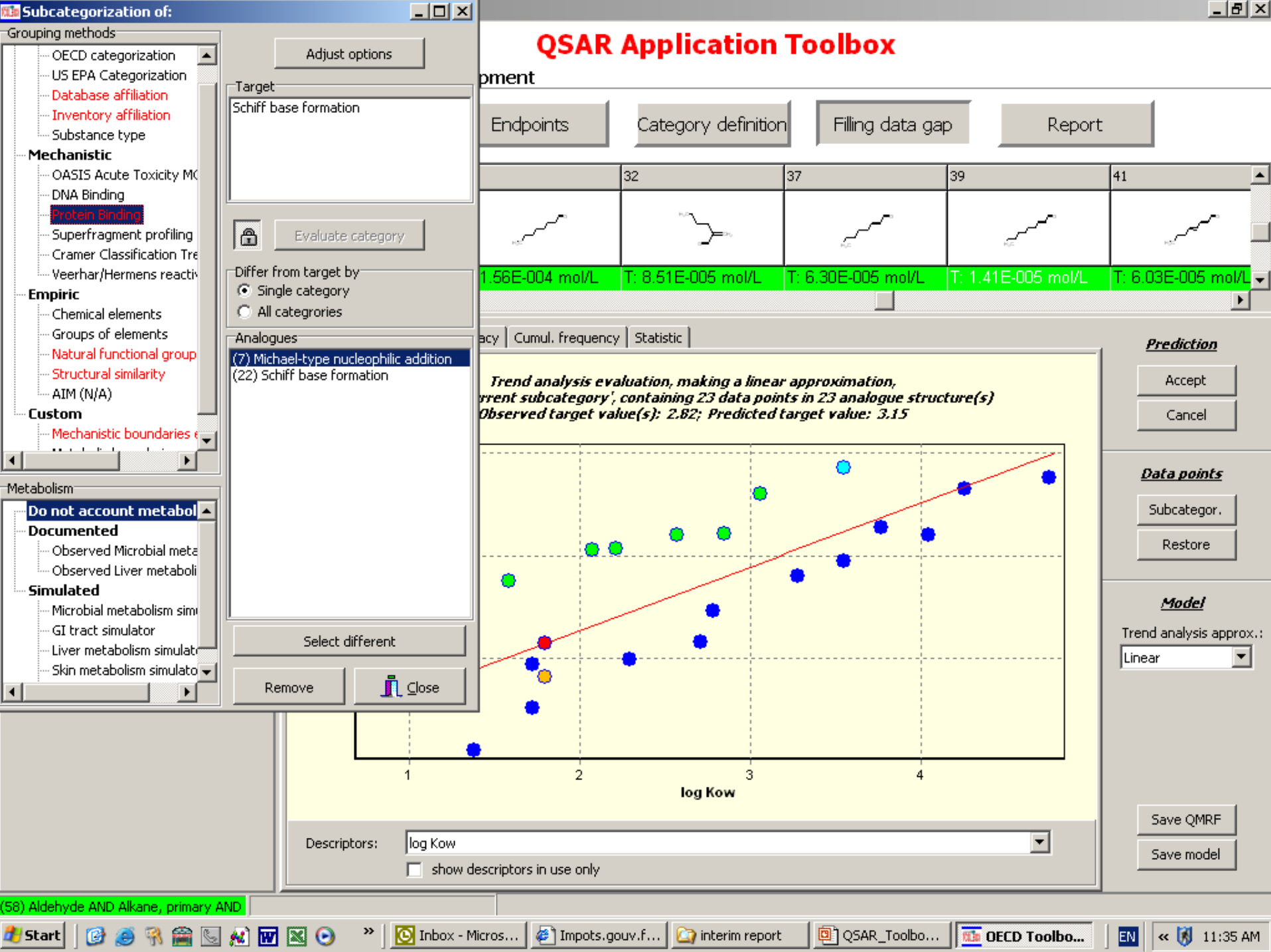
Model

Trend analysis approx.: Linear

Save QMRF | Save model

(58) Aldehyde AND Alkane, primary AND


Subcategorization of:
Grouping methods
OECD categorization
US EPA Categorization
Database affiliation
Inventory affiliation
Substance type
Mechanistic
OASIS Acute Toxicity M
DNA Binding
Protein Binding
Superfragment profiling
Cramer Classification Tre
Veerhar/Hermens reacti
Empiric
Chemical elements
Groups of elements
Natural functional group
Structural similarity
AIM (N/A)
Custom
Mechanistic boundaries
Metabolism
Do not account metabol
Documented
Observed Microbial meta
Observed Liver metaboli
Simulated
Microbial metabolism simu
GI tract simulator
Liver metabolism simulat
Skin metabolism simulato
Adjust options
Target
Aldehydes
Evaluate category
Differ from target by
Single category
All categories
Analogues
(22) Aldehydes
(1) Reactive unspecified
Select different
Remove
Close
QSAR Application Toolbox
pment
Endpoints
Category definition
Filling data gap
Report
28
32
35
37
3.43E-004 mol/L
T: 1.56E-004 mol/L
T: 8.51E-005 mol/L
T: 1.29E-005 mol/L
T: 6.30E-005 mol/L
acy
Cumul. frequency
Statistic
Trend analysis evaluation, making a linear approximation,
rrent subcategory', containing 24 data points in 24 analogue structure(s)
Observed target value(s): 2.82; Predicted target value: 3.24
1
2
3
4
log Kow
Descriptors:
log Kow
show descriptors in use only
Prediction
Accept
Cancel
Data points
Subcategor.
Restore
Model
Trend analysis approx.:
Linear
Save QMRF
Save model
(58) Aldehyde AND Alkane, primary AND
Start
Inbox - Micros...
Impots.gouv.f...
interim report
QSAR_Toolbo...
OECD Toolbo...
EN
11:33 AM


Subcategorization of:
Grouping methods
OECD categorization
US EPA Categorization
Database affiliation
Inventory affiliation
Substance type
Mechanistic
OASIS Acute Toxicity MC
