

HUMAN HEALTH

ENVIRONMENTAL HEALTH

ChemBioFinder 14.0 User Guide

Table of Contents

Chapter 1: About ChemBioFinder	1
About BioViz	1
About this guide	1
Serial Number and Technical Support	2
Chapter 2: Getting Started	4
The ChemBioFinder User Interface	4
More UI features	10
Opening ChemBioFinder	11
Using ChemBioFinder with databases	11
Chapter 3: Tutorials	14
Sample databases	14
Tutorial 1: Creating forms	14
Tutorial 2: Opening a database	16
Tutorial 3: Creating a database	17
Tutorial 4: Searching a database	19
Tutorial 5: Reaction searches	22
Tutorial 6: Creating a BioViz plot	25
Tutorial 7: Working with subforms	29
Chapter 4: Forms	31
Creating forms automatically	31
Saving a form	32
Creating forms manually	33
Setting box properties	39
Editing forms	46
Chapter 5: Databases	58
Selecting a database	58
Opening databases	61

Browsing databases	63
Creating a database	69
Creating a portal database	75
Chapter 6: Working with Data	76
Entering data	76
Editing data	78
Sorting data	83
Resetting the database	84
Changing the database scheme	84
Chapter 7: Importing and Exporting Data	86
Supported file formats	86
Importing data	86
Exporting data	92
Chapter 8: Queries	96
Text searches	96
Numeric searches	97
Molecular formula searches	97
Date searches	99
Find list	99
Structure searches	100
Reaction searches	105
Combined searches	109
SQL searches	110
Query procedures	110
Refining a search	117
Special structure searches	120
Managing queries	124
Search examples	130
Chapter 9: Compound Profiles	135

Creating a compound profile	136
Editing a compound profile	138
Customizing	142
Chapter 10: Chemical Properties	143
ChemBioFinder Chemical Properties	143
About Molecular Networks	144
ChemBio3D Chemical Properties	153
ChemBioDraw Properties	153
Chapter 11: Relational Data and Subforms	155
Access relational data	155
Creating a subform	155
Working with subforms	157
Searching a subform	157
Chapter 12: BioViz	160
Creating a plot	160
Viewing Plots	163
Plotting queries	164
Changing the display	166
Statistical analysis	167
Customizable styles	172
Chapter 13: Clustering in BioViz	174
Features	174
Working with Clustering	174
Modifying Clustering	176
CS Clustering Dialog features	178
Chapter 14: ChemBioFinder/Oracle	180
Setup	180
Opening an Oracle database	181
Searching	182

Sorting	182
Handling lists	183
Setting Oracle preferences	184
Updating and adding data	184
Loading	185
Indexing	186
CAL	188
Chapter 15: BioSAR	189
Starting the BioSAR browser	189
Performing searches in BioSAR	189
Exporting the data form to ChemBioFinder	189
ChemBioFinder file types	189
Subforms and autolinks	190
Multiple sort	191
Query hit lists	192
Domains	195
Creating an mdb database	196
Chapter 16: Customizing ChemBioFinder	199
Setting preferences	199
Favorites tree	205
Toolbars	206
Periodic table	208
Embedding ActiveX controls	209
Scripting	210
Methods of communication	218
Chapter 17: Structural Query Features	222
General properties	222
Changing the scope of a search	230
Atom properties	230

Bond properties	235
Chapter 18: Formula Input Rules	238
Rules	238
Examples	238
Chapter 19: Similarity Rules	240
Complete structure similarity	241
Substructure similarity	241
Chapter 20: CAL Commands	242
CAL help	242
Menu commands	242
Box creation commands	245
Box manipulation commands	254
Program execution commands	258
General commands	262
File commands	277
Variable commands	280
Environment variables	287
Script-only commands	288
Chapter 21: ChemScript	295
Why use ChemScript?	295
How ChemScript works	296
Getting Started	296
Editing Scripts	297
Introducing the ChemScript API	298
Tutorials	299
Useful References	302
Chapter 22: CS Oracle Cartridge	303
Fast-move caching scheme	303
Configuring via CF_SETTINGS table	304

Searching305

About ChemBioFinder

ChemBioFinder is a database management system for anyone who works with chemical information. It provides a place to store chemical structures, physical properties, notes, tables of data, and plots. With ChemBioFinder, you can quickly find and organize your data.

About BioViz

BioViz is a visualization add-in for ChemBioFinder that lets you plot numeric data from ChemBioFinder databases. You can identify trends and correlations in your data without exporting to another application. You can have as many plots as you like, each showing a different visualization of data from the current form.

ChemBioFinder & BioViz is integrated with these applications:

ChemBioDraw. Lets you draw two-dimensional structures and reactions.

ChemBio3D. Lets you view three-dimensional structures.

ChemBioDraw/Excel. Lets you export data to MS Excel.

About this guide


The topics in this user guide are organized by task. They are intended to help you familiarize yourself with the ChemBioFinder applications and start using them as quickly and efficiently as possible. New users should read the Getting started chapter for an overview of the product and how it works. The Tutorials chapter demonstrates most of the features of the application. Perform the tutorials in the order they are presented. Experienced users can skip to subsequent chapters, which provide more detailed information. This guide assumes that you are familiar with the basics of your Windows operating system. If you are not, refer to your system manual. Some of the material describes tasks that must be performed in conjunction with other integrated ChemBioOffice products. The material on the addins describes tasks that must be performed in conjunction with Microsoft Excel or Word. If you are not familiar with these applications, consult the relevant user's manual.

Conventions

The following notations are used throughout this guide:

 **Caution:** *Cautions warn you of situations that might cause a possible loss of data.*

 **Note:** *Notes highlight important information.*

 **Tip:** *Tips are helpful hints that supplemental the main text.*

Shortcut key sequences are indicated with a + sign, for example: "Use the command: CTRL+H to toggle hidden Hydrogens and lone pairs."

A bold font is used to refer to the User Interface, for example "Go to **File>Save**".

Serial Number and Technical Support

The serial number is a unique key that helps you install and activate the ChemBioOffice application on your computer. This is the same number you entered when you launched your ChemBioOffice application for the first time. If you have lost your installation instructions, you can find the serial number from the application. Go to **Help>About <ChemBioFinder>**. The serial number appears in the dialog box.

For more information on obtaining serial numbers and registration codes, see the [online Services Web site](#).

Technical support is available to registered users through the Internet and our Technical Support department. When contacting Technical Support, always provide the serial number of your ChemBioOffice application. Our Technical Support Web pages contain answers to frequently asked questions (FAQs) and other information.

To access our Technical Support in ChemBioFinder, go to **Online>Browse CambridgeSoft Technical Support**. If you are unable to find a solution on the Web site, follow these steps before you contact Technical Support:

1. Check the system requirements for the software.
2. Read "[Troubleshooting](#)" on page 2 and follow the possible resolution tactics outlined there.
3. If all your attempts to resolve a problem fail, contact [Technical support](#).
 - Try to reproduce the problem before contacting us. If you can reproduce the problem, please record the exact steps.
 - Record the exact wording of any error messages that appear.
 - Record what you have tried to correct the problem.

Additional sources of ChemOffice information are:

- ChemBioFinder online Help
- The [PerkinElmer Informatics Web site](#)
- [PerkinElmer Informatics Services](#)

Troubleshooting

This section describes steps you can take that affect the overall performance of CS ChemBioOffice Applications, as well as steps to follow if your computer crashes when using a CS software product.

Performance

Below are some ways you can optimize the performance of ChemBioOffice Applications:

- In the Performance tab in the System control panel, allocate more processor time to the application.
- Install more physical RAM. The more you have, the less ChemBioOffice Applications will have to access your hard disk to use Virtual Memory.
- Increase the Virtual Memory (VM). Virtual memory extends RAM by allowing space on your hard disk to be used as RAM. However, the time for swapping between the application and the hard disk is slower than swapping with physical RAM.

Applications and Drivers

As with most complex software applications, there may be unusual circumstances in which ChemBioFinder may become unresponsive. Below are some recommended steps for you to follow to try to resolve software and driver issues.

1. Restart Windows and try to reproduce the problem. If the problem recurs, continue with the following steps.
2. The most common conflicts concern video drivers, printer drivers, screen savers, and virus protection. If you do need to contact us, be sure to determine what type and version of drivers you are using.
 - Video Driver related problems: If you have problems displaying any ChemBioOffice Application, try switching to the VGA video driver in the display Control Panel or System Setup, and then retest the problems. If using a different driver helps, your original driver may need to be updated—contact the maker of the driver and obtain the most up-to-date driver. If you still have trouble, contact us with the relevant details about the original driver and the resulting problem.
 - Printer Driver related problems: Try using a different printer driver. If using a different driver helps, your original driver may need to be updated—contact the maker of the driver and obtain the most up-to-date driver. If you still have trouble contact us with the relevant details about the original driver and the resulting problem.
3. Try reinstalling the software. Before you reinstall, uninstall the software and disable all background applications, including screen savers and virus protection. See the complete uninstall instructions on the Technical Support web page.
4. If the problem still occurs, use our online [contact form](#) and provide the details of the problem to Technical Support.

Getting Started

In this section, we describe the ChemBioFinder working environment and introduce the basic features. After you become familiar with the fundamental concepts, you will be able to perform database queries and review the results. Afterwards, you will learn to create your own query forms, plot data, and perform other tasks using ChemBioFinder. The general steps for performing a query in ChemBioFinder are:

1. Open (or create) a form.
2. Create or open a database and link it to the form.
3. Enter query parameters in the form.
4. Perform a search.
5. Browse the search results.

These steps are described in the ChemBioFinder Tutorials, and in the detailed reference material in the manual.

The ChemBioFinder User Interface

The ChemBioFinder User Interface (UI) consists of menus, toolbars, and the form window. The central part of the form window contains the work space where you create a form, enter query parameters, and view query results. The ChemBioFinder UI is shown below:

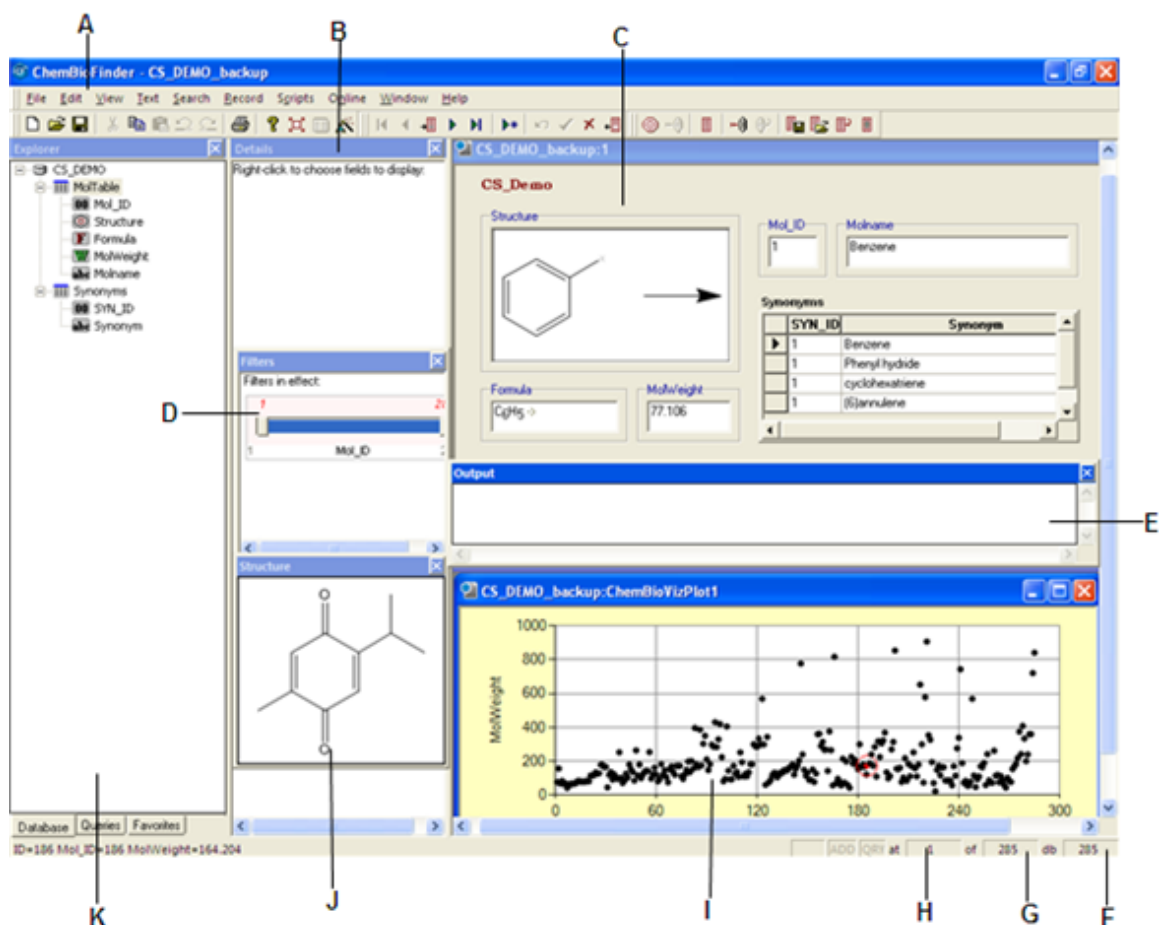


Figure 2.1: The ChemBioFinder user interface; A) Menu bar; B) Details window; C) Main form; D) Filters window; E) Output window; F) Total database size; G) Current list size; H) current record; I) BioViz window; J) Structure window; K) Explorer window.

The table given below describes the components of the ChemBioFinder UI.

Screen Element	Description
Main form	Displays data contained in one record of the database. To keep the display from getting cluttered, additional tabs may be added for less important data.
Data box	A viewing area in a form linked to a field in the database associated with that form, and displaying data from that field. See also: Control

Screen Element	Description
	box, Picture box, Profile box, Subform, Text box.
Frame	An enclosure around a data box or group of boxes containing a label.
Framed data box	A data box surrounded by a labeled frame.
Picture box	Displays a graphic. Picture boxes may be either Fixed, that is, they remain the same for all records in the database, and are linked to a particular graphic file, or associated with a picture Field, in which case there could be a different graphic in each record.
Profile box	Displays Compound Profiles.
Structure box	Displays the chemical structure of the current record.
Text box	Displays text. Text boxes are Fixed (text fields are displayed in text data boxes) and may be formatted in the Box Properties dialog box.
Menu bar	Contains all the commands specific to the application for managing forms, tables, databases, and their contents.
Details window	Displays details of selected (or moused-over) point when a plot window is displayed. May be displayed as tabbed with other windows.
Explorer window	Displays information on three tabbed (overlapped) windows: <ul style="list-style-type: none"> ▪ The Database field hierarchy. ▪ The Queries list that is associated with the form. ▪ A Favorites list of frequently used files.

Screen Element	Description
BioViz window	Displays plotted data from ChemBioFinder databases.
Filter window	Displays the filters that modify the BioViz plot. May be displayed as tabbed with other windows.
Output window	Displays Python scripts output. May be displayed as tabbed with other windows.
Structure window	Displays the chemical structure of the selected (or moused-over) point when a plot window is displayed. May be displayed as tabbed with other windows.
Subform	Displays relational information from a different database, or a different data table within the same database.
Main toolbar	Contains general-purpose commands such as copying, saving, and printing.
Form toolbar	Contains the form-creation tools.
Record toolbar	Contains commands from the Record menu.
Search toolbar	Contains commands from the Search menu.
Text Format toolbar	Used to format text in Text and Memo fields.
Control box	A data box containing an ActiveX control.
Status bar	Displays information about the item opened in the ChemBioFinder window as well as recent actions and user-interface elements.
New Record indicator	Displays ADD when you add a new record that has not been committed to the database.
Read-only indicator	Displays READ if the database is read-only.

Screen Element	Description
Query indicator	Displays QRY when you enter a search query.
Current record	Displays the position within the current hit list of the record you are viewing.
Current list size	Displays the number of records through which you are browsing. Might be less than the total database size if you have recently performed a search.
Total database size	Displays the total number of records in the current table.

Toolbars

The ChemBioFinder toolbars help you to create and manage forms and the database records that are displayed. Toolbars are normally docked at the top and left side of the UI, but they can be “torn off” and placed anywhere on the screen for your convenience. You can view the toolbars displayed in the ChemBioFinder window through **View>Toolbars** submenu. You can also use the Toolbars submenu to customize the toolbars. A description of toolbar customization is beyond the scope of this manual. (See Microsoft help for information on customizing toolbars.)

Main toolbar

Use the 'Main toolbar' to perform basic functions such as opening and closing forms, copying/pasting objects, and modifying a form.

- Layout Mode—switches between Layout (Edit) Mode and View Mode. When you click **Layout Mode**, the Form Toolbar and an alignment grid appear on the form.
- Switch to Table—switches between Form View and Table View.
- Database Wizard—activates the Database Wizard to connect a database to a form.

Form toolbar

Use the Form Toolbar to add an object to an existing form.

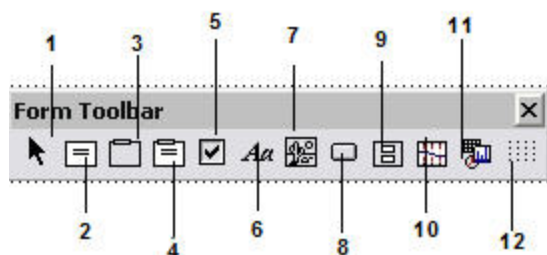



Figure 2.2: The Form Toolbar: 1) Selection tool; 2) Data box; 3) Frame; 4) Framed box; 5) Check box; 6) Text box; 7) Picture; 8) Button; 9) Subform; 10) Profile box; 11) Control box; 12) Grid.

 **Note:** The Form Toolbar docks on the left side of the ChemBioFinder window by default, but can be docked anywhere or “torn off”, as shown.

The tools are described in "Forms" on page 31.

Search toolbar

The Search Toolbar contains the tools you need to query the database and work with the hit lists that the query produces.

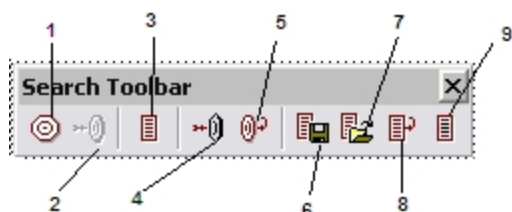


Figure 2.3: The Search Toolbar: 1) Enter Query; 2) Find; 3) Retrieve All; 4) Find Current Molecule; 5) Previous Query; 6) Save List; 7) Restore List; 8) Restore Previous List; 9) Find List.

Record toolbar

The Record Toolbar contains the tools you need to browse a database or hit list. It also has tools for adding, deleting, and changing records.

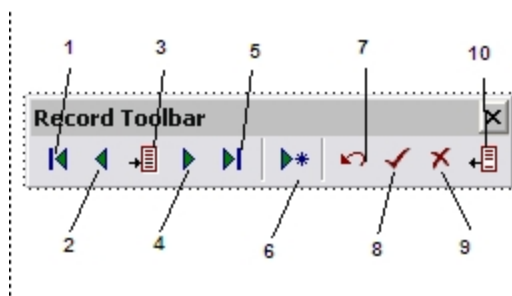


Figure 2.4: The Record Toolbar: 1) First Record; 2) Previous Record; 3) Go to Record; 4) Next Record; 5) Last Record; 6) Add a Record; 7) Undo Changes; 8) Commit Changes; 9) Delete Record; 10) Omit Record.

Text format toolbar

The Text Format Toolbar contains standard text formatting tools that you can use when entering or editing information in data and text fields.

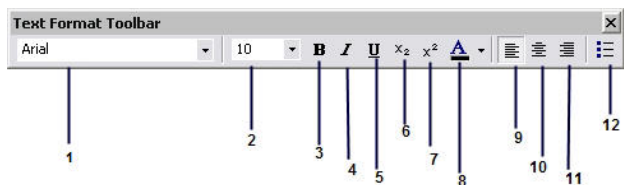
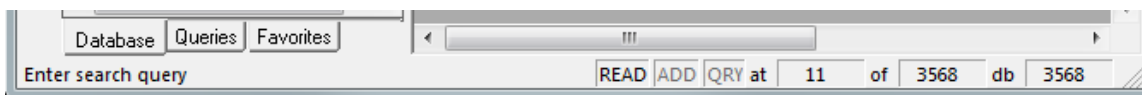


Figure 2.5: The Text Toolbar: 1) Font; 2) Point size; 3) Bold; 4) Italics; 5) Underline; 6) Subscript; 7) Superscript; 8) Color; 9) Left alignment; 10) Center alignment; 11) Right alignment; 12) Bullets

The status bar

While you move among records, counters in the lower-right corner of the window change to indicate the current record, the current list size, and the total size of the database. The lower-left corner of the window displays help for menu items and other information.



When you first open a form, the current list size equals the total database size. The total database size changes when you add or delete records. If you search to find a subset of the entries in the database, then the current list size changes to indicate the number of hits in the search.

To the left of these counters are three other indicators that show the general status of the database:

- READ indicates you are using a read-only database, a read only form file, or if you have selected **Open as read-only** on the open file dialog box.
- ADD appears when you enter a new record.
- QRY when you enter a query.

To hide or show the Status Bar, select or deselect **View>Status Bar**.

More UI features

Style button for text and frame boxes

The Box Properties dialog box makes the Style button available for frames and plain text labels. This button opens the Box Style dialog box, where you can set frame type, printability, and other display characteristics.

Read-only structure box option

The Box Style dialog box has a check box **Box is read-only**. This is applied to text, numeric data boxes and structure boxes. If you mark a structure box read-only, the structure cannot be edited except in query mode.

Automatic frame label updates

The label of a framed box is updated if you change the field it displays.

Resizable dialog boxes

The Box Properties, Plot Properties, Data Import and Data Export dialog boxes, and the script editor, can be resized, allowing you to see more in text windows, tree controls, and so forth. The new position and size is applicable only to the current ChemBioFinder session.

Opening ChemBioFinder

In default mode, ChemBioFinder opens with the ChemBioFinder startup dialog. The various tabs in the dialog box lets you do the following:

From the **New** tab

- Blank Form—Open a blank form to create your own.
- Database Connection—Open a blank form and connect it to a database. You can create the form manually or automatically. For more information, see ["Connecting a database to a form" on page 16](#).
- Database Wizard—Helps you set up a database connection.

From the **Existing** tab

- Browse to an existing form in your file system.

From the **Recent** tab

- Open a form on the recently used file list.

You can choose to display/hide the ChemBioFinder startup dialog when ChemBioFinder opens by doing the following:

1. Go to **File>Preferences**.
2. In the **Preferences** dialog box, click the **General** tab.
3. Select or deselect **Show startup dialog**.

If you choose to hide the ChemBioFinder startup dialog box, ChemBioFinder will open a new blank form by default when you open the application.

To open the ChemBioFinder startup dialog box when the preference is set to hide, go to **File>New**.

Using ChemBioFinder with databases

ChemBioFinder comes with the Microsoft Jet database engine. ChemBioFinder maintains its own table of chemical information—structure, formula, and molecular weight—and relies on Jet to create a database for managing the rest of the data. The data created by Jet is stored in an MS Access database file (MDB file). If you have MS Access installed, you can use it with the relational part of the database.

You can use ChemBioFinder to add chemical structures to a database you have already developed in Microsoft Access.

The database model

A database is a collection of information. In ChemBioFinder the information is organized into levels of increasing complexity. At the simplest level is the data item itself, for example a molecular weight value as shown in the table below:








MOL_ID	Structure	Formula	MolWeight	Molname
157		C ₂₁ H ₃₀ O ₅	362.46	Hydrocortisone
158		C ₁₉ H ₃₀ O ₂	290.44	Androsterone
159		C ₂₁ H ₃₀ O ₂	314.462	Progesterone
160		C ₁₆ H ₁₇ N ₃ O	267.326	Lysergamide
161		C ₁₆ H ₁₆ N ₂ O ₂	268.31	Lysergic acid
163		C ₁₇ H ₂₀ N ₄ O ₆	376.364	Riboflavin
164		C ₁₂ H ₁₇ N ₄ S	265.355	Thiamine

Figure 2.6: A data item in the database

At the next level is a field, a group of data items defining one type of data. Fields are generally set up once and rarely modified.







MOL_ID	Structure	Formula	MolWeight	Molname
157		C ₂₁ H ₃₀ O ₅	362.46	Hydrocortisone
158		C ₁₉ H ₃₀ O ₂	290.44	Androsterone
159		C ₂₁ H ₃₀ O ₂	314.462	Progesterone
160		C ₁₆ H ₁₇ N ₃ O	267.326	Lysergamide
161		C ₁₆ H ₁₆ N ₂ O ₂	268.31	Lysergic acid
163		C ₁₇ H ₂₀ N ₄ O ₆	376.364	Riboflavin

Figure 2.7: A data field in a database

At the next level is the record, a set of data items (one for each field) defining a single entry.

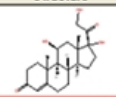
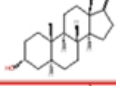
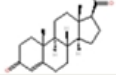
MOL_ID	Structure	Formula	MolWeight	Molname
157		C ₂₁ H ₃₀ O ₅	362.46	Hydrocortisone
158		C ₁₉ H ₃₀ O ₂	290.44	Androsterone
159		C ₂₁ H ₃₀ O ₂	314.462	Progesterone

Figure 2.8: A record in a database

A collection of records is a table. A table is defined by a set of fields, and a set of records which grows as entries are added.

In spreadsheet terminology:

- A data item is found in a cell.
- A field corresponds to a column.
- A record corresponds to a row.
- A table corresponds to a worksheet.

A database is a storehouse for one or more tables. A database containing only one table is known as a simple (or flat) database. Databases containing multiple tables are called relational. For more on relational databases see ["Relational Data and Subforms " on page 155](#)

A form displays data from a single table, but may contain subforms that display data from other tables. If the tables have a field in common, then any record retrieved in the form calls up the related records in the subform.

Understanding forms and databases

A database is where data is stored whereas a form displays but does not store the data. The form acts like a window into the database, letting you select which fields and tables you want to view. You can create more than one form to access the database. For example, you can create one sample form for working with structural data and another to include literature or lab data. By switching between forms, you can look at just those fields you want to see.

While no data is stored in a form, some things are saved with it. BioViz plots, database queries (depending on preference settings), file paths of linked graphics, and certain settings are saved with each form.

Tutorials

The general steps for using ChemBioFinder are:

1. Create or open a form.
2. Create or open a database and link it to the form.
3. Add or manage data.
4. Perform a search.


The tutorials introduce you to ChemBioFinder basic functions. We recommend performing the tutorials in the sequence they are presented because each tutorial builds on and refers to the previous ones.

Sample databases

The Samples directory contains several small databases, forms, and sample scripts. By default, these files are in the Samples folder at:

C:\Documents and Settings\All Users\Application Data\CambridgeSoft\ChemOffice2014\ChemFinder\Samples

The CS_Demo database contains 285 structures covering a range of structural types. To browse the Cs_Demo database, open the form CS_Demo.cfx. This is the form used in the tutorials in this guide. You may also find the form Demo.cfx useful. This is a simpler form that displays only the main table of the CS_Demo database.

 **Caution:** When you open a sample database file, ChemBioFinder gives you the option to create a backup of the file. We recommend you make a backup copy of the database before experimenting with it. If you change the data or structures in the database, the examples in the tutorials may no longer give the documented results.

Tutorial 1: Creating forms

Forms let you view and browse through your data in a display format of your own design.

To begin building a form, first create a new one using the option **File>New** or the startup dialog, then select the Layout mode button on the Main toolbar. For information about each tool, see [Toolbars](#).

Creating data boxes

Data and structures from a database are displayed in boxes on the form.

To create a data box:

1. In the Form toolbar, click the **Data Box** tool.
2. In the form window, click drag diagonally to create a box.
3. Draw two more boxes in the same way. You will edit them in later steps.

A frame, a framed box, and a text box are different types of boxes and require a label.

To draw a framed box:

1. Click the **Framed Box** tool and drag diagonally to create a framed box. A box, labeled **Data** is created.

2. Right-click on the **Data label** and select **Label**. The 'Box Text or Label' dialog box appears. In the 'Box Text or Label' dialog box, type **Frame Box** and click **OK**. The label appears above the Framed Box.

To draw a frame around several existing boxes:

1. Click the **Frame** tool.
2. Place the pointer at the corner of the upper left data box and drag to create a border around the group of three data boxes. The Enter the Label dialog box appears.
3. In the Enter the Label dialog box, type **Frame** and click **OK**. The label appears above the box.

To place a picture in your form:


1. Click the **Picture** tool and drag in the form. The Open dialog box appears.
2. Browse to a graphic file, then click **Open**. The picture appears in the area you dragged in the form.


 **Note:** ChemBioFinder supports EMF, WMF, BMP, GIF, JPG, PNG, and TIF formats.

Editing data boxes

To edit one or more data boxes you must first select them:

1. Click the Selection tool in the Form toolbar.
2. To select the upper left box, click it. A selected data box is designated by four black squares at its corners.

 **Note:** To de-select the box, click an empty space in the form with the Selection tool. To select multiple boxes, press **SHIFT** and click in each box.

 **Tip:** To select multiple boxes, you can also click-drag a rectangle around the boxes. To select all boxes on the form, go to **Edit>Select All**.

The frame and the box behave as one object when you select, move, resize, or delete. To separate them into two objects:

1. Click on the frame of the Framed box to select it.
2. Go to **Edit>Bring to Front**.

Here are other editing options:

- To resize a box, select the box and drag a side or corner.
- To reposition a box, click-drag the center of the box to a new location.
- To move multiple boxes, select them and click-drag the center of the selection to a new location.

Many of the Windows commands you use on text can also be applied to data boxes:

- **Edit>Cut** or press **DELETE**.
- **Edit>Paste**

- **Edit>Undo**
- **Edit>Redo**

Saving a form

To save a form:

1. Draw another box on the form using the **Framed box** tool and label it as Structure.
2. Go to **File>Save As**.
3. In the Save As dialog box, save the form as tut1.cfx in the directory of your choice.
4. Go to **File>Close**.


Tutorial 2: Opening a database

After you create a form, you can use it to connect to a database. In this tutorial, you will use the form you saved in the previous tutorial. After you connect to the database, data will appear on the form.

Connecting a database to a form

To connect a database to your form:

1. Go to **File>Open**. The Open dialog box appears.
2. Select tut1.cfx and click **Open**. The form you created in the previous tutorial appears, with its fields blank.
3. Select **Layout mode** on the main toolbar if it is not already selected.
4. Right-click on the Structure framed box and select **Data Source**. The Box Properties dialog box appears with the Database tab displayed.
5. Click **Open Database**. The Open dialog box appears.

 **Note:** The **Open** and **Create Database** buttons work with ChemBioFinder databases only. To access data in other types of databases, use the **Attach Table** or the **Oracle** button (if available). For more information about data sources, see ["Attaching tables from other applications" on page 70](#).

6. Select CS_demo.mdb from the Samples folder (\ChemOffice 2014\samples folder) and click **Open**. The database opens, and the Box Properties dialog box appears displaying the Database tab. A list of the tables in the database and the fields in each table are shown.

 **Note:** The field you select from the table determines what type of data appears in the box you selected in step 4.


Assigning fields to data boxes

To display structures from the CS_Demo database in the Structure Framed box on your form:

1. Right-click on the Structure Framed box on your form and select **Properties**.
2. From the list of fields in the Box Properties dialog box select **Structure** and click **OK**. The Structure field is linked to the Structure Framed box. In the Structure data box, the data of the Structure field for the first record in the CS_Demo database appears.

To assign fields to the other data boxes:

1. Right-click another box, click **Properties**, and select **Molname** from the properties list. Click **OK**. The Molname field in the database is linked to this box, and the data for the first record appears in it.

 **Note:** After you open a database and table, the shortcut menu displays the database fields.


2. Right-click in another box and click **Formula**.
3. Right-click a third box and click the **MOL_ID**.
4. Go to **File>Save As**.
5. In the Save As dialog box, save the form as tut2.cfx in the directory of your choice.
6. Go to **File>Close**.

Congratulations! You have created your own customized form for viewing the CS_Demo database.


Tutorial 3: Creating a database

In this tutorial you will create a new database and use automatic form generation method.

1. Create a new form.
2. Right-click in the form and click **Data Source**. The Form Properties dialog box appears with the Database tab displayed.
3. Click **Create Database**. The Save As dialog box appears.
4. Type **mydb** and click **Save**. The name of the database appears in the Properties box. ChemBioFinder creates one data table (MolTable) that contains four fields: Structure, Formula, MolWeight, and MOL_ID.
5. Click the **Field** tab. The options for creating and deleting fields appears. You are now going to add two new fields to the data table.
6. Click **Create Field**. The Create Field dialog box appears.
7. Type **MolName** in the text box, change the width to 254, and click **OK**.

 **Note:** You can enter a maximum of 254 characters in a text field. If you want a text field to contain more than 254 characters, choose Memo/Rich Text from the Type drop-down list.

8. Click **Create Field** again. Type **Boiling Point** in the text box, change the type to **Double**, and click **OK**. The data table tree displays your changes.

 **Note:** Use the field type **Double** to create a field containing real numbers (such as -123.7 and 43.242).

9. Click the **Form** tab. The various form options appear.
10. Click the **Generate form** check box. Click **Style**. In the Form Generation dialog box, make sure the check box for **Structure in upper left of form** is checked.
11. Click **OK** in the Form Generation dialog box and again in the Form Properties dialog box. The Mydb form is generated.

Next, check the field assignments in the data boxes:

1. Right-click in the structure data box. There should be a check next to **Structure** indicating that the Structure field is linked to the structure data box.
2. Right-click in the Boiling Point data box. There should be a check next to **Boiling Point**.

Adding records

Now that you have created new fields and assigned them to data boxes, you can add data to your database.


This tutorial demonstrates the use of the ChemBioDraw control (the default structure mode). Before you begin to add structures, check whether the structure mode is set appropriately.

To check the current setting:

1. Right-click on the Structure data box, and click **Properties**.
2. In the Box Style section of the Box Properties dialog box, select **ChemBioDraw style**.
When you select ChemBioDraw style, ChemBioFinder defaults to the ChemBioDraw control. This lets you edit structures in the Structure data box in ChemBioFinder. ChemBioFinder style opens ChemBioDraw in its own window for editing.
3. To exit from the Form editor, deselect the **Layout** tool.

Now you are ready to begin adding records to your database. In the steps below, we create records for benzene, n-pentane, and cyclohexane:

1. Double-click in the Structure box. The ChemBioDraw toolbar appears.


 **Tip:** With ChemBioDraw style selected, you can also open ChemBioDraw in its own window by right-clicking in the structure box and selecting **Edit>Edit in ChemBioDraw**.

2. Select the benzene tool, and draw benzene in the structure box.
3. Click anywhere outside the Structure data box to finish editing the structure. Benzene appears in the Structure Data box. ChemBioFinder calculates the molecular formula, and assigns an ID number of 1.
4. Click the MolName box and type Benzene.
5. Click the Boiling Point box and type 80.1.
6. Go to **Record>Commit Changes** create the new record.

You have now entered the first record in your database. The size of your database is indicated in the Status Bar.

Enter two more records:

1. Go to **Record>Add New Record**.
2. Add a record for n-Pentane, with boiling point = 36.1.
3. Repeat step 1 and then add a record for Cyclohexane with boiling point= 80.7.

 **Note:** After you have two or more records in your database, you can commit changes by moving to a different record using the Record tools.

4. Go to **File>Save**.

 **Note:** Choose **Save** to save the form. Choose **Commit Changes** to save the data in your database.

Tutorial 4: Searching a database

You can often find useful information in smaller databases by browsing through records one at a time. However, for larger databases, you will likely want to enter specific search criteria to find what you want. When you search a database, you find only those few records that have the information you want. After you have this smaller set of records (a hit list), you can then browse it much more efficiently than you could the whole database.

In this tutorial, we discuss text and structure searches. We will cover reaction searches in ["Tutorial 5: Reaction searches" on page 22](#).

Opening the sample database

To open the CS_Demo database:

1. Start ChemBioFinder and click the **Existing** tab on the Open dialog box. If ChemBioFinder is already running, go to **File>New** to open the dialog box, or access CS_Demo directly from the **Favorites** tab of the Explorer Window.¹
2. Navigate to the samples directory: C:\ProgramData\CambridgeSoft\ChemOffice2014\ChemFinder\Samples.
3. In the Samples directory, select CS_demo.cfx and click **Open**. The CS_Demo database opens in ChemBioFinder.

Before you begin searching, open the Explorer window, if it is not already open.

4. Go to **View>Explorer Window**.
5. Click the **Queries** tab.

Formula searching

To find compounds in the CS_demo database with six carbons and one or two nitrogen atoms:


1. Go to **Search>Enter Query**. The form is cleared so that you can enter a new query.
2. Click in the **Formula** box and type C6N1-2.
3. Click **Search**.

The Queries tree in the Explorer window displays one query as a “child” of the full list. Queries can be saved with the form and can be reviewed at any time. For more information see ["Managing queries" on page 124](#).

There are three methods for browsing the search results:

- In the Form View—Use the **Record** menu commands or toolbar to go through the records.
- In the Data Table View—Browse the table to view the records.
- In the Continuous Forms View—view multiple records in their own forms.

To toggle between the different views, go to **View>Switch Views** or Ctrl+W.

 **Note:** As you view the records, counters in the Status bar indicate the current record, the current list size, and the total size of the database.

¹The Explorer window is available in Pro and Ultra versions but not in Standard version.

To view the database in a table that shows the records in a list, do one of the following:

- Go to **View>Data Table>In Current Window**.
- Click the **Switch to Table** tool.
- Type <Ctrl>+<T>.

 **Note:** *The Switch to Table tool is a toggle. Select it again to return to the Form view.*

The Table view appears and displays all the records of the current list (in this case, the records that were returned by the search) in a table. The table view provides several options for you to examine the search results. For example:

- You can sort the records by a specific field in Table View. For example, to sort by the MolWeight field, double-click on the MolWeight table header.
- You can change the column widths of your table. Position the pointer over a table header divider and drag to the width you want.
- You can adjust the height of any form. Drag the bar divider on the left.

To use continuous forms to browse your records:

- Go to **View>Continuous Forms**. The Continuous Forms view appears. By default, the Continuous Forms view shows the same form as the Form view.


To retrieve all the records in your database:


- Double-click **Full List** in the Explorer window.

Name searching

The query feature provides several options for searching for text in a database. For example, to find all compounds in the CS_Demo database with molecular names starting with “benz:”

1. Switch to Form View, if you are not already in it.
2. Go to **Search>Enter Query**. The form appears in the Query mode.
3. Click the **MolName** box, and type benz*.
4. Press <Enter>. All records are returned with names starting with “benz”.
5. Go to **View>Data Table>In Current Window**. The Table view appears. Browse to verify that the molecular names are correct.


 **Tip:** *Although the tutorials describe the use of the Search menu, you may find using the Search toolbar more convenient. The icons on the toolbar match those you have already seen on the Search menu. Clicking the Find icon is equivalent to pressing the ENTER key when you are ready to begin your search. See ["Queries" on page 96](#) for information on more advanced use of the Search toolbar.*

 **Note:** *Notice that this search gave you “benzene” but not “bromobenzene.” The query you entered above is an “anchored substring” and only gives you strings starting with the indicated substring. For more information on how to specify text searches, see ["Text searches" on page 96](#).*

Numerical searching

You can perform range searches in ChemBioFinder numeric fields. For example, to search in the CS_Demo Database for compounds with molecular weights between 90 and 100:

1. Switch to Form View, if you are not already in it.
2. Go to **Search>Enter Query**. The form is cleared so that you can enter a new query.
3. Click in the **Molecular Weight** box and type 90-100.
4. Press <Enter>. Your search returns all records with molecular weights between 90 and 100.
5. Go to **View>Data Table>In Current Window**. The Table view appears.

 **Note:** A molecular weight query is a decimal value or range. The precision of the search depends on the number of significant digits entered. For more information on molecular weight searching, see ["Numeric searches" on page 97](#).

Substructure searching

To search for a substructure:


1. Go to **Search>Enter Query**. The form is cleared to allow you to enter your search terms.
2. Double-click in the Structure box. A blue box is displayed around the Structure box and the ChemBioDraw control appears.
3. Draw benzene.


To set the correct options for a substructure search:

1. Select the **Search>Substructure** menu option, if it is not already selected.
2. Deselect the **Search>Similarity** menu option, if it is selected.
3. Click **Search** or go to **Search>Find**.

ChemBioFinder begins searching. The progress of the search is indicated by counters in the status bar at the bottom of the window.

When the search is complete, the number of hits is displayed in the Current List Size window of the Status Bar, and the form displays the first hit. In a substructure search, the matched portion of each molecule is highlighted in red. The above mentioned search returns all structures that contain a benzene ring.

 **Note:** You can also set search preferences on the Search tab of the Preferences dialog box. For more information, see ["Customizing ChemBioFinder" on page 199](#).

 **Note:** If a search gets no hits, an alert appears and you are returned to the query mode with the query on display.

Combined searching

In some cases, you may want to combine structure searching with text searching. For example, you may want to find all compounds in the database that have a benzene substructure and a molecular weight greater than 400.

To perform a combined search:

1. Go to **Search>Enter Query**. The form is cleared so that you can enter a new query.
2. Double-click in the Structure box. The ChemBioDraw control appears.
3. Draw benzene.
4. Click the **Molecular Weight** box and type >400.
5. Press <Enter>.

Your search returns all records that include a benzene ring and have a molecular weight greater than 400.

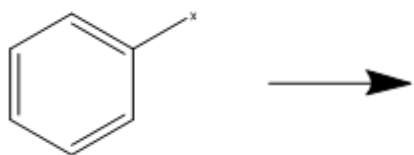
Tutorial 5: Reaction searches

In addition to helping you organize information about individual substances, ChemBioFinder lets you store and search for chemical reactions.

Searching for reactants

In this tutorial, you will learn how to search substructures and text in the Infochem Database. This database is included with ChemBioFinder and has approximately 250 reaction records. Searching for reactants is useful if you have a known starting material and you are interested in learning more about what substances it can produce.

1. Go to **Search>Enter Query**. The form clears.
2. Double-click in the Structure box. The ChemBioDraw toolbar appears.
3. Draw the structure below:



This structure represents a carbon atom bonded to a generic halogen atom (x). The arrow at the right indicates that you are looking for this substructure as a reactant.

4. Select the **Search>Substructure** if it is not already selected.
5. Deselect the **Search>Similarity** if it is selected.
6. Go to **Search>Find**.

ChemBioFinder begins searching. The progress of the search is indicated by counters in the status bar at the bottom of the window.

When the search is complete, the number of hits is displayed in the Current List Size window of the Status Bar, and the form displays the first hit. In a substructure search, the matched portion of each molecule is highlighted in red.

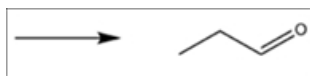
Searching for products

In this exercise, you search for information on a particular reaction product. Searching for products of reactions is very common in syntheses, where you know what you are aiming for but you do not know how to produce it. In this

example, we look for reactions that close a ring alpha to a carbonyl.

To perform a reaction product search:

1. Go to **Search > Enter Query**. The form is cleared so that you can enter a new query.
2. Double-click on the Structure data box. The ChemBioDraw toolbar appears.
3. Draw the following:



4. Switch to any **Selection** tool (Lasso or Marquee) and select the single bond next to the double bond.
5. Right-click and select **Topology> Ring** from the context menu.
6. Select the remaining single bond.
7. Right-click and select **Topology > Ring**. Right-click again and select **Reaction Center > Make/Break** from the context menu. The arrow at the left indicates that you are looking for this substructure as a product.
8. Press <Enter>. The hit list appears.

Searching by reaction type

In many cases, you have some idea of both your starting materials and your products, but are looking for some information on how to get from one to the other.

For example, to search for reactions that reduce a carbonyl to an alcohol:

1. Go to **Search>Enter Query**. The form is cleared.
2. Double-click in the Structure box.
3. Draw the following:

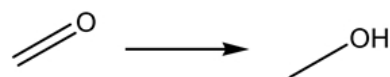


Figure 3.1: A reaction type query structure

4. Press ENTER.

Browse through the search results. Each of the reactions shows the transformation of a carbon-oxygen double bond to a carbon-oxygen single bond.

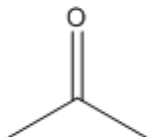
This hit list includes reductions of aldehydes, acids, and ketones.

Search over list

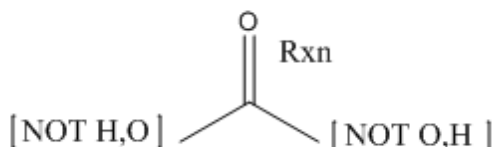
In this example, you start with the previous list (see ["Searching by reaction type" on page 23](#)) and search again to get only ketone reduction reactions.

To search over a list:

1. If you are not continuing directly from the last example, double-click the previous query in the Query Tree to make sure it is the active hit list.
2. Go to **Search>Over Current List**.
3. Double-click in the Structure box.
4. Draw the following:



- a. Using a selection tool, right-click on the double bond and set **Reaction Center** bond properties to **Change**.
- b. Using the **Text** tool, select one of the carbon atoms. Type **[NOT O,H]** (all uppercase, with brackets as indicated) and press ENTER to replace it.
- c. Using the **Text** tool, select the label, right-click and select **Repeat Last Label** from the context menu.
- d. Double-click on the other carbon atom to reproduce the label. The structure now looks like this:



You have specified a search for ketones by adding the restriction that the atoms adjacent to the carbon must not be oxygen atoms or hydrogen atoms. You also specified that the double bond had to change, and not be broken.

5. Press <Enter>.

The query returns results based on the query structure and the previous hit list. Each reaction shows the reduction of a ketone to an alcohol. In the Queries Tree, this search displays as a “child” of the first query.

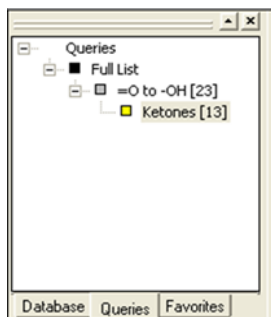



Figure 3.2: A child search in the queries tree

 **Tip:** If you plan to run several searches over a hit list, use the **Set Domain to Current List** command rather than **Search Over Current List**. Setting a domain eliminates the need to keep restoring the original hit list.

Tutorial 6: Creating a BioViz plot

Using BioViz, you can plot the data in your search results based on any numerical field in the database.

BioViz plots can consist of line or scatter plots (with either one or two variables) or Histograms (with X-axis data)


For more information on BioViz plots, see "[BioViz](#)" on page 160.

Single variable plots

A single variable plot displays a Y-coordinate variable vs. the record number.

To create a single variable plot:

1. Go to **View>ChemBioViz Plots>New**. The ChemBioViz Plot Properties dialog box opens.
2. Click **One variable** in the Dimension box.
3. Select any numerical field from the drop-down list for the Y-coordinate.
4. (optional)Enter a name for the plot.

 **Tip:** Note the **Point Limit** field on the lower right side of the Properties dialog box. Plotting a large dataset can be very slow. As a safety precaution, the default is set to 500, which means that only the first 500 points in the dataset will be plotted. On most computers, this will take no more than a few seconds. To plot a larger dataset, reset this value as necessary.

5. Click **OK**. The Plot appears in a new window.
6. (Optional) Click the **Notes** button to view additional information. For example, points will not be plotted if there are missing Y values. This is reported in the Notes, if any.

 **Tip:** To create a single variable scatter plot, right-click in a numerical field and select **BioViz Plot> ID Plot**.

Histogram plots

A histogram plot is a variation of a single variable plot, where the X-coordinate is the variable value, and the Y-coordinate shows the count in each histogram 'bin'.

To create a Histogram plot:

1. Go to **View>ChemBioViz Plots>New**.
2. Click the **One Variable** radio button in the **Dimension** box of the BioViz Plot Properties dialog box.
3. Click the **Histogram** radio button in the **Style** box.
4. Click either the **Number** or **Size** radio button in the **Bins** box and enter a value (optional).
 - **Number** sets how many bars will appear in the plot
 - **Size** sets the range of the values for each bar.

5. Select any numerical field from the drop-down menu for the X-coordinate.
6. Enter a name for the plot (optional).
7. Click **OK**. The high and low values of each bar appears under the plot.

Two-variable plots

In two variable plots, you choose both the X and the Y variables.

To create a two-variable plot:

1. Go to **View>ChemBioViz Plots>New**.
2. Click the **Two Variable** option in the **Dimension** box.
3. Select fields for both the X and Y axes.
4. Click the **Scatter plot** option in the **Style** box.
5. Enter a name for the plot (optional).
6. Click **OK**.

Removing a plot

To remove a plot:


1. Go to **View>ChemBioViz Plots>Remove**.
2. Select the plot in the dialog box and click **Remove**.
3. Click **OK**.

Filters

Filters can be used in two different ways:

- Filter a plotted variable to reduce its range
- Filter on other variables to add other dimensions to the plot.

To activate filters, go to **View>Filter Window** or use the context menu on a plot. When the filter window opens, right-click in it to select or deselect fields. There is a limit of 31 on the number of filters that can be applied to a plot.

 **Note:** *Filters act on all active plots, not just the one displayed.*

Autoscale

When you filter a plot, you may want to expand the scale to better display the remaining points. If you select **Autoscale** on the BioViz context menu, the scale will readjust automatically every time you adjust a filter. Autoscale is a toggle that remains in effect until you cancel it.

Alternately, you can make all of your filter adjustments without changing the scale, then select **Rescale to All Points** on the BioViz context menu.

Zoom

The BioViz zoom feature is a toggle activated from the BioViz context menu.

To expand a portion of a BioViz plot:

1. Select **Zoom on Drag** from the BioViz context menu. A check mark appears next to the command when it is selected. To deselect, click the command again.
2. Drag over a section of the plot.

When you release the mouse button, the scale expands to show the points selected.

While the initial view is the area you selected, you may use the scroll bars to view any part of the plot at the expanded scale. Clicking either the **X or Y Zoom** Button will return that axis to the original scale, while leaving the other axis at the expanded scale. Clicking **Unzoom** on the BioViz context menu will restore the original scale to both axes simultaneously.

You may repeat zoom, that is, zoom in on an area in an already expanded scale. If you have performed multiple zooms, the **Zoom** button acts as Undo control, reversing one step at a time.

Statistical analysis

You can perform statistical analyses on plots using the **Analysis** tab on the BioViz Plot Properties dialog box. The tab is available only for two-variable plots.

The following statistical parameters may be calculated:

- Minima and maxima of the X or Y variable.
- Mean or median of the X or Y variable.
- Standard deviation of the X or Y variable.

When you select curve fitting, the R squared value also becomes available for display. The calculated values appear on the plot.

You may also display a linear, quadratic, or cubic least-squares fit of the data points, with or without a confidence interval of one to three standard deviations.

