U.S. EPA Use of QSAR and Category Approaches in Profiling Hazards of Industrial Chemicals

Tala Henry
Office of Pollution Prevention and Toxics
U.S. Environmental Protection Agency
Washington, DC

QSAR 2008
Syracuse, NY
June 11, 2008

Office of Pollution Prevention and Toxics

- Review Pre-Manufacture Notices (PMN) for new industrial chemicals
- Testing, assessment, and risk reduction of existing industrial chemicals
- Management of “national chemicals” (e.g. PCBs, Hg)
- International chemical issues (e.g. POPs)
- Pollution prevention advocacy
- Partnership programs, e.g. HPVC Challenge, Green Suppliers Network, DfE and Green Chemistry

Chemical Management Programs and Workflow

- New Chemicals
- Existing Chemicals
  - High Production Volume (HPV)
    - Manufacture or Import @ > 1M lbs/yr
  - Medium Production Volume (MPV)
    - Manufacture or Import @ 25K – 1M lb/yr

Overview of *In silico* and SAR Tools/Models

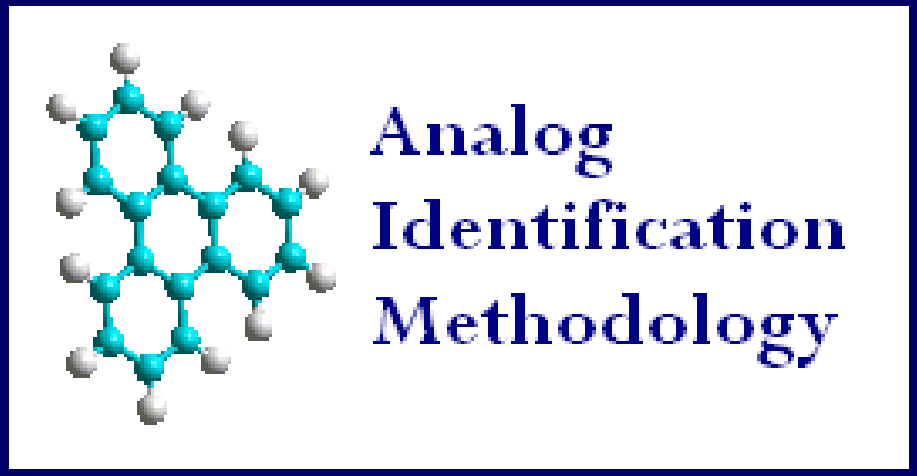
- Analogs & Categories
- AIM
- EPI Suite™
- ECOSAR
- Oncologic®
- PBT Profiler

- ChemSTEER
- E-FAST

Analogs & Categories

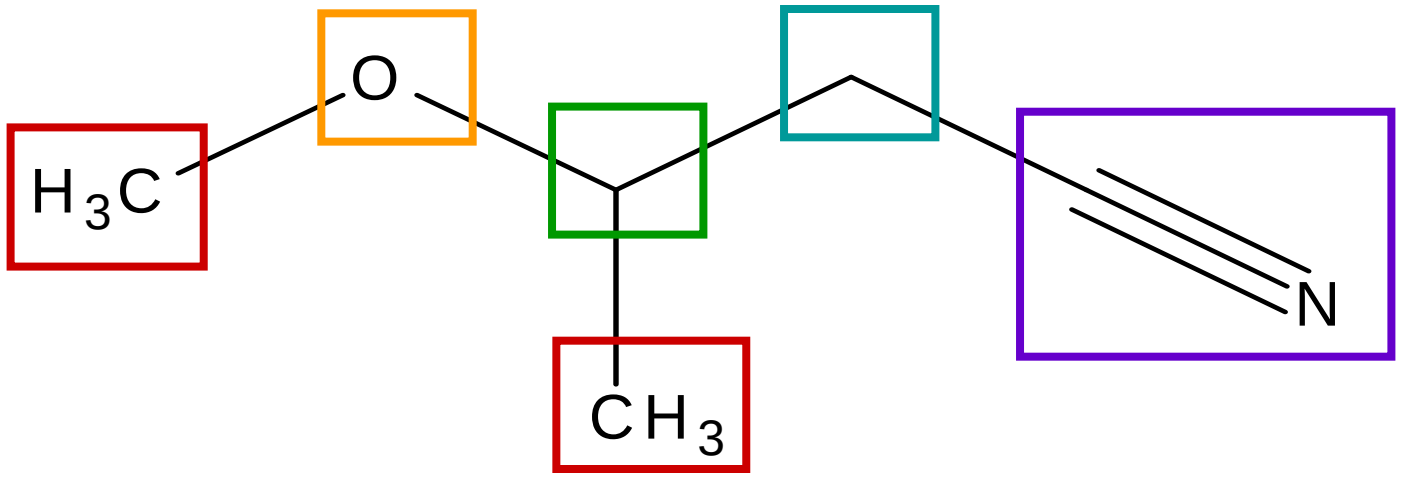
- Analogs - chemicals whose physicochemical and toxicological properties are likely to be similar as a result of structural similarity.
- Category - A group of analogs
- Similarities based on:
  - a common functional group (e.g., aldehyde, epoxide, ester, etc.)
  - common precursors and/or breakdown products
  - similar biological mechanisms of action
- 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

