Agilent G6825AA MassHunter Pathways to PCDL Software

Quick Start Guide

For Research Use Only. Not for use in diagnostic procedures.

What is Agilent Pathways to PCDL?

Agilent MassHunter Pathways to PCDL is stand-alone software designed to facilitate the creation of custom Agilent PCDLs from metabolites present in pathway content culled from popular databases such as BioCyc, KEGG, and WikiPathways. Pathways to PCDL is also part of the MassHunter VistaFlux software.

You can create a personal compound database (PCD) or a personal compound database and library (PCDL) from one or more compounds (metabolites/lipids) that participate, or are associated with metabolic reactions, in selected pathway(s). The PCD/PCDL is created by finding compound names and their associated formulas in a reference METLIN database, that is based on a selected pathway(s). You can filter and select pathways based on database, organism, and/or custom text entries.

The pathway metabolites can be filtered based on their relationship with the pathway (member of the pathway or reaction partner) or through custom text entries.

Agilent Technologies

Pathways to PCDL currently supports pathway content from BioCyc, KEGG (with the required license), and WikiPathways. The databases created with Pathways to PCDL can be used to query acquired data in a targeted data mining workflow, the results of which can be mapped onto biological pathway maps.

Features of Pathways to PCDL

- Create a custom MassHunter METLIN compound PCD/PCDL from pathway information contained in BioCyc, KEGG, and WikiPathways. The pathways may be selected based on:
  - one or many organisms
  - custom search text entries
  - combinations of organism and text entries
- The list of target compounds that you can extract can be filtered based on the compound being present:
  - in the selected pathway(s)
  - as a reaction partner to compounds involved in the selected pathway(s)
- The pathway PCD/PCDL that you create may be used for identification of metabolites as a source database in other MassHunter software (for example, MassHunter Qualitative Analysis and ID Browser).

Terminology

PCD An accurate mass compound database. The database may or may not contain retention time data.

PCDL An accurate mass compound database that also contains an MS/MS accurate mass spectral database (spectral library).

Compound Name Many compounds are commonly known by their salts. The mass spectrometer, however, detects the anion or cation portion of the salt, rather than the neutral salt. PCD/PCDL entries may contain the familiar compound names, but the empirical formulae reflect the detectable cation or anion portion of the molecule rather than the formula of the neutral compound salt. For example, the full name Vecuronium bromide may be used for identification in the PCD/PCDL, even though the mass/formula only includes the Vecuronium cation.

METLIN database The MassHunter METLIN PCD/PCDL (METLIN database) is one of the best-known and most-comprehensive metabolite and lipid databases in the world today. The METLIN database contains endogenous and exogenous metabolites, as well as lipids, drug metabolites, and di- and tripeptides. Each entry may include mass, chemical formula, and structure information, as well as ID numbers that link to more information about the compound, such as the CAS and ChemSpider IDs.

Where to Find More Information

Go to www.agilent.com for the most current information on Agilent products.

Getting Started

How do I get started?

This *Quick Start Guide* helps you launch Pathways to PCDL, become familiar with a typical workflow, and use the software to create a custom PCD/PCDL metabolite database from a reference METLIN Metabolite PCD/PCDL database.

1 Install Pathways to PCDL. Follow the instructions in "Pathways to PCDL Installation" on page 15.

2 Double-click the **Pathways to PCDL** icon located on your desktop, or click **Start > All Programs > Agilent > MassHunter Workstation > Pathways to PCDL > Pathways to PCDL**.

3 The first time you run Pathways to PCDL you are prompted to set a reference METLIN database path and filename.

   a Click **Select** in the **Select Reference METLIN PCD/PCDL** dialog box (see Figure 1).

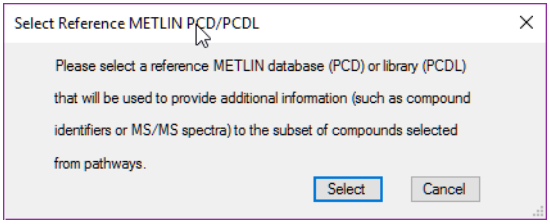


Figure 1 **Select Reference METLIN PCD/PCDL** dialog box

   b Select your METLIN database in the **Please specify the reference METLIN PCD or PCDL (.cdb) file** dialog box (see Figure 2 on page 5).

