TerraQSAR™ - FHM

Fathead minnow 96-hr LC50 Estimation Software, vs. 2.1

User Manual

Purpose

TerraQSAR™ - FHM is a holistic, probabilistic-algorithm-based neural network software program, designed and optimized solely for the computation of acute (96-hr) median lethal concentrations (LC50) of organic (carbon-containing) substances with a defined chemical structure to the fish fathead minnow (*Pimephales promelas*).

Table of Contents

Introduction

Welcome to **TerraQSAR™**, a breakthrough development in toxicity estimation software, developed by TerraBase Inc.

TerraQSAR™ prediction software is based on the probabilistic neural network methodology using the molecular structure of the substances under investigation. The **TerraQSAR™ - FHM** program estimates the 96-hr lethal concentrations to 50% of a population (LC50) of the North American fish fathead minnow (*Pimephales promelas*), a widely used test species.

TerraQSAR modules use as input a chemical's SMILES code (2-D or 3-D), which is an international code for the representation of chemical structures and amenable to computer analysis.

The **TerraQSAR™ - FHM** module computes the LC50 in both mg/L and pT (log [L/mmole]) units, as well as the molecular weight (MW) of substances entered.

Theory

The field of artificial intelligence and neural network application is experiencing rapid growth in all aspects of technology. From elevator control to drug design, neural network methods have lots to contribute to product development, operating improvements, and frequently enable the customer to tackle problems, which were inaccessible hitherto.

The **TerraQSAR** products exploit the neural network methodologies developed in recent years by researchers and programmers both within and outside the company. In contrast to linear methodologies, such as simple regression methods, principal components analysis and others, neural networks make use of non-linear relationships, which makes them particularly useful for chemical/biological problems where different and/or unknown modes of action are known or likely to be present, in addition to linear relationships.

Important information on both the theory and specific aspects of this software can be found in the references given in the Literature.

Computation Process

Data Set

The **TerraQSAR™ - FHM** fathead minnow toxicity estimation program is based on a data set of measured values for 886 organic (carbon-containing) compounds. These data are widely available, both from non-commercial sources, such as the US Environmental Protection Agency's Ecotox database, or from commercial sources, such as TerraBase Inc.'s **TerraTox™: Explorer** database. In fact, these data form the basis for most fish toxicity estimation programs.

Fragments

Major fragments used in the **TerraQSAR** modules have been described in detail in several publications listed in the Literature, especially the works by Kaiser *et al.* An overview of basic fragment types considered is given in Table 1 below.

Table 1. Examples of fragments used in **TerraQSAR**.

Fragment type	Examples
Acidity fragment	C(=O)O, S(=O)(=O)O
Aliphatic ring fragment	C1CCCCC1, C1CCCC1
Aromatic ring fragment	c1ccccc1, c1ccccn1
Atom fragment	C, H, N, O
Bond fragment	C-C, C=C, C#C
Group fragment	C-O-H, C-O-C, O=C-O-C
Hydrophobicity fragment	C(C)(C)C, CCCC
Ionisation fragment	[O-], [Na+]
Polarity fragment	O=N(=O)CC(O)
Reactivity fragment	C=CC=O
Stereo fragment	Cl[C@H](C)N, Cl[C@@H](C)N
Weight fragment	molecular weight

Input Query

All **TerraQSAR** modules use the SMILES string code as input. For additional comments about the SMILES code, refer to the paragraph on SMILES Notation.

Computation

The computer evaluates the number and type of fragments present in the query string and computes the resulting estimate on the basis of the same types of fragments present in a data set of 886 compounds for which measured values have been published in the literature. Computation time varies with the complexity of the query structure and speed of the computer. Typically, for compounds without chiral centers, and molecular weight of <200, computation time on a 2 GHz machine takes <5 seconds.

Results

Figure 1 shows the measured vs. predicted values for all 886 compounds used in the development of the **TerraQSAR** fathead minnow estimation program, as obtained from the program. The data cover approximately ten orders of magnitude, ranging from pT= -3 to pT= 7, where pT is the negative logarithm of the millimolar LC50 concentration.

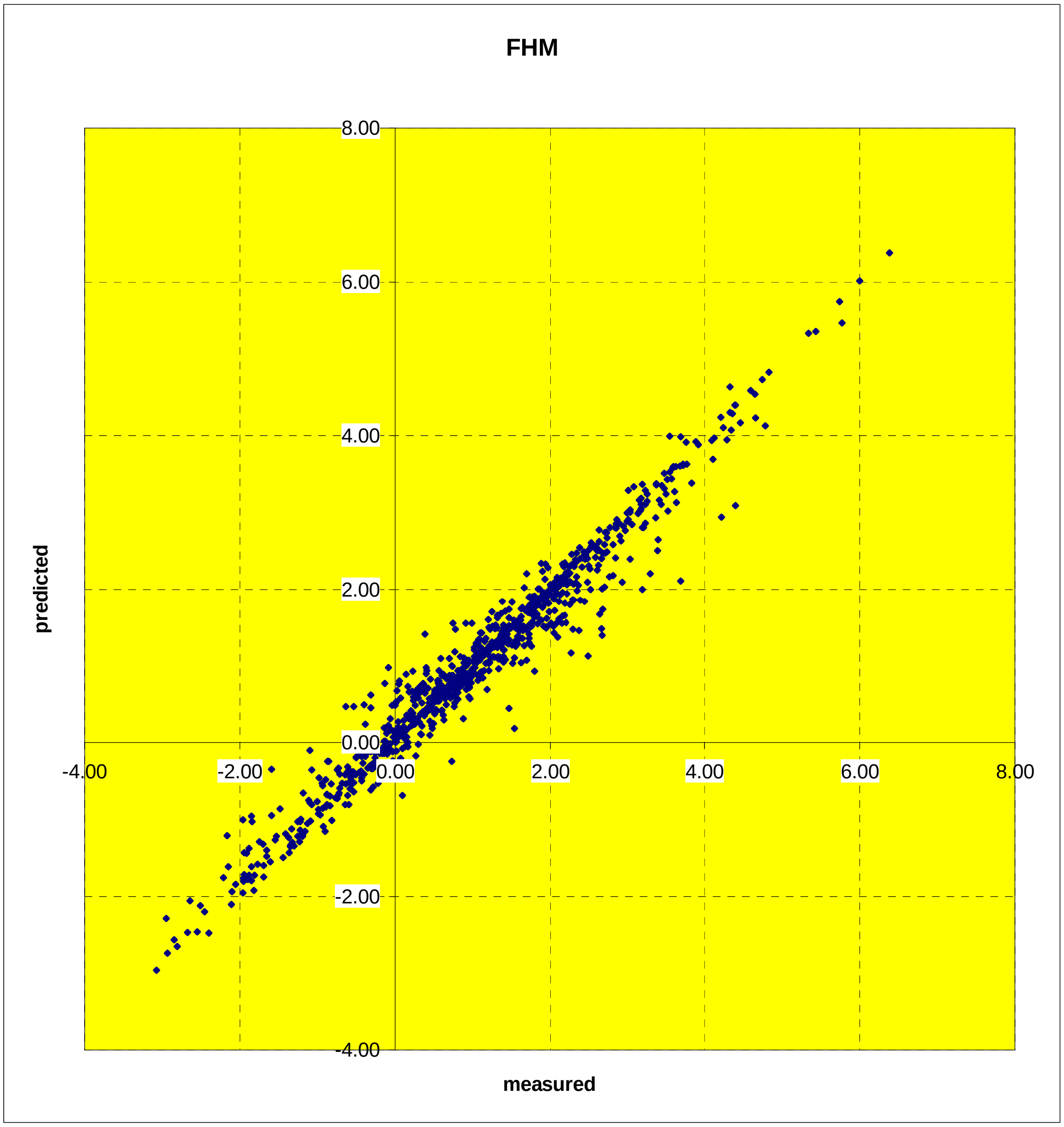


Fig. 1. Plot of the measured vs. predicted fathead minnow LC50 values for all 886 compounds in the training set; units are pT, equal to log(L/mmol).

SMILES Notation

The Simplified Molecular Input Line Entry System (SMILES) has been developed by D. Weininger at the beginning of the 1980's. It is far superior to the previously used Wiswesser Line Notation (WLN) for coding and depicting chemical structures, by being simple, intuitive, and machine readable. For an excellent tutorial on the SMILES notation, refer to the Daylight Corp. web site.

Recently, Accelrys Inc. introduced a variety of software modules allowing the visualization of SMILES codes as chemical structure drawings. In this process, Accelrys introduced changes to the common (Daylight Corp.) interpretation of SMILES codes by their software. As a result, lower case "c", formerly only interpreted as sp^2 carbon, is now interpreted as either as sp^2 or sp^3 carbon, depending on its surrounding and connections to other atoms. The determinant here is whether or not the carbon atom is part of an aromatic ring, as defined by the Hueckel rules. This has ramifications for the correct interpretation of SMILES strings by the **TerraQSAR** programs, as they are built on the backbone of the Accelrys software. Therefore, all users are cautioned to ascertain that their SMILES codes follow the rules of the Accelrys software, i.e., to ascertain that only sp^2 carbons in ring systems which satisfy Hueckel conditions for aromaticity are given in lower case "c"; all other sp^2 carbons, whether in rings or not, must be entered as capital "C". Some examples of valid and not valid SMILES strings are listed below in Table 2.