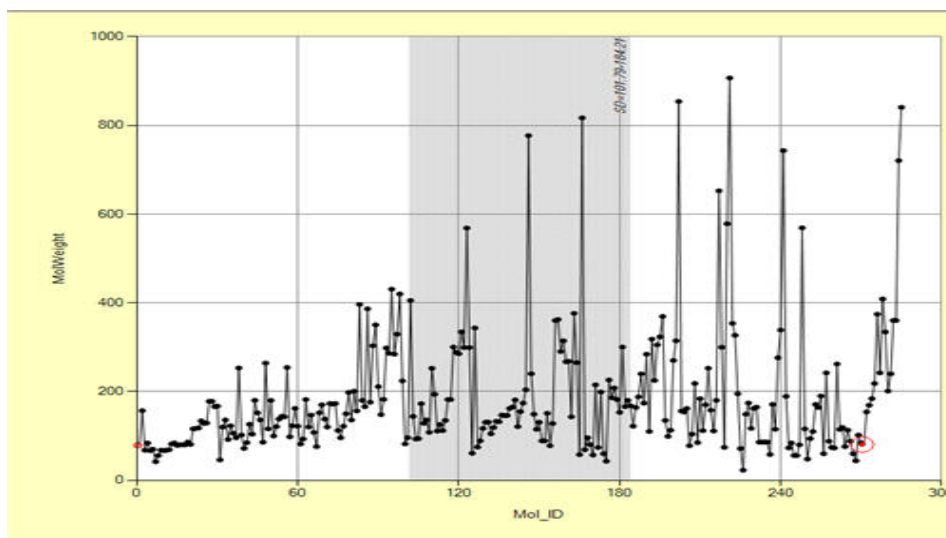


Figure 3.3: Plot with statistical analysis

The following table describes how to perform various functions in BioViz.

if you want to...	then	result
See which data point corresponds to a particular record...	navigate to the record using the ChemFinder Record Toolbar.	The data point is highlighted.
View the structure or data values of a data point...	mouse over the data point.	The Structure window displays the structure for the data point; the Details window displays the values of selected fields. (The current record in the form does not change.)
Select a database record from the plot...	click on a data point.	The form displays the matching data record.
Display plotted variable (s) for a query...	select (or create) the query.	The plot displays the current query.
Create a query from the plot...	On the BioViz context menu: 1. If a check mark appears next to Zoom on Drag, click to deselect it. 2. Drag over the points of interest to select them. 3. Click Selection to List on the context menu.	The selected points become a new query.
Prevent a plot from changing when you select queries or change variables...	check the Lock option in the BioViz Plot Properties dialog box.	The plot is locked to a given dataset. Note that filters will be locked as well.
Expand the plot to view individual data points...	Select Zoom on Drag on the BioViz context menu.	Dragging over a section of the plot enlarges it, rather than selecting the points.
Change the color of the data points...	In the Queries control, select Change Color on the Queries context menu. Note that you cannot change the color of the Full List,	The data points are displayed in the color you select.

if you want to...	then	result
	only of daughter lists.	
Plot multiple datasets (overlay)...	With the Full List selected, click Color on Plot on the Queries context menu for each query you want to overlay. alternately, click the query or the colored box next to it.	A check mark appears in the box next to the query name, and the data points are displayed in the color assigned to the query.

Tutorial 7: Working with subforms


Subforms allow you to display relational data. If you have a database containing two or more data tables, and these data tables share a common, linking field, then you can display data from both tables. Whenever the value of the linking field changes in the main form, the subform only displays those records from its table which have the same value in the linking field. You can also use subforms to display data from different databases.

In this tutorial, you open the CS_Demo database and display the MolTable in the main form. Then you display the Synonyms table of the CS_Demo database in a subform. By defining the MOL_ID field as the linking field, you display the two sets of data relationally.

Creating a subform

To create a subform:

1. Open your ChemBioFinder form tut1.cfx. Opening this form connects you to the CS_Demo.mdb database.
2. Make some space on the right side of the form by rearranging the data boxes and making them smaller.
3. Click the **Subform** tool.
4. In the form, drag to create a large subform.

 **Tip:** If you cannot make room on the right side of your form, you can put the subform at the bottom of the form.

When you release the mouse button the Subform Properties dialog box appears.

5. Click the **Form** tab, then select **Synonyms** from the hierarchical tree display.
6. Click the **Generate form** check box, then click **Style**.
7. Select the **Plain** form, deselect all of the fields to be included except **Synonym**, and click **OK**. You are returned to the Subform Properties dialog box.
8. Click the **Subform** tab.
9. **SYN_ID** already appears as the default in the **Link to SYNONYMS** field section. Using the drop-down menu, choose **MOL_ID** for the **Link from MOLTABLE** field.
10. Click **OK**.

You have just selected the data source for your subform and linked it to the main form. The MOL_ID field links the main form and subform. Clicking in either form activates it.

To test this:

1. Click anywhere outside the subform box, and browse your database using the **Record** tools. As each molecule record of the MolTable is displayed, the subform shows the first matching ID from the Synonyms table.
2. Click inside the subform box. Now when you browse, you are browsing only the entries in the subform.

To display all of the synonyms:

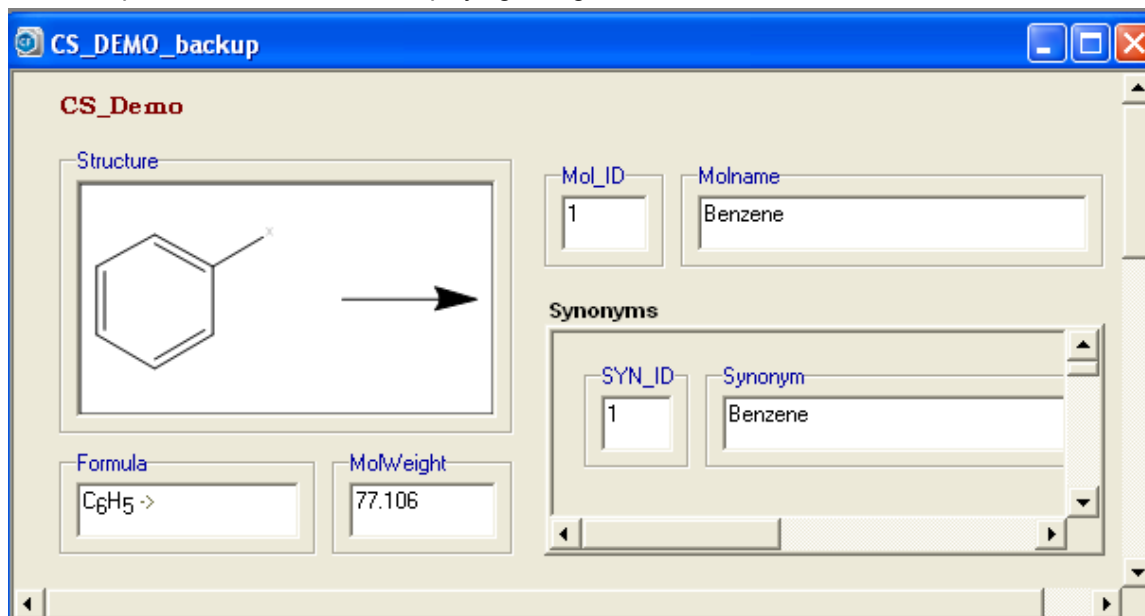
1. Double-click in the subform box to switch to **Table view**. Save and close the file:
 - a. Go to **File > Save**.
 - b. Go to **File > Close**.

Congratulations! You have completed the ChemBioFinder Tutorials.

Forms

You use forms to interact with information in a database. A ChemBioFinder form is composed of data boxes for viewing or modifying data items, such as structures, numbers, text, or pictures. A form can also contain subforms for relational access to different data tables and different databases.

The example below shows a form displaying a single record of information from a database.



The screenshot shows a window titled "CS_DEMO_backup" with a form titled "CS_Demo". The form contains several data boxes:

- Structure:** A chemical structure of benzene with an arrow pointing to the right.
- Mol_ID:** A text box containing the value "1".
- Molname:** A text box containing the value "Benzene".
- Formula:** A text box containing the value "C₆H₅ ->".
- MolWeight:** A text box containing the value "77.106".
- Synonyms:** A subform containing a table with two columns: "SYN_ID" and "Synonym". The table has one row with the values "1" and "Benzene".

Figure 4.1: A form showing a single record.

You can create a form any of three ways:

- Automatically, using the Form Generator dialog box.
- Manually, using the Form tools.
- With the Database Tree.

Creating a form consists of creating a layout (see ["Creating forms manually" on page 33](#)), and setting box properties (see ["Setting box properties" on page 39](#)), tabs (see ["Creating and editing tabs" on page 39](#)), and security (see ["Securing forms" on page 50](#)).


You can edit your form at any time (see ["Editing forms" on page 46](#) and ["Changing form layout" on page 50](#)).

To begin designing a form, you first select the database and deciding which fields to include.

Creating forms automatically

Use the Form Generation dialog box to create a form or change the layout of an existing form. You can choose pre-defined form style and properties.

To create a form:

1. Right-click a blank area of the current form and select **Data Source**.
2. In the Form Properties dialog box, select the **Form** tab.
 **Note:** If the **Generate form** check box and **Style** button are disabled, you must connect your current form to a database. For more information, see "[Connecting a database to a form](#)" on page 16.
3. Select the **Generate form** check box, then click **Style** . The Form Generation dialog box appears. In the left panel, all the fields are selected by default.
4. Deselect those fields in the **Choose fields to be included** section that you want to exclude from the form.
5. To have a Structure box in the upper left corner of the form, select the **Structure in upper left of form** check box. If you do not select this option, the boxes are generated in the order they appear in the list.
6. Select the **Form style** you want:
 - Framed-Each form field has a label above it and a frame.
 - Titled-Each form field has a label above and no frame.
 - Labeled-Each form field has a label to the left and no frame.
 - Plain-Each form field is displayed without a title or frame.
7. Use the **Columns** scrollbar to select the number of columns you want in your form.
8. Use the **Grid size** drop-down list to choose the size and spacing of the boxes and grid:
 - Large-larger and spaced further apart.
 - Medium-medium, relative to the Large and Small settings.
 - Small-smaller and spaced closer together.
9. Click **OK** in the Form Generation dialog box. Your form settings are saved.
10. Click **Apply** and then click **OK** in the Form Properties dialog box. A warning dialog box appears prompting you to create a new form or replace the existing form. Click **No** to create a new form, or **Yes** to replace the existing form.


Saving a form

In most non-database applications—including ChemBioDraw and ChemBio3D—when you edit data on the screen, your changes are not made permanent until you select **File>Save**, or type **CTRL+S**.

In database programs such as ChemBioFinder, your changes are automatically and permanently saved to the database when you switch records. This is known as “**committing**” the changes. You then have the opportunity to revert to the original data by selecting **Record>Undo Changes**.


The **File>Save** menu command refers only to changes made on the form itself, such as the position of boxes. Choosing the **File>Save** menu command has no effect on changes you make to data stored in the database.

 **Note:** Saving a form also saves subforms and changes that you make to subforms.

 **Note:** After you create a form, you can save it. When you retrieve the saved form, it automatically opens a connection to the database defined in the forms.

To save a form as .cfx file:

1. Go to **File>Save**. To save the form under a new file name, go to **File>Save As**. The Save dialog box appears.
2. Choose the directory in which you want to save the form, type a filename, and click **Save**. ChemBioFinder saves the form with a CFX extension.

 **Note:** While saving, if you select **All Files** in the **Save as type** dialog box, you will see files with **.msf** and **.msk** extensions. A file with the extension **.msf** is a multiple sequence file. A file with **.msk** extension is a mask file.

Creating forms manually

You create a form and define it by using the **Form** tools to create objects and the Box Properties dialog box to set the form properties.

To create a new form, do one of the following:

- Go to **File>New**.
- Click the New File icon on the Main Toolbar.

The Database Tree is the same tree as in the Database Properties dialog, showing the tables (or views) and columns of the currently-open database. When you open or activate a form or subform, or change databases, the tree updates to display the current database, and expands to show the columns of the currently-selected table.

To view the Database Tree (if it is not visible):

1. Go to **View>Explorer Window**.
2. Click the **Database** tab.

To use the Database Tree to build a form:

1. Open a new form.
2. Right-click on the form and select **Data Source**.
3. Click **Open Database...** on the **Database** tab.
4. Browse to a database, and click **Open**.
5. Click **OK** in the Form Properties dialog box. The Database Tree displays the database.

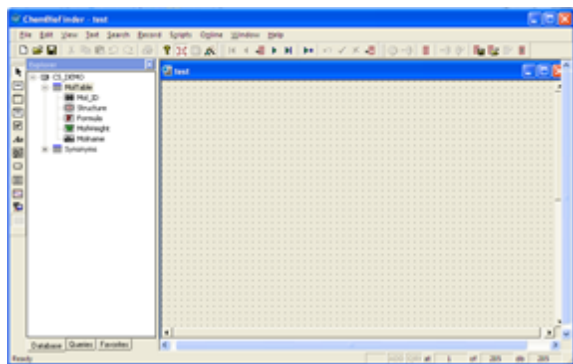


Figure 4.2: Creating a form with the database tree

6. Double-click the Structure field in the Database Tree. A structure box appears in the upper left of the form.

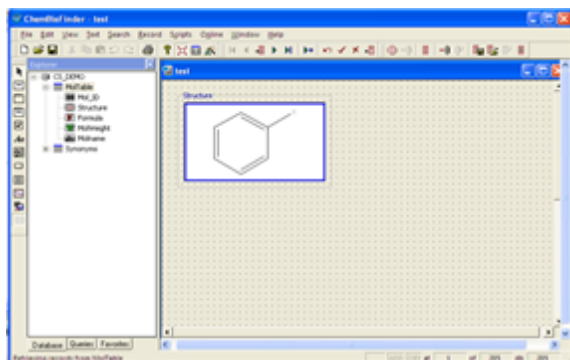


Figure 4.3: Structure box created from the database tree

7. Continue double-clicking the other fields.
8. Use the **Selection** tool to move the data boxes to where you want them in the form.

Using the form tools

You must use the Form Tools toolbar to design a form. Box creation commands are not available from the menu. To display the Form tools, do one of the following:

- Click the Layout icon on the Main toolbar.
- Go to **View>Toolbars>Form**. The Form toolbar appears providing tools for you to create and edit a form.

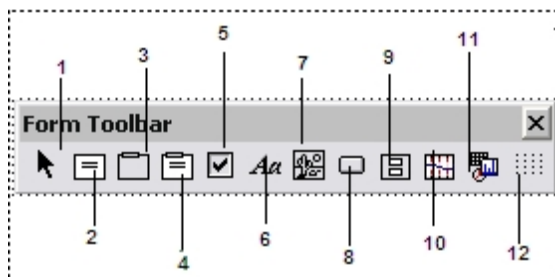


Figure 4.4: The Form toolbar: 1) Selection tool; 2) Data box; 3) Frame; 4) Framed box; 5) Check box; 6) Text box; 7) Picture; 8) Button; 9) Subform; 10) Profile box; 11) Control box; 12) Grid.

Using the grid

In a new form, the snap grid is turned on by default. This grid helps you align boxes on the form. Turning on the grid forces objects to snap to the grid as they are drawn.

To toggle the grid on or off, do one of the following:

- Click the **Grid** tool on the Form toolbar.
- Select/deselect the **View>Grid** menu option.

 **Note:** You can change the grid spacing using the Preferences dialog box. See ["Setting preferences" on page 199](#).

Creating boxes

A ChemBioFinder form is composed of a collection of boxes that display data from the associated database, or other information. Each box displays one data item of the current record. To create a data box, use one of the tools on the Form toolbar.

 **Note:** *The Form toolbar is not visible unless the **Layout** button is selected.*

The Data Box displays a data item or structure. Use it to display any data from a database, including text, numbers, dates, molecules, reactions, or pictures. It is the only box that allows you to edit data in the database. Data boxes do not have labels, but you can label them by adding a Frame or Plain Text. You can also use the Data Box to display static data not from a database.

To draw a Data Box:

1. Click the Data Box tool.
2. On the form, drag diagonally to create a box.

To specify what type of data is displayed in the box, see ["Setting box properties" on page 39](#). If the form already has a database open, you can right-click in the data box, select **Field...** and choose from a list of the available fields.

Creating frames

The Frame tool lets you create an enclosure with a label surrounding a data box or group of boxes. When you draw a new frame, you are automatically prompted for a text label to be attached to it.

To create a frame:

1. Click the Frame tool.
2. On the form, drag diagonally to create a frame. A framed box appears with the label **Data**.

To change the label:

3. Right-click on the label and choose **Label** from the drop-down menu. The Enter the label dialog box appears.
4. Type the label you want to attach to the frame.
5. Click **OK**. A fixed label appears on the frame.

Another way to change a frame's label is in the Box Properties dialog. For more information, see ["Setting box properties" on page 39](#).

You can use a frame to show "live" data from the database. If a frame is connected to a database field, it is updated automatically as you browse or search. For example, you might place a frame around a structure field, where the label on the frame represents the name of the current structure. The label changes every time the structure does. You must use Box Properties to set up a live data connection for a frame.

Creating boxes with frames

The Framed Box tool lets you create a data box and a frame simultaneously. You name the frame in the Enter the Label dialog box.

To create a framed data box:

1. Click the Framed Box tool.
2. On the form, drag to create a box.
3. In the Enter the Label dialog box, type the label you want to attach to the frame, and click **OK**. The framed data box is labeled.

To edit the label of a framed data box:

1. Select the frame by carefully right-clicking near the edge of the framed box and select **Label**.
2. In the Enter Box Label dialog box, type the new label. The label does not have to match the name of the field displayed in the box.
3. Click **OK**.

When you create a data box and frame with the Framed Box tool, these two items are connected. To manipulate them separately:

Select the Framed box, then go to **Edit>Bring to Front**.

Automatic labels

You can let ChemBioFinder automatically label your framed box with the name of the field from the data tree.

To automatically label a Framed Box:

- Right-click inside the data box and select the field from the drop down list. The data appears in the data box and the frame label changes to match the field name.

You cannot automatically label a frame that was created separately from a data box.

Non-data boxes

The Form toolbar contains two tools, **Plain text** and **Picture**, to create static displays (though they may also be used to display data) and three others with special functions.

Adding plain text

The **Plain Text** tool lets you display text without a box or a frame. You can use it to place a static label on the form, but, like a data box, you can also use it to display live data if desired.

To use the Plain Text tool:

1. Click the **Plain Text** tool.
2. On the form, drag to create an area to specify where the text will appear. The Enter the Text dialog box appears.
3. Type the text you want to display.
4. Click **OK**.

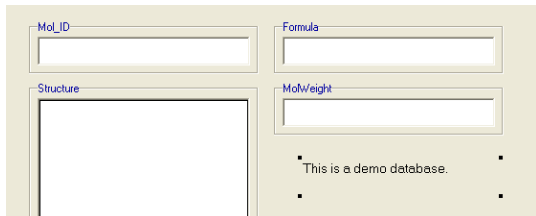


Figure 4.5: Creating text on a form

To change the font or color of the text, select it with the Form toolbar Selection tool. Use the text formatting toolbar to customize the text.

To use plain text to display data:


1. With the Select tool, right-click on the text and select **Properties**. The Box Properties dialog box appears with the **Box** tab displayed.
2. Select a field from the data table. The **Data**: text box becomes grayed out, and the **Data Source** selection switches to **Field**.
3. Click **OK**. The contents of the field are displayed in plain text.

Adding pictures


The **Picture** tool lets you create a Picture Box to display a bitmap (BMP, GIF, etc.) or Windows metafile (WMF, EMF) picture. It can display a static file, or a live one stored in a picture column of a database.

To create a static picture:

1. Click the **Picture** tool.
2. On the form, drag to create an area for the picture to occupy. The Open dialog appears.
3. In the Open dialog box, select the file you want to display, then click **Open**. The picture appears in your form.

 **Note:** The file is not actually embedded in the form. Only its filename is saved with the form. If you move or rename the file, the form will not display the corresponding picture. However, if you make changes to the picture and keep its filename the same, the picture in the form will be updated.

To create a Picture box linked to a field.


 **Note:** You must have a picture field in the database. See ["Creating fields" on page 71](#) for more information.

1. Click the **Picture** tool.
2. On the form, drag to create an area for the picture to occupy. The Open dialog appears.
3. Click **Cancel**. A blank box appears in the form.
4. Right-click in the box and click **Properties**.
5. In the Box Properties dialog box, select the picture field you want to link to the box and click **OK**. The picture field item appears in the box.

To read in a picture file or to replace a picture in a Picture Box:

1. Right-click a picture and click **Read File**. The Open dialog box appears.

2. In the Open dialog box, select the file you want to display in the Picture Data Box, then click **Open**. The new picture is displayed in the Picture Box.

 **Note:** If the Picture box is linked to a Picture field, you must click the Commit changes tool to save the picture in the database.

To save a file from a Picture Data Box:

1. Right-click on the Picture Box and select **Save File**.
2. In the Save As dialog box, type a file name, then click the **Save**. The picture is saved.

Adding a check box

The check box tool lets you add check boxes to your form. A check box can be assigned to either a Boolean or an Integer field.

To add a check box:

1. Click the **check box** tool.
2. On the form, drag an area large enough for the text next to the check box. The Enter the label dialog box appears.
3. Type in the text and click **OK**.
4. Right-click, and select a Boolean or Integer field from the field list to assign to the check box.

Adding a button

Buttons on the form are used to activate scripts.

To place a button on a form:

1. Write a script using the ChemBioFinder Automation Language (CAL).
2. Click the **Button** tool on the Form toolbar.
3. On the form, drag to draw a button. The Enter the Label dialog box appears.
4. Enter a label. The text can be descriptive ("Click to browse") and does not have to correspond to the name of the script.
5. Click **OK**.

For more information about using a button to run a script, see ["Using scripts" on page 218](#).

Adding a subform

The **Subform** tool is used to create a subform box to display data from a related data table. See ["Relational Data and Subforms" on page 155](#) for more information

Adding a profile

The **Profile box** tool creates a profile box on a form. Compound profiles are a new feature in ChemBioFinder. See ["Compound Profiles" on page 135](#) for more information.

Adding a control

The **Control box** tool is used to embed an ActiveX control in a form. You can use this to display a calendar, spreadsheet, chart, spectrum, picture, web link, and so forth. See ["Embedding ActiveX controls" on page 209](#) for more information.

Creating and editing tabs

If you have a large number of fields in a database, you can access information more easily, and make the form less cluttered, by using tabs to divide a long form into smaller parts contained on separate tabs.

To create tabs on a form:

- Right-click an empty part of a form and select **Add Tab context**. Two tabs appear in the lower left corner of the form. Clicking on the second tab will present you with a new empty form.

To rename a tab:

1. Right-click the tab and click **Rename Tab**. The Tab Name dialog box appears.
2. Type a name for the tab and click **OK**. The new name appears on the tab.


To remove a tab:

- Right-click the tab and click **Delete Tab context**. The tab is removed.

Setting box properties

You can set box properties such as data source, display type, font, and box style.

To set the Box properties, right-click a data box and select **Properties**.

 **Note:** The “Properties” option is available only in Layout mode.

Setting data box styles

You can change the style that is applied to a data box on a form. For example, you can change a data box to a frame or picture.

Choose from these styles:

- Data Box—displays alphanumeric data. Can display multiple lines of data.
- Frame—displays alphanumeric data. Can display a single line of data.
- Plain Text—displays alphanumeric data. Text data can be displayed in multiple lines. (Use CTRL+ENTER to start a new line.) You set the field size when you create the field, up to a maximum of 254 characters.
- Picture—displays Windows metafile data. Pictures are scaled to fit within picture boxes, with fixed height to width ratios.
- Formula—displays any kind of text but numbers are subscripted. Formulas are presented in modified Hill order, as follows: If a substance contains both carbon and hydrogen, the carbons are listed first, followed by the hydrogen atoms, followed by other elements alphabetically by element symbol. If a substance contains carbon, the carbons are listed first, followed by other elements alphabetically by element symbol. Otherwise, all elements are listed alphabetically by element symbol.
- Button—used to create buttons to actuate scripts. Buttons display button labels.
- Rich Text—displays styled data in multiple lines. (press CTRL+ENTER to start a new line.) You can format the text with the Text Format toolbar. If the data in a field contains rich text, markup characters are displayed.
- Structure (*Chem & Bio Draw* style)—displays chemical structures or reactions. Reactions are laid out for best fit with the box and the layout may change if the box is reshaped. The box defaults to using the ChemBioDraw Control to

- create and edit structures, but you can reset the preference. For more information, ["Setting preferences" on page 199](#)
- Structure (ChemBioFinder style)—as above, but defaults to opening *Chem & Bio Draw* to create and edit structures.
- Structure (*Chem & Bio 3D* style)—as above, but defaults to display 3D structures.

To set the box style:

1. Right-click in the data box you want to change and click **Properties**. The Box Properties dialog box appears.
2. Choose the box style from the **Box Style** drop-down list.
3. Click **OK**.

Viewing structures

You can display structures in a structure box in three styles:

- ChemBioFinder
- ChemBioDraw
- ChemBio3D

In ChemBioFinder style, the ChemBioDraw application opens when you want to edit structures or create a query. (This mode of operation was used in all earlier versions of ChemBioFinder 11.0.)

In ChemBioDraw style, you edit directly in the structure box using the ChemBioDraw Control.

The ChemBio3D style is a display-only format that lets you rotate and analyze structures within a ChemBioFinder structure window. See the ChemBio3D User's Manual for information on using the ChemBio3D Control.

The ChemBioDraw format displays more graphically-rich drawings than the ChemBioFinder format. However, ChemBioFinder does not use all of the ChemBioDraw drawing features when searching. For example, textual annotations are ignored, and R-group tables are not recognized. Certain graphics, such as rectangles and orbitals, are not transferred, nor does ChemBioFinder consider their chemical implications.

To choose how to view structures:

1. Right-click in a structure data box and choose **Properties...** The Box Properties dialog box appears.
2. On the **Box** tab, in the **Box Style** section, select the structure format to view from the menu:

If you want to	Then choose
edit in the Structure box using the ChemBioDraw Control	Structure (ChemBioDraw style).
edit in ChemBioDraw	Structure (ChemBioFinder style).

If you want to	Then choose
insert 3D images	Structure (<i>ChemBio3D</i> style).

Viewing structures in ChemBio3D format

Unlike the ChemBioDraw and styles, ChemBio3D style is view only. You cannot edit the structures in ChemBio3D style.

To activate the ChemBio3D Control:

- Double click in the 3D Structure Box. Click outside the box to deactivate the control.

To rotate the structure:

- Drag in any direction in the box. Unlike in ChemBio3D, you do not have pure (locked) X- or Y- axis rotation, but you can approximate them by dragging horizontally or vertically.

The ChemBio3D Control displays different context menus, depending on whether you right-click while pointing to the model or while pointing elsewhere in the Structure Box. Either can be used to change the display mode.

To change the display, method one:

1. Double-click in the 3D Structure Box to activate the ChemBio3D display.
2. Point to an atom or bond in the model and Right-click.
3. Select a display mode from the **Display Mode** submenu.

Method two:

1. Double click in the 3D Structure Box to activate the ChemBio3D display.
2. Right-click anywhere in the 3D Structure Box, but not on the model itself.
3. Select a display mode from the **View>Model Display>Display Mode** submenu.

For more information on the ChemBio3D Control, see the ChemBio3D User's Manual.

Setting fixed and live data

In the Box Properties dialog box, you can designate whether the data source is fixed or from a field. Fixed data is attached to the form and does not change as you browse. Data from a field is stored in the database and is different for each record. As you browse, the record display updates according to the contents of the field.

To specify live data for a data box:

1. In the **Box** tab, select the field to associate, and then click the **Field** radio button.
2. Click **OK**. The box display updates as you change records.

To specify fixed data for a data box:

1. In the **Box** tab, select the field to associate, and click the **Fixed** radio button.

2. Type a label in the text box. If you choose a structure or picture field, click the browse tool to browse to the file to display.
3. Click **OK**. The box display is fixed.

Adding a data box menu

You can add a menu containing a list of choices to a data box so you can choose an item from it to appear in the data box.

To add a menu:

1. Right-click on a data box and click **Properties**. The Box Properties dialog box appears.
2. On the **Box** tab, click the **With choices or auto-link** check box. The **Choices** tab appears.
3. Take the appropriate action:

If you want to	Then click
use a list from a database table as the data box menu	From table and follow the instructions at the bottom of the dialog box.
create your own list for the data box menu	Fixed list of choices and follow the instructions at the bottom of the dialog box.
use a list from a database table and add your own items to the data box menu	<ol style="list-style-type: none"> a. From table. A list appears. b. Click Fixed list of choices, and edit the list.

4. Click **OK** to close the dialog box, then click in the data box. A button appears in the lower right corner of the box when you click in the box.
5. Click the button to display the menu.
6. Select any item on the menu—the text appears in the top left corner of the box and the button disappears. Click within the data box to make the button reappear.

Adding scroll bars

Scroll bars are used to accommodate lengthy data when there is limited space for the data box.

To add a scroll bar:

1. Right-click a data box and click **Properties**. The box properties dialog box appears.
2. In the **Box style** section, select the **With scroll bars** check box.
3. Click **OK**. The scroll bar appears on the bottom right in the data box.

Hiding data boxes

You can show or hide boxes when switching between query and browse mode. For example, you can hide non-searchable data boxes during query entry, or show only those data boxes that display information useful for query entry. When you are in form layout mode, all boxes are visible, regardless of the setting.

To hide data boxes:

1. Right-click on the box and click **Properties**.
2. In the **Box Style** area, select the **Hidden in Browse Mode** check box or the **Hidden in Query Mode** check box.
3. Click **OK**.
4. Deselect the **Layout Mode** to view the results.

Customizing text

Use the Text Format toolbar to customize text fonts, sizes, styles, colors, and alignment in form boxes. All of the standard text formatting options are available.

To display the Text Format toolbar:

- Go to **View>Toolbars>Text Format**. The Text Format toolbar appears.

 **Note:** Text copied from a word processor such as Microsoft Word retains its styles when pasted into a Rich Text box.

The text format toolbar is active:

- When you are editing a memo field in a rich text box. In this case, font changes apply only to the current selection within the text.
- When you are editing in a regular data (or structure) box. Font changes apply to the entire contents of the box.
- When you are in form edit mode, with the selector tool active, and one or more boxes are selected. Font changes apply to all selected boxes (except frames around data boxes).

Customizing fonts

All labels on ChemBioFinder objects can be edited to your specifications. You can edit any label's font, font style, size, and color.


To set the font for a label:

1. Click on the label to select it.
2. Right-click and select **Properties**.
3. On the **Box** tab of the Box Properties dialog box, click **Font**.

The Font dialog box appears.

4. Set the text font, style, size, and color. The sample area previews the font.

5. Click **OK** to commit the changes.

 **Note:** To change the color of the atom labels in the structure box, you must use the Periodic Table. "[Periodic table](#)" on page 208

Customizing numbers

Using the Box Properties dialog box, you can specify how numeric data is displayed in forms you can customize the following properties:

- Currency symbol
- Decimal position
- Scientific notation

To specify the numeric format:

1. Right-click in a data box containing numeric data and click **Properties**. The Box properties dialog box appears.
2. Click **Format**. The Numeric Format dialog box appears.

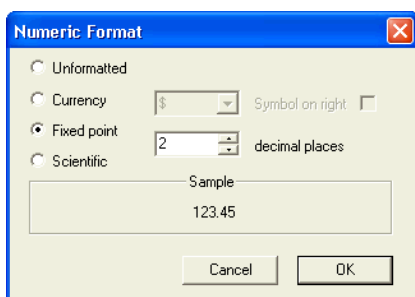


Figure 4.6: The Numeric Format dialog box

3. Select the appropriate option:

If you want to display	Then click
a varying number of decimal places	Unformatted.
a currency symbol	Currency and select the symbol from the drop-down menu.
the currency symbol to the right of the number	Symbol on right when a currency symbol is selected.

If you want to display	Then click
a standard floating-point value with a selected number of decimal places	Fixed point and select the number of decimal places.
the numbers displayed in scientific notation	Scientific and select the number of decimal places.

4. Click **OK**. The format is applied to the current data box.

Setting color

You can set the following color options:

- The form background color
- The color of the form in query mode
- The default background for all forms

To set the form background and form query colors:

1. Right-click in a data box and click **Properties**. The Box Properties dialog box appears.
2. Click the **Form** tab to display the Form properties.

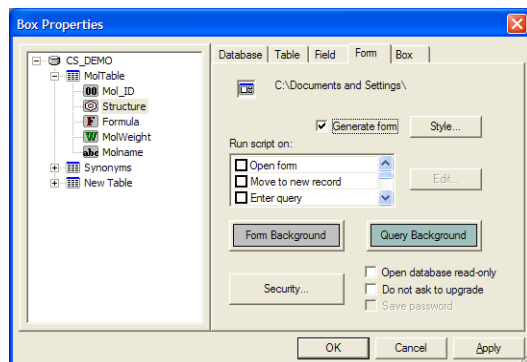


Figure 4.7: Form properties dialog box

3. Take the appropriate action:

To set the background color of the...	Click...
database form	the Form Background .
query form	the Query Background .

The Color dialog box appears.

Figure 4.8: The Color picker dialog box

4. Select a color.
5. Click **OK**. The **Form or Query Background** changes to reflect the color you choose.

To set the default background color:

1. Go to **File>Preferences**. The Preferences dialog Box appears.
2. Click the **Color** tab.
3. On the **Color** tab, click **Form Background**.
4. In the Color dialog box, select a color and click **OK**. The color you choose becomes the default form background color.

Editing forms

You can move or resize boxes and other objects on your form with the Selection tool and the items in the Edit menu.

Selecting objects on a form

Use the **Selection** tool to choose data boxes and other objects on a form that you want to move, resize, or delete.

When the selection tool is active, the Edit menu commands apply to the boxes themselves. When the selection tool is not active, the Edit Menu commands apply to the contents of the active data box.

To select an object on a form:

- Click the **Selection** tool on the Form toolbar, then click an object on the form to select it. A highlight appears around the object indicating that it is selected.

To select multiple objects on a form:

1. Click the **Selection** tool on the Form toolbar.
2. Do one of the following:
 - Press the SHIFT key and click on the multiple objects.
 - Drag the **Selection** tool around the boxes you want to select.

To select all the objects in the form:

- Go to **Edit>Select All**.

Moving objects

Use the Selection tool to move data boxes or other objects to a different place on a form.

To move objects on a form:

1. Click the **Selection** tool on the Form toolbar.
2. Select the object to move.
3. Drag the object to a different place.

 **Note:** To move a subform, you must drag the title bar of the subform, not its contents.


If an object is nearly in the right place, you can “nudge” it to get it exactly aligned.

To move an object one pixel at a time:

1. Select the object you want to move.
2. Press an arrow key in the direction you want to move the objects. The objects move one pixel in the direction of the arrow key.

Resizing objects

Use the **Selection** tool to resize a data box or any other object on the form.

 **Note:** You cannot resize more than one object at a time or simultaneously resize a box and its frame, unless you create a framed box using the Framed Box tool. For more information, see ["Creating boxes with frames" on page 35](#).

To resize objects on a form:

1. Click the **Selection** tool on the Form toolbar.
2. Position the pointer over an edge or corner of the object until the pointer is a double-headed arrow.
3. Drag in the direction you want to resize.

If an object is nearly the right size, you can make small adjustments to get it exactly the size you want.

To resize objects one pixel at a time:

1. Select the object you want to resize.
2. Place the mouse pointer over an edge or corner of the object until it is a double-headed arrow.
3. Press an arrow key in the direction you want to resize the object. The edge of the object moves one pixel in the direction of the arrow key.


Deleting objects

Use the **Selection** tool to remove data boxes and other objects from a form.

To remove objects from a form:

1. Click the **Selection** tool on the Form toolbar.

2. Select the object you want to delete.

 **Tip:** Use **SHIFT+click** to select multiple objects.

3. Do one of the following:

- Go to **Edit>Clear**.
- Use the Backspace key.

The object is removed from the form.

Reversing and restoring changes

Use the Main Toolbar icons or the Edit menu to reverse or restore recent changes to a form.

To reverse or restore changes:

- Go to **Edit>Undo** to reverse a change., or click the Undo icon.
- Go to **Edit>Redo**. to reverse a change., or click the Redo icon.

 **Note:** **Undo** and **Redo** work over multiple changes. Either command can be used repeatedly to reverse multi-step operations.

Ordering objects

The **Bring to Front** and **Send to Back** commands place data boxes in a specific sequence. They have two effects:

- Select which overlapping data box is on top.
- Set the order in which the cursor moves when you press the Tab key.

To put a box on top or bottom:

1. Click the box to select it.
2. Do one of the following:
 - Go to **Edit>Bring to Front** to put it on top, or
 - Go to **Edit>Send to Back** to put it on the bottom

The order in which data boxes are created is the order in which the cursor moves as you press the Tab key. You can set which data box the cursor moves to first or last.

To set the cursor to move to a data box first:


1. Select the data box.
2. Go to **Edit>Send to Back**.
3. Press the **Tab** key. The cursor moves to the box you set.

To set the cursor to move to a data box last:

1. Select the data box.
2. Go to **Edit>Bring to Front**. The box you set is the last box the cursor moves to when you press Tab repeatedly.

Aligning and distributing objects

You can line up objects relative to each other by aligning them and space objects at an equal distance apart by distributing them.

 **Note:** In framed dialog boxes, the inner data box is aligned, not the frame.

To line up objects on a form:

1. Press SHIFT and click the objects to align.
2. Use the **Edit>Align** submenu to take the appropriate action:

If you want to line up along the	Then choose
top edge of the first object you place on the form	Top.
middle of an object in a vertical line	Vertical Center.
bottom edge of an object	Bottom.
left edge of an object	Left.
middle of an object in a horizontal line	Horizontal Center.
right edge of an object	Right.

The form objects line up according to your choice.

To space objects on a form evenly:

1. Press SHIFT and click the objects to distribute.
2. Use the **Edit>Distribute** submenu to take the appropriate action:

If you want to space objects evenly between the	Then choose
outer boxes of the form	Vertically.
left and right of the form	Horizontally.

The form objects are spaced according to your choice.

Changing form layout

Changing a form layout is the same as automatically creating a new form. When you have created a new form, you can use it to replace the old form or save it as an alternate form. ["Creating forms automatically" on page 31](#) for details.

Securing forms


You can control the options available to users of your forms by setting the security options.

You can also provide the database connection information used to log on to an MS Access database. MS Access provides a security system that allows the creation and management of usernames and passwords, and the assignment of permissions to those usernames. MS username password account information is stored in the Workgroup Information file (.mdw, .mda). For more information about securing an MS Access database, ["Opening a secured MS Access database" on page 60](#)

Setting security options

To set what ChemBioFinder form options are available to users:

1. Right-click in the form you want to secure, and click the **Properties**. The Box Properties dialog box appears, with the Form tab displayed.
2. Click **Security**. The Form Security dialog box appears.

 **Note:** The **Security** button does not appear if the security on your form has been set to not allow access to the Security options.

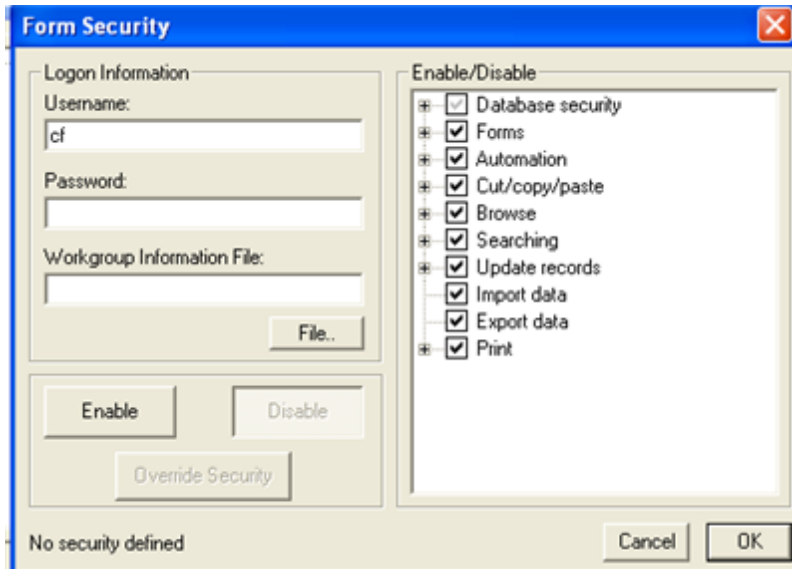



Figure 4.9: The Form Security dialog box

 **Tip:** The lower left corner of the dialog box indicates whether security is defined.

3. Type the **Username** to use for logon in the Username box. The **Enable** button becomes available.
4. Type the **Password** to use for logon in the Password box.

 **Note:** Use of a password is optional.

5. To enforce Workgroup Security in an MS Access Workgroup Administrator database you must specify a Workgroup Information File. Take the appropriate action:

 **Note:** Enforcing workgroup security is optional.

If you want to	Then
enter a workgroup file name	Type the name of the MS Access Workgroup database in the Workgroup Information File box.
browse for the workgroup file	Click File and select the file.

The following steps refer to the options in the **Enable/Disable** section of the dialog box.

1. Click the plus sign (+) next to **Database Security** to display the options; take the appropriate action:


If you want users to...	then select...
be prompted to log on with a username password to open a form	Password to open form
use a username password to log on to the MS Access database (MDB files)	Logon to MS Access database
be prompted to log on to a database with a username password when the form is open	Prompt for database logon
use a username password to open the molecule database (MST files)	Protect molecule database
be able to access the Form Security dialog box	Security dialog available.
encounter all of the above options	Database security.

2. Click the plus sign (+) next to **Forms** to display the options; take the appropriate action:

If you want users to...	Then select...
open or create databases	Open/create database
create forms	Create new

If you want users to...	Then select...
	forms
open forms	Open other forms
edit forms	Save changed forms
change the layout of a form	Change form layout
use all of the above options	Forms

3. Click the plus sign (+) next to **Automation** to display the options; take the appropriate action:

If you want users to	Then click
access OLE automation writing and programming	Allow CAL/OLE Automation access  Note: <i>Must be checked to enable export to Excel</i>
use scripts	Can use scripts menu
edit scripts	Edit scripts
use all of the above options	Automation

4. Click the plus sign (+) next to **Edit** to display the options; take the appropriate action:

If you want users to	Then click
clear forms	Clear

If you want users to	Then click
delete relational and structural data	Cut
copy relational and structural data and multiple table rows to the clipboard	Copy
paste relational and structural data from the clipboard	Paste
Perform all of the above functions	Cut/copy/paste

5. Click the plus sign (+) next to **Browse** to display the options; take the appropriate action:

If you want users to	Then click
browse databases	Browse database records
view data in forms in a table	View form as table
view data in continuous view	View in continuous forms view
edit and view structures in ChemBioDraw	Edit/view structures in ChemBioDraw
view structures in ChemBio3D	View structures in ChemBio3D
perform all of the above options	Browse

6. Click the plus sign (+) next to **Search** to display the options; take the appropriate action:

If you want users to	Then click
submit queries	Query database records
use current hit list for searching	Search over current list
search the entire database	Search over full database
save and manipulate hit lists	Enable hit list tools
search exact structure only	Full structure
search substructures only	Substructure
search similar structures only	Similarity
use all of the above options	Searching

7. Click the plus sign (+) next to **Update records** to display the options; take the appropriate action:

If you want users to be able to	Then click
update database records	Update records
add database records	Add new records
delete database records	Delete records

8. To allow users to import data SDFiles or RDFiles select **Import data**.

9. To allow users to export data, including delimited text files select **Export data**

10. Click the plus sign (+) next to **Print** to display the options; take the appropriate action:

If you want users to	Then click
print a single record only.	Single record
Print several records	Several records
Use print preview mode	Print preview
perform all of the above functions	Print


To complete setting security:

1. Click **Enable**.
2. Click **OK**. The options you choose are applied to the form.

Disabling security

To disable security and reset defaults:

1. Right-click in the form and click **Properties**.
2. In the **Form** tab of the Form Properties dialog box, click **Security**. The Form Security dialog box appears.
3. In the Form Security dialog box, click **Disable**. The Validate Security dialog box appears.

 **Note:** This form only appears if you entered a password when you enabled security. ["Setting security options" on page 50](#)


4. Enter your password, then click **OK**. Security is disabled and the defaults are reset. "Security Disabled" appears in the lower left corner of the Form Security dialog box.

Overriding security

In order to edit security options, you must temporarily remove security. In this case, you should use the Override option rather than the Disable option.

To temporarily disable security:

1. Right-click in the form and click **Properties**.
2. In the **Form** tab of the Form Properties dialog box, click **Security**. The Form Security dialog box appears.
3. In the Form Security dialog box, click **Override Security**. The Validate Security dialog box appears.

 **Note:** This form only appears if you entered a password when you enabled security. ["Setting security options" on page 50](#)

4. Enter your password, then click **OK**. You are allowed to edit the security options. "Security Overridden" appears in the lower left corner of the Form Security dialog box.

5. When you have completed your changes, click **Override Security** to toggle security back on. "Security Enabled" appears in the lower left corner of the Form Security dialog box.

Databases

You work with data within ChemBioFinder by entering it into a database, editing it, and interfacing with ChemBioDraw and ChemBio3D.

You can perform all of these functions on data using the commands in the menus, or you can use the buttons in the Record toolbar for most of these functions.

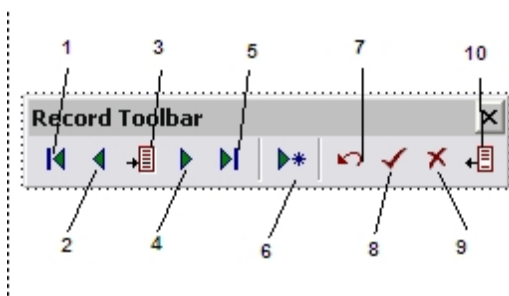


Figure 5.1: Database functions on the Record toolbar: 1) First Record; 2) Previous Record; 3) Go to Record; 4) Next Record; 5) Last Record; 6) Add a Record; 7) Undo Changes; 8) Commit Changes; 9) Delete Record; 10) Omit Record.

Selecting a database

A ChemBioFinder form does not store data directly, but is simply the window through which you look at data. When creating a form, you need to specify the source of the data to display. A form displays data from a single table in a database. To specify the data source for a form, you must open or create a database, and then select a table.

If you want to view data from more than one table, you must create a form for each table. Typically, you create a main form for one main table and a subform for each other table. For more information, see ["Relational Data and Subforms" on page 155](#).

You can connect a database to the form before creating boxes, or any time after. If the form is already connected to a database, you can change the database or the data source. If the form has boxes with fields connected to them, then the boxes automatically connect to fields of the same name in the newly-opened database.

Opening an existing chemical database

To open a database to associate with a form:

1. Right-click on the form and click **Data Source**. The Form Properties dialog box appears.
2. Click **Open Database**. The Open dialog box appears.

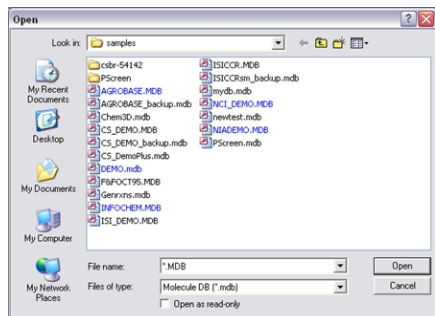


Figure 5.2: The Open dialog box

3. In the Open dialog box, choose a database with an .MDB file extension, then click **Open**. The Form Properties dialog box displays the data source tree. The CS_Demo database is shown below.

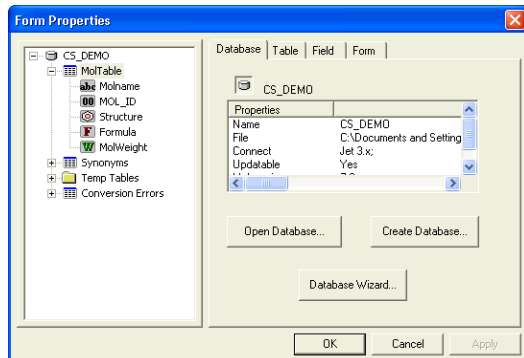


Figure 5.3: The CS Demo database Form properties

4. Click **OK**. The database is now associated with the form.

Selecting the data to display

You indicated the source of the data to display on the form, but the form is still blank. To see the data, you need to indicate what data to show in each box. You can use the Box Properties dialog box, as described in ["Setting box properties"](#) on page 39.

To choose the data box to display:

1. Right-click in a box or frame, and select **Field**. The Field properties appears.
2. From the data source tree, select the field to display.

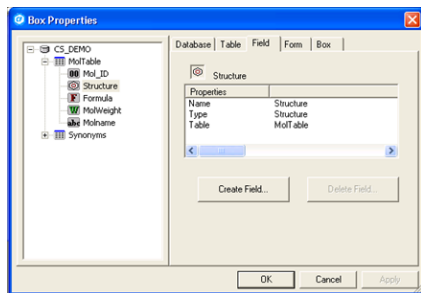


Figure 5.4: Selecting a field in the database tree

3. Click **OK**. The field data is displayed in the box.

If you use this method with a framed box, the box label will remain unchanged. If you want to use the field name as the box label, you can do it in one step as follows:

1. Right-click in a framed box and select the field name from the list at the bottom of the context menu. The frame receives the field name and the data is displayed in the box.

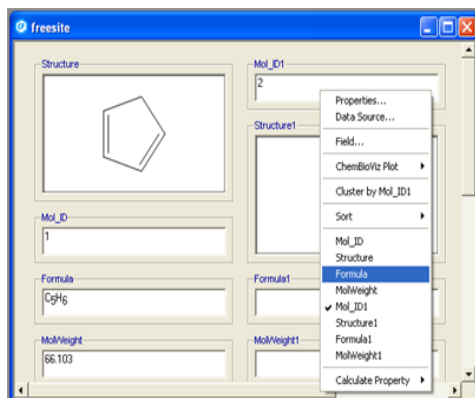


Figure 5.5: Selecting a field name from the context menu.

Opening a secured MS Access database

To open a secured Microsoft Access database in ChemBioFinder, you need to set the appropriate permissions in MS Access.

To set the permissions:

1. Open the desired database in MS Access.
2. Go to **Tools > Security > User and Group Permissions**.

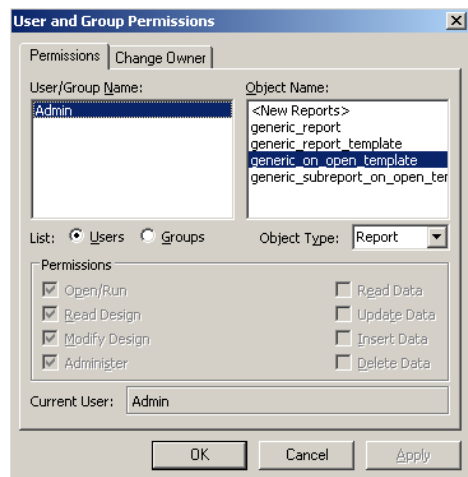


Figure 5.6: User and Group Permissions dialog box

3. Select the Users or Groups radio button to assign permissions to individuals or groups.
4. Select an Object Type from the drop-down menu.
5. Select the permissions you want to assign for that object type to each user or group.
6. Click **Apply**.
7. When you have finished assigning permissions, click **OK**.
8. Open the database in ChemBioFinder.

