
Overview of EPI Suite™: Software for Chemical Property and Fate Estimation

**Neil Patel and Bob Boethling
22 February 2006**

**U.S. Environmental Protection Agency
Office of Pollution Prevention and Toxics
(OPPT)
Washington, DC 20460**



Office of Prevention, Pesticides, and Toxic Substances

- Develops national strategies for toxic substances control and promotes pollution prevention and the public's right to know about chemical risk
- Office of Pesticide Programs (OPP)
- Office of Pollution Prevention and Toxics (OPPT)
- Office of Science Coordination and Policy (OSCP)

Office of Pollution Prevention and Toxics

- OPPT is responsible for assuring that industrial chemicals for sale and use in the U.S. do not pose unacceptable risks to human health or the environment
- Accomplished through pollution prevention, safer chemicals, risk reduction, risk management and public understanding

Office of Pollution Prevention and Toxics

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

Estimation Programs Interface (EPI Suite™)

- Estimates physical/chemical properties and environmental fate and transport
- Runs estimation programs sequentially with chemical structure as only input
- Includes PHYSPROP, a database of measured p/chem and fate properties for >40,000 chemicals
- Considered a screening-level tool; not applicable to all substances
- Intended for use only in absence of measured values

Things to Look for in an Estimation Program

- (1) How comprehensive is the methodology?**
- (2) How accurate are the methods?**
- (3) Have the methods been validated?**
- (4) Have the methods been published / peer-reviewed?**
- (5) Does the computer program that implements the methods have necessary features?**
 - (a) Easy to use?**
 - (b) Runs batches of structures?**

EPI Suite™ Chemical Structure Entry

- Runs from SMILES representation of chemical structure
- SMILES can be entered directly or using a chemical's CAS number
- Chemicals can be run batchwise
- Accepts MDL Mol files (generated by Isis Base/Draw)



Chemical Property and Fate Programs in EPI Suite™



AOPWIN	atmospheric oxidation
BCFWIN	bioconcentration factor (BCF)
BIOWIN	biodegradability
HENRYWIN	Henry's law constant
HYDROWIN	aqueous hydrolysis
KOWWIN	octanol-water partition coefficient
MPBPVP	melting point, boiling point, vapor pressure
PCKOC	soil sorption coefficient (Koc)
WSKOW	water solubility from log Kow
WATERNT	water solubility from fragments
STPWIN	removal in activated sludge treatment
LEVEL III	transport/distribution by fugacity
WVOLWIN	volatilization from water

EPI Suite™ P/Chem Property Estimation Programs

- MPBPVP
 - Estimates properties at 25 C; if run as a standalone, VP can be estimated for any temperature of interest
 - Various methods used, including Joback, Stein and Brown, Antoine, others
- KOWWIN
 - If run as a standalone, can provide an estimate using the “experimental value adjusted” (analog) method
- WSKOW
- WATERNT
 - If run as a standalone, can provide an estimate using the “experimental value adjusted” method

EPI Suite™ P/Chem Property Estimation Programs, cont.

- HENRYWIN
 - Henry's Law constant from bond/group contribution methods (improved Hine-Mookerjee)
 - If run as a standalone, can estimate H_c as a function of temperature, and using the “experimental value adjusted” method
- BCFWIN
 - Metabolism not considered
- PCKOC
 - Organic carbon partition coefficient from molecular connectivity
 - One of EPI's oldest methods

EPI Suite™ Transformation (Degradation) Programs

- AOPWIN
 - Rate constants for hydroxyl radical and ozone oxidation in the atmosphere
- BIOWIN
 - Biodegradability using six predictive models
 - Semi-quantitative results from two models, probability of fast degradation from the other four
- HYDROWIN
 - Hydrolysis rate constants for acid- and base-catalyzed reactions, but only for a small number of classes
 - Neutral hydrolysis rate not estimated

EPI Suite™ Modeling Programs

(1) Rate of volatilization from water

EPI adaptation of model from Lyman Handbook

Results: volatilization half-lives for model river & lake

(2) Sewage treatment plant model

EPI adaptation of Mackay's Toronto (STP) Model

Results: percent removal (air, biodegradation, sludge)