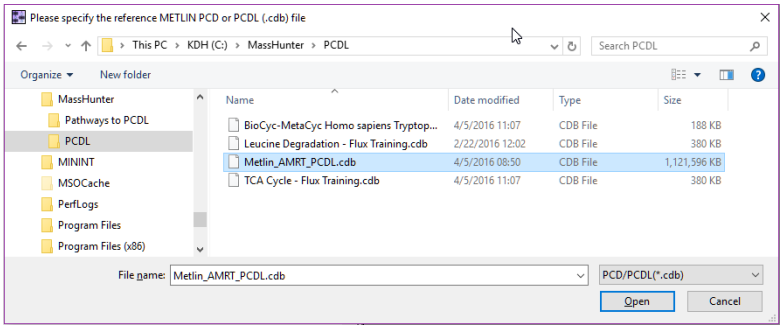


Figure 2 **Please specify the reference METLIN PCD or PCDL (.cdb) file** dialog box

4 Study the Pathways to PCDL Workflow shown in Figure 3.

5 Review the rest of this *Quick Start Guide* and follow the steps using your METLIN database.

Pathways to PCDL Workflow

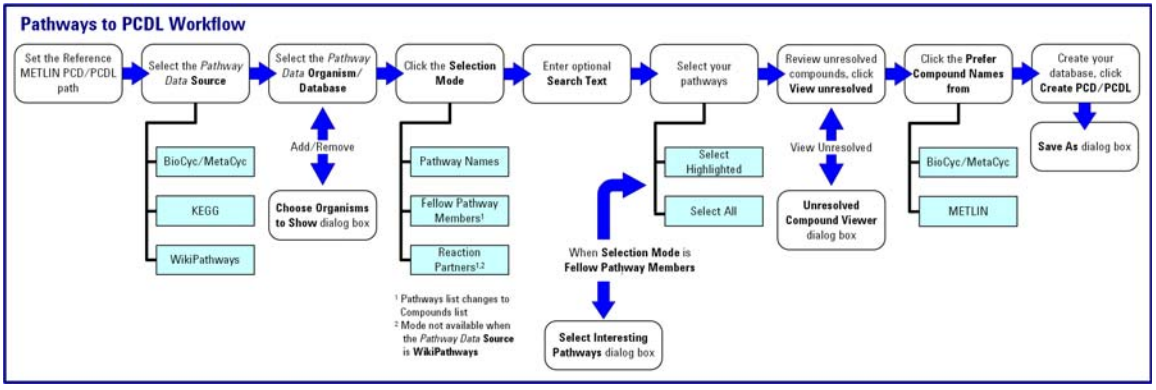


Figure 3 Typical Pathways to PCDL Workflow

User Interface

The main functional areas of Pathways to PCDL software are shown in Figure 4.

The Pathways to PCDL window consists of two parts: the Menu Bar and the Display Pane. The Display Pane is divided into the *Pathway List* and the *Selected Pathways List* to help you visualize your progress as you select pathways to create your PCDL. The number of pathways and compounds that meet your criteria are shown above each table in the *Display Pane*.

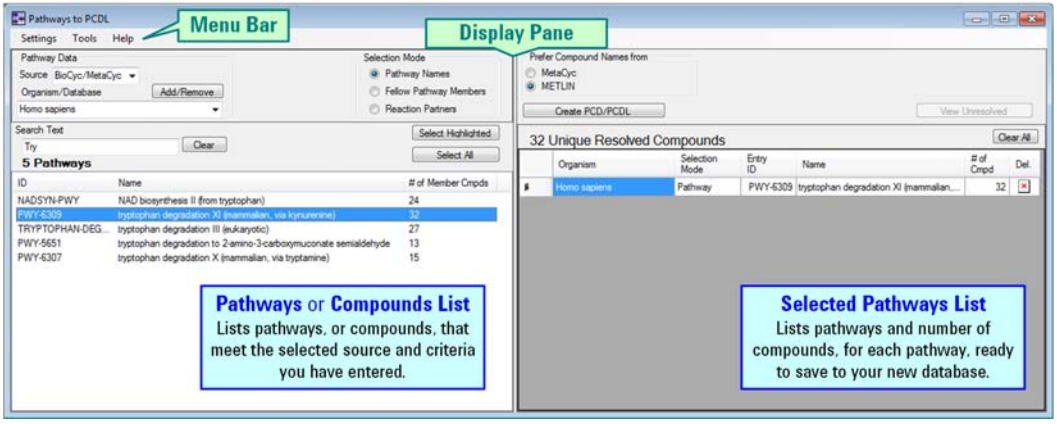


Figure 4 Window areas of Pathways to PCDL

You can update the pathway databases from **Tools** located on the Menu Bar (see Figure 5).