Creating a database

You can create a new, empty database and associate it with a form.

To create a database:

1. Open an existing form, or create a new one.
2. Right-click in the form and select **Properties**. The Box Properties dialog box appears.
3. Click the **Database** tab, and then click **Create Database....** The Save As dialog box appears.
4. Type a name for your database, and then click **Save**. A data source tree appears, containing the database and its tables and fields.
5. Click **OK**.

Opening databases

When you open a database, you can access the data in any of three ways:

- Normal access—you can read and write data.
- Read-Only access—you can read but not write data.
- Secured access—the database is password-protected and you are subject to the security restrictions applied to the database.

 **Note:** Versions of ChemBioFinder earlier than 11.0 cannot read .cfx files.

Read-only access

Any of the following conditions determine whether a database opens in read-only mode:

- If the component files (extensions CFX, MDB, LDB, MST, MSX) have read-only attributes.
- If the files are on read-only media, such as CD-ROM.
- If you select Read-only in the File Open dialog box when you open a form or database.

When you open a read-only database, the following conditions apply:

- READ appears in the status line.
- You can not modify data in text or structure boxes.

 **Note:** If only the form (CFX) file is set to read-only, READ does not appear on the status line and the database can be modified.

Multi-user access

If a database resides on a network, more than one ChemBioFinder user may access it at the same time. Each user can view, print, or modify records independently of the others.

The following table shows how the type of user affects the other users in the group.

User type	Affect on other users
Multiple readers in read-only access	User actions do not affect each other.
Multiple writers in normal access	All users can read and write. A user may see another user's edits. If two users try to edit the same area at the same time, one is alerted that the database is temporarily locked.
One writer, multiple readers	The writer, in normal access, is not affected by the readers. The readers, in read-only access, may see the writer's edits.

If two people are viewing the same record within a database, and one person changes the data in that record, the second person will not see the changes immediately. The changes will be visible when the second user switches to another record, and then back to it.

Secured access

ChemBioFinder supports the following types of security:

- Form access—ChemBioFinder enforces form access security by requiring a user name and password, which are stored in the form (CFX) file.
- Molecule file access—ChemBioFinder enforces molecule file access security by requiring a user name and password, which are stored in the database's MST file.
- Microsoft Jet Relational Database Engine—the database engine that underlies Microsoft Access enforces security. The user names and passwords are stored in an MS Access Workgroup Administrator database.

For more information on Form access, see ["Securing forms" on page 50](#). For more information on MS Access security, see ["Opening a secured MS Access database" on page 60](#).

When you open a form or database that has security options applied to it, you are prompted to enter a username and password to log on.

When you open a secured MS Access database from within ChemBioFinder, a username, password, and workgroup information file is required. Contact your Access Database Administrator for this information.

To open a secured form:

1. In the Validate Security dialog box, type your username and password, and then click **OK**.

If the form you access includes a secured molecule database, the Database Logon dialog box appears.

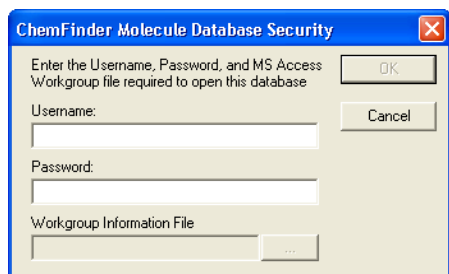


Figure 5.7: ChemBioFinder Molecule Database Security dialog box

2. Type your user name, Password, and Workgroup Information File name (if applicable), then click **OK**.

The form opens.

Browsing databases

You can browse a database using the buttons on the Record toolbar.

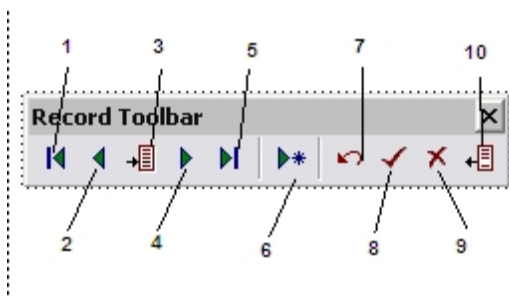


Figure 5.8: Browse functions on the Record toolbar: : 1) First Record; 2) Previous Record; 3) Go to Record; 4) Next Record; 5) Last Record; 6) Add a Record; 7) Undo Changes; 8) Commit Changes; 9) Delete Record; 10) Omit Record.


To Browse:

- Click **First Record** to display the first record in the database.
- Click **Previous Record** to display the record before the currently shown record.
- Click **Next Record** to display the record after the currently shown record.
- Click **Last Record** to display the last record in the database.

 **Note:** The current record number is displayed in the Status Bar.

Go directly to a specific record as follows:

1. Go to **Record>Go To Record**. The Record Number dialog box appears.
2. Type the number of the record (within the current list) to display, and click **OK**. The specified record is displayed.

 **Note:** Record numbers are temporary, referring only to positions within the current list in its current sort order. For example, if you have done a search that found ten hits, then the only valid record numbers are 1–10. If you want to move to a specific absolute location in a database, you must run a search for a value in a field that identifies that location.

The data table

Sometimes browsing through a data record set is more convenient if the data is presented in tabular format.

 **Note:** You cannot add or modify records while viewing data in a table.

To display data in a table:

- Go to **View>Data Table>In Separate Window** (or type **Ctrl+T**). A list window containing the fields and records in the form appears.

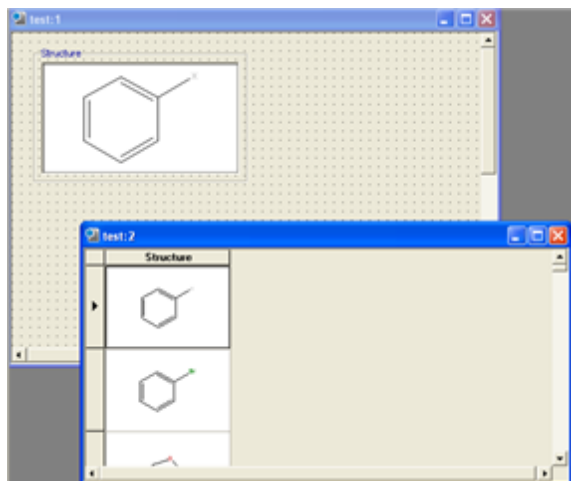


Figure 5.9: Data Table view in a separate window

To display a particular record in the form:

- Click the record's entry in the list window. By default, the fields in the data table are displayed in the order in which they were created.

To reorder the columns in the Data Table:

- Click one of the column headers to select the column.




MOL_ID	Structure	Formula	MolWeight	
1		C ₆ H ₆	78.11184	Benzene
2		C ₆ H ₅ Br	157.0079	Bromobenzene
3		C ₄ H ₄ O	68.07396	Furan

Figure 5.10: Column selection indicator in the Data Table view

- Drag the header to a new position. The cursor changes to the icon shown below. The new position is indicated with a red vertical line.



Figure 5.11: Moving a column

To resize the column widths in the Data Table:

1. Place the cursor on the dividers in the top header. The cursor changes to the icon shown below.


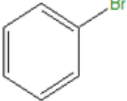
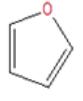

Structure	Mol_ID	Formula	MolWeight	↔	Molname	SYN_ID	
	1	C6H6	78.11184		Benzene	1 1 1 1	Benzene Phenyl hydri cyclohexatri (6)annulene
	2	C6H5Br	157.0079		Bromobenzene	2 2 2	Bromobenzene Phenyl brom Monobromo
	3	C4H4O	68.07396		Furan	3 3 3 3 3 3 3	Oxole Tetrole Divinylene c Furan Oxacyclope 1,4-epoxy-1,

Figure 5.12: Changing the width of a column

2. Drag to adjust the column width.

 **Note:** To hide the column, right-click and choose Hide Column from the context menu.

To resize the row heights in the Data Table:

1. Place the cursor on the dividers in the left header. The cursor changes to the icon shown below.

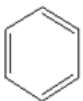
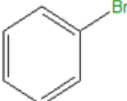

Structure	Mol_ID	Formula	MolWeight	Molname	SYN_ID	
	1	C6H6	78.11184	Benzene	1 1 1	Benzene Phenyl hydri cyclohexatri (6)annulene
	2	C6H5Br	157.0079	Bromobenzene	2 2 2	Bromobenzene Phenyl brom Monobromo
	3	C4H4O	68.07396	Furan	3 3 3 3 3 3 3	Oxole Tetrole Divinylene c Furan Oxacyclope 1,4-epoxy-1,

Figure 5.13: Changing the height of a row

2. Click-drag to adjust the row height.

 **Note:** To resize all rows, right-click in a column header and choose Resize Rows to Fit from the context menu.

R-group tables

In Data Table display, you can display a substructure search as table of R-group substituents.

1. Run a substructure search.


 **Note:** You cannot prepare a table from a full structure or similarity search.



Figure 5.14: Substructure query to set up R-Group table

2. Switch to Data Table view.

3. Right-click in the Structure column, and click **R-Group Table**. The following changes occur:

- the topmost structure is replaced with the template – similar to the query you used, but with Rs attached at all positions which found substituents.

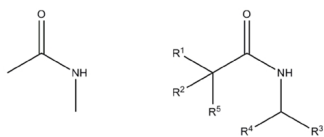




































































Figure 5.15: R-Group query and template

- new columns are generated for all the Rs.
- the other rows are populated with the Rs hit.

SYN_ID	Synonym (E name)	Structure	R2	R3	R4	R5
65	1H-benzole-1,3(2H)-dione					
65	1,3-benzodioxole					
65	1,3-benzodioxole					
65	formic acid					
65	1H-combendrone-1,3(2H)-dione					
69	Phenoxymethylpenicillanic acid					
69	2-phenoxymethylpenicillanic acid					
69	Phenoxymethylpenicillin					
113	5-Methyl-2,4(1H,3H)-pyrimidinedione					
113	Thymine					
113	2,4-Dihydroxy-5-methylpyrimidine					
113	5-Methyluracil					
113	2,4-dicy-5-methyl pyrimidine					
114	Uracil					
114	2,4-Dihydroxyuracil					
114	2,4(1H,3H)-pyrimidinedione					
153	2-Amino-1,7-dihydro-6H-purin-6-one					
153	Guanine					
153	2-amino-6-one-purine					
153	2-amino-1,7-dihydro-6H-purin-6-one					
153	2-Aminopurinomilone					
153	2-Amino-6-hydroxypurine					
155	2,4,6(1H,3H)-S(1H)-Thiazolidone					
155	Thiazole acid					
155	Methylthio					

 **Note:** The data in this table is just for display. It is not saved with the form. The table will continue to be displayed when you toggle the Table view until you change the hit list by running another query or restoring the full list.

To sort an R-Group Table:

- Right-click the header of any R column (R1, R2, etc.).
- Go to **Sort Column>[command]**. The options are:

- Ascending
- Descending
- Multiple

Ascending and Descending sort by the molecular weight. Multiple opens a dialog box that lets you sort by another parameter or group of parameters.

To set up a multiple sort:

1. Select fields in the order in which you want them to apply.
2. Right-click selected fields to change the sort order (ascending, descending) or to remove an unwanted parameter.

You can cut and paste R-group cells containing either structures or text using the clipboard.

R-Group plots

Once you've generated an R-group table, you can then plot the various substituent groups against any other plotable field.

1. With the table view still displayed, go to **View>BioVizPlots>New**.
2. On the General tab, click the X or Y variable to display the drop-down menu. The R-Group positions are now included in the menus, and may be used like any other numerical field to create a plot.

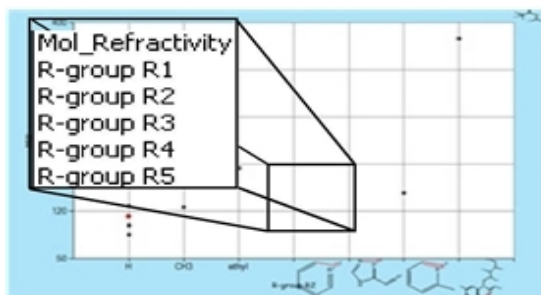


Figure 5.16: Setting up an R-Group plot.

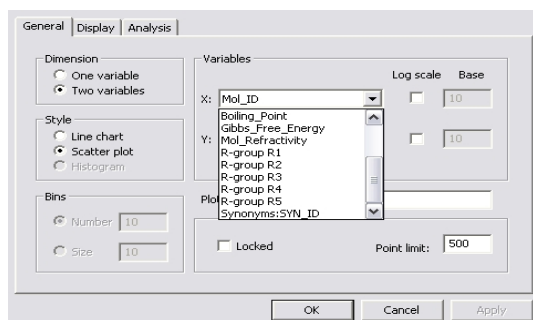


Figure 5.17: R-Group 2 vs. Molecular Weight

The R-Groups are along the X-axis as shown in the preceding figure. The structure on the upper right corner of the plot is the template.





Creating a database

Instead of opening an existing database, you may want to create a new, empty one.

To create a new database using the automatic form generation technique, see ["Tutorial 3: Creating a database" on page 17](#). The following procedure results in a simple database with one table for storing structures.

To create a database:

1. Right-click in the form and click **Properties**.
2. On the **Database** tab, click the **Create Database**. The Save As dialog box appears.
3. Choose the directory to which you want to save the database.
4. Specify a name for the database, and click **Save**. ChemBioFinder creates the database containing one table (called *MolTable*) with four fields:

-  Structure
-  Formula
-  Molecular weight
-  Mol_ID

The Mol_ID field corresponds to a column in the table where a numeric ID is automatically entered as each structure is registered. The other fields represent information stored in the structural portion of the database, linked to the assigned ID.

The Formula and Molecular weight data are automatically calculated from a structure. These internal fields cannot be edited or deleted.

After creating a database, you can create fields for storing other types of data. It is not necessary to create the entire set of fields before working with the database. You may add more fields later.

Creating tables

There is no limit to the number of tables a database can contain. You may want multiple tables in a database in order to manage relational data, or simply to organize different collections of information in the same place.

A table must contain at least one column. When you create a new table, ChemBioFinder creates a numeric column called ID. You can use this column to store integer data, or delete it and replace it with your own columns.

Table names must:

- Begin with a letter
- Not contain punctuation characters
- Not be the same as a table already on display

To create a new, empty table:


1. Right-click in the form and click **Properties**. The Box Properties dialog box appears.

2. Click the **Table** tab and then click **Create Table**. The Create Table dialog box appears.
3. Type a name for the new table, then click **OK**. The table appears in the Data Source tree.

Deleting tables

To delete a table from the database:

1. Right-click in the form and click **Properties**. The Box Properties dialog box appears.
2. Click the **Table** tab.
3. In the data source tree, select the table to delete.
4. Click **Delete Table**. A Warning message appears.
5. Click **Yes** to delete the table and **Cancel** to keep the table. The table and all data it contains are immediately deleted from the database.

 **Caution:** You cannot undo a deleted table. Before you delete a table, create a backup copy of the database to prevent accidental loss of data. For more information, see ["Backing up databases" on page 74](#).

Attaching tables from other applications

You can use ChemBioFinder to add chemical structures to a database you have already developed. If your database was developed in Microsoft Access, you can open it directly in ChemBioFinder.


You can also work with data from other file-based databases such as dBASE and xBase clones, but in this case you must open or create a ChemBioFinder database and attach the table to link to tables in the data source. After you attach a table, it appears and functions as if it were part of your local ChemBioFinder database.

The procedure you use depends on whether the tables you want to attach are from a file-based (Access, FoxPro, dBASE) or non file-based (Oracle) database system.

Attaching files from a file based database

To attach a file-based database table:

1. Right-click in the form and click **Properties**.
2. Click the **Table** tab, and then click **Attach Table**. The Attach Table dialog box appears.
3. Click **Open MS Access Database**. The Open dialog box appears.

 **Note:** The list of file types available in the Attach Table Open dialog box varies from one computer to another. The files for which you have drivers installed shown. For example, if Paradox is not installed, you may not have Paradox drivers on your system or in the drop-down list. If the database type you wish to access is not on the list, it may not be a file-based system, and you need to connect using ODBC. For detailed instructions, see ["Attaching files from a non file based database" on page 71](#).

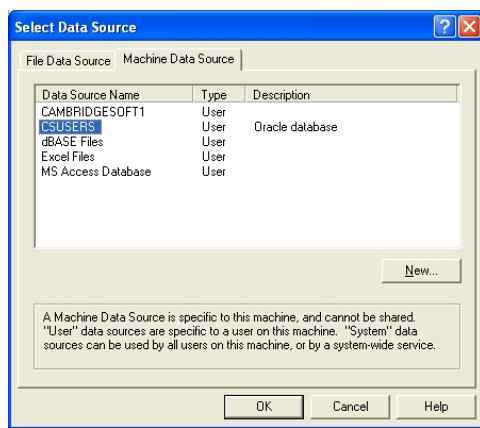
4. Choose the database to access, then click **OK**.
5. In the tree diagram, click the table you want to attach, then click **OK**. The newly-attached table appears in the tree diagram in the Form Properties dialog box.
6. Click **OK**.


Attaching files from a non file based database

If the database you want to attach is not file-based, such as Oracle, you can attach it using Microsoft's Open Database Connectivity (ODBC).

To attach a non-file-based database table:

1. Right-click in the form and click **Properties**.
2. Click the **Table** tab, then click **Attach Table**. The Attach Table dialog box appears.
3. Click **Open Oracle/ODBC Data Source**. The Select Data Source dialog box appears.
4. From the **Machine Data Source** tab, select the ODBC data source to access, then click **OK**.



 **Note:** The ODBC data source dialog box shows all data sources known to ODBC, including ChemBioFinder, Access, and other file-based data sources. If you attempt to open one of these through the ODBC dialog box, you get an error message. File-based data sources must be opened using Open Database, as described above.

If the ODBC data source dialog box does not show the database you seek, you may need to create a new data source. You can do this using **New** option. For details, click **Help** in the ODBC dialog box.


5. If you are prompted for a user name and password, enter it. If you do not know what name and password to use, see your System Administrator.
6. In the tree diagram, click the table you want to attach, then click **OK**. The newly-attached table appears in the tree diagram in the Form Properties dialog box.
7. Click **OK** in the Form Properties dialog box.

The source of the data to display on the form is indicated, but the form is still blank. To access the data, you need to draw some data boxes, and indicate what data to show in each of those boxes.

Creating fields


Whether you create a new database or are working in an existing one, you can add or remove fields in the selected table.


You can choose from the following types of fields:


 **Text field**—lets you enter text such as names, comments, and references. A text field is fixed in length and you choose the maximum length (“width”) for any data item to be stored in the field. If you enter a width of 50, then you can’t store any item with more than 50 characters in that field.


You can specify widths of text fields as large as 254 characters. If this is insufficient, you need to create a Memo/Rich Text field instead.

Because the text field width cannot be modified after a field is created, it is often wise to err on the side of caution and make it longer than you need initially. On the other hand, larger field widths also create larger files and slower search times.


 **Note:** *Non-ASCII characters will not display correctly in a data table text field.*

 **Integer field**—Used for whole numbers such as ID’s. All integers in ChemBioFinder are long, so can accommodate billions of values (2^{32} of them).

 **Double field**—Used for real numbers such as physical constants and unit prices. Real numbers in ChemBioFinder are double-precision. In ChemBioFinder, double fields can serve as ID fields.

 **Picture field**—Allows you to store a bitmap file (BMP, GIF, JPEG, PNG, or TIFF) or Windows metafile (WMF, EMF), such as a spectrum or experimental setup, as a data item in a database.

A Picture box on a form may be used to display a static picture attached to the form, such as a logo, or it may be used to show pictures stored in the database which change as you move from record to record. If you import a database containing pictures (type Long Binary or BLOB) they will be interpreted as pictures and the form generator will create a picture box for it.


 **Memo/Rich Text field**—Used to display text. Memo fields can be of any length. Because memo fields are less structured, searching them can be slower than searching a text field. Additionally, memo fields cannot be sorted. To search for text:

1. Right-click in a memo or plain text field and select **Find Text**. The **Find text** dialog box appears.
2. Enter the text and the search criteria.
3. Click either **Forward** or **Backward** to search.
4. To search for the next occurrence, do one of the following:
 - Click **Find Next**.
 - Use the F3 key.
5. To search for the previous occurrence, do one of the following:
 - Click **Find Prev**.
 - Use the SHIFT+F3 key.

To replace text:

1. Right-click in a memo or plain text field and select **Replace**. The **Find and Replace Text** dialog box appears.
2. In the **Text to find** text box, enter the text to be replaced.
3. In the **Replace with** text box, enter the text that will replace the text specified in the **Text to find** text box.

4. Click **Replace**.

 **Note:** *The Find Text option is available only if the length of the text is greater than the size of the text box and the entire text is not visible. If the entire text is visible in the text box, Find Text option will not be available in the context menu.*

Memo fields can store Styled Text. For more information, see ["Styled text" on page 82](#).



Date field—lets you store dates. The dates are displayed according to the settings in the Windows Regional Settings control panel.



Structure field—Consists of four fields: a numeric ID stored in the relational database, plus three fields (Structure, Formula, MolWeight) that take data from the ChemBioFinder structure database files.



Boolean—Used with check boxes. When searching Boolean fields, you can only search for ON. This is because ChemBioFinder automatically clears all fields when activating the Queries form.

You can create more than one set of structure fields in a table. Each is assigned a unique set of names, and each refers to its own ID column in the table, although all structural data is taken from the same structure database files. To create a field:


1. Right-click in the form and click **Properties**. The Box Properties dialog box appears.
2. Click the **Field** tab. The Field properties appear.
3. Click **Create Field**. The Create Field dialog box appears.
4. In the **Name** box, type a name.
5. In the **Type** box, choose a data type from the drop-down menu.
6. For text fields only, type a width (number of characters) in the **Width** box, or use the thumb wheels to select a value.
7. Click **OK**. The name of the new field appears in the data source tree.

Deleting fields

Just as you can create any field and assign it to a data box at any time, you can modify the database by deleting fields from the selected table.

To delete a field:

1. Right-click in the form and click **Properties**. The Box Properties dialog box appears.
2. Click the **Field** tab. The Field properties appear.
3. In the field list, select the field to be deleted.
4. Click **Delete Field**.
5. A warning message appears, click **OK** to delete or click **Cancel** to leave it unmodified.

 **Caution:** *When you delete a field, all data contained in the field is also deleted. You are not warned explicitly about this.*

Adding multiple structures

You can include more than one structure on the same form by creating multiple structure columns in a table. Each structure column you create represents four types of data:

- The structure displayed as a diagram.
- The Molecule ID, which connects the structure to a record in a relational table.
- The formula derived from the structure.
- The molecular weight derived from the structure.

To create multiple structure columns in a table:

1. Go to **File>Database**. The Form Properties dialog box appears.
2. Create or open a database.
3. Click the **Field** tab.
4. Select the Table in which you want to create new structure columns.
5. Click **Create Field**. The Create Field dialog box appears.
6. Select **Structure** from the **Type** drop-down list. An un-editable name is assigned to the field.
7. Click **OK**. Four new fields appear in the list and are named to belong to the same set of structure columns.
8. Connect the new fields to boxes by right-clicking on the appropriate data box and choosing the field from the context menu field list.

Non-chemical databases

If you open a database that you created using a program other than ChemBioFinder and you want to add structures to it, you can create structure fields in any modifiable (non-attached) table.

To create a structure field:

1. Open a database.
2. Right click and select **Properties**.
3. In the database tree, click the table to which you want to add structures.
4. Click the **Field** tab.
5. Click **Create Field**. The Create Field dialog box appears.
6. Choose **Structure** from the Type drop-down list, then click **OK**.

You do not provide a name or other details when you create Structure columns. ChemBioFinder automatically creates four columns and names them. These columns contain no data until you enter structures into the database.

Backing up databases


To back up the current ChemBioFinder database:

1. Go to **File>Copy Database**. A **Save As** dialog box appears. By default, your copy is named "YourDatabase_copy.cfx"

2. Rename your copy (if you wish), select a location, and click **OK**. A complete set of ChemBioFinder database files is created in the selected location.

Moving databases

ChemBioFinder saves only the definition of the form and information for connecting to the database in the CFX file. The actual data are stored in files with MSX and MST extensions for structure data, and in files with an MDB extension for non-structural data. If you want to move a database to another computer, you must move (for a default ChemBioFinder database) at least four separate files. This number might be greater if you have several forms that access the same database.

 **Note:** *The MSX, MST, and MDB files all have the same file name, but the name of the CFX files might be different depending on how you saved them.*

After you move the database to its new location, open your forms to make sure the data source links have been retained. If ChemBioFinder cannot locate the data source and displays an empty form, you will need to use the **File>Database** menu command to reconnect to the data source.

If the form is connected to a remote data source on a network, you have fewer files to move. The data source can remain on the remote machine, and you only move the CFX file that contains your form for accessing the data. You may need to reconnect it to the data source if it is not done automatically. This situation is common in large organizations where several people access the same central data source.

Creating a portal database

To carry out certain operations, such as sorting structural data, ChemBioFinder needs to create temporary tables in a database. This cannot be done in a read-only source such as a data CD.

Instead of accessing the read-only source directly, you can create a portal database—a local, writable database with attachments to the external tables of interest. The portal looks and behaves just like the target database, but without the limitations.

ChemBioFinder creates a portal database when needed. For example, if you are using a database CD, perform a search, and attempt to sort by formula, a message appears that offers to create a local database, attach the current table, save the form, and proceed with the sort. If you create the local database, you can use the new form to get the data with full functionality and performance.

Working with Data

Entering data

You enter data into a database by adding a new record. Adding a new record consists of three steps:

- [Clearing the form](#)
- [Adding new data](#)
- [Committing the new entries](#)

Clearing the form

To begin adding a new record:

1. Create a form and link it to a database. For more information, see ["Creating forms manually" on page 33](#).
The form should contain all the data boxes you want to view and edit. The boxes should be assigned to their appropriate fields.
2. Do one of the following:
 - Go to **Record>Add New Record**.
 - Click the Add Record icon on the Record Toolbar.
3. All of the boxes in the form are cleared to prepare for entering new data. The Status Bar is updated to show that you are in record addition mode:



Figure 6.1: Status bar in record addition mode.

Adding new data

To add alphanumeric data:

- Click in a box with an alphanumeric field and type the data.

To add a structure:

1. Right-click in a box with a structure field.
2. Do one of the following:

If you want to...	Then...
draw a structure, ChemBioDraw mode	<ol style="list-style-type: none"> 1. Double-click in the structure box. The ChemBioDraw Control appears. 2. Draw the new structure


If you want to...	Then...
	in the Structure box. 3. Click outside to box to complete the structure.
draw a structure, ChemBioFinder mode	1. Double-click in the structure box, or Right-click and select Edit in ChemBioDraw . ChemBioDraw opens. 2. Draw the new structure in the ChemBioDraw window. 3. Go to File>Exit and Return to , or type CTRL+W to return to ChemBioFinder and insert the structure.
import a structure, ChemBioFinder or ChemBioDraw mode	1. Right-click in the structure box, and click Read Structure . The Open dialog box appears. 2. Browse to a structure file, then click OK . The file is inserted into ChemBioFinder.

 **Note:** See the *ChemBioDraw User's Guide* for information about using ChemBioDraw.

Committing the new data

When you finish entering all of the data items, do one of the following:

- Click **Commit** on the Add Record Toolbar.
- Click the Commit Changes icon on the Record Toolbar.
- Go to **Record>Commit Changes**.
- Perform another action such as moving to another record or printing. The new record is added to the database.


 **Note:** Selecting *Commit Changes* (or moving to another record if you have two or more records) saves the data to the database. Do not use **File>Save**; clicking **File>Save** saves changes to the form layout, not changes to data in the database.

Duplicating records

You can create a new record by modifying an existing one.

To duplicate a record:

- Go to **Record>Duplicate Record**. You are in Add Mode and the form fills with data from the previously displayed record. When you commit the changes, you create a new record whose fields contain the data displayed. Before committing the duplicate, you can modify fields or structure and commit changes just as you would with any other new record.

 **Note:** *When duplicating records, only those fields that are visible on the form are duplicated. Data in fields present in the database but not visible on the form are not copied into the new record. The new record has a new Mol_ID.*

Undoing data entry

Before committing a new data entry, you can revert the contents of the form to its previous unmodified state.


To undo your changes, do one of the following:

- Go to **Record>Undo Changes**.
- Click the Undo Changes icon on the Record Toolbar.

After you commit the changes, they cannot be undone.

Editing data

Modifying the data in a database is performed by directly changing the data items on the form and committing those changes.

 **Note:** *The Formula and MolWeight fields are automatically calculated by ChemBioFinder from the Structure field and cannot be edited by the user. The MOL_ID field is also set automatically by ChemBioFinder and cannot be edited.*

To edit data:

- Click the data box whose data item you want to edit. If you click on a Structure data or a picture box it is highlighted. If you click on a data box containing alphanumeric data, a cursor appears in the data box.

To edit alphanumeric data:

1. Replace it with the text or number you want.
2. Do one of the following:
 - Go to **Record>Commit Changes**.
 - Click the Commit Changes icon.
 - Move to a different record.

 **Note:** *Moving to a different record always commits changes first.*

Editing structures

You can edit structures by using the ChemBioDraw drawing tools. You can also neaten the appearance of a structure by cleaning it up. The **Clean Structure** command is used to neaten the appearance of molecules by regularizing bond lengths and angles. Since the degree of change required cannot be determined a priori, the Clean Structure command begins gently. You may need to repeat the command to get the changes you wish. For more details about how the command works, see “Structure Clean Up” in Chapter 4: Advanced Drawing Techniques of the *ChemBioDraw User’s Guide*.

Before committing a data entry, you can revert to the previous unmodified record by choosing Undo Changes from the Record menu.

To edit structural data:

1. Do one of the following:
 - In ChemBioDraw mode: Double-click the structure box. The ChemBioDraw Control appears.
 - In ChemBioFinder mode: Double-click the structure box. ChemBioDraw opens and the structure appears in the ChemBioFinder window.
 - In either mode: Right-click in the structure box and click **Edit in ChemBioDraw**. ChemBioDraw opens and the structure appears in the ChemBioFinder window.
2. Edit the structure.
3. Click in the ChemBioFinder form or close the ChemBioDraw window by typing CTRL+W when you are finished.
4. Go to **Record>Commit Changes**. Click the Commit Changes icon, or move to a different record. The changes are stored.

To clean up a structure:


1. Right-click in the box containing the structure, and click **Clean Structure**.
2. To save the cleaned up structure, do one of the following:
 - Go to **Record>Commit Changes**.
 - Click the Commit Changes icon.

Working with structures using ChemBioDraw

Enter or edit ChemBioDraw structures in data boxes of the Structure field type. You can work in ChemBioDraw directly, or work in the Structure data box using the ChemBioDraw Control. You choose the default in the Structure data box Box Properties.

To set the preference:

1. Right-click in the Structure box.
2. Click **Properties**.
3. Use the drop-down menu in the **Box Style** section to choose your default:
 - ChemBioDraw style to use the ChemBioDraw Control.
 - ChemBioFinder style to edit in ChemBioDraw.

 **Tip:** When ChemBioDraw style is the default, you still have the option of editing directly in ChemBioDraw. Just right-click in the Structure box and choose **Edit in ChemBioDraw**.

To edit a structure:

1. Double-click in the structure data box. ChemBioDraw opens or the ChemBioDraw Control appears, depending on your default. If you already have a structure in a data box, that structure appears in the edit window. You can modify and manipulate the structure just like any other ChemBioDraw structure. If there is no structure in the Structure data box, the edit window is blank. You can draw a structure to store in the ChemBioFinder database.

When you have finished, do one of the following:

- If you are editing in ChemBioDraw, go to **File>Exit** and Return to Structure or type CTRL+W.
- If you are editing with the ChemBioDraw Control, click outside the Structure box.

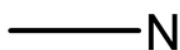
The new or edited structure is now displayed, but has not yet been added to the database.

2. Go to **Record>Commit Changes**, then click the Commit Changes icon, or move to a different record. The changes are stored.

With a few exceptions, you can store any chemically meaningful structure or reaction that can be drawn in ChemBioDraw. Within ChemBioDraw, you can confirm that a structure is chemically meaningful by selecting it and using the **Structure>Check Structure** menu command. For more information, see the *Chem & Bio Draw User's Guide*.

Two exceptions are:

- ChemBioFinder does not support importing molecules with multiple or variable points of attachment such as ferrocene.
- ChemBioFinder does not recognize "bare" heteroatoms. For example, if you draw the following structure:



ChemBioDraw reports an illegal valence when you choose the Check Structure command. ChemBioFinder automatically infers hydrogen atoms to main-group elements as necessary to fill their lowest acceptable valence. The structure above is registered in as methylamine, CH₃NH₂.

Structures and reactions drawn with query properties are generally meaningful only in the context of a query. Query structures can be stored in a ChemBioFinder database, but they are not treated as Markush structures and are not guaranteed to be hit by all valid search queries. For more information, see ["Structural Query Features" on page 222](#). *Chem & Bio Draw* allows you to draw many objects that have no chemical meaning. These include boxes, circles, arrows, orbitals, and others. It also allows you to assign non-chemical styles (color) to objects that have chemical meaning. ChemBioFinder ignores these properties, and stores only objects with chemical meaning in structure fields.

To store a ChemBioDraw drawing exactly as drawn, store it as a picture. You can copy and paste a ChemBioDraw drawing into a Picture field.

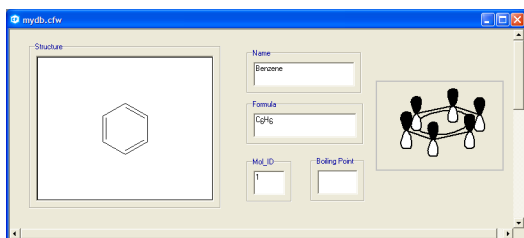



Figure 6.2: Form with picture field

 **Note:** Objects stored in a Picture field have no chemical significance and cannot be searched.

Viewing models using ChemBio3D

ChemBioFinder also provides access to the ChemBio3D Control. This means a Structure data box can be designated as ChemBio3D style.

To view the molecular model for a structure:

1. In the **Database** pane of the Explorer window, double-click the **Structure** field to add another Structure field to the form.
2. Right-click in the field and click **Properties**.
3. In the **Box Style** section, select **Structure (ChemBio3D style)** from the drop-down menu. The Structure field displays a 3D structure.

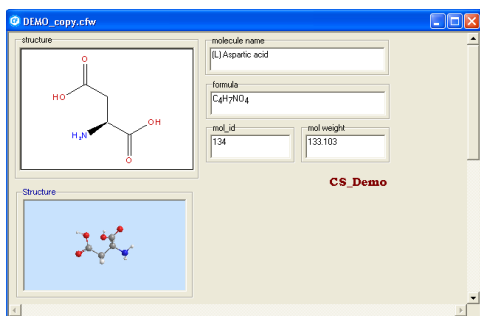



Figure 6.3: Form with 3D structure field.

To rotate the model or modify the display, double-click in the field. The default tool is the rotation tool, but it also can be used to select atoms or bonds. If you right-click in the field, the context menu displayed is the ChemBio3D control context menu, giving you access to ChemBio3D commands.

See the *ChemBio3D User's Guide* for more information about using ChemBio3D.

 **Note:** Unlike interaction with ChemBioDraw, changes you make to a model within ChemBio3D are not transmitted back to ChemBioFinder, and thus are not saved. You can put a picture of the ChemBio3D model into a data-

base by saving the ChemBio3D model as a bitmap or metafile and inserting it into a Picture field. The picture does not retain a connection table and cannot be edited.

Styled text

Make changes to text fonts, sizes, styles, colors, and alignment with the tools on the Text menu or the Text Format toolbar.

To display the Text Format toolbar, if it is not visible:

- Go to **View>Toolbars>Text Format**. The Text Format Tool appears.

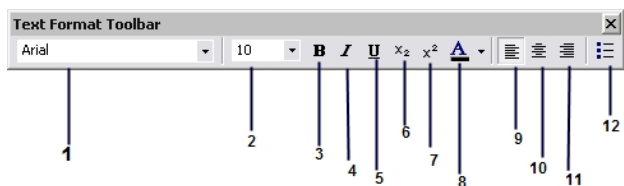


Figure 6.4: 1) font; 2) point size; 3) bold; 4) italics; 5) underline; 6) subscript; 7) superscript; 8) color; 9) left alignment; 10) center alignment; 11) right alignment; 12) bullets

Use the toolbar to format text in Plain Text or Memo/Rich Text data fields.

Undoing changes

Before committing your text changes, you can revert to the previous unmodified state.

To undo your changes, do one of the following:

- Go to **Edit >Undo**.
- Click the Undo icon on the Main Toolbar.

After you commit the changes, they cannot be undone.

Redoing changes

When you undo an action, the Redo command becomes active. You can reverse the effect of the Undo command by choosing the Redo command.

To redo the last action performed, do one of the following:

- Go to **Edit>Redo**.
- Click the Redo icon.

The last action undone is reinstated.

Deleting data

You can delete the contents of individual fields by using the Delete or Backspace keys. You cannot delete a structure, formula, molecular weight or Mol_ID.

To delete an entire record:

1. Move to the record that you want to delete using the Record commands or toolbar.
2. Do one of the following:
 - Go to **Record>Delete Record**.
 - Click the Delete Record icon.


The record is permanently removed.

 **Note:** You can delete multiple records if you use the Table view.

Sorting data

You can sort the database by most types of data in a form. Sorting on the Mol_ID field, the molecular weight field, or any other numeric field arranges the current list in increasing or decreasing numeric order according to that field. Sorting on the structure field arranges the list by increasing number of atoms contained in the structure. Implicit hydrogens are not counted. Sorting on the formula field orders the records by increasing C-H-N count (for compounds containing carbon) followed alphabetically by any other elements. Formulas for compounds without carbon are sorted in alphabetical order. Sorting by text fields such as molecular name sorts the records alphabetically.

After sorting, the database must be reset to its original state before you can update it.

 **Note:** You cannot edit records sorted by formula or molecular weight. However, if you attempt to sort on one of these fields, ChemBioFinder offers to set up a database (a portal database) which allows the operation. For details, see ["Attaching tables from other applications" on page 70](#).

To sort on a field:

- In the Form View, right-click in the field you want to sort by, and go to **Sort>Ascending** or **Sort>Descending**.
- In the Data Table View, double-click on the table header of the field you want to sort by.

The table is sorted in ascending order by the column you selected. Double-clicking again will sort in descending order.

 **Tip:** There is also a context menu option for sorting in Data Table view. Right-click in a column and choose **Sort Column**.

Sorting in languages other than English

Text that you sort must be in the same language as the default language of the system on which the database was created, or an incorrect sort order can occur.

The ChemBioFinder sample databases, created in English, can be sorted correctly using the following languages:

- English
- German
- French
- Portuguese
- Italian
- Modern Spanish

If you want to sort text in a different language than the one in which the database was created, perform the following procedure:

1. On a computer using the same language as the text you want to sort, open **MS Access**.
2. Open the MDB file you want to sort.
3. Compact the database: Go to **Tools>Database Utilities>Compact Database**.
4. After the database is compacted, close **MS Access**.

The correct sort information is written into the database. When you perform a sort in ChemBioFinder, the sort order will be correct.

Resetting the database

To reset the database to its original state:

- Go to **Search>Retrieve All**.

Changing the database scheme

You can create or remove fields or tables in an existing or new database if the database is not read-only and you have permission to modify it. You may also attach, or link, tables from external data sources.

Use the Box Properties dialog box to perform the following procedures:

Do one of the following:

- Go to **File>Database**.
- Right-click any empty space in the form window and click **Data Source**.

The Box Properties dialog box appears, with the **Database** tab displayed.

To add a field:

1. Click the **Field** tab.
2. If there is more than one table, select the table that will contain the new field.
3. Click **Create Field**. The Create Field dialog box appears.
4. Replace the default name (New_Field), select a data type, and adjust the field width (optional). Click **OK**.

To delete an existing field:

1. Select a field.
2. Click **Delete Field**. A warning box appears and asks you to confirm this deletion.
3. Click **Yes** to delete the field.

 **Caution:** All data entered in the field, will be lost as soon as you click **Yes**.

To add a table:

1. Click the **Table** tab. Do one of the following:
2. Click **Create Table**.
3. Enter a name in the Create Table dialog box. The new table is created, with a default field named ID.

To remove a table:

1. Select a table.
2. Click **Delete Table**.
3. Click **Yes** in the warning box.

 **Caution:** *All data in all fields in the table will be lost as soon as you click **Yes**.*

To attach a table from another database:

1. Click the **Table** tab.
2. Click **Attach Table**. The Attach Table dialog box appears.
3. Select the type of database – MS Access or Oracle/ODBC.
4. For MS Access, browse to the MDB file containing the table. Select a table, and click **OK**.

Importing and Exporting Data

You can move data into and out of a database if the data is in a supported file format. You can import single files, import or export databases, or add data to an existing database.

Supported file formats

ChemBioFinder lets you work with individual chemical structures and reactions in various file formats. The supported formats are:

- ChemBioDraw (cdx)
- ChemBioDraw XML (cdxml)
- Connection Table (ct)
- Delimited text (csv, txt)
- MDL Molfile and V3000 (mol)
- MDL RXNfile and Rxn V3000 (rxn)
- MDL RDFfile (rdf)
- MDL SDFfile (sdf)
- MDL Sketch (skc)
- MDL Graphic (tgf)
- SMD 4.2 (smd)

Supported formats for output files only:

- Bitmap (bmp)
- XML (cfxml)
- Encapsulated Postscript (eps)
- GIF (gif)
- Microsoft Word (doc)
- TIFF (tif)
- Windows metafile (wmf)

For information on the .cdx, .ct, and .cdxml file formats, see the *ChemBioDraw Users Guide*.

Importing data

ChemBioFinder lets you import ChemBioDraw structures, Structure Data files (SDFfiles) and Reaction Data files (RDFfiles) directly into a database. In addition, you can import text files that are comma or tab delimited.

Because ChemBioFinder uses the Access (Jet) database standard, names of fields, controls, and objects must conform to Access rules. These are:

- May be up to 64 characters long.


- May include any combination of letters, numbers, spaces, and special characters except a period (.), an exclamation point (!), an accent grave (`), or brackets ([]).
- Must not begin with leading spaces.
- Must not include control characters (ASCII values 0 through 31).
- Must not include a double quotation mark (") in table, view, or stored procedure names.


Importing structures

ChemBioFinder imports structure files directly into a database. You can choose to replace the existing records, append new records, or merge the data, eliminating duplicates.

The general procedure for importing structures is:

1. Use the **File>Import** submenu, to choose the type of structure file to import. The Open Chemical Structures dialog box appears.
2. Browse to a folder containing structures.
3. Select a file or files (using Ctrl+click or SHIFT+click) and click **Open**. The Data Import dialog box appears.
4. Click **File**, and select a database to import into.


 **Note:** You can also type in a name to create a new database file.

5. Choose the type of import: Overwrite, Append, or Merge. If you have selected Merge, click the **Merge** tab and choose the matching field, match options, and logging options.
 6. Select the fields to be imported with the check boxes.
 - Optional: Double-click the field name to change the name of the field, field type, and width in the Data Field Import dialog box.
 - Optional: Click the **Logging** tab to change details of the log file.
 - Optional: Click the **Form Style** tab to select the style of boxes that will be created for new fields.
 - Optional: Click the **Advanced** tab to set the Advanced options. Import only part of a file by setting the starting point. You may indicate a starting point by line number, byte number, or record number. Also, select the version of the ChemBioFinder database to be created. Versions back to 7.0 are supported. In addition, import structures only by selecting the check box and enter duplicate reagent values by checking the **Store redundant reagents** check box.
-  **Note:** In SD and RD files, reagents and catalysts are specified in each record. ChemBioFinder stores them in subtables. Thus, if the same reagent appears in more than one record, ChemBioFinder only needs to store it once. You can over-ride this default with the check box.
7. When you have finished selecting your options, click **Import**. The structures are imported to the specified database.

Structure and reaction

ChemBioFinder lets you import **Structure Data files (SDFiles)** and **Reaction Data files (RDFiles)** directly into a database. Since these files contain both structures and data, ChemBioFinder creates fields in the database to accommodate the incoming data.

In RDF files, incoming data may be hierarchical and complex. When loading RDF files, ChemBioFinder converts the data to a relational form, creating new tables as necessary and generating linking data.

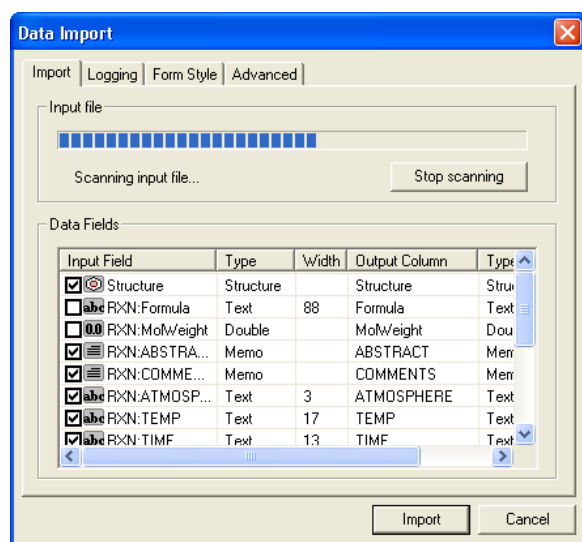
 **Note:** When importing RD files, boxes and subforms are created and positioned automatically.

If you have a blank form when you import, ChemBioFinder creates boxes on the main form. If you have a form with boxes when you import, new boxes are created on tabbed forms.

You can import with a form linked to a database, a blank form, or no form. If you import with no form, ChemBioFinder creates a new form. Selections you make on the **Form Style** tab of the Data Import dialog box determine what the form will look like.

To import structure and reaction data files:


1. Use the **File>Import** submenu, to choose the type of structure file to import. The Open dialog box appears.
2. Choose the file to import and click **Open**. The Data Import dialog box appears.



The Import dialog box scans the file to determine what data fields are present and how much space to allow for them in the database. The number of records scanned is shown in the status bar. If the input file is large, the scan may take a while.

To interrupt the scan, press the Esc key. A message box appears to confirm whether you really want to stop. The scan continues until you click **Yes**.

To abort the scan, click **Stop scanning**.

 **Note:** You should let the scan go to completion so that all needed fields are created in the database before loading.

Using log files

During the import, ChemBioFinder creates a log file of the actions taken. The log file (.log) is a text file that is created alongside the input file and overwrites any previous log file. The log records data of your choice from the following options during the import.

- Errors: Errors and warnings
- General data: Information about the input file and import process
- Records processed: Logs an entry for each record of the input file
- Structures registered: Logs an entry for each structure stored in the database
- Data registered: Shows each data item stored in the database
- Database schema: Shows input fields found and columns generated

The log file for an RDFFile import additionally presents an overview of the data table hierarchy within the RDFFile.

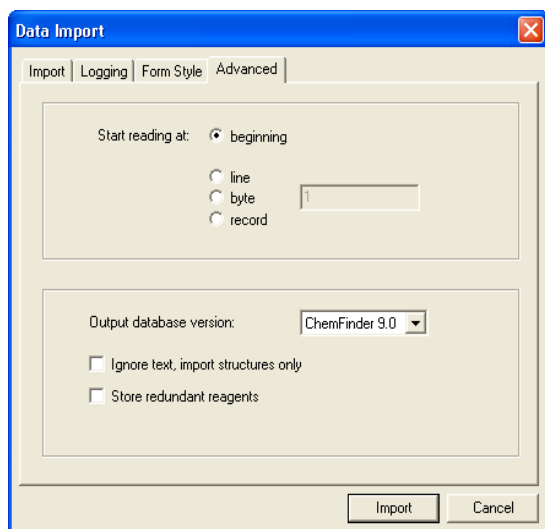
You can choose whether to create the log file or not. If you choose to create the log file, you can save it with a different name or append to an existing log file.

Importing from a specified location

You can start an import from an arbitrary location in an SDFFile or RDFFile, such as a byte or line number.

To set the location from which to import:

1. Click the **Advanced** tab of the Data Import dialog box.
2. Select the appropriate option:



To start from	Click
the beginning of the SDFFile or RDFFile	beginning.
a specified line number	line and type the line number.
a specified byte	byte and type the

To start from	Click
number	byte number.
a specified record	record and type the record number.

3. Click **Import**.

Reading a structure

In addition to importing, ChemBioFinder can read files of any of the supported structural formats. For example, if you have structures stored in ChemBioDraw format, you can open them directly from ChemBioFinder without having to import or redraw them.

To read a structure from a file into a structure box:

1. Right-click in the structure box, and click the **Read Structure** context menu command. The Open dialog box appears.
2. Choose the file to read.
3. Click **Open**. The structure is read into the structure box. The database is not affected by this operation until you choose Commit Changes or move off of the record.

Drag and drop

You can drag certain types of files from the Windows Explorer or Desktop onto ChemBioFinder. The following table summarizes the ChemBioFinder Drag and Drop options.

type of file	supported file types	comments
Structure	<ul style="list-style-type: none"> ▪ CDX ▪ SKC ▪ MOL ▪ MST 	Must be dropped into a structure box.
Graphic	<ul style="list-style-type: none"> ▪ WMF ▪ EMF ▪ BMP ▪ GIF ▪ PNG ▪ TIF 	Must be dropped into a picture box.

type of file	sup-ported file types	comments
Form	<ul style="list-style-type: none"> ▪ CFW 	Does not need to be dropped onto an existing form.

Multiple files can be dropped at once, but you cannot mix types (as defined in column 1 of the table). If several structure or graphic files are dropped at once, they are loaded into successive ChemBioFinder records. If only one is dropped, you are asked if you wish to append the structure/graphic in a new record or overwrite the existing record. If more than five structures/graphics are dropped, an alert will ask you if you wish to proceed. Dropping one or more form files opens the forms, without affecting the current form and database.

Importing text

You can import comma or tab delimited text files. The Import Delimited Text command uses the same dialog box and has most of the same features as Import SD File. Delimited text (DT) differs from SD files in that:

- DT files cannot contain structures.
- DT files are not as well defined, and require interpretation in order to determine field names and delimiter type.

To import text records into a new database:

1. Open a new, blank form.
2. Go to **File>File Import>Delimited Text**.
3. In the File Open dialog box, browse to select the input file.
 - The input file is scanned to determine the field delimiter. If the file extension is “.csv,” the delimiter is taken to be comma. If the extension is “.txt,” the first few lines of the file are examined for the presence of commas or tabs. If one or the other is found, it is taken as the delimiter. If not, an alert is presented.
 - This alert will appear in two cases:
 - The input file contains only a single column. The alert can be ignored (click **OK**).
 - The input file has some delimiter between fields other than tab or comma. In this case you must specify the delimiter using the File Export dialog box.

To specify the delimiter using the Role Export dialog box:

1. Click **Cancel** to dismiss the alert and abort the import.
2. Open a form containing data, and retrieve a small hit list.
3. Go to **File>Export>Other**. The File Export dialog box appears.
4. Click the **Text Options** tab.
5. Choose the **Other** option, and specify a delimiter.

