(Q)SAR APPLICATION TOOLBOX

VERSION 1.1

Guidance on Importing Databases

JANUARY 2009

TABLE OF CONTENTS

1. INTRODUCTION

While the Toolbox is installed with a number of databases of experimental results, a user friendly import function for adding other experimental results is available for the Toolbox. These new data can be either for endpoints already defined by in the Toolbox or for endpoints not yet defined in the Toolbox.

The format of a database to import can be in a simple spreadsheet format or tab-delineated text.

Based on the nature of the data to be imported one or the other of two protocols must be followed; either vertical or horizontal importing.

Vertical importing is used when for each result for a given endpoint the same metadata[1] (i.e., author, test conditions, protocol, source, etc.) applies. An example of a vertical import would be for results obtained with the same assay in the same laboratory. The file can contain more than one column of results data, but for each column (i.e., specific endpoint) the metadata is identical.

Horizontal importing is used when the results column has mixed data. In this case the metadata may be different for different rows. For example the column of results data may contain a mixture of *Daphnia magna* 48-h EC50 values in mg/l, *Daphnia magna* 24-h EC50 values in mM/l and *Scenedesmus subspicatus* 72-h algal growth inhibition EC50 values in M/l. Clearly, the protocol and dimensions are different in the three cases thus the metadata will be different.

In the following guidance, examples for importing horizontal and vertical databases are outlined. Furthermore, a separate chapter on tips and tricks as well as trouble shooting is added.

> NOTE: Since the importer of a new database is responsible for quality assurance of chemical identification information both a valid CAS number and a valid SMILES notation is needed for each chemical in the database to be imported.

Please also note that while the screen shots provided in this document were taken from OECD Toolbox 1.0 or OECD Toolbox 1.1 (BETA) they apply to OECD Toolbox 1.1.

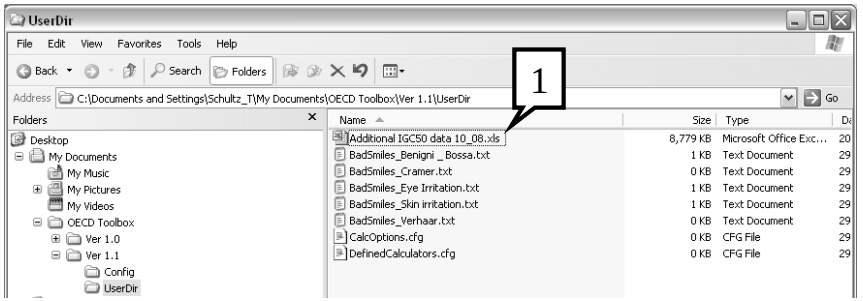
[1] [1]Metadata is the information that informs the Toolbox user about the conditions under which the endpoint results were obtained as well as any references

2. VERTICAL IMPORT

Vertical importing is done when the chemicals are distributed along the rows and each endpoint value is defined by a column. While there may be more than one endpoint for each compound in the database (i.e., more than one column with results), the metadata is the same for an entire column. Therefore, each record in the database is defined by a combination of row and column information. Specifically, the chemical identification information (i.e., the CAS number and SMILES notation) come from the row, and endpoint value(s) and corresponding metadata come from the column(s). For vertical file layout the minimum data for each compound is a set endpoint correspondence for each value column and a valid CAS number and SMILES notation for each chemical (see file layout below).

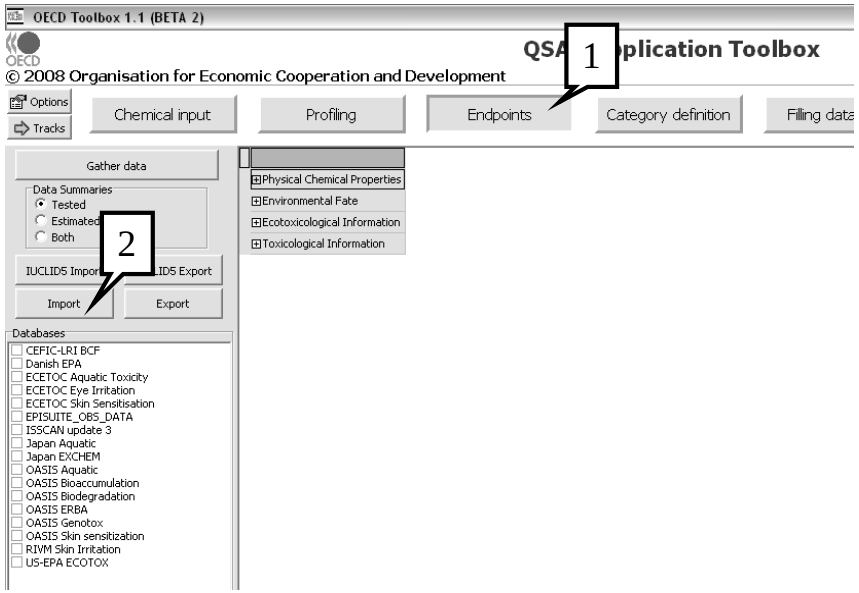
	CAS	NAME	SMILES	VALUE
2	100-06-1	4'-methoxyacetophonone	C(C)(=O)c1ccc(OC)cc1	203.28
3	1003-09-4	2-bromothiophene	C1(Br)=CC=CS1	131.06
4	104-57-4	benzyl formate	c1(COC=O)ccccc1	435.99
5	106-92-3	allyl glycidyl ether	C(=C)COCC1CO1	383.92
6	107-08-4	1-iodopropane	C(I)CC	417.43
7	107-11-9	allyl amine	C(=C)CN	383.05
8	1073-05-8	1,3-propanediol cyclic sulfate	C1COS(=O)(=O)O1	51.21
9	108-91-8	cyclohexylamine	C1(N)CCCCC1	366.07
10	109-66-0	n-pentane	C(C)CCC	113.26
11	109-69-3	1-chlorobutane	C(Cl)CCC	745.56
12	110-54-3	n-hexane	C(C)CCCC	16.28
13	1120-06-5	2-decanol	C(C)(O)CCCCCCCC	16.14
14	1154-59-2	3,3',4',5-tetrachlorosalyicylanilide	C(=O)(c1c(O)c(Cl)cc(Cl)c1)Nc1cc(Cl)c(Cl)cc1	0.11
15	1192-62-7	2-acetylfuran	C(C)(=O)C1=CC=CO1	1215.76
16	123-96-6	DL-2-octanol	C(C)(O)CCCCCC	136.29
17	13243-65-7	2,3-dibromo-1,4-naphthoquinone	C1(Br)C(=O)c2c(C(=O)C=1Br)cccc2	0.31

An example of a vertical database import is outlined below. The database to be imported is in spreadsheet format and was placed in the folder \My Documents\OECD Toolbox\Ver. 1.1\UserDir. In this example we will be adding the database **Additional IGC50 data 10_08 (1)** (see screenshot below).



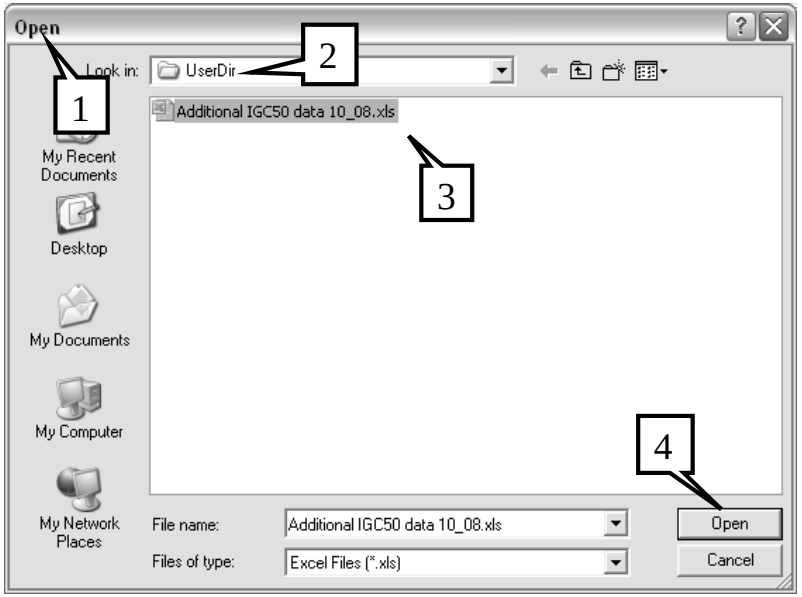
Since the installer of a new database is responsible for quality assurance of chemical identification information both a valid CAS number and a valid SMILES is needed for each chemical in the database.

With this background information, launch the Toolbox and choose the **Flexible Track** among the three tracks. Wait while your computer opens the resident databases and then click on the **Endpoints** button **(1)**. Press the **Import** button **(2)** and follow the import wizard (see screenshot below).