Table 2. Valid and not valid examples of Accelrys' SMILES code.

Substance	SMILES not valid	SMILES valid
cyclopentadiene [a]	c1cccC1	C1=CC=CC1
coumarin [a]	c1cc2OC(=O)ccc2cc1	c1cc2OC(=O)C=Cc2cc1

[a] The SMILES strings shown as "not valid" are valid *per se*, however, the interpretation of these codes are the hydrogen-saturated compounds cyclopentane and 3,4-dihydrocoumarin, respectively.

European (SIDS) Evaluation

The SIDS Evaluation program by the European Community undertook a comparative evaluation of six fish toxicity QSAR programs, including the **TerraQSAR - FHM** (fathead minnow) program. The following text outlines the main steps and results, for a complete copy of the report, see Pavan et al. (2005).

Method outline

1. Preliminary analysis of SIDS acute fish toxicity data.
2. Generation of molecular structure files for the SIDS chemicals (Smiles, mol files), for further calculation of both two-dimensional molecular descriptors and three-dimensional descriptors. An excel file containing chemical names, CAS numbers and SMILES for 177 chemicals was kindly provided by Eva Wedebye (DK).
3. Development of a list of literature-based models to make predictions of SIDS endpoints. The focus was on models for fish toxicity.
4. Selection of transparent and reproducible models: recovery of the training set used to develop the models and checking of the test method used to generate it; identification of the molecular descriptors used and assessment of the transparency of the algorithm.
5. Estimation of predictive ability by internal validation techniques (cross-validation, bootstrap, response randomization).
6. Evaluation of QSAR applicability domains by making predictions of SIDS test data: checking the domain of applicability with respect to descriptor ranges and any structural rules defining the group of substances for which the models are valid.
7. Application of the models to the SIDS chemicals.
8. Evaluation of predictive performance in terms of explained variance (Q^2_{ext}) and the prediction reliability (order of magnitude between estimated and experimental data). Predictive

performance was assessed for the full set of SIDS substances, and for subsets based on different hypotheses about the applicability domain.
9. Comparative analysis of the model quality.

Results of Method

Table 3, below, is an excerpt from “Comparative Assessment of QSAR Models for Aquatic Toxicity” (Pavan et al., 2005), comparing six different QSAR models for the “Danish dataset” of 177 SIDS compounds. As is evident from Table 3, the TerraQSAR™ - FHM program delivers estimates for all (177) test compounds (versus 125 compounds by the nearest competitor), and has the highest Q^2*ext* of all tested programs. A copy of the complete report can be accessed through the reference (Pavan et al., 2005).

Table 3. Model performance comparison.

Model	N. Train	SIDS Train	Q^2_{LOO}	$Q^2_{boostrap}$	SDEP	R^2	Test MOA	N.Test	Unknown SIDS predictions	Total SIDS predictions	Q^2ext
NPN	58	8	91.51	91.66	0.421	92.18	NPN	14	37	51	89.06
							Mixed	28	97	125	90.86
PN	86	5	89.59	89.64	0.336	90.07	PN	2	4	6	N.A.
							Mixed	25	98	123	86.66
N	144	13	87.06	87.11	0.461	87.55	NPN + PN	13	41	54	92.18
							Mixed	24	97	121	91.63
MIXED	114	9	75.94	75.83	0.495	77.57	Mixed	22	51	73	87.10
E-State	121	8	68.28	9.30	0.505	84.04	Mixed	17	69	86	89.43
Terra QSAR	886	N.A	N.A	N.A	N.A	94.56	Mixed	57	120	177	99.39

Literature

Accelrys Inc. (2002); http://www.accelrys.com/.

Carpenter, W.C., and J.F. Barthelemy. Common misconceptions about neural networks as approximators**. Proc. 3rd Int. Conf. Applic. Artificial Intelligence to Civil and Structural Engineering in Neural Networks and Combinatorial Optimization in Civil and Structural Engineering**. Topping, B.H.V. (Ed.), pp. 11-18 (1993).

Daylight Chemical Information Systems, Inc. (2002); http://www.daylight.com/smiles/f_smiles.html.

Eldred, D.V., C.L. Weikel, P.C. Jurs, and K.L.E. Kaiser. Prediction of fathead minnow acute toxicity of organic compounds from molecular structure. **Chem. Res. Toxicol., 12:** 670-678 (1999).

Grabec, I. Self-organization of neurons described by the maximum-entropy principle. **Biol. Cybern., 63:** 403-409 (1990).

Grabec, I. Optimization of kernel-type density estimator by the principle of maximal self-consistency. **Neural Parallel & Scientific Computations, 1:** 83-92 (1993).

Hecht-Nielsen, R. Nearest matched filter classification of spatiotemporal patterns. **Appl. Optics, 26:** 1892-1899 (1987).

Hecht-Nielsen, R. **Neurocomputing**. Addison-Wesley Publishing Co., Inc. (1990).

Hertz, J., A. Rogh, and R. Palmer. **Introduction to the Theory of Neural Computation**. Addison-Wesley, Redwood City, California (1991).

Johansson, E.M., F.U. Dowla, and D.M. Goodman. Backpropagation learning for multi-layer feed-forward neural nets using the conjugate gradient method. **Lawrence Livermore National Laboratory** (1990).

Kaiser, K.L.E., S.P. Niculescu, and M.B. McKinnon. On the simple linear regression, the multiple linear regression and the elementary probabilistic neural network with Gaussian kernel's performance in modeling toxicity values to fathead minnow based on Microtox data, the octanol/water partition coefficient and various structural descriptors for a 419 compound data set. **QSAR in Environmental Sciences - VII**, F. Chen and G. Schüürmann (Eds.), SETAC Press, Pensacola, FL, pp. 285-297 (1997).

Kaiser, K.L.E., S.P. Niculescu, and G. Schüürmann. Feed forward backpropagation neural networks and their use in predicting the acute toxicity of chemicals to the fathead minnow. **Water Quality Res. J. Canada, 32**: 637-657 (1997); http://www.cciw.ca/wqrjc/32-3/32-3-637.htm.

Kaiser, K.L.E., S.P. Niculescu, and K.M. Gough. Neural network modeling of *Vibrio fischeri* and fathead minnow acute toxicity data with molecular indicator variables and physico-chemical bulk parameters. Poster, **Workshop on Computational Methods in Toxicology**, Dayton, OH, April 20-22, (1998), http://www.ccl.net/ccl/toxicology/abstracts/abs14.html .

Kaiser, K.L.E., and S.P. Niculescu. Using probabilistic neural networks to model the toxicity of chemicals to the fathead minnow (*Pimephales promelas*): A study based on 865 compounds. **Chemosphere**, **38:** 3237-3245 (1999).

Kaiser, K.L.E., and S.P. Niculescu. Probabilistic neural network (PNN) methodology for the prediction of acute toxicity of chemicals to fathead minnow based solely on chemical structure-derived input parameters. **National Water Research Institute Contribution, No. AEP-TN99-001**, 39 p. (1999).

Kaiser, K.L.E., and S.P. Niculescu. Modeling the acute toxicity of chemicals to *Daphnia magna*: a probabilistic neural network approach. **Environ. Toxicol. Chem.**, **20:** 420-431 (2001).

Kaiser, K.L.E., S.P. Niculescu, and T.W. Schultz. Probabilistic neural network modeling of the toxicity of chemicals to *Tetrahymena pyriformis* with molecular fragment descriptors. **SAR & QSAR Environ. Res., 13:** 57-67 (2002).

Kaiser, K.L.E., and S.P. Niculescu. On the PNN modeling of estrogen receptor binding data for carboxylic acid esters and organochlorine compounds. **Water Qual. Res. J. Canada**, **36:** 619-630 (2001); http://www.cciw.ca/wqrjc/36-3/36-3-619.htm .

Kohonen, T. **Self-Organization and Associative Memory**, 2nd ed. Springer-Verlag, Berlin, (1988).

Kühne, R., K.L.E. Kaiser, S.P. Niculescu, and G. Schüürmann. Neural network to estimate acute toxicity of chemicals to the fathead minnow and *Vibrio fischeri* bacteria. **SETAC 99 - Europe**, Leipzig, May 1999, poster.

Martin, T.M., and D.M. Young. Prediction of the acute toxicity (96-hr LC50) of organic compounds to the fathead minnow (Pimephales promelas) using a group contribution method. **Chem. Res. Toxicol., 14:** 1378-1385 (2001).

Masters, T. **Practical Neural Network Recipes in C++**. San Diego, Academic Press (1993).

Niculescu, S.P., K.L.E. Kaiser, and G. Schüürmann. Influence of data preprocessing and kernel selection on probabilistic neural network modeling of the acute toxicity of chemicals to the fathead minnow and to *Vibrio fischeri* bacteria. **Water Quality Res. J. Canada, 33:** 153-165 (1998); http://www.cciw.ca/wqrjc/33-1/33-1-153.htm .

Niculescu, S.P., and K.L.E. Kaiser. Modeling the relative binding affinity of steroids to the progesterone receptor with probabilistic neural networks. **Quant. Struct.-Act. Relat.**, **20:** 223-226 (2001).

Niculescu, S.P., K.L.E. Kaiser, and T.W. Schultz. Modeling the toxicity of chemicals to *Tetrahymena pyriformis* using molecular fragment descriptors and probabilistic neural networks. **Archiv. Environ. Toxicol. Chem., 39:** 289-298 (2000).

