FATE, THE ENVIRONMENTAL FATE CONSTANTS INFORMATION SYSTEM DATABASE

by

Heinz P. Kollig[1]
Karen J. Hamrick[2]
Brenda E. Kitchens[1]

[1] Environmental Research Laboratory
U.S. Environmental Protection Agency
Athens, GA 30613-0801

and

[2] Computer Sciences Corporation
c/o Environmental Research Laboratory
U.S. Environmental Protection Agency
Athens, GA 30613-0801

DISCLAIMER

The information in this document has been funded wholly or in part by the
United States Environmental Protection Agency. It has been subject to the
Agency's peer and administrative review, and it has been approved for
publication as an EPA document. Mention of trade names or commercial products
does not constitute endorsement or recommendation for use by the U.S.
Environmental Protection Agency.

FOREWORD

As environmental controls become more costly to implement and the penalties of judgement errors become more severe, environmental quality management requires more efficient analytical tools based on greater knowledge of the environmental phenomena to be managed. As part of this Laboratory's research on the occurrence, movement, transformation, impact, and control of environmental contaminants, the Measurements Branch provides physical, chemical, and microbial rate and equilibrium constants for use in mathematical models of pollutant behavior.

Assessment of potential risk posed to humans by man-made chemicals in the environment requires the prediction of environmental concentrations of those chemicals under various scenarios. Models and other risk assessment techniques frequently require the use of physical and chemical process data to estimate the transport and transformation of specific chemicals. To meet this data need, an online database, called FATE, has been developed to provide the user with reliable and environmentally realistic fate constants.

Rosemarie C. Russo, Ph.D.
Director
Environmental Research Laboratory
Athens, Georgia

iii

ABSTRACT

A new online database, designated the FATE database, has been developed for the interactive retrieval of kinetic and equilibrium constants that are needed for assessing the fate of chemicals in the environment. The database contains values for twelve parameters, but may not contain a value for each parameter for each chemical. As of July 1991, the database contained values for about 200 chemicals. Unique features of the database include experimental data that are extracted only from primary references, and pertinent experimental conditions that are entered into the database to assure the user of the credibility and applicability of a value. A newly developed computer program is used to extrapolate hydrolysis rate constants to a standard format. Acidic, basic and neutral contributions are combined to calculate the overall hydrolysis rate constant, k_h, and the half-life of the chemical at 25°C and pH 7. The data are also reported as second-order acidic and basic rates and a first-order neutral rate at 25°C. Products of transformation are listed for degradation processes when available. A newly developed computerized expert system will be applied to compute accurate fate constant values. The expert system has the capability of crossing chemical boundaries to cover all organic compounds.

TABLE OF CONTENTS

INTRODUCTION

Assessment of potential risk to human health and to the environment posed by chemicals requires the prediction of environmental concentrations of those chemicals under different scenarios. Risk assessment frequently requires the use of physical and chemical process data to estimate the transport and transformation of specific chemicals in the environment. This information is subsequently used for regulating the allowable concentrations of specific chemicals in ground water wells, air emissions, hazardous waste sites, etc.

The need for specific rate and equilibrium constants for chemicals that have potential environmental impact has grown in tandem with the production of new chemicals by the chemical industry. Approximately 1600 new chemicals or formulations are submitted to EPA's Office of Toxic Substances (OTS) each year for premanufacturing review, and approximately 75,000 chemicals are potential candidates for review under the OTS existing chemicals program. Relatively few of the required fate constants have been measured experimentally and published, and many of the published fate constants are of questionable reliability or applicability. Data being used for environmental and human risk assessment for regulatory purposes must be of known reliability for the assessment to have validity.

As laboratory instrumentation has improved and experimental procedures have become more sophisticated, the environmental fate constants that have been measured in the laboratory have become more reliable. However, no matter what the objectives of the research, the experimental protocol, the sampling procedures, and the chemical and statistical analyses of the data remain critical aspects of the research and affect the reliability of fate constants[1]. Therefore, literature values of fate constants vary, sometimes

1

considerably[2,3], and often several values are reported for the same chemical parameter by different laboratories. Data evaluation systems have been suggested[3-5] as a mechanism to assist the user in deciding which one of many values might be the most reliable. However, the investigator who measured the fate constant is the only one who can actually ensure the accuracy of the data. The reliability of the fate constant can be evaluated by others only by examining the research protocol and experimental conditions provided by the author.