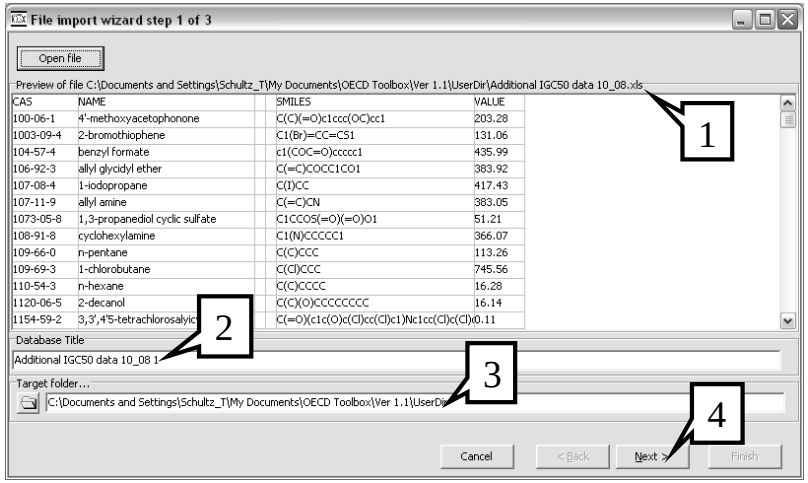


In step one of the wizard, press the **Open file** button **(1)** (see screenshot below).

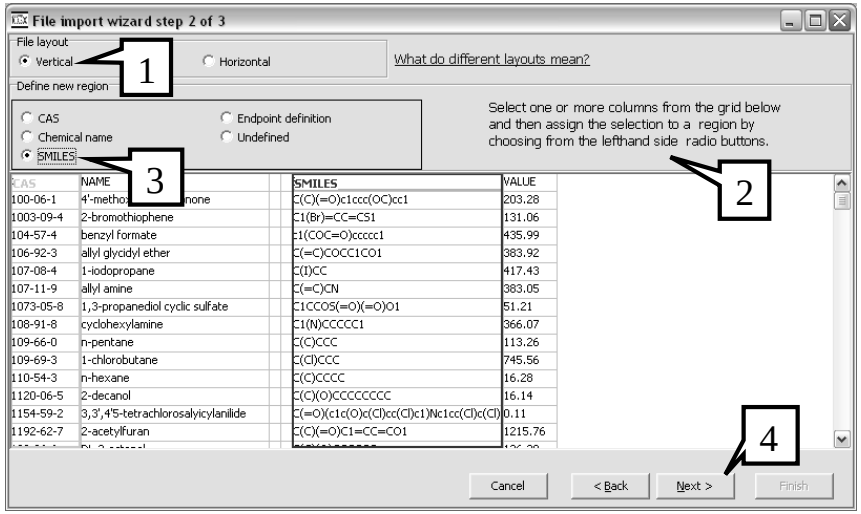
This reveals the window **Open (1)**, which allows you to find the data file you wish to import into the Toolbox. The database we wish to import is stored as a spreadsheet preferably in \My Documents\OECD Toolbox\Ver 1.1\UserDir. Go to **UserDir (2)**, highlight the file of choice (in this example it is **Additional IGC50 data 10_08.xls (3)**. Clicking on **Open (4**) (see screenshot below) reveals another window.



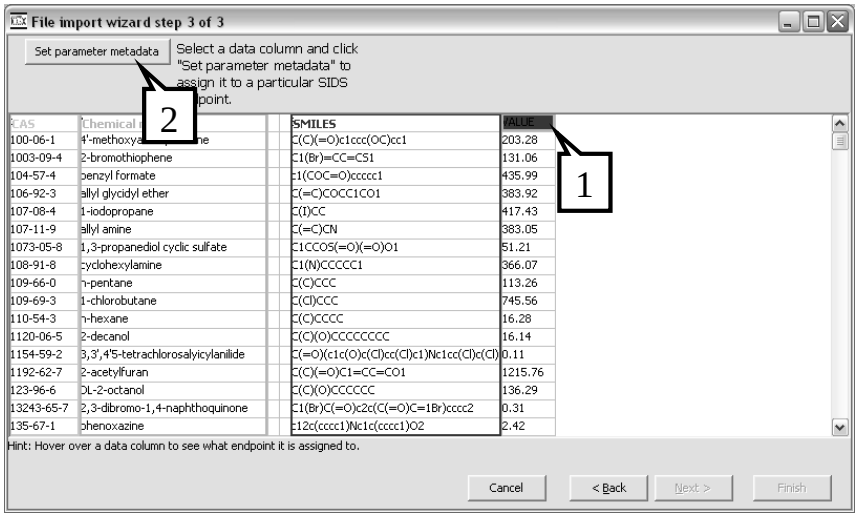
Upon clicking **Open** in the screenshot above the window **File import wizard step 1 of 3** appears (see screenshot below). This window is filled in with the relevant information including **Preview of the file (1)**,which in this case contains 48-hr IGC50 data in mg/l for *Tetrahymena pyriformis*, the **Database title (2)** and the **Target folder…(3).** While the Toolbox does not require headers for the columns of the imported database, it is recommended they be provided. The columns for CAS number, SMILES and NAME (if provided) are obvious from their format, but the endpoint and units for the data columns are not as obvious; you will be asked to provide this information as you map the database to the Toolbox.



Clicking **Next (4)** in the screenshot above brings the window for step 2 into view. **File import wizard step 2 of 3** is viewed in the screenshot below. In a **Vertical (1)** file layout, each column (noted in a color outline) is assigned a definition **(2)**. In this screenshot the **red outlined SMILES** is linked to **SMILES (3)** in **Define new region**. Similarly, CAS is linked to CAS, NAME is linked to Chemical name, VALUE is linked to Undefined. Once all columns have been assigned (even if several are listed as Undefined) click on the **Next** button **(4)**.

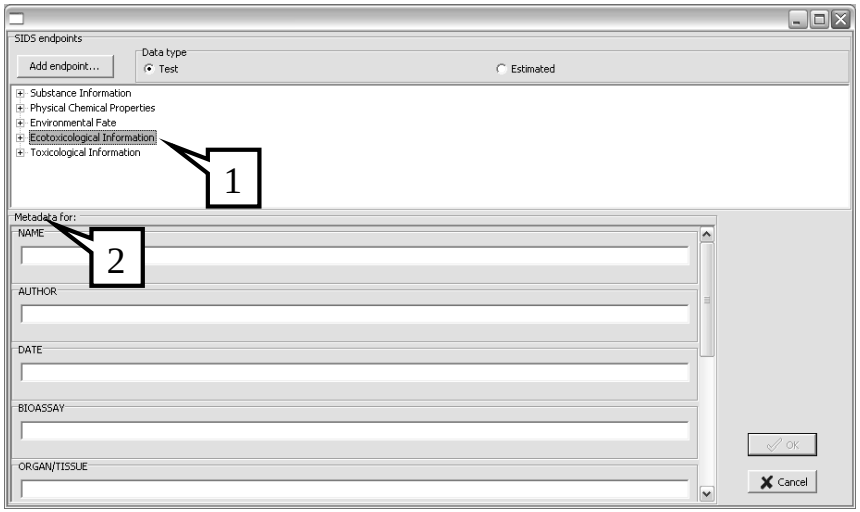


This brings into view the **File import wizard step 3 of 3** window (see screenshot below).



Since the data being imported is in a vertical layout the metadata for each endpoint result in a particular column is identical. Thus, the metadata only has to be entered once. In this example the data column with the red background, VALUE needs to be mapped to the Toolbox by setting its metadata. After clicking on **VALUE (1)** click the **Set parameter metadata** box **(2)**.

This brings into view the window **SIDS endpoints** (see screenshot below).



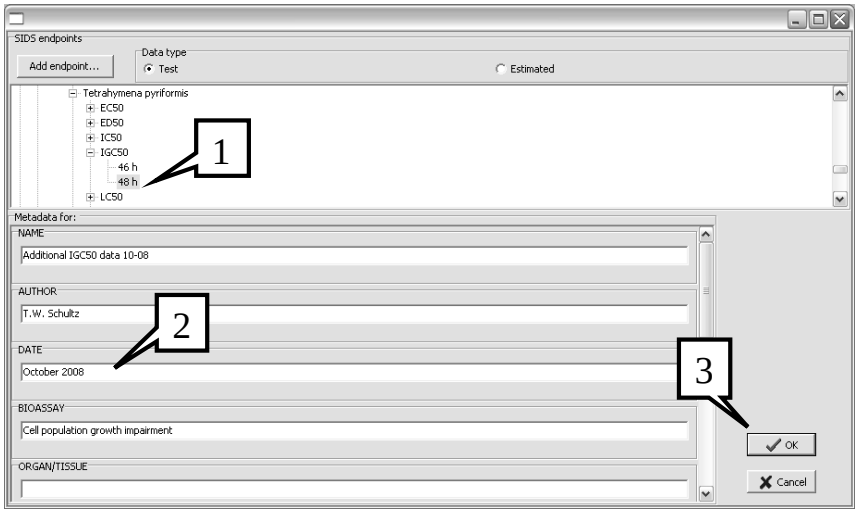
In this example, we are adding data for an endpoint (*Tetrahymena pyriformis* 48-hr IGC50), which is already in the data matrix under **Ecotoxicological Information (1)**. By clicking on Ecotoxicological Information, Aquatic Toxicity, Protozoa etc. the nodes of the data tree appear.