This *Quick Start Guide* and the software version are available from Help located on the Menu Bar (see Figure 5).

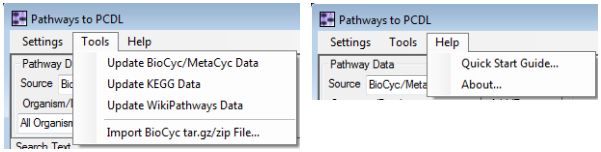
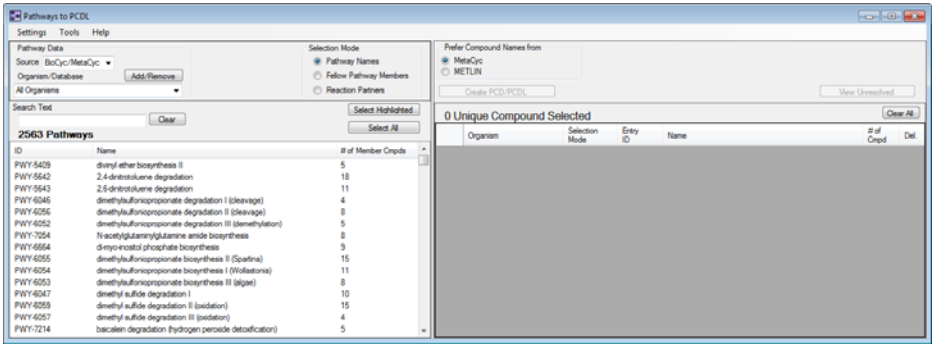


Figure 5 Tools and Help options on the Menu Bar

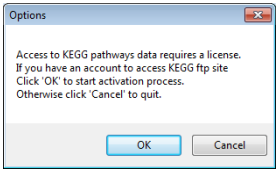
Creating a Custom PCDL

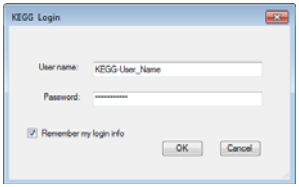
Creating a custom PCD/PCDL from metabolites present in pathway content follows a simple workflow as shown in Figure 3 on page 5.

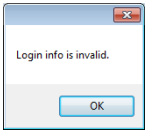
Steps	Detailed Instructions	Comments
1 Start Pathways to PCDL.	**a** Click the **Pathways to PCDL** icon on your desktop.	• You can also click **Start > All Programs > Agilent > MassHunter Workstation > Pathways to PCDL > Pathways to PCDL**.
2 Select **Source** under **Pathway Data**.	**a** Select **BioCyc/MetaCyc** for the **Source** under **Pathway Data**. If you select **KEGG**, go to step 3c on page 8. **b** Review the updated list of pathways in the *Pathways List*, and then go to step 3 to if you are using the BioCyc/MetaCyc pathway database.	• Pathways to PCDL supports pathway content from **BioCyc**, **KEGG**, and **WikiPathways**. • When you select **KEGG**, you are immediately prompted for your KEGG database license.



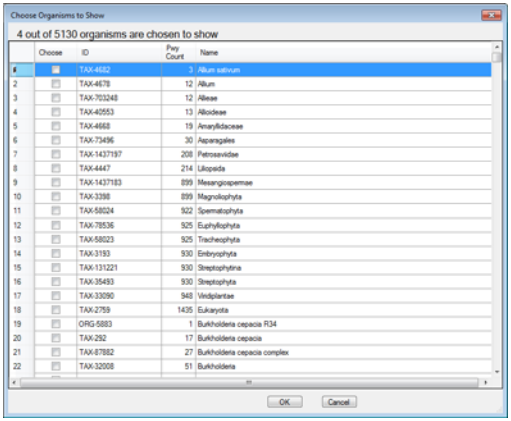
Steps	Detailed Instructions	Comments
KEGG pathway database license:	**c** Click **OK** in the **Options** dialog box. **d** Enter your license information if you select the **KEGG** pathway database. **e** Click **OK**.	• If the KEGG Login **User name** or **Password** is incorrect a dialog box notifies you.