6. Click **Export** to save the choice (and carry out an export)
7. Go back to step 1 of the import procedure to begin the import again.

 **Tip:** Once you have specified a delimiter, it will be remembered and you will not have to repeat this operation.

The input file scan also determines the names and types of the data fields in the input file. Field names must appear in the first line of the file, in the correct order, separated by the same delimiter as the data items. Field types are determined by examining all the data in each column, using the same rules as for SD import: if all items in one column are integers, then the field type is set to integer, and so on. Widths of text fields are set to accommodate the largest item in the file.

Exporting data


You can create a new SDF file, RDF file, ASCII text, or MS Word file by exporting records from an existing database. Text files can be saved in text (*.txt) format, or in the Comma Separated Value (*.csv) format readable by most spreadsheets. When you export a file, all records in the current hit list are exported. You can include or exclude fields to be exported from a checklist.

The Data Export dialog box is similar to the Data Import dialog box. It gives you options for file naming and file type selection, and lets you select which fields will be exported. If you save to an existing file, you have the option to replace or append the records.

The Text Options tab of the dialog box contains options relevant to both text file export and structured text export. It also has an option (Prefix output filenames) relevant to Oracle subform data, and an option to select only the fields on the current tab (for tabbed ChemBioFinder forms).

To export a file:

1. Use the **File>Export** submenu to select the output file type. The Data Export dialog box appears.
2. Type a name and path for the file, or click **File** and browse to an existing file.
3. Select/deselect the fields you want to export.
4. Click the **Text Options** tab and choose appropriate options.

 **Note:** Choosing Comma delimited does not automatically set the file type. Text files (*.txt) may be either tab delimited or comma delimited. If you want to export in the Comma Separated Value (*.csv) format, you must name your file accordingly—either by typing the name or selecting the file type from the Save As dialog box.

5. Click **Export**. The file is exported.

Saving structures


Saving structures is similar to exporting them, without the fuss.

To save the structure on display:

1. Right-click in the structure data box, and click the **Save Structure** context menu command. The Save As dialog box appears.

2. Choose the destination directory.
3. Type the file name and choose a file format.
4. Click **Save**.

The structure is saved to the indicated file. You can read these files with any application that supports the specified file format.


 **Note:** To save most of the file types listed above, you must have ChemDraw ActiveX Control Pro installed.

Using copy-paste with structures

You can copy structures and paste them into ChemBioDraw or ISIS/Draw. You can also use the **Copy As** command to copy the structure as a text string.

The structure **Copy As** command has three options.

- The Name command invokes the CS Struct=Name™ utility. Struct=Name generates systematic chemical names with proper CIP stereochemistry descriptors.
- The SMILES command creates a SMILES string. A SMILES string is a standard way to describe a chemical structure in a line of text. Several software packages use SMILES strings as a way to enter and store chemical structure information.
- The InChI™ command invokes the CambridgeSoft implementation of the IUPAC InChI™ algorithm. The InChI™ algorithm takes a chemical structure and converts it into an alphanumeric string. A key aspect of the InChI™ algorithm is that the generated string is independent of the way that a structure is drawn, producing a consistent alphanumeric representation of different structural representations of the same structural formula (canonicalization).

 **Note:** InChI™ is a registered trademark of the International Union of Pure and Applied Chemistry. InChI™ Material in ChemBioFinder is © IUPAC 2008.

All three commands place the data on the Clipboard, allowing it to be copied into other applications.

To copy a structure as a name or InChI™ string:

1. Click in the Structure box to select it.
2. Right-click, and click **Copy As>Name** or **Copy As>InChI** submenu command.
3. Paste the data into the other application using **CTRL+V** or the **Edit>Paste** menu command.

Exporting an ASCII file

Delimited text files may be the most universal file format used by non-chemical applications. In a delimited text file, each line represents a record, and all fields in the record are listed in order from left to right, separated by some character such as a tab or a comma. Individual records are separated by carriage returns. ChemBioFinder lets you include a header line at the start of the file that lists the names of all of the fields. It also allows you to append to, or overwrite, existing text files.

You can import a delimited text file saved by ChemBioFinder into many other applications, including spreadsheets and external database systems. Choose the .csv file format to export to spreadsheets, and the .txt format for general export.

 **Note:** *In delimited text files, structures are automatically exported as SMILES strings.*

Exporting a Word file

The export to MS Word option is similar to the other export options, however a couple of points are worth noting:


- Export to Word is much slower than to other formats. For this reason, you should probably limit the use of this option to relatively short hit lists.
- You can speed up the export by not exporting the structure field.
- Structures are exported as OLE objects, and can be edited in Word with the ChemBioDraw Control.
- If you should, by accident, begin exporting a large database, you can terminate the export by bringing ChemBioFinder to the front and pressing the <Esc> key.

Exporting to SDFFile

When you export to SDFFile, you are allowed to export subform fields. Each subform field is exported to a separate record in the SDFFile, but separate subform records are concatenated into a single multi-line text block.

To export subforms:

1. Go to **File>Export>SDFFile**. The **Data Export** dialog box appears.
2. Click the **Text Options** tab, and select the **Export subform data** check box.
3. Return to the **Export** tab. In the **Field** column of the **Data Fields** table, all subform fields now appear with checks in the check box next to the field name. Deselect any subform fields you do not wish to export (and select those main form fields you do wish to export).
4. Click **Export**.

 **Note:** *Re-importing a file of this type will not regenerate the original database! It will bring back the data, but not in the subtables and subforms you started with. To do this, export to cfxml.*


When exporting to SDFFile, you may append to, or overwrite, existing SD files. This option is only available for SD files and delimited text exports.

Printing


You can print out all or part of your database.

To print a database:

1. Before printing, decide if you want to print a table view or a form view. Set ChemBioFinder to the appropriate view.
2. Go to **File>Print Setup**.
3. In the Print Setup dialog box, set the printer, paper size (default: US Letter), and page orientation desired. Click **OK**.

 **Tip:** Settings in this dialog box affect the Page Setup, so it's best to start here.

4. Go to **File>Page Setup**.
5. In the Page Setup dialog box, set the margins and units (cm/inch), layout, and other options.
 - The Header option prints the file path at the top of the page.
 - The Footer option prints the date, time and record number at the bottom of the page.Click **OK**.
6. Go to **File>Print Preview**.
7. Go to **File>Print** . The Print dialog box appears.
8. Set the print range and number of copies. You may also select a different printer than the default.

 **Note:** ChemBioFinder defaults to printing a single page. This is indicated by the option labeled "Selection". If, for example, you are printing in Form view and you set the layout in Page Setup to three per page, you will print the record selected and the next two records. If your layout is one per page (the default), you will print only the selected record. Note also that if a subform is selected, the entire record with the subform is printed.

9. Click **OK**.

Queries

You can search the database by querying any field or combination of fields. You can specify a chemical structure and/or text as the query. When searching text or numbers, you can use wildcards or specify a numerical range. When structure searching, you can search by sub- or full structure, search for an exact match, similarity, or tautomerism, and specify how stereochemistry will be matched. You can also combine structure searching with text or numerical searching. A **Find Structure** command performs “duplicate”, “error”, “isotopic label”, and other special structure searches.

The chapter describes the different types of searches and how to set up and manage searches. It concludes with some examples of advanced techniques.

Text searches

ChemBioFinder lets you search any alphanumeric field for strings of text. Use text searching, for example, to find chemical names or comments in a reference field.

For text searches, ChemBioFinder interprets the query, then passes it as Structured Query Language (SQL) to the relational database.

These rules apply to text searches:

- When searching a normal text or memo field, a plain text string is taken as an “unanchored substring”, if you are not performing a Full Word search. An asterisk (*) is added automatically at each end of the string, and the hits include any string containing the search string. Thus, if you search for ‘benz’, it becomes ‘*benz*’ and returns benzene and fluorene.
- When searching a normal text or memo field with wildcard, the string is taken as an “anchored substring”. Thus, if you search for ‘benz*’ it gets anchored to the left and returns words beginning with ‘benz’.
- Text strings may contain wild cards or Boolean operators. Standard wild cards are % and *. These characters are equivalent and hit any string. Other wild cards are also possible. If wildcards are present, they override the above defaults.
 - 📌 **Note:** Using wildcards or a “=” in memo fields may lead to unexpected results. This is because a memo field may contain formatting information surrounding the text you see, so, for example, a search for benz* may yield no results. Using *benz*, however, will yield the same results in a memo field as in a plain text field.
- A check box in the Search Preferences dialog box lets you request a full-word match. When this box is checked, a hit must contain the query as a complete word, not embedded within a larger string.
- Boolean operators are NOT, OR, and AND. They may be used to combine search terms within one field.

✔ **Tip:** Use NOT* to search for empty fields.

Examples of queries:

Entry	Possible Hits	Will Not Hit
benz*	Benzene, Benzoic acid	Bromobenzene
benz* OR *bromo*	Benzene, Benzoic acid, Bromobenzene, Dibromobenzene	Azobenzene
=benzene	Benzene	Benzoic acid, Benzene-d ₆ , Bromobenzene

Numeric searches

Use numeric searches to find information such as boiling points, molecular weights, and other properties. A numeric query can be either a decimal value or range. When you enter a single value, the number of significant digits determines the precision of the search. A hit is any value that rounds off to the query. When entering a range, enter two values separated by a hyphen.

Examples of numeric queries:

Entry	Possible hits
90	values from 89.5 to 90.5
90.1	values from 90.05 to 90.15
90–100	values from 90 through 100, inclusive
>90 and <100	values from 90 through 100, exclusive

Molecular formula searches

Formula searching lets you search molecular compositions. Searches can be inclusive, exact (designated with an equals sign) or with an element range you specify.

ChemBioFinder allows formula searching based on special atoms, such as A, Q, X, and M. The definitions for these atom types can overlap. For example, the structure, BiCl₃ will be hit by either the query AX₃ or the query A₄.

However, it will not be hit by the query A4X3 even though it has four atoms that match A and three atoms that match X.

Formula queries consist of element symbols and element counts or ranges.

The following rules apply:

- Symbols may be one or two letters. Symbols may be in upper or lower case; if there are ambiguities, the program resolves them according to rules described in ["Formula Input Rules" on page 238](#).
- Capitalize the symbols properly and insert spaces between elements.
- Use the Periodic Table to enter formulas. For more information, see ["Periodic table" on page 208](#) and ["Formula Input Rules" on page 238](#)
- Element counts are single integers or ranges (two integers separated by a hyphen). If a count is omitted, it is assumed to be 1.
- If the formula query is preceded by = (for example =C6H6) then the search requires an exact formula match, containing no elements other than those indicated. If there is no =, then the search is a partial match: other elements may be present in the hits.
- Symbols may be repeated. For example, CH3CH3 is interpreted as C2H6.
 - 📌 **Note:** *Formula searches are non-structural: CH3CH2OH matches both dimethyl ether and ethyl alcohol because both compounds have the same condensed formula: C2H6O.*
- Parentheses may be used to group elements and apply a count to the entire group. For example, (CH2)3 is interpreted as C3H6.
- Spaces or non-alphanumeric characters other than parentheses are ignored.

This table shows query examples.

Entry	Possible hits
C6H6	compounds with 6 carbons and 6 hydrogen atoms, plus any number of other elements.
=C6H6	compounds with 6 carbons and 6 hydrogen atoms and no other elements.
C6N0	compounds with 6 carbons, no nitrogens, plus any number of other elements.
C6 N1-3	compounds with 6 carbons and one to three nitrogen

Entry	Possible hits
	atoms, plus any number of other elements.

Date searches

You can use date searches to track individual reaction runs, purchasing histories, and so on.

You specify date ranges using a hyphen between the values at either end of the range. Ranges may also be indicated using inequality operators (<, >) with the AND operator.

Dates are always displayed according to the preferences set in the operating system's International control panel, but need not be input in that format.

This table shows query examples.


Entry	Possible hits
Apr 11, 1971	the exact date April 11, 1971
March 31, 1971– July 1, 1971	any date in the second quarter of the year 1971
>=4/11/71	any date after April 11, 1971 or on April 11, 1971

Find list

Use the **Find List** command to retrieve records having specific values in any non-structure field.

Type parameters into the Find List box as follows:

- If the column you search is an integer, use hyphens to indicate ranges.
- If the column you search is a number or text type, hyphens are not allowed; you must search for an exact match.

 **Note:** To search for a list of records using partial text (a “wildcard” search) or range of numbers, perform a search on the relevant field.

- If the column you search is text, use quotes around items that contains commas. For example, type 1,2-Pyran as “1,2-Pyran”.
- You can save a list as a text file and retrieve it later using **File**.

To find a list:

1. Go to **Search>Find List**. The Find List dialog box appears.
2. Choose a field from the **Column** menu.

3. Paste or type a list of field parameters into the **Find List** text box.
4. Click **OK**. The list of matching records is retrieved.

Structure searches

You can search a ChemBioFinder database using sub- or full structure queries. To refine your search, use the query functions in ChemBioDraw. These are described in the *ChemBioDraw User Guide*, and in "Structural Query Features" on page 222 in this guide.

ChemBioFinder matches structures in three ways:

- Search type (Normal/Exact/Similarity)
- Structure mode (sub- or full structure)
- Tautomerism

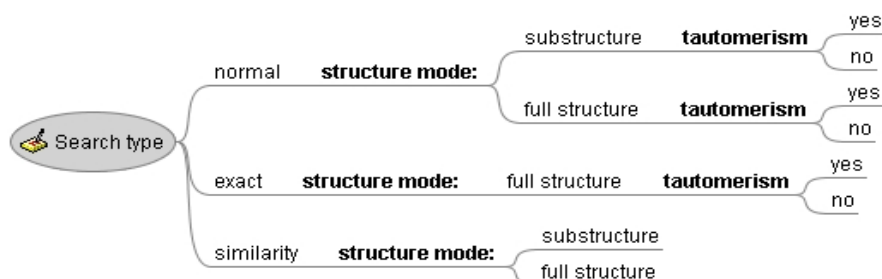


Figure 8.1: Structure search tree

Search types

ChemBioFinder supports three structure search types:

- Normal
- Exact
- Similarity

Normal and similarity searches may be in either of two modes: full structure or substructure. Normal and exact searches may also search for tautomerism.

Normal searches

A Normal search finds structures that either contain (Substructure) or match (Full Structure) the query. When drawing a structure query, you can attach different features to a query, such as atom lists and variable bond types, to perform a narrower or broader search. You can also modify the results of a Normal search by selecting options in the **Details** tab of the **Preferences** dialog box.

The query structure is highlighted in red in the hit list structures to visualize the match.

Exact searches

The Exact search type, also known as an Identity search, is intended for use in compound registration, when you must know if an identical copy of your query compound is already present in the database. The target must be chemically identical to the query, including stereochemistry, charges, and isotopy. It is a convenient shorthand for full structure, same-stereochemistry, and appropriate settings of all the other options. Thus, generic atom and bond types such as R, A, and Double-or-Aromatic in the query will match corresponding atom and bond types in the target only if they are also of the same generic type. In other words, an atom labeled “R” in the query must find a matching atom labeled “R” in the target to produce a hit. If the stereochemistry of the query is unspecified, it will only hit targets that do not specify stereochemistry. The only variable in an exact match search is whether or not it is tautomericly flexible. The **Details** tab in the **Preferences** dialog box is unavailable for Exact searches.

Similarity searches

A Similarity search finds targets that “look like” the query. Similarity searches are by their nature “fuzzy”. What “looks like” means is obviously subject to interpretation and depends on the application. In medicinal applications the drug absorption properties are relevant. In a toxicological context the metabolism is of interest. For database queries, the similarity in functional groups is what is measured. A quantitative measure of similarity, called the Tanimoto algorithm, is calculated. For more information, see ["Similarity Rules" on page 240](#).

Similarity searches may be either full- or substructure. In a substructure similarity search there is no penalty if the target contains extra, non-similar aspects. In a Full Structure similarity search, the results are guaranteed to include all hits you would obtain from a full-structure search with the same query. Usually, they include additional hits. For this reason, similarity searches are useful if you have a general idea of the types of compounds you are looking for, but don't have an accurate description of the target compound.

Unlike exact searches, similarity searches do not highlight matched portions of the target compounds. Similarity searching matches general structural features and not specific atoms and bonds, so highlighting specific areas would not be appropriate.

You can adjust the degree of similarity (from 0 to 100%) necessary to produce a hit by using the slider on the **Search Type** tab of the Preferences dialog box.


Structure modes

The structure mode determines whether the query defines a substructure of the compounds in the database being searched, or the full structure. In this context, it is important to note that “full structure” is not “exact”. When you specify the “Exact” search type, the structure mode is always full structure, but you may choose the full structure mode with any of the three search types. See ["Full structure searches" on page 102](#) and ["Exact searches" on page 101](#) for more details.

Substructure searches

A substructure search finds structures that contain the query, plus any additional attachments at the open positions. This is the default search mode. The substructure is highlighted in red in the hit-list structures. Using the ChemDraw ActiveX toolbar, you can attach different features to a query, such as atom lists and variable bond types, to perform a

narrower or broader search. For more information about what query features you may use and how they affect a search, see ["Changing the scope of a search" on page 230](#).

 **Note:** In searching substructures, ChemBioFinder finds the substructure query regardless of its orientation or drawing presentation in the targeted molecules. Bonds shown in bold above are highlighted in red in ChemBioFinder.

Full structure searches

A full structure search finds structures that completely match the query. If you do not specify stereochemistry or use generic atom or bond labels, you may get more than one hit. You may also get more than one hit if there are duplicates in the database. For more information, see ["Stereochemistry" on page 103](#) and ["Exact searches" on page 101](#).

Fragment searches

You can draw more than one structure or structure fragment in your substructure search query. In the **Details** tab of the Preferences dialog box, you may allow fragments to overlap in the target structure.

For example, if you perform the following substructure query with **Query fragments may overlap in target** selected:

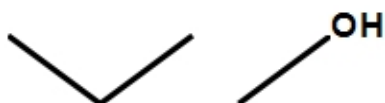


Figure 8.2: Fragment query

The hit list will include this structure:

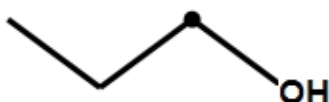


Figure 8.3: Example of substance hit in a fragment query

The dot indicates an atom shared between the two fragments. The hit list will not include this molecule if the overlap option is deselected.

 **Note:** The option, **Query fragments may overlap in target** is useful only in substructure search mode.

Tautomeric searches

In a tautomeric search, the location of single and double bonds may vary from query to target. A hit occurs if, in addition to the atom types matching, the target possesses a tautomeric form whose bond orders match the query's. Which tautomers are recognized is under some degree of control. Tautomers of the form:



Figure 8.4: Tautomeric structure

are always recognized when at least one of X and Z is not carbon. The lone pair may include a migrating proton. It is also possible to include tautomers in which X and Z may both be carbon. The choice must be made at the time the database is created or upgraded, because it affects how structures are stored.

Mesomers (resonance structures) are also picked up in tautomeric searches. Tautomerism involving formation or cleavage of sigma bonds is not currently supported. The principal categories are ring-chain tautomerism, as in the open vs. cyclic forms of sugars, and valence tautomerism.

Tautomeric searching is slower than non-tautomeric, but is the mode of choice when you think the target might have a slightly different layout of double bonds.

 **Note:** *If the database has not been indexed, the tautomer searches may be slow.*

Stereochemistry

You can specify whether or not you want a structure search to consider stereochemistry. Given a structure with one or more stereocenters, you can store four different possibilities:

- A given absolute configuration. To specify this, centers must be drawn with stereo bonds, and the entire structure marked Abs (CHIRAL).
- A racemic mixture of the drawn configuration and its mirror image. This is drawn as above, but without the Abs (CHIRAL) mark.
- An unmarked structure, representing unknown stereochemistry, or a mixture of all possible stereo-isomers.
- A given relationship between centers. That is, a known orientation of the substituents with respect to each other, rather than a known absolute configuration. To specify this, centers are drawn, not with the standard hashed and wedged bonds, but with thick stereo bonds.

If these options are checked in the Search tab of the Preferences dialog box, then any stereochemistry indicated on the query must be matched by the target. For more information on changing searching preferences, see ["Setting search details preferences" on page 112](#).

3D Properties

3D queries are useful in pharmacophore searching where you are looking for a particular 3D relationship among atoms and bonds, for example in a series of potential receptor ligands. You can create a 3D query in ChemBioDraw Pro by adding geometries (lines, planes, etc.) and constraints (specified as ranges) to a query structure. For example, you might specify that two atoms must be between 4Å and 5Å apart, or that two planes must be separated by 80-100°. ChemBioFinder can use these properties to refine a search. See the *ChemBioDrawUser's Guide* for information on how to add 3D properties to structures.

Searching with R-groups

ChemBioFinder supports queries using R-tables, including multiple R-tables. Queries containing large numbers of R-groups or large numbers of R-values are not advisable however, because ChemBioFinder pre-processing expands the query, substituting all combinations from the R-group table(s).

Alternative groups

Instead of submitting multiple queries on structures that share a common substructure, you can submit a single query with the parent structure and variable functional groups or substructures. The parent structure is drawn with attachment point(s) that refer to a list of alternative groups that you define.

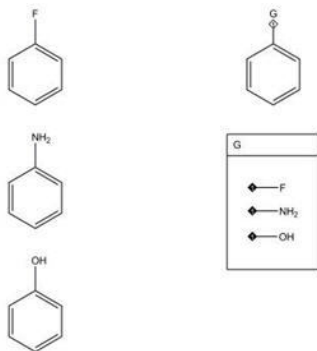



Figure 8.5: alternative groups

Defining an alternative group

To define an alternative group:

1. Click the query icon to begin the search.
2. Right-click in the structure window and click the **Edit in ChemBioDraw**. ChemBioDraw opens to a new page.

 **Tip:** You want to open ChemBioDraw because you will need more space than the structure window provides.

3. Draw the parent compound.
4. Label the atom where the alternate groups will attach with a generic label such as R.
5. Click the alternative Group query tool.
6. Drag with the tool in an open part of the page to create an area large enough to draw the alternative groups.
7. Type a title in the alternative Group Title box. The title must match the atom label in the parent structure, but we'll use R in this example).

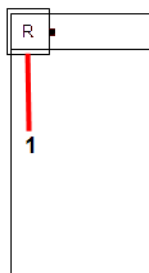


Figure 8.6: Adding a title to an alternative group: 1) Title box.

8. Draw the substructure fragments in the alternative Group box. Use the single bond tool to define a fragment, and the text tool, atom HotKeys, or Nicknames to create atom labels for each variable.

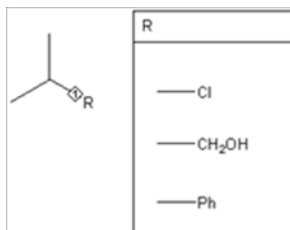


Figure 8.7: Adding the alternative groups

9. Add the Attachment Points:

- Click the diamond shaped Attachment Point tool on the Chemical Symbols palette.
- Click the open atom position on each substructure fragment.

 **Tip:** alternatively, the HotKey “.”(period) can be used to add attachment points.

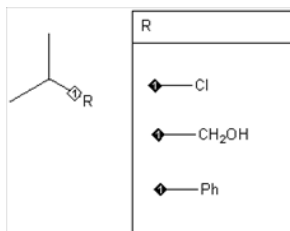



Figure 8.8: Completed alternative group query

 **Note:** In the above procedure, the parent structure was drawn, then the R-Group table was added. In fact, the order doesn't matter. You may create the R-Group table first, then draw the parent structure. In either case, as soon as a generic atom label in the structure matches the group title in the R-Group table, a hollow attachment point symbol appears next to the label in the parent structure.

- Type **CTRL+W** to return to ChemBioFinder. The structure and the R-Group table appear in the structure window.
- Continue the query as usual.

Reaction searches

You can search and store reactions. In a reaction, one or several compounds (reactants) are transformed into other compounds (products). Individual reactants (or products) are separated from each other with plus signs. The reactants are separated from the products with an arrow.

Reactions may have multiple steps, for example

(A)→(B)→(C)→(D). Here, (A) is the reactant and (D) is the product. (B) and (C) are intermediates for the complete reaction.

A multi-step reaction is actually a shorthand notation for many related reactions. In the example above, (B) is an intermediate for the complete reaction, but it is also a reactant relative to (C) or (D). It is also a product relative to (A). The complete reaction implies many subreactions, such as:

(B)→(D) and (A)→(B).

Reaction centers

The most important part of a reaction is the part that changes from the reactants to the products. This part, which probably includes a number of atoms and bonds, is called the reaction center. For example, only the bold bond in the figure below (and the two atoms on either side) is part of the reaction center. The rest of the structure is unchanged from the reactant to the product:

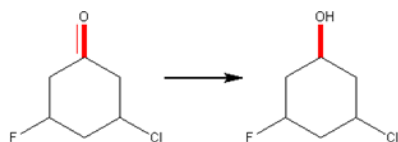
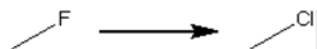


Figure 8.9: Reaction center example

By default, ChemBioFinder considers reaction centers whenever you search for reactions. ChemBioFinder assumes that any atoms and bonds that change in the query must be part of the reaction center of the target.

For example:



does not hit

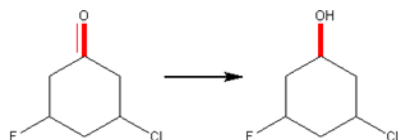


Figure 8.10: Query does not hit because it does not contain the reaction center.

when the **Reaction query must hit reaction center** preference is selected. Even though there is a C-F bond in the target reactant and a C-Cl bond in the target product, these bonds do not participate in the reaction, which really affects another part of the compound. If you deselect the **Reaction query must hit reaction center** preference, this query hits the target above.

When creating reaction queries, it is important to consider what sort of information you are really looking for. Suppose you want to convert *n*-decanal to *n*-decanol:

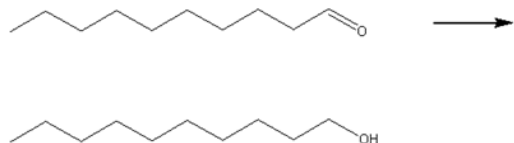


Figure 8.11: Reduction of *n*-decanal

Are you really interested only in these two compounds? You might be interested in any reaction that converts a straight-chain aldehyde to the alcohol:

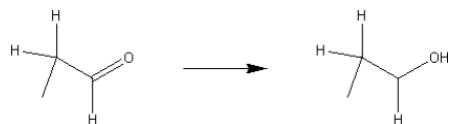



Figure 8.12: Generic aldehyde reduction query

Since the corresponding *n*-octanal→*n*-octanol reaction would probably occur under very similar conditions, it is a reasonable thing to look at. Generally, you want to use substructure queries that include little beyond the reaction center in question when you are searching for reactions.

 **Note:** ChemBioFinder supports several query properties to let you specify exactly how a bond participates in a reaction center. For more information about these properties, see ["Setting search details preferences" on page 112](#).

Atom-to-atom mapping

The second most important part of reaction searching is the atom-to-atom map. You can specify maps in ChemBioDraw, where they are stored as part of the data about a reaction. Maps are used during searching to resolve certain types of structure search hits. Atom-to-Atom mapping does not affect the search result when a query reaction is done with or without reaction map.

Consider a simple esterification reaction:



Figure 8.13: Esterification example

Does the ester oxygen come from the acid or the alcohol? You specify the fate of individual atoms through an atom-to-atom map. In reality, the ester oxygen in this reaction originates in the alcohol, so the atom-to-atom map looks like this:



Figure 8.14: Atom map of esterification


By matching numbers across the arrow, you can see where atoms move during the course of the reaction. The other reaction (not observed experimentally), where the ester oxygen comes from the acid, would be mapped like this:



Figure 8.15: alternate possible atom map for esterification

ChemBioFinder uses atom-to-atom map information to determine reacting centers for reactions. If only some atoms are mapped, ChemBioFinder uses that information and does not worry about the specific fates of the other atoms. For example, if you don't know (or don't care) about the mapping of some atoms, you can leave them unspecified in the atom-to-atom map.

By default, atom-to-atom mapping is not displayed within ChemBioFinder. To turn on this display, select **Atom-to-Atom Map** on the Structure sub-menu of the View menu. For more information, see ["Changing options with the view menu" on page 203](#).

 **Note:** For information about specifying atom-to-atom maps with ChemBioDraw, see the *ChemBioDraw User's Guide*.

Searching for reactants

If you know what starting materials you are interested in but don't know their products, you might perform a reactants query. A reactants query is very similar to a reaction search, except that there is nothing to the right of the arrow. For example, consider the query:

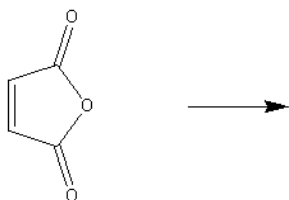


Figure 8.16: Sample reactants query

If you are doing a substructure search, this finds any reactions in which maleic anhydride or a compound containing a maleic anhydride substructure is consumed or transformed.

Searching for products

If you know the desired end product but not how to get there, you can do a products query. A products query is similar to a reaction search, except that there is nothing to the left of the arrow. For example, consider the query:



Figure 8.17: Sample products query

If you are doing a substructure search, this finds any reactions in which bicyclo[2.2.1]heptane or a compound containing this substructure is produced.

Searching for intermediates

Rarely, you may be looking for reactions for which you only know something about an intermediate. An Intermediate Search is very similar to a normal reaction search, except that there are arrows on both sides of the target structure. For example, consider the query:

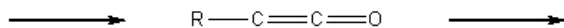



Figure 8.18: Sample intermediates query

This finds any reactions containing the Ketene structure shown as an intermediate.

 **Note:** ChemBioFinder cannot predict products or intermediates of reactions. It finds this information only if it is already present in the database.

Combined searches

You can combine structure searching with text searching to find a specific class of compounds. For example, you may want to find all compounds in the database whose names end in mycin and whose structures contain a phenyl ring. Because you are entering multiple queries in different data boxes, there is an implicit AND condition between data items in different fields.

To perform a combined search:

1. Do one of the following to clear the form:
 - Go to **Search>Enter Query**.
 - Click the Enter Query icon on the Search toolbar.
2. Enter the structural query, if any.
3. Enter the text and/or numeric queries in the appropriate boxes.


4. Do one of the following:
 - Go to **Search>Find**.
 - Click the Find icon on the Search Toolbar.

Combined searching example:

Suppose you want to find all molecules in the CS_Demo database that contain a benzene ring and which have names related to penicillin.

1. Go to **Search>Enter Query**.
2. Draw a benzene ring.
3. Enter *penicillin* in the Molname field.
4. Check that the **Search>Substructure** option is selected.
5. Go to **Search>Find**.

You get 2 hits of molecules whose names contain penicillin and whose structures have an aromatic ring of six carbon atoms.

 **Note:** In a combined search, progress reports are given only during the structure search part. When the counters at the bottom of the screen are advancing, structures are being searched. You can press Esc to end a search during the structure searching. You cannot end a search during the SQL searching by pressing Esc.

SQL searches

SQL (Structured Query Language) can be used to create powerful searches. The details of the language are not discussed here.

Queries beginning with a backslash (\) are taken as straight SQL, and passed directly to the database as the WHERE clause of the SQL query. These must contain column names and punctuation as dictated by SQL. For example, a valid query might be

```
\[molname] like 'benz*' and [bpoint] > 200.
```

A straight SQL query may be entered in any box of the form which contains non-structural data. The SQL query is not associated with the box in which it is entered.

Query procedures

Searching includes the following steps:

- Setting the search type preferences (Structure searches only).
- Entering the query in the form.
- Submitting the query to the search engine.

To perform a search, use the **Search** menu or the corresponding tools in the Search toolbar.

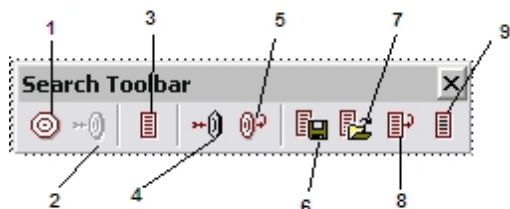


Figure 8.19: The Search toolbar: 1) Enter Query; 2) Find; 3) Retrieve All; 4) Find Current Molecule; 5) Previous Query; 6) Save List; 7) Restore List; 8) Restore Previous List; 9) Find List.

The following procedures describe searching with the **Search** menu. You may use the Search toolbar instead for most procedures.

Setting search type preferences

You can set preferences for the type of structure for which you want to search.

To set Search Type preferences:

1. Select Preferences from the Search menu. The **Preferences** dialog box appears, with the **Search Type** tab displayed.

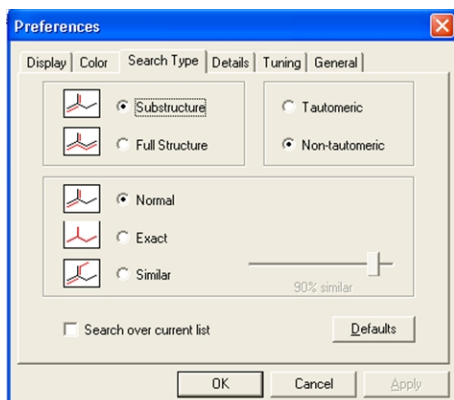




Figure 8.20: Search type preferences

2. Take the appropriate action:

If you want to	Then click
find structures containing a common structural	Substructure.

If you want to	Then click
component	
find a particular structure	Full Structure.
search with variability in the query	Normal
search for a precise match	Exact
find structures having features similar to the query.	Similar and set the degree of similarity.  Note: <i>The higher the value, the fewer hits found.</i>
find structures that are tautomeric forms of the query	Tautomeric.
match bonds as drawn, without tautomeric forms	Non-tautomeric

3. Click **OK**.



 **Tip:** You don't always have to access the Preferences dialog box to change search preferences. Switches on the Search menu let you toggle Substructure/Full Structure and Normal/Similarity.

Setting search details preferences

You can set preferences for the details of each search.

To set search details preferences:

1. Go to **File>Preferences**. The Preferences dialog box appears.
2. Click the **Details** tab. The Details tab appears.
3. Take the appropriate action:

If you want to	Then click
<p>Allow uncharged non-carbon atoms in the query to match charged atoms in the target.</p> <p> Note: <i>Charged atoms in the query must always match charged atoms in the target, regardless of this setting.</i></p>	<p>Hit any charge on heteroatom.</p>
<p>Allow uncharged carbon atoms in the query to match charged carbon atoms in the target.</p> <p> Note: <i>Charged atoms in the query must always match charged atoms in the target, regardless of this setting.</i></p>	<p>On carbon.</p>
<p>Allow hits to contain molecular fragments in addition to the structure hit by the query.</p>	<p>Permit extraneous fragments in full structure searches.</p>
<p>Run a search that ignores differences merely due to salt state and presence of solvents.</p>	<ul style="list-style-type: none"> ▪ Ignore salts and solvents. ▪ Hit any charge on heteroatom. ▪ On carbon. ▪ Permit extraneous fragments in full structure searches.

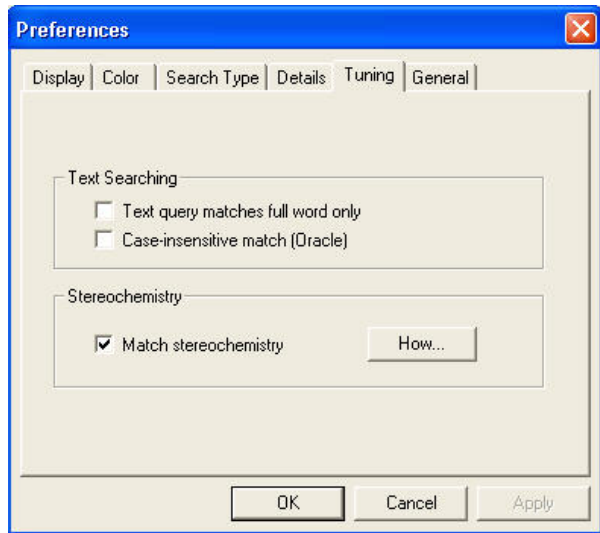
If you want to	Then click
Allow fragments in the query to overlap (share one or more atoms) in the target.	Query fragments can overlap in target.
Require that any reaction center present in the query overlap with reaction centers in the target. This preference applies only to reaction searching.	Reaction query must hit reaction center.
Prohibit generic structures from hitting any other structures in a query.	Generics hit only generics.
Require full-word text matching. If you do not check this box, the query will hit any matching text fragment.	Text: Match full word only
Require that the stereochemistry of the target structure match that of the query structure.	Match stereochemistry and click How to set how the stereochemistry is matched.

Tuning

The Tuning tab of the Preferences dialog box has settings for special query conditions. On this tab, there are check boxes for:

- full word text matching
- case-insensitive matching (for Oracle)
- matching stereochemistry

The only box that is checked by default is the **Match Stereochemistry** box. See the figure below:



Setting stereochemical search preferences

The Stereochemical Search Preferences dialog box is activated from the **Tuning** tab of the **Preferences** dialog box. To set Stereochemical preferences:

1. Go to **File>Preferences**. The Preferences dialog box appears.
2. Click **How** on the **Tuning** tab of the **Preferences** dialog box. The Stereochemical Search Preferences dialog box appears.

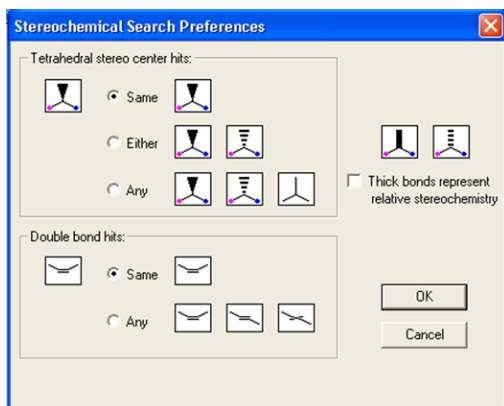



Figure 8.21: Stereochemistry search preferences

3. Set the Tetrahedral stereo center hits:

If you want a tetrahedral stereocenter in the query to...	Then...
match the target exactly.	in Tetrahedral stereo center hits , click Same .
match same or opposite configuration at the center of the target.	in Tetrahedral stereo center hits , click Either .
match any target.	in Tetrahedral stereo center hits , click Any .
match a relative relationship between centers	click Thick bonds represent relative stereochemistry .

 **Note:** When **Thick bonds represent relative stereochemistry** box is cleared (default), thick bonds are interchangeable with hash/wedge bonds. When the box is checked, and **Tetrahedral stereo center hits** is set to **Same**, a query marked with thick bonds will only hit a target that has the same relative relationship between centers.

4. Set the Double bond hits:

If you want the configuration of a double bond in the query to	Then in
match the target structure exactly.	Double bond hits , click Same .
match any configuration in the target.	Double bond hits , click Any .

5. Click **OK**.

Entering query mode

To clear the form and enter Query Mode:

- Go to **Search>Enter Query**. You are in Query mode.


In Query mode:

- The form background color changes.
- QRY appears in the status bar.
- Form boxes which do not permit data entry become editable for entry of queries.
- Form boxes are displayed according to their visibility properties. For more information, see ["Hiding data boxes " on page 43.](#)

Entering and submitting a query

To enter the query:

1. Position the cursor over the field you want to search and click to select it.
2. Enter the query in the data boxes.
3. Click **Search** on the Query Mode box. The number of hits is shown in the Status Bar at the lower right corner of the window.

 **Note:** *If a search gets no hits or an error occurs, an alert appears and you are returned to Query mode to enter a different query or modify the current one.*


The hit list is a subset of the complete database. You can browse it as you would any database using the Record commands in the Record menu or toolbar.

For example, to view the hits in tabular form:

- Go to **View>Data Table**.

Stopping a search

To stop a structure search in progress, press the Esc key. The query stops and you return to browse mode.

 **Note:** *Only queries that involve structural data (structure, molecular formula, and molecular weight searches) can be stopped in this manner. SQL searches and searches that involved non-structural data cannot be aborted.*

Refining a search

You can refine a hit list of one or more records by entering a query that searches only the hit list, not the entire database.

To refine a search:

1. Verify that the **Search>Over Current List** menu option is selected.
2. Do one of the following:
 - Go to **Search>Enter Query**.
 - Click the Enter Query icon on the Search toolbar.

ChemBioFinder searches over the previously retrieved hit list. If no hits are found, a message appears asking whether to search over the entire database.

As long as the **Over Current List** switch is on, each search further refines the current list. Use the **Retrieve All** command to reset. You may also use the **Restore Previous List** command, which acts like a “Back” button. **Restore Previous List** goes back one in the history, including any “**Retrieve All**”s you might have done.

Entering a structural query

To begin a structural query:

1. Do one of the following to clear the form:
 - Go to **Search>Enter Query**.
 - Click the Enter Query icon on the Search toolbar.
2. Do one of the following to place a structure into a data box:
 - Double-click in the structure box and edit with the ChemBioDraw Control.
 - Right-click and click **Edit in ChemBioDraw**.
 - Right-click and click **Read Structure** to open an existing molecule file.
 - Use **Edit>Paste** to insert a structure from the clipboard.
3. Go to **Search>Preferences**.
4. Click the **Search Type** tab.
5. Select the appropriate options. Optionally, select options on the **Details** tab as well.
6. Do one of the following:
 - Go to **Search>Find**.
 - Click the Find icon on the Search Toolbar.

The status bar counters indicate search progress. When the search is complete, the form displays the first hit. The list you can browse is limited to the hits. For Normal or Exact searches, the hit portion of each molecule is highlighted in red. If a search doesn't find any hits, you are returned to query mode.

 **Tip:** You can change the highlight color in the preferences dialog box. For example, to show no highlighting, choose black as the highlight color. See ["Color preferences" on page 201](#) for more information.

Using the current molecule as a query

As you browse through a database, you may submit any structure on the screen as the structural query. Often, you use similarity or substructure searches (see above) with this type of query to find related compounds.

To use the current molecule for a query:

1. Using the **Record** commands, go to the record containing the structure that you want to use as a query.
2. Go to **Search>Current Mol As Query**.

All boxes of the form are cleared except the structure so that the molecule on display can be used as part of the query. You can then continue with the query as if you had drawn the structure from scratch.

Finding the current molecule

The Find Current Molecule command, located in the Search menu and on the Search toolbar, is similar to the Current Mol as Query command discussed above. As you browse through the database you may find a molecule that interests you, and want to find other related records.

The Find Current Molecule feature lets you perform a quick structural search of the structure currently being displayed on the form. However, this feature does not let you enter other search terms. The type of structure search (complete structure, substructure, and/or similarity) is determined by what you have selected in the Search menu.

To find the current molecule:

1. Browse to the record containing the structure of interest.
2. From the **Search** menu, choose the type of structure search you want to perform.
3. Do one of the following:
 - Go to **Search>Find Current Mol**.
 - Click the Find Current Molecule icon on the Search Toolbar.

ChemBioFinder begins the search and displays the search status on the lower right corner of the status bar. When the search is complete, the form displays the first hit, and the list you can browse is restricted to the records hit by the query.


Entering a reaction query

The general procedure for creating a reaction query is very similar to creating a structural query.

1. Do one of the following to clear the form:
 - Go to **Search>Enter Query**.
 - Click the Enter Query icon on the Search toolbar.

The form clears.

2. Enter the query:
 - Double-click in the Structure box. The ChemBioDraw Control appears.
 - Draw the structure or substructure or reaction in the From ChemBioFinder window.
 - Click somewhere in the form outside the Structure box.
3. Optional: enter more query terms in the other data boxes for a combined search.
4. Select or deselect the **Search>Substructure** menu option, as appropriate for your query.

 **Note:** The Similarity option is not available for reaction queries.

5. Do one of the following to search:
 - Go to **Search>Find**
 - Click the Find icon on the Search Toolbar.

The search proceeds as with simple structure searching.

Special structure searches

The **Find Structures** submenu on the **Search** menu provides for several common types of structure searches can not be efficiently conducted using regular means.

As with other searches, all of these searches operate within the current record set if the **Over Current List** option is in effect. Otherwise, the whole list is searched, regardless of how many records are currently displayed. Special structure searches are intended to apply to all data sources, including Oracle, but presently only work for MST sources.

The following figure shows the menu options under Find Structures submenu:

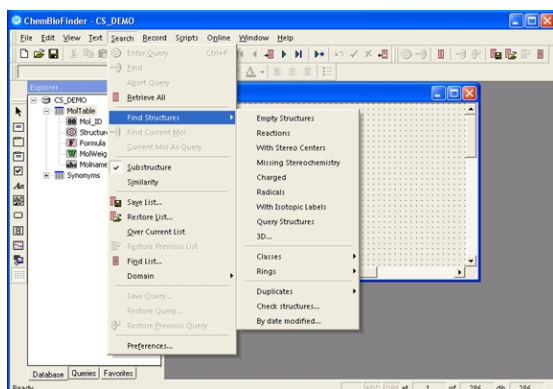


Figure 8.22: Find Structure menu


These are generic searches, made without entering any query information. The options are:

- **Empty Structures:** Locates records containing no structure. It is equivalent to specifying a molecular weight of zero in a regular search. It is provided because the latter method is a bit obscure.
- **Reactions:** Locates structures that are reactions. The same effect can be achieved by entering a bare arrow as a structure query.
- **With Stereo Centers:** Finds structures containing tetrahedral stereochemistry.
- **Missing Stereochemistry:** Finds structures with atoms capable of tetrahedral or double bond stereochemistry, but drawn as “unspecified”.
- **Charged:** Finds one or more atoms with a charge.
- **Radical:** Finds one or more atoms with a radical.
- **With Isotopic Labels:** Retrieves records containing a structure with an atom that either has a mass number, or has an **Isotopic Abundance** attribute.
- **Query Structures:** Locates structures with any generic variability such as atom types (R, A, M, Q, X, Alkyl, Aliphatic, EDG, EWG), bond types (single-or-double, etc.), variable charge, element lists ([C,N,O]), ring bond count, etc.
- **3D:** Locates structures that could be targets in a 3D search. These are structures scaled in Angstroms and having Z coordinates. Both conditions must be met. Structures that meet only one of these conditions can be found by using the **Check structures** search.

 **Note:** The 3D feature is available in the Ultra version, but not in the Standard version.

- **Classes:** Provides these search options (selectable from a submenu):
 - **Hydrocarbons:** Finds structures with Carbon and Hydrogen only. If there are more than one fragments, they must all be hydrocarbons.
 - **Aliphatics:** Finds structures with a carbon-carbon bond (or methane) and non-aromatic rings.
 - **Alicyclics:** Finds structures having a carbocyclic ring that is not aromatic or antiaromatic.
 - **Aromatics (4n+2):** Finds structures containing one or rings with a 4n+2 electron system. Includes pyrrole and charged rings.
 - **Antiaromatics (4n):** Finds structures containing one or more rings with a 4n electron system.
 - **Aromatics (Cyclically alternating):** Finds structures containing one or more rings that have 4n+2 bonds, distributed as alternating single and double bonds. This is the criterion for "delocalized" used since ChemFinder 10.
 - **Natural products:** Finds structures containing a carbon and at least nitrogen or oxygen.
 - **Alkaloids:** Finds structures with MW>200, and containing carbon and basic nitrogen.
 - **Heterocycles:** Finds structures having a ring containing a heteroatom.
 - **Metals:** Finds structures containing a metal bond.
 - **Organometallics:** Finds structures containing a metal-carbon bond.
 - **Organics:** Finds structures having a carbon atom.
 - **Inorganics:** Finds structures lacking carbon.
 - **Salts:** Finds structures that can be considered as salts.
 - **Rings:** Finds any of the following types of ring structures (selectable from a submenu):
 - **Bridged:** Systems. For example, two rings share more than one bond.
 - **Spiro:** Systems. For example, two rings sharing just one atom.
 - **Fusion:** Systems. For example, structures with a ring fusion (two rings sharing one bond).
 - **Highly condensed:** Rings sharing at least four bonds with others in the group, where the group has at least seven rings.
 - **Necklace:** A ring bridging at least one other ring half its size or less.
 - **Polyhedra:** A group of rings each of which is fused to at least two others in the group.
 - **Duplicates:** Locates records whose structures duplicate at least one other.

If **Clustered List** is selected, the query will be a list of duplicates only, clustered together (that is, a sorted list). This makes duplicates easier to compare, but you cannot edit records displayed in a sorted list.

If **Unique Structures Only** is selected, the query will be the same as the parent list, but with only one unique copy of each duplicate structure.
-  **Note:** The Duplicate search must be performed in **Full Structure** mode. If the Search type is set to **Substructure**, you will get an error message reminding you to reset the search type preference to Full Structure.
- **By date modified:** Locates structures updated recently or some other point in time. It applies to structures only, not to the other database fields. It is not implemented for Oracle at this time. Also, while it applies to the last time a structure was added or edited, it misses deletions. The way it works is to read the modification date written within the structure. Deleted records have no such date.

- Check structures: Locates records by checking the properties specified in the Check structures in database dialog box. Depending on the specified properties, and whether the database is read-only, it may be possible to automatically re-register the structure with a corrected version. By default, this scan begins at the currently displayed record and proceeds to the end of the record set, or until the scan is aborted. You can change this by unchecking the box at the bottom of the dialog box.

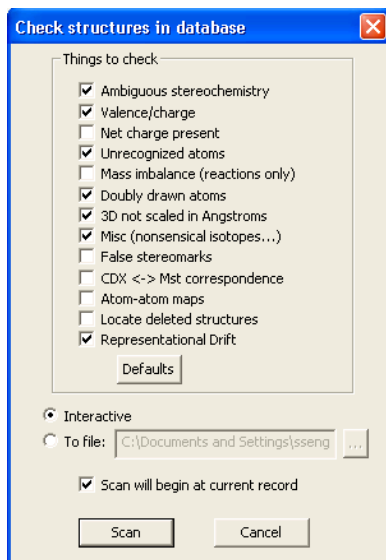



Figure 8.23: Check structures in database dialog box

The following properties can be checked:

- Ambiguous stereochemistry: Flags structures with ambiguously drawn tetrahedral or double bond stereochemistry. Not correctable.
- Valence/charge: Flags structures containing an atom with an impossible valence or charge. Sometimes correctable.
- Net charge present: Flags structures containing a net charge. This option is not checked by default because it is not normally useful. Not correctable.
- Unrecognized atoms: Flags structures containing an atom that was not interpreted by ChemBioFinder. Such atoms typically have atom type R with a grayed text label containing question marks. Not correctable.
- Mass imbalance: Flags reactions in which the molecular formulae of reactants is different from products. Not correctable.
- Doubly drawn atoms: Flags structures in which two atoms are very close together and at least one of which lacks a label. (Also, to qualify as a doubly drawn pair, the combined valence of explicit bonds must not exceed the capacity of either atom.) You get this sort of structure when attempting to draw a bond from an existing atom, but miss and accidentally create a new atom. ChemBioDraw flags these cases with a wavy red box.
- 3D not scaled in Angstroms: Flags structures with Z coordinates or 3D not scaled in Angstroms. Such structures will not be struck in a 3D search.
- Miscellaneous: Checks whether isotopic masses are reasonable.