In the Vertical import **Metadata (2)** include the following:

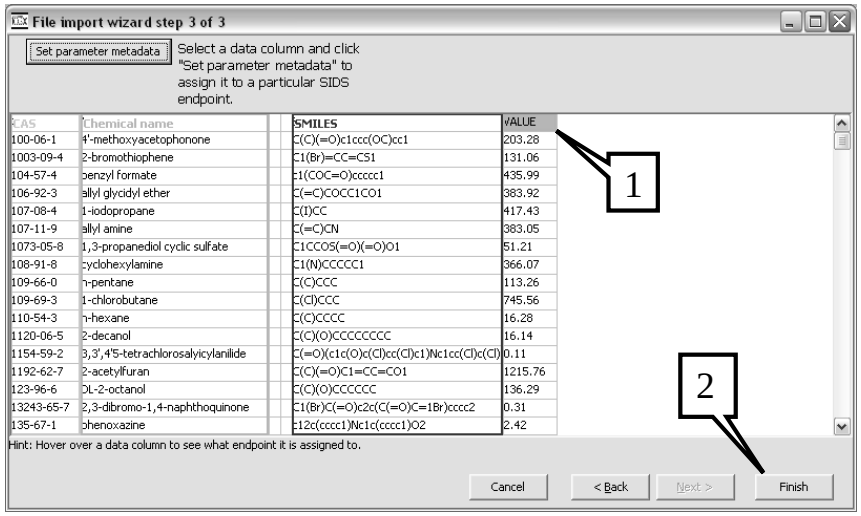
NAME	the name of the database
AUTHOR	the compiler of database
DATE	the date the database was complied
BIOASSAY	Information on the bioassay used to generate the endpoint results
ORGAN/TISSUE	target organ or tissue for the endpoint
CONDITIONS	information on test conditions
ADMINISTRATION	the route by which the chemical was administered
SOURCE	Reference for the endpoint results
DIMENSIONS	the units of endpoint values
COMMENT	a field where additional information may be noted

Of these ten fields DIMENSIONS is critical to mapping the database to the Toolbox. This is because the Toolbox recognizes DIMENSIONS only when expresses in particular forms (see Annex 1). For example, the Toolbox recognizes mg/l but not ppm

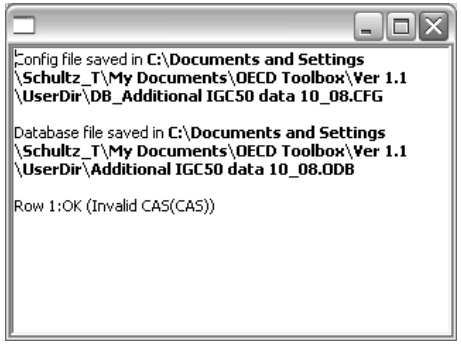
Under Protozoa click **Tetrahymena pyriformis**, **IGC50**, and **48 h (1**). One can enter the metadata in the 10 boxes under **Metadata for:** In **(2**) we observe **October 2008** has been entered in the **DATE** box as an example of metadata entry. Click the **OK** button **(3)** returns the **File import wizard step 3 of 3** screen.



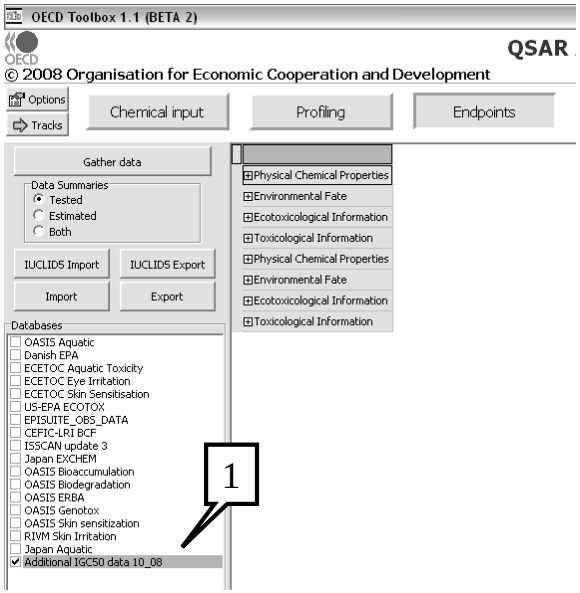
In viewing the **File import wizard step 3 of 3** window this time the **VALUE** box **(1)** has a green background noting it is ready to be mapped to the Toolbox. Clicking the **Finish** button **(2)** starts the configuring and loading of the database into the Toolbox.



After some time, while the Toolbox configures and saves the new database, a window appears confirming that the database has been configured and saves (see screenshot below).

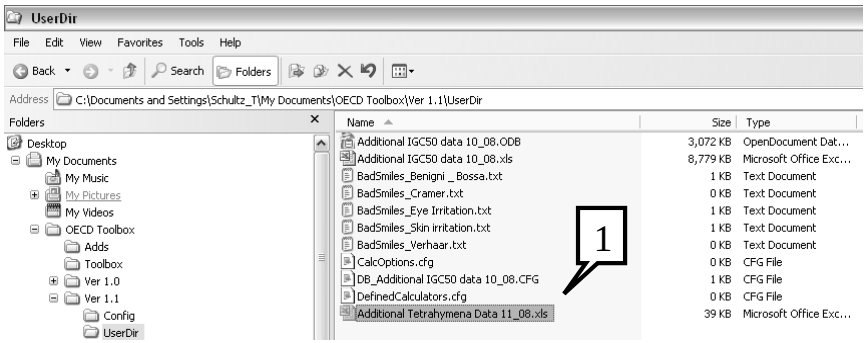


Upon closing the above window the Toolbox formats the new data matrix before returning to the Toolbox workflow screen where the **Additional IGC50 data 10_08** appears in the Databases box **(1)**.



3. HORIZONTAL IMPORT

As with a vertical import, the database to be imported should be in spreadsheet format and placed in a folder on the hard drive, preferably in the folder \My Documents\OECD Toolbox\Ver. 1.1\UserDir. In this example we will be adding the database **Additional Tetrahymena Data 11_08 (1)** (see screenshot below).

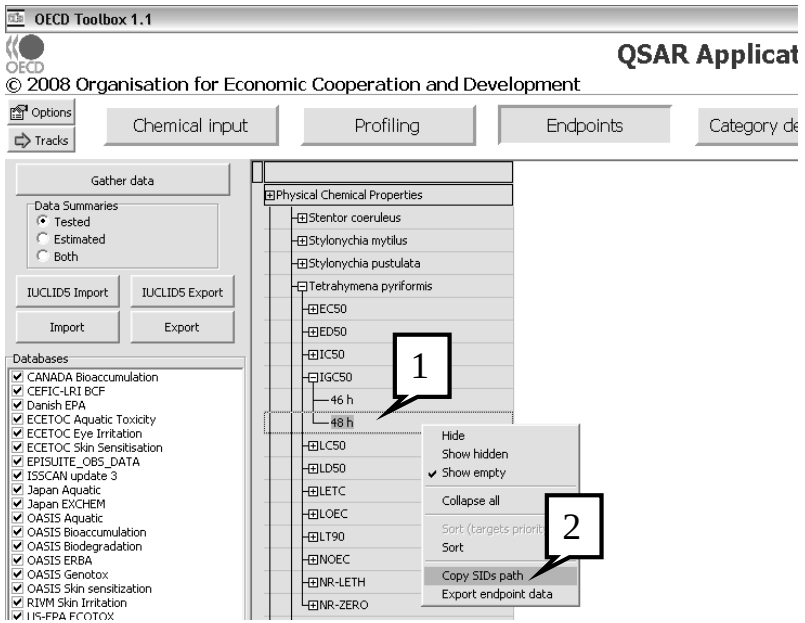


With horizontal importing each record is defined by a single row. In each row there is chemical identification information (at minimum a CAS number and a SMILES notation), endpoint values and metadata. The metadata includes the information as to which leaf in the data-tree of the Toolbox each endpoint value is to be attributed (SIDS path). There are two variants for this attribution. In the single column variant, the data nodes along the SIDS path are delimited by the # symbol (e.g., Ecotoxicological Information#Aquatic Toxicity#Protozoa etc). In the multi-column variant there is a column for each node along the SIDS path (e.g., column 1 = Ecotoxicological Information, column 2 = Aquatic Toxicity, column 3 = Protozoa, etc. Remember a CAS number and valid SMILES are required for each chemical). In this example the database **Additional Tetrahymena Data 11_08**, is in a single column variant layout as the nodes along the **SIDS path in column D (1)** are separated by # symbols (see screenshot below).