Analog Identification Methodology (AIM)



AIM Methodology

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



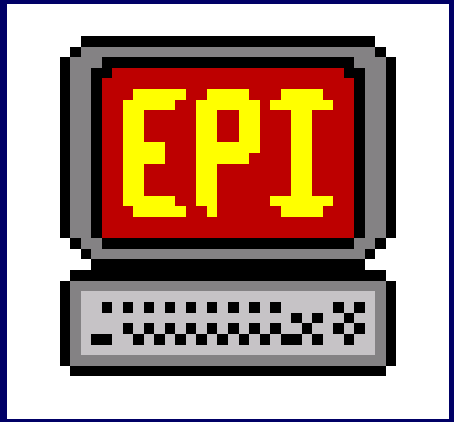
- Web-based program provides user with a list of chemical analogs linked to publically available toxicity data.
- From a database of 31,031 chemicals/structures that are indexed to publicly available toxicity sources such as: TSCATS, AEGLS, IRIS, HPV Challenge, ATSDR, NLM-HSDB, NTP, RTECS, IUCLID

AIM Clustering Tool/Category Builder

- AIM fragment matching algorithm expanded to organize data sets to identify “structural clusters” of chemicals
- Applying to multiple EPA chemical ‘universes’ (PMNs, HPV, 8(e), IUR) to formulate structure-based “clusters”

EPISuite™

Estimation Program Interface (EPI) Suite – a suite of physical/chemical property and environmental fate estimation models



See talks/posters by: Meylan, Boethling, Arnot

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

See talks/posters by: Meylan, Boethling, Arnot

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"

See talks/posters by: Meylan, Boethling, Arnot

ECOSAR

Aquatic toxicity estimates based on
Structure Activity Relationships (SAR)

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

ECOSAR
Aquatic Toxicity Profile

6 Endpoints
Fish LC_{50}
Daphnid LC_{50}
Green Algae LC_{50}
Fish Chronic Value (ChV)
Daphnid Chronic Value (ChV)
Green Algae (ChV)

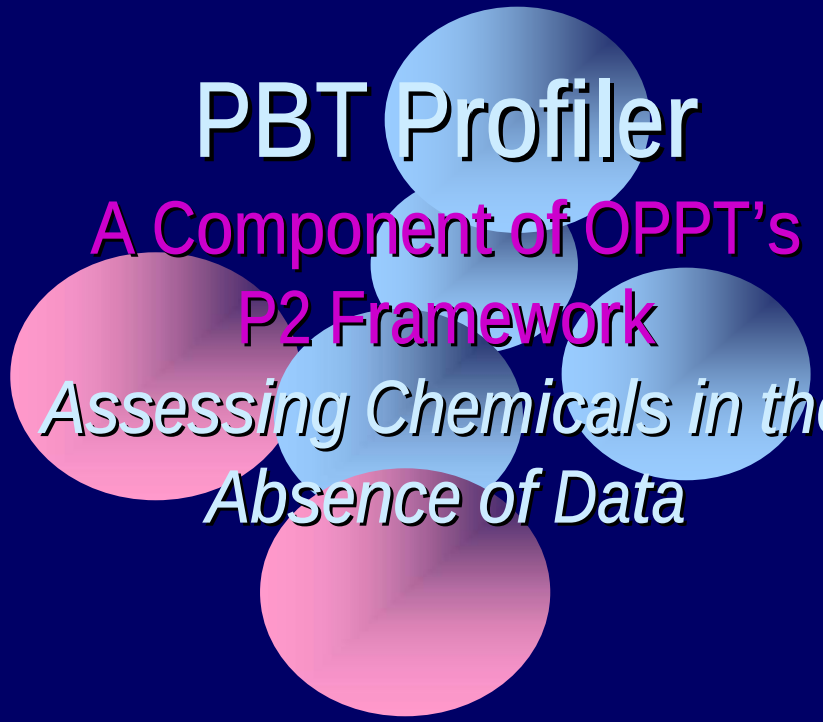
Aquatic Toxicity - ECOSAR

- ECOSAR v 0.99h – Currently Available
  - 44 Chemical Classes
  - 160 QSARs for the 44 classes based on confidential and publicly available experimental data
  - A ‘New Chemical’ Standard Aquatic Toxicity Profile is not available for all chemical classes.