The inherent complexity in measuring physical and chemical properties, especially those of hydrophobic chemicals, makes the measurement process a difficult and, therefore, expensive one. Even if the prohibitively high costs of the measurement processes could be ignored, the need for data will never be completely satisfied through laboratory studies because of the time involved. For these reasons, reliable computational techniques are needed to estimate physical and chemical properties. Computational techniques will generate values more rapidly at a fraction of measurement costs and will eventually satisfy the need for much of the required data.

In this project, we developed an online database that provides the user with reliable and environmentally realistic fate constants. We ensure the quality of the data by: a) applying objective screening criteria to determine whether a value should be entered into the database, and giving the user pertinent experimental conditions, b) entering into the database literature values from primary sources only, and c) entering into the database computed values based on estimation techniques that use both fundamental chemical structure theory and conventional techniques based on property-reactivity correlations that are carefully screened for applicability before use.

2

Potential users should write to Heinz P. Kollig, U.S. EPA, College Station Road, Athens, GA, requesting a user's application form.

PROBLEMS WITH CURRENTLY AVAILABLE DATA

Experimental Data

Currently, computerized databases are being constructed to provide experimental fate constant data that are more conveniently accessible to the scientific community. As these databases have been compiled or used, it has become obvious that data cannot simply be taken from the scientific literature and used without extensive knowledge of the process under investigation and the conditions under which the investigations were conducted[6,7].
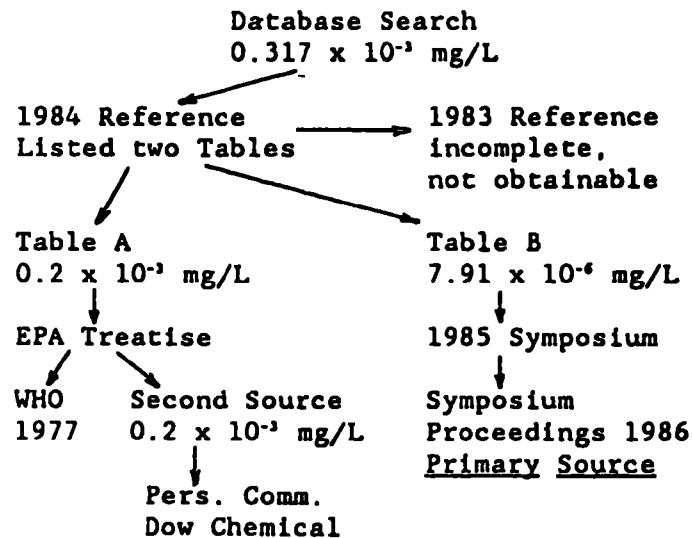
Many early experiments were conducted using investigative criteria that were different from the stringent criteria needed for assessments today. For example, some hydrolysis rate studies were conducted under uncontrolled or undocumented hydrolytic conditions, using phrases such as 'distilled water' or 'room temperature', or conditions that could not be extrapolated to environmental situations, such as with co-solvents. This criteria problem results partially from the fact that most data in the literature were obtained without standard protocols and in mechanistic research where absolute and precise fate constant values were not the major objectives of the research. The result has been that much published data cannot be used to generate valid environmental assessments without considerable mathematical manipulation, if they can be used at all. Often, when the toxic chemical in question has been in use for decades, reported aqueous solubility values, partition coefficients, and other parameters will range over several orders of magnitude[2,8].

3

In any reporting situation, there is always the chance that data were miscalculated, transposed, published with less than (<), greater than (>), positive (+) or negative (-) signs missing, decimal point placed incorrectly, or any number of other problems resulting from repeated citation of other-than-primary sources. For example, we needed the octanol/water partition coefficient, K_{ow}, for acenaphthene, CAS number [83-32-9]. A quick search in a secondary publication revealed a value of -2.02 for the log K_{ow}, and included the primary reference. Because of our knowledge of the K_{ow} values for structurally similar compounds, we viewed the value of -2.02 as questionable. When we obtained the primary source, it listed a value of 3.98 for the log K_{ow}. Clearly, without knowledge of the process, one might use a frequently reported and cited value that is off by six orders of magnitude, demonstrating the importance of obtaining values from primary sources.