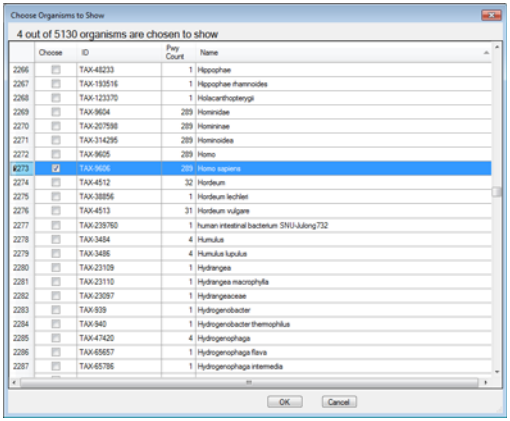




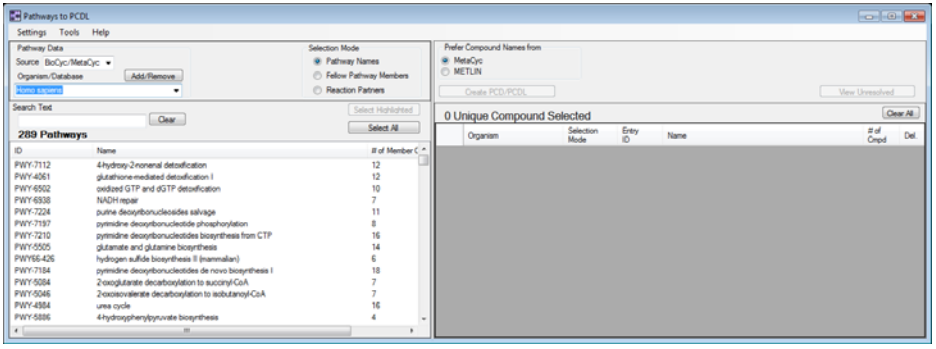


3 ***Pathway Name:*** Select the organism(s) you want to include within the pathway data source.	**a** Click **Add/Remove** under **Pathway Data**. **b** Click the column heading *Name* to sort the organisms. For example click the **Name** column heading. **c** Mark **Homo sapiens**. Mark a check box in the **Choose** column and in the row for each organism that you want to include in your **Organism/ Database** selection list. **d** Click **OK**.	• In the **Choose Organisms to Show** dialog box, the **Pwy Count** column is the number of pathways associated with each organism. • Each of the organisms that you select in the **Choose Organisms to Show** dialog box are available in the **Organization/Database** selection list. The selection is persistent the next time you run **Pathways to PCDL**.

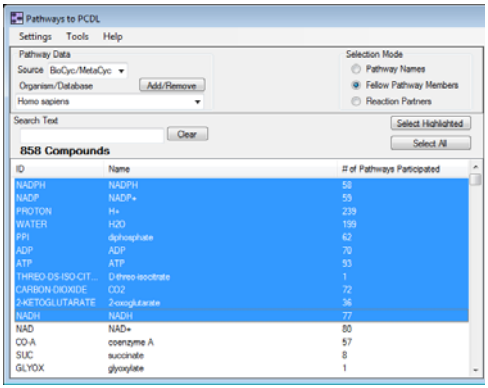




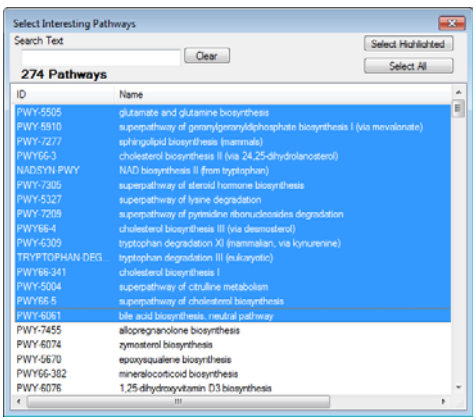
Steps	Detailed Instructions	Comments
	e Select **Homo sapiens** from the **Organism/Database** list. **f** Review the updated list of pathways in the *Pathways List*.	• Organisms may be removed from the **Organism/Database** selection list. Click **Add/Remove** and clear the organism check boxes.