- New Version of ECOSAR v 1.00 – Release June 2008
  - 130 SAR Classes
  - 440 QSARs for the 130 classes based on confidential and publicly available experimental data
  - An EPA New Chemical Standard Aquatic Toxicity Profile will be created using a combination of QSAR and/or alternative approaches such as acute-to-chronic ratios and alternate class QSARs based on scientific judgment.

See Poster by: Mayo-Bean

The PBT Profiler



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 and formats results in color-coded output
  - *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

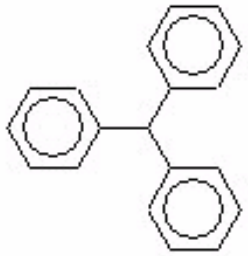
Toxicity	Low Concern	Moderate Concern	High Concern
Fish ChV (EPA New Chemical Program Criteria)	> 10 mg/L Or No Effects at Saturation	0.1-10 mg/L	< 0.1 mg/L

P2 and PBT Profiler

PBT Profiler Estimate = PBT

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

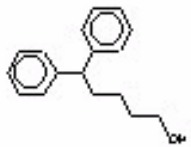
Media	Half-Life (days)	Percent in Each Medium	BCF	Fish ChV (mg/l)
Water	38	8%	2,700	0.027
Soil	75	53%		
Sediment	340	39%		
Air	1	1%		



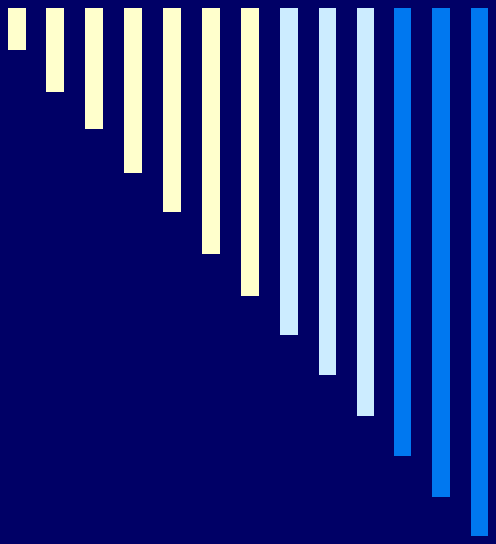
P2 Considerations and more information

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”

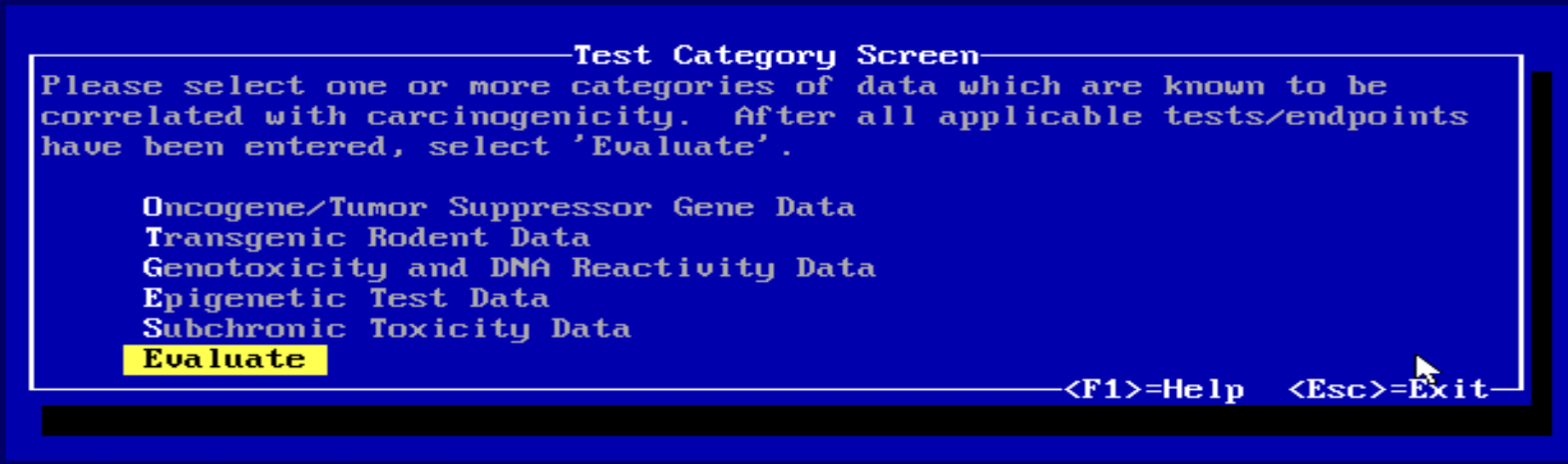
OncoLogic™ - Cancer Expert System

- Knowledge based rules for chemical classes to predict cancer concern
- Evaluates how substituents on the chemical may affect carcinogenicity
- Assigns a concern level ranging from low to high

Data Sources for Rules: IARC, NCI/NTP, Survey of Chems Tested, Non-CBI submission data from EPA Offices