Obtaining a primary publication can be a frustrating process, however. The frustration involved in locating the primary reference is illustrated by the following example. A value was needed for the water solubility of 2,3,7,8-tetrachloro-dibenzo-p-dioxin (TCDD). A fate constant database showed a value of 0.317×10^{-3} mg/L and gave a 1984 reference notation with the comment that the authors obtained the value from a second author who published in 1983. No journal was referenced for the second (1983) author. A search with CAS ONLINE for the second author revealed no listing under the author's name published in 1983. The referenced 1984 publication, in this instance, showed two tables of properties for TCDD. Neither table gave the value that appeared in the fate constant database (0.317×10^{-3} mg/L) for the water solubility. The first table, Table A, gave a value of 0.2×10^{-3} mg/L and referenced an EPA treatise; the treatise, in turn, referenced the World Health

Organization 1977 and a second source as the suppliers of all data in Table A. The treatise, however, did not differentiate between the two sources for individual values. The second source identified the value found in Table A (0.2×10^{-3} mg/L) but referenced personal communication with workers at Dow Chemical Company. We did not contact the Dow Chemical Company. Table B gave a value of 7.91×10^{-4} mg/L for the water solubility of TCDD and referenced a symposium held in Germany in 1985. Proceedings of the symposium were published in 1986 and this reference revealed that the water solubility had actually been determined experimentally. Figure 1 shows a flow diagram of this search to provide a graphic understanding of the effort involved. The paradox concerning this search was that the paper containing the information extracted for Table B was published in 1984, but referenced a paper that was presented in 1985 and published in 1986.

Figure 1. Flow of search for the primary source of TCDD

Database Search
0.317×10^{-3} mg/L

1984 Reference ———→ 1983 Reference
Listed two Tables   incomplete,
                      not obtainable

Table A   Table B
0.2×10^{-3} mg/L 7.91×10^{-4} mg/L

EPA Treatise   1985 Symposium

WHO   Second Source   Symposium
1977 0.2×10^{-3} mg/L   Proceedings 1986
                       Primary Source

Pers. Comm.
Dow Chemical

Additionally, an inconsistent use of exponents and significant figures gives the impression that the reported water solubility values, 0.2×10^{-3} and

7.91×10^{-6}, are about three orders of magnitude apart. The difference, of course, is only 25 times; nevertheless, the difference is significant. The important question becomes, which value is correct? A third reference corroborated the 7.91×10^{-6} mg/L value from Table B. Because both values were verifiable experimentally, the 7.91×10^{-6} mg/L value was considered more reliable. It is noteworthy to mention that corroborative information often is not available and that adequate experimental, analytical or statistical information is sometimes not provided by authors. Whatever the reason, if information is missing, the value's assessed credibility decreases.

What can be said about the quality, accuracy, or applicability of data? Consider the dilemma of an exposure or risk assessment modeler who is confronted with the information provided in Table 1. Because these reported values span several orders of magnitude, the range for the octanol/water partition coefficient is 8.2×10^{6} for p,p'-DDT and 1.6×10^{6} for dieldrin and each set of values has a standard deviation larger than the mean. A modeler would prefer to be given one "best estimate" fate constant value for each chemical because he may not have the resources available for determining which value would be the best to use. Unfortunately, the accuracy of a reported value cannot be fully evaluated. Accuracy depends on the skill and expertise of the researcher, on the maintenance condition and precision of the equipment, and on the repeatability of the experiment. The thoroughness of the documentation of a report can be evaluated, however, and thus suggest the confidence with which a reported value can be used[6]. Applicability can only be determined by knowing the conditions under which the experiment was conducted. If the value is to be used in an environmental risk assessment, then the value must have been determined in a manner that permits

extrapolation to environmentally realistic conditions for the assessment to have validity.