A	B	C	D	E	F	G
14920-89-9	2,3-dimethylfuran	c1(C)ccoc1C	Ecotoxicological Information#Aquatic	309.07	mg/l	static 48-hour
3208-16-0	2-ethylfuran	c1ccoc1CC	Ecotoxicological Information#Aquatic	319.98	mg/l	static 48-hour
78-59-1	3,5,5-trimethyl-2-cyclohexen-1-one	CC1=CC(=O)CC(C)(C)C1	Ecotoxicological Information#Aquatic	513.87	mg/l	static 48-hour
583-60-8	2-methyl cyclohexanone	O=C1C(C)CCCC1	Ecotoxicological Information#Aquatic	723.17	mg/l	static 48-hour
1120-75-5	2-methyl cyclopentanone	O=C1C(C)CCC1	Ecotoxicological Information#Aquatic	2.12	mg/l	static 48-hour
2758-18-1	3-methyl-2-cyclopenten-1-one	O=C1C=C(C)CC1	Ecotoxicological Information#Aquatic	23.1	mg/l	static 48-hour
79-39-0	methacrylamide	C=C(C)C(=O)N	Ecotoxicological Information#Aquatic	32.2	mg/l	static 48-hour
623-48-3	ethyl iodoacetate	ICC(=O)OCC	Ecotoxicological Information#Aquatic	0.012	mM	static 2-day
70-11-1	2-bromoacetophenone	c1ccccc1C(=O)CBr	Ecotoxicological Information#Aquatic	0.0015	mM	static 2-day
816-40-0	1-bromo-2-butanone	CCC(=O)CBr	Ecotoxicological Information#Aquatic	0.0025	mM	static 2-day
624-75-9	iodoacetonitrile	N#CCI	Ecotoxicological Information#Aquatic	0.0034	mM	static 2-day
144-48-9	2-iodoacetamide	NC(=O)CI	Ecotoxicological Information#Aquatic	0.0142	mM	static 2-day
814-75-5	3-bromo-2-butanone	CC(=O)C(Br)C	Ecotoxicological Information#Aquatic	0.0142	mM	static 2-day
78-95-5	chloroacetone	ClCC(=O)C	Ecotoxicological Information#Aquatic	0.0206	mM	static 2-day
7205-91-6	2-chloromethyl phenyl sulfide	c1ccccc1SCCl	Ecotoxicological Information#Aquatic	0.0214	mM	static 2-day
32807-28-6	methyl 4-chloroacetoacetate	C(Cl)C(=O)CC(=O)OC	Ecotoxicological Information#Aquatic	0.0475	mM	static 2-day

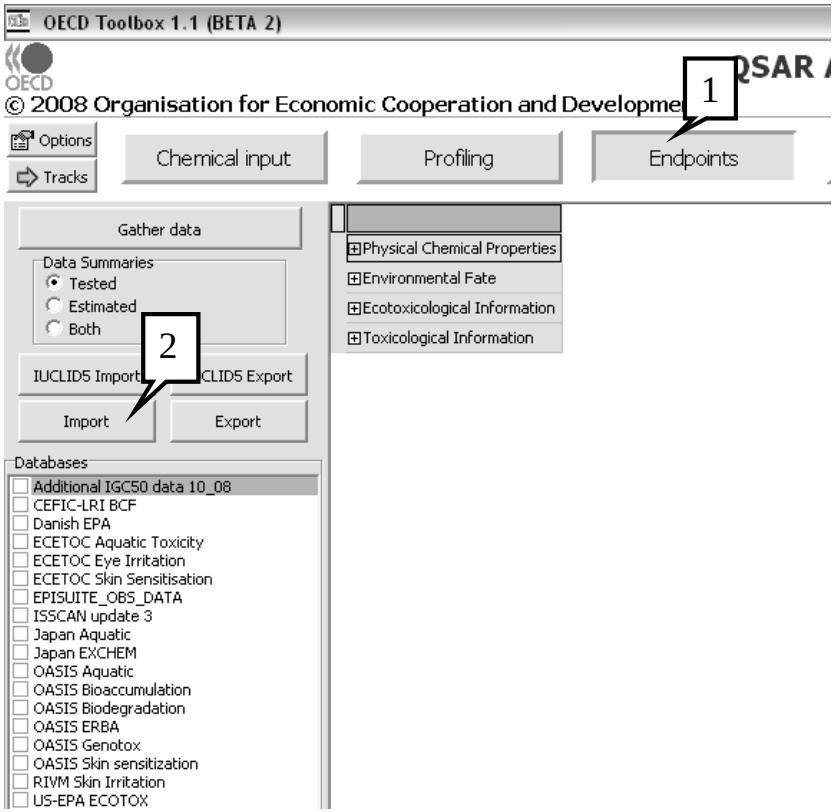
1

TIP: The SIDS path can be quite long and small mistakes can block the import. To make sure the correct SIDS path is listed in the file, it can be copied from the SIDS tree in the Toolbox by right-clicking on the leaf of the SIDS tree **(1)** and then selecting **Copy SIDS path** from the drop-down menu **(2)** (see screenshot below).

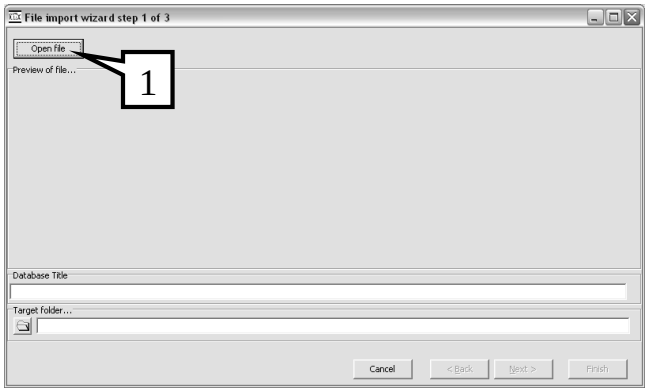


Since the installer of a new database is responsible for quality assurance of chemical identification information both a valid CAS number (column A) and a valid SMILES (column C) is needed for each chemical in the database.

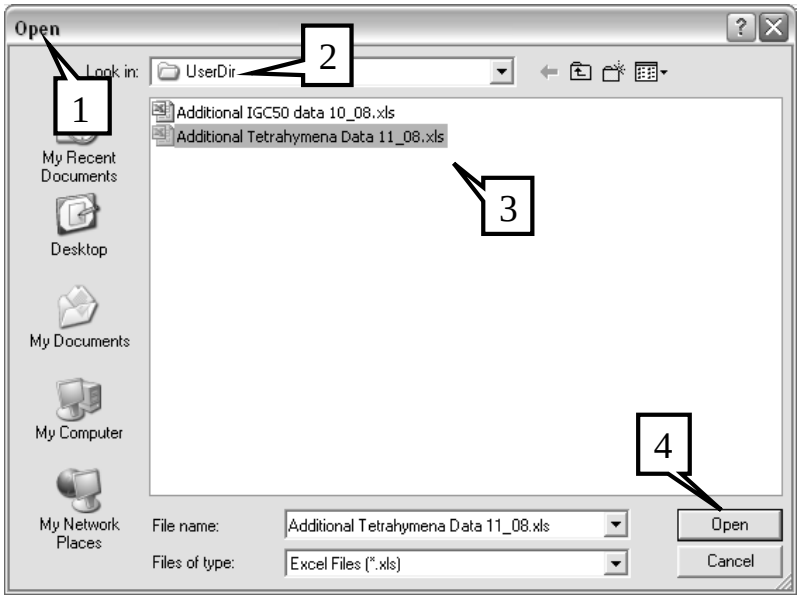
After launching the Toolbox in the Flexible Track, click the **Endpoints** button **(1)** and then press the **Import** button **(2)** and follow the import wizard. Note that database imported in the first exercise Additional IGC50 data 10-08 is listed in the Databases box (see screenshot below).



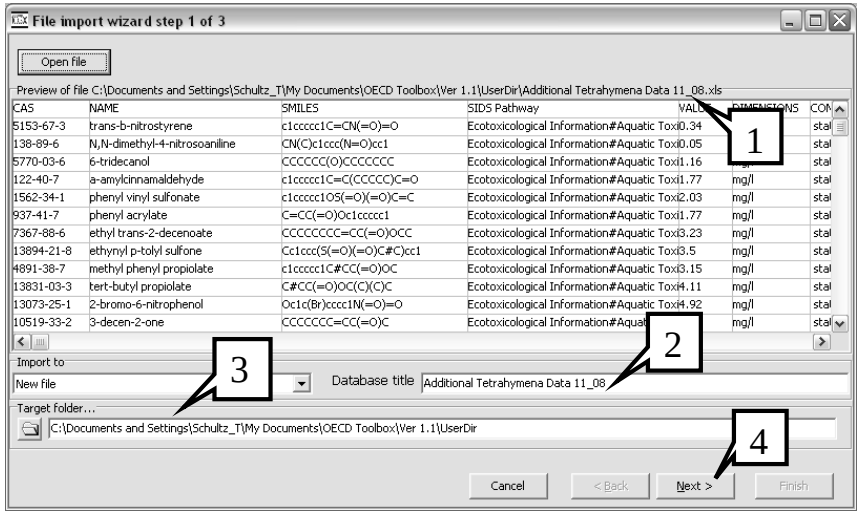
In step one of the wizard, press the **Open file** button **(1)** (see screenshot below).



This reveals the window **Open (1)**, which allows you to find the data file you wish to import into the Toolbox. Since the database we wish to import was stored as a spreadsheet in \My Documents\OECD Toolbox\Ver 1.1\UserDir, go to **UserDir (3)**, highlight the file of choice (in this example it is **Additional Tetrahymena Data 11_08.xls (3).** Clicking on **Open (4**) (see screenshot below) reveals another window.