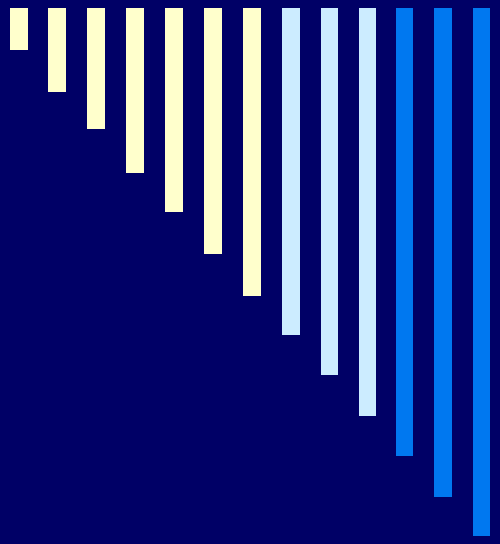
OncoLogic™

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



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



Application of Profiling Tools within Chemicals Programs Workflow

New Chemicals – TSCA Section 5

- 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 of chemical
- PMN Exemptions: R & D chemicals, Certain polymers, Export only, Low release/exposure (LoREX), Test Marketing (TME)
- Designed to prevent health and/or environmental risks before they occur
- Regulatory decisions must be made within 90 days; often in the absence of data
- Proven track record: ~1,500 PMNs per year; over 41,000 Section 5 notices reviewed to date

PMN Review Process

Chemical Review / Search Strategy (CRSS) [Analogs, P/Chem Properties, Use, etc.]	Day 8-12
Structure Activity Team (SAT) [Hazard evaluation]	Day 9-13
Exposure / Release Profiles Review [Exposure and releases evaluated]	Day 10-19
Focus Meeting [Decision to drop or regulate]	Day 15-20
Standard Review [Detailed risk assessment]	Day 21-70

Structure Activity Team Meeting

- Key technical meeting to consider potential hazard to human health and the environment
- Hazard Profile based on:
  - PChem properties
  - Routes of absorption
  - PMN data
  - Structure activity relationship (SAR) analysis: analogs & categories
  - QSAR estimates

EPI

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

Analog Identification Methodology

ECO SAR

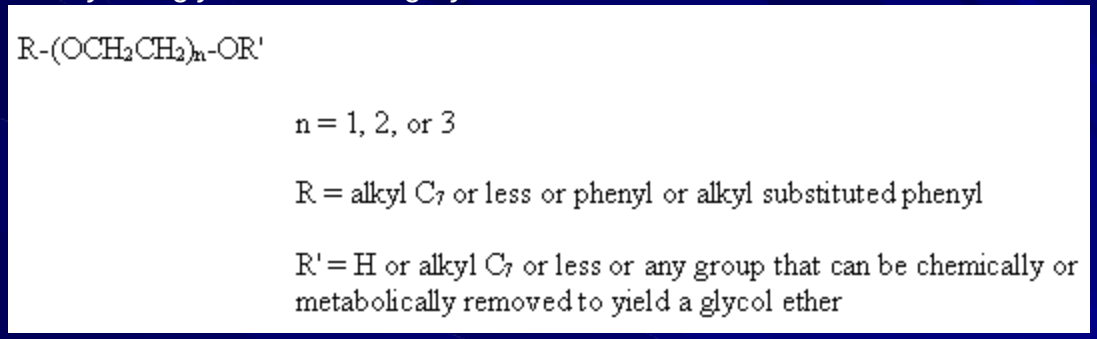
OncoLogic™

Analogs/Categories

EXAMPLE - NEW CHEMICALS CATEGORY – HUMAN HEALTH

Category: Ethylene Glycol Ethers Human Health

Definition. The ethylene glycol ether category is defined as follows:



Hazard Concerns. Short-chain ethylene glycol ethers are absorbed by all routes of exposure and have caused irritation of skin, eyes, and mucous membranes; hemolysis, bone-marrow damage, and leukopenia of both lymphocytes and granulocytes; direct and indirect kidney damage; liver damage, immunotoxicity, and central nervous system (CNS) depression. Short-chain ethylene glycol ethers are also developmental and reproductive toxicants. 2-Phenoxyethanol is known to cause hemolysis and eye irritation.

Boundaries. There is evidence that developmental toxicity is reduced going from the methyl to the butyl ether, and that it is reduced going from the ethylene glycol to the triethylene glycol. However, there is still a concern for maternal toxicity as reflected in developmental and subchronic toxicity studies. The systemic toxicity of longer-chain glycol ethers and alkylphenyl glycol ethers is uncertain because data are not available. The alkyl chain length of C_7 or less was chosen as a boundary for short-chain ethylene glycol ethers based on the available data.