- False stereomarks: Flags stereochemical indicators such as a wedged bond on an atom that is incapable of stereochemistry because of its geometry. In a normal search, such indicators are ignored. The option lets you find all such stereochemical indications and correct any errors.
- CDX <-> Mst correspondence: Checks whether the chemical interpretation of the original CDX drawing remains same in the latest version of the software. It also checks whether the stored internal structures need to be regenerated from the CDX.
- Atom-atom maps: Checks whether the quality of auto-generated maps is satisfactory. However, user-defined maps are not checked, and are considered legitimate.
- Locate deleted structures: Identifies structures that have been deleted from structure boxes.
- Representational Drift: Checks several invisible aspects of stored chemical structures and ensures that the stored chemical structures are up to date with the latest software version.

 **Note:** When the Interactive option is selected, the application displays dialog boxes with warnings and errors, if any. However, if you select the To file option, you need to provide the path for storing the log file containing warnings and errors. In this case, errors and warnings are not displayed in dialog boxes.

To perform a **Check structures** search:

1. Select the categories of interest.
2. Click **Scan**. When a structure is flagged, its record is displayed, the offending atom or bond (if any) is “flashed” and an alert offers the user several choices.

The options are:


if you want to...	then click...
ignore the error and continue checking the same structure.	Don't know or ignore
rewrite the MST record, thus correcting the problem.	is specified or has stereo
Continue the scan from the next structure.	Skip to next structure
stop the scan. This leaves you on the last record that exhibited a problem	Stop Scanning
“flash” the problematic atom or bond.	Flash

if you want to...	then click...
This button only appears when the error revolves around an atom or bond. Clicking the button creates a circle that moves toward the atom/bond, drawing the eye to it.	
View details about a warning message.	Help
Fix the error automatically.	Auto-Fix

Managing queries

The Queries Tree Control in the Explorer window maintains a list of search queries from the current and previous sessions. Queries are associated with forms, and the query list is saved when you save the form. When you open or activate an existing form, the tree updates to show only the queries for that form.

Each time you carry out a search or list operation, a new query is generated in the tree. A name is assigned to the query, and it is displayed in the tree along with the size of the list (number of hits) and a brief description. The generated name is "Q<n>*", where n is a sequential number, and * indicates that the query has not been named or marked for saving.

 **Caution:** Depending on your preference settings, unnamed queries may be automatically discarded. See ["Saving and restoring lists" on page 128](#) for details.

To activate the Queries tree control:

1. Go to **View>Explorer Window**.
2. Click the **Queries** tab on the Explorer window to display the Queries tree.

Queries are listed as "children" of the database (displayed as **Full List**). When you select **Search Over Current List** from the **Search** menu before performing a query, it is displayed as a "child" of the previous list. (The current list is indicated with a frame around the colored box.).

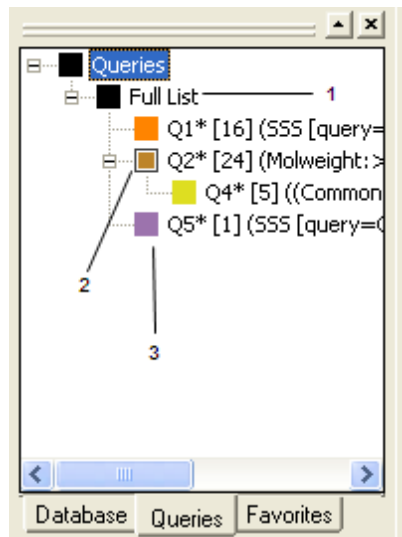



Figure 8.24: The Queries tree: 1) Database (full list); 2) Current List (marked with a frame); 3) Search over Current list query (child list).

To restore a query:

- Double-click a query.

 **Tip:** Double-clicking the root item, **Full List**, has the same effect as choosing **Retrieve All** from the **Search** menu.

The Queries context menu

When you right-click on a query, you display a menu with commands to manage the query.

if you want to...	then choose...
repeat the search using the selected query. (This is the same as double-clicking on the query.)	Rerun
merge the hits of the selected query with the current list. See "Restoring and merging hit lists" on page 129 for details.	Restore List

if you want to...	then choose...
apply the query's color to a plot of the intersection of the selected query with the current list. (This is the same as single-click on the query.)	Color On Plot
choose a different query icon color from the color picker dialog box.	Change Color
modify the query and/or rerun it. Restore Query goes to Query mode and displays the structure. You may modify the query, then rerun it manually.	Restore Query
save the query to a separate file, with extension .cfq.	Save Query...
mark the query so it will be saved (as named) when the form is saved. This has no effect unless the Query Saving preference has been set to Discard unnamed queries . See "Saving and restoring lists" on page 128 .	Keep
delete the selected query from the tree. The next time you save the form, this query will be lost. If	Remove

if you want to...	then choose...
you delete by mistake, you can close and reopen the form without saving.	
rename the query. When the Query Saving preference has been set to Discard unnamed queries , this has the same effect as Keep .	Rename
for all queries at the same depth in the tree, sort alphabetically by name.	Sort Folder

You also use the context menu to display the Query Properties dialog box. In this dialog box, you can modify the query name, add a comment, or change the color used in plots. Text entered in the **Comments** box will replace the standard description displayed along with the number of hits.

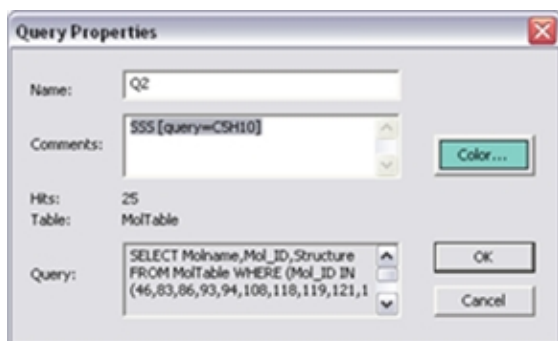


Figure 8.25: The Query Properties dialog box


Domains

If you are browsing a list other than the full database, you can make it seem that the database consists only of the current “child” list.

- Go to **Search>Domains>Set Domain to Current List**. The current list acts like a full database. For example, a **Retrieve All** command will retrieve only this list; without selecting **Search Over Current List**, searches will go over it only, etc.

To reset searches to the full database:

- Go to **Search>Domains>Reset to Full Database**.

 **Note:** *Set Domain to Current List* remains in effect until canceled. If you save the form, it will be saved with the form file and will still be in effect when you open the form again.

Saving and restoring lists

You can save queries individually as .cfq files or automatically when saving the form. The old commands for saving and restoring lists (as .cfq files) still appear on the **Search** menu, and on the context menu displayed when you right-click on a specific query. If you close a form without having saved your queries (or other changes), you are prompted to save or discard changes.

On the **General** tab of the **Preferences** dialog box, there is an option to save all queries (the default), or only those that have been renamed. If you select this option, there are two ways to save a query with the form:

- Rename the query, either with the **Rename** command in the context (right-click) menu, or from the query properties dialog box (also accessed from the context menu.)
- Mark the query with the **Keep** command on the context menu. Query saved in this way retain their standard name.

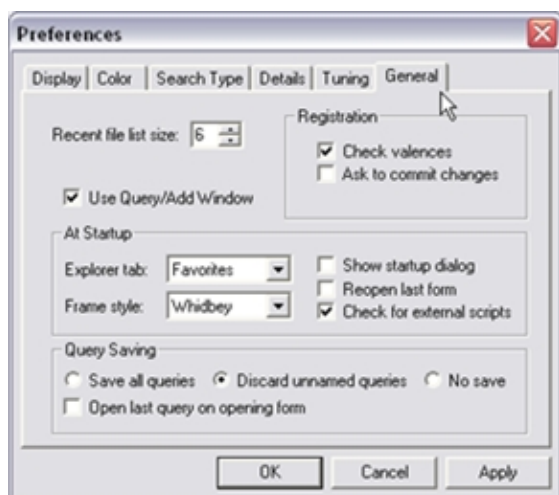


Figure 8.26: General preferences

Saving a hit list

Create a query and perform the search. See ["Entering and submitting a query "](#) on page 117 and ["Entering a structural query"](#) on page 118 for more information.

- When you have received the hit list, go to **Search>Save List**. The Save As dialog box appears.
- Type a filename for the hit list, and click **Save**.

The file consists of a list of values of the primary key of each record in the hit list. If there is no primary key, the molecule ID is used.

To edit a saved hit list, open the text file with a text editor such as Notepad and edit the values you want changed.

Restoring and merging hit lists

Once you have saved a hit list, you can perform another search and then integrate the results from the second search with the results from the first. Thus, you can perform Boolean operations on different searches.

There are two ways to integrate the results of two searches:

Browsing to the saved list:

1. Perform a search and save the hit list.
2. Perform another search.
3. Go to **Search>Restore List**. The Open dialog box appears.
4. Browse to the previously saved list you want to integrate with the current hit list, and click **Open**. The Restore/Merge List dialog box appears.

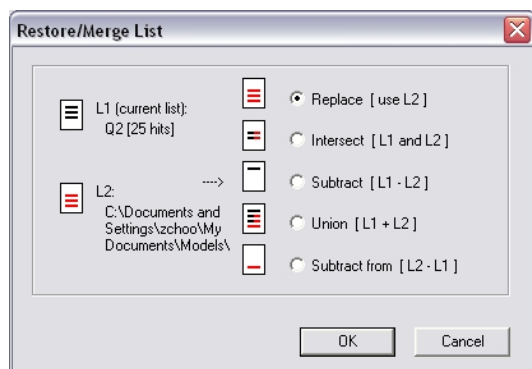


Figure 8.27: The Restore/Merge List dialog box

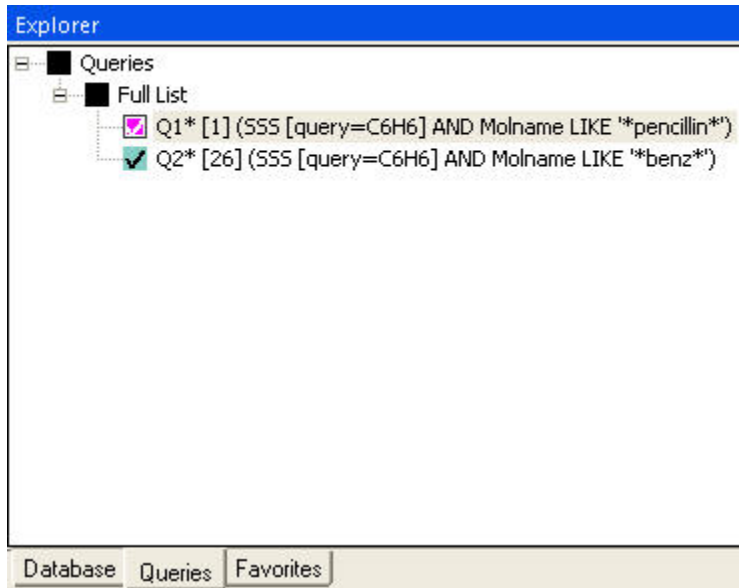
5. Select the integration option and click **OK**.
 - **Replace**: Discards the current hit list and displays the records from the hit list you chose to open.
 - **Intersect**: Displays only those records that appear in both the current hit list and in the saved hit list. This is a Boolean AND operation.
 - **Subtract**: Displays only those records that are in the current hit list but are not in the saved hit list. For example, if the current list contains records 1, 2, 3, 4, and 5, and the restored list contains records 4, 5, 6, and 7, then only records 1, 2, and 3 will be displayed. This is a Boolean NOT operation; it shows the records in the current list that are not also in the restored list.
 - **Union**: Displays all records in either the current list or the saved hit list. This is a Boolean OR operation.
 - **Subtract from**: Displays only those records that are in the saved hit list but are not in the current hit list. This is the same as Subtract, but with the order of the lists reversed.

The hit lists are integrated. You can browse and save this new hit list just as any other hit list.

Drag and Drop list merge:

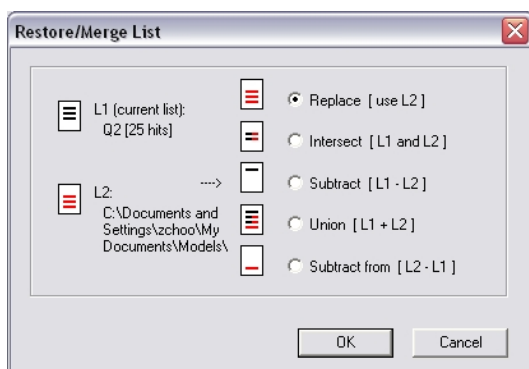
1. Perform a search and save the hit list.
2. Perform another search and save the list.

Two saved queries are shown in the figure below:




3. Simply drag one query item in the tree on top of another.

Restore/Merge dialog box comes up allowing you to choose the logic to be applied to the merge: replace, intersect, subtract, union, and subtract from.



A successful merge results in a new list.

 **Note:** If you merge two lists and get an empty result, an alert appears and the list reverts to the previous list.

Search examples

The following are examples of searching options. By specifying atom and bond properties, you see how to use the query functions in ChemBioDraw to search the database more effectively. All substructure search query properties recognized by ChemBioFinder are listed and described in ["Structural Query Features" on page 222](#)

Multiple hit lists

A detailed example of how you can use hit list management to perform specific, sophisticated searches follows.

Suppose you want to search for all compounds in the CS_Demo database that contain a benzene substructure and have molecular weights between 50 and 200. After you perform this combined search, you want to find which of these compounds do not contain a carbon-nitrogen bond. By integrating hit lists, you can perform this search.

To create the hit list for the first part of this search:

1. Open the **CS_Demo** database.
2. Create the first query by drawing benzene in the Structure data box and entering 50-200 in the Molecular Weight Data Box.
3. Search the query. You get 75 hits.
4. Go to **Search>Save List**.
5. Save the hit list as **benz.txt**.

To create another hit list for the second part of the search:

1. Draw the query shown below in the Structure Data Box.



2. Search the query. You get 103 hits.
3. Go to **Search>Save List**.
4. Save the hit list as c-n.txt.

Integrate the two hit lists:

1. Go to **Search>Restore List**.
2. In the **Open** dialog box, open the **benz.txt** file.
3. In the **Restore List** dialog box, click the **Replace** radio button, then click **OK**. You return to your first list of 75 hits.
4. Go to **Search>Restore List**.
5. In the **Open** dialog box, open the c-n.txt file.
6. In the **Restore List** dialog box, click the **Subtract** radio button, then click **OK**.

The hit list is reduced to the 60 records with compounds containing a benzene substructure, not containing a C-N substructure, and having molecular weights between 50 and 200.

If you had chosen the **Replace** option in the last step, then only the C-N records would be displayed. Choosing **Intersect** would display those records in either list. Finally, choosing **Union** would display all records from both lists.

Using atom lists

Using the text tool in ChemBioDraw, you can enter in a structural query a list of possible atom types, one of which must match in the target compound.

If you want to search for molecules containing a benzene ring with an ether, amine or phosphane group, a query might look like the following:

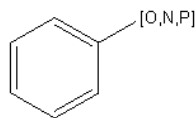


Figure 8.28: Atom list query

Atom types and bond types

To broaden or narrow a search query, you can define the properties of the atoms and bonds in a structural query. These properties are definable in ChemBioDraw Pro using the Atom Properties and Bond Properties dialogs accessible from the Structure menu. See ["Structural Query Features" on page 222](#) for more information. Suppose you want to find all molecules that contain a non-oxygen chalcogenide bonded to another atom, not necessarily carbon. You also want the bond type between the chalcogen and the other atom to be a single or double bond. The query, drawn in ChemBioDraw Pro, may look like this:

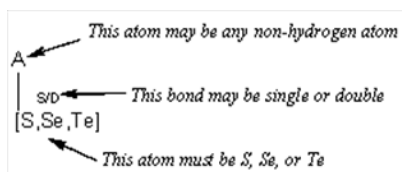


Figure 8.29: Specifying bond and atom types in a query

The “A” label denotes that the atom may match any atom except hydrogen. The indicator near the bond indicates that the bond has been defined; in this case, in the Bond Properties dialog of ChemBioDraw, you specified that the bond type may be single or double, S/D. Finally, by entering “S,Se,Te” enclosed in brackets, you specified that one of these elements must match in the target molecules.

Searching fullerenes

A search of fullerenes can illustrate the restrictions you can place on a formula search. Suppose you want to find all fullerenes containing 20 to 80 carbon atoms, but you also do not want to include large organic molecules.

The formula query could be:

C20-80 H0

By designating zero hydrogen atoms, you exclude hydrocarbons from the hit list. By clearly capitalizing the elements and spacing the query, you avoid searching ambiguities, although there were no ambiguities in this example.

Link nodes and multivalent Rs

A link node is a placeholder for zero or more unspecified atoms. It is especially useful when searching for targets with any of several ring sizes, or chains substituted in particular ways at either end. While link nodes are traditionally divalent, ChemBioFinder places no constraints on the upper limit of the range. Multi-valent R (and A) atoms are functionally similar to link nodes. They may be thought of as “higher valent” link nodes.

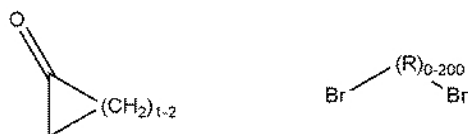


Figure 8.30: Link nodes

The traditional representation of a link node is CH₂, but this is misleading because any atom type can substitute for the link node. For this reason, represents link nodes as R. R (without a repeat count) can match any number of atoms, including zero.

although any atom type of a target atom matches a link node, you can impose two restrictions the on the link node:

- Unsaturation
- Ring topology.

Using ChemBioDraw's Atom Properties menu to apply either of these properties to the link node restricts the range of target atoms to the specified topology. For example, specifying the link node to be aromatic would give the following results.

:

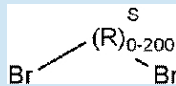
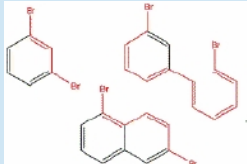
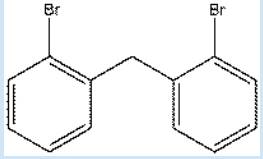
Query	Results
	Will hit 
	Will not hit 

Figure 8.31: Results of a link node query

Searching more than one substructure

A substructure search may contain more than one substructure unit. Suppose you want to find all compounds in the CS_Demo database which contain a benzene substructure and another substructure unit containing chlorine bonded to any atom. A structure query, as drawn in ChemBioDraw would look like:

A—Cl

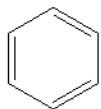


Figure 8.32: Structure query with two substructures

You should get five hits. Browse the hits. By specifying the query as above, you obtain hits such as benzyl chloride, shown below, where the substructure units are not connected.

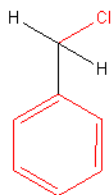


Figure 8.33: Benzyl chloride, hit by query with two substructures

The substructure units can overlap; they can share a common atom. Examples of this overlap are shown below.

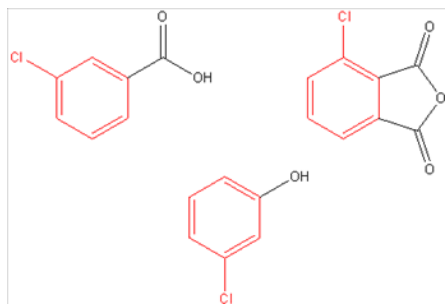



Figure 8.34: Overlapping substructures in the hit list

Compound Profiles

Compound Profiles, a way of visually comparing and ranking structures based on values of selected properties. A compound profile is a graph, or, more typically, a series of graphs, each of which plots the value of a numerical field in the database on a vertical bar color-coded to show regions of “cost”, or undesirability. These regions are defined by a cost profile associated with the field. In the profile, costs are normalized to score between 0 and 1, where the ideal is assigned the value of zero, and the slope of the deviation to maximum (1) is user-defined.

 **Note:** The function is a normalizing linear spline.

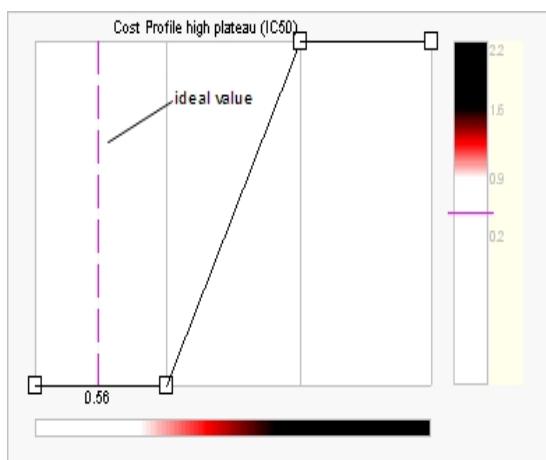


Figure 9.1: IC50 profile.

The cost profile is designed by the user to reflect their opinion of what constitutes a “good value” for a property. The user can adjust the cost profile to reflect how unhappy they would be when a compound deviates from the ideal value by changing the range of values at zero and the slope of the line to one.

For example, a user might want a drug candidate to have a good solubility (in units of mg/ml) and good biological activity (in units of nm). They might be willing to sacrifice some activity to get better solubility, so the cost function for solubility would be adjusted to penalize any deviation from some “ideal” range more severely than the corresponding cost function for activity.

The default display preference is to have white regions in the bars represent ideal values, darker regions increasing deviation from the ideal.

The normalization function permits properties of different types to be compared. A score, the total deviation from “ideal” of all fields, is displayed. Each property can be assigned a weight to adjust its effect on the overall calculation. The score is the same for a given compound regardless of whether or not you’ve done a sort, or where the compound appears within a hit list. A “bad” compound has a high score.

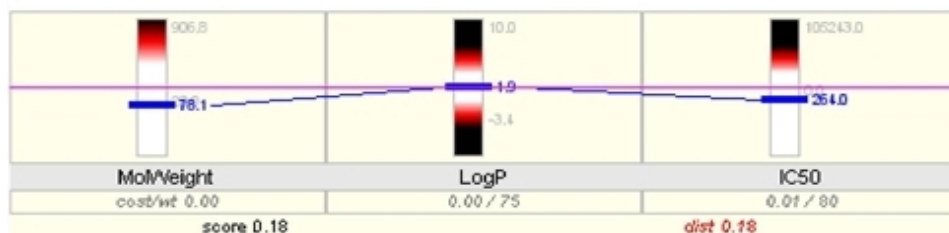


Figure 9.2: A compound profile for MW, LogP, and IC50.

A compound profile is useful as a visual indication of the “quality” of each compound. As such, it is most useful in table view, where you can see and compare several profiles at once. The highest benefit comes from using profiles in sort/search operations. You can pose one compound as the query, and then sort the remaining ones by deviation from that query.

Creating a compound profile

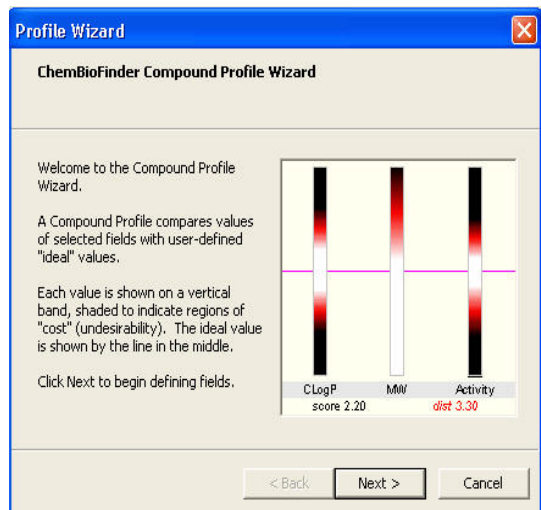
Profiles are displayed on the form in a special data box, created with the Profile tool.

To create a Profile box:

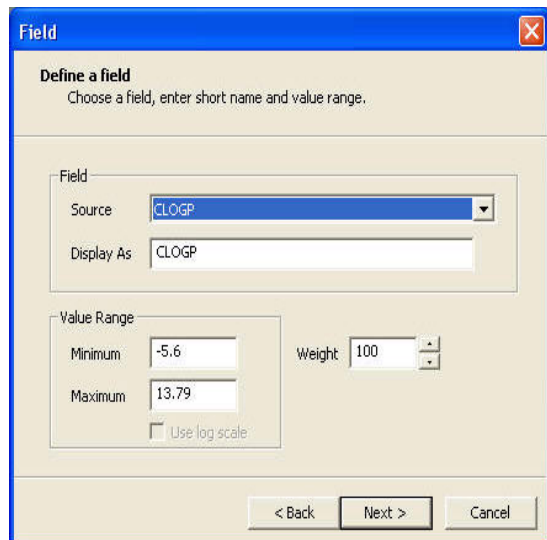
The database connected to the form or subform you are working with must have at least one numeric field.

 **Note:** You can also include structure field in profile.

1. Click the Profile tool on the Form toolbar.
2. Drag a box on the form. The Profile Wizard appears then click Next.

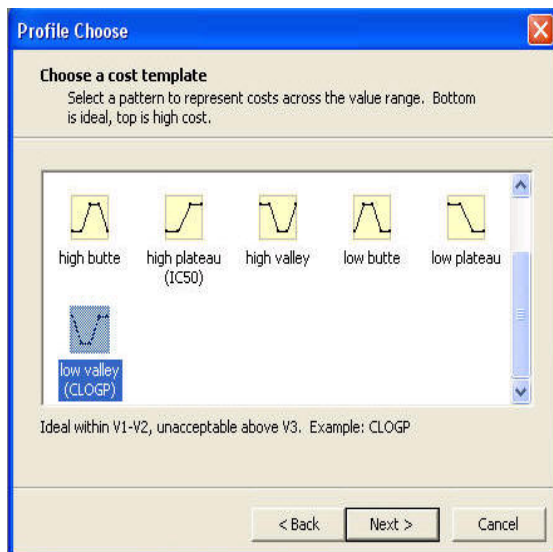


Define Field dialog box is displayed.



3. On the Field page, choose a numeric field from the drop-down menu of Source.
 - You may edit the name automatically entered in the **Display As** text box.
 - You may edit the Value Range.

The default display is the actual low/high values for the first 1000 records in the database.
 - You may choose a log scale, if no part of the Value Range is negative.
 - You may weigh this profile relative to other profiles; where 100 is important, 0 is not.
4. On the Profile Choose page, choose a cost profile.



Choosing from this gallery is just for convenience. Any pattern you choose can be later modified to any other shape.

Tip: As you click each profile, a line of descriptive comments is displayed below the gallery window, to aid your selection.

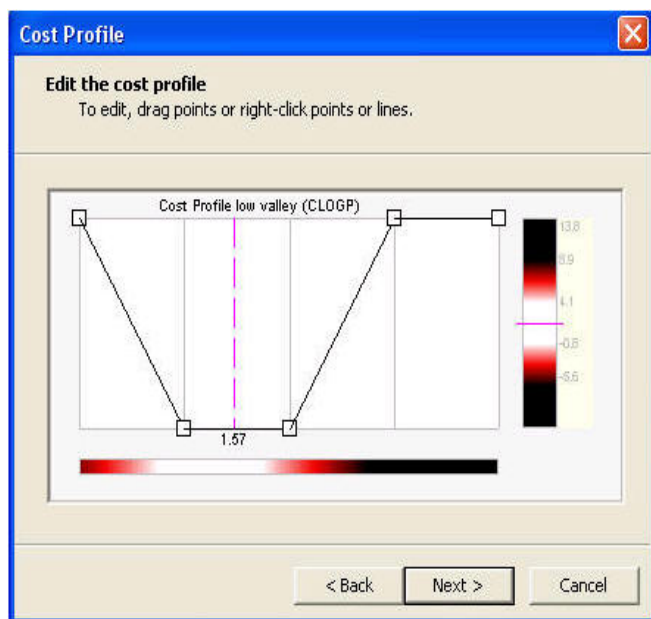
- On the Cost Profile page, you may edit your profile. See ["To change a cost profile:" on page 139](#) for details.
- The Profile page of the wizard has an “add another field” check box that changes the **Finish** button to a **Next** button. Click **Next** to repeat the procedure for each profile you want to add. When you have added all of your profiles, click **Finish**. The profiles are displayed in the data box.

Editing a compound profile

You can edit any of the characteristics of a compound profile with the Compound Profile Editor.

To edit a profile:

- Double-click in the profile data box. The Compound Profile Editor appears.



To add a new profile:

- Right-click in the editor window, then click the **Add Field** context menu command.

To delete a profile:


- On the **Profile** tab, select the required profile.
- Right-click in the editor window, then click **Remove Field** in the context menu.

To change the order of the profiles:

- On the **Profile** tab, click one of the profiles to select it.
- Right-click in the editor window, then click the **<Move Left** or **>Move Right** context menu command.

To change the field associated with a profile:

1. On the **Profile** tab, click one of the profiles to select it.
2. Click the **Field** tab. Change the **Source** with the drop-down menu. Change any of the other variables as required.
3. Click the **Cost Profile** tab and select a new profile.

 **Note:** You must select a new profile when you change the source field. If you forget, you will be prompted with an error message.

4. Click **OK**.
The cost profile is displayed, and may be edited.
5. Click **OK** to complete the change.

To change a cost profile:

1. On the **Profile** tab, click one of the profiles to select it, then click the **Cost Profile** tab.

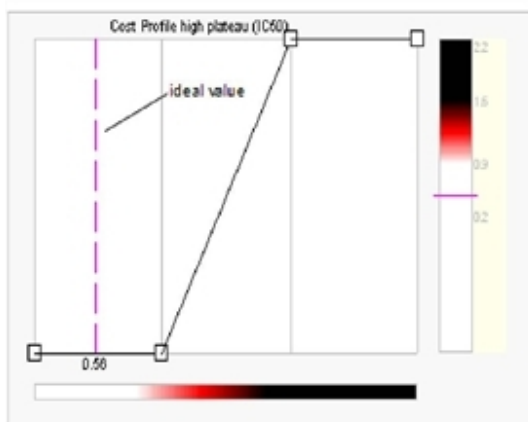


Figure 9.3: Cost profile, showing the drag points and ideal value line.

2. To change an inflection point, drag it to a new location.
As you drag, the X and Y coordinates of the point are displayed on the grid.

 **Tip:** You cannot drag past a neighboring point.

3. To add a new inflection point, right click on the line segment where you want the new point, and click the **Add point here** context menu command.
4. To delete a point, right-click on the point and click **Remove point** in the context menu.
5. To change the ideal value, drag the pink dashed line on the grid to a new location.
6. To choose a different template, right-click anywhere in the editor window, click the **Choose** context menu command, and select a new template.
7. Click **OK** to save changes and exit the editor, or **Cancel** to exit without saving.

Display preferences

You can change display characteristics of a profile with the Profile Display Preferences dialog box. Changes you make apply only to a selected profile box, unless you choose to reset the defaults.

To change the profile display preferences:

1. Right-click a profile box and click the **Compound Profile > Preferences...** context menu command. The Profile Display Preferences dialog box appears.
2. To change the base color of the cost band (default = red), select a primary color in the **Cost Shading Color** section.
3. To reverse the “polarity” of the cost band, deselect the **Ideal = white** check box. When checked, light shades are “good”, dark are “bad”; when unchecked, the reverse is true.
4. To change the color of the value marks and the lines between them, click **Values** and select a color from the color picker.
5. To change the color of range labels, click **Labels color** and select a color from the color picker.
6. To turn off the range labels (to the right of the cost band), deselect the **Labels** check box.
7. To use the current settings as the defaults for all subsequent profile creations, click **Use As Defaults**.
8. When you have finished editing the settings, click **OK** to save the changes, or **Cancel** to discard.

Search/sort

You get the most advantageous use of profiles in comparing compounds by using the search/sort function. Select one compound as your base (for example an “ideal” compound with an overall score of zero), then sort the database by increasing deviation from that score.

To use Search/Sort to compare profiles:

You must be in Form View to begin this procedure.

1. Browse to the compound you want to use as your base.
2. Right-click in the Profile data box, then select **Compound Profile>Search/Sort>Sort relative to current** from the context menu. The database is sorted.
3. For easiest comparison, switch to Table View.

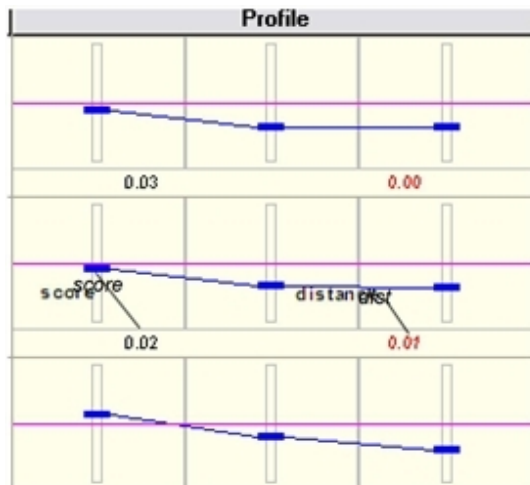



Figure 9.4: Profiles in the table view

The dist value in the profile (displayed in red at the bottom of the chart) changes when you sort. Before sorting, it matches the [score](#) value. After sorting, it displays the deviation of that compound from the base. If you sort relative to a compound, then that compound gets dist = 0 since it defines the reference point. The dist value for other compounds will increase as they become less similar to the base. You can also sort relative to the ideal value. In this case, the dist value for each compound is the same as its score, and the records are ordered on that basis.

 **Tip:** To reset the database, right-click in the Mol_ID field and click the **Sort Column>Ascending** context menu command.

Example:

Do the following:

1. In the CS Demo database, create a profile with one field, Molweight, and a cost profile such that values below 100 are ideal, and above 100 increasingly bad.
2. Browse through the compounds.

Benzene (mw 78) falls in the ideal region, so it has a cost of zero and a score of zero. Buckyball fullerene (mw 800+) falls way outside the ideal, so it has a high cost and a score of 1.
3. Right-click in the profile box.
4. Click the **Compound Profile>Search/Sort>Sort relative to ideal** context menu command.
5. Click the Switch to Table icon to make it easier to compare the profiles.

The compounds are ordered by dist value from 0.00 to 1.00. Note that the score and dist values are the same.
6. Browse to bromobenzene (mw 157).

The score for this compound is not ideal. It will be approximately 0.1, depending on how you set the profile.

 **Tip:** You can sort by Molname or MW to find the compound more easily.

7. Click the Switch to Table icon again to return to form view.
8. Right-click in the Profile data box, then select **Compound Profile>Search/Sort>Sort relative to current** from the context menu.
9. Switch back to Table View and browse the results.
Bromobenzene is at the top of the list. Other compounds with similar scores follow. Note that the score and dist values are no longer equal.

Customizing

The gallery of cost profile templates shown in the Cost Profile Chooser dialog box was designed to help you get started. Some are based on user experience, other chosen to provide the most common cost profile patterns. As you work with the compound profiler, you will develop your own templates and may want to substitute them in the chooser.

The templates are taken from files stored in the ...\\ChemBioFinder\\system\\profiles\\ folder. The files have a CPX extension, and the name of each file is the name given to the template. The Cost Profile Chooser displays all files in this directory having the CPX extension. The internal format is XML, readable in Notepad.

To remove a template from the chooser, delete the corresponding file, or move it to another location, or change its extension.

To add or modify a template:

1. Double-click in the profile data box to open the Profile Editor, then click the **Cost Profile** tab.
2. Edit the profile until it looks the way you want.
3. Right-click in the edit window, then select **Save...** from the context menu.
4. When prompted, enter a line of descriptive comments to be displayed below the chooser window, and click **OK**.
5. Type the name you want to appear in the chooser, and click **Save**.

Chemical Properties

ChemBioOffice lets you determine more than 100 different chemical, topological, physical, and thermodynamic properties. This guide describes the properties you can calculate using ChemBioFinder and lists the property calculations that are available in other ChemBioOffice applications.

- [ChemBioFinder Chemical Properties](#)
- [ChemBio3D Chemical Properties](#)
- [ChemBioDraw properties](#)

ChemBioFinder Chemical Properties

To view the properties for a given structure:

1. Right-click in the form and select **calculate properties** from the context menu.
2. Select the desired property in the Property Populator window.
3. Click **Populate**.

The properties listed below can be calculated in ChemBioFinder Ultra for ChemBioOffice. For more information on these properties and how to calculate them, see the ChemBioFinder online help.

B	H	P
Balaban Index	Heat of Formation	pKa
Boiling Point	Henry's Law Constant	Polar Surface Area
C	I	Principal Moment
Chemical Name	Ideal Gas Thermal Capacity	R
Cluster Count	L	Radius
Connolly Accessible Area	LogP/Partition Coefficient	S
Connolly Molecular Area	LogS	Shape Attribute
Connolly Solvent Excluded Volume	M	Shape Coefficient
Critical Pressure	Mass	Sum Of Degrees
Critical Temperature	Melting Point	Sum of Valence Degrees
Critical Volume	Mol Formula	T
E	Mol Formula HTML	Topological Diameter
Elemental Analysis	Mol Refractivity	Total Connectivity
Exact Mass	Mol Weight	Total Valence Connectivity
F	Molecular Topological Index	V
Formal Charge	m/z	Vapor Pressure
G	N	W
Gibbs Free Energy	Num Rotatable Bonds	Water Solubility
	O	Wiener Index
	Ovality	

About Molecular Networks

*MOSES.pKa, MOSES.logS and MOSES.logP for predicting acid dissociation constants, aqueous solubility and octanol/water distribution coefficients of chemical compounds are computational calculator modules based on Molecular Networks' cheminformatics platform MOSES. MOSES is developed, maintained and owned by Molecular Networks GmbH, Erlangen, Germany. For further information please visit www.molecular-networks.com/moses. All rights reserved. Molecular Networks GmbH, Erlangen, Germany (www.molecular-networks.com).

Limitations

Only the following atom types and hybridization states are parameterized:

C _{sp} ³	C _{sp} ²	C _{aromatic}	C _{sp}
N _{sp} ³	N _{sp} ²	N _{amide}	N _{aromatic}
N _{sp}	O _{sp} ³	O _{sp} ²	P _{sp} ³
S _{sp} ³	S _{sp} ²	S _{sulfoxide}	S _{sulfone}
F	Cl	Br	I

Balaban Index

Is represented as, J and is defined by the formula:

$$J = \frac{q}{\mu + 1} \sum_{edges\ ij} (S_i S_j)^{-1/2}$$

Where:

- q is the number of edges in the molecular graph.
- $\mu = (q - n + 1)$ is the cyclomatic number of molecular graph.
- n is the number of atoms in the molecular graph.
- S_i is the sum of all entries in the i^{th} row (or column) of the topological distance matrix of the molecule. The distance matrix stores the shortest path between all pairs of atoms.

Boiling Point

The boiling point is report in Kelvin.

A solution boils at a slightly higher temperature than the pure solvent. The change in boiling point is calculated using the formula:

$$\Delta T_b = K_b m$$

Where K_b is the molal boiling point constant, m is the concentration of the solute expressed as molality, and ΔT_b is the change in temperature.

Chemical Name

This is the name of the chemical.

Cluster Count

Is the number of paths of a given length in the distance matrix.

Connolly Accessible Area

This is the locus of the center of a spherical probe (representing the solvent) as it is rolled over the molecular model.

Connolly Molecular Area

This is the contact surface created when a spherical probe (representing the solvent) is rolled over the molecular model.

Connolly Solvent Excluded Volume

This is the volume of space bounded by the solvent accessible molecular surface.

Critical Pressure

Reported in bars, this is the least applied pressure required at the critical temperature to liquefy a gas.

Critical Temperature

Reported in Kelvin, this is the temperature above which a gas cannot be liquefied, regardless of the pressure applied.

Critical Volume

Reported in cm^3/mol , this is the volume occupied by one mole of a substance at the critical temperature and pressure.
To report the boiling point:

Elemental Analysis

This is a process in which a material is analyzed for its elemental and isotopic composition. Elemental analysis can be qualitative (determining what elements are present) or quantitative (determining how much of each are present).

Exact Mass

This is the sum of the masses of the individual isotopes in the molecule. For example, the exact mass of water containing two hydrogen-1 (1H) and one oxygen-16 (16O) is $1.0078 + 1.0078 + 15.9994 = 18.0106$. The exact mass of heavy water, containing two hydrogen-2 (deuterium or 2H) and one oxygen-16 (16O) is $2.014 + 2.014 + 15.9994 = 20.027$.

Formal Charge

This is the charge assigned to an atom in a molecule, assuming that electrons in a chemical bond are shared equally between atoms, regardless of relative electronegativity.

The formal charge of any atom in a molecule can be calculated using the formula:

$$FC = V - N - B/2$$

Where 'V' is the number of valence electrons of the atom in isolation (atom in ground state); 'N' is the number of non-bonding electrons on this atom in the molecule; and 'B' is the total number of electrons shared in covalent bonds with other atoms in the molecule.

Gibbs Free Energy

Reported in KJ/mole, Gibbs free energy is defined as:

$$G(p, T) = U + pV - TS$$


which is same as:

$$G(p, T) = H - TS$$

Where:

- U is the internal energy (SI unit: Joule)
- p is pressure (SI unit: Pascal)
- V is volume (SI unit: m³)
- T is the temperature (SI unit: Kelvin)
- S is the entropy (SI unit: joule per Kelvin)

- H is the enthalpy (SI unit: Joule)

 **Note:** *H and S are thermodynamic values found at standard temperature and pressure.*

Heat of Formation

Reported in KJ/mole, the heat of formation is the increase in enthalpy resulting from the formation of one mole of a substance from its constituent elements at constant pressure.

Henry's Law Constant

A unitless value, Henry's Law Constant can be expressed as:

$$k_{H,pc} = \frac{p}{c}$$

Where 'p' is the partial pressure of the solute in the gas above the solution, 'c' is the concentration of the solute and ' $k_{H,pc}$ ' is a constant with the dimensions of pressure divided by concentration. The constant, known as the Henry's law constant, depends on the solute, solvent, and temperature.

Ideal Gas Thermal Capacity

The thermal capacity at constant volume of an ideal gas is:

$$C_V = \left(\frac{\delta U}{\delta T} \right)_V = \hat{c}_V N k_B = \hat{c}_V n R$$

Where:

- \hat{c}_V is a constant dependent on temperature
- U is the internal energy
- T is the absolute temperature
- V is the volume
- n is the amount of substance of the gas
- R is the gas constant ($8.314 \text{ JK}^{-1} \text{ mol}^{-1}$ in SI units)
- N is the number of gas particles
- k_B is the Boltzmann constant ($1.381 \times 10^{-23} \text{ JK}^{-1}$ in SI units)

The thermal capacity at constant pressure of an ideal gas is:

$$C_p = \left(\frac{\delta H}{\delta T} \right)_p = (\hat{c}_V + 1)Nk$$

Where 'H' is the enthalpy of the gas, calculated as: $H = U + pV$.

LogP/Partition Coefficient

The partition coefficient is a ratio of concentrations of un-ionized compound between two solutions. To measure the partition coefficient of ionizable solutes, the pH of the aqueous phase is adjusted so that the predominant form of the compound is un-ionized. The logarithm of the ratio of the concentrations of the un-ionized solute in the solvents is called LogP.

Normally, one of the solvents chosen is water; the second is hydrophobic, such as octanol.

The formula to calculate logP is:

$$\log P_{oct/wat} = \log \left(\frac{[solute]_{octanol}}{[solute]_{un-ionized water}} \right)$$

LogS

The logarithm of the solubility of a substance, measured in mol/liter.

Mass

This is the mass of one molecule of a substance, relative to the unified atomic mass unit u (equal to 1/12 the mass of one atom of ^{12}C). Also called molecular mass.

Melting Point

Is the temperature at which a solid becomes a liquid at standard atmospheric pressure. At the melting point, the solid and liquid phase exist in equilibrium. The formula to calculate melting point is:

$$T = \frac{\Delta H}{\Delta S}$$

Where 'T' is the temperature at the melting point, ' ΔS ' is the change in entropy of melting, and ' ΔH ' is the change in enthalpy of melting.

Mol Formula

This is the molecular formula of the compound.

Mol Formula HTML

This is the molecular formula written in HTML format. You can copy and paste the output to an HTML file.

Here is an example of the output for a model of aniline:

C₆H₇N

Mol Refractivity

Is a measure of the total polarizability of one mole of a substance. It is dependent on the temperature, index of refraction, and pressure.

The molar refractivity, 'A' is expressed as:

$$A = \frac{4\pi}{3} N_A \alpha$$

Where 'N_A' is the Avogadro constant and α is the mean polarizability.

Mol Weight

This is the sum of the atomic weights of all atoms in a molecule.

Molecular Topological Index


Is a graph index defined by:

$$MTI = \sum_{i=1}^n E_i$$

Where E_i are the components of the vector:

$$E = (A + D)d$$

Where, A is the adjacency matrix and D is the graph distance matrix, and d is the vector of vertex degrees of a graph.

 **Note:** An atom's degree is the number of non-hydrogen atoms to which it is adjacent.

m/z

The mass-to-charge ratio is a physical quantity used in the electrodynamics of charged particles. Two particles with the same m/z ratio move in the same path in a vacuum when subjected to the same electric and magnetic fields.

Num Rotatable Bonds

Is the number of rotatable bonds in a molecule. A rotatable bond is defined as an acyclic bond drawn as a single bond (amide C-N bonds, O=C-N, are rotatable) where the atom on each end has at least one other non-hydrogen substituent. This topological parameter is a measure of molecular flexibility. Unsaturated bonds, and single bonds connected to hydrogens or terminal atoms, single bonds of amides, sulphonamides, and those connecting two hindered aromatic rings (having at least three ortho substituents) are considered non-rotatable.

Ovality

This refers to a measure of how the shape of a molecule approaches a sphere (at one extreme) or a cigar shape (at the other). Ovality is described by a ratio of volume to area:

$$O = A / (4 \times \pi \times ((3 \times V) / 4 \times \pi)^{2/3})$$

Where 'A' is area, 'V' is Volume, and 'O' is Ovality.

The ovality of a helium atom is 1.0 and HC₂₄H (12 triple bonds) is ~1.7 (equivalent to 1.7).

pKa

The pKa describes the tendency of compounds or ions to dissociate in solution. If B is defined as the conjugate base of the acid HB, then pKa is calculated as:

$$\text{pKa} = -\text{Log}_{10}(\text{Ka})$$

where

$$\text{Ka} = [\text{H}^+][\text{B}] / [\text{HB}]$$

Polar Surface Area

The Polar Surface Area (PSA) is defined as the surface sum over all polar atoms (usually oxygen and nitrogen) including attached hydrogens.

Principal Moment

This refers to the principal moments of inertia about the principal axes of a molecule. The moments of inertia are computed for a series of straight lines through the center of mass using the formula:

$$I = \sum_i m_i d_i^2$$

Where, 'I' is the moment of inertia, 'm_i's are point masses whose distances from the rotation axis are denoted by 'd_i's.

- Distances are established along each line proportional to the reciprocal of the square root of I on either side of the center of mass. The locus of these distances forms an ellipsoidal surface. The principal moments are associated with the principal axes of the ellipsoid.
- If all three moments are equal, the molecule is considered to be a symmetrical top. If no moments are equal, the molecule is considered to be an unsymmetrical top.

Radius

The eccentricity of an atom is the largest value in its row (or column) of the distance matrix, and represents how far away from the molecular center it resides.

The diameter (D) is the maximum such value for all atoms, and is held by the most outlying atom(s). Examples: Diameter of methane = 0; ethane = 1; propane = 2; n-butane = 3.

The radius R is the minimum such value, and is held by the most central atom(s). Examples: Radius of methane = 0; ethane = 1; propane = 1; n-butane = 2.

Shape Attribute

The shape attribute (kappa) measures the branching of a molecule, and is scaled so as to fall between the minimum and maximum values possible for the given order. The first-order shape attribute counts the number of one-bond paths. The second-order attribute counts the number of two-bond paths, and so on. The first three orders (1..3) are available.

Shape Coefficient

The shape coefficient, I is given by:

$$I = (D - R) / R$$

Where the diameter (D) is the maximum such value for all atoms, and is held by the most outlying atom(s). The radius (R) is the minimum such value, and is held by the most central atom(s).

Examples values for D: methane = 0; ethane = 1; propane = 2; n-butane = 3.

Example values for R: Radius: methane = 0; ethane = 1; propane = 1; n-butane = 2.

Sum Of Degrees

Is the sum of degrees of every atom. An atom's degree is the number of nonhydrogen atoms to which it is bonded.

Sum of Valence Degrees

Is the sum of degrees of every atom. An atom's valence degree is equal to the sum of its adjacent bonds' orders, including hydrogens.

Topological Diameter

Is the longest dimension of a molecule.

Total Connectivity

Is the connectivity considered over all the heteroatoms.

Total Valence Connectivity

Is the valence connectivity considered over all the heteroatoms.

Vapor Pressure

Is the pressure exerted by a vapor in equilibrium with its solid or liquid phase. Vapor pressure measures the concentration of solvent molecules in the gas phase.

Vapor pressure can be calculated using Raoult's law. The formula to calculate vapor pressure is:

$$P_{tot} = \sum_i p_i \chi_i$$

Where 'p' is vapor pressure, 'i' is a component index, and ' χ ' is a mole fraction.

Water Solubility

Is the maximum amount of a substance that can be dissolved in water at equilibrium at a given temperature and pressure. Water solubility is measured in mg/L.

Wiener Index

Provides measure of branching defined as:

$$W = \frac{1}{2} \sum D_{ij}$$

Where D_{ij} are the off-diagonal elements of the distance matrix.

ChemBio3D Chemical Properties

B	G	Mol Weight	Spin Density
Balaban Index	Gibbs Free Energy	Molecular Mass	Sum Of Degrees
Boiling Point	H	Molecular Surfaces	Sum of Valence Degrees
C	Harmonic Zero Point Energy	Molecular Topological Index	T
C_p	Heat Capacity	Molecular Volume	Thermodynamic Energy
C_v	Heat of Formation	Mulliken Charges	Topological Diameter
Cluster Count	Henry's Law Constant	Mulliken Populations	Total Connectivity
Connolly Accessible Area	Hyper Polarizability	m/z	Total Energy
Connolly Molecular Area	Hyperfine Coupling Constants	N	Total Valence Connectivity
Connolly Solvent Excluded Volume	I	Num Rotatable Bonds	V
Critical Pressure	Ideal Gas Thermal Capacity	O	Vapor Pressure
Critical Temperature	Internal Energy	Ovality	W
Critical Volume	Ionization Potential	P	Water Solubility
D	K	Partition Coefficient	Wiener Index
Dipole	Kinetic Energy	pKa	Z
E	L	Polar Surface Area	Zero-Point Energy
Electron Density	LogP	Polarizability	
Electrostatic Potential	LogS	Potential Energy	
Elemental Analysis	Lowdin Charges	Principal Moment	
Enthalpy	Lowdin Populations	R	
Entropy	M	Radius	
Exact Mass	Mass	RMS Force	
F	Melting Point	S	
Formal Charge	Mol Formula	SCF Energy	
Frequencies	Mol Formula HTML	Shape Attribute	
	Mol Refractivity	Shape Coefficient	

ChemBioDraw Properties

The properties listed below can be calculated in ChemBioDraw Ultra for ChemBioOffice. For more information on these properties and how to calculate them, see the ChemBioDraw online help.