Clicking **Open** in the screenshot above reveals the window **File import wizard step 1 of 3** appears (see screenshot below). This window is filled in with the relevant information including **Preview of the file (1)**,which in this case contains *Tetrahymena pyriformis* 48-hr IGC50 data, the **Database title (2)** and the **Target folder…(3).**

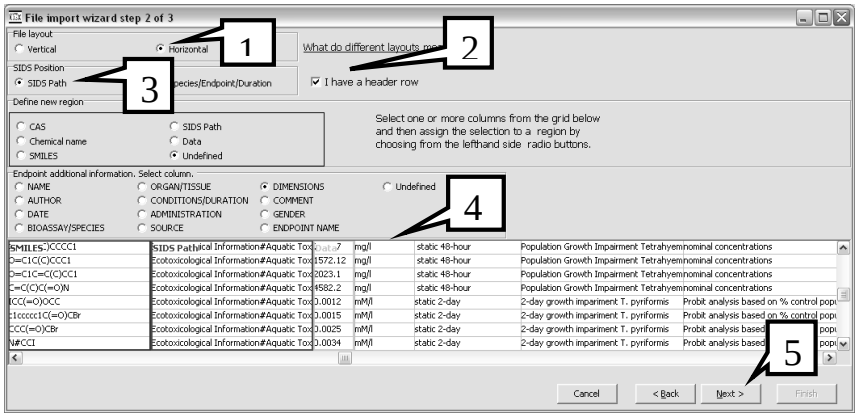


In the horizontal import metadata includes the following:

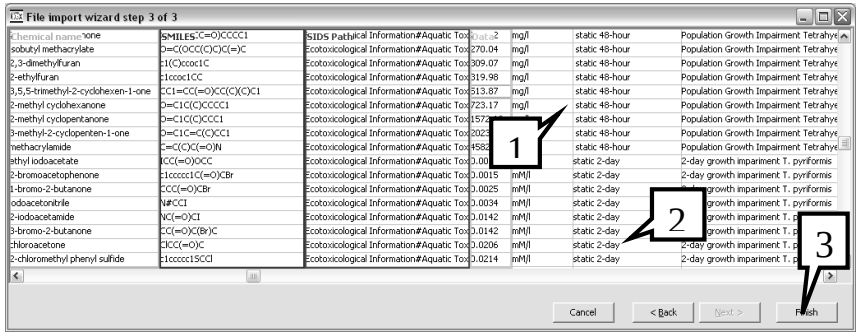
NAME	the name of the database
AUTHOR	the compiler of database
DATE	the date the database was complied
BIOASSAY	information on the bioassay used to generate the endpoint results
ORGAN/TISSUE	target organ or tissue for the endpoint
CONDITIONS	information on test conditions
ADMINISTRATION	the route by which the chemical was administered
SOURCE	reference for the endpoint results
DIMENSIONS	the units of endpoint values
COMMENT	a field where additional information may be noted
GENDER	the gender of the species tested
ENDPOINT NAME	the name of the endpoint

Of these twelve defined fields DIMENSIONS is critical to mapping the database to the Toolbox. This is because the Toolbox recognizes DIMENSIONS only when expresses in particular forms (see annex 1). For example, the Toolbox recognizes mg/l but not ppm. Clicking on **Next (4)** in the screenshot above brings the window for step 2 into view.

File import wizard step 2 of 3 is viewed in the screenshot below. Click the **Horizontal** button **(1)** Since our database has a header row this box is checked **(2)**. Click the **SIDS Path** radio button **(3)** in the **SIDS Position** box. In this file layout, each column noted in a color outline is assigned a defined radio button in the **Define new region** box. In this screenshot we observe the red outlines **SMILES**, purple outlines **SIDS path**, and aqua outlines **Data**. Similarly but not shown, CAS is linked to CAS and NAME is linked to Chemical name. Since the import is a horizontal one, all other columns in the database must be linked to a metadata field in the **Endpoint additional information** box. In this window, column **(4)** is linked to the radio button **DIMENSIONS** in the **Endpoint additional information** box. Since **DIMINSIONS** is not defined in the **Define new region** box, the **Undefined** radio button is on in the **Define new region** box. Note that dimensions in column 4 vary some are in mg/l while others are reported in mM/l. If the column does not fit one of the titled fields in the **Endpoint additional information** box then the column is linked to the **Undefined** button in the **Endpoint additional information box**. Once all columns have been assigned (even if several are listed as **Undefined**) click on the **Next** button **(5)**.

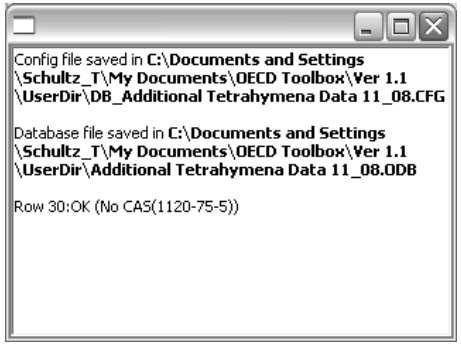


This brings into view the **File import wizard step 3 of 3** window (see screenshot below).

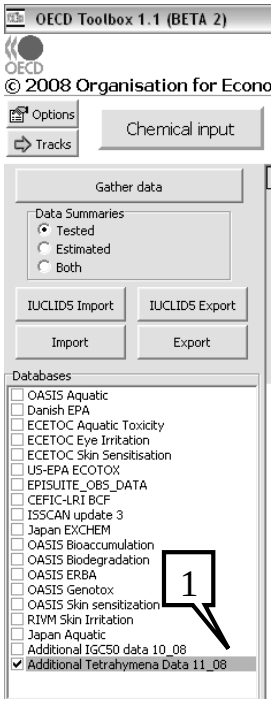


Remember since the data being imported is in a horizontal layout the metadata must be entered for each row. In this example the **CONDITIONS** column is reported as either static 48-hr **(1)** or static 2-day **(2)**. Clicking the **Finish** button **(3)** brings up the next window.

After some time while the Toolbox configures and saves the new database window appears confirming that the database has been configured and saves (see screenshot below).

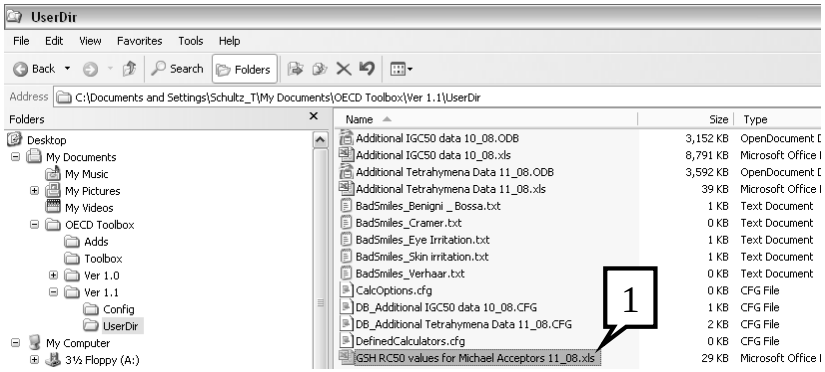


Upon closing this window the Toolbox formats the new data matrix before returning to the Toolbox workflow screen where the **Additional Tetrahymena Data 11_08** appears in the Databases box **(1)**.



4. IMPORTING A DATABASE WITH A NEW ENDPOINT

While the Toolbox is installed with a number of endpoints mapped to it, the Toolbox allows for adding experimental results for endpoints not yet defined in the Toolbox. As in other cases, the database to be imported should be in spreadsheet format and placed in a folder on the hard drive, preferably in the folder \My Documents\OECD Toolbox\Ver. 1.1\UserDir. In this example we will be adding the database **GSH RC50 values for Michael Acceptors 11_08.xls (1)** which contains 2-hr RC50 data in mM/l for glutathione reactivity (see screenshot below).



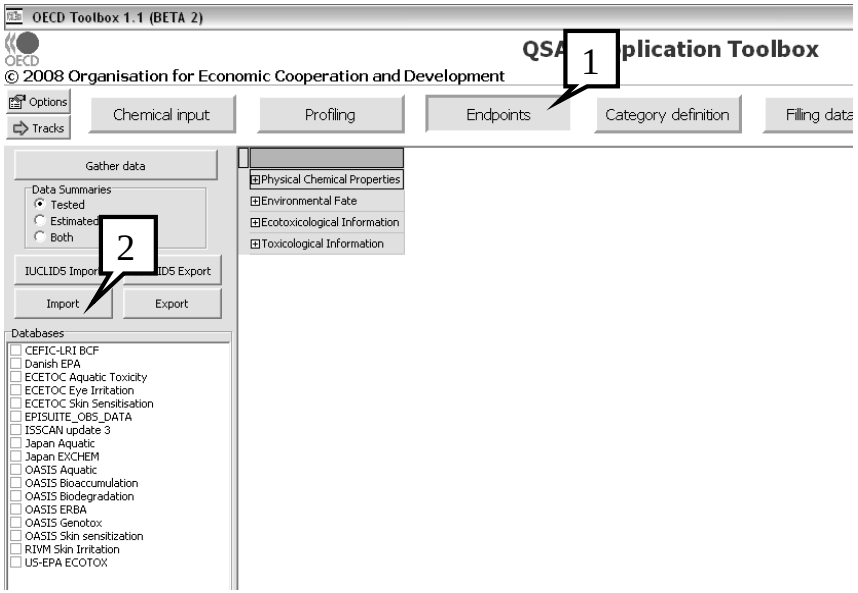
Since the installer of a new database is responsible for quality assurance of chemical identification information both a valid CAS number and a valid SMILES is needed for each chemical in the database.

As noted before, based on the nature of the data being imported importing can be vertical or horizontal. A vertical import is performed when the data is homogenous (i.e., all the data in each results column, whether there is one or more results column, has the same metadata). Remember, metadata is the information that informs the Toolbox user about the database. In contrast, a horizontal import is performed when the data is not homogenous; that is the data in a results column does not have the same metadata.