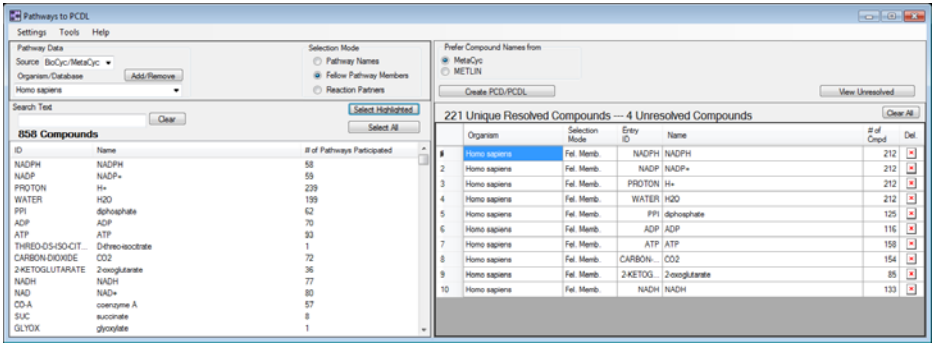
Steps	Detailed Instructions	Comments
4 ***Fellow Pathway Members:*** Select pathways related to compounds when **Fellow Pathway Members** is the **Selection Mode**.	**a** Click **Fellow Pathway Members**. **b** Review the updated list of compounds in the Pathway List. **c** Select the first eleven (11) compounds in the *Compounds List*. **d** Click **Select Highlighted**.	Three options are available as the **Selection Mode**: • **Pathway Name:** Build your personal compound database by selecting pathways individually (as shown in this step) or by clicking **Select All** from the *Pathways List* as shown on the prior page. • **Fellow Pathway Members:** Build your personal compound database by selecting compounds within a pathway that you select based on the pathways all being related to each other (fellow pathways) to the compound you previously select from the *Compounds List* as shown on this page. • **Reaction Partners:** Build your personal compound database by selecting compounds associated with pathway reactions involving the compound you select from the *Compounds List* in the *Display Pane*. • Compounds associated with each pathway are added to the *Selected Pathways List* each time you click **Select Highlighted** or **Select All**.



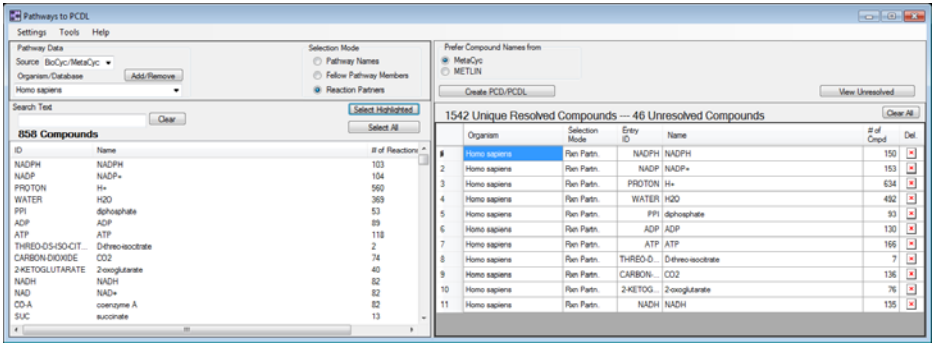
e Select one or more pathways, with or without filtering the pathway list using custom search text. Select the first fifteen (15) pathways.
f Click **Select Highlighted**.