General Testing Strategy

The New Chemicals Program considers the following tests to be the most appropriate for ethylene glycol ethers with sufficient exposure to potentially pose an unreasonable risk:

Tier 1 - Combined Repeated Dose Toxicity Study with the Reproduction/Developmental Toxicity Screen Test (OECD Guideline 422). If signs of hematuria are seen red and white blood cell counts should be taken 2 days later except for female animals during pregnancy and lactation.

Tier 2 - The need for further testing would be determined by the results of Tier 1. This could include any of the following tests.

Prenatal Developmental Toxicity via the most appropriate route (40 CFR 799.9370)

2-Generation Reproduction Study via the most appropriate route (40 CFR 799.9380)

90-Day Subchronic Study via the most appropriate route (40 CFR 799.9346 - inhalation; 870.3250 - dermal; 870.3150 - oral)

Immunotoxicity Study via the most appropriate route (OPPTS 870.7800)

June 1992, revised December 1997

Security and Prosperity Partnership (SPP) & ChAMP

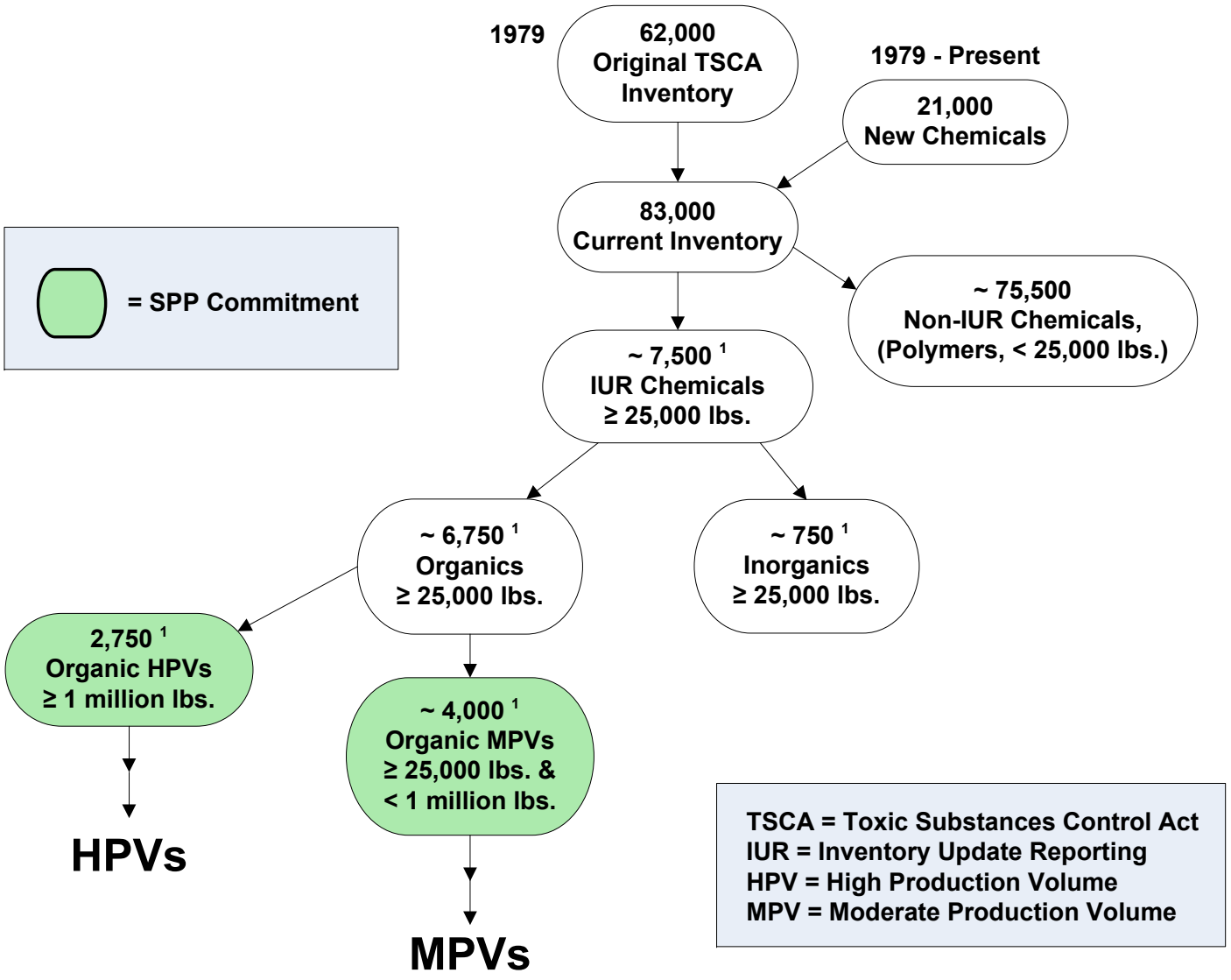
- North American Chemical Cooperation (Announced August 21, 2007)
  - Goal: enhance regulatory cooperation among Canada, Mexico, U.S.