Pavan, M., A.P. Worth, and T.I. Netzeva. Comparative assessment of QSAR models for aquatic toxicity. European Commission. EUR 21750 EN, 120 pp. (2005).

Prechelt, L. (Ed.). Frequently asked questions (FAQ) on neural networks. (1995); http://www.ipd.uka.de/~prechelt/FAQ/nn7.html.

Ramirez, M.R., and D. Arghya. A faster learning algorithm for back-propagation neural networks in NDE applications. **Proc. 2nd Int. Conf. on AI**, pp. 275-283 (1991).

Samaad, T. Backpropagation improvements based on heuristic arguments, theory track, neural and cognitive sciences. **Track. Int. Joint Conf. Neural Networks**, Vol. 1, Washington, D.C., (1990).

Sarle, W.S. (Ed.). Neural network FAQ, part 1 of 7: Introduction. (2002); http://accpc.com/nnfaq/FAQ.html.

Serra, J.R., P.C. Jurs, and K.L.E. Kaiser. Linear regression and computational neural network prediction of *Tetrahymena* acute toxicity for aromatic compounds from molecular structure. **Chem. Res.Toxicol.**, **14:** 1535-1545 (2001).

Shanno, D.F. Conjugate gradient methods with inexact searches. **Math. Oper. Res., 3:** 244-256 (1978).

Specht, D.F. Probabilistic neural networks for classification, mapping or associative memory. **ICNN, Conference Proc.** (1988).

Yeh, Y., Y. Kuo, and D. Hsu**.** Building KBES for diagnosis PC pile with artificial neural network. **J. Comp. Civil Eng., 7:** 71-93 (1993).

Acknowledgement

TerraBase Inc. acknowledges with thanks the contributions of its contractors and associates, especially Dr. A. Krajnc, M.B. McKinnon, and Dr. S.P. Niculescu to the development of its products.

Security Features

As other software providers, TerraBase Inc. is interested is preventing un-authorized duplication and use of its highly sophisticated software products. Therefore, all copies of this **TerraQSAR** program are personalized, i.e., they contain encrypted information of the name and affiliation of the rightful user. This information is displayed on the Start-up screen and can be reviewed at any time by activating the About box, prominently displayed on the main program panel; see Figure 2. Hence, any un-authorized propagation of this software can be traced back to the original source. Anyone knowing of any such un-authorized copies is requested to contact TerraBase Inc. We offer generous rewards for information of any illegal duplication or use of this program, as described in the About box.

In order for the **TerraQSAR** program to function properly, the TerraQSAR CD has to be present in a CD-ROM drive on the computer system which is used to run the program.

Other than for archival copies, to be used solely for backup and protection against loss by the rightful user, duplication of this software is not permitted.

Legal

License Agreement

The rightful **TerraQSAR** user is hereby granted a non-exclusive use license, subject to the agreement as described in the text below.

TerraQSAR™
End User License Agreement

NOTICE TO USER: PLEASE READ THIS CONTRACT CAREFULLY. BY USING ALL OR ANY PORTION OF THE SOFTWARE YOU ACCEPT ALL THE TERMS AND CONDITIONS OF THIS AGREEMENT, INCLUDING, IN PARTICULAR THE LIMITATIONS ON: USE CONTAINED IN SECTION 2; TRANSFERABILITY IN SECTION 4; WARRANTY IN SECTION 6 AND 7; LIABILITY IN SECTION 8; AND SPECIFIC EXCEPTIONS IN SECTION 14. YOU AGREE THAT THIS AGREEMENT IS ENFORCEABLE LIKE ANY WRITTEN NEGOTIATED AGREEMENT SIGNED BY YOU. IF YOU DO NOT AGREE, DO NOT USE THIS SOFTWARE. IF YOU ACQUIRED THE SOFTWARE ON TANGIBLE MEDIA (e.g. CD) WITHOUT AN OPPORTUNITY TO REVIEW THIS LICENSE AND YOU DO NOT ACCEPT THIS AGREEMENT, YOU MAY OBTAIN A REFUND OF THE AMOUNT YOU ORIGINALLY PAID IF YOU: (A) DO NOT USE THE SOFTWARE AND (B) RETURN IT, WITH PROOF OF PAYMENT, TO THE LOCATION FROM WHICH IT WAS OBTAINED WITHIN THIRTY (30) DAYS OF THE PURCHASE DATE.

1. Definitions.
"Software" means (a) all of the contents of the files, disk(s), CD-ROM(s) or other media with which this Agreement is provided, including but not limited to (i) TerraBase Inc. or third party computer information or software; (ii) digital images, stock photographs, clip art, sounds or other artistic works ("Stock Files"); (iii) related explanatory written materials or files ("Documentation"); and (b) upgrades, modified versions, updates, additions, and copies of the Software, if any, licensed to you by TerraBase Inc. (collectively, "Updates"). "Use" or "Using" means to access, install, download, copy or otherwise benefit from using the functionality of the Software in accordance with the Documentation. "Permitted Number" means one (1) unless otherwise indicated under a valid license (e.g. volume license) granted by TerraBase Inc.. "Computer" means an electronic device that accepts information in digital or similar form and manipulates it for a specific result based on a sequence of instructions. "TerraBase" means TerraBase Inc., a Canadian corporation, 1363 King St. West, Suite 130, Hamilton, Ontario L8S 4S3, CANADA.

2. Software License.
As long as you comply with the terms of this End User License Agreement (the "Agreement"), TerraBase grants to you a non-exclusive license to Use the Software for the purposes described in the Documentation. Some third party materials included in the Software may be subject to other terms and conditions, which are typically found in a "Read Me" file located near such materials.

2.1. General Use. You may install and Use a copy of the Software solely on your compatible personal computer.

Server Use. You may not install the Software on a computer file server for the purpose of downloading, installing, or using the Software onto or with other computers.

2.3. Backup Copy. You may make one backup copy of the Software, provided your backup copy is not installed or used on any computer. You may not transfer the rights to a backup copy unless you transfer all rights in the Software as provided under Section 4.

3. Intellectual Property Rights.
The Software and any copies that you are authorized by TerraBase to make are the intellectual

property of and are owned by TerraBase Incorporated and its suppliers. The structure, organization and code of the Software are the valuable trade secrets and confidential information of TerraBase Incorporated and its suppliers. The Software is protected by copyright, including without limitation by Canadian and United States Copyright Law, international treaty provisions and applicable laws in the country in which it is being used. You may not copy the Software, except as set forth in Section 2 ("Software License"). Any copies that you are permitted to make pursuant to this Agreement must contain the same copyright and other proprietary notices that appear on or in the Software. Without exception, you agree not to modify, adapt or translate the Software. You also agree not to reverse engineer, decompile, disassemble or otherwise attempt to discover the source code of the Software. Trademarks shall be used in accordance with accepted trademark practice, including identification of trademarks owners' names. Trademarks can only be used to identify printed output produced by the Software and such use of any trademark does not give you any rights of ownership in that trademark. Except as expressly stated above, this Agreement does not grant you any intellectual property rights in the Software.

4. Transfer.

You may not rent, lease, sublicense or authorize all or any portion of the Software to be copied onto another users computer except as may be expressly permitted herein. You may, however, transfer all your rights to Use the Software to another person or legal entity provided that: (a) you also transfer each this Agreement, the Software and all other software or hardware bundled or pre-installed with the Software, including all copies, Updates and prior versions, and all copies of font software converted into other formats, to such person or entity; (b) you retain no copies, including backups and copies stored on a computer; and (c) the receiving party accepts the terms and conditions of this Agreement and any other terms and conditions upon which you legally purchased a license to the Software. Notwithstanding the foregoing, you may not transfer education, pre-release, or not for resale copies of the Software.

5. Updates.

If the Software is an Update to a previous version of the Software, you must possess a valid license to such previous version in order to Use the Update. You may continue to Use the previous version of the Software on your computer after you receive the Update to assist you in the transition to the Update, provided that: the Update and the previous version are installed on the same computer; the previous version or copies thereof are not transferred to another party or computer unless all copies of the Update are also transferred to such party or computer; and you acknowledge that any obligation TerraBase may have to support the previous version of the Software may be ended upon availability of the Update.

6. LIMITED WARRANTY.

Except as may be otherwise provided in Section 14, TerraBase warrants to the person or entity that first purchases a license for the Software for use pursuant to the terms of this license, that the Software will perform substantially in accordance with the Documentation for the ninety (90) day period following receipt of the Software when used on the recommended hardware configuration. Non-substantial variations of performance from the Documentation does not establish a warranty right. THIS LIMITED WARRANTY DOES NOT APPLY TO UPDATES, FONT SOFTWARE CONVERTED INTO OTHER FORMATS, PRE-RELEASE (BETA), TRYOUT, PRODUCT SAMPLER, OR NOT FOR RESALE (NFR) COPIES OF SOFTWARE (See Section 14). To make a warranty claim, you must return the Software to the location where you obtained it along with proof of purchase within such ninety (90) day period. If the Software does not perform substantially in accordance with the Documentation, the entire liability of TerraBase and your exclusive remedy shall be limited to either, at TerraBase's option, the replacement of the Software or the refund of the license fee you paid for the Software. THE LIMITED WARRANTY SET FORTH IN THIS SECTION GIVES YOU SPECIFIC LEGAL RIGHTS. YOU MAY HAVE ADDITIONAL RIGHTS WHICH VARY FROM JURISDICTION TO JURISDICTION. For further warranty information, please see the jurisdiction specific information at the end of this Agreement, if any, or contact TerraBase.