B	H	P
Boiling Point	Heat of Formation	pKa

C

Critical Pressure
Critical Temperature
Critical Volume
CLogP/CMR

G

Gibbs Free Energy

Henry's Law Constant

L

LogP
LogS

M

Melting Point
MR

T

tPSA

Relational Data and Subforms

Subforms lets you work with multiple data tables in a relational database. You should be familiar with working with ChemBioFinder forms and databases before proceeding to subforms.

Access relational data

A subform is a special box within a form that behaves like a separate form file. A subform is used to display information stored within a database, but it usually displays data from a different database (or different table within the same database) than the main form. A subform is usually linked to the main form, so that retrieval of data in the main form also retrieves related records in the subform.

Like other types of boxes, you can move select and resize a subform on the main form. However, you cannot select it by clicking inside it. You must click the title bar. When you click inside a subform, you activate the miniature form inside the box and can work within it.

Below are examples of subform use:

- You have data that is associated in a “one-to-many” relationship. For example, you want to view a chemical structure together with its physical properties, stored in a separate table.
- You are running a stockroom and you want to store package sizes and prices in the same place because they are related. Each chemical may be available in lots of package sizes, and you wouldn't want to re-draw the structure every time you added a new package. You can use a subform so the physical property information is entered once in the main form and the package sizes and prices are entered many times in the subform.

By linking these two data tables with a linking field you can make these tables relational; one table can interact with the other. The contents of a linking field are not important, as long as they are different for each record in the main form. Various forms of ID numbers are often used as linking fields. As you browse through the main form, corresponding records in the subform appear.

Creating a subform

You create a subform, place form objects on it, and connect the subform to a database just as you do a regular form. For detailed information on these procedures, see ["Forms" on page 31](#).

Creating a subform manually


To create a subform:

1. Start ChemBioFinder and open or create a form. The form should have at least one data box, and be linked to an existing database.
2. Click **Subform** on the Form toolbar.
3. Click somewhere in the form and drag to create a data box. When you release the mouse button the **Subform Properties** dialog box appears.

By default, newly created subforms are associated with the same database as the main form.

To use a different database as the data source for the subform:

1. Click **Open Database**. The Open dialog box appears.
2. Choose a database and click **Open**.
3. Select the table to be linked (in this case, **Synonyms**).
4. Click the **Form** tab. Select the **Generate form** check box. If you want to set the style, click **Style**. See "[Creating forms automatically](#)" on page 31 for details on Form style.
5. Click the **Subform** tab. The tab should contain reasonable data (**Link To SYNONYMS (subform) = SYN_ID, Title = Synonyms**) based on the choices made so far.
6. Choose **Mol_ID** from the **Link from MOLTABLE (main form)** drop-down menu.

 **Note:** To use the subform relationally to the main form, the two forms must have one field in common. This Linking Field must share the same data type (text, integer, real) as a field in the main form. The fields do not need to have the same name.

7. Modify box properties as desired. You may want to change the box title, font, or presence of scrollbars.
8. Click **OK**. The new subform appears showing one synonym.
9. Optional: double-click in the subform to switch it to table view. The Subform is connected to the database you chose.

In this example, when a record is retrieved in the main form, its **Molname** is used to search the **Synonym** field of the Synonyms table of CS_Demo. The subform displays the hits from this search.

Subform generator

Automatic subform generation is part of automatic form generation (see "[Creating forms automatically](#)" on page 31).

To generate a subform:

1. Right-click on a blank area of the current form and click the **Data Source...** context menu command. The Form Properties dialog box appears.
2. In the Form Properties dialog box, select the **Form** tab.
3. Select the **Generate form** check box, then click **Style...** The Form Generation dialog box appears.
4. Click **Subforms**. The Subform Generation dialog box appears.
5. For each subform you want to include:
 - a. Choose a field in the **Link from** box to specify the linking field of the parent form.
 - b. Expand the table you want to display in the subform in the **Link to** box, and select the linking field of the subform.
 - c. Click **Add** to append the selected subform to the list being generated.
6. When you have added all subform links, click **OK** to close the dialog box and return to the Form Generation dialog box.

Change existing subform layout

You can use the Form Generator to automatically change the overall layout of an existing subform.

To change the layout of an existing subform:

1. Right-click on the title bar of the subform data box and click the **Properties...** context menu command. The Box Properties dialog box appears.
2. In the Box Properties dialog box, select the **Form** tab.
3. Select the **Generate form** check box, then click **Style...**. The Form Generation dialog box appears. In the left panel, all of the Fields are selected by default.
4. Deselect those fields in the **Choose fields to be included** section that you want to exclude from the form.
5. In the **Form style** section, select the form options you want. For a detailed description of the options, see "[Creating forms automatically](#)" on page 31.
6. Click **OK**.
7. Click the **Properties** dialog box.

A warning dialog box appears allowing you to create a new form or replace the existing form. Click **No** to create a new form, **Yes** to replace the existing form. The subform is automatically changed.

Working with subforms

You must select a subform to work with it. The subform title bar is highlighted when the subform is selected. After a subform is selected, the toolbars affect the subform. You can use the Record tools to browse the subform and add records.

To select a subform:

- Click anywhere in the subform.

To return to the main form:

- Click the main form.

To select a subform box:


- With the selection tool, click in the title bar.

Searching a subform

You search a subform the same way you search the main form. You can search a subform and the main form simultaneously.

To search a subform:

1. Go to **Search>Enter Query**. The subform and the main form is cleared.

 **Tip:** You can also click the Enter Query icon on the Search Toolbar or type CTRL+F.

2. Enter a query in the subform data box. See "[Queries](#)" on page 96 for details on entering queries.
3. Go to **Search>Find**.

If the subform is linked to the main form the hit list contains the main form records related to the matches in the subform search. You can browse and save this hit list just as you would a hit list from a main form search.

If the subform is not linked to the main form, it behaves independently of the main form and is searched separately.

Viewing subform data in a table

If you have more than one record in the subform associated with a single record in the main form, it may be more convenient to view the subform as a table while you browse through the main form. The table view shows you the complete list of related records for each entry in the main form.

To display a subform in Table view:


- Select the subform and do one of the following:
 - Double-click on a record in the table.
 - Go to **View>Data Table>In Current Window**.
 - Type Ctrl+T.
 - Click the **Switch to Table** tool. A table view of the records in the subform appears.

To return to the Form view, repeat the above action.


Subform plots

A third type of view is the plot view. You can use BioViz to create a 1-D, 2-D, or histogram plot of any numerical field (s) of the data table connected to the subform.


To create a subform plot:

 **Note:** You must be working with a form that has a subform with at least one column of numeric data.

1. Click the subform box to activate it.
2. Go to **View>Subform Plot View**. The BioViz Plot Properties dialog box appears.
3. Set up the plot. See "[Creating a plot](#)" on page 160 for more information. The plot appears in the subform data box.

 **Note:** For 2-D plots, both numeric fields must be in the data table associated with the subform data box.

Subform plots have less functionality than regular BioViz plots. While mousing over a point will highlight it, selection, filtering, and coloring are disabled. Since selection is disabled, subform plots are always in “zoom on drag” mode. Use Control+W to toggle between the form, table, and plot views. The **Switch to Table** tool only toggles between form and table, as does double-clicking in the subform.

 **Tip:** The subform remains selected until you click the main form, so you can only browse in the subform. If a plot is displayed, browsing does nothing.

If you toggle the main form to table view when a subform plot is displayed, the plot is displayed as a graphic rather than an active plot.

Using scripts in subforms

You can specify a CAL script to be executed when you click an item in a subform in table view. The subform can then be used as a list selection box. The script can use CAL commands to retrieve the clicked item, and perform an action

on it. For more information about CAL, see ["CAL Commands" on page 242](#).

To specify a CAL script:

1. Right-click on the subform header and select **Properties...** from the context menu. The Subform Properties dialog box appears.
2. In the **Table script** box, type the name of a CAL script or click **Edit** and select a script.
3. Click **OK**. In Table view, the entries are blue and underlined indicating that they are hot linked to a script.
4. Click any of the hot links to run the script.

BioViz

With BioViz, you can identify trends and correlations in your data, and within subsets of your data, without the extra step of exporting to another application. You can have as many plot windows as you like, each showing a different visualization of data from the current form.

BioViz is designed to be used with the Structure Window and the Queries Tree. While not necessary to creating visualizations, they offer considerable convenience when working with plots. Before working with BioViz, go to the **View** menu and activate the **Structure** and **Explorer** windows. The Structure window gives you a way to display structures without taking up room on the form. The Queries Tree in the Explorer Window gives you an easy way to track multiple hit lists.

To create a new plot, view an existing plot, or delete a plot, use the BioViz **Plots** submenu of the **View** menu. Plots are saved with the form and will come back when it is opened, though they may be hidden.

Creating a plot

To begin working with BioViz, select a set of data to work with. You can plot an entire database or the results of database queries. You can filter the plot to limit the data range within a given dataset.

To create a new plot:

1. Go to **View>BioViz Plots>New**. The BioViz Plot Properties dialog box appears.
2. Select a **Dimension** (one or two variables) and a **Style** (line, scatter, or histogram).
3. Select the variable(s) from the drop-down list(s).
4. Optional: select other options, change the name.
5. Click **OK** to view the plot. The plot appears in a separate window.

In a 2D plot, you may plot any numeric field on the X axis and any other on the Y. For a 1D plot, you choose only the value to be plotted on the Y axis, where the X axis represents record number. For a histogram, choose a single value to be plotted on the X axis, with the frequency plotted on Y.

Another way to select a field to be a variable, or create a 1-D plot is to right-click in a property field and select BioViz Plot from the context menu.

When you right-click in a numeric data box, you may choose that value to be plotted on either axis, or as a 1D plot against record number. In the latter case, the plot appears immediately when you release the mouse button. For a 2D plot, you must first select a data box for the other axis. Use the **Reset X,Y command** to clear both X and Y settings and start again.

Histograms

Histogram plots are only available when you select a one-variable plot. When you select a Histogram plot in the BioViz plot properties dialog box, the plotting axis switches to X, and the Bins section becomes available.

Each bin in a histogram is labeled with its high-low cutoff points. Note also that the data for the bin is displayed in the status bar as you mouse over it.

When you overlay queries, the histogram bins display the partial membership of the selected queries within the bars of the histogram.

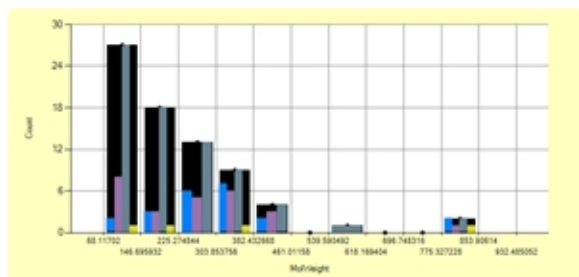


Figure 12.1: Histogram display with overlaid queries

Plot options

BioViz options are available from the BioViz Properties dialog box and from the context (right-click) menu. Some of the context menu commands repeat options in the Properties dialog box.

General tab

Dialog Option	Description
Dimension	Number of variables plotted. One and two dimensional plots are currently supported.
Style	Currently supported plot styles are Line, Scatter, and Histogram.
Bin	By number: Sets up n bins evenly spread across the whole range of values. n must be ≤ 500 . By size: Sets up bins of size n. The number of bins will be as many as are necessary to cover the range.
Log scale	Select this option to display a log scale.
Base	Base of the log scale. You must enter a number, thus for a base e log scale, you would enter 2.718281828.
Name	BioViz automatically names plots with sequential numbers (BioViz1, etc.) Edit this field if you want to give your plot a different name.
Locked	Locks the display. When the display is locked, the following have no effect: filters

Dialog Option	Description
	<p>changes in the current list</p> <p>list coloring</p> <p>BioViz Plot Properties dialog box options (except Name) are blocked. The Display and Analysis tab options are all blocked, even though they are not grayed out.</p> <p>Selection and mouse-overs are not affected.</p>
Point limit	Sets the maximum number of points in the plot. Used to limit the plotting time with very large data-sets.

Display tab

Dialog Option	Description
Point shape	Selects the point shape – circle, square, diamond, triangle, cross, or star.
Point size	Selects the point size. Values from 2 - 20 may be dialed in.
Comment	Text box for a comment or legend.
Background	Selects the background color of the plot.
Label	Text box for custom label for both X and Y axes.
Show grid	Option to display a grid. X and Y axes may be selected separately.
Font	Selects the font. Different fonts may be used for the Comment, X and Y axes.

Context menu options


Menu option	Description
Prop- erties...	Opens the BioViz Properties dialog box.
Filter Win-	Opens a filter window. See " Filtering " on page 169.

Menu option	Description
dow	
Zoom on Drag	A toggle which selects the mode of operation when you drag a rectangle on the plot. If Zoom On Drag is checked, the rectangle defines the area you want to examine, and releasing the drag will cause the plot to zoom in on that area. When unchecked, dragging a rectangle selects the points within it without changing the scale. You can use zoom to reset the scale even on locked displays.
Unzoom	Resets the scale after a zoom, or a series of zooms.
Rescale to All Points	Scales the plot so that the visible points fill the plot window.
Autoscale	Automatically rescales every time you change the list (that is, perform a search, restore a list, etc.).
Locked	Locks the display. See " Plot options " on page 161 for details.
Copy Image	Copies the plot to the Clipboard, allowing it to be pasted into other documents.
Selection to List	Creates a hit list from selected data points. The list is displayed in the Queries Tree and is treated and saved like any other search query. You cannot use Restore Query on a list selected this way, but you can use Restore List to retrieve the records.

Viewing Plots

ChemBioFinder provides three windows to view data in a BioViz plot:

- Structure — displays structures of data points as you mouse over them; used when the form is not visible or is displaying a different record.
- Details — displays data for selected fields, either from the database or from child tables, as you mouse over data points; used when the form is not visible or is displaying a different record.
- Filters — displays filter bars for those fields being used to modify the plot.

 **Tip:** To see some of the features, you must “tile” the windows to view the form and the plot simultaneously.

Displaying plot details

The Details Window displays field values (other than Structure) when viewing a BioViz plot.

To display the Details window, go to **View>Details Window**. The window floats by default, but can be docked with other windows.

To add new fields to the Details Window:

1. Right click in the Details Window.
2. Select the field to be displayed. A check mark appears next to the selected field. Fields appear in the order you select them.
3. To remove a field from the Details Window, right-click and deselect it.

Selecting data points

Each point on the Bioviz plot represents a record in your database. You can select a point in the plot to view data it represents. When selected, the point is displayed as a double hollow red circle. When you hover over a point, the center circle is solid. Points that belong to queries appear as solid colored points without circles.

Removing empty points

A point which has an X value, but not a Y value, is an empty point on the plot with a distinctive marker. If there is a Y value, but no X value, the point is simply omitted. BioViz omits empty X and Y points. All omitted points are noted. To view omitted points, click **Notes** at the bottom of the chart.

Plotting queries

When you perform queries (see "[Queries](#)" on page 96), and have one or more hit lists attached to your form, the BioViz plot displays the results of the current hit list. You can display overlays by selecting hit lists rather than making them current.

For example, the following figure shows results of a screening test relative to a control.

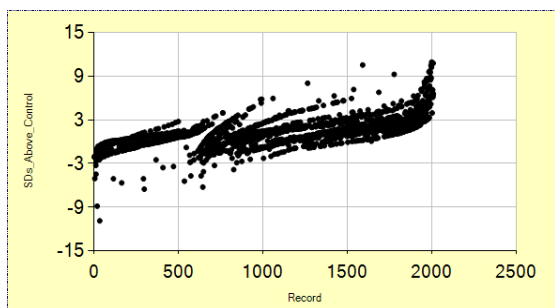



Figure 12.2: Screening test results

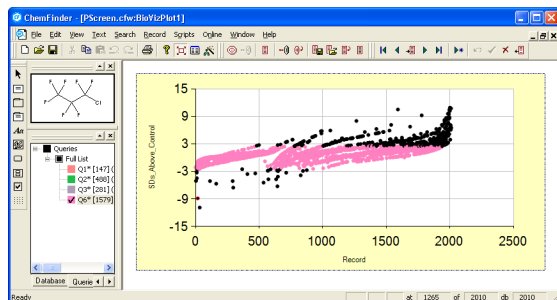
To remove outliers more than 2.5SD from the control:

- Search for “<2.5 and >-2.5” in the **SDs Above Control** field.
- The plot displays the results for the hit list.

 **Note:** By default, the hits are shown on the same scale as the original list, which may cause them to cluster in one sector of the window. If this happens, right-click the plot and choose **Rescale to All Points** or **Auto-scale**. **Rescale to All Points** operates on the current data points only; **Auto-scale** sets a switch so that all subsequent searches and list operations will automatically rescale the plot.

To display the results of an overlay on the full list:

1. Open the Queries pane of the Explorer Window.
2. Double-click on the Full List to make it the current list.
3. Click the query to select it for display, over the current list.



Selecting a list by single-clicking in the queries tree (or choosing **Color On Plot** from the tree context menu) causes points on the plot to be re-colored. The re-colored points are those which are on both the selected list and the plotted one. You may color more than one list at a time. When a list is colored on the plot, its item in the queries tree is shown with a check mark; clicking it a second time removes color and also removes the check mark.

Synchronization of plots

Multiple plots attached to a form are synchronized. If you select (mouse over) a point in one plot, the same point is highlighted in all other plots (and the related structure is displayed in the Structure Window).

This behavior is slightly modified if one of the plots is a histogram. You must select (not mouse over) a histogram bar to highlight points in other plots. Mousing over a point in a line or scatter plot will highlight the point in the center of the corresponding histogram bar.

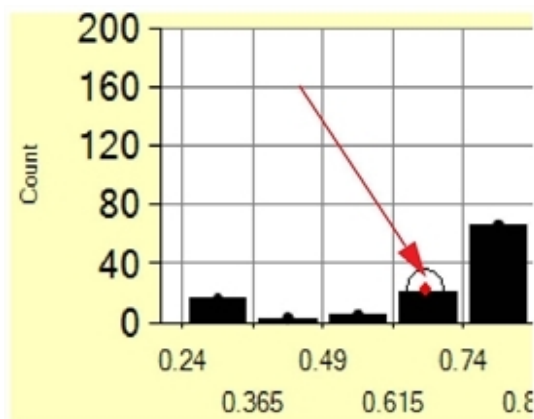


Figure 12.3: Synchronized histogram

Changing the display

Once you have selected data points and produced a plot, there are several options for modifying the display.

- Changing Colors—to change the color of a plot:
 - a. Open the Queries pane of the Explorer window.
 - b. Right-click a list and click the **Change Color...** context menu command.
 - c. Select a new color from the color menu.
- Selection and Zooming—to select a part of the plot and zoom in:

To select part of the plot:

 - a. Drag a rectangle across a portion of the plot. The portion of the plot is displayed in red.
 - b. To select more than one section, use SHIFT+drag or CTRL+drag.

To zoom-in on part of the plot:

 - a. Right-click in the plot and click the **Zoom on Drag** context menu command. This is a toggle switch—choose it again to deselect.
 - b. Drag a rectangle across a portion of the plot.

This time, when you release the mouse button, the plot display zooms in to show only the points selected. When the plot is zoomed, scrollbars appear. Use these to reposition your view of the zoomed plot. Click the “circled-point” icon on a scrollbar to restore that axis to the previous zoom. Right-click and click the **Unzoom** context menu command to restore to the full view.
- Changing the Style—you can change the style of an existing plot. For example, you can turn a scatter plot into a histogram.
 - a. Right-click in the plot and click the **Properties...** context menu command.
 - b. In the properties dialog box, select the style and click **OK**.

The plot displays in the new style.
- Changing Fields—you can change the field plotted on the X or Y axis with the context menu.
 - a. Point in the area of the numbers just below or to the left of an axis.
 - b. Right-click. The context menu displays a list of plottable fields.
 - c. Select a different field to plot on that axis.

 **Note:** You can only change fields when the chart is not locked.

In Data Table display, you can display a substructure search as table of R-group substituents.

Statistical analysis

You can perform statistical analysis on BioViz plots. When you specify two variables, a third tab appears on the BioViz Properties dialog box.

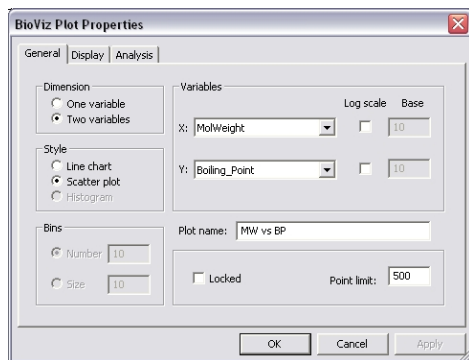
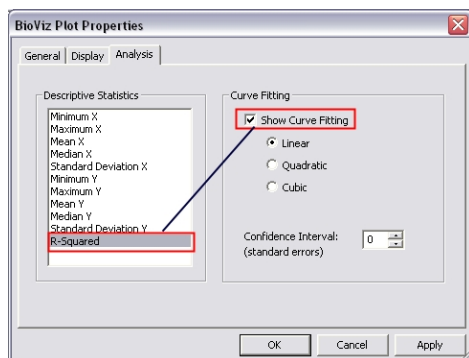


Figure 12.4: The Analysis tab appears when two variables are specified.

For descriptive statistics, minimum, maximum, mean, median, and standard deviation can be calculated for both X and Y variables. R-Squared is available only when curve fitting is chosen.




 **Note:** R-square is available when curve fitting is selected.

Figure 12.5: BioDraw Analysis

In Version 12, results of descriptive statistics are displayed on the plot.

Version 10 plot

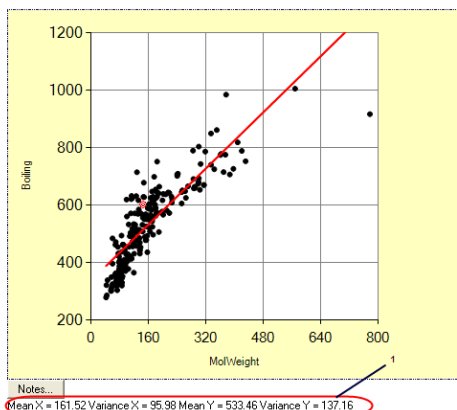


Figure 12.6: 1) Descriptive statistical analysis

Version 11 plot

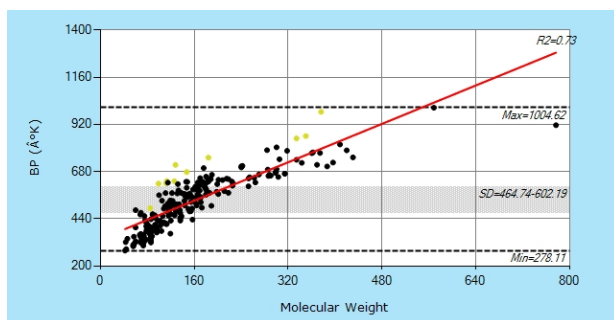



Figure 12.7: Displaying Statistical Analysis Results

You may choose from linear, quadratic, or cubic curve fitting. A confidence interval of one to three standard deviations may also be displayed. Both descriptive statistics and the curve fit update dynamically when the recordset changes or filters are applied.

 **Note:** Curve fitting gives incorrect results if one of the axes is logarithmic. For example, you do a linear fit on the original data and then transform that fit to the log plot thereby creating a curve.

Once you have created a plot, you might want to clean up the display. There are two ways to do this: sorting the data on a field, and using filters.

Sorting

The standard ChemBioFinder sort feature is the first step in cleaning up a graph. You may sort into ascending or descending order. There are two ways to sort a data set on a given field:

- Right-click in the field box and click the **Sort** context menu command.
- Display the Data table view, and double-click on the table header of the field you want to sort.

Filtering

BioViz filters limit the display of a variable by trimming either end of the range, that is, you can eliminate high points, low points, or both from the display. Filters can be set for either or both plotted variables. You may also filter the data-set on other variables, up to a total of 31 filters per chart.



Figure 12.8: Filter sliders

Filters are set in the filter window. If the window is not visible, check the box next to **Filter Window** on the **View** menu. When a filter is set:

- All plots associated with a given form are affected, unless the plot is locked before applying the filter.
- Points missing a value for a given field are hidden whenever the filter for that field is visible.

The range is displayed in black below the slider, and the set points in red above the slider. Filter name is centered, rather than on the right. Operation of the sliders is unchanged.



Figure 12.9: Filter window, stacked with details and structure windows

To activate a filter:

- Right-click in the Filter window, and select a variable from the list. When you view the context menu again, the variable will appear with a check mark next to it.

To remove a filter:

- Right-click in the Filter window to display the context menu, then click the variable to deselect it.

The following bio-assay example shows two plots of the octanol/water partition coefficient (CLogP) plotted against an activity measure called Fold_Above_Control. The second is filtered on molar refractivity. Note that the upper chart has been locked to prevent the filter from affecting it.

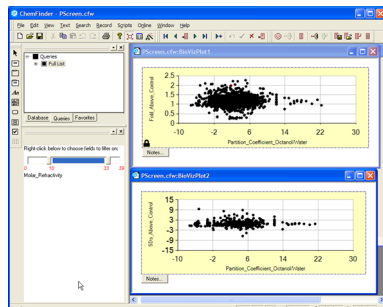


Figure 12.10: Locking a chart

To apply the filter:

- Press the thumb at one end of the slider and drag along the slider.
- In the bottom diagram, the plot is limited to Molar_Refractivity from 10 to 33.

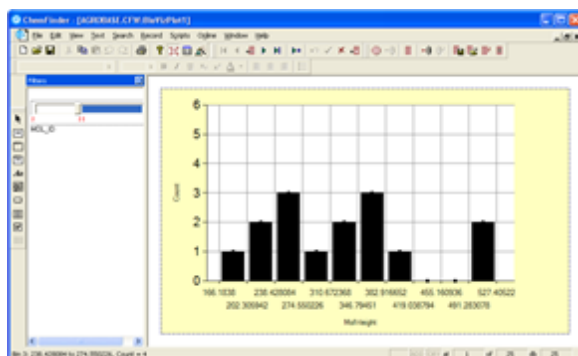


Figure 12.11: Histogram Filtering

Filter slider adjustment

Once the top and/or bottom of the range has been set, you can move the entire range. For example, the Boiling Point filter has been set to display a range of about 250° —from 501.3°K to 753°K.



Figure 12.12: Filter Slider Adjustments

By dragging the blue part of the slider, you can view any 250° range of the chart that you want. The range being displayed has been moved to 651.5°K—903.2°K. The size of the range hasn't changed, just its location.

Customizable styles

Point shapes

You can select one of the six shapes for data points: square, circle, triangle, diamond, cross, or star. Selected or mouse-over points retain the same display as in earlier versions: the selected point appears as a double hollow red circle, and points being mouse-over (pointed at) appears as red solid with a black circle.

Point shape selection is from the **Display** tab of the BioViz Plot Properties dialog box.

Axes

You can set both the content and the font used for each axis. The default X and Y labels are the names of the fields plotted, but you can now enter any text in any font to override the default. You may also choose whether or not to display the X and/or Y grids by using the check box next to each label.

Tip: Labels are “sticky”, that is, the labels used for the previous plot are the defaults for the next plot created. If you entered axis labels for Plot 1 and want to use default labels for Plot 2, you must go to the **Display** tab to delete the label entries when you create Plot 2.

Legend

You can add a legend or comments to be displayed on the plot.

1. Right-click in the plot, then click the **Properties**. The BioViz Plot Properties dialog box appears.

Tip: You must right-click either in the plot area or in the top border. Clicking any other border brings up different context menus. See ["Changing the display" on page 166](#) for details.

2. On the **Display** tab, enter your legend in the **Comment** text box. You may also enter a font and point size for the legend.
3. Click **OK** to display the legend.

The default position for the legend is in the upper right corner of the plot. After you have created it, you may position it where you want by dragging it.

To remove the legend, open the BioViz Plot Properties dialog box and delete the text in the **Comment** box.

Background color

You can choose any background color you want for a chart. A **Color...** button on the Display tab of the BioViz Plot Properties dialog box opens a color picker dialog box.

Other changes

Extended state saving. Filter slider positions and colored-by state of plots are now saved with the form.

Improved Filter display. See ["Filtering" on page 169](#) for details.

Improved statistical display. See ["Statistical analysis" on page 167](#) for details.

Improved Histogram bar coloring. See ["Histograms" on page 160](#) for details.

Clustering in BioViz

Clustering in ChemBioFinder/BioViz is a method for analyzing biological and chemical data. It is smoothly integrated with ChemBioFinder databases and forms, and presents to the user visually appealing, highly interactive display.

Features

- One-click clustering over current hit list on any numeric field or structure box.
- Multi-dimensional clustering via compound profile; allows detailed specification of fields and conditions.
- Fast new clustering and projection algorithms from the core chemistry group.
- Resizable, zoomable two-pane display showing dendrogram and cluster visualization.
- Drop-down control panels for display and behavior customization.
- Unique cluster window allows visualization by adjustable partitioning of dendrogram tree.
- Dendrogram tree may be oriented in any direction.
- Sliding adjustment varies dendrogram node spacing between linear and log scale.
- Shaded coloring may be applied to clusters based on value of a selected field.
- User choice of text labels may be applied independently in each pane.
- Tooltips provide full information about nodes and clusters.
- Intuitive highlight and selection mechanisms are synchronized between the two panes.
- Selected set of nodes or clusters can be converted with one click to ChemBioFinder hitlist.
- Pop-up structure window displays structures and data of highlighted nodes.
- Structure clustering based on Tanimoto similarity between structure screens.

Working with Clustering

This section describes the steps for carrying out clustering operations.

Clustering on single field

1. Connect your form to the CS_Demo database. On the form there are four fields: Mol_ID, Molname, MolWeight, Structure and Formula. In this section clustering is performed on MolWeight.

 **Note:** You can perform clustering on any numeric or structure field.

2. Perform search on C6. You will get 42 hits.

 **Tip:** Perform analysis on fairly small set of records otherwise the display becomes cumbersome.

3. Right click the MolWeight box and then click Cluster by MolWeight.

It takes few minutes for the operation to perform as it gathers the data, perform clustering and data projections then calculates the display co-ordinates.

You will get screen as shown below displaying the results of clustering operation.

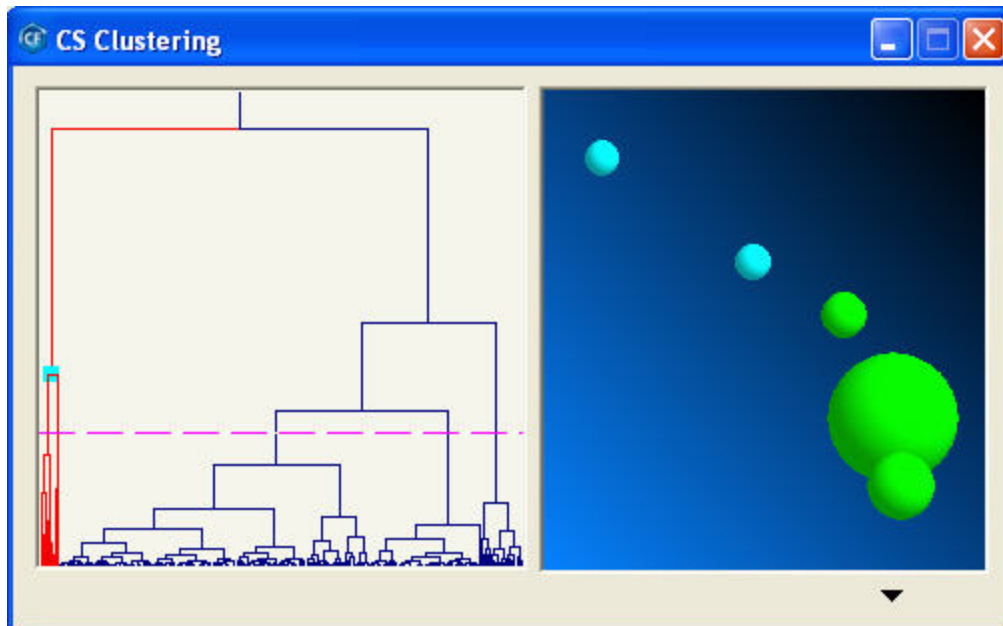


Figure 13.1: CS Clustering dialog box

In the above figure, the display is divided into two parts. On the left is the dendrogram, a standard graph for displaying the results of the clustering operation. Each point along the baseline ("leaf") represents one of the input records. In the clustering figure, you get 42 hits and 42 of them are evenly spaced along the x-axis. A horizontal line above the base represents a "node" or a merge of the two objects underneath it. The position of the node under y-axis is proportional to the distance between its children, where distance is the measure of similarity between two objects. Thus the node just above the baseline connects pairs of compounds with very similar (in this case usually) identical molecular weights.

On the right is the cluster window, which provides another way of viewing the data.

The cluster window screen depends on the position of the dotted pink cutoff line across the dendrogram. Every node directly below the line is represented as a single ball in the cluster window, with its size scaled to the number of leaves it contains. Thus a large ball represents a large group of compounds with similar molecular weights, where the smallest ball is a single compound of a weight all its own. By moving the cut off line up and down, you vary between a coarse-grained view with few clusters, to a more detailed view with many.

To modify the cluster dialog box, click the black drop down arrow in the lower right corner. See ["Dendrogram Properties"](#) on page 177 and ["Cluster Properties"](#) on page 178 for more information.

Clustering on multiple fields:

To perform clustering on multiple fields, you need to create a profile box, which contains user defined set of fields and conditions for comparing data records. See "[Creating a compound profile](#)" on page 136.

After the profile box is created on the form, save the form.

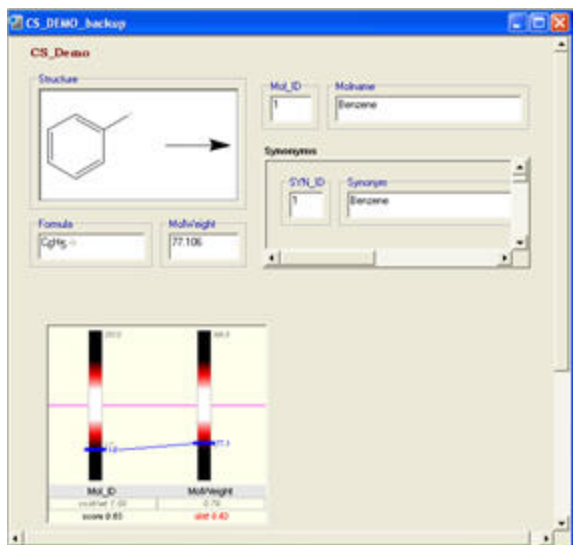


Figure 13.2: Profile box form

Each field in the profile box form is assigned a cost and weight. Right-click on the profile box, CLOGP, and choose Cluster By Profile, the clustering engine uses all the three values of each record in the hit list together with their weight and cost and displays a normalized distance in the records.

Modifying Clustering

To modify the cluster dialog box, click the black drop-down arrow that appears below the 3D structure in the CS Clustering dialog box. This drop down control panel has two tabs: **Dendogram** and **Clusters**.

 **Note:** The **Dendogram** and **Clusters** tabs may appear by default as well in the CS Clustering window.

Dendrogram Properties

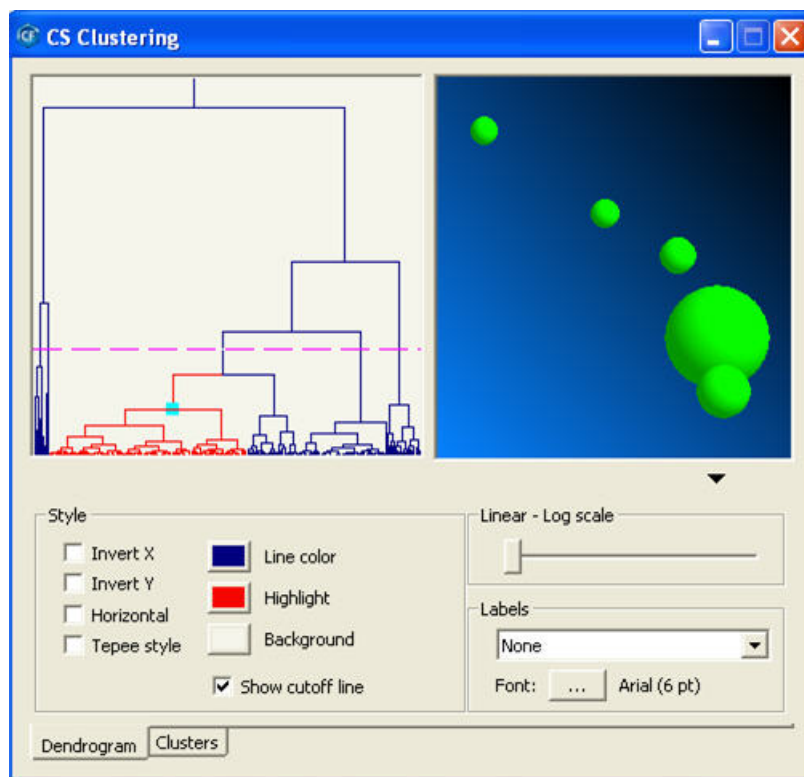


Figure 13.3: Dendrogram options

- In the Style Frame, you can change the orientation of the diagram using the **Invert X**, **Invert Y**, and the **Horizontal** check boxes. In Teepee style, lines moves from nodes to subnodes i.e. it displays the dendrogram in a non-standard way.
- Clicking the **Line color**, **Highlight**, or **Background** button opens the Colors dialog box. By default, the dendrogram lines are in navy blue color, highlighted items are in red color, and background is in light grey color.
- **Show cutoff line** is a pink line. It describes the relationship between the dendrogram an the cluster window.
- **Linear-Log scale** adjusts the attenuation of the tree by plotting the vertical/horizontal node position on the scales.
- In the Label frame, click the drop down arrow and select the label to be displayed on the dendrogram nodes or leaves.
- Choose the Font for the label display.

Cluster Properties

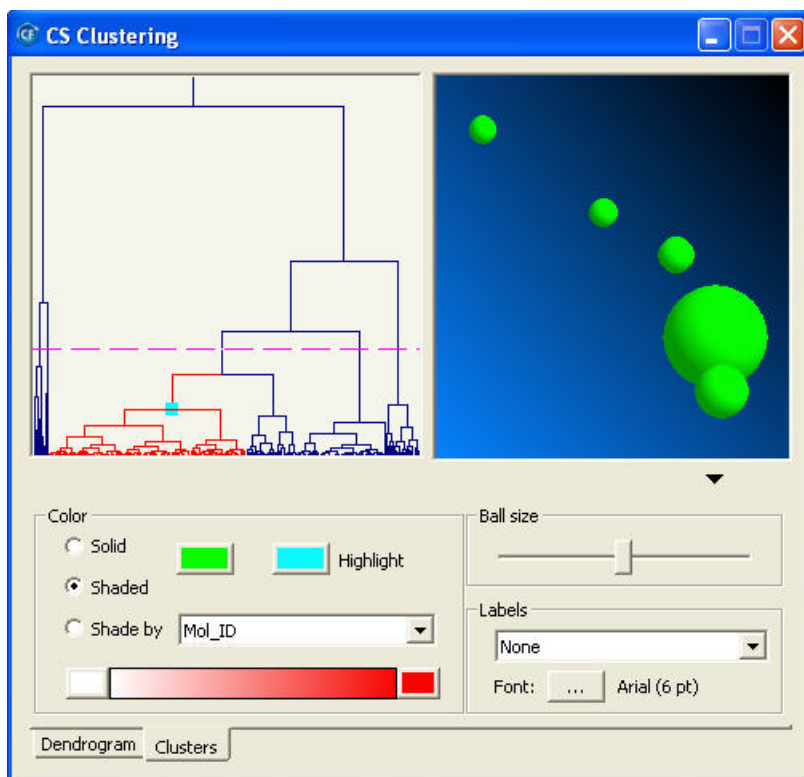


Figure 13.4: Cluster properties

- In the color frame, **Solid** and **Shaded** specifies how balls to be colored in the display. For Shade By, select the data type from the drop down menu and choose a range of colors using a button at the end of the color gradient bar.
- **Ball Size**, adjusts relative sizes of the ball.
- In Labels frame, click the drop down arrow and select the text to be displayed on the ball.
- Choose the Font for the label display.

CS Clustering Dialog features


The various actions that you can perform on the CS Clustering dialog box to make the dendrogram and cluster window look the way you want, are as follows:

1. **Hover:** As you move mouse across the cluster dialog box, the nodes or the balls it touches gets highlighted like the lines in dendrogram are highlighted to red and the ball in the Cluster window to blue.
2. **Click:** As you move the mouse across the two panes, click on any node or ball. This click will freeze the selected node or the ball (color changes on freeze) and you can easily move your mouse across the dialog box. To unfreeze the selection, click on any empty space in any of the window (get original color).
3. **Drag:** You can change the shape and size of the dialog box.

4. Re-project cluster: When you move the pink dotted line up or down in the left window (dendogram pane) the cluster window (right pane) will change the diagram accordingly.
5. Show Properties: Click the black arrow in the lower-right corner of the cluster dialog box. You will see the properties of the dendogram and the cluster.
6. Right click: On freezing any node/ball you can easily perform actions on that particular node/ball. Right click on it, it shows menu options: **Zoom**, **Show structure window**, and **Selection to list**.
 - Zoom: It changes the size of the data display beneath the viewing window.
 - Show structure window: The highlighted or freezed node/ball structure is displayed correspondingly.
 - Selection to list: It generates a hit list of the selected node/ball.

ChemBioFinder/Oracle

Users familiar with ChemBioFinder will find that ChemBioFinder/Oracle operates much the same way. You can open a database, browse, search, register, load and so forth, without worrying about what sort of database you use. In practice, however, there are several visible and invisible differences.

 **Note:** *ChemBioFinder/Oracle is available in ChemBioFinder Pro and Ultra versions only.*

When you open an Oracle database in ChemBioFinder/Oracle, you enter a new body of code which connects directly to Oracle, and carries out all searches and transactions on the server. This new mode of operation requires a different underlying technology, based on Microsoft ActiveX Data Objects, and a different philosophy in some aspects of usage, such as the handling of lists, and new features, such as index management. All of these are described in this document.

In ChemBioFinder/Oracle:

- Searches are carried out on the server, including structure searches.
- Search results are stored in tables in the Oracle database.
- Hit lists are saved to other tables in Oracle rather than in files.
- Saved lists can be annotated and, on restore, can be selected from a pick-list directory.
- When working with large databases, you can choose a new mechanism which eliminates certain slow operations associated with large record sets (e.g., Move Last). This is a user choice because there are trade-offs involved.
- You cannot create an Oracle database, but if you have the privileges you can create a table within one. This causes import operations to work somewhat differently.
- To improve performance, you can create indexes on selected columns.

Other more subtle differences are noted in the sections below.

Structure search speed is very fast in ChemBioFinder. The fast speed comes from two factors: improvement of the SQL so that hits are added much faster to the hitlist table, and an optimization of certain frequent operations and improved screening efficiency that improves structure search speed in all PerkinElmer Informatics' products.

Setup

Before you can make use of the ChemBioFinder/Oracle, your machine must be configured as an Oracle client. If you are already a user of the CS Oracle Cartridge, you are probably all set. If not, you will need to enlist the help of an Oracle administrator to set up the server and set you up as a client.


For ChemBioFinder installation instructions, see the readme.txt included on the distribution media.

Ideally, ChemBioFinder/Oracle should access any existing Oracle database in any format. In practice, there are some limitations. See "[CS Oracle Cartridge](#)" on page 303 for notes and recommendations about preparing an Oracle database for use with ChemBioFinder/Oracle.

Opening an Oracle database


To open an Oracle database:

1. Right-click in the form and click the **Data Source...** context menu command to open the Database tab of the Box Properties dialog box.
2. Click **Oracle Database**. The **CS Oracle Connection** dialog box opens.


 **Note:** The **Oracle Database** button is available to any user. There is no check to see whether the machine is a valid Oracle client.

3. Enter values for **Host name**, **User name**, and **Password** — the same values you use to log in from any Oracle client. OR:

- Choose a name from the drop-down list of recently-used items. When you choose an item from this list, it automatically fills in the user name, and the password if **Save password** was checked when the item was created.


 **Note:** The first time you bring up the Oracle Connection dialog box there is no list of recent items, so ChemBioFinder offers to create one for you. If you accept, it generates a list of all available hosts, as found in the file `tnsnames.ora` created during Oracle client configuration.

- The database is opened and the list of tables displayed.

 **Note:** The list of tables shown in the tree consists of all tables and views owned by the current user, plus those to which the user has been granted certain privileges. To see the details, you can turn on SQL tracing while a database is being opened.

4. Proceed as you would for any ChemBioFinder database: click to select the table you want to display, set other desired options, and check the box on the **Form** tab if you want to generate a form automatically.

The **Oracle** tab of the dialog box, described in "[Setting Oracle preferences](#)" on page 184. The features on this tab are mainly for advanced users and do not require adjustment.

 **Note:** The Oracle tab appears in the Properties dialog box only when an Oracle database is open.

5. Click **OK**. The database opens, in a display form if you requested one.

The *Database Wizard* can also be used to open an Oracle database. A button on the Wizard labeled **CS Oracle Cart-ridge** brings up the appropriate Oracle parts of the process.

When you have opened an Oracle database, note that some **Properties** dialog box tabs show entries not available in ChemBioFinder, including:

- **Database** tab: CS Oracle Cartridge version if any.
- **Table** tab: Table owner name; name of primary index if any.
- **Field** tab: Oracle native data type name; name of associated index if any.

When you open an Oracle database and select a table, ChemBioFinder/Oracle gathers information about the columns. If you open a table you know to contain structures, and ChemBioFinder/Oracle does not show a structure


(or formula or molweight) column, then it may be necessary to set up some configuration information about the table. For details, see ["Configuring via CF_SETTINGS table" on page 304](#).

Searching


Searching in ChemBioFinder/Oracle works basically the same way as in ChemBioFinder: you enter a query, search, then work with the hit list. Internally, however, the ChemBioFinder/Oracle machinery is quite different. There are three types of table involved in handling hit lists. All are created in your own tablespace. One is global, applying to all lists saved from any table; others are connected to the particular table or view being searched.

- **CF_HITLISTS**. The directory of all saved lists. This table is created the first time any list is saved.
- **SAVED_tablename**: All lists which have been explicitly saved from a given table or view. This is created the first time a list is saved.
- **HITS_tablename**: All lists automatically saved after every search over a given table or view. This table is created the first time a search or list operation is carried out and is deleted at the end of the session.

Here's what happens when you present a query to ChemBioFinder/Oracle:

1. If the query contains a structure, it is converted to a text representation and copied to a temporary table.
 -  **Note:** The table is called **temp_queries**. It is created in the **CSCartridge** tablespace, and removed as soon as the search is finished or interrupted. ChemBioFinder does not yet handle the case of multiple structure boxes where more than one contains a query.
2. The query is converted to a SQL select statement. Query components in form boxes are ANDed together (just as in ChemBioFinder), where the structural parts are calls into the CS Oracle Cartridge structure search functions.
3. The hits table is created, if it does not already exist.
4. A unique ID is assigned to the new list which will result from the search.
5. The select statement is wrapped in a larger SQL statement which will cause the results to be deposited directly into the hits table.
6. The SQL is executed.
7. When the search is complete, the results are new rows in the hits table. Each row contains the new list ID alongside the ID of a record from the searched table.
8. The final list is prepared by a join, selecting rows from the main table which have record ID's matching those of the new list in the hits table.

The resulting list is ready to browse, save, export, etc.


-  **Note:** Text searches in Oracle are case sensitive. You will get different hits from the query "benz*" than from "Benz*."

Sorting


Certain sort operations require a ChemBioFinder system table, **CF\$_SORT**, that may not exist in your current form. If the table does not exist, you will be alerted the first time you attempt such a search in an Oracle database.

To create the **CF\$_SORT** table:

1. Open a new form, then go to **File>Database**. The Form Properties dialog box opens.
2. On the **Database** tab, click **Oracle Database...**
3. Log in as the owner of the schema where your ChemBioFinder system tables are located. Typically this will be ChemBioFinder

 **Note:** If you do not have privileges to log in as this user, consult a system administrator.

4. Click the **Oracle** tab of the dialog box, then click **Prepare Schema**. ChemBioFinder will create the one new table needed, without modifying any of the others.

 **Note:** If this button is dimmed, it means you are not logged in as the owner of the current schema.


5. Go back to your original form and repeat the sort.

 **Tip:** You may have to click **Refresh Session Data** on the Oracle tab before the sort will work.

Subform value aggregate

In an Oracle database, ChemBioFinder allows sorting the main list over subform field aggregate values. If, for example, you have a subform showing a column of test values where there are multiple records in the subform for each main record, you can sort the main hit list based on the sum or average of the set of test values for each record. To sort by aggregate:

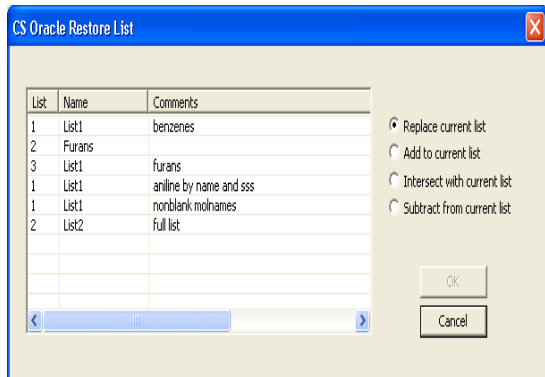
1. Go to **Record>Sort**. The **Sort Multiple** dialog box appears.
2. Click a subform value in the left-hand box to sort by.
3. In the right-hand box, right-click the **Fxn** column and choose a function from the popup menu: MIN, MAX, AVG, or SUM.
4. Click **OK**. The hit list is sorted by the selected value(s).

 **Note:** AVG and SUM are available only for numeric fields; MIN and MAX can also be used for text fields.

Handling lists

There are differences between ChemBioFinder and ChemBioFinder/Oracle in working with hit lists:

- **Save List** does not bring up a file dialog box in ChemBioFinder/Oracle; instead, it presents a dialog box in which you enter a name and a line of comments for each list.
- **Restore List** presents a dialog box in which you choose from the available saved lists (see screen shot below).



- **Restore Previous List** is no longer dimmed after you have done a single search; it is available, and will return you to the full list. (This is true also in ChemBioFinder.)
- **Over Current List, Omit From List, and Find List** work as before.

Setting Oracle preferences

When you are working with an Oracle database, the **Properties** dialog box contains a tab for Oracle settings. Items currently in this tab are as follows:

- Auto-sort on. Displays the default sort field. All lists will be sorted in ascending order over this field if no other sort criterion has been specified. At present, the auto-sort field is always the primary key, and cannot be changed.
- Primary key. Names the primary key of the current table, if any.
- Saved hits. Displays the name of the table in which saved hit lists are stored for the current form, if one has been created.
- Records per retrieve. Sets the size of the ADO recordset cache, that is, the number of records brought to the client for each retrieve operation. This value can be adjusted for better performance, but in practice it does not have much effect.
- Cache ID's for faster moves. Checking this box enables a new record-retrieval scheme designed for more efficient browsing of large lists; details are given in ["Fast-move caching scheme" on page 303](#). The state of this check box is saved with the formfile. Currently, when you check or uncheck this box, *you must save and reopen the form* before proceeding.
- SQL Trace to file. Check this box to generate a text file of the SQL being sent from ChemBioFinder to Oracle. To specify an output file, enter a pathname or use **browse**. The file is never overwritten: new data are always appended. Tracing begins when this box is checked, and ends when it is unchecked. If you exit ChemBioFinder with tracing in effect, it will be in effect the next time you start up, and you will be given a quick warning about it on the status bar.