(3) Level III fugacity model

EPI adaptation of Mackay's EQC Model

Results: overall persistence; mass distribution in air, water, soil, sediment

EPI Suite™ v3.12 Input Screen (30 Nov 04)

EPI v3.12 [_] [] [X]

File Edit Functions BatchMode ShowStructure Output Fugacity STP Other Help

PhysProp Previous Get User Save User CAS Input **CALCULATE** ClearInputField What's New

Enter SMILES:

Chem NAME:

NameLookup

Henry LC (atm-m3/mole): Wat Sol (mg/L): MP:

Vap Pr (mm Hg): BP:

River: Lake: Log Kow :

Water Depth (meters): 1

Wind Velocity (m/sec): 5

Current Velocity(m/sec): 1

Output

Summary

Full

 The EPI (Estimation Programs Interface) Suite™ was developed by the EPA's Office of Pollution Prevention Toxics and Syracuse Research Corporation (SRC). Important information on the application of the individual models contained within the EPI Suite™ is included in the EPI Suite™ User's Guide.

© 2000 U.S. Environmental Protection Agency

EPI Suite™ v3.12 Input Screen (8 Dec 05)

EPI v3.12 [Minimize] [Maximize] [Close]

File Edit Functions BatchMode ShowStructure Output Fugacity STP Other Help

PhysProp Previous Get User Save User CAS Input **CALCULATE** ClearInputField What's New

Enter SMILES:

Chem NAME:
NameLookup

Henry LC (atm-m³/mole): Wat Sol (mg/L): MP:
Vap Pr (mm Hg): BP:

River: Lake: Log Kow :

Water Depth (meters): 1
Wind Velocity (m/sec): 5
Current Velocity(m/sec): 1

Output
 Summary
 Full



The Estimation Programs Interface (EPI) Suite™ was developed by the US Environmental Protection Agency's Office of Pollution Prevention and Toxics and Syracuse Research Corporation (SRC). It is a screening-level tool and cannot be used for all chemical substances. Like other such tools, it is intended for use in screening-level applications such as to quickly screen chemicals for release and exposure potential, and "bin" chemicals by priority for future work. Estimated values should not be used when experimental (measured) values are available.

Important information on the performance, development and application of the individual estimation programs within EPI Suite™ is included in the User's Guide.

© 2000 United States Environmental Protection Agency

EPI Suite™: Method Details, Accuracy, Validation

- Method details
 - Estimation methods for chemical properties and degradation are based on standard regression techniques
 - Most use correction factors
 - Method details are summarized online in the Help files
 - Full reference citations are also given so that users can examine methods in more detail, if they desire

PhysProp

Previous

Get User

Save User

CAS Input

CALCULATE

ClearInputField

What's

Enter SMILES:
000123-31-9Chem NAME:

NameLookup

Henry LC (atm-m3/mole): Wat Sol (mg/L): MP: Vap Pr (mm Hg): BP:

River:

Lake:

Log Kow:

Output

 Summary FullWater Depth (meters): Wind Velocity (m/sec): Current Velocity(m/sec): 

The Estimation Programs Interface (EPI) Suite™ was developed by the US Environmental Protection Agency's Office of Pollution Prevention and Toxics and Syracuse Research Corporation (SRC). It is a screening-level tool and cannot be used for all chemical substances. Like other such tools, it is intended for use in screening-level applications such as to quickly screen chemicals for release and exposure potential, and "bin" chemicals by priority for future work. Estimated values should not be used when experimental (measured) values are available.

Important information on the performance, development and application of the individual estimation programs within EPI Suite™ is included in the User's Guide.

© 2000 United States Environmental Protection Agency

EPI Suite User Guide

SMILES Help

What's New in EPI v3.12

AOPWIN User Guide

AOP Accuracy Doc

BCFWIN User Guide

BIOWIN User Guide

ECOSAR User Guide

HENRYWIN User Guide

HYDROWIN User Guide

KOWWIN User Guide

MPBPWIN User Guide

PCKOCWIN User Guide

WSKOWWIN User Guide

WATERNT User Guide

Fugacity Model Help

STP Model Help

Water Volatilization Model Help

About EPI Suite...