DNA Binding
Protein Binding
Superfragment profiling
Cramer Classification Tre
Veerhar/Hermens reactiv
Empiric
Chemical elements
Groups of elements
Natural functional group
Structural similarity
AIM (N/A)
Custom
Mechanistic boundaries
Metabolism
Do not account metabol
Documented
Observed Microbial meta
Observed Liver metaboli
Simulated
Microbial metabolism simu
GI tract simulator
Liver metabolism simulat
Skin metabolism simulato
Adjust options
Target
Schiff base formation
Evaluate category
Differ from target by
Single category
All categories
Analogues
(7) Michael-type nucleophilic addition
(22) Schiff base formation
Select different
Remove
Close
QSAR Application Toolbox
Endpoints
Category definition
Filling data gap
Report
32
37
39
41
1.56E-004 mol/L
T: 8.51E-005 mol/L
T: 6.30E-005 mol/L
T: 1.41E-005 mol/L
T: 6.03E-005 mol/L
Cumul. frequency
Statistic
Trend analysis evaluation, making a linear approximation,
rrent subcategory', containing 23 data points in 23 analogue structure(s)
Observed target value(s): 2.82; Predicted target value: 3.15
log Kow
Descriptors:
log Kow
show descriptors in use only
Prediction
Accept
Cancel
Data points
Subcategor.
Restore
Model
Trend analysis approx.:
Linear
Save QMRF
Save model
(58) Aldehyde AND Alkane, primary AND
Start
Inbox - Micros...
Impots.gouv.f...
interim report
QSAR_Toolbo...
OECD Toolbo...
EN
11:35 AM

OECD Toolbox v 0.5

QSAR Application Toolbox

Organization for Economic Co-operation and Development

Options | Tracks

Chemical input | Profiling | Endpoints | Category definition | Filling data gap | Report

Read-across | Trend analysis | (Q)SAR models | Apply

27	28	44	51	58
T: 3.43E-004 mol/L	T: 1.56E-004 mol/L	T: 7.75E-003 mol/L	T: 1.13E-004 mol/L	T: 6.20E-005 mol/L

Target endpoint:
Ecotoxicological Information Aquatic Toxicity
Protozoa Tetrahymena pyriformis