 **Caution:** The trace file can become very large if you turn on this feature and forget about it.

Updating and adding data

If you have privileges to add, update, and delete data in the table connected to your form, then you should be able to do these operations on records in the database just as in ChemBioFinder. However, there are some cautions:

- Do not modify data in an existing *Web Server*, *RegDB*, or *E-Notebook* database. ChemBioFinder/Oracle does not automatically prevent you from doing this, but you are sure to foul up those systems unless you go through their normal registration procedures, or know exactly what you're doing.

Most database alterations are possible, as in ChemBioFinder: you can add, modify, and delete tables and fields, assuming you have the appropriate privileges. Differences include:


- When you create a structure field, you can give it any name you like.
- Structure columns in tables managed by the CS Oracle Cartridge can be in any of several formats, but ChemBioFinder/Oracle does not allow you to choose one: it defaults to a character (CLOB) column storing text-encoded CDX.
- ChemBioFinder/Oracle does not allow you to create an Oracle database in the same way ChemBioFinder creates an Access MDB file. The equivalent in ChemBioFinder/Oracle is to create a table.

Loading

You can build ChemBioFinder/Oracle databases by loading from SDFiles or using Import Structures. (RDFiles should work also, but have not been tested.) However, because you cannot actually create a new database, the procedure is somewhat different from ChemBioFinder. You must already have a database open, and then you can import to an existing table or have a new one created.

To import data and structures into a new table:

1. On an empty form, right-click and choose **Data Source**.
2. Click **Oracle Database** and proceed as described under "[Opening an Oracle database](#)" on page 181. It doesn't matter what table you select, since you will be creating a new one.
3. Click **OK** to return to the blank form. You are now ready to import.

 **Note:** When you click **OK** to dismiss the properties dialog box, you will get a warning if the selected table does not have a primary key defined. You can ignore this if you are about to create or load a new table. Otherwise, you should consider using an Oracle tool to define a primary key for the table. Choose **Import SDFile** or **Import Structures** and select the source file(s).

The **Data Import** dialog box appears:

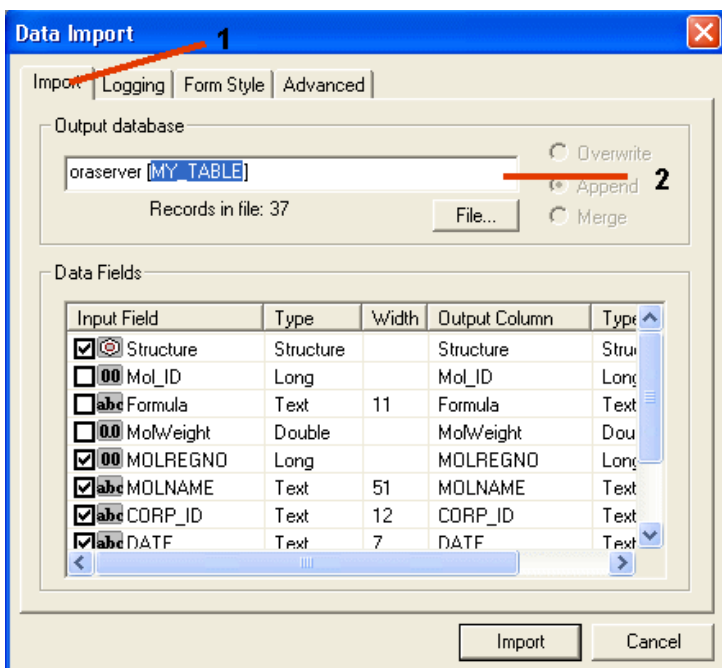



Figure 14.1: 1) Import tab; 2) Output database text box

- In the **Output database** box, you see the name of the database followed by the table name in brackets. To specify the table to load, **edit the table name** by typing between the brackets. If you want to create a new table, enter the name you wish to give it. If you want to append to an existing table, enter its name.
- Set other options as desired, and click **Import** to begin the process.

Tip: for text fields being imported, make them wider than the value determined by the input scan. (To do this,  double-click the name under **Input Field** and enter a larger **Width** value.) Otherwise you might have problems if you append more records later.

- After importing, you should create a primary key and index; see "[Indexing](#)" on page 186.

To **append** to an existing table follow the procedure above except:

- Step 1: You can start with an existing form instead of a blank one
- Step 2: You should **select** the target table. When you do, there will be no need to edit the table name in Step 5.

Indexing

As every Oracle administrator knows, a key to good performance is to index certain columns so that searches over them become fast lookups. If you intend to create, load, or manage tables, then you are an administrator, and need to know something about indexing. ChemBioFinder/Oracle provides a few tools to assist you.

There are two types of index of interest to ChemBioFinder/Oracle:

- Structure indexes. Any column containing structures should have an index created by the CS Oracle Cartridge. If there is no index, searching is still possible, but very slow.
- Primary keys. A table containing a column of structures should also have a column of unique record identifiers for use in various list-handling operations. Preferably this column is of type INTEGER and is designated as the *primary key* of the table.

ChemBioFinder provides information about these indexes, and allows you to create or recreate them if necessary. To create a structure index:

1. Right-click on the structure box of an Oracle-connected form, and choose **Properties...**
2. Go to the **Field** tab of the dialog box. If the selected field is indexed, the index name is shown in the properties box. In the example below, the index is named IX50.

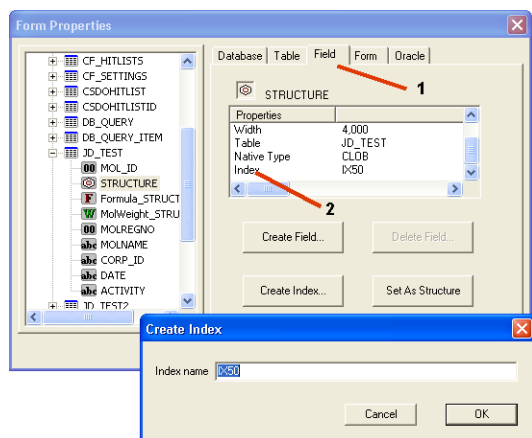



Figure 14.2: 1) Field name; 2) Index name

 **Note:** If there is no index, or the field is not indexable, the Index line does not appear in the list.

3. If there is no index, click **Create Index**.

 **Note:** If this button does not appear, it means the selected field is not a candidate for indexing.

4. In the **Create Index** dialog box, provide a name for the new index and click **OK**. The index is created.


You will get an error message from Oracle if:

- You provide an index name which is already in use, or
- The column already has an index.

ChemBioFinder/Oracle does not provide a direct way to delete or overwrite an index, but the CAL SQL command can be used for this purpose.

To create a primary key index:

- The process is the same as for a structure index, except that you carry it out on a field of type **Long**. This creates an index on the field and also designates it as the primary key. It will fail unless the column already contains unique, non-null values.

 **Note:** A new button, **Set As Structure**, may appear if you have selected a certain type of column. This is an advanced feature described in "[Configuring via CF_SETTINGS table](#)" on page 304.

-
- ✓ **Tip:** If you want to create a primary key, but do not have a column of unique integer values, you can create one using a CAL script. For example (assuming you've already created a column and form box called "ID"):
-

```
loop
putdata ID $index

record commit

record next

endloop
```

CAL

Some new features have been introduced into ChemBioFinder Automation Language (CAL) to support Oracle. Some are not specific to Oracle, but of general utility:

- **SQL command.** When you are connected to an Oracle database, you can pass any SQL command which does not return records. The format is simply SQL <command> where the command is not quoted.

Example:

```
SQL create table mytable (id integer).
```
- **CURR_TABLE variable.** This variable contains the name of the table connected to the current form. This works whether or not you are in Oracle.

Example:

```
MSG Current table is $CURR_TABLE
```
- **OPENDB** allows change of table. Normally this command takes the name of a database, followed optionally by the name of a table in angle brackets. In ChemBioFinder, it can now take just the table name, and causes the form to become attached to a different table in the current database — as if you had brought up **Properties** and clicked a different table name.

Example:

```
OPENDB <CHEM_STRUCTS>
```

causes the current form to be connected to the table CHEM_STRUCTS.
- **STATMSG** command displays a message on the status bar. Format: STATMSG <message>.

Example:

```
STATMSG Search is now in progress...
```
- **SEARCH SAVE LIST, RESTORE LIST** work with Oracle. Normally you can follow either of these commands with the name of a file; in ChemBioFinder/Oracle you can instead provide a list name.

BioSAR

This section explains how to export a data form from BioSAR and view it as an equivalent form in ChemBioFinder. Once in ChemBioFinder, you can manage the form using a variety of features.

Starting the BioSAR browser

To use BioSAR, you must have Internet Explorer 9.0 (with SP2) or greater installed.

To start BioSAR:

1. In Internet Explorer, log into the ChemOffice Enterprise portal.
2. In the Enterprise screen, under BioSAR Enterprise, open a query form using either of these links:
 - Open Form
 - Manage your forms

Performing searches in BioSAR

Before exporting a BioSAR data form to ChemBioFinder, you need to search on the data in the BioSAR database.


1. In BioSAR, open a query form of your choosing.
2. In the form, enter your search criteria and click **Search**.

Exporting the data form to ChemBioFinder

After the search is complete, you are ready to export the data form to ChemBioFinder.

To export the form:

1. Click **Display In BioViz**. If the button is not visible, you may need to edit the BioSAR form properties. See your administrator.
2. When prompted to either save the form file or open it, select **Open**. ChemBioFinder will launch.
3. At the Oracle login, provide your user name and password.

 **Note:** If you don't want to view the form now, simply save the file. You can open the file in ChemBioFinder at any time.

ChemBioFinder file types

Whenever you export a data form, BioSAR creates a file with the extension .bsbxml. ChemBioFinder reads this file, and opens the form, linking it directly into the database.


The .bsbxml file is read-only. although ChemBioFinder doesn't let you modify the .bsbxml file itself, you can still edit the form and save your changes. When saved, the form file is assigned the ChemBioFinder extension cfx. Unlike .bsbxml files, cfx files are not read-only. You can open any cfx file in ChemBioFinder, edit it, and save it again. Meanwhile, the .bsbxml file is still available whenever you need to open the original form.

Setting “save password”

If an Oracle-connected form contains an invalid (or blank or no) password, the program prompts you for a name and password on opening the form. The same is true if the **Save password** box has not been checked in the form properties dialog box. This is now the default when a form comes from BioSAR. You are prompted for username and password every time, until you check the **Save password** box and save the form.

 **Note:** *In Ultra and Pro versions, Save Password defaults to ‘off’.*


Subforms and autolinks

 **Note:** *Autolink boxes are available in Pro and Ultra versions, but not in Standard version.*

When you create a form in BioSAR, you select from the database which fields you want to include in the form. Those same fields appear in the form when you import it into ChemBioFinder. After you import your form, you can continue to build on it in ChemBioFinder using additional fields. The fields you use can also include child tables and their fields. Child tables that you add appear as subforms; fields from child tables are displayed as autolink boxes.

To add a subform:

1. In the Explorer Window, select the BioSAR tab.
2. Select a child table from the Child Tables directory.
3. Double-click the child table. The child table appears in the form as a subform.

 **Note:** *Not all fields in the child table are necessarily displayed in the subform. The subform fields are limited to the same fields you selected in BioSAR and defined in the default form group.*

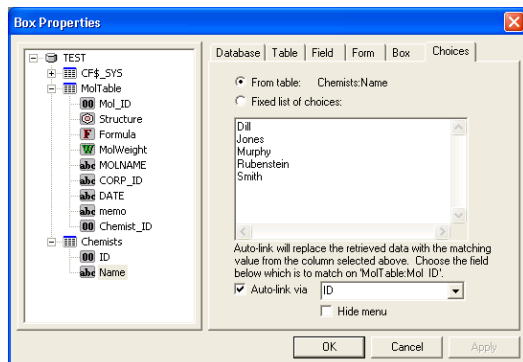
To add an autolink box (automatically):

1. Select a child table from the Child Tables directory
2. In the child table, double-click an individual field. The child table field appears in the form as an autolink box.

To create an autolink box (manually):


Assume you have a structure table **Moltable** with a column of non-unique integers **Chemist_ID**. In a separate table **Chemists** there is a text column to be displayed **Name** and a column of unique integer record identifiers ID. Each Chemist_ID in MolTable matches one ID in Chemists. To set up an auto-link box on the main form which shows, for each structure, the name of the associated chemist.

1. Create a data box.
2. Right click on the data box and connect it to the ID field to be retrieved. In this example, connect the form box to the Chemist_ID field.
3. To connect the box to the ID field, right-click on the box and choose Properties. The Box Properties dialog box appears.
4. Check the option **With choices or auto-link**. Checking this will automatically bring up the **Choices** tab.



In the Choices tab:


- If the "Auto-link via" box is unchecked, the dialog box lets you define a set of choices for data input.
- If the box is checked, then you have to define both a set of input choices and an auto-link connection.

 **Tip:** To prevent the drop-down choice menu from appearing, check the "Hide menu" box. This applies only if you have set up auto-link.

5. Select the **"From table"** radio button. In the data source tree, locate the subtable and click the column of data to be displayed in the box (in this case, "Name" under "Chemists").
6. The label, Name, after the radio button, From table, updates to show your choice, and the **"Auto-link via"** check box becomes available.

 **Note:** An auto-link cannot be set up if you are entering a fixed list of choices.

7. Check the **"Auto-link via"** box. From the drop-down menu, choose the column of the subtable (in this case "Name" under "Chemist") which is to be matched against the column ID.

 **Note:** The columns shown in the drop-down menu should match the type of the fields to be linked, i.e., if the ID is numeric, then only numeric fields may be linked.

8. The form box displays a chemist's name.

Multiple sort

ChemBioFinder's multiple sort feature lets you rearrange the hit list based on one or more form fields you specify. For example, you can first sort the list by Reg number and then further arrange the list in order of Full reg number.

Sorting by main form fields

To sort the hit list using fields in the main form:

1. Go to **Record>Sort...**. The Sort Multiple window appears.
2. In the Sort Multiple window, select from the left pane the field by which you want to sort. To sort by multiple fields, select the fields in the order you want to sort.

3. By default, each field is sorted in ascending order. To reverse the order, right-click the field name in the right pane and click **Descending** in the pop-up menu. To remove a field, right-click the field and select **Remove**.
4. Click **OK**.

Sorting by subform fields

Just as in the main form, you select in the subform the fields you want and the order you want them to be sorted. You also indicate which value in each field you want ChemBioFinder to sort by—the minimum value, the maximum, average of the values, or the sum.

To sort the main form using subform fields:

1. Go to **Record>Sort...**
2. In the Sort Multiple window, select from the left pane the subform field by which you want to sort the hit list. The subform field is displayed in the format, *subform name: field name*.
To sort the hit list by more than one field, select the fields in order of priority.
3. To reverse the order in which values will be sorted, right-click the field name in the right pane and select **Descending** in the pop-up menu.
4. Also in the pop-up menu, select the value in the field you want ChemBioFinder to sort by—**max**, **min**, **average**, or **sum**. The default is Min.
5. Click **OK**.

Query hit lists

The Queries tab in the Explorer window maintains a list of search queries from the current and previous sessions. Queries are associated with forms. So, the query hit list is saved when you save the form.

Viewing the query list

When you open an existing form, the tree updates to show only the queries for that form.

The Explorer Window must be visible to view the Queries control. If it isn't visible, do the following:

1. Go to **View>Explorer Window**.
2. Click the **Queries** tab to display the Queries tree.

Query list operations

Each time you perform a search or list operation, a new query is generated in the tree. ChemBioFinder assigns a name to the query and displays it in the tree along with the size of the list (number of hits) and a brief description. The generated name is Q<n>*, where n is a sequential number and * indicates that the query has not been renamed or marked for saving. ("[Saving query lists](#)" on page 195)

To perform list operations, begin by performing a simple search. For example, you can search the database for all compounds that have a molecular weight less than 250g/mol:

1. Click the Enter Query icon on the Search toolbar or menu. The form goes into query mode.

- Enter <250 in the **MolWeight** field of the query form and click **Search**. A query labeled Q1 appears as a child of the Full List.



Changing the query name

If you want, you can change the name of a query to one that is more descriptive. Renaming a query also ensures that the query is saved with the form. (Depending on your Preferences settings, queries not renamed may be automatically discarded when you save the form).

To rename a query:

- Point to the query and right-click. The context menu is displayed.
- Click **Rename** on the context menu. The name (Q1) is highlighted.
- Enter a new name and press ENTER.

Changing the display color

The color of the box next to the query name is the same color to be used to display the points should you plot the results. You may change any query to any color so that the plots look the way you want.

To change the display color:

- Point to the query and right-click. The context menu is displayed.
- Click **Change Color** on the context menu. The color picker context menu appears.
- Select a new color and click **OK**. The color in the box changes.

If you have plotted data with BioViz, you will see that the points that correspond to the query also change color.

Child lists

You can add a new query either by searching the Full List or searching another query. When you create a new query from another one, the result is a **child list**. For comparison purposes, it is often an advantage to create multiple child lists, each from a single search attribute, rather than one child list based on multiple attributes.

You can create a child list for only the currently selected list. If you are not sure that the list you want is selected, double-click it.

Before you create a child list, select **Over Current List** on the Search menu. Note that, once selected, **Over Current List** has a check mark next to it.

You are now ready to create a child list query. The following sample procedure describes how to create a child list by searching for carboxyl structures:

1. Click the **Query** icon to put the form in query mode.
2. Double-click in the Structure window. The ChemDraw toolbar appears.
3. Using the bond tool, create a carboxyl structure.
4. When you have finished creating the structure, click outside the Structure window.
5. Click **Search**. The new list is created as a child of the previous list.

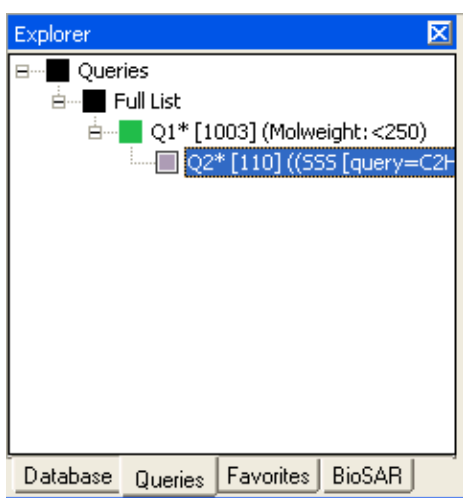


Figure 15.1: In this example, Q1 is a child list of a parent list named MW Filter.

Creating multiple child lists

You can run similar queries, each with different attributes and displayed as a child of the parent query.

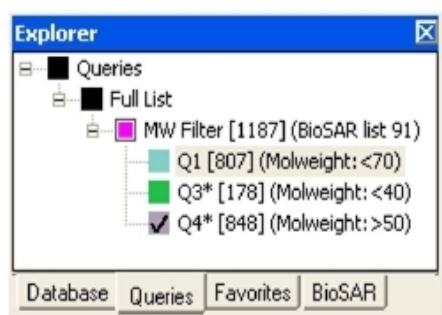


Figure 15.2: Several child lists of the parent list MW Filter.

To create another child list, you must first select the parent list. To select the parent, do either of the following:

- Double-click the parent list
- With a child list selected, click **Restore Previous List** on the Search menu or toolbar.

You can verify that you have selected the parent list by viewing the status bar information. With the parent list selected, you can add another child list to your query.

ADD QRY at 1 of 1187 db 38547

Figure 15.3: The value “1187” in the status bar indicates that the parent list in figure 5 has been selected. “38547” is the number of entries in the full list.

Merging lists with drag-and-drop

In some situations you might want to merge two lists using drag-and-drop. The result is a new list that has attributes based on the merged lists.

1. Select a list and drag it onto another list. The **Restore/Merge List** dialog box opens.
2. Select the type of merge you want, in the Restore/Merge List dialog box.
3. Click **OK**. A new list is displayed with the merged criteria you selected.

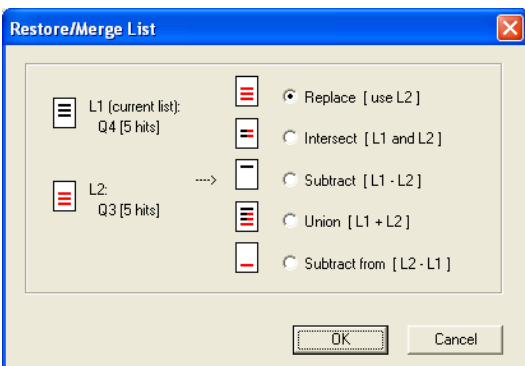


Figure 15.4: Restore/Merge List dialog box

Saving query lists

When you transfer data to ChemBioFinder from BioSAR Browser, ChemBioFinder opens a file of type .bsbxml and creates a form to simulate the BioSAR setup. At the same time, it restores whatever hit list was in effect in BioSAR. In ChemBioFinder, the BioSAR hit list is automatically saved when you save, close and reopen a file. This is different from the default in previous versions.


Domains

The Domain feature in ChemBioFinder lets you create a subset of a database. Once created, a domain acts as if it is a full database. Any database searches you perform are limited to only the domain. Any hit lists you create will be added

to the query tree as if you are working with the full database.

Creating a domain

Before you create a domain, you must run a query to define the subset of the database with which you want to work. This subset becomes the domain.

 **Tip:** If you want your domain to have a particular order, sort before creating the domain. The order will be preserved for all subsequent use.

To create a domain:

1. Select a query by double clicking on it.
2. Go to **Search>Domain>Set Domain to Current List**.

Query control display

When you run queries in a domain, they will not be displayed as child queries of the domain unless you select **Over Current List** on the Search menu and specify the domain as the active list.

Cancelling the domain

If you no longer want to use a domain, go to **Search>Domain>Reset to Full Database** again to cancel it.

Creating an mdb database

When working with a large, Oracle-based database on a server, it is sometimes convenient to store a subset of the database (dataset) as a local database. ChemBioFinder local databases are Microsoft Access-based (file extension: MDB), so the exported dataset must take into account Access naming rules and limitations.

Access database rules

Because ChemBioFinder uses the Access (Jet) database standard, names of fields, controls, and objects must conform to Access rules. These are:

- May be up to 64 characters long.
- May include any combination of letters, numbers, spaces, and special characters except a period (.), an exclamation point (!), an *accent grave* (`), or brackets ([]).
- Must not begin with leading spaces.
- Must not include control characters (ASCII values 0 through 31).
- Must not include a double quotation mark (") in table, view, or stored procedure names.

Creating a local copy of a database

If you frequently use a set of data that never changes, consider exporting the data as a local ChemBioFinder database. Having a local database often makes many common tasks run faster; tasks such as browsing, searching, plotting, and analyzing.

Creating a local database requires two steps:

1. Exporting the dataset as a CFXML file and
2. Importing the file back into ChemBioFinder. Once you import the file, save it as a CFX file. It will then be available for you to use whenever you need it.


Exporting the dataset

To export a dataset:

1. Run a query to create the dataset.
2. Go to **File>Export>CFXML** . The Data Export window opens.
3. Select the options you want to apply to the exported dataset:
 - In the **Data Export** dialog box, select a path and file name for the new database. Use **File** to browse for a location.
 - In the lower half of the dialog box, all data fields are selected for inclusion in the exported database by default. Clear the appropriate check boxes to deselect fields you want to omit.
 - Check the field names to ensure each one complies with Access database rules (If necessary, edit the field name by double-clicking it).
 - On the **Text Options** tab, select a delimiter and other options you want to include: subform data, empty records, and headers. You also have an option to save structures in CDX or CDXML format.
4. After you have made your selections, click **Export**.

Importing the cfxml file

Once you have saved the dataset in CFXML format, you must bring it back into ChemBioFinder to save it as an MDB file.

 **Note:** To avoid conflicts with field name/type definitions, close the master database before importing and open a blank form.

1. In a blank form, go to **File>Import> CFXML** and open the file. The Data Import dialog box appears.
2. Select/deselect the fields you want to import and the options you want to apply to the file.
3. After you make your selections, click **Import**. Your dataset is imported into an Access database.

Exporting to an Excel spreadsheet

When you export a BioSAR-derived form with multiple subforms into an Excel spreadsheet, you must take precautions to get all of the subforms correctly exported. This is because the exporter normally uses only the field names, and the BioSAR subforms show the same set of fields from different assay tables. On export, they will collide, and the data will not come over correctly.

To prevent this, ChemBioFinder has an option on the Text Options tab of the Data Export dialog box — **Prefix output fieldnames**. When this box is checked, the output names will match the input names, which are prefixed for subform fields.

Tips for optimizing performance

When working with a large database or multiple forms, you will want to ensure that you can still work efficiently. To do so, there are a few techniques you can use to improve ChemBioFinder's performance, even when you are working with a lot of data.

Use domain-linked forms

Using a domain offers a significant speed advantage compared to working with the full database. If you choose Retrieve All after performing some queries, you will get only the domain as the full database. Since the domain is typically much smaller than the database it was created from, the query and restore commands are executed much faster.

Another advantage is that the domain is "non-volatile", that is, the view remains on the server until you delete it. If you routinely work with a particular subset of a large database, creating a domain eliminates the need to perform the initial routine query or create a local database. You only need to save the domain information by saving the form. The next time you open that form, you see only the domain.

A third advantage is that the sort order at the time of creation is preserved. Sorting is a time consuming operation. By using a domain, you can sort once and save the preferred order.

To more information, including how to create a domain, see ["Saving query lists" on page 195](#).

Minimize child tables on a single tab

Whenever a subform appears on screen, ChemBioFinder automatically updates the data from the server. If you have a lot of subforms on a single tab, ChemBioFinder must update all of them at once, potentially slowing down your work. So, if you have subforms you need to see only occasionally, put them on other tabs of the form, so they will not be updated until you activate those tabs.

Remove plots that you are not using

Plots that are visible all take at least some processing and closing a plot window still doesn't remove it from memory. If you have generated plots you no longer need, use the **View>BioViz Plots>Remove** menu command to delete them, then save the form.

Customizing ChemBioFinder

You can customize ChemBioFinder in the following ways:

- Customize display of your molecules, fonts, pictures and forms
- Customize the Favorites tree
- Design the toolbars to your specifications
- Perform automated tasks, such as interfacing with Microsoft Excel, or by using CAL, the ChemBioFinder Automated Scripting Language

Setting preferences

The Preferences dialog box allows you to customize the display of molecules, pictures, and forms, and set options for searching and exporting.

General instructions for using the Preferences dialog box:

1. Click the tab containing the preferences to set.
2. Select the preferences, and click **OK**.

Display preferences

To set the Display preferences:

Click the **File>Preferences...** menu command.

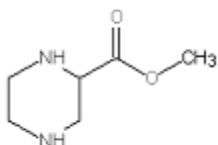
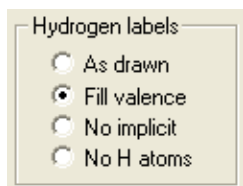
The Preferences dialog box appears with the Display tab on top.

Structure display

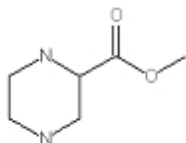
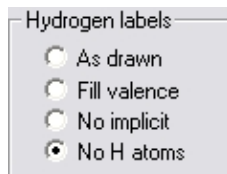
To display carbon atoms on methyl groups or on interior aliphatic or aromatic chains, check the relevant check boxes in the Carbon labels section.


To display hydrogen atoms on heteroatoms or on terminal carbons:

Select the **Fill valence** option in the Hydrogen Labels section.



Select No H atoms in the Hydrogen Labels section.



 **Note:** Selecting **No implicit** means these types are displayed without implicit hydrogens.

Using keyboard shortcuts

When using a form, you can use keyboard shortcuts to show or hide atom-to-atom maps, reaction centers, atom numbers, and bond numbers.

To use keyboard shortcuts:

Select the **Enable keyboard shortcuts** check box on the **Display** tab of the Preferences dialog box.

When keyboard shortcuts are enabled, the following keys toggle these properties:

- A: show/hide atom numbers
- B: show/hide bond numbers
- M: show/hide atom-to-atom maps
- R: show/hide reaction centers
- S: show/hide stereochemistry

Scaling structures

To scale each structure so that it is as large as possible within its structure box:

- Click the **Fit to box** radio button in the **Structure Scaling** section of the Display tab.

To display all structures with a constant bond length:

1. Click the **Uniform bond length** radio button in the **Structure Scaling** section of the Display tab.
2. Select the bond length percentage.

With **Uniform bond length** selected, structures may be reduced in size if they are too large to fit within the structure box, but they will never be enlarged.

Framing pictures

To select whether the pictures in a form are surrounded by a border:

- Select the **Framed** radio button in the **Pictures** section of the Display tab.

Grid spacing

To set the grid spacing (in pixels) on a form:

- Type in a number, or press the up and down arrows to change the current value by one unit in the **Grid Spacing** section of the Display tab.

Choosing a small grid spacing allows you to place objects more precisely by snapping to a tighter matrix.

Color preferences

To set the color preferences:

- Click the **Color** tab of the **Preferences** dialog box.

The color tab allows you to specify the color of various interface elements.

To set a color:

1. Click the button corresponding to the interface element you want to change. The color dialog box appears.
2. Select the new color.
3. Click **OK**.

Tip: Some users prefer not to see red highlighting because it masks the atom colors. You can change the highlight color on the **Color** tab of the Preferences dialog box. Selecting "**black**" will effectively cancel highlighting, leaving atoms to display in their normal colors.

Structural query matching

Search Type, Search Details, and Tuning help you define a query to give the kind of hits you are looking for. The settings are discussed in the chapter on queries under "[Search types](#)" on page 100, "[Setting search details preferences](#)" on page 112, and "[Tuning](#)" on page 114.

General preferences

The general preferences let you set:

- Recent file list size
- Startup defaults
- Frame styles
- Query save defaults

To set the general preferences:

1. Go to **File>Preferences....** The Preferences dialog box appears.
2. Click the **General** tab. The General tab appears.

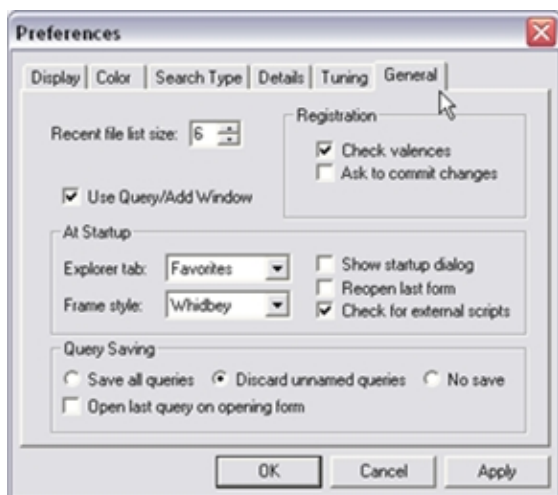


Figure 16.1: Setting general preferences

Structure registration options

To have ChemBioFinder present an alert when attempting to enter a structure with an atom in a non-standard valence state:

- Click the **Check valences** check box in the **Registration** section.

To have ChemBioFinder confirm when you are about to modify data in the database:

- Click the **Ask to commit changes** check box in the **Registration** section.

ChemBioFinder opening options

You can set ChemBioFinder to open with the ChemBioFinder Opening dialog box or to open the last form you were using.

To set the options for what ChemBioFinder displays when it starts up:

1. On the **General** preferences tab:

If you want to...	...then
start with the opening dialog box...	check the box next to Show startup dialog .
start by opening the last form you used...	check the box next to Reopen last form .
use a default script...	check the box next to Check for

If you want to...	...then
	external scripts
change which window is on top in the Explorer window ^a	select an entry from the Explorer tab: drop-down menu.
select a different theme	select an entry from the Frame style: drop-down menu.

2. Click **OK**.

Setting the recent file list size

You can set the number of files you opened recently that ChemBioFinder shows.

To set the list size for the most recent files opened:

- Type in a number (from 0 to 16), or press the up and down arrows to change the current value by one unit in the **Recent file list size:** option on the **General** preferences tab.

Changing options with the view menu

Several structure display options can be set from the View menu, including Atom Numbers, Reaction Centers, Stereochemistry, and Atom Maps in reactions.

To display reaction centers:

Go to **View>Structure>Reaction Centers**.

^aExplorer window is available in Pro and Ultra versions but not in Standard version.

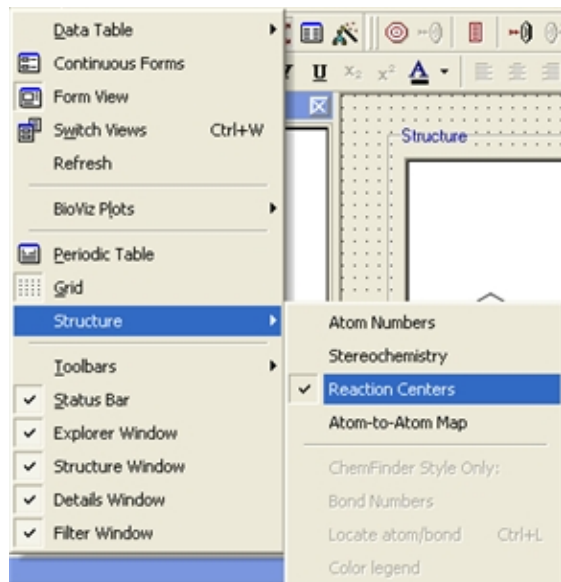


Figure 16.2: Viewing reaction centers

With reaction centers shown, any bond that changes in the course of a reaction is colored. Additionally, any atoms that participate in reaction centers are circled if none of their adjacent bonds participate in the reaction center.

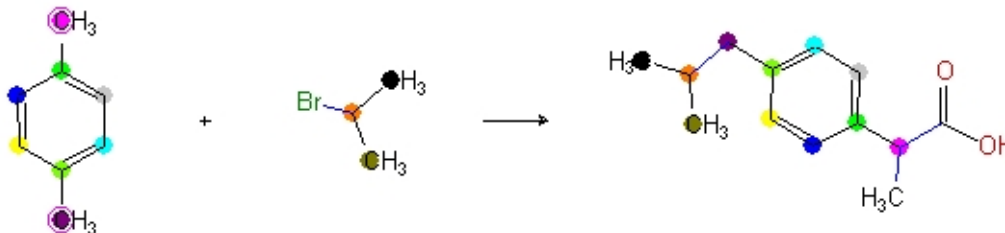


Figure 16.3: Visualization of reaction centers

To display atom-to-atom maps:

Go to **View>Structure>Atom-to-Atom Map**.

With atom-to-atom maps shown, equivalent atoms in reactants and products are colored the same.

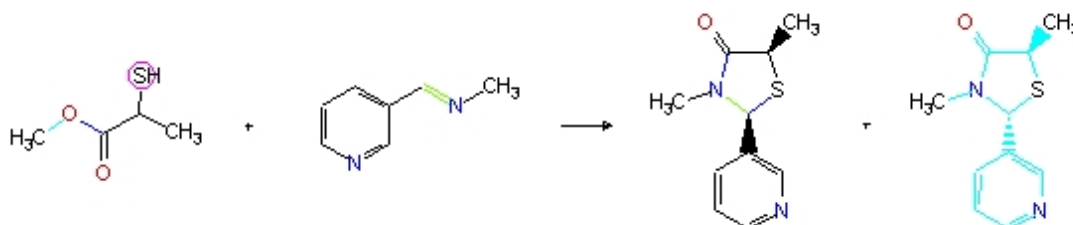


Figure 16.4: Atom-to-atom maps

These two preferences affect only the display of reactions. Not checking the boxes means these types are displayed as all other atoms and bonds.

Favorites tree

The Favorites Tree is a user-constructed collection of file system objects. It may include folders, subfolders, ChemBioFinder forms, structure files, and documents of all kinds. Its purpose is to allow you to collect, organize, and access the data and documents you use regularly.

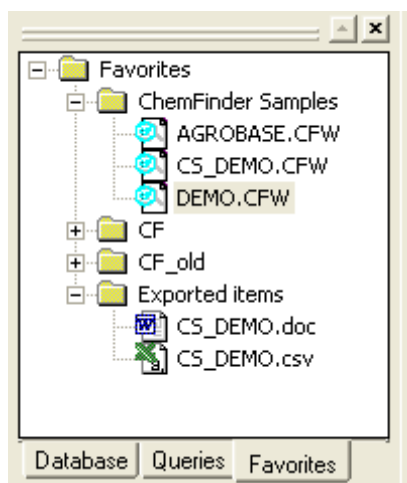


Figure 16.5: The Favorites tree

You can open items in the Favorites Tree from within ChemBioFinder by double-clicking them:

if the file is...	double-clicking will...
a ChemBioFinder form,	open it.
a ChemBioFinder script,	execute it.

if the file is...	double-clicking will...
an SD- or RDFFile,	load it.
a ChemBioFinder query,	restore it.
any other type of document,	open it in its source application.

There are two ways to add items to the Favorites Tree:

- Drag one or more files from Windows Explorer into the Favorites window.
- Use the context menu.

To create a new folder in the Favorites Tree:

1. Point to the “parent” folder. (The new folder will be a subfolder of this folder.)
2. Right-click and click the **New Folder** context menu command.
3. Type in a new name for the folder.

To add an item to a folder:

1. Point to a folder.
2. Right-click and click the **New Item...** context menu command. A File dialog box opens.
3. Browse to the file you want to add, select it, and click **Open**. The item is added to the folder.

To rearrange items or folders in the tree, drag the item to a new location.

To resort a folder or subfolder alphabetically, right-click on the folder and click **Sort Folder** from the context menu.

Toolbars

ChemBioFinder lets you format your toolbars. You can Customize the toolbars by dragging buttons on or off.

To open the **Customize** dialog box:

- Go to **View>Toolbars>Customize**. The **Customize** dialog box appears.

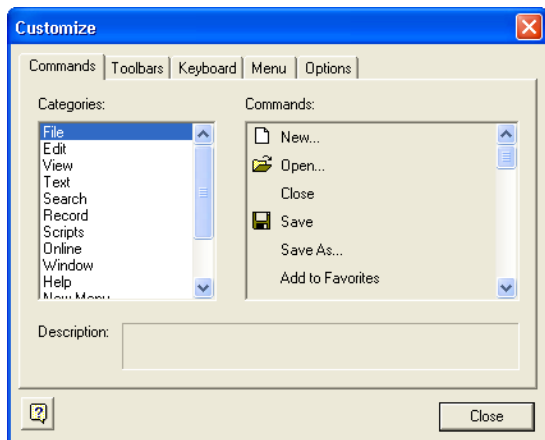


Figure 16.6: Customizing toolbars

To add an option to a toolbar that is already in the ChemBioFinder window:

1. Click the **Commands** tab in the **Customize** dialog box.
2. Locate the command and drag an option from the **Commands** window to a toolbar in the ChemBioFinder window. The option appears where you drop it on a toolbar.

You can delete a button by dragging it off the toolbar.

To return a toolbar to the default settings:

1. Click the **Toolbars** tab in the **Customize** dialog box. The **Toolbars** tab appears and shows all of the toolbars that currently appear in the ChemBioFinder window.

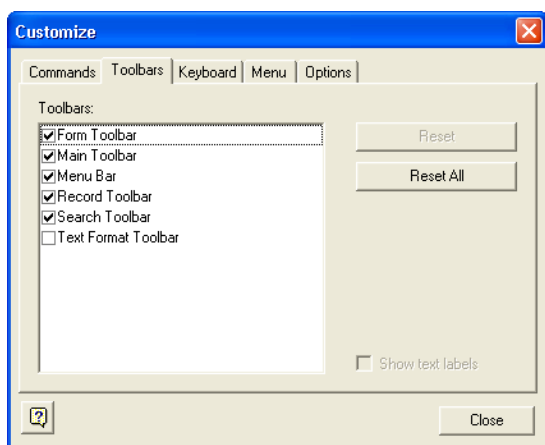


Figure 16.7:

2. Click the toolbar you want to return to default settings, and then click **Reset**. The toolbar in the ChemBioFinder window changes to the default settings.

To return all the toolbars to the default settings:

1. Click the **Toolbars** tab in the **Customize** dialog box. The **Toolbars** tab appears and shows all of the toolbars that currently appear in the ChemBioFinder window.
2. Click **Reset All**.

Periodic table

ChemBioFinder features a periodic table for data display and formula entry. Selecting an element displays physical and historical data.

To display the Periodic Table window:

1. Go to **View>Periodic Table**.

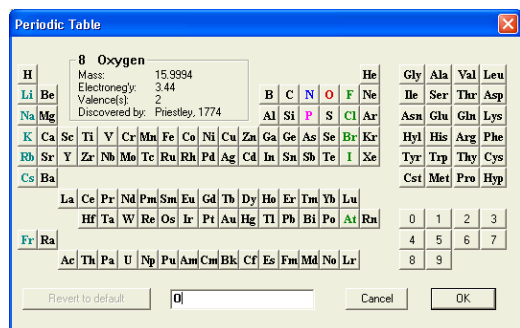


Figure 16.8:

2. Click an element to display its name, mass, and other properties in the top box, and to display its symbol in the bottom edit box.

By clicking different elements and numbers sequentially, you can create a molecular formula in the bottom editable text box. You can then paste this formula into the form for a formula query.

To copy a formula from the Periodic Table to the form:

1. Ensure that the cursor is placed in the editable text box.
2. Select the text to copy using the SHIFT and left arrow keys of the keyboard, or press CTRL+A to copy the entire text displayed in the editable text box.
3. Press CTRL+C.
4. Click **OK** in the Periodic Table window to close it.
5. In the form, click the formula box into which you want to paste the text and press CTRL+V.

To display data about the selected element in the Element Editor, do one of the following:

- Display the desired element and click the display box at the top of the Periodic Table window.
- Double-click the desired element button.

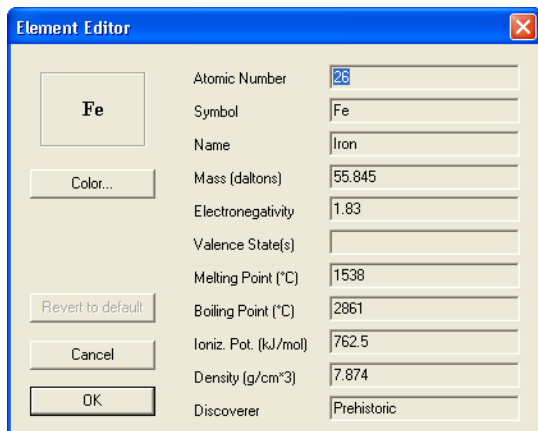



Figure 16.9:

The Element Editor displays properties of the selected element. The color of the element is shown in the periodic table for that element and in any structure data boxes in which that element is present.

To change the color of an element:

1. Click the **Color** button.
2. Select a color from the Color dialog box and click **OK**.
3. Click **OK** in the Element Editor.

To reset an element's color to the default, click **Revert to Default** in the Element Editor.

 **Note:** The data of the Periodic Table is stored in a tab-delimited ASCII file called *CS ChemBioFinder 14.0 CS_Isotopes_Table.txt* in your ChemBioFinder 14.0 system directory. You can edit with a text editor or spreadsheet program to change default data values.

Embedding ActiveX controls

You can embed most windowed ActiveX controls on a form. For stand-alone controls, you can embed the control and use it immediately. If the control needs programming, or if you want the control to manage ChemBioFinder data, you will need to customize its behavior using Python scripts.

To embed a control:

1. Select the **Control box** tool on the Forms toolbar.
2. Drag a rectangle on the form. The Insert Control dialog box appears.
3. Select a control from the list. The control appears in the Control box.

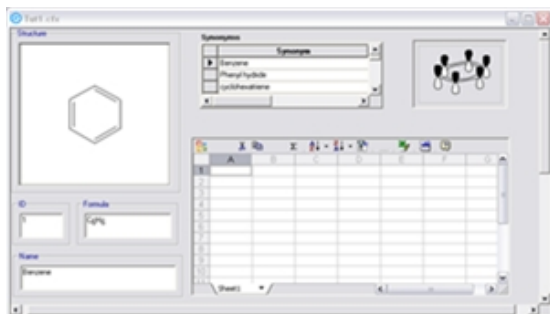



Figure 16.10: Adding a spreadsheet to a form

 **Note:** Not all controls listed are suitable for use in a Control box, but appropriate error messages should be displayed in such cases.

Customizing controls

The behavior of an embedded control can be customized by adding Python scripts to the form to perform custom actions at given times. The control can be accessed using the Automation interface through the Control property of the corresponding Box object, which is obtained from the Boxes collection of the Document object.

Contact [PerkinElmer Informatics support](#) for information on using the Automation interface.


Scripting

CAL

ChemBioFinder is equipped with its own scripting language, the ChemBioFinder Automation Language (CAL). CAL is used to operate the program from the keyboard, or to create custom scripts for automating simple operations such as switching between forms or sending data to Microsoft Excel.

CAL scripts are stored in text files, with their pathnames stored in forms. In ChemBioFinder, you have the option of storing scripts directly in form files. Scripts created in ChemBioFinder are stored internally by default. When you open a form which references external script files, you will be prompted to store the script internally.

If you choose to convert the external script files to internal, the scripts will be saved when the form is saved, and the original script files may be deleted. Scripts originally saved as internal may be also be saved externally.

 **Tip:** You can save the form with a different name, and the original will remain untouched (in which case deleting the script files would not be a good idea).

If you choose to leave the form unmodified, the alert will continue to show up every time the form is opened.

To suppress the prompt:

1. Go to **File>Preferences...**
2. Click the **General** tab.
3. Deselect the **Check for external scripts** check box, then click **OK**.

To execute CAL commands:

1. Go to **Scripts>Command Line**. The Enter CAL Command dialog box appears.
2. Type in a single CAL command. Click **Help** to see [CAL Scripting Help](#).
3. Click **Execute** to run the command. You can keep entering commands, one by one.
4. When you have finished running commands, click **Done**.

The Command drop-down list contains previously entered commands.

To rerun a previously entered command:

1. Select the command from the list.
2. Click **Execute**.

For more information on the command language, see ["CAL Commands" on page 242](#).

Getting CAL help

To display information about the CAL scripting language:

- Click **Help** in the Enter CAL Command dialog box. The CAL Scripting Help window appears containing commands, variables, and syntax notes.

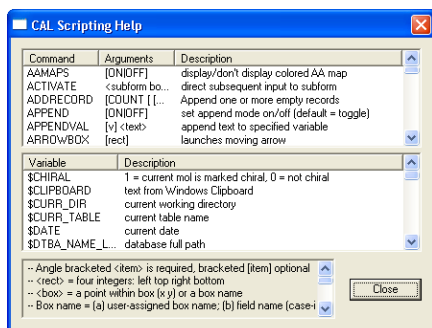



Figure 16.11: CAL help window

 **Tip:** You can re-size this window In ChemBioFinder.

Creating a script

To create a script:

1. Create a button on a form.
2. Label the button.
3. Right-click and click the **Edit Script** context menu command. The CAL editor appears.

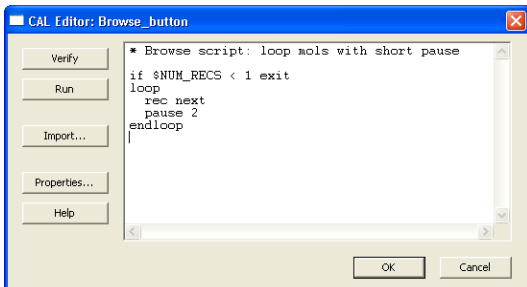


Figure 16.12: The ChemBioFinder script editor

4. Type in your script commands, or use the **Import** button to import an existing script.
5. Click **OK** to create a new file.

Tip: You can still create scripts in Notepad or another text editor if you wish. Save the file with extension `.cfs` in the `...C:\Documents and Settings\All Users\Application Data\CambridgeSoft\ChemOffice 2014\ChemBioFinder\System (windows)` if you want the name of the script to appear on the **Scripts** menu.

The CAL editor is a simple, resizable text-entry window. It accepts carriage returns and tabs. To copy /paste, use CTRL+C/CTRL+V. To undo or redo (last change only) use CTRL+Z.


The **Verify** button runs the script through the CAL command parser. The parser checks only that lines begin with recognized keywords, so just because a script is parsed without error does not mean it will run correctly.

The **Properties** button displays a dialog box used to specify whether the script is to be stored in an external file or internally, and to provide a file path or script name. You can assign a script any name you like, but the name must be unique among scripts on the current form.

The **OK** button saves the script. The **Run** button saves the script and runs it.

To execute a script not assigned to a button:

- Choose the script from the **Scripts** menu.

 **Note:** To execute a script that does not appear on the menu, use the **Command line** and enter “Call <script file-name>”.

Debugging a script

You can step through a script line-by-line when debugging it.

To view a script line-by-line:

1. Go to **Scripts>Command Line**. The Enter CAL Command dialog box appears.
2. Type “step on” and click **Execute**. This turns on the step mode, where each step is displayed.
3. Run a CAL script by doing one of the following:

If your script	Then
appears in the Scripts menu.	choose the appropriate script.
does not appear in the Scripts menu.	type call and the name of your script in the Enter Script Command dialog box, then click Execute .

4. Press any key except Escape to execute the command and go to the next command. As each step is encountered, it is displayed in the status line.
5. Press the Esc key to stop debugging the script.
6. Type step off in the **Enter Script Command** dialog box, to exit the debugging mode.

Trigger scripts

Trigger scripts run in response to certain predefined events. A trigger script can, for example, automatically load a box with data calculated from the contents of another box whenever you move to a new record.

You access a trigger script from the **Run script on** listbox in the **Form** tab of the Properties dialog box. The listbox shows the available trigger events, and allows you to create, edit, enable, or disable scripts for each event.

To run a script on an event:

1. Right-click in the form and click the **Properties...** context menu command.
2. Click the **Form** tab.
3. Click the check box of the desired event in the **Run script on:** list box. If an event is not checked, no script will run on that event, even if one is available.
4. Click the event name to highlight the row of the listbox. (Clicking in a check box does not select the row.)
5. Click **Edit** to write or edit the script in the CAL Editor.
6. Click **OK** to return to the form. The script will automatically be executed at the specified event.
7. Click **OK** in the Box Properties dialog box.
8. Save the form if you wish to keep the changes.

Python™

A Python interpreter is embedded in ChemBioFinder—the language components reside outside ChemBioFinder as well as available from within it and appear to be built in. In the same way you can develop complex macros inside Microsoft Word or Excel using Visual Basic, you can develop scripts inside ChemBioFinder using Python.

The Output window facilitates Python in ChemBioFinder. The script editor is used with either CAL or Python. It's default configuration is Python interactive.

About Python

Python is a full-featured, cross-platform programming language used to develop applications, controls, DLLs, command-line utilities, and so forth. The language is Open Source, maintained and supported by a large community of developers.