Example KOWWIN Output

Kowwin Results
Print Save Results Copy Remove Window Help

Log Kow(version 1.65 estimate): 0.86

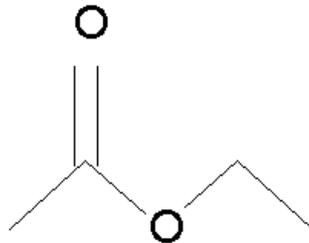
Experimental Database Structure Match:
Name : Ethyl acetate
CAS Num : 000141-78-6
Exp Log P: 0.73
Exp Ref : Hansch,C et al. (1995)

SMILES : CC(=O)OCC
CHEM : Ethyl Acetate
MOL FOR: C4 H8 O2
MOL WT : 88.11

TYPE	NUM	LOGKOW FRAGMENT DESCRIPTION	COEFF	VALUE
Frag	2	-CH3 [aliphatic carbon]	0.5473	1.0946
Frag	1	-CH2- [aliphatic carbon]	0.4911	0.4911
Frag	1	-C(=O)O [ester, aliphatic attach]	-0.9505	-0.9505
Const		Equation Constant		0.2290

Log Kow = 0.8642

Structure
File Edit Structure Help



Log Kow (estimated): 0.86
Ethyl Acetate

EPI Suite™: Method Details, Accuracy, Validation

- Method details
- Accuracy
 - EPA considers the accuracy acceptable for a screening-level tool
 - Information on method error is summarized online in the Help files
 - Full reference citations are also given so that users can examine the statistics in more detail, if they desire

KOWWIN Program - $\log K_{ow}$ ($\log P$)

Methodology - Atom/Fragment

Contributions:

170 Fragments

290 Correction Factors

Log Kow used by:

BCFWIN

DERMWIN

ECOSAR

WSKOWWIN

Statistical Accuracy:

	<u>number</u>	<u>Corr (r^2)</u>	<u>Std Dev</u>	<u>Mean Error</u>
Total	13229	0.954	0.436	0.316
Training	2467	0.981	0.219	0.162
Validation	10762	0.943	0.473	0.354