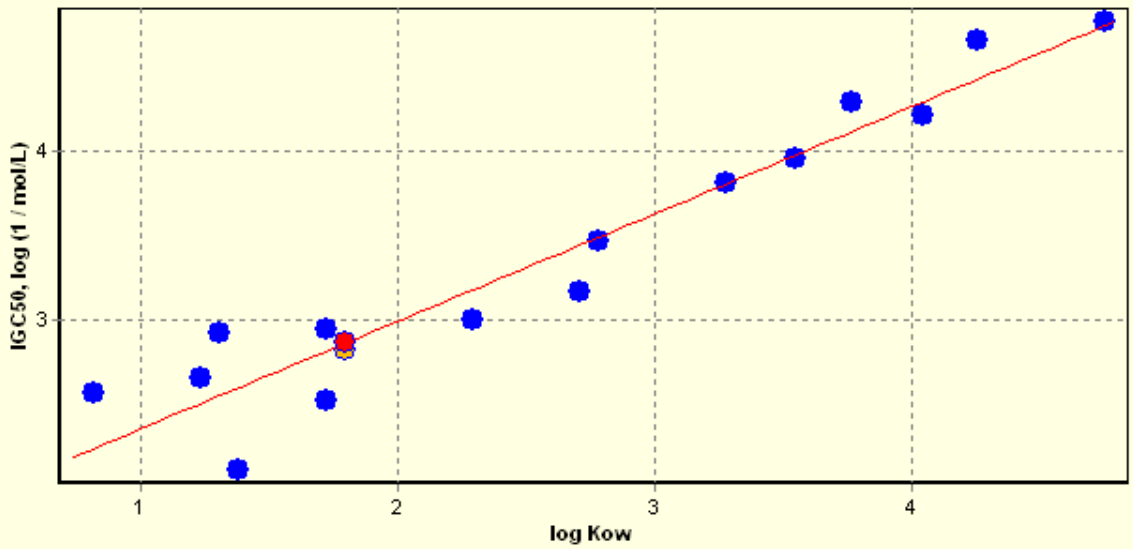
Select data points...

Endpoint
- As it is
- log(Endpoint)
- log(1/Endpoint)

Data
- All values
- Average value
- Min value
- Max value

Descriptors | Endpoint | Adequacy | Cumul. frequency | Statistic

Trend analysis evaluation, making a linear approximation,
based on 'Current subcategory', containing 16 data points in 16 analogue structure(s)
Observed target value(s): 2.82; Predicted target value: 2.86



Descriptors: log Kow

show descriptors in use only

Prediction

Accept | Cancel

Data points

Subcategor. | Restore

Model

Trend analysis approx.:
Linear

Save QMRF | Save model

(58) Aldehyde AND Alkane, primary AND

Start | Inbox - Micros... | Impots.gouv.f... | interim report | QSAR_Toolbo... | OECD Toolbo... | EN | 11:35 AM

Logical sequence of components usage

Chemical input	Profiling	Endpoints	Category Definition	Filling data gap	Report

Report the results

- **QMRF/QPRF/TERF**
- **Harmonized Templates**
- **SIDS Dossiers (Data matrix)**

Harmonised Templates
C & L Summaries
SIDS Dossiers
User-Defined Reports
RIVM Suggested Reports

History
Show history
Print setup
Print
Print preview

OECD Toolbox Application study history

Report created: 09.05.07 - 11:36

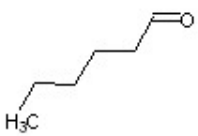
Chemical input *09.05.07 - 11:10*

Imput single target chemical by: SMILES

CAS # :66-25-1

Chemical name: hexaldehyde

SMILES: C(=O)CCCCC



Profiling *09.05.07 - 11:10*

Schedule

- September 2007: beta testing of proof-of-concept version:
- March 2008: Release of proof-of-concept version (free of charge)
- 2008-2011: Phase 2 (initial plans):
  - Other functionalities
  - Expansion of categorisation mechanisms
  - Addition of databases
  - Addition of QSARs
  - Communication with other tools
  - Long-term maintenance
  - Training material

www.oecd.org/env/existingchemicals/qsar