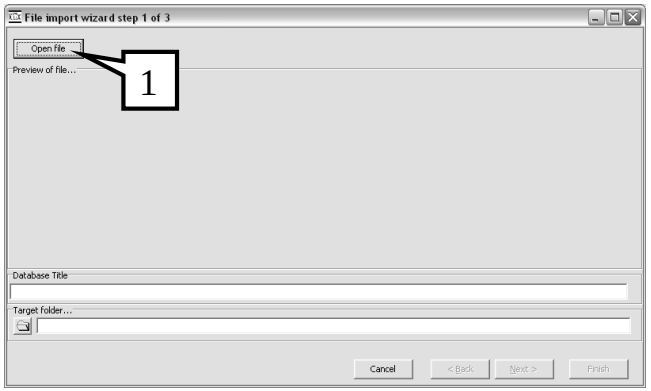
In this example we have homogenous data (all results were obtained with the same protocol and test conditions) and vertical importing is made. Note this database has the chemical identification information (i.e., CAS, NAME and SMILES) distributed along the rows and the endpoint VALUE reported in column D (see screenshot below). Remember the metadata is the same for the entire column. While the Toolbox does not require headers for the columns of the imported database, it is recommended they be provided. The columns for CAS number, SMILES and Name (if provided) are obvious from their format, but the endpoint and units for the data columns are not as obvious; you will be asked to provide this information as you map the database to the Toolbox.

	A	B	C	D
1	**Chemical**	**CAS**	**SMILES**	**RC50 (mM/l)**
2	1,2-Naphthoquinone	524-42-5	O=C(c(c(C=C1)ccc2)c2)C1=O	0.022
3	Phenylpropargyl aldehyde	2579-22-8	c1ccccc1C#CC=O	0.025
4	1-Nitro-1-cyclohexene	2562-37-0	C1CCCC=C1N(=O)=O	0.028
5	1,4-Benzoquinone	106-51-4	O=C1C=CC(=O)C=C1	0.046
6	Ethynyl-p-tolyl sulfone	13894-21-8	Cc1ccc(S(=O)(=O)C#C)cc1	0.060
7	Phenyl vinyl sulfonate	1562-34-1	c1ccccc1OS(=O)(=O)C=C	0.081
8	Acrolein	107-02-8	C=CC=O	0.086
9	Methyl propiolate	922-67-8	C#CC(=O)OC	0.12
10	3-Hexyn-2-one	1679-36-3	CC(=O)C#CCC	0.12
11	Phenyl vinyl sulfone	5535-48-8	c1ccccc1S(=O)(=O)C=C	0.14
12	Diethyl fumarate	623-91-6	CCOC(=O)C=CC(=O)OCC	0.20
13	2-Cyclohexen-1-one	930-68-7	O=C1C=CCCC1	0.37
14	Methyl acrylate	96-33-3	C=CC(=O)OC	0.45
15	2-Vinyl pyridine	100-69-6	n1c(C=C)cccc1	2.8
16	4-Ethynylpyridine	352530-29-1	n1ccc(C#C)cc1	4.5
17	Phenyl vinyl sulfoxide	20451-53-0	c1ccccc1S(=O)C=C	17
18	2-Methyleneglutaronitrile	1572-52-7	N#CCCCC(=C)C#N	21
19	2-Hydroxyethyl methacrylate	868-77-9	C=CC(C)C(=O)OCCO	33
20	*trans*-2-Pentenoic acid	13991-37-2	OC(=O)C=CCC	37

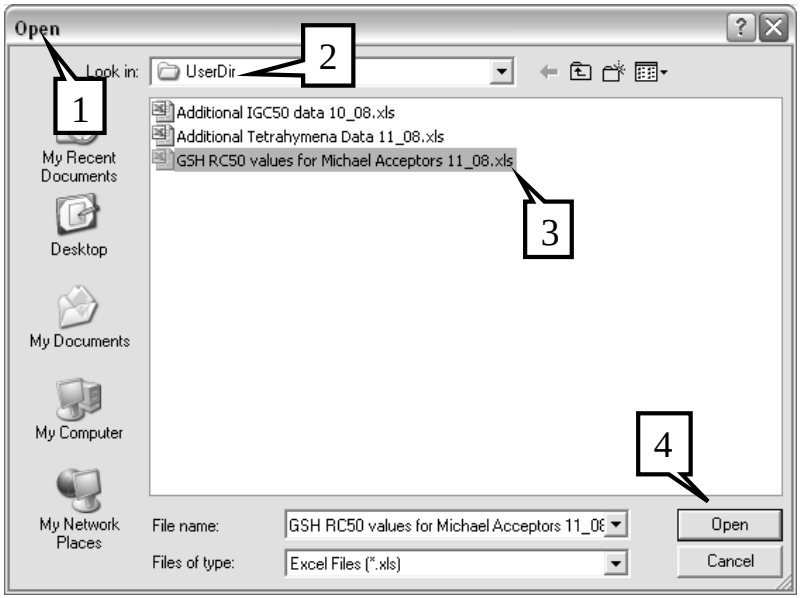
After launching the Toolbox in the Flexible Track, click on the **Endpoints** button **(1)** and then press the **Import** button **(2)** and follow the import wizard (see screenshot below).



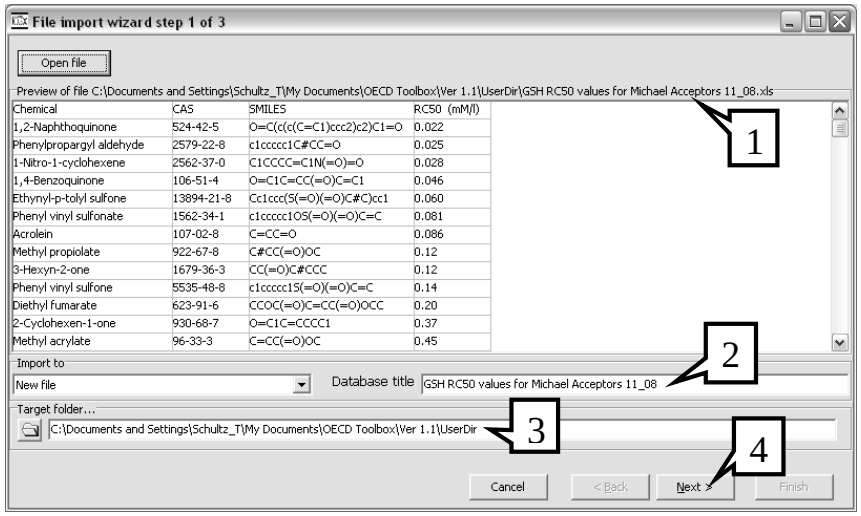
In step one of the wizard, press the **Open file** button **(1)** (see screenshot below).



This reveals the window **Open (1)**, which allows you to find the data file you wish to import into the Toolbox. Since the database we wish to import was stored as a spreadsheet in \My Documents\OECD Toolbox\Ver 1.1\UserDir, go to **UserDir (2)**, highlight the file of choice (in this example it is **GSH RC50 values for Michael Acceptors 11_08.xls (3)**. Clicking on **Open (4**) (see screenshot below) reveals another window.

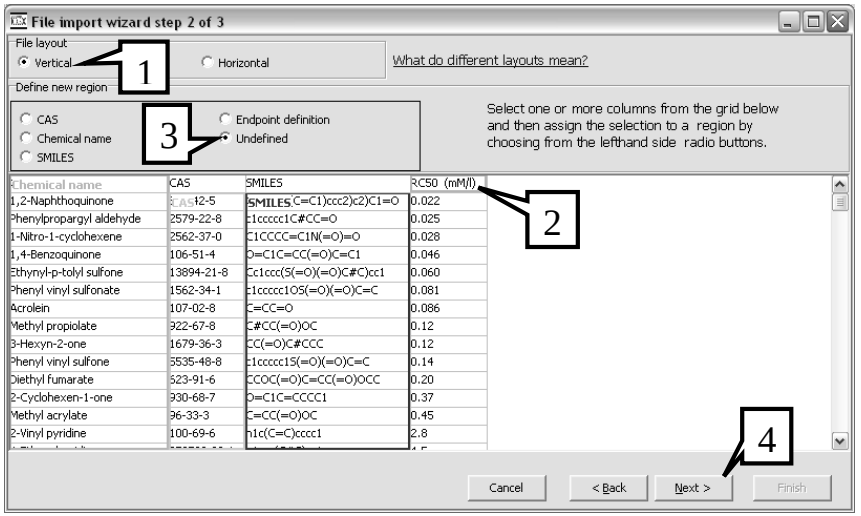


Upon clicking **Open** in the screenshot above the window **File import wizard step 1 of 3** appears (see screenshot below). This window is filled in with the relevant information including **Preview of the file (1)**, the **Database title (2)** and the **Target folder…(3).**

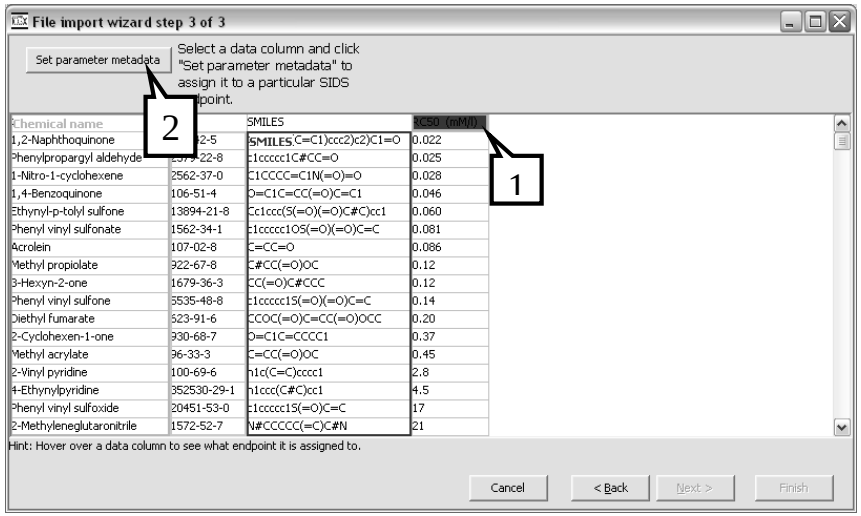


Clicking **Next (4)** in the screenshot above brings the window for step 2 into view.