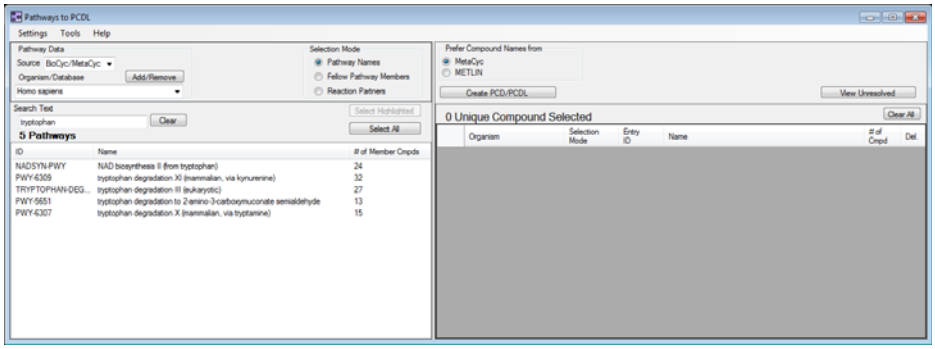
Steps	Detailed Instructions	Comments
	g Review the list in the *Selected Pathways List*. Go to step 6 on page 12	

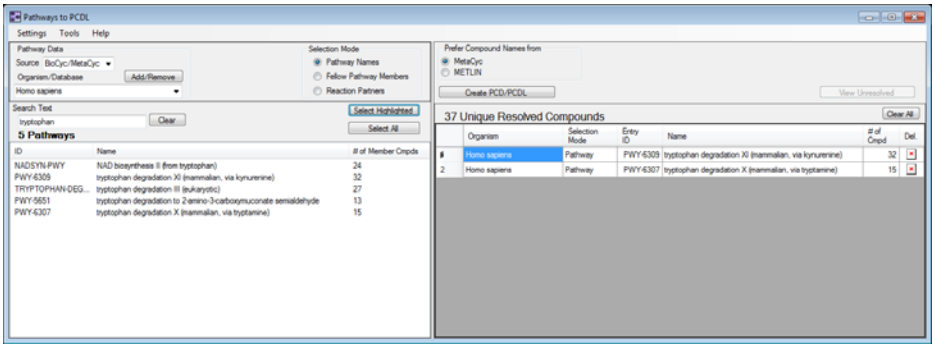
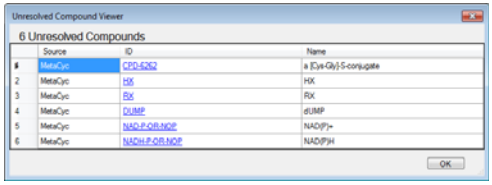


Steps	Detailed Instructions	Comments
5 ***Reaction Partners:*** Select pathways related to compounds by reaction when **Reaction Partners** is the **Selection Mode**.	**a** Click **Clear All** in the *Selected Pathways List* to start a new pathways list. **b** Click **Reaction Partners**. **c** Review the updated list of compounds in the *Compounds List*. **d** Select the first eleven (11) compounds in the *Compounds List*. **e** Click **Select Highlighted**. **f** Review the list of pathways in the *Selected Pathways List*.	• **Reaction Partners** is not an option when KEGG or WikiPathways is the **Pathway Data**, **Source**. • After you click **Select Highlighted** or **Select All** the original selection in the *Pathways List* is no longer highlighted.

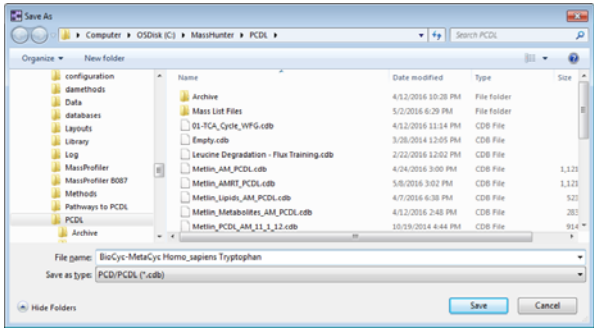


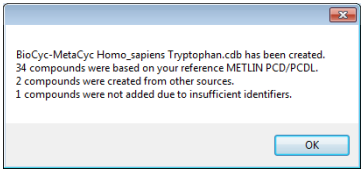
Steps	Detailed Instructions	Comments
6 Select compounds using custom **Search Text** entry when any **Selection Mode** is selected.	**a** Click **Clear All** in the *Selected Pathways List* to start a new pathways list. **b** Type `tryptophan` for the **Search Text**. **c** Review the list of pathways in the *Pathways List*. The number of pathways is smaller than when viewed in step 3 on page 8.	• Pathways can be filtered by several text searches.