7. DISCLAIMER.
THE FOREGOING LIMITED WARRANTY STATES THE SOLE AND EXCLUSIVE REMEDIES FOR TERRABASE'S OR ITS SUPPLIER'S BREACH OF WARRANTY. TERRABASE AND ITS SUPPLIERS DO NOT AND CANNOT WARRANT THE PERFORMANCE OR RESULTS YOU MAY OBTAIN BY USING THE SOFTWARE. EXCEPT FOR THE FOREGOING LIMITED WARRANTY, AND FOR ANY WARRANTY, CONDITION, REPRESENTATION OR TERM TO THE EXTENT TO WHICH THE SAME CANNOT OR MAY NOT BE EXCLUDED OR LIMITED BY LAW APPLICABLE TO YOU IN YOUR JURISDICTION, TERRABASE AND ITS SUPPLIERS MAKE NO WARRANTIES, CONDITIONS, REPRESENTATIONS OR TERMS, EXPRESS OR IMPLIED, WHETHER BY STATUTE, COMMON LAW, CUSTOM, USAGE OR OTHERWISE AS TO ANY OTHER MATTERS, INCLUDING BUT NOT LIMITED TO NON-INFRINGEMENT OF THIRD PARTY RIGHTS, INTEGRATION, SATISFACTORY QUALITY OR FITNESS FOR ANY PARTICULAR PURPOSE. The provisions of this section 7 shall survive the termination of this Agreement, howsoever caused, but this shall not imply or create any continued right to Use the Software after termination of this Agreement.

8. LIMITATION OF LIABILITY.
IN NO EVENT WILL TERRABASE OR ITS SUPPLIERS BE LIABLE TO YOU FOR ANY DAMAGES, CLAIMS OR COSTS WHATSOEVER OR ANY CONSEQUENTIAL, INDIRECT, INCIDENTAL DAMAGES, OR ANY LOST PROFITS OR LOST SAVINGS, EVEN IF AN TERRABASE REPRESENTATIVE HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH LOSS, DAMAGES, CLAIMS OR COSTS OR FOR ANY CLAIM BY ANY THIRD PARTY. THE FOREGOING LIMITATIONS AND EXCLUSIONS APPLY TO THE EXTENT PERMITTED BY APPLICABLE LAW IN YOUR JURISDICTION. TERRABASE'S AGGREGATE LIABILITY AND THAT OF ITS SUPPLIERS UNDER OR IN CONNECTION WITH THIS AGREEMENT SHALL BE LIMITED TO THE AMOUNT PAID FOR THE SOFTWARE, IF ANY. Nothing contained in this Agreement limits TerraBase's liability to you in the event of death or personal injury resulting from TerraBase's negligence or for the tort of deceit (fraud). TerraBase is acting on behalf of its suppliers for the purpose of disclaiming, excluding and/or limiting obligations, warranties and liability as provided in this Agreement, but in no other respects and for no other purpose. For further information, please see the jurisdiction specific information at the end of this Agreement, if any, or contact TerraBase.

9. Export Rules. You agree that the Software will not be shipped, transferred or exported into any country or used in any manner prohibited by the any law governing such activities either in Canada or the United States, such as the United States Export Administration Act or any other export laws, restrictions or regulations (collectively the "Export Laws"). In addition, if the Software is identified as export controlled items under the Export Laws, you represent and warrant that you are not a citizen, or otherwise located within, an embargoed nation (including without limitation Iran, Iraq, Syria, Sudan, Libya, Cuba, North Korea, and Serbia) and that you are not otherwise prohibited under the Export Laws from receiving the Software. All rights to Use the Software are granted on condition that such rights are forfeited if you fail to comply with the terms of this Agreement.

10. Governing Law.
This Agreement will be governed by and construed in accordance with the substantive laws in force: (a) the Province of Ontario, if the license to the Software is purchased when you are in Canada, the United States, Canada, or Mexico; or (b) in Japan, if a license to the Software is purchased when you are in Japan, China, Korea, R.O.C, or other Southeast Asian country where all official languages are written in either an ideographic script (e.g., hanzi, kanji, or hanja), and/or other script based upon or similar in structure to an ideographic script, such as hangul or kana; or (c) the Netherlands, if a license to the Software is purchased when you are in any other jurisdiction not described above. The respective courts of Ontario when Canadian law applies, Tokyo District Court in Japan, when Japanese law applies, and the courts of Amsterdam, the Netherlands, when the law of the Netherlands applies, shall each have non-exclusive jurisdiction over all disputes relating to this Agreement. This Agreement will not be governed by the conflict of law rules of any jurisdiction or the United Nations

Convention on Contracts for the International Sale of Goods, the application of which is expressly excluded.

11. General Provisions.
If any part of this Agreement is found void and unenforceable, it will not affect the validity of the balance of the Agreement, which shall remain valid and enforceable according to its terms. This Agreement shall not prejudice the statutory rights of any party dealing as a consumer. This Agreement may only be modified by a writing signed by an authorized officer of TerraBase. Updates may be licensed to you by TerraBase with additional or different terms. This is the entire agreement between TerraBase and you relating to the Software and it supersedes any prior representations, discussions, undertakings, communications or advertising relating to the Software.

12. Notice to U.S. Government End Users.
The Software and Documentation are "Commercial Items," as that term is defined at 48 C.F.R. §2.101, consisting of "Commercial Computer Software" and "Commercial Computer Software Documentation," as such terms are used in 48 C.F.R. §12.212 or 48 C.F.R. §227.7202, as applicable. Consistent with 48 C.F.R. §12.212 or 48 C.F.R. §§227.7202-1 through 227.7202-4, as applicable, the Commercial Computer Software and Commercial Computer Software Documentation are being licensed to U.S. Government end users (a) only as Commercial Items and (b) with only those rights as are granted to all other end users pursuant to the terms and conditions herein. Unpublished-rights reserved under the copyright laws of the United States by TerraBase Incorporated, 1363 King St. West, Suite 130, Hamilton, Ontario L8S 4S3, CANADA.

13. Compliance with Licenses.
If you are a business or organisation, you agree that upon request from TerraBase or TerraBase's authorised representative, you will within thirty (30) days fully document and certify that use of any and all TerraBase Software at the time of the request is in conformity with your valid licenses from TerraBase.

14. SPECIFIC EXCEPTIONS.
14.1 Limited Warranty for Users Located in Germany or Austria. If you purchased a license to the software in Germany or Austria then Section 6 does not apply, instead, TerraBase warrants to the person or entity that first purchases a license for the Software, that the Software will perform substantially in accordance with the Documentation for a period of six (6) months following receipt of the Software when used on the recommended hardware configuration. Non-substantial variations of performance from the Documentation does not establish a warranty right. THIS LIMITED WARRANTY DOES NOT APPLY TO UPDATES, FONT SOFTWARE CONVERTED INTO OTHER FORMATS, PRE-RELEASE, TRYOUT, PRODUCT SAMPLER, NOT FOR RESALE (NFR) COPIES OF SOFTWARE, OR TO SOFTWARE THAT HAS BEEN ALTERED BY YOU, TO THE EXTENT SUCH ALTERATIONS CAUSED A DEFECT. To make a warranty claim, you must return the Software, at our expense, to the location where you obtained it along with proof of purchase within such six (6) month period. If the Software does not perform substantially in accordance with the Documentation, the entire liability of TerraBase and your exclusive remedy shall be limited to either, at TerraBase's option, the replacement of the Software, the reduction of the license fee, or a refund of the license fee you paid for the Software. THE LIMITED WARRANTY SET FORTH IN THIS SECTION GIVES YOU SPECIFIC LEGAL RIGHTS. For further warranty information, please contact TerraBase.

14.2 Limitation of Liability for Users Located in Germany and Austria. If you purchased a license to the software in Germany or Austria then Section 8 does not apply, instead, TerraBase may be liable without limitation for damages you have incurred under or in connection with this Agreement only if the damage has been caused by the willful or grossly negligent act of TerraBase or its agents. TerraBase is liable only to the extent of the typically foreseeable damage for such damages which have been caused by any other negligent breach of a substantial contractual duty by TerraBase or its agents. These aforementioned limitations apply irrespective of their legal basis, in particular with regard to any pre-contractual or auxiliary contractual claims. The limitations shall not apply, however, to any mandatory liability under the applicable German or Austrian Product Liability Act, nor to any damage which is caused due to the breach of an express warranty to the extent that such express warranty was intended to protect the user against the specific damage incurred. The obligations under the second sentence shall not apply if any damage is caused by your having altered the Software, or if the respective data carriers (media) have been damaged by accident, misuse or inappropriate use, or if the damage concerned has been caused by the use of the Software in connection with other software for which the Software was not designated to be used according to the Documentation.