Table 1. Experimental octanol/water partition
coefficients for p,p'-DDT and dieldrin
(literature survey)

Name	K_{ow}
p,p'-DDT	9.5×10^3
	1.2×10^5
	8.2×10^6
Dieldrin	1.2×10^4
	2.5×10^3
	1.6×10^6

Estimated Data

Because of the high cost of laboratory measurement (estimated to be more than $10,000 per parameter), there has been a recent trend toward estimating fate constants[9]. In addition, the inherent complexity in measuring physical and chemical properties, especially of hydrophobic chemicals, makes the measurement process difficult. Even if the prohibitively high cost of the measurement process could be ignored, there will always be a shortage of experimental data because of time constraints. For these reasons, computational techniques to estimate physical and chemical properties have been developed. Computational property estimation techniques can generate values at a small fraction of measurement costs, and it is likely that much of the published data of the future will come from the application of these methods.

The largest compilation of property estimation methods was made by Lyman et al.[10] Most of Lyman's methods are based on property-reactivity

7

correlations and allow estimation of a number of constants, but relationships often hold only within limited families of chemicals. If the estimation is done for a chemical within a family for which a relationship was established, the value can be very reliable. For a non-established relationship, however, the estimated value can be off by one or more orders of magnitude. As stated earlier, so can experimental values. Thus, a user evaluating data from application of predictive techniques must fully appreciate the range and limitations of the techniques themselves no matter how easy it is to generate such estimates. It is noted that some of these estimation methods have been fully automated[11,12] and only require the input of a CAS number for most chemicals.

A promising new computational method for predicting chemical reactivity is the computer expert system SPARC Performs Automated Reasoning in Chemistry (SPARC)[13], being developed by scientists at ERL-Athens and the University of Georgia. This system uses algorithms based on fundamental chemical structure theory to estimate parameters and uses an approach that combines principles of quantitative structure-activity relationships, linear free energy theory, and perturbation theory from quantum chemistry. The goal for SPARC is to compute a value that is as accurate as a value obtained experimentally for a fraction of the cost required to measure it. Once established, the expert system should be able to estimate environmental fate constants with remarkable accuracy because the computation will be based on molecular theory with an increasing database to "train" the system and refine its algorithms. This contrasts with conventional estimation techniques that are based on correlations or other relationships that have been shown to incorporate inherent errors. Reliable experimental data with good documentation are still

necessary, however, for further testing, training, and validation of SPARC.

THE DATABASE MANAGEMENT SYSTEM

A prototype of FATE was developed with d-BASE III Plus using a relational file structure as a preliminary step in the analysis of system size and database design requirements. The prototype contained the CAS number and chemical name information of 2000 chemicals and a complete reference database of 320 references. The objective of the analysis was to determine an efficient file structure and programming language that could be used on the VAX to develop the programs and menus necessary to manipulate the chemical database files.

An evaluation was performed to determine the requirements of FATE by defining the relationships of the data elements to one another, the required field types and dimensions, the key fields that were necessary for indexing, and the scope of the reporting and maintenance requirements.

A design specification for FATE was prepared by the Computer Sciences Corporation, under EPA contract, without regard to specific hardware and software implementation platforms. They determined that the design specifications for data storage, interactive data entry, and data retrieval were of a standard nature and that the database could be implemented with any one of many fourth-generation database management systems. The recommendation was that a fourth-generation language would be preferable to a third-generation language for database development because of advantages in development time, ease of enhancement, and ease of maintenance. It was further recommended that FOCUS be used as the software to develop the chemical database because FOCUS software was already installed on the Athens VAX and

9

because FOCUS was currently the EPA standard database management system.

FOCUS is a fourth-generation language composed of several languages and utilities that are used for specific operations on the data files. FOCUS allows the use of scientific notation and the development of relational or hierarchical files, and contains keyed indexes with pointers for the linkage of the files.

The FOCUS-based FATE database is installed on the ERL-Athens VAX and operates within the VMS (Virtual Memory System) operating system.

FATE DATABASE

Currently, FATE contains published fate data for approximately 200 chemicals obtained only from primary references plus selected fate constants derived from computational techniques applied at ERL-Athens. It is hoped that reliable data computed with SPARC will soon close the void of missing fate data that exists today, at a fraction of the cost and time of obtaining the values experimentally.

Database Files

The FATE database system consists of three data files. The CAS file contains CAS numbers (Chemical Abstract Service), molecular formulae, SMILES notations (Simplified Molecular Identification and Line Entry System)[14,15], and chemical and common names. The REF file contains reference numbers and complete citations. The FATES file is cross-referenced to the other files and contains the data for the fate parameters.