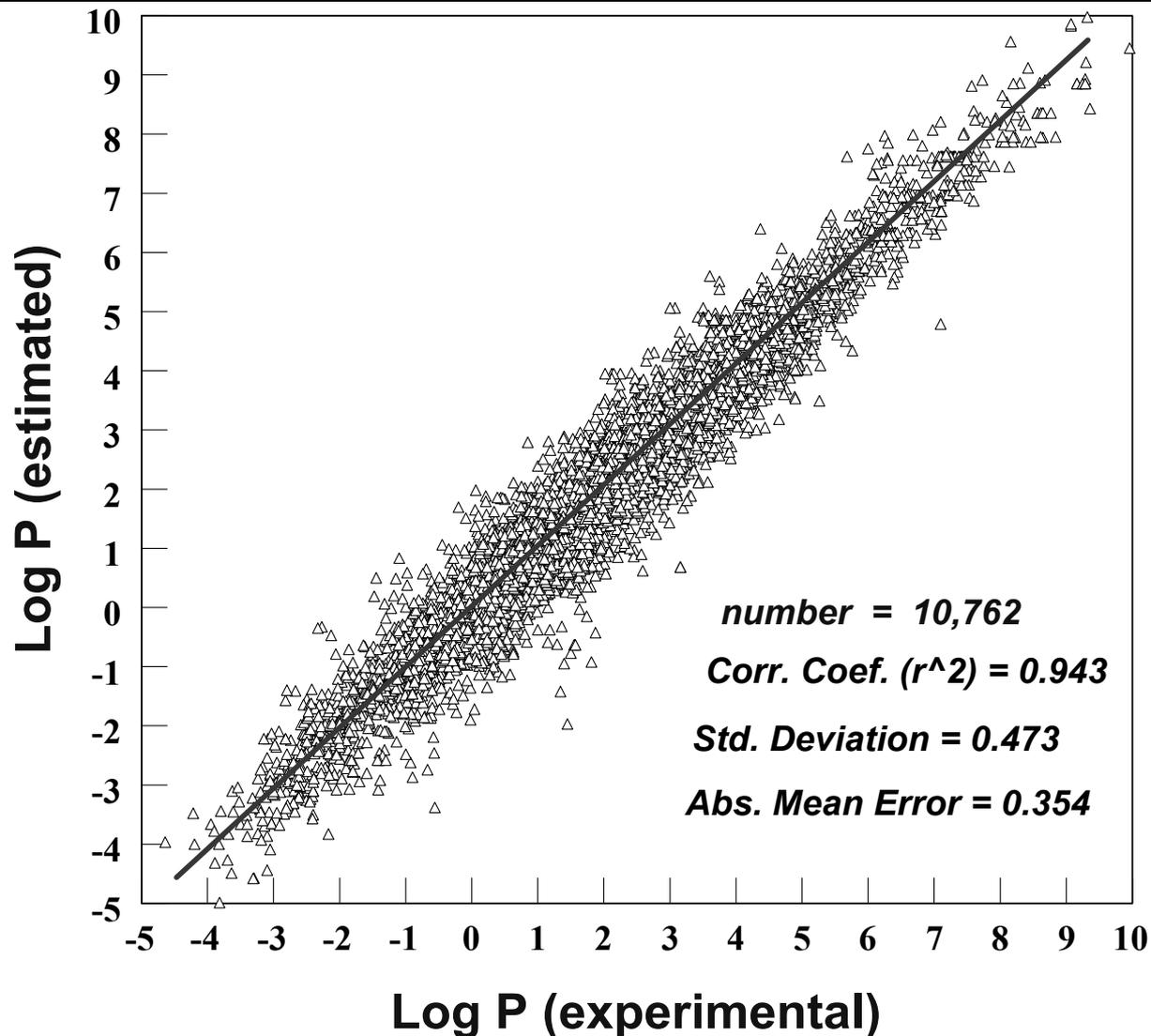
KOWWIN includes the experimental database of 13,229 recommended $\log P$ values

Journal Article Description: J. Pharm. Sci. 84(1): 83-92 (1995)

EPI Suite™: Method Details, Accuracy, Validation

- Method details
- Accuracy
- Validation
 - Most methods have been validated using independent (external) validation sets
 - Information on validation is summarized online in the Help files
 - Full reference citations are also given so that users can examine the statistics in more detail, if they desire

KOWWIN: Validation Data Set



PCKOC - Soil Sorption K_{oc}

Methodology:

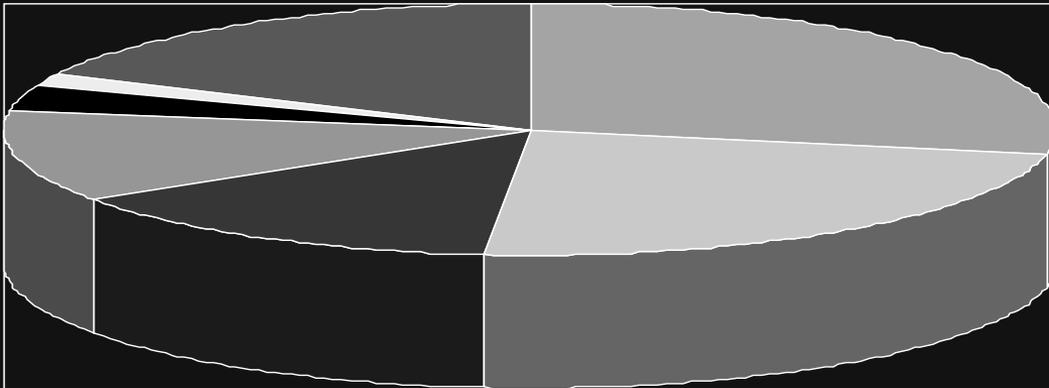
Estimates soil sorption coefficient (K_{oc}) from the first-order molecular connectivity index ($^1\chi$) and 27 polar, structural correction factors by the following equation:

$$\log K_{oc} = 0.53 \ ^1\chi + 0.62 + \sum \text{correction factors}$$

	<u>number</u>	<u>Corr (r^2)</u>	<u>Std Dev</u>	<u>Mean Error</u>
Training Data Set	189	0.955	0.230	0.182
Validation Data Set	205	0.856	0.462	0.344

Journal Article Description: Environ. Sci. Technol. 26: 1560-1567 (1992)

Who is Using OPPT Exposure Models ?



Industry 28%

**University/Research
23%**

Consultants 14%

**Fed. Government
12%**

State/Local Gov. 4%

Non-OPPT EPA 1%

Other 18 %

Note: based on 2003 data

Use of EPI Suite™ in OPPT Programs

- Understanding of basic physical/chemical properties and environmental fate and transport is fundamental to assessing exposure and hazard, and ultimately risk
- EPI Suite™ developed by EPA and Syracuse Research Corporation for use in EPA's New Chemicals Program
- Also widely used in other EPA programs
- **Reliable test data are always used in preference to estimates**

EPI Suite™ and the NCP

- Under EPA's New Chemicals Program (NCP), a Premanufacture Notification (PMN) must be filed in advance of manufacture or import
- After first steps in the PMN process, e.g. identification of the chem structure, EPI Suite is run to provide estimated properties. This must be done quickly!
- EPI is a screening-level tool and is not applicable to, e.g., polymers. Consistent with this, it is run only for discrete organics with MWt < approx. 500
- EPI-generated data are only part of the picture. Submitted (measured) data, literature data, expert judgment all contribute

EPI Suite™ and Data-Intensive Programs

- For HPVCs, estimated values acceptable for $\log K_{OW}$
- For fate endpoints other than biodegradation, estimates acceptable for
 - Photodegradation (atmospheric oxidation)
 - Transport/Distribution
 - Stability in water
- Even for BP, VP and water solubility, estimates may be acceptable under certain conditions

EPI Suite™ and the HPVC Program

- Solve this equation:
User-friendly estimation software
+ HPVC policy allowing some estimates
= ???
- One answer: submit EPI data to EPA instead of collecting measured values, even when no new testing is required
- What to do?
 - Option #1: Make EPI Suite™ harder to use
 - Option #2: Add info on proper use, data quality, etc.
Done--see next 2 slides
 - Option #3: Any ideas from SAB??

EPI Suite™ v3.12 Input Screen (8 Dec 05)

EPI v3.12



File Edit Functions BatchMode ShowStructure Output Fugacity STP Other Help

PhysProp Previous Get User Save User CAS Input CALCULATE ClearInputField What's New

Enter SMILES:

Chem NAME:

NameLookup

Henry LC (atm-m³/mole):

Wat Sol (mg/L):

MP:

Vap Pr (mm Hg):

BP:

River:

Lake:

Log Kow :

Output

Summary

Full

Water Depth (meters):

1

1

Wind Velocity (m/sec):

5

0.5

Current Velocity(m/sec):

1

0.05



The Estimation Programs Interface (EPI) Suite™ was developed by the US Environmental Protection Agency's Office of Pollution Prevention and Toxics and Syracuse Research Corporation (SRC). It is a screening-level tool and cannot be used for all chemical substances. Like other such tools, it is intended for use in screening-level applications such as to quickly screen chemicals for release and exposure potential, and "bin" chemicals by priority for future work. Estimated values should not be used when experimental (measured) values are available.

Important information on the performance, development and application of the individual estimation programs within EPI Suite™ is included in the User's Guide.