14.3 Pre-release Product Additional Terms. If the product you have received with this license is pre-commercial release or beta Software ("Pre-release Software"), then the following Section applies. To the extent that any provision in this Section is in conflict with any other term or condition in this Agreement, this Section shall supercede such other term(s) and condition(s) with respect to the Pre-release Software, but only to the extent necessary to resolve the conflict. You acknowledge that the Software is a pre-release version, does not represent final product from TerraBase, and may contain bugs, errors and other problems that could cause system or other failures and data loss. Consequently, the Pre-release Software is provided to you "AS-IS", and TerraBase disclaims any warranty or liability obligations to you of any kind. WHERE LEGALLY LIABILITY CANNOT BE EXCLUDED FOR PRE-RELEASE SOFTWARE, BUT IT MAY BE LIMITED, TERRABASE'S LIABILITY AND THAT OF ITS SUPPLIERS SHALL BE LIMITED TO THE SUM OF FIFTY DOLLARS (U.S. $50) IN TOTAL. You acknowledge that TerraBase has not promised or guaranteed to you that Pre-release Software will be announced or made available to anyone in the future, that TerraBase has no express or implied obligation to you to announce or introduce the Pre-release Software and that TerraBase may not introduce a product similar to or compatible with the Pre-release Software. Accordingly, you acknowledge that any research or development that you perform regarding the Pre-release Software or any product associated with the Pre-release Software is done entirely at your own risk. During the term of this Agreement, if requested by TerraBase, you will provide feedback to TerraBase regarding testing and use of the Pre-release Software, including error or bug reports. If you have been provided the Pre-release Software pursuant to a separate written agreement, such as the TerraBase Incorporated Serial Agreement for Unreleased Products, your use of the Software is also governed by such agreement. You agree that you may not and certify that you will not sublicense, lease, loan, rent, or transfer the Pre-release Software. Upon receipt of a later unreleased version of the Pre-release Software or release by TerraBase of a publicly released commercial version of the Software, whether as a stand-alone product or as part of a larger product, you agree to return or destroy all earlier Pre-release Software received from TerraBase and to abide by the terms of the End User License Agreement for any such later versions of the Pre-release Software. Notwithstanding anything in this Section to the contrary, if you are located outside Canada or the United States of America, you agree that you will return or destroy all unreleased versions of the Pre-release Software within thirty (30) days of the completion of your testing of the Software when such date is earlier than the date for TerraBase's first commercial shipment of the publicly released (commercial) Software.

14.4 Tryout, Product Sampler, NFR, Additional Terms. If the product you have received with this license is a tryout, product sampler, or NFR copy of the Software ("Tryout Software"), then the following Section applies until such time that you purchase a license to the full retail version of such product. To the extent that any provision in this Section is in conflict with any other term or condition in this Agreement, this Section shall supercede such other term(s) and condition(s) with respect to the Tryout Software, but only to the extent necessary to resolve the conflict. YOU ACKNOWLEDGE THAT THE TRYOUT SOFTWARE CONTAINS LIMITED FUNCTIONALITY AND/OR FUNCTIONS FOR A LIMITED PERIOD OF TIME. TERRABASE IS LICENSING THE SOFTWARE ON AN "AS IS" BASIS, SOLELY AS A DEMONSTRATION MODEL. If the Tryout Software is a timeout version, then the program will terminate operations after a designated period of time (e.g. 15. 30. or 45 days) following installation (the "Time Out Date"), which is specified in the Software. Upon such Time Out Date, the license hereunder shall be terminated, unless extended by

TerraBase upon your purchase of a full retail license from TerraBase. You acknowledge that such Tryout Software shall cease operation upon the Time Out Date and accordingly, access to any files or output created with such Tryout Software or any product associated with the Tryout Software is done entirely at your own risk. TERRABASE DISCLAIMS ANY WARRANTY OR LIABILITY OBLIGATIONS TO YOU OF ANY KIND, INCLUDING. WHERE LEGALLY LIABILITY CANNOT BE EXCLUDED FOR PRE-RELEASE SOFTWARE, BUT IT MAY BE LIMITED, TERRABASE'S LIABILITY AND THAT OF ITS SUPPLIERS SHALL BE LIMITED TO THE SUM OF FIFTY DOLLARS (U.S. $50) IN TOTAL.

15. This concludes the License Agreement for the TerraQSAR software. If you have any questions regarding this Agreement or if you wish to request any information from TerraBase please use the address and contact information included with this product to contact TerraBase.
Gen_WWCombined_USEnglish_8.7.00_11:14

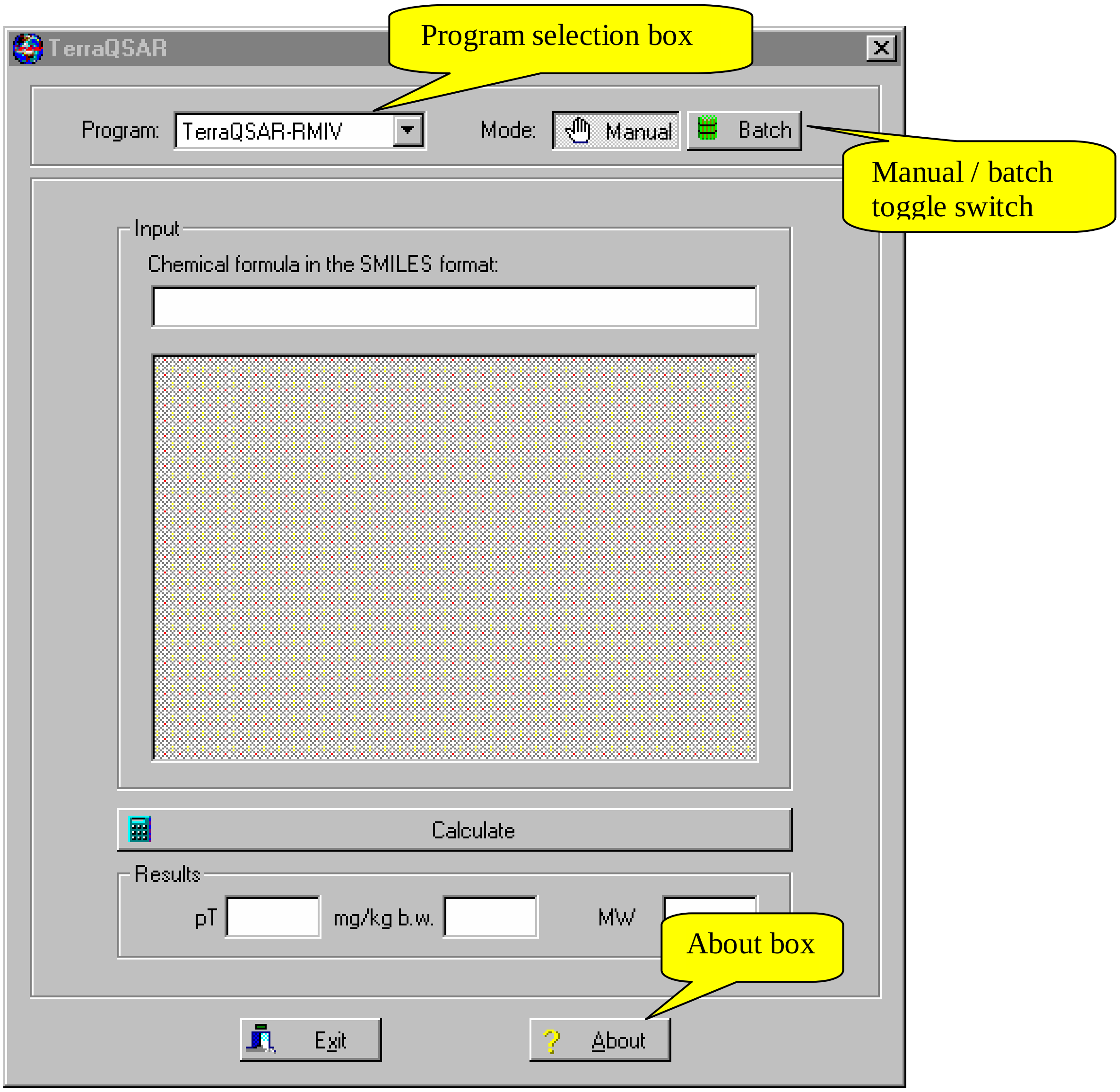
Trademarks and Copyrights

Windows 95, Windows 98, Windows NT, Windows 2000, Windows ME, and Windows XP are trademarks of Microsoft Corporation, Redmond, Washington, USA.

Other copyrights and trademarks are those of their respective owners.

Program Interface

The program interface of the **TerraQSAR** toxicity prediction modules is shown in Figure 2. It is simple, intuitive, and highly functional.