Five fields in the CAS file are indexed: CASNUMBER, CASFORMULA, CASSMILES, CASNAME and CASCOMMON. The database can be searched for an entry

in any one of the indexed fields.

Two fields in the REF file are indexed: REFNUMBER and REFAUTHOR1. The database can be searched for an entry in either one of the fields, or for one of the secondary authors.

The FATES file is cross-referenced to the CAS and REF files, and can, therefore, be searched for an entry by all of the indexed fields.

Fate Constants

The database contains fields for entry of the following twelve fate constants:

fate code	symbol	fate parameter
04	H_o	Henry's law constant
05	k_h	hydrolysis rate constant
06	pK_a	ionization constant
07	K_{ow}	octanol/water partition coefficient
08	K_{oc}	organic carbon normalized sediment/water partition coefficient
10	k_d	direct (aqueous) photolysis rate constant
11	K_p	sediment/water partition coefficient
13	P_v	vapor pressure
14	S_w	water solubility
21	ϵ_λ	molar absorptivity
26	ϕ_r	aqueous photolysis reaction quantum yield
27	k_{bio}	biodegradation rate constant

Sources of Data

The open literature is the source of primary references for the experimental data included in the FATE database.

Fate constant data for some of the twelve processes are estimated by our staff with computational techniques, using the SMILES notation to define the molecular structure of a chemical. We use the QSAR[12] (Quantitative Structure-Activity Relationships) System and SPARC[13] (SPARC Performs Automated Reasoning

in Chemistry) for estimating data.

The QSAR system contains estimation routines that have been modified from the routines written by Lyman et al.[10] It can be searched by CAS number or SMILES notation and provides the estimated values in a table format.

The expert system SPARC uses computational algorithms based on fundamental chemical structure theory and allows estimation of values for any parameter that depends upon molecular structure. Unlike methods based on property-reactivity correlations, this capability crosses chemical family boundaries to cover all organic compounds. SPARC eventually will contain estimation routines for most, if not all, of the twelve parameters that are included in the FATE database.

Rate Program

Measured hydrolysis data are analyzed initially with RATE, a FORTRAN program that was developed at ERL-Athens. RATE is used to extrapolate data to a standard format.

The RATE program requires the entry of several first-order rate constants over a range of pH values. Individual data points consist of three parameters, the first-order rate constant, the temperature, and the pH at which the rate was measured. The program uses the Arrhenius equation, a standard temperature of 25°C and an assumed energy of activation of 20,000 cal/mol to transform the data and produce a plot of pH versus log k. The plot of the data is used to determine whether there are acidic, basic or neutral contributions to the rate constant. Superimposed on the plot are the lines of slope − +1 and slope − -1. Data points that fall parallel to the line of slope − -1 contribute to the acidic portion of the rate constant. Data points

that fall parallel to the line of slope - +1 contribute to the basic portion of the rate constant. Data points that are horizontal on the plot are neutral to pH.

The transformation section of the program determines the data points that will be included in the array to be analyzed as acid, base or neutral, and the energy of activation that will be used for each analysis. If the data array for each pH category contains first-order rate constants that have been measured at different temperatures, a linear regression is used with the Arrhenius plot to calculate a more accurate energy of activation. The data are ultimately reported as second-order acidic and basic rates and a first-order neutral rate at 25°C.

In the half-life section of the program, the acidic, basic and neutral contributions are combined to calculate the overall hydrolysis rate constant, k_h, and the half-life of the chemical at 25°C and pH 7.

Methods of Data Retrieval

Interactive users of the FATE database system can query the database by CAS number, by SMILES notation, by molecular formula, by a substring of the preferred chemical name, by a substring of the common names, by the REFerence number, and by primary or secondary authors from the reference publications.

a) CAS Number

The most efficient method of data retrieval from the FATE database is by entry of the CAS number. The CAS Registry number has no chemical significance, and the numbers have been assigned in sequential order as substances have been entered into the CAS Registry System for the first time.