© 2000 United States Environmental Protection Agency

1.4. Limitations

It is important to stress that EPI Suite™ is a screening-level predictive tool and cannot be used for all chemical substances. Like other such tools, it is intended for use in screening-level applications such as to quickly screen chemicals for release and exposure potential, and "bin" chemicals by priority for future work. The estimation methods in EPI Suite™ have been developed by government, academic, and private sector researchers over many years and represent some of the best techniques currently available. Nevertheless, EPI Suite™ is a tool that, like all tools, has strengths, weaknesses, and limitations. These limitations should be considered before using EPI Suite™. For example, predicted data should not be used in place of experimental data. Additional model limitations are described in the Users Guide for each individual program, and in the original publications referenced therein.

1.5. Data Quality Considerations

The User Guides for the individual estimation programs contain detailed information on the estimation methods, including sources of experimental data, predictive algorithms and method error (accuracy). The User Guides also reference publications in peer-reviewed journals in which further details are given. Most EPI Suite™ methods have been published in peer-reviewed journals. In general, measured values used to develop models were selected based on a multi-step review by senior scientists at Syracuse Research Corporation. In some cases, such as for the KOWWIN and WSKOW programs, most data came from highly regarded sources (e.g. Hansch et al. 1995) for which the data had already been carefully evaluated using explicit data quality criteria. For the KOWWIN program full reference citations are also given for all of the training set data. This allows users to check measured values themselves.

Measured values in the PHYSPROP file (accessed within EPI Suite™) are periodically uploaded from the PHYSPROP database maintained at Syracuse Research Corporation. This latter file has been actively built by SRC over the last two decades. It started as a database of physical properties for chemicals being evaluated by SRC for the Hazardous Substances Data Bank (HSDB), available from the National Library of Medicine (NLM)(<http://toxnet.nlm.nih.gov/>). Initially data were entered by junior and senior scientists using many sources for which the data had already been carefully evaluated (see Boethling RS, PH Howard and W Meylan. 2004. Finding and estimating chemical property data for environmental assessment. Environ. Toxicol. Chem. 23: 2290-3308). Data were then checked by senior scientists as described for the CHEMFATE file of the Environmental Fate Data Base (EFDB). For all records the QC process includes evaluation of the record to determine if the value makes sense scientifically (correct units, appropriate value given the chemical structure, etc). In addition, for approx. 10% of the records, a senior scientist checks the original source of the data. Additional quality control is performed by comparing measured values to estimated values from structure/property relationships (as a possible means of identifying outlying observations); and/or by comparing the values for one property (e.g. Henry's Law constant) to estimates

PhysProp

Enter SMILES:

Chem NAME:

NameLookup

Henry LC (atm-m3/

Water Depth (met

Wind Velocity (m/s

Current Velocity(m/



EPI Updates for v3.12

(30 Nov 04)

- Chemical name lookup function added
 - Type name: program gets SMILES and enters it on main screen
 - No need to know either SMILES or CAS, if the chemical name is on list
- STPWIN
 - New option: estimate activated sludge biodeg half-lives from chemical structure via BIOWIN
 - Default mode is still no biodeg in STP

EPI Updates for v3.12

30 Nov 04 enhancements, cont.

- BLOWIN
 - Ready biodegradability now estimated using a Bayesian model battery (BLOWIN3 and 5)
- WATERNT (formerly WATERFRAG)
 - Help information now available
- Level III EPI (multimedia model)
 - Default ratios of half-lives for water:soil:sediment changed from 1:1:4 to 1:2:9
 - Corrects inconsistency with PBT Profiler™

Chemical Name Lookup X

Enter a chemical name ... the Name database will be searched for possible matches with corresponding SMILES notations.

Chemical Name: hydroquinone

OK

Cancel



The Estimation Programs Interface (EPI) Suite™ was developed by the US Environmental Protection Agency's Office of Pollution Prevention and Toxics and Syracuse Research Corporation (SRC). It is a screening-level tool and cannot be used for all chemical substances. Like other such tools, it is intended for use in screening-level applications such as to quickly screen chemicals for release and exposure potential, and "bin" chemicals by priority for future work. Estimated values should not be used when experimental (measured) values are available.