Program: TerraQSAR-RMIV

Mode: Manual

Program selection box

Manual / batch toggle switch

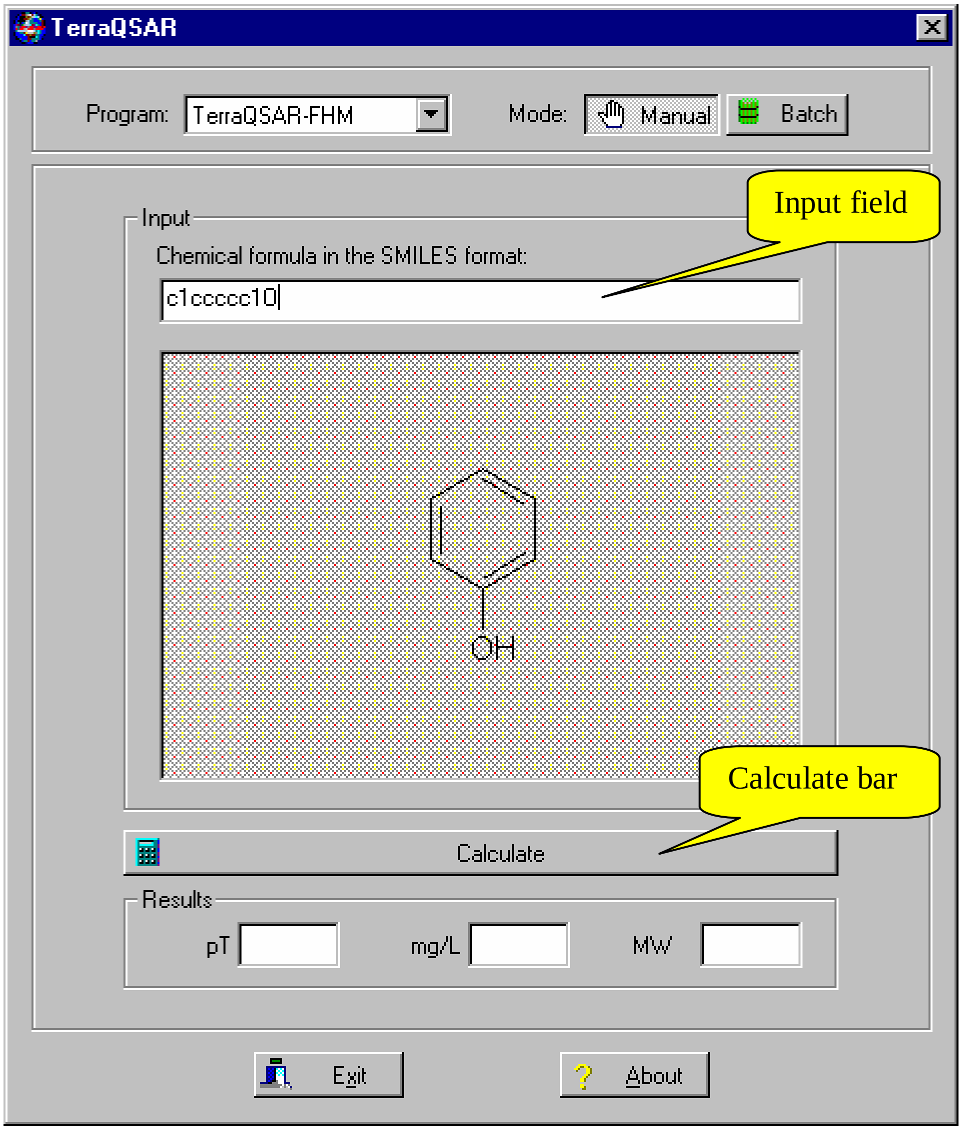
About box

Fig. 2. The Program Interface as it appears on startup.

The Program selection box allows the user to the selected program, e.g. FHM for “Fathead minnow LC50”, “RMIV for “Rat / Mouse *iv*. LD50”, etc. The second box provides a toggle switch between switch between Manual mode and Batch mode (available in the Professional versions only). The About box contains important user information.

In the Manual mode, the user enters a SMILES (Simplified Molecular Input Line Entry System) string, either by typing it into the Input field, or by pasting it from memory, when copied into memory from another source, such as a TerraBase Inc. database output. Please consult the section on SMILES Notation for important advice.

Example 1: Phenol has the SMILES string “c1ccccc1O”. Copying this string into memory, for example from this text (making sure the quotation marks are omitted), and pasting it into the Input field, will result in the appearance of the chemical structure of phenol in the shaded, rectangular field below, as shown in Figure 3.



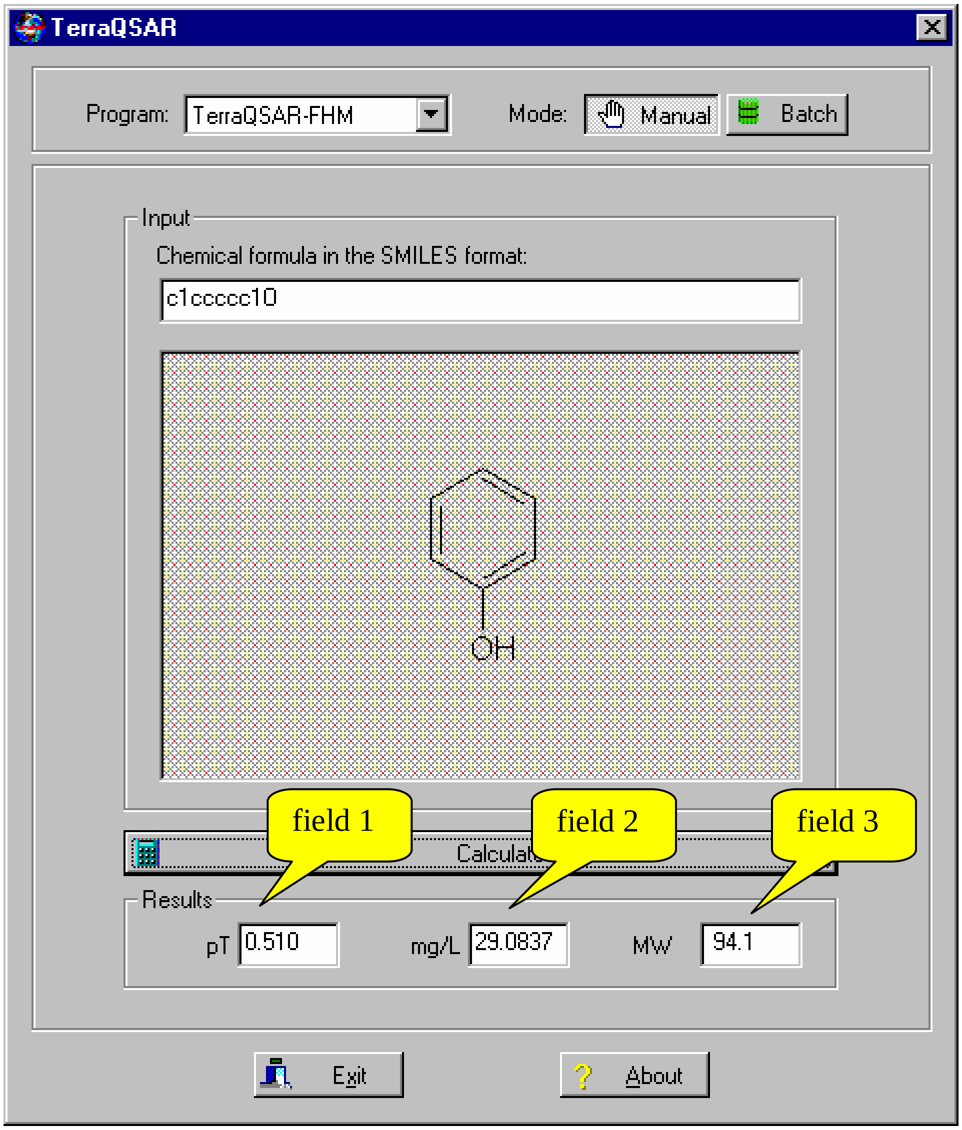
Program: TerraQSAR-FHM

Mode: Manual

Chemical formula in the SMILES format: c1ccccc1O

Fig. 3. Result of entering or pasting the SMILES string for phenol (c1ccccc1O) into the input field: the structure of phenol (excluding hydrogen atoms) will appear in the field below.

Once the user has ascertained that the structure of the compound is that of the desired chemical, a simple click of the Calculate bar below the structure field will result in the three fields below the bar to be filled with the predicted values for the compound, as shown in Figure 4. Field 1 (pT) is the negative logarithm of the LC50 concentration in mmol/L; field 2 (mg/L) is the LC50 value in mg/L; and field 3 (MW) shows the molecular weight of the compound. Computation time varies with the complexity of the structure and the computer specifics, ranging from a nearly instantaneous result (~1 sec) for small structures to minutes for large molecules with highly complex structures.