- U.S. Commitments Under SPP
  - By 2012, assess and initiate needed action on the over 6,750* existing chemicals produced above 25,000 lbs/yr in the U.S.

- Chemical Assessment and Management Program (ChAMP)
  - Created to implement commitments the United States made under the SPP
  - Includes assessment of HPVs and MPVs and other related actions:
    - Risk-based Prioritizations for HPVs
    - Hazard-based Prioritizations for MPVs
    - Inorganic HPV Chemicals
    - Inventory Reset

* Based on preliminary statistics from 2006 IUR data

U.S. SPP Commitments = ChAMP



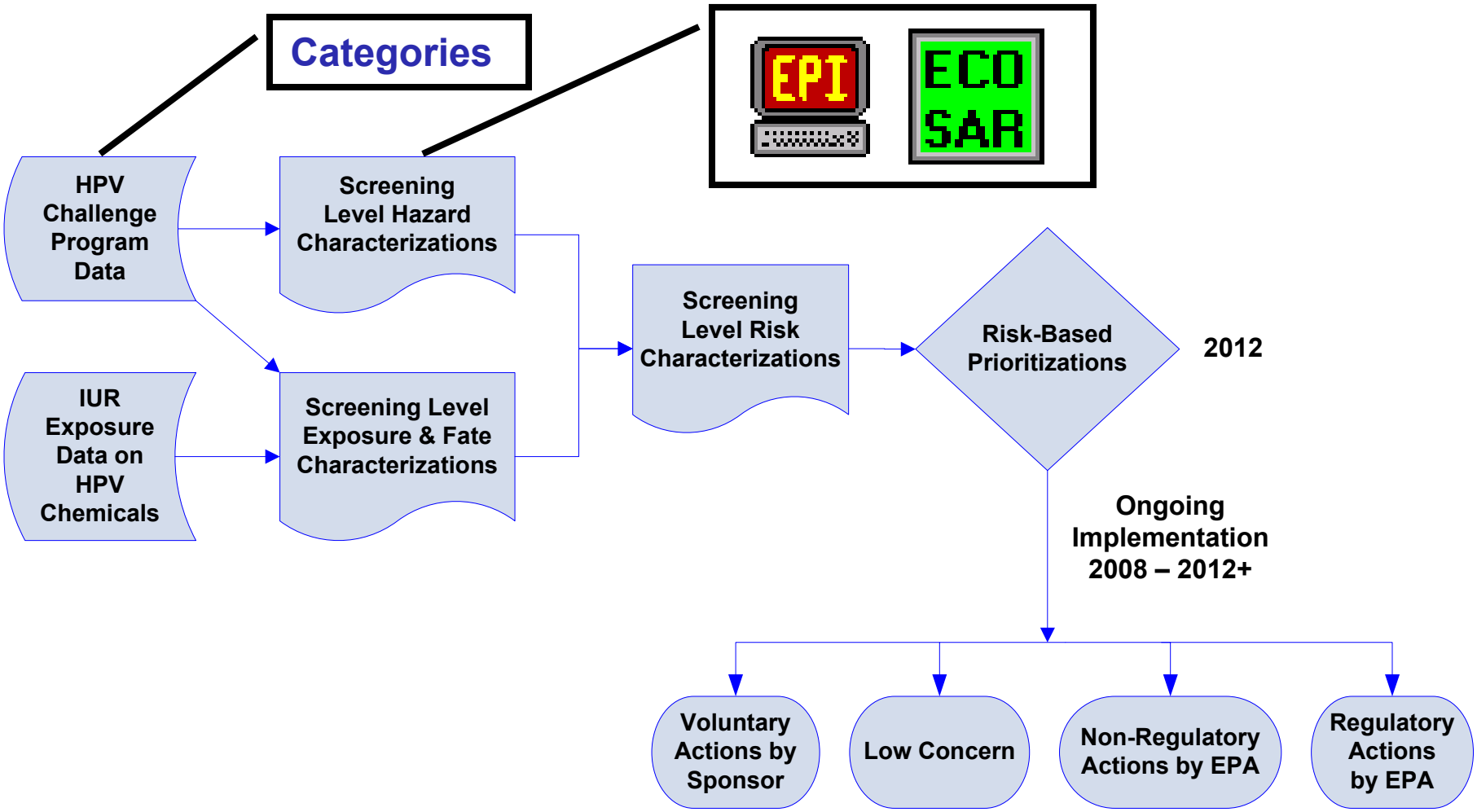
[1] Statistics are based upon preliminary 2006 IUR data; the actual numbers may change slightly when official statistics are available.

Note: The 2006 IUR introduces new reporting thresholds.