Important information on the performance, development and application of the individual estimation programs within EPI Suite™ is included in the User's Guide.

© 2000 United States Environmental Protection Agency

Names Found - Highlight and Select: [X]

- HYDROQUINONE /000123-31-9
- HYDROPEROXIDE, HEPTYL /000764-81-8
- HYDROPEROXIDE, HEXYL /004312-76-9
- HYDROPEROXIDE, PENTYL /000074-80-6
- CYCLOHEXANOL, 1- (1-HYDROPEROXYCYCLOHEXYL)DIOXY - /000078-18-2
- HYDROPEROXYL /003170-83-0
- HYDROPRENE /041096-46-2
- HYDROQUINIDINE /001435-55-8
- HYDROQUININE /000522-66-7
- 2-HYDROQUINOLINE /070254-42-1
- HYDROQUINONE /000123-31-9
- HYDROQUINONE26DIMETHOXY /000002-79-0
- HYDROQUINONE2METHYL6BROMO /000002-88-6
- HYDROQUINONE, DIACETATE /001205-91-0
- HYDROQUINONE DIETHYL ETHER /000122-95-2
- HYDROQUINONE DIGLYCIDYL ETHER /002425-01-6
- HYDROTHIAZIDE /023141-82-4
- BORATE(1-), HYDROTRIS(1-METHYLPROPYL)-, LITHIUM, /038721-52-7

OK

Cancel

© 2000 United States Environmental Protection Agency

Enter SMILES: Oc(ccc(O)c1)c

0001: PhysProp Data

Print Copy TransferDataToEpiScreen

Chem Name Experimental Data from PhysProp Database:

CAS Number: 000123-31-9
 Chem Name: HYDROQUINONE
 MP (deg C): 172.3
 BP (deg C): 287
 Log Kow: 0.59
 Kow ref: HANSCH,C ET AL. (1995)
 Water Sol: 7.2E+004 mg/L
 WS temp: 25 deg C
 WS ref: GRANGER,FS & NELSON,JM (1921)
 Vapor Pr: 2.4E-005 mm Hg (extrapolated)
 UP temp: 25 deg C
 UP ref: DAUBERT,TE & DANNER,RP (1991)
 Henry LC: ---
 pKa: 10.85
 pKa temp: 25 deg C
 pKa ref: PEARCE,PJ & SIMKINS,RJJ (1968)
 OH Rate: ---

File Edit Functions BatchMode ShowStructure Output Fugacity STP Other Help
 PhysProp Previous Get User Save User CAS Input CALCULATE ClearInputField What's New

Enter SMILES:
 000123-31-9

Chem NAME:
 NameLookup

Henry LC (atm-m3/mole):
 Wat Sol (mg/L): MP:
 Vap Pr (mm Hg): BP:

River: Lake: Log Kow:
 Water Depth (meters):
 Wind Velocity (m/sec):
 Current Velocity(m/sec):
 Output
 Summary
 Full



The Estimation Programs Interface (EPI) Suite™ was developed by the US Environmental Protection Agency's Office of Pollution Prevention and Toxics and Syracuse Research Corporation (SRC). It is a screening-level tool and cannot be used for all chemical substances. Like other such tools, it is intended for use in screening-level applications such as to quickly screen chemicals for release and exposure potential, and "bin" chemicals by priority for future work. Estimated values should not be used when experimental (measured) values are available.

Important information on the performance, development and application of the individual estimation programs within EPI Suite™ is included in the User's Guide.

© 2000 United States Environmental Protection Agency

EPI Suite™: 2005/6 enhancements

(italics-status as of 2/22/06)

- Subcooled liquid vapor pressure-**done**
- Dimensionless Henry constant (K_{AW})-**done**
- Octanol/air partition coefficient (K_{OA})-**done**
- Fraction of airborne substance sorbed to particulates (ϕ)-**done**
- Bioaccumulation in aquatic food webs (BAF)-**done**
- Anaerobic biodegradation potential
- Hydrolysis
 - All-new version of HYDROWIN