Program: TerraQSAR-FHM

Mode: Manual

Chemical formula in the SMILES format: c1ccccc1O

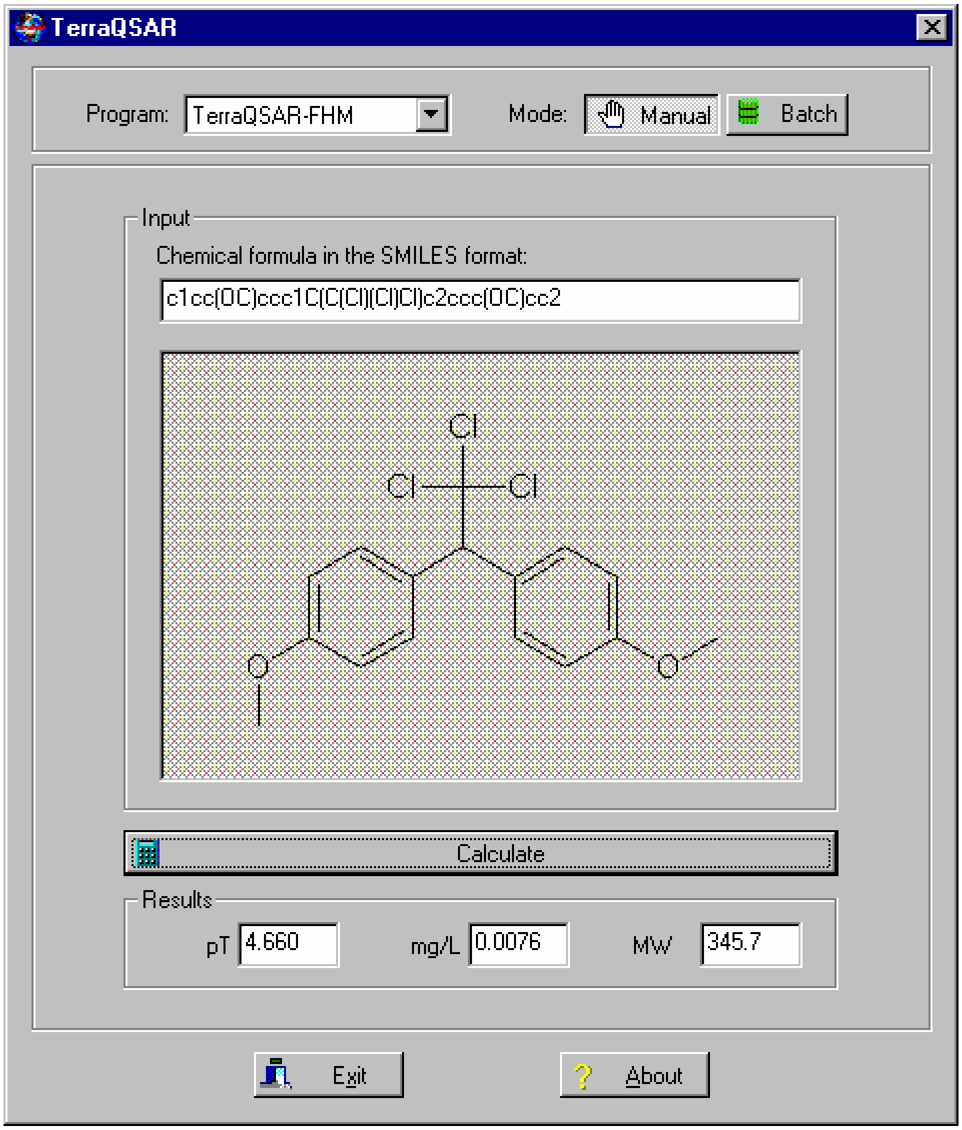
pT 0.510

mg/L 29.0837

MW 94.1

<u>Fig. 4</u>. Result of the execution of “Calculate”. The fields show the following values: Field 1 (pT) is the negative logarithm of the millimolar LC50 concentration, field 2 (mg/L) is the LC50 value in mg/L, and field 3 shows the molecular weight of the compound. Computation time ca. 2 sec at 2 GHz.

Example 2: A more complex molecule, the insecticide p,p'-Methoxychlor, with the SMILES "c1cc(OC)ccc1C(C(Cl)(Cl)Cl)c2ccc(OC)cc2" and structure shown in Figure 5, has a computed LC50 value of 0.0076 mg/L, as shown in Figure 5.



Program: TerraQSAR-FHM

Mode: Manual

Chemical formula in the SMILES format: c1cc(OC)ccc1C(C(Cl)(Cl)Cl)c2ccc(OC)cc2

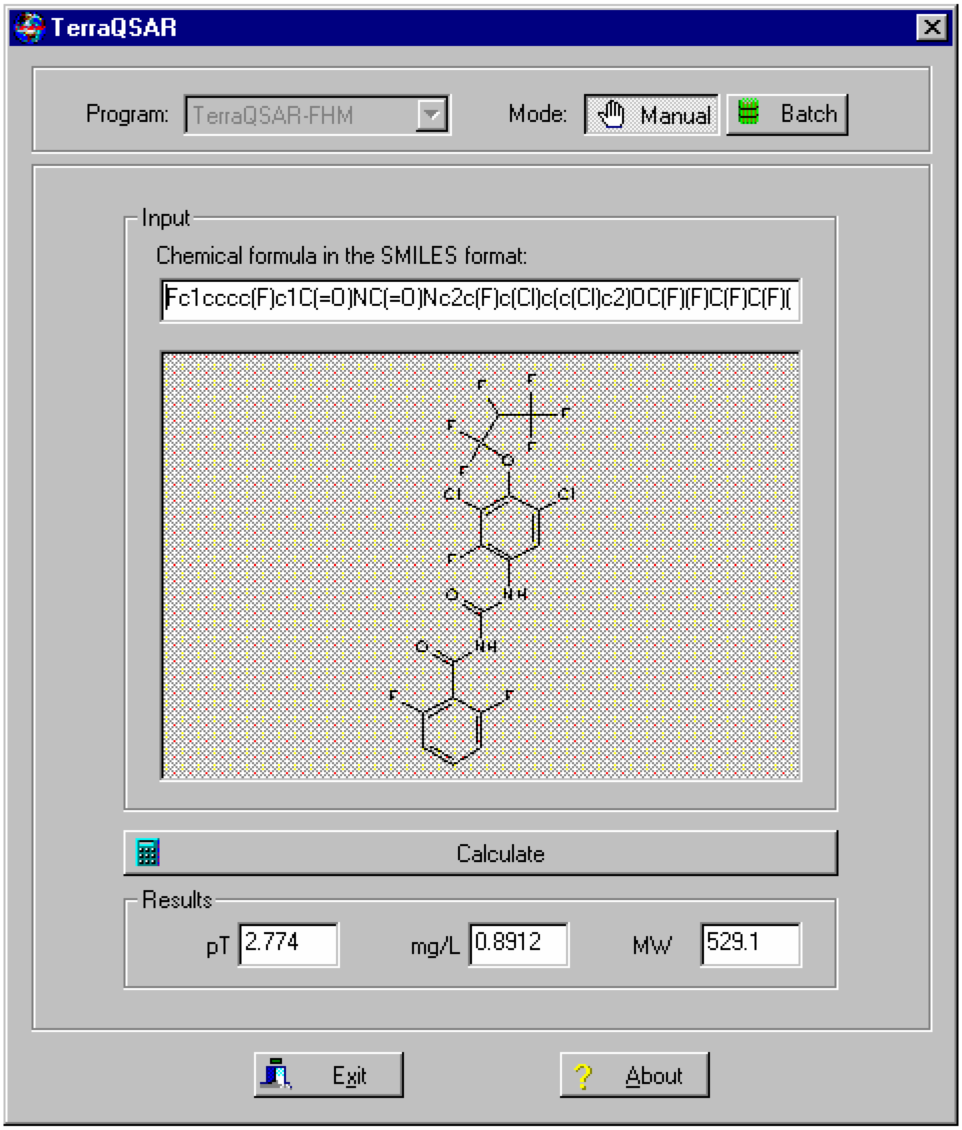
pT 4.660

mg/L 0.0076

MW 345.7

<u>Fig. 5</u>. Result of computation for the insecticide p,p'-Methoxychlor, SMILES string shown in the input field, depicting the compound p,p-Methoxychlor, with its computed 96-hr LC50 value for fathead minnow of 0.0076 mg/L (pT: 4.66).

Example 3: The insecticide Noviflumuron, CAS 121451-02-3, has a predicted fathead minnow 96-hr LC50 of 0.89 mg/L, as shown in Figure 6.



Program: TerraQSAR-FHM

Chemical formula in the SMILES format:

```
Fc1cccc(F)c1C(=O)NC(=O)Nc2c(F)c(Cl)c(c(Cl)c2)OC(F)(F)C(F)C(F)(
```