File import wizard step 2 of 3 is viewed in the screenshot below. In a **Vertical (1)** file layout, each column (noted in a color outline) is assigned a definition. In this screenshot the RC50 (mM/l) is linked to **Undefined (3)** in **Define new region**. Once all columns have been assigned (even if several are listed as Undefined) click on the **Next** button **(4)**.

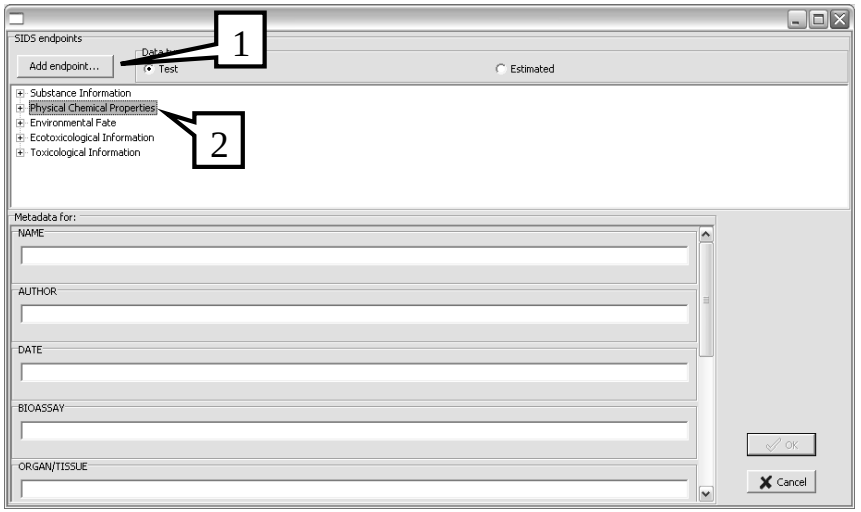


This brings into view the **File import wizard step 3 of 3** window (see screenshot below).

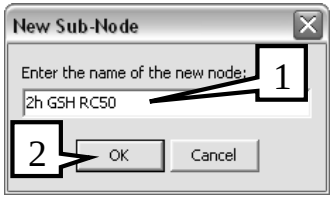


Since the data being imported is in a vertical layout the metadata for the endpoint data in a particular column is identical. Thus, the metadata only has to be entered once. In this example the column headed with the cell with the red background, RC50 (mM/l) needs to be mapped to the Toolbox by setting its metadata. After clicking on **RC50 (mM/l) (1)** click the **Set parameter metadata** box **(2)**.

This brings into view the window **SIDS endpoints** (see screenshot below).



In this example, we are going to add the endpoint (2-hr RC50 for GSH reactivity), which is a new endpoint and not in the data matrix of the Toolbox. By clicking on **Add endpoint…** button **(1)** with the **Physical Chemical Properties (2) highlighted**. The window **New Sub-Node** appears and the name of the node is entered **(1)** before clicking the **OK** button **(2)** (see screenshot below).

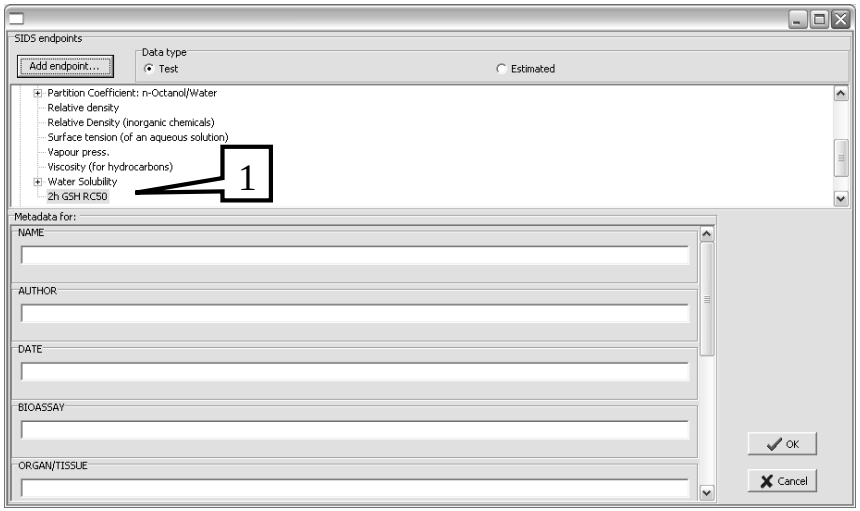


This brings into view the **SIDS endpoints** window (see screenshot below). The new endpoint, **2h GSH RC50 (1)** has been added. The metadata must now be entered. Remember this is a vertical import so metadata can be entered in the ten boxes under **Metadata for:**.

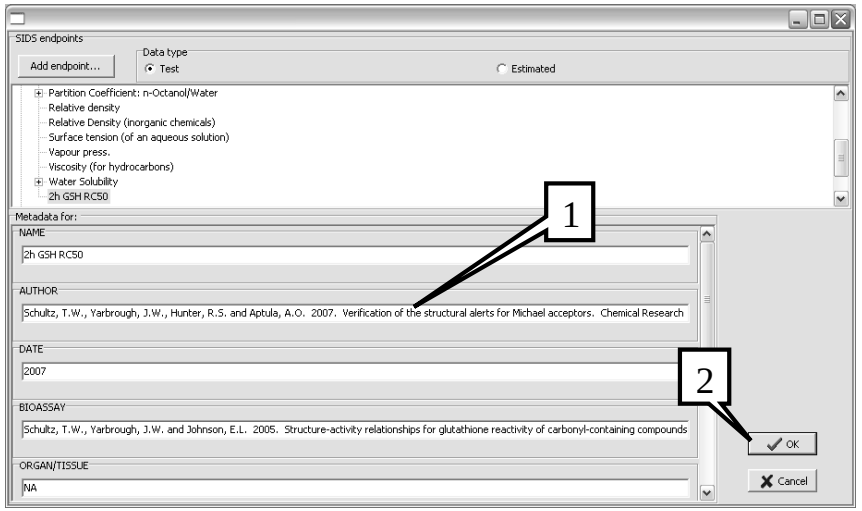
Remember in the Vertical import **Metadata (2)** includes the following:

NAME	the name of the database
AUTHOR	the compiler of database
DATE	the date the database was complied
BIOASSAY	Information on the bioassay used to generate the endpoint results
ORGAN/TISSUE	target organ or tissue for the endpoint
CONDITIONS	information on test conditions
ADMINISTRATION	the route by which the chemical was administered
SOURCE	Reference for the endpoint results
DIMENSIONS	the units of endpoint values
COMMENT	a field where additional information may be noted

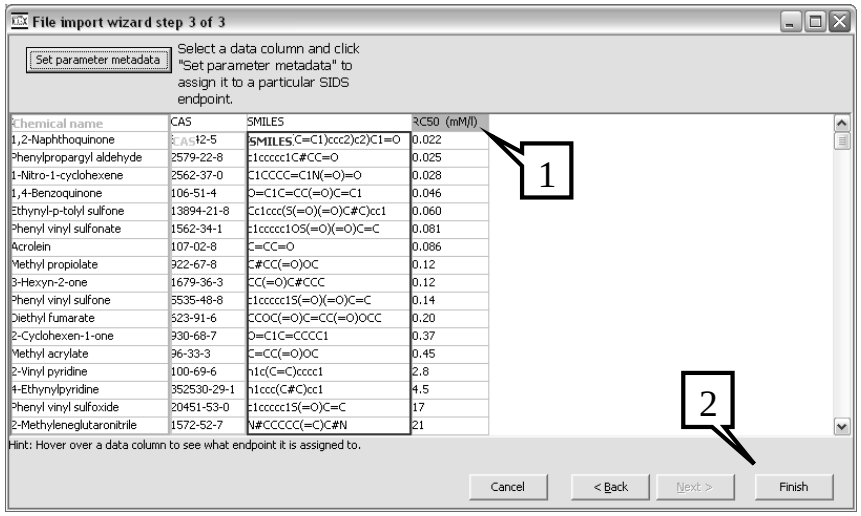
Of these ten fields DIMENSIONS is critical to mapping the database to the Toolbox. This is because the Toolbox recognizes DIMENSIONS only when expresses in particular forms (see Annex 1). For example, the Toolbox recognizes mg/l but not ppm



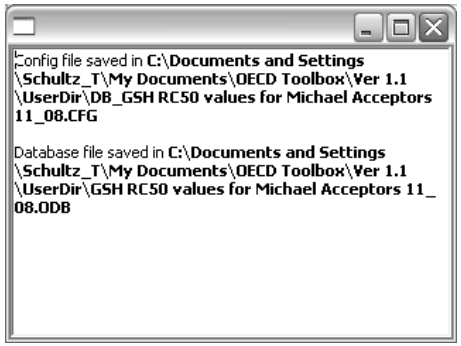
The metadata is typed into each of the ten defined boxes **(1)** before clicking the **OK** button **(2)** (see screenshot below).