Steps	Detailed Instructions	Comments
7 Add one or more pathways to your compound database.	**a** Click the pathway you want to add to your compound database. **b** Click **Select Highlighted** to move the pathway to the *Selected Pathways List.* **c** Repeat steps a and b to continue adding pathways to your compound database.	• Click **Select Highlighted** or **Select All** after each search to build your custom compound database based on the selected pathways.
		
8 Review unresolved compounds in your *Selected Pathways List*.	**a** Click **View Unresolved**. **b** Review the list of unresolved compounds in the **Unresolved Compound Viewer** dialog box. **c** Click **OK**. **d** Review the list of compounds in the Compound List. **e** Remove unresolved compounds from your Compound List. Mark compounds to delete under the **Del.** column.	• Pathways to PCDL provides you with an opportunity to manually review the compounds in your *Selected Pathways List* that may not have a strong correlation with the pathway(s) you have selected (unresolved compounds). • The **View Unresolved** button is available when you have unresolved compounds in your Compound List. • After reviewing the unresolved compounds, you may retain the related pathways in your *Selected Pathways List* or remove them from the *Selected Pathways List* by clicking the compound row under the **Del.** column. When you click the compound under the **Del.** column it is immediately removed.
		

Steps	Detailed Instructions	Comments
9 Select the preferred source of the final database compound names.	**a** Click **MetaCyc** as the **Prefer Compound Names from** source.	
10 Create your custom PCD/PCDL.	**a** Click **Create PCD/PCDL**. **b** Select the folder to save your PCD/PCDL database. **c** Type the name for your PCD/PCDL database in **File name**. **d** Click **Save**.	• This step creates your custom database based on the compounds involved in the pathway selections you have made. • If your reference METLIN metabolite database is a PCD, your custom database does not contain MS/MS data. If your reference METLIN metabolite database is a PCDL, your custom database contains MS/MS data.
	e Click **OK** after you have reviewed the summary information describing your custom PCD/PCDL.	
11 Create another PCD/PCDL or exit Pathways to PCDL.	**a** Create another PCD by returning to "Select Source under Pathway Data." on page 7. **b** Click the **Close** button to exit Pathways to PCDL if you are done creating your custom PCD/PCDL database.	• You have completed creating your first custom PCD/PCDL database. You can create another database or exit the software.





Pathways to PCDL Installation

Pathways to PCDL is installed from a Setup Wizard, which you run from the main installation program..

Double-click **Pathways to PCDL.msi**, or right-click the file and then click **Install**.

Acknowledgments and Citations

BioCyc Pathway/Genome Databases

Includes BioCyc Pathway/Genome databases from the Bioinformatics Research Group at SRI International®, used under license.

http://www.biocyc.org/

Citation based on use of BioCyc databases or the Pathway Tools software

If you use BioCyc databases or the Pathway Tools software in your research, cite relevant publications as described on the BioCyc website:

http://biocyc.org/publications.shtml

For example, users who publish research results in scientific journals based on use of data from the EcoCyc Pathway/Genome database should cite:

Keseler et al., *Nucleic Acids Research* **39**:D583-90, 2011.

Users who publish research results in scientific journals based on use of data from most other BioCyc Pathway/Genome databases should cite:

Caspi et al., *Nucleic Acids Research* **40**:D742-53, 2012.

KEGG Database

Includes KEGG (Kyoto Encyclopedia of Genes and Genomes) databases developed by Kanehisa Laboratories.

http://www.genome.jp/kegg/

Citation based on use of KEGG Database

If you use the KEGG database in your research, cite relevant publications as described on the KEGG website:

http://www.genome.jp/kegg/kegg1.html

This page intentionally left blank.

www.agilent.com

In this book

The *MassHunter Pathways to PCDL Software Quick Start Guide* presents the first steps to use the MassHunter Pathways to PCDL Software.

This Quick Start Guide applies to MassHunter Pathways to PCDL B.07.00 and later until superseded.

For Research Use Only. Not for use in diagnostic procedures.

Revision B, May 2016

G6825-90008

Agilent Technologies