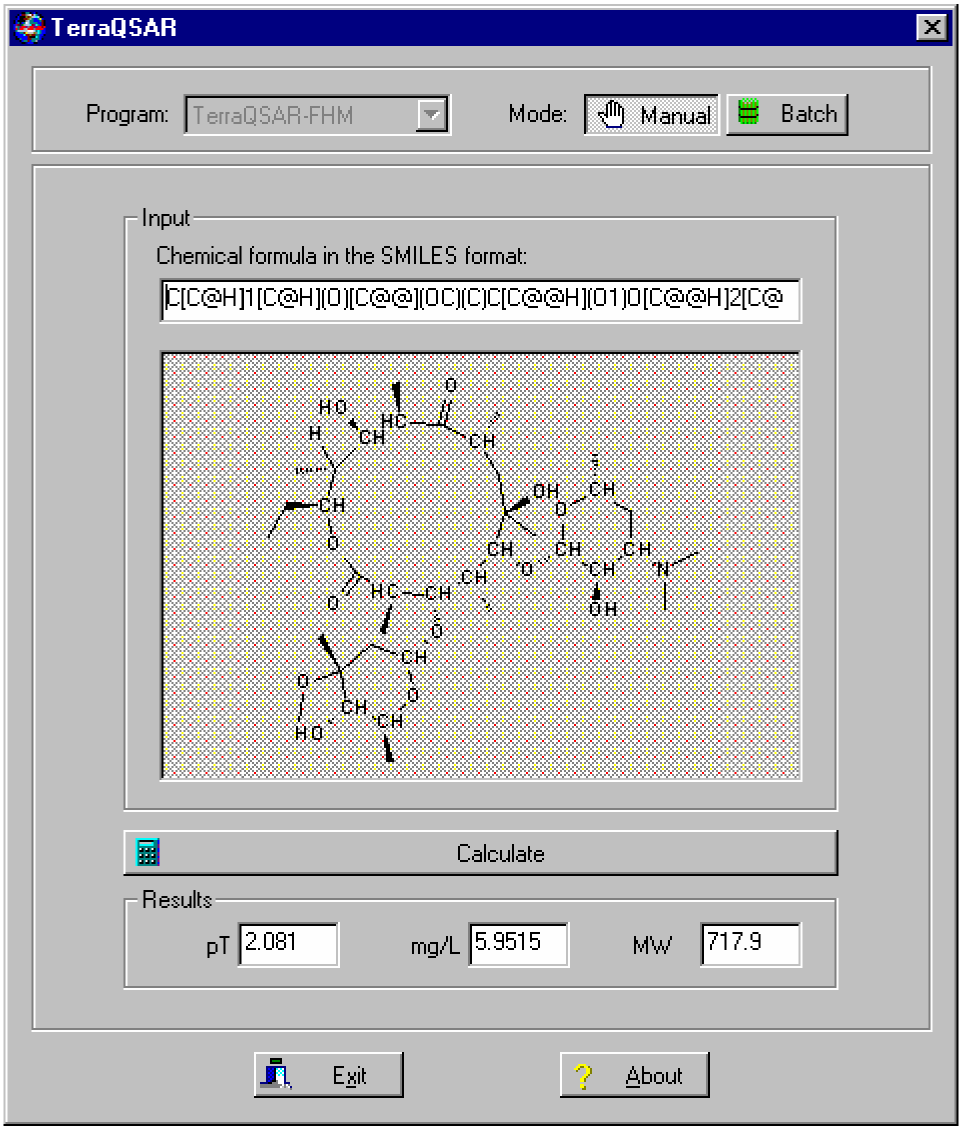
pT 2.774

mg/L 0.8912

MW 529.1

Fig. 6. Result of computation for the SMILES string shown in the input field, depicting the insecticide Noviflumuron, with its computed 96-hr LC50 value for the fathead minnow of 0.89 mg/L.

Example 4: The antibiotic compound Erythromycin B, CAS 527-75-3, with the 3D-SMILES code C[C@H]1[C@H](O)[C@@](OC)(C)C[C@@H](O1)O[C@@H]2[C@@H](C)C(=O)O[C@@H](CC)[C@@](C)([H])[C@@H](O)[C@@H](C)C(=O)[C@H](C)C[C@](O)(C)[C@@H]([C@@H]2C)O[C@@H]3O[C@H](C)C[C@H](N(C)C)[C@H]3O (only a part of the SMILES code is visible in the Input field) has a predicted fathead minnow 96-hr LC50 of 5.95 mg/L, as shown in Figure 7; computation time ca. 50 sec at 2 GHz.



Program: TerraQSAR-FHM

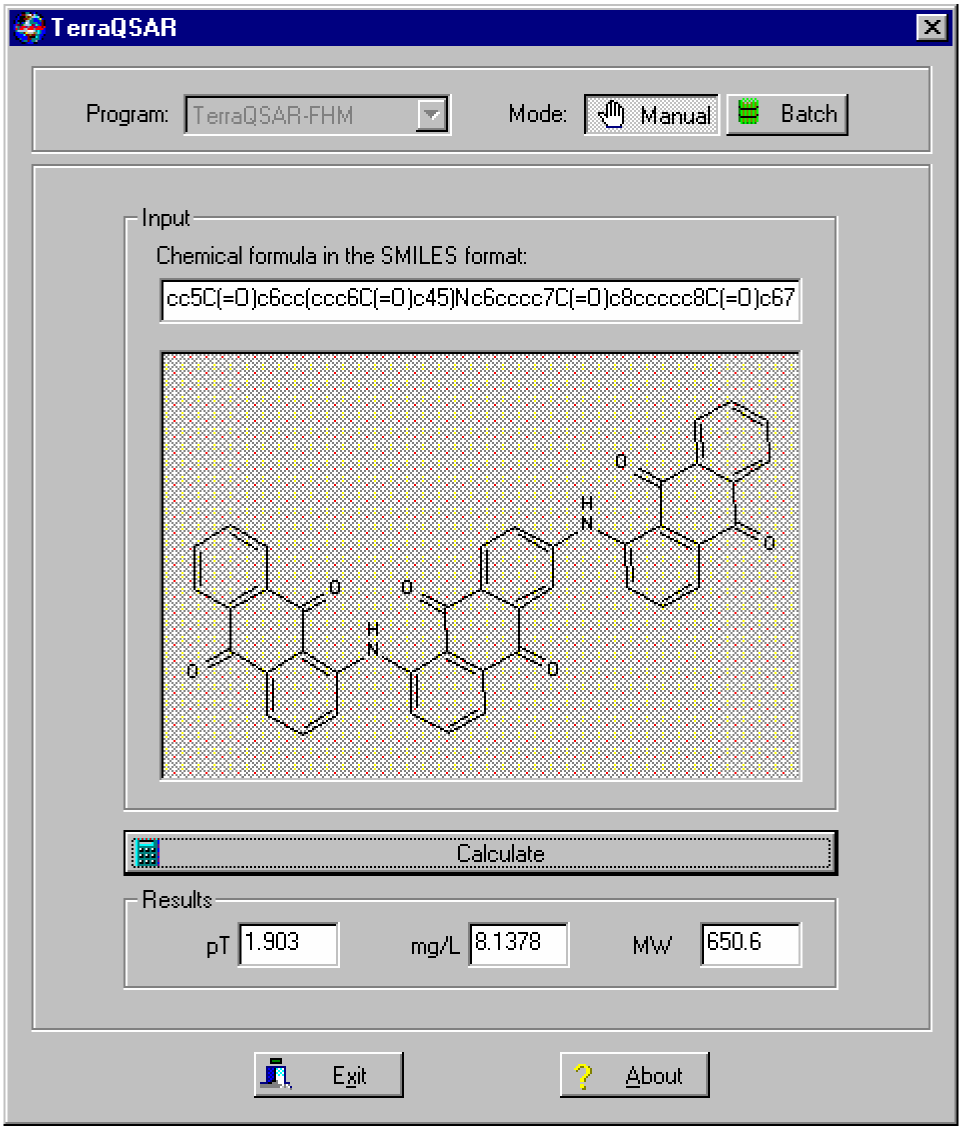
Mode: Manual

Chemical formula in the SMILES format: C[C@H]1[C@H](O)[C@@](OC)(C)C[C@@H](O1)O[C@@H]2[C@

pT 2.081 mg/L 5.9515 MW 717.9

Fig. 7. Result of computation for the SMILES string shown in the input field, depicting the antibiotic Erythromycin B, with the computed 96-hr LC50 value for the fathead minnow of 5.95 mg/L.

Example 5:
An anthraquinone dye, CAS 4478-06-2, with the molecular formula C42H22N2O6, and the SMILES code:
c1cccc2C(=O)c3c(cccc3C(=O)c12)Nc4cccc5C(=O)c6cc(ccc6C(=O)c45)Nc6cccc7C(=O)c8ccccc8C(=O)c67 (only a part of the SMILES code is visible in the Input field) and the structure shown, has the predicted fathead minnow 96-hr LC50 of 8.1 mg/L, as shown in Fig. 8.



Program: TerraQSAR-FHM

Mode: Manual Batch

Input

Chemical formula in the SMILES format:

cc5C(=O)c6cc(ccc6C(=O)c45)Nc6cccc7C(=O)c8ccccc8C(=O)c67

Results

pT 1.903 mg/L 8.1378 MW 650.6

<u>Fig. 8</u>. Result of computation for the SMILES string shown in the input field, depicting an anthraquinone dye, with the computed 96-hr LC50 value for the fathead minnow of 8.1 mg/L.

Technical Requirements

Operating system: PC with Windows 95, 98, NT, 2000, or ME, XP (SP2), or Vista, operating system.
Central processor unit (CPU): No specific requirement, duration of computations will increase with decrease in CPU speed; 2.0 GHz or higher recommended.
Mouse or other pointing device: required.
Screen setting: Variable, 800 x 640, or higher.
CD-ROM drive: required.
Other: Presence of the **TerraQSAR** CD in the CD-ROM drive is required for program execution.

Installation Instructions

The **TerraQSAR** software is a fully functional, stand-alone system, easy to install or uninstall. It consists of two installation parts; we recommend installation in the order mentioned.

Part 1. Installation of the Accelrys software

To install/un-install the Accelrys supporting software, necessary for the proper functioning of the **TerraQSAR** program, use the Windows – Control – Install/Remove Software command to run the “setup.exe” in the Accelrys folder and follow the instructions.

Part 2. Installation of TerraQSAR

To install/un-install **TerraQSAR**, use the Windows – Control – Install/Remove Software command to run the “setup.exe” in the TerraQSAR folder and follow the instructions.

Customer Support

TerraBase Inc. is committed to effective customer support. With the rapid change in PC technology, operating systems and other software and hardware changes, the occasional hiccup is bound to happen. We will try our best to help customers with problems related to our products, in most cases free of charge. Contact our help department with any question and concern about our products, either by EMAIL, FAX, or MAIL.

TerraBase Inc.
1063 King St. West, Suite 130
Hamilton, ON, L8S 4S3, Canada
Fax: 905-527-0263
Internet: http://www.terrabase-inc.com/