For this reason, CAS numbers provide unique identification of chemicals that are independent of nomenclature. CAS numbers are separated by hyphens into three groups in the Registry System, but the FATE database has been designed to contain numerical characters without hyphens. A CAS-checking algorithm is incorporated in the code of the FOCUS program that allows the CAS file to be updated. Whenever a new CAS number is entered in the database, it is verified for validity with the CAS-checking algorithm.

b)   SMILES Notation

Data may be retrieved from the FATE database by use of the SMILES notation to define a chemical. SMILES notation is based on the principles of molecular graph theory and is a chemical notation language specifically designed for computer use by chemists[14,15]. SMILES notation provides unique identification of a chemical substance based on a connection table that represents the topological structure of the chemical. Therefore, the SMILES notation does have chemical significance. Computer programs are available to draw the chemical structure, based on the SMILES notation. In SMILES notation, the hydrogen atoms are suppressed, aromatic atoms are represented by lower case characters and non-aromatic atoms are represented by upper case characters. The FATE database contains a subroutine that will translate legitimate SMILES notation entered by the user on the data selection screen into the unique code that has been stored in the database before the data search is initiated.

c)   Molecular Formula

The FATE database may be searched by molecular formula, but formulas are

not unique and all chemicals with the same formula will be retrieved. Element symbols are arranged within the total formula according to the Hill System[16] for carbon-containing compounds, where C for carbon appears first, followed by H for hydrogen, then the remaining symbols are alphabetically detailed. For non-carbon-containing compounds, all symbols are arranged alphabetically.

d) Preferred and Common Names

Chemical-Abstracts-preferred index names, up to 250 characters in length, have been entered in the CAS file as they appear in the *Chemical Abstracts Service Registry Handbook*. These names were based on the Chemical Abstracts Eighth, Ninth, Tenth and Eleventh Collective Index Period nomenclature policies. The CAS file can be searched by any string up to 20 characters in length contained in a CAS name. If the first character of the search string is upper case and the rest lower case, the program will look for that string at the beginning of each CAS name. If all characters of the search string are lower case, the entire CAS file will be searched for the presence of the string imbedded in all CAS names. The report will contain all hits for the 20-character string.

Some chemical names are preceded by numeric or positional designations, such as D- or L-, .alpha.-, etc. The prefix characters have been included in a separate field for each CAS number. The prefix is not included in the search for the CAS preferred index name.

The CAS file can also be searched by common name. The search string will be an exact match of any string up to 50 characters in length. Upper and lower case rules apply as detailed above, except that prefixes are included in the search. The database may contain up to ten common names for a given

15

chemical. The report will contain all hits for the 50-character string.

e)   Reference Number

The reference citations contained in the REF file are indexed by a unique REF number. The REF number consists of one upper case letter, the first initial of the primary author's last name, and a sequential number. The database may be searched by a single reference number, a range of reference numbers, the name of the primary author, or any of the names of the secondary authors.

DATABASE MENU SYSTEM

Update of FATE

FOCUS programs that allow the database to be updated have restricted access, and can only be used by the database management team.

Data Retrieval from FATE

The FOCUS menu and reporting program can be accessed by any outside user who obtains an account and a password for the FATE database. Initial contact with the database is established by logging onto the restricted FATE account on the ATHENS VAX. Transactions from the FATE database can be captured on and printed from the personal computer of the user. Printing capabilities are not available for any users from the ATHENS VAX because of the restricted access to the FATE account. After logging onto the ATHENS VAX, the user obtains the following menu:

ENVIRONMENTAL FATE CONSTANTS
INFORMATION SYSTEM

FATE CONSTANTS REPORTS
- -
CAS File Report

Reference File Report

Fate Constants, Reference Report

Quit
- -
ENTER: C, R, F or Q

        To respond to the menu, select the first letter of the function desired

and press Enter or Return, or use the arrow key to move the highlight bar.

When C for CAS File Report is selected the following screen will appear:

                    CAS FILE LISTING SELECTION SCREEN
              Use the TAB key to move down the screen

CAS Number:

   Formula:

     Smiles:

  CAS Name:

    Common:

- Press
RETURN for the report   PF1: Help Menu
PF2: Clear the screen   PF3: Return to Main Menu

        The three selection screens in FATE have access to the same Help Menu.