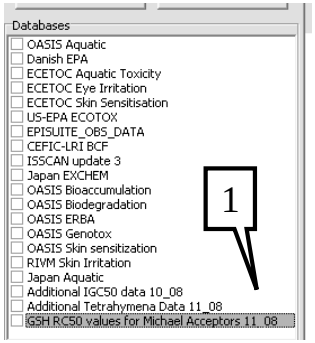
This brings into view again the **File import wizard step 3 of 3** window. This time the RC50 (mM/l) cell has a green background **(1)** indicating the endpoint has been mapped to the Toolbox (see screenshot below).



Click the **Finish** button **(2).** After some time, while the Toolbox configures and saves the new database, a window appears confirming that the database has been configured and saves (see screenshot below).



Upon closing the above window the Toolbox formats the new data matrix before returning to the Toolbox workflow screen where the **GSH RC50 values for Michael Acceptors 11_08** appears in the **Databases** box **(1)**.



5. TIPS AND TRICKS

5.1 How to share imported databases with other Toolbox users

When importing a new database into a toolbox, to files are created and stored locally:

- A database file named after the imported database and with the extension ODB, e.g. **Additional IGC50 data 10_08.ODB** in the example outlined in chapter 2. The default location of this file is c:\...\My Documents\OECD Toolbox\Ver 1.1\UserDir

- A configuration file named after the imported database and with the prefix DB extension CFG, e.g. **DB_Additional IGC50 data 10_08.CFG** in the example outlined in section XXX.. The default location of this file is c:\...\My Documents\OECD Toolbox\Ver 1.1\UserDir

[Note, the default location can be changed at installation. If the defaults were changed at installation, the files will be found in the location indicated at installation.]

These two files can be shared with any other local installation of the Toolbox, i.e. by copying them into the folder c:\...\My Documents\OECD Toolbox\Ver 1.1\UserDir of another installation, the new database will automatically appear in the application after restarting it.

It is nevertheless necessary to edit the CFG file first before copying it into another installation. In the CFG file, replace

"DBName=c:\...\My Documents\OECD Toolbox\Ver 1.1\UserDir\Additional IGC50 data 10_08.ODB"

by

"DBName= Additional IGC50 data 10_08.ODB"

5.2 Trouble shooting

Missing USER folder.

The default USER folder for the OECD Toolbox is placed in the **My Documents** folder. That, however, leads to a possible scenario where the My Documents folder is configured to be on a network drive that is for some reason temporarily unavailable.

In the case the Toolbox is started with no USER folder defined, a number of issues may arise, including:

- Inability to write QSARs

- Inability to create profiling schemes

- Inability to import databases

To avoid these issues you need to select a local folder to house the Toolbox USER folder.

If you've already installed it, and you want to avoid future problems, you could use the Options to change the USER folder to one that is on the local drive, and move the files from the old USER folder to the new one.

InterBase has insufficient rights.

The OECD Toolbox has a number of databases that come with the installation package that use the InterBase RDBMS. In order for them to work, the InterBase server needs to be operational and have access to the database files.

It is possible that the InterBase service runs in a security context that does not allow it access to the database otherwise available to the currently logged-in user. In this case you may get one of the following errors:

- Cannot establish connection to [FILENAME].

- Win32 error 5.

- Access is denied.

One way of solving that is to assure that the security context, that the InterBase service runs on, has the necessary access permissions for reading and writing to the database file.

Another work-around is to start the InterBase server in the same security context as your user. Assuming that you have full control over the DB files, so would the InterBase Server. To do that you need to uncheck the "Run InterBase server as a service on Windows NT" checkbox (figure 2), part of the InterBase Server Manager and restart your computer. You can find the InterBase Server Manager under Start Menu->Programs->InterBase 6 Open Edition - 6.0.2.0. (figure 1)

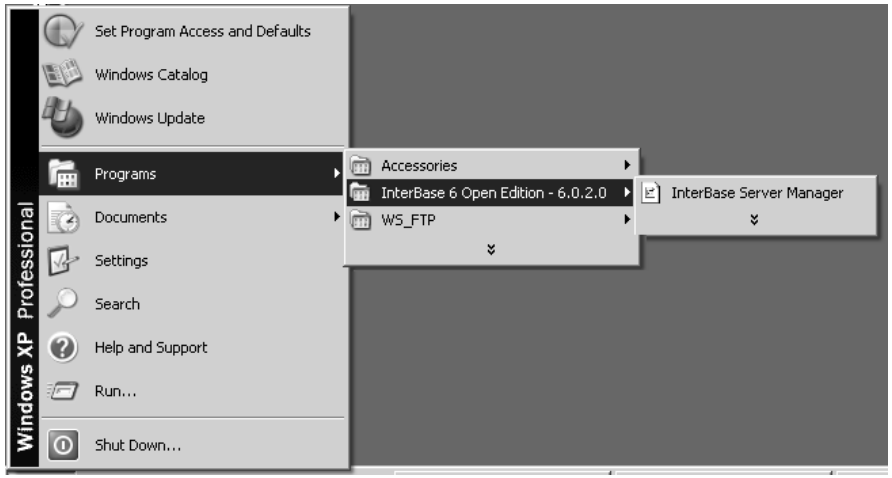


Figure 1

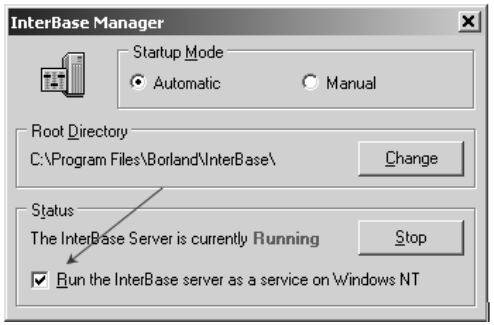


Figure 2

You are trying to access an InterBase database from a network drive.

The InterBase server can only access databases on the local machine. If you try to open a database located on a network drive you would receive the following message:

Cannot establish connection to [FILENAME] cannot attach to password database

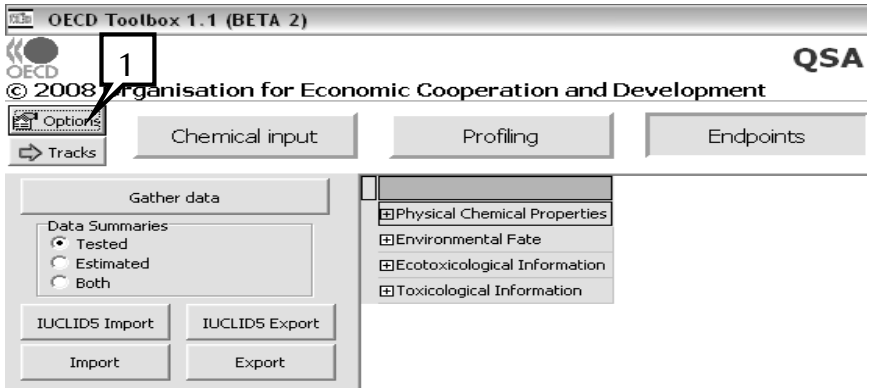
There is no fix for this issue in the current version of the Toolbox, so in order to open the database you would need to move the ODB file to one of your local drives and edit the corresponding Config file's **DBName** row to point to the new database file's position.

ANNEX 1: DIMENSIONS

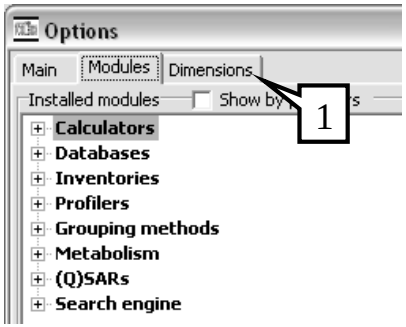
As indicated in the chapters above, when defining the units of the endpoint values imported into the Toolbox, it is necessary to choose a unit that the Toolbox recognizes so that it can perform conversions if necessary.

To find out which units are recognized by the Toolbox, the module **Options** can be consulted.

When the works flow window of the Toolbox is opened, in the upper right hand corner of the window is the **Options** button **(1)** (see screenshot below).



Clicking on the **Options** button opens the options window. Clicking on the **Dimensions** Button **(1)** opens the **Dimensions** window.



The **Dimensions** window lists all the dimensions recognized by the Toolbox for different types of endpoint values, for example for **Concentration (1)**, 18 units can be used and are automatically converted by the Toolbox to allow for a homogeneous use of all the concentration results.

In this example the **Data matrix** concentrations dimension (i.e. the unit used for estimated results which are inserted into the datamatrix) is noted as **mg/l (2),** while the **Gap filing** concentration (i.e. the unit used when performing read-across or trend analysis) is noted as **mol/l (3).** Other dimensions are observed by clicking the down arrow in the box where **Concentration** is currently listed **(1)**. Click the **Cancel** button **(4)** and return to the work flow screen (see screenshot below).