HPV Existing Chemicals - Challenge Program

- 1998 – High Production Volume (HPV) “Challenge”
  - a voluntary initiative aimed at developing and making publicly available screening-level health and environmental effects information on HPV chemicals
  - ~ 2,200 HPVs sponsored - ~ one-third via OECD
  - producers and importers identify and provide initial assessment of the adequacy of existing toxicity data/information and conduct new testing if data do not exist
- Screening Information Data Set or “SIDS” (18 internationally agreed endpoints):
  - 10 P-Chem Properties & Environmental Fate Parameters
  - 5 Human Health: Acute, Repeated-Dose, Repro, Develop and Genetic Toxicity
  - 3 Ecological: Acute Fish, Invertebrate, Plant Toxicity); Chronic Toxicity (case-by-case)

HPV Chemicals: Risk-Based Prioritization Process



* The first 200 Hazard Characterizations on HPV chemicals have been posted to EPA's website. The first set of Risk-Based Prioritization documents will be posted soon.

MPV Hazard Characterization Workflow

1986–2006 IUR

New Chemicals

Identify Structural Clusters

Cluster 1 …

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

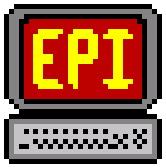
Fate // Environmental Fate – HPV, PhysProp, EPISuite

Ecotoxicity // Ecotoxicity – HPV, ECOTOX, ECOSAR

Human Health Endpt 1 // Human Health Endpt 2 // Human Health Endpt 3

// - breakpoints/sub-category boundaries

Analog Identification Methodology

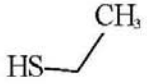
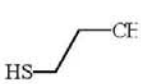


OncoLogic™

Read-Across

Table A

P-Chem and Fate Properties Table

n-Alkyl mercaptans								
82								
CAS #	75081	107039	109795	111319	111886	143102	112550	2885009
Molecular Weight	62.13	76.16	90.18	118.24	146.29	174.35	202.4	286.56
Persistence	L	L	L	L	L	L	L	L
Bioaccumulation	L	L	L	L	L	L	L	L
Melting Point	-1[illegible]	-133	-116	-81	-49	-26	-7	30
Boiling Point	35	68	98	151	199	240	274	360
Vapor Pressure	529	[illegible]	45.[illegible]	3	0.42	0.037	0.008	0.0000004
Water Solubility	15600	1900	600	70.4	5.7	0.693	0.1	0.24
Log Kow	1.18	1.81	2.28	3.23	4.21	5.2	6.18	9.12
Koc	24	44	81	275	935	3179	10820	430000
pKA	9.93	9.93	10.8	9.93	9.93	9.93	9.93	9.93
Henry's Law	0.00453	0.00408	0.0045	0.0[illegible]935	0.[illegible]139	0.0196	0.0186	0.323
Hydrolysis								
Photolysis								
Photooxidation	0.27	0.25	0.24	0.23	0.21	0.2	[illegible]	0.17
Ready Biodeg	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Rapid Biodeg	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ultimate Biodeg	weeks	weeks	days-weeks	weeks	weeks	weeks	weeks	weeks
Other Biodeg								
Other Fate								
Fugacity Air	7.49%	7.33%	6.98%	3.39%	1.45%	0.7%	0.29%	0.15%
Water	80.9%	80%	67.8%	39.5%	21.1%	12.6%	6.04%	3.75%
Soil	11.4%	12.4%	25%	56.1%	73.2%	64.8%	38.5%	28%
Sediment	0.17%	0.23%	0.3%	1.02%	4.25%	21.9%	55.2%	68.1%
BCF	2	4	11	61	35	200	361	3
BAF								

Source Quality Key

Experimental:
- ● Value from guideline study, clear weight of evidence, or evaluated data base
- ■ Value from non-guideline but valid experimental study
- ◆ Value reported without supporting experimental details

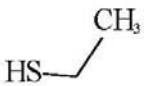
Estimated:
- ◆ Value from SAR/QSAR analysis
- ○ Value obtained using read-across