Python is available in ChemBioFinder in the same ways as CAL. In both cases, a script can:

- be connected to and activated from a button, form trigger event, or hyperlink in a subform table.
- be developed and executed in command mode, one line at a time.
- provide direct access to data in boxes on the ChemBioFinder form.
- be saved to an external file or stored internally in the form.

But Python is far superior to CAL in many ways:


- Python is a complete programming language, not just a set of commands. It allows calls to system functions, access to databases, OLE Automation, file I/O, program flow control, data structures, object-oriented and structured programming, and all other features offered by a modern language.
- Python can be developed interactively in the ChemBioFinder Script Editor, a development environment similar to IDLE, PythonWin, and others. This environment combines line-by-line execution with program construction, convenient for development and debugging.
- Python scripts can operate ActiveX controls that is also available in ChemBioFinder form boxes.
- Python communicates with ChemBioFinder via OLE Automation, allowing it to access documents, record sets, data, form boxes, structures, CAL commands.
- Python in ChemBioFinder is integrated with ChemScript[®], the chemical structure processing language from CambridgeSoft. ChemScript is an extension to Python which allows carrying out detailed analysis or manipulation of molecules, reactions, and models.

Creating a script

A ChemBioFinder form can contain a collection of scripts, both CAL and Python. Each is connected to the form in one of three ways, depending on how the script will be executed. To create a script, first decide how it will be called.

To create a script activated by a button:

1. Create a button on the form.
2. Right-click the button and click the **Edit script** context menu command. The script editor window opens for script entry.
3. Click **OK** when you have completed the script. The script is executed when the button is clicked.

 **Tip:** You must toggle out of layout mode to use the button.

To create a script connected to a trigger event:

1. Open the Box Properties dialog box, and click the **Form** tab.
2. Select an event in the **Run script on:** box by clicking the check box next to it.
3. Click **Edit**. The script editor window opens for script entry.
4. Click **OK** when you have completed the script. The script is executed when the event occurs.


To create a script linked to a subform table:

1. Right-click on the subform name to open the Subform Properties dialog box.
2. Enter a name in the **Table script:** text box, then click **Edit**. The script editor window opens for script entry.
3. Click **OK** when you have completed the script. The subform contents are displayed in hyperlink style (blue, underlined). Clicking any link executes the script.

The script editor

There is a single script editor for all scripts, CAL and Python. The contents change to match the script being edited. This is a change from the CAL editor in earlier versions, where you could have a separate window for each script. The editor is modeless, that is, it stays on the screen when you switch to work in ChemBioFinder. To close the script editor, click **OK** or **Cancel**.

The editor has three modes:

- CAL – for developing CAL scripts. Use the **Verify** button to check syntax; **Run** to test the script.
 **Note:** To work with CAL interactively, do not use the script editor. Use the command box that appears when you click the **Scripts>Command Line** menu command.
- Python – for editing Python scripts. As in CAL mode, you can view or edit a script and click **Run** to test it. When you switch into this mode, prompts are removed.
- Python Interactive – for interactive development of Python scripts. In this mode, you can enter Python commands and execute them one at a time, or click **Run** to test the script.

The buttons on the script editor are as follows:

- Verify – checks a CAL script for syntax errors. This button does not apply to Python.
- Run – executes a script; used for testing.
- Import... – loads a new script from a text file.
- Properties... – Opens a dialog box with two options: save script internally (with the ChemBioFinder form) or save script as an external text file.
- Help – opens the CAL Help window or Python manual.

Interactive mode

The Python Interactive Mode is modelled on standard Python environments such as IDLE and PythonWin. In this mode:

- Type `cfhelp` for a page of tips and samples of ChemBioFinder-specific programming.
- At each new line, the window shows the Python prompt (`>>>`). After the prompt, you can enter any Python command, then press `<Enter>`. The command is executed immediately, and a new prompt is displayed on the next line.

- The first time you execute a command in the window, there is a short delay while Python is initialized. If it is not already visible, the Output Window then opens, showing a ChemBioFinder header and the output of your command.
- If the line you enter ends with a colon, then it is taken as the beginning of a compound (multi-line) statement. After you press <Enter>, the next line begins with the continuation prompt (...) and is indented. The statement is not executed until you press <Enter> twice in a row, that is, you must enter a blank line to terminate the compound statement.
- You can execute commands other than the one you just typed. Position the cursor on any line and press <Enter>. The command under the cursor is executed. If the line is part of a compound statement, the entire statement executes. Afterwards you are not given a new prompt, and the cursor does not move (unless it was positioned at the end of the line, in which case it inserts a new line and moves to it).
- If you copy text into the window, it is not automatically executed. You can execute pasted code a line at a time by positioning the cursor on each line and pressing <Enter>.
- If you switch from interactive to standard Python mode, all prompts are removed and the text is shifted left. Line-by-line execution is then disabled, and you can edit the text freely.

The output window

The output window is a dockable window for displaying output from scripts. The window comes up automatically the first time it is needed. The text in the window is not editable, but can be copied to the Windows Clipboard for pasting into other applications. Standard keyboard accelerators such as select all (<Ctrl>+<A>), or copy (<Ctrl>+<C>) can be used.

To generate output in the window from a CAL script, use the OUTPUT_STR command.:

```
OUTPUT_STR string or phrase
```

which echoes *string or phrase* into the window, followed by a carriage return. The string should not be quoted.

The string goes through the CAL translator and may be modified with regard to spacing or internal quoting. To prevent translation and have the string taken verbatim, precede it with a backslash character:

```
OUTPUT_STR \use this as written
```

Output from a Python script uses the same mechanism.

To copy text or clear the output window:

- Right-click anywhere in the window and choose the desired command from the context menu.

Programming in Python

Python programming is possible in Pro and Ultra versions but not in Standard version.

A description of the Python language is beyond the scope of this manual. However, to learn more about Python, you will find numerous books and Internet resources on Python, as well as the Python Help file. This section covers programming topics with specific application to ChemBioFinder.

Python operates on your data through the ChemBioFinder OLE Automation interface. This interface provides external access to a library of the objects you work with in ChemBioFinder – application, document (form), recordset, fields, boxes, and others. The interface is documented on the ChemBioFinder SDK site. You may find it more convenient, however, to use an object browser to see the library in its current form.

To illustrate, we will develop an example Python script. Suppose you have a form open, you perform a search, and get a hit list. You would like to generate a report that contains a set of records in formatted blocks showing specific data items.

You can perform this task using the script shown below. You can copy and paste the script into the Script Editor to try it or use it as a template:

```
# ChemFinder Python script 'Generate Report'
# Created Wednesday, November 01, 2006, 10:55 AM
# Loop records printing name and formula
from win32com.client import GetActiveObject
cfapp = GetActiveObject("ChemFinder11.Application")
cfdoc = cfapp.ActiveDocument
cfrset = cfdoc.Recordset
namefield = cfdoc.Fields("molname")
fmlafield = cfdoc.Fields("formula")
recno = 1
cfrset.MoveFirst()
while (not cfrset.EOF):
    molname = cfrset.Value(namefield.index)
    formula = cfrset.Value(fmlafield.index)
    print 'Rec',str(recno),': Name =',molname,'; Formula =',formula
    cfrset.MoveNext()
    recno = recno + 1
    namefield = 0
    fmlafield = 0
    cfrset = 0
    cfdoc = 0
    cfapp = 0
```

Notes

- Lines 1-2: Comments beginning with # are ignored. ChemBioFinder generates lines like these whenever you create a new script.
- Line 3: It's a good idea to add your own comment telling what a script is designed to do.
- Line 4: Import statements are used to declare and execute code from external Python modules. This one says we want to use a specific routine (GetActiveObject) from a module in the Win32 extension library (win32com.client).
- Lines 5-7: Obtain the currently-running ChemBioFinder application object -- the one from which the script was launched -- along with its active document and recordset. Assign them variable names.
- Lines 8-9: Identify the fields to be included in the report, and get corresponding field objects from the document.
- Lines 10-11: Prepare to loop records. Calling MoveFirst ensures that you loop the entire list. If you omit this or comment it out, the loop begins at the current record.
- Line 12: Begin the loop, go until the recordset reaches EOF after the last record. A colon at the end of a line indicates it is the first of a compound statement.
- Lines 13-14: Retrieve data, using the Value method of the recordset.
- Line 15: Format and output the record data. Our simple example just prints a line to the output window.
- Lines 16-17: Proceed to the next record.

- Lines 18-22: Objects created in a script must be freed before leaving. Without this, you may find it impossible later to exit ChemBioFinder. An object can be freed by setting it to zero.

```
Rec 1 : Name = Bromobenzene ; Formula = C6H5Br
Rec 2 : Name = 3-Bromobenzoic acid ; Formula = C7H5BrO2
Rec 3 : Name = 6,6'-Dibromoindigo ; Formula = C16H8Br2N2O2
Rec 4 : Name = 3-Bromophenol ; Formula = C6H5BrO
```

Caveats

- Subforms are not supported by the CF Automation interface. If you need to work with subform data, you will have to use CAL for the time being.
- Running Python scripts cannot currently be interrupted. Be warned before trying the above script on a long hit list! Once it starts running, it will not stop until finished or until you kill ChemBioFinder in Task Manager. We are working on this problem.
- If after you have run a script, you try to exit ChemBioFinder and it refuses to exit, it means the script created a ChemBioFinder object and failed to release it.

Methods of communication

There are two general methods of communicating with other Windows applications such as Microsoft Excel: by using a script within ChemBioFinder, or by using a Visual Basic procedure within the other application. The following is an overview of each method.

Using scripts

A ChemBioFinder (CAL) script can communicate with other Windows applications using either of two commands: EXEC—to start an application and possibly pass information on the command line.

DDE—to communicate using Dynamic Data Exchange with a DDE-ready application.

Using EXEC is straightforward, but limited. You can start all Windows applications by this command. Most can be passed a filename on the command line, such that the specified file is opened (or printed) on startup. A few applications can accept more detailed instructions. Consult the application's manual for information about how it can be operated using the command line.

If you have Visual Basic or similar programming language, you can extend the power of EXEC. You can write an application using the advanced features of Visual Basic, and then call the application from within ChemBioFinder using the EXEC command.

Using DDE is more complicated. You can operate most Microsoft Office components and many other programs to varying extents with DDE. For example, practically every command on the Excel menu can be executed by DDE.

The syntax is rather difficult, but can usually be worked out by experimenting and consulting ChemBioFinder help. An example is given below.

DDE is the most direct way of using Excel to view data from ChemBioFinder.

To use MS Excel to view ChemBioFinder data:

1. Start MS Excel. You can start it manually using a CAL script or by starting the application in Windows.

2. In ChemBioFinder, obtain the hit list you want to transmit to Excel. If you want to work with the entire database, go to **Search>Retrieve All**.
3. Execute a short CAL script (below) which exports the hit list as comma-delimited text to a temporary file, then instructs Excel with DDE to load that file into a spreadsheet.
4. Activate Excel to work with the data in the spreadsheet.

This procedure takes data one way, from ChemBioFinder to Excel. Returning modified data from Excel to ChemBioFinder can be done using other techniques described in this chapter.

Here is a script to start up Excel:

```
*RUNEXCEL.CFS - script to start Excel  
*
```

```
EXEC "c:\msoffice\excel\excel.exe"
```

If the Excel program is on your search path, you can eliminate the complete pathname and just give the executable name ("exec excel.exe"); if not, you may need to modify this script to indicate where EXCEL.EXE is located on your system.

Here is a script to transfer the current hit list from ChemBioFinder to Excel:

```
*TOEXCEL.CFS - script to send data  
*to Excel  
*
```


```
WRITETEXT C:\DATA.TMP
```

```
DDE Excel System [OPEN("C:\DATA.TMP")]
```

```
DDE Excel
```

```
System [COLUMN.WIDTH(1,"C1:C4",,3,1)]
```

The first line writes out the current ChemBioFinder hit list as a temporary delimited ASCII file. By default, all fields that appear in boxes on the current form—except structure, but including formula and molecular weight—are written. The second line instructs Excel to open the file. Excel can automatically recognize the file format as tab-delimited. The third line instructs Excel to auto-size column widths 1–4 to fit their contents.

 **Note:** This example requires that the text export delimiter be set to TAB. otherwise Excel may not read the file correctly. To check this, go to **File>Export**, then open the Text Options tab of the Data Export dialog box.

You can include either or both of these scripts on the Scripts menu, and you can activate them with buttons on the form. To include a script on the Scripts menu, give it a filename with extension CFS, and place it in the ChemBioFinder System directory, or in the directory containing the ChemBioFinder application. To activate a script from a button, label the button with a script filename or string which can be converted into a filename. For example, if

TOEXCEL.CFS exists in the ChemBioFinder System directory, label a button "ToExcel" to start the script. For more information, see ["Adding a button" on page 38](#).

Using Visual Basic

The second method of communicating between ChemBioFinder and other applications such as Excel is using OLE Automation. ChemBioFinder is an OLE Automation server, meaning that it offers a collection of data management capabilities to outside programs capable of communicating with OLE objects. While this collection is currently fairly small, it is adequate for a variety of data retrieval and search tasks. This feature allows you to write a custom Visual Basic procedure that directly retrieves and manipulates data from ChemBioFinder.

The general procedure for accessing ChemBioFinder data from within a Visual Basic script is as follows.

1. Create a ChemBioFinder Document object, typically passing a filename so that you open a form complete with its database connection.
2. Use methods of the Document object to move through the database, search, and access data. Document methods include some that access Field objects, used to query the data in the database, and Molecule objects, for accessing details of molecular structures.

For more information, see the [CambridgeSoft SDK web site](#).

Using Microsoft Access

The methods described above for communicating between ChemBioFinder and Excel apply also to Access. You can start Access using the EXEC command. You can send it DDE commands contained in a CAL script, although Access provides fewer capabilities with DDE than does Excel. Or you can write programs using Access Basic that rely on the OLE Automation methods found in ChemBioFinder. In addition, you can use Access directly to operate on a ChemBioFinder database.

A ChemBioFinder molecule database consists of three components: the structure storage files (with file extensions MST and MSX), the data storage files, which include a Microsoft Access database (file extensions MDB and LDB), and the forms (with file extension CFX) used to view the structures and the data. For more information, see ["Database File Types" on page 221](#).

If you have Access on your system, double-clicking an MDB file in Explorer starts up Access and opens the specified database.

When you open a ChemBioFinder database in Access, you will see the same tables as displayed in the ChemBioFinder Database dialog box, including the main structure table (usually named "MolTable"), but you will not see columns for structure, formula, or molecular weight. These fields cannot be manipulated directly using Access. The following are some of the operations you can perform on a ChemBioFinder database using Access. Most of these capabilities are not available through the current version of ChemBioFinder:

- Compress or repair the database.
- Change column (field) or names or formats.
- Change table names.

- Add or delete columns or tables.
- Import or export tables.
- Load non-structural data from various file types, including delimited ASCII, Excel, or Word.
- Move quantities of data from one column or row to another.
- Carry out complex queries on non-structural data.
- Permanently change the sort order of a table.

For more information about these actions, please consult the Microsoft Access User's Guide.

Do not add or delete records to the MolTable within Access, because the data component of the database will become out of synchrony with the structure component.

Database File Types

ChemBioFinder databases are supported by a library called MstLib. This library contains a number of files for supporting the structure column. In addition, ChemBioFinder uses additional screening files for improving performance of search operations.

The following table lists the various file types used by ChemBioFinder:

Extension	Purpose
.MST	Structure Storage
.MSI	Screen: Traditional
.MSX	Screen: Traditional
.MSF	Screen: Full Exact
.MSK	Screen: Skeletal
.MSS	Screen: Similarity
.MS3	Screen: 3-D
.MSH	Screen: Hash
.CFW	ChemBioFinder Form
.CFX	ChemBioFinder Form
.CFS	ChemBioFinder Script
.MDB	MS Access database
.LDB	Locks on database

 **Note:** When .MSS extension is missing, the .MSX extension is used, but the quality of the search is lower.

Structural Query Features

To perform a substructure search in ChemBioFinder, you must first draw the query structure itself. Query structures can be drawn in many different programs, but we recommend using ChemDraw. This chapter is focused on using ChemDraw to draw query structures. For more information about the structure drawing and query capabilities of ChemDraw, please see the *ChemBioDraw User Guide*.

ChemBioFinder does its best to follow your instructions even if those instructions are contradictory. For example, you can create a query such as the following:

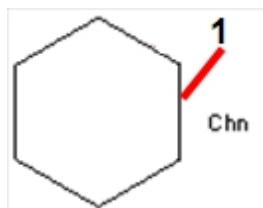


Figure 17.1: 1) This bond must not be in ring

That bond is already in a ring, so ChemBioFinder returns no hits for this query.

General properties

ChemBioFinder allows the following general properties to be assigned to a query:

- Atom
- Bond
- Substituents
- Charges and radicals
- Isotopes
- Stereochemistry
- Normalization

Atoms

Atom types specified in the query must match atoms at corresponding positions in the target. Hydrogen is an exception—see ["Substituents" on page 223](#).

Bonds

All bonds explicitly drawn in the query must match in the target. For certain caveats, see ["Stereochemistry" on page 226](#) and ["Normalization" on page 229](#). ChemBioFinder recognizes the following standard bond types:

Bond Type	Description
Single	target must have single bond here
Dashed	same as Single
Hashed	same as Single
Thick	same as Single
Wedged Hashed	specifies stereochemistry down from the point end to the wide end
Wedged	specifies stereochemistry up from the point end to the wide end
Wavy	specifies stereochemistry "either" at both ends
Hollow Wedged	same as Wedged
Dative	same as Single
Double	target must have double bond; stereo dictated by geometry
Double Either	target must have double bond; any stereochemistry ok
Double Bold	same as Double
Triple	target must have triple bond here

Substituents

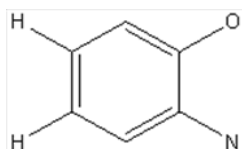
In ChemBioFinder, a substituent is defined as a non-hydrogen atom connected by a bond of any order. For example, a carbonyl oxygen is a substituent of the carbonyl carbon.

All unfilled valences in the query may be filled by hydrogen atoms or by non-hydrogen substituents. The normal valence of an atom is determined from data in the Periodic Table window. For example, carbon has a valence of 4, while sulfur has valences of 2, 4, and 6. Any explicit charges, radicals, or query properties modify the normal valence. For example, a carbocation has a valence of 3.

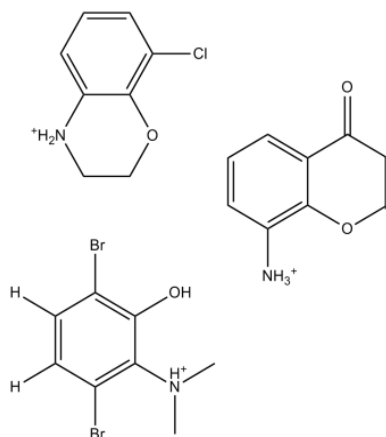
Hydrogen atoms in the query may match non-hydrogen substituents in the target if the hydrogen in the query is implicit on an unlabeled carbon atom or heteroatom. This is the default setting.

Hydrogen atoms in the query must match hydrogens in the target when the query hydrogen is at the end of an explicit bond. The matched hydrogen in the target may be implicit in an unlabeled carbon atom.

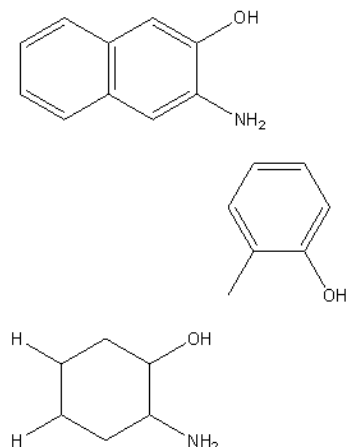
The following query with explicit hydrogens:



Finds any of:



Does not find any of:

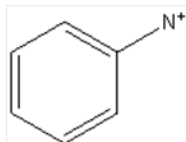


Charges and radicals

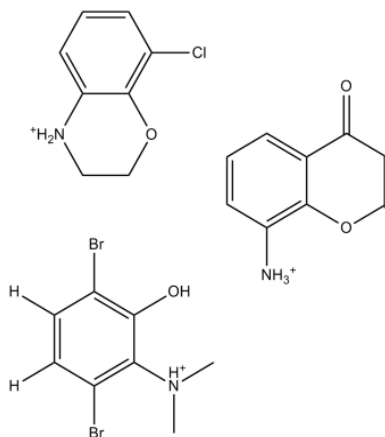
Charges or radicals specified on atoms in the query must match those in the target. Uncharged atoms in the query may or may not match charged atoms in the target, depending on the state of the appropriate check box in the Search tab of the Preferences dialog box.

The valence of a charged atom is taken to be the valence of the isoelectronic neutral atom.

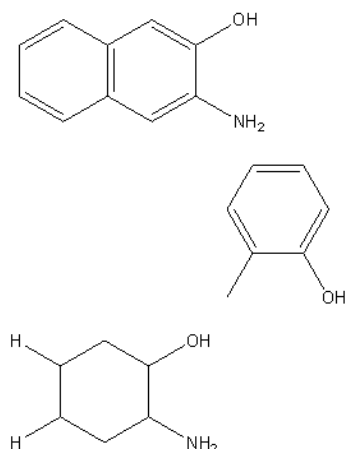
With a substructure search, the query with charge or radical:



Finds any of:



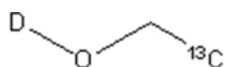
Does not find any of:



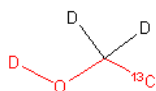
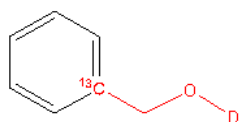
Isotopes

Isotopic labels specified in the query must match the target. Unlabeled atoms in the query match unlabeled or isotopically labeled atoms in the target. Additionally, D is treated interchangeably with ^2H , and T is treated the same as ^3H .

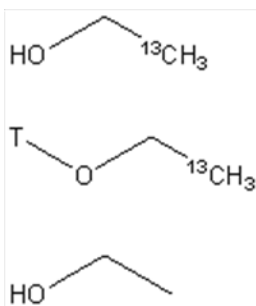
With a Substructure Search, the Isotopic query:



Finds any of:



Does not find any of:



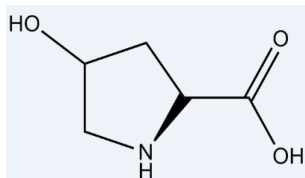
Stereochemistry

Stereochemistry specified on the query must match the target if the relevant Match Stereo item is selected in the Search tab of the Preferences. Stereochemistry is specified at a tetrahedral site by using stereo bonds (up, down, either). Stereochemistry about a double bond is specified by the geometry of the drawing. ChemBioFinder cannot currently interpret other stereochemistry types (allenic, square planar, octahedral, etc.) and ignores them during a search. When evaluating a possible match, the following rules are applied:

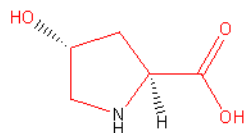
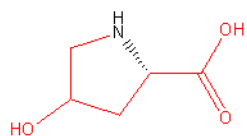
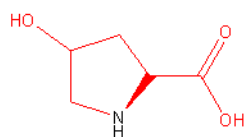
- Unspecified stereochemistry (a plain bond) may match any stereochemistry (either a wedged, hashed, bold, or a plain bond).
- Specific bond types need not match as long as the overall stereochemistry at a given atom does match.

- Implicit hydrogens are taken into consideration in both the query and the target if doing so helps to determine the chirality of a stereocenter.

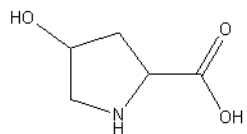
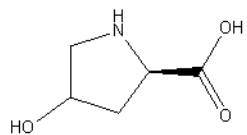
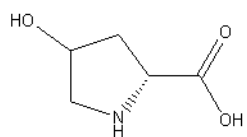
With a full structure search and the Match Tetrahedral stereo option selected, the query with stereochemistry:



Finds any of:



Does not find any of:



Relative tetrahedral stereochemistry

Relative Tetrahedral Stereochemistry, or RTS, specifies a given relationship between the centers. That is, a known orientation of the substituents with respect to each other, rather than a known absolute configuration. To specify this, centers are drawn, not with the standard hashed and wedged bonds, but with thick (bold) stereo bonds.

Relative and absolute configurations may be registered separately in ChemBioFinder, and can be distinguished by various search options.

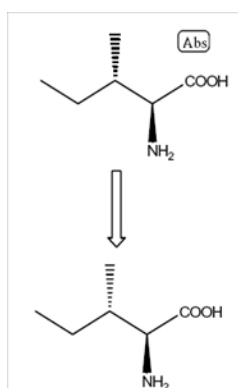
The basic ChemBioFinder search philosophy can be summarized as follows: a more specific query is a more precise request, and should get fewer hits than a more general query. If the query represents a particular absolute configuration, it should hit only that; if it represents a mixture, it should hit any of the components.

If you choose *Same* in the stereo search choices, you are requesting that whatever stereochemistry is specified in the query must match that of the target. If the query has relative bonds and RTS is activated, a hit must have the same relationship between centers.

You can use the following three flags in searches based on relative stereochemistry:

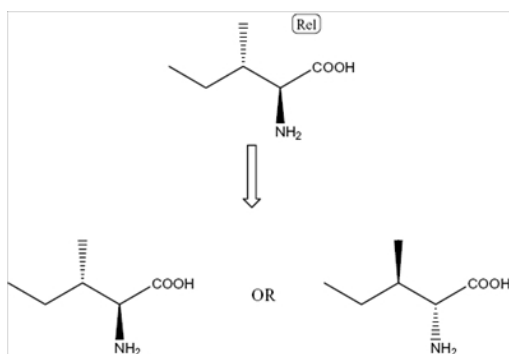
- **Abs:** Refers to the Absolute flag. It matches the exact stereoisomer, as drawn.

Example:



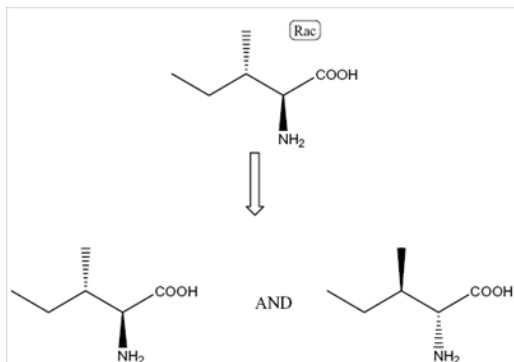
- **Rel:** Refers to the Relative flag. It matches the exact stereoisomer, as drawn, or its enantiomer.

Example



- **Rac:** Refers to the Racemic flag. It matches a mixture of the exact stereoisomer, as drawn, and its enantiomer.

Example



MDL file formats

All bond styles available in ChemDraw are retained in CDX files. The same is not true of major MDL structure files: MolFiles, SDFfiles, and RDFfiles. In these formats, only three stereo bond types are available: hash, wedge, or either (squiggly bond); thick bonds are not recognized. This presents a problem for the ChemBioFinder user wishing to import or export structures having relative stereocenters.

ChemBioFinder addresses this problem by using a proprietary tag in the file, recognized only by ChemBioFinder. The tag is BOND_RELS, and is applied to every stereo bond of thick type. For example, here is an excerpt from the bond table of a molfile with relative bonds.

...

```
4 5 1 0 0 0 02 6 1 1 0 0 0      <- bond 5, marked as UP4 7 1 1 0 0 0      <- bond 6, marked as UPM CFW 5
BOND_RELS      <- overrides type of bond 5 to be UP/THICKM CFW 6 BOND_RELS      <- same for bond 6
```

The consequences of having a ChemBioFinder-only tag are:

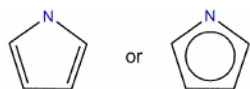
- When any of the MDL file formats is saved from within ChemBioFinder, the tag is written. Reading these files back into ChemBioFinder will retain the thick bond types.
- If a structure is drawn in ChemDraw and saved as MDL Molfile, thick bonds are lost, and are converted to normal hash/wedge bonds. (This is not true if the file is saved as CDX).
- If a structure is saved as molfile from ChemBioFinder, reading it into ChemDraw loses the thick bonds.
- Mol- or SDFfiles obtained from ISIS or other programs do not know about this tag, and thus cannot convey thick bonds. However, such files may be edited by hand to include the ChemBioFinder tag.
- An MDL file written by ChemBioFinder, containing tags, should in principle be readable by any program which can import such a file. The unrecognized tags should simply be ignored.

Normalization

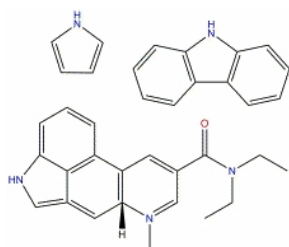
Closed rings of alternating single and double bonds are “normalized”. This is required to make their substituents equivalent. (Ortho-dimethylbenzene is remains same whether there is a single or double bond between the two substituted

ring atoms.)

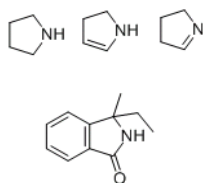
In a substructure, non-tautomeric query, the following query:



will get the following hits:



and will NOT hit the following:



Normalized query

Changing the scope of a search

When preparing a chemical structure to be used in a query, there are two ways to broaden or narrow the scope of the search:

- Specify [atom](#) or [bond](#) properties
- Specify [atom](#) or [bond](#) Special Types

In general, properties narrow a search and special types may either narrow or broaden the search.

Atom properties

ChemBioFinder allows special atom properties to be assigned to an atom in a query. These properties are usually only meaningful during a search. They generally serve to broaden or narrow the scope of the search.

Special atom types

ChemBioFinder recognizes six special atom types that can match any one of a predefined set of elements:

- A matches any non-hydrogen atom.
- Q matches any heteroatom (non-hydrogen, non-carbon).

- R matches any atom, including hydrogen. Also used to indicate a link node. See "[Link nodes and multivalent Rs](#)" on page 132 for details. A link node may be any number of atoms, including zero.
- X matches any halogen (F, Cl, Br, I, At).
- M matches any metal atom, shaded in the periodic table below:

H																	He																												
Li	Be											B	C	N	O	F	Ne																												
Na	Mg											Al	Si	P	S	Cl	Ar																												
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr																												
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe																												
Cs	Ba	La	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn																												
Fr	Ra	Ac	Rf	Db	Sg	Bh	Hs	Mt	Ds	Rg																																			
<table border="1" style="margin-left: auto; margin-right: auto;"> <tbody> <tr> <td>Ce</td><td>Pr</td><td>Nd</td><td>Pm</td><td>Sm</td><td>Eu</td><td>Gd</td><td>Tb</td><td>Dy</td><td>Ho</td><td>Er</td><td>Tm</td><td>Yb</td><td>Lu</td> </tr> <tr> <td>Th</td><td>Pa</td><td>U</td><td>Np</td><td>Pu</td><td>Am</td><td>Cm</td><td>Bk</td><td>Cf</td><td>Es</td><td>Fm</td><td>Md</td><td>No</td><td>Lr</td> </tr> </tbody> </table>																		Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr
Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu																																
Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr																																

Figure 17.2: Metals in the periodic chart

Atom lists

As with the predefined special atom types, an atom list is a list of atoms, one of which must match the target atom.

For example:

[Cl,Ag,N] atom must be Cl or Ag or N

Atom lists may contain only elements. Special atom types, nicknames (Ph), and structural fragments (NH₂, OCH₂CH₃) may not be included in an atom list. ChemBioFinder recognizes a maximum of five atoms in an atom list.

Atom not-lists

The opposite of an atom list is a list of atoms, none of which must match the target atom. For example:

[NOT O,S,Se] atom must not be O or S or Se (but may be any of the 100 other elements)

Atom not-lists have the same restrictions as atom lists.

Substituents: exactly

This property specifies a precise value for the number of substituents on an atom, including those explicitly drawn.

This property is only meaningful in a substructure search.

With a substructure search, the query:

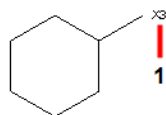
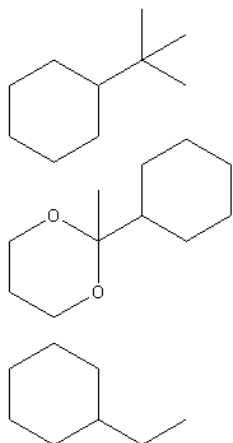
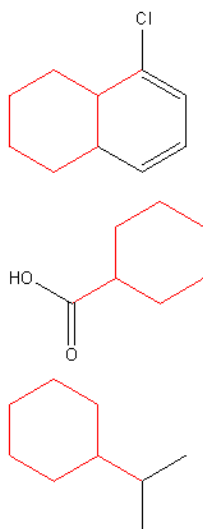


Figure 17.3: 1) This atom is marked with the atom property, *Substituents Exactly:3*.

Finds any of:



Does not find any of:



Substituents: up to

This property specifies a maximum value for the number of substituents on an atom, including those explicitly drawn.

This property is only meaningful in a substructure search.

With a substructure search, the query with substituent up to property:

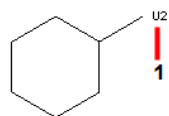
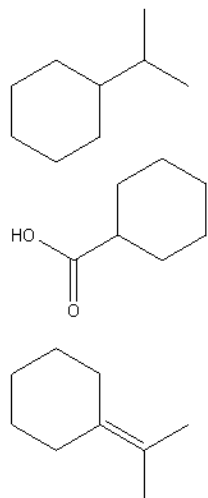
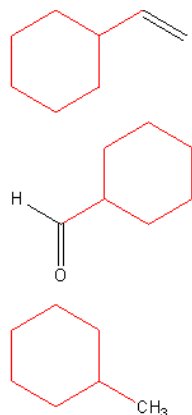


Figure 17.4: 1) This atom is marked with the atom property *Substituents Up to:2*.

Finds any of:




Does not find any of:



Substituents: free sites

The Substituents: Free Sites property specifies the maximum number of additional substituents that may be present on an atom. This property is only meaningful in a substructure search.

 **Tip:** Specifying Free Sites: 0 is a quick way to indicate that you want no further substitution at a site. Target structures will match the query structure as drawn, with no additional ligands.

With a substructure search, the query with free sites property:

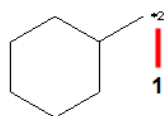
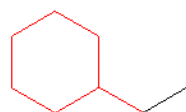
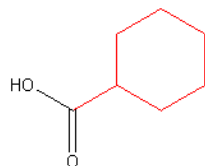
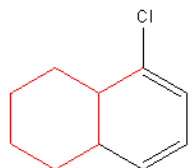
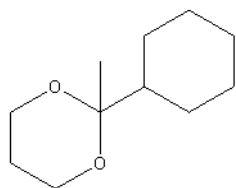
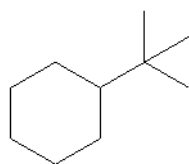


Figure 17.5: 1) This atom is marked with the atom property Substituents Free Sites:2.

Finds any of:




Does not find any of:



Implicit hydrogens

This atom property may have either of two values: Allowed (default) or Not Allowed. If implicit hydrogens are Not Allowed, the atom must be fully substituted in the target.

This property is only meaningful in a substructure search.

 **Note:** This atom property does not affect the display of implicit hydrogens, only their presence in a search. For more information about displaying implicit hydrogens, see ["Setting preferences" on page 199](#).

Unsaturation

Sometimes it is useful to specify that an atom must or must not be attached to unsaturated (aromatic, double, or triple) bonds. ChemBioFinder allows searches for atoms whose unsaturation Must Be Absent (all bonds to the atom are single). It also allows searches for atoms with at least one multiple (double, triple, or aromatic) bond. The default value, Undefined, finds targets without regard to the hybridization of the atom.

This property is only meaningful in a substructure search.

With a substructure search, the query:

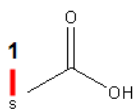
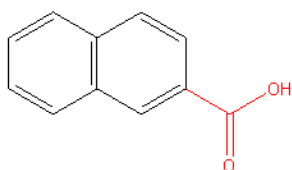
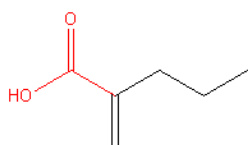
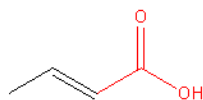
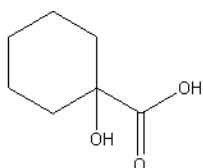
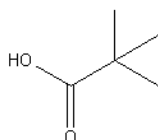
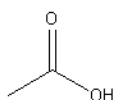


Figure 17.6: 1) This atom is marked with the atom property *Unsaturation: must be present*.

Finds any of:



Does not find any of:



Bond properties

ChemBioFinder allows special properties to be assigned to bonds in a query. These properties usually will only be meaningful during a search. They generally serve to broaden or narrow the scope of the query.

Special bond types

The table below describes the special bond types that ChemBioFinder allows.

Bond Type	Description
Aromatic	target bond must be aromatic as defined by the Hückel $4n+2$ rule
Any	target can have any bond type here
S/D	target must have a single bond or a double bond here
Tautomeric	same as S/D
D/A	target must have a double bond or an aromatic bond here
S/A	target must have a single bond or an aromatic bond here

Topology

If Ring or Chain is chosen, the target bond must or must not be in a ring, respectively.

Reaction center

The reaction center refers to those bonds that are directly affected by a reaction. This property allows you to specify just how a given bond is affected.

Property	Description
Unspecified	target must have a bond here, but the bond can participate in the reaction in any fashion, or not at all
Center	target must have a bond here that directly participates in the reaction in some way
Make/Break	target must have a bond here that is either made (if

Property	Description
	in a product) or broken (if in a reactant)
Change	target must have a bond here whose bond order changes over the course of the reaction, but is not made or broken
Make&Change	target must have a bond here that is either made (if in a product) or broken (if in a reactant) or whose bond order changes over the course of the reaction
Not Center	target must have a bond here, and that bond must not participate in the reaction
Not Modified	target must have a bond here, and regardless of whether it is part of the reaction center or not, its order must not change over the course of the reaction

This property is only meaningful when searching a reaction database.

Formula Input Rules

In formula queries, you are not specifically required to enter symbols in any case. This can give rise to ambiguities.

This chapter describes how ChemBioFinder interprets formulas and resolves these ambiguities.

For example, in a formula query, when you type “cooh”, most chemists recognize this as a carboxyl group: carbon, two oxygens, hydrogen. However, it might be interpreted as cobalt (Co), oxygen, hydrogen. Similarly, “CLI” might be carbon-lithium or it might be chlorine-iodine. “PHE” might be phosphorus-helium or, since ChemBioFinder allows 3-character amino-acid symbols as special atom types, it might be phenylalanine (Phe).

One way to avoid this ambiguity is to separate letters with spaces. For example, the string “c o” cannot possibly be interpreted as cobalt. Another way to avoid ambiguity is to enter formulas using the Periodic Table instead of typing them. ChemBioFinder lets free-format input, and attempts to make the most reasonable interpretations of them by using the rules described below.

Rules

If a symbol is properly capitalized (first letter upper case, followed by zero, one, or two lower-case letters), then the *longest* valid symbol which matches is preferred. Thus “Phe” matches phenylalanine rather than phosphorus; “Co” matches cobalt.

Because of this rule, if you properly capitalize all symbols, no ambiguities will arise.

If symbols are not properly capitalized (such as all lower-case), then, with the exception noted below, the shortest symbol which matches is preferred. Thus “co” is taken as carbon-oxygen instead of cobalt.

The exception is: if two characters represent a valid two-letter symbol and also a valid one-letter symbol followed by an invalid one, then the two-letter symbol is favored. Thus “cl” is not taken as carbon followed by the (invalid symbol) L, but instead is taken as chlorine.

Examples

There is an easy way to experiment with formula interpretation: use the Periodic Table. Type a formula into the text box at the bottom, then click anywhere outside that box. ChemBioFinder interprets the formula and redisplay it with correct capitalization.

ChemBioFinder interprets some ambiguous formulas as follows:

Formula	ChemBioFinder Interpretation
cooh	COOH
Cooh	CoOH
cnosi	CNOSI

Formula	ChemBioFinder Interpretation
cNoSi	CNoSi
bru	BrU
b ru	BRu

Similarity Rules

A common situation is that you are not looking for a specific compound, but any compound that is “close enough” will do. ChemBioFinder uses the Tanimoto equation to determine if compounds are similar. You can specify “how similar” in the Searching tab of the Preferences dialog box.

Exact searching relies on the notion of an atom-and-bond table: a specific set of atoms joined to each other in a certain order by a given set of bonds. If some atoms or bonds are missing, added, or different, the query structure and the target structure do not match.

Most similarity searching, on the other hand, relies instead on the notion of molecular descriptors. Each compound can be represented by a collection of qualitative terms that describe general aspects of the structure.





For example, benzoic acid might be described as:

- organic acid
- contains 6-membered ring
- contains de-localized ring
- contains C-double-bond-O

As you can see, the descriptors can be very broad, and they can overlap. The set of descriptors used by ChemBioFinder is very large.

When you draw a compound, ChemBioFinder compares it against all of the descriptors it knows about. Some descriptors will be present in your compound and some will not. Since each descriptor for a given compound is either Present or Not Present, they are often stored as bits, and another name for a compound’s set of descriptors is its bitscreen.

Consider the following query and target compounds:

Bitscreen	Desc.	#	Name
	Query	23	Q
	Target	24	T
	Q and T	17	<i>Q ∩ T</i>
	Q or T	30	

Both have similar numbers of descriptors present - the query has 23, while the target has 24.

Complete structure similarity

One of the hallmarks of a good similarity algorithm is whether it is *commutative*. That is, two compounds should have the same similarity value no matter which you compare to which. The full structure Tanimoto similarity test is commutative. It compares the number of descriptors they have in common (in the intersection of the query and the target) to the number of descriptors they have in total (in the union of the query or the target). The ratio of these two values is known as a Tanimoto coefficient, and is always a value between 0% and 100%.

$$\frac{Q \cap T}{Q \cup T} \quad \textit{Tanimoto Coefficient}$$

For the two compounds above, the Tanimoto coefficient is 17/30, or about 57%. This is not very similar. although ChemBioFinder will allow you to specify any Tanimoto value down to 0%, for most cases you will likely be looking for compounds that have Tanimoto coefficients of 90% or higher.

Substructure similarity

Unlike full structure similarity, substructure similarity is not commutative: you are comparing a portion of one structure against the entire other structure, and so it does matter which you compare to which. In considering substructure similarity, ChemBioFinder finds what percentage of descriptors in the query are also present in the target. This value will always be at least as large as the complete-structure Tanimoto coefficient for the same two compounds, and usually it will be larger. The two compounds above are 17/23, or about 74% similar by substructure similarity. For a given coefficient value, a substructure similarity search will always return all of the hits in a full structure similarity search, and will often return additional ones as well.

CAL Commands

ChemBioFinder Automation Language (CAL) is a set of commands that control many ChemBioFinder operations.

You invoke CAL commands in either of two ways:

- Interactively: Type commands and execute them one at a time on a command line.
- Scripts: Use ASCII files containing a series of CAL commands to be executed automatically in sequence.

Arguments to commands consist of numbers and text strings. You can enter a text string without punctuation unless the argument is not the last one on the command line and contains multiple words or spaces. In this case, you must enclose the argument in quotation marks.

CAL help

Information about the CAL command and variables is available in the CAL Scripting Help window.

To access CAL Help:

1. From the Scripts menu select **Command Line**. The **Enter Cal Command** dialog box appears.
2. In the **Enter Cal Command** dialog box, click **Help**. The CAL Scripting Help window appears:

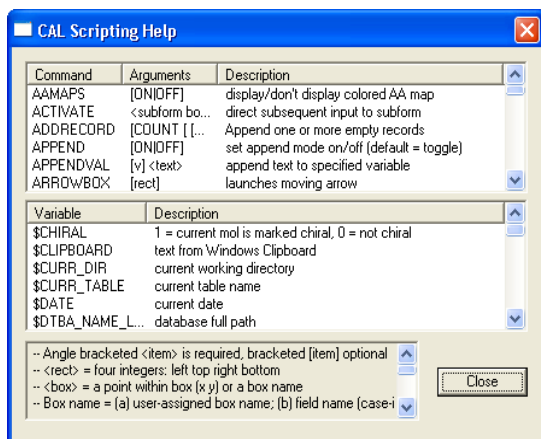


Figure 20.1: The CAL scripting window

Menu commands

CAL commands MENU syntax

Menu commands consist of a two-word command from the main menu and in some cases an optional argument.

Following are examples of commands:

file new

edit paste

search find

record next

You do not need to spell out menu commands. You can enter just enough letters to uniquely identify the menu option.

For example, to execute the Record Next command, type:

```
rec n
```

Some menu commands open a dialog box and wait for user feedback. To avoid this, a limited number of menu commands can take explicit arguments:

FILE OPEN [*filename*]

FILE SAVE [*filename*]

RECORD GO TO RECORD [*recno*]

The command:

```
menu menuitem [param]
```

executes any menu command as if selected from the menu.

The menu commands have the following arguments:

Part	Description
menu	Stringstring_definition containing the name of the menu in which the command resides.
menuitem	Stringstring_definition containing the name of the specific command within the menu.
<i>param</i>	Stringstring_definition containing additional information required by the command.

Settings

The settings for param vary depending on the specific menu command. Most menu commands use no param and ignore it if present. The following commands accept a value for param, and suppress the appropriate dialog box if it is present.

Command	Param
File Open	Name of the file to open.

Command	Param
File Save	Name of the file to save.
Record Go to	Index of the destination record within the current list. list.current_list_definition
Search Sub-structure	ON or OFF, corresponding to the checked/unchecked state of the menu.
Search Similarity	ON or OFF, corresponding to the checked/unchecked state of the menu.
Search Match Stereo	ON or OFF, corresponding to the checked/unchecked state of the menu

Remarks

Menu commands need not be spelled out in full; you can enter just enough letters to identify the menu option. For example, to execute the Record next command, you can enter “rec n”.

For File Open, if the file to be launched is in ChemBioFinder’s standard search path, only the actual file name (CS_DATA) is required. Otherwise, the complete file path (c:\some\long\path\CS_DATA.CFX) is necessary.

For File Save, if the specified file does not exist, it will be created; otherwise it will be overwritten.

For File Save, if the path is omitted from the file name, the file will be saved in the current directory.

Example 1

LOOP Example

This example uses the **LOOP** command and the Menu Command **Record Next** to display all of the records in the current list.

The **PAUSE** command keeps each record on screen long enough for people to see it.

```
looprec nextpause 2endloop
```

Example 2

Different Form Example

This example shows how you can use a script to open a second form with some of the same fields as the original form but, for example, a different layout. This same sort of procedure can be used to display related records if you do not want to use an explicit subform.

The APPEND, SETVAL, and GETDATA commands build a search string based on the contents of the link_field field on the original form. The Menu Command File Open opens the second form, or switches to it if it is already open.

Finally, the Menu Command Search Enter Query enters search mode, the PUTDATA command enters the search string generated earlier in the script, and the Menu Command Search Find executes the search.

```
append offsetval 1 "="append ongetdata 1 link_fieldappend offfile open d:\-
chemfndr\otherform.cfxsearch enter queryputdata link_field $v1search find
```

Example 3

Saving Compounds Example

This example uses the LOOP...ENDLOOP commands and the Menu Command Record Next to step through each record in the current list. For each record, it builds a unique file path based on the current \$INDEX using SETVAL and APPENDVAL commands. Finally, it saves each compound to disk with the WRITEMOL command. Since the file path ends in .cdx, this script saves all of the compounds in ChemDraw format.

```
append offloopsetval 1 c:\mydir\appendval 1 $indexappendval 1 .cdxwritemol $v1rec
nextendloop
```

Example 4

GOTO Example

This example allows the user to batch-enter several new compounds and input one field for each. It starts with a MSG command to tell the user what to do. Then it has a line starting with a colon (:) that serves as a label to start the loop. The INPUT command gets the name of the new entry from the user. The IF command tests to see if the user entered a blank string, and EXITS the script if so. After the Menu Command Record Add creates a new record, the PUTDATA command enters the user's string into the molname field.

Closing the loop, the GOTO command jumps execution back to the start label.

```
msg Enter new compound names. Enter a blank name when complete
:startinput lif $v1 = "" exitrecord addputdata molname $v1goto start
```

Box creation commands

```
DBOX coords [fieldname]
```

```
FRAME coords [text]
```

```
TEXT coords [text]
```

```
PICT coords [filename]
```

```
BUTTON coords [scriptname]
```

```
SUBFM coords
```

```
ARROWBOX coords
```

```
FRAMEDBOX coords [fieldname] [text]
```

To create a new box on the form:

1. Specify the box type with the appropriate keyword.
2. Type four integers giving the rectangular coordinates of the box after the keyword.
3. If desired, type a text string giving further information appropriate to the box type, as shown below:

Box Type	String
Data box	name of database field to be displayed in the box
Frame	static label for upper left
Text	static text
Picture	pathname of a Windows metafile
Button	pathname of a script file, or name as it appears on Scripts menu
Subform	<nothing>
Arrowbox	<nothing>
Framedbox	name of database field to be displayed in the box, and static label for upper left

If the text string is omitted, you can supply it later using SETFIELD or SETTEXT. For a button, you must use SETTEXT if you want its visible label to be different from the name of its script.

Box coordinates are specified in this order: left, top, right, bottom. Units are in pixels; the origin is at the upper left, with coordinates increasing from left to right and top to bottom, so that coordinates will range from left = 0 to right = 640 or 1024 or whatever fits your screen, and from top = 0 to bottom = 480 or 768 or similar.

When a new box is created, it adopts the current font. To specify a particular font for a box, use the FONT command prior to creating the box.

When specifying a script name, give one of the following:

- The complete pathname of the script file.
- A simple name without extension, if the script is stored in the standard scripts subdirectory with the standard file extension. If the script name appears on the Scripts menu, you can use the name from the menu.

The table below shows examples of Box Creation Commands:

COMMAND	ACTION
Frame 10 10 300 60 Molecu-	create frame with label at

COMMAND	ACTION
Iar Formula	upper left
DBOX 20 20 290 50 for- mula	create data box for formula in above frame
TEXT 100 100 300 150 "A Label"	create static text string
PICT 500 300 600 400 C:/LOGO.WMF	put specified picture at lower right
BUTTON 500 10 550 40 DEMO	create button at upper right to run script DEMO.CFS

CAL commands DBOX syntax

DBOX creates a new data box on the active form.

```
DBOX left top right bottom [fieldname]
```

The DBOX command has the following arguments:

Part	Description
left	Integer value representing the distance, in pixels, between the left edge of the box and the left edge of the active form.
top	Integer value representing the distance, in pixels, between the top edge of the box and the top edge of the active form.
right	Integer value representing the distance, in pixels, between the right edge of the box and the left edge of

Part	Description
	the active form.
bottom	Integer value representing the distance, in pixels, between the bottom edge of the box and the top edge of the active form.
<i>fieldname</i>	Database field name definition that will provide the data that is displayed in the data box.

Remarks

When a new box is created, it adopts the current font settings. These settings can be controlled with the FONT command.

You can assign a fieldname even if there is no data source for the form. If a data source is assigned at a later point and a field within that data source has the same name as assigned to the data box, the data box will automatically display the contents of that