The Help Menu, PF1, provides information about database fields, PF key

functions, etc. The PF2 key can be used to clear the screen of data that have

been entered previously.

        When R for Reference File Report is selected from the Main Menu, the

17

following selection screen will appear:

REFERENCE SELECTION SCREEN
Use the TAB key to move down the screen

Single Reference Number or Range: To:

Primary Author:

Secondary Author:

--   Press
RETURN for the report   PF1: Help Menu
PF2: Clear the screen   PF3: Return to Main Menu

When F for Fate Constants, Reference Report is selected from the Main

Menu, the following selection screen will appear:

FATE CONSTANTS, REFERENCE SELECTION SCREEN
Use the TAB key to move down the screen
CAS Number:
Formula:
Smiles:
CAS Name:
Common:

REF Number:

pK(a):   S(w):   K(ow):   E(1):
H(c):   K(p):   k(d):   k(h):
P(v):   K(oc):   O(r):   k(bio):

Press RETURN for the report   PF1: Help Menu
PF2: Clear the screen   PF3: Return to Main Menu

A sample report for the hydrolysis rate constant k_h, for

1,1,2,2-tetrachloroethane, CAS number [79-34-5] follows:

CAS Number: 79345   FATE Code: 05   FATE Reference: C00000165
Analytical Method: GLC   Estimating Program:
Medium: buffered dist. H2O   pH: see comments
Experimental Temperature: 95.00 C
Products: [79-01-6]
Comments: 1st order rates were meas. over the pH range 5 to 9 at 11 temp.
          Data were extrapolated to 1st and 2nd order rates at 25°C with the
          RATE program. Ea(base) was estimated as 21.2 kcal/mol or 88.8
          kJ/mol.

```
                -------------Fate Values---------------
          half-life   :  98 day          (25 C)
          k(acid)     :  0.0/M-yr         (25 C)
          k(base)     :  2.6E7/M-yr       (25 C)
          k(h), pH 7  :  2.6/yr           (25 C)
          k(neutral)  :  0.0/yr           (25 C)
```

FATE Reference: C00000165
Authors: Cooper, William J.; Mehran, Mostafa;
         Rlusech, David J.; Joens, Jeffrey A.

Date: 1987
Title: Abiotic transformation of halogenated organics. 1. Elimination
         reaction of 1,1,2,2-tetrachloroethane and formation of 1,1,2-
         trichloroethene.

Citation: Environ. Sci. Technol. 21(11):1112-1114.

DISCUSSION and CONCLUSION

The development of the FATE database grew out of the need for fate

constants in chemical risk assessment. The literature searches that were

conducted in response to EPA requests for data revealed that few values were

available, that many publications lacked sufficient documentation to determine

data credibility, and that many data were determined under environmentally

unrealistic conditions. More importantly, very few authors determined the

products of the degradation processes. Risk should be assessed for the

"persistent" chemical(s) and not for the "transient" parent compound or intermediate products alone[17].

A support activity was organized at ERL-Athens to provide equilibrium and kinetic constants for critical chemicals (and their transformation products) whose environmental transport and transformation must be assessed. This activity involves conducting literature searches for measured data, postulating transformation pathways and products, performing laboratory measurements of fate constants, and estimating values using computational techniques as required.

The FATE database was developed to eliminate a number of the problems that were experienced in this support. As a result, FATE users will find values that can be used with confidence for up to twelve rate and equilibrium constants. Data are screened for applicability to environmental assessment, and only data from primary sources are entered. If a value was determined in a manner that prevents extrapolation to environmental conditions or lacks sufficient documentation to ascertain environmental applicability, it will not be entered into the database. Transformation products are listed when available. Chemical hydrolysis rate constants are extrapolated to a standard format with a computer program developed at ERL-Athens. Acidic, basic and neutral contributions to the rate constant are combined to calculate the overall hydrolysis rate constant, k_h, and the half-life of the chemical at 25°C and pH 7. In addition, hydrolysis data are reported as second-order acidic and basic rates and a first-order neutral rate at 25°C. For critical chemicals, when measured data cannot be located, laboratory measurements may be performed at ERL-Athens. Data are also computed whenever an applicable technique is available, using both conventional techniques, e.g. QSAR, and the

newly developed SPARC expert system.