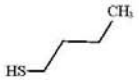
Table D

Hazard Characterization Summary Table

n-Alkyl mercaptans

82





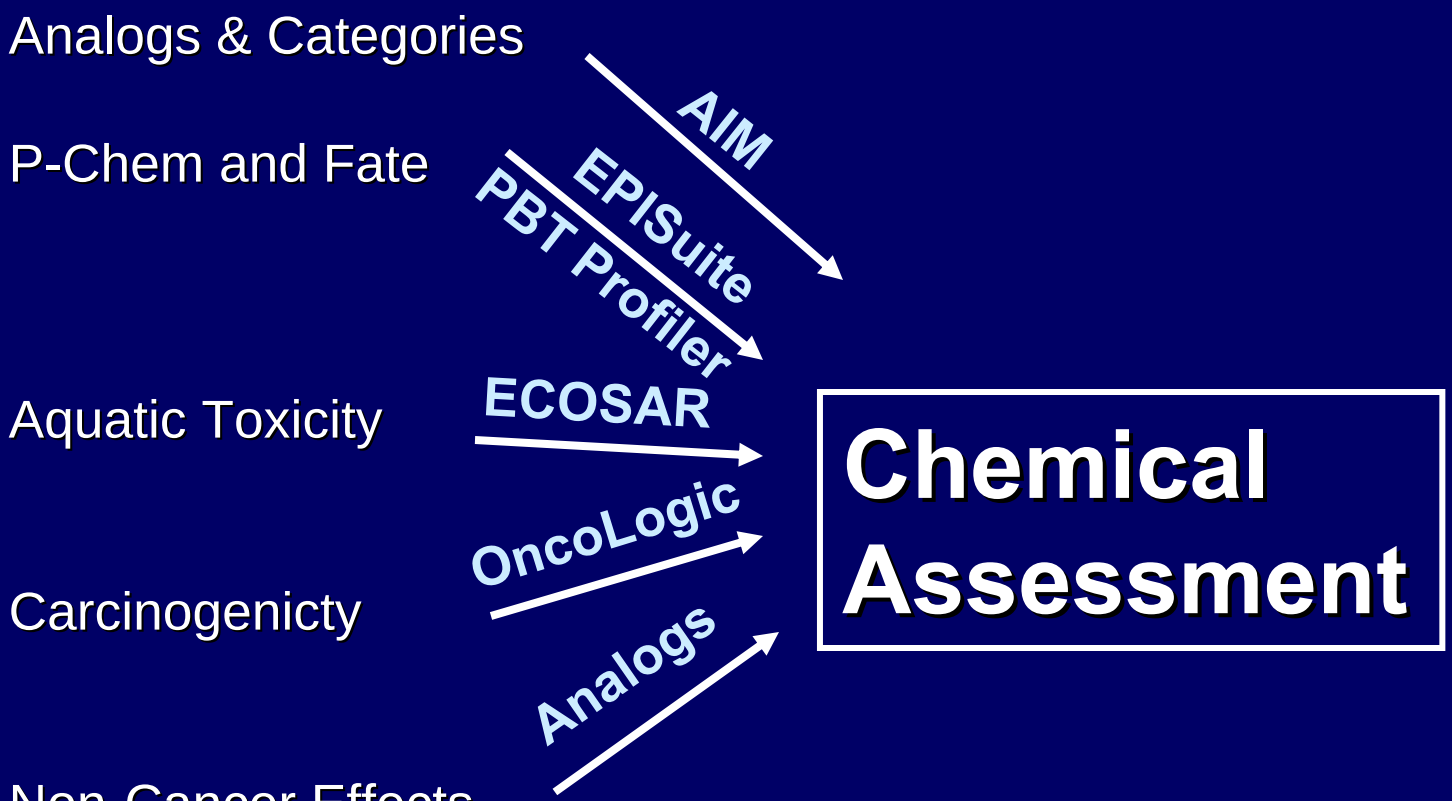
CAS #	75081	107039	109795	111319	111886	143102	112550	2885009
Molecular Weight	62.13	76.16	90.18	118.24	146.29	174.35	202.4	286.56
Environmental								
State	Liquid	Liquid	Liquid	Liquid	Liquid	Liquid	Liquid	Liquid
Persistence	L	L	L	L	L	L	L	L
Bioaccumulation	L	L	L	L	L	L	L	L
Aquatic Toxicity								
Acute	H	H	H	H	H	H	H	H
Chronic	H	H	H	H	H	H	H	H
Human Health Effects								
Acute Toxicity	M	M	M				L	
Repeated-Dose	H							
Reproductive			H					
Developmental			H				H	
Genotoxicity	±	+					–	
Cancer Hazard	–	–	–	–	–	–	–	–
Eye Irritation	–				–	+	+	
Skin Irritation	–				–	–	+	
Skin Sensitizer					+		+	

Source Quality Key

Experimental:
- Value from guideline study, clear weight of evidence, or evaluated data base
- Value from non-guideline but valid experimental study
- Value reported without supporting experimental details

Estimated:
- Value from SAR/QSAR analysis
- Value obtained using read-across

In Silico Chemical Profiling – Integral Part of EPA TSCA Chemicals Assessments



Models Publicly Available @ http://www.epa.gov/oppt/sf/

- EPISuite™
  - http://www.epa.gov/oppt/exposure/pubs/episuite.htm
- ECOSAR:
  - "Stand-alone", available for download at: www.epa.gov/oppt/newchems/21ecosar.htm
  - Integral part of EPI Suite™, available for download at: http://www.epa.gov/oppt/exposure/pubs/episuite.htm
- OncoLogic™
  - http://www.epa.gov/oppt/newchems/tools/oncologic.htm
- PBT Profiler
  - http://www.epa.gov/oppt/sf/tools/pbtprofiler.htm