Future emphasis with the FATE database will be on data computed with SPARC. This expert system has the capability of crossing chemical boundaries to provide estimates for all organic chemicals and will generate reliable values for a fraction of the cost and time it takes to determine an experimental value.

ACKNOWLEDGMENT

The authors would like to express their appreciation for the assistance and recommendations in the development of the database to Computer Sciences Corporation, especially Mr. Matthew P. Holway.

REFERENCES

1. ACS Committee on Environmental Improvement, 1980. Guidelines for data acquisition and data quality evaluation in environmental chemistry. *Anal. Chem.* 52:2242-2249.

2. B. T. Bowman and W. W. Sans, 1983. Determination of octanol-water partition coefficients (K_{ow}) of 61 organophosphorus and carbamate insecticides and their relationship to respective water solubility (S) values. *J. Environ. Sci. Health* B18(6):667-683.

3. W. R. Mabey, J. S. Winterle, T. Podoll, H. Jaber, D. Haynes, D. Tse, V. Barich, A. Liu and T. Mill, 1984. Elements of a quality database for environmental fate assessment. Final Report, EPA Contract No. 68-03-2981, SRI International Project 2073, Work Assignment No. 6. Unpublished report.

4. P. H. Howard and G. W. Sage, 1986. Development of indicators of data documentation quality for environmental measurements data. Developed by the Syracuse Research Corporation, New York, for the Chemical Manufacturer's Association, Washington, D.C. under Contract 010114. 185, SRC F0089-01.

5. S. M. Creeger, N. K. Whetzel and C. L. Fletcher, 1985. Standard evaluation procedures: Hydrolysis, aqueous photolysis, aerobic soil metabolism, soil photolysis, soil column leaching. Report No. EPA 540/9-85-013, -014, -015, -016, -017. U.S. Environmental Protection Agency, Washington, D.C.

6. H. P. Kollig, 1988. Criteria for evaluating the reliability of literature data on environmental process constants. *Toxicol. and Environ. Chem.* 17(4):287-311.

7. T. Mill and B. T. Walton, 1987. How reliable are data-base data? *Environ. Toxicol. and Chem.* 6:161-162.

8. C. T. Chiou, V. H. Freed, D. W. Schmedding and R. L. Kohnert, 1977. Partition coefficient and bioaccumulation of selected organic chemicals. *Environ. Sci. and Technol.* 11(5):475-478.

9. S. R. Heller, 1988. Property prediction. *Ind. Chemist,* July.

10. W. J. Lyman, W. F. Reehl and D. H. Rosenblatt, 1982. *Handbook of Chemical Property Estimation Methods: Environmental Behavior of Organic Compounds.* McGraw-Hill Book Company, New York, New York.

11. The Graphical Exposure Modeling System (GEMS) is an interactive computer system located on the VAX cluster in the National Computer Center in Research Triangle Park, North Carolina, under management of EPA's Office of Toxic Substances.

12. Quantitative Structure-Activity Relationships (QSAR) is an interactive chemical database and hazard assessment system designed to provide basic information for the evaluation of the fate and effects of chemicals in the environment. QSAR was developed jointly by the U.S. EPA Environmental Research Laboratory, Duluth, Minnesota, Montana State University Center for Data Systems and Analysis, and the Pomona College Medicinal Chemistry Project.

13. S. W. Karickhoff, L. A. Carreira, C. Melton, V. K. McDaniel, A. N. Vellino, and D. E. Nute, 1989. Computer prediction of chemical reactivity---the ultimate SAR. U.S. Environmental Protection Agency, Athens, Ga. EPA/600/M-89/017.

14. E. Anderson, G. D. Veith, and D. Weininger, 1987. SMILES: a line notation and computerized interpreter for chemical structures. U.S. Environmental Protection Agency, Duluth, MN. EPA/600/M-87-021.

15. D. Weininger, 1988. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inf. Comput. Sci.* 28:31-36.

16. E. A. Hill, 1900. On a system of indexing chemical literature; adopted by the Classification Division of the U.S. Patent Office. *J. Amer. Chem. Soc.* 22(8):478-494.

17. H. P. Kollig, 1990. A fate constant data program. *Toxicol. and Environ. Chem.* 25(2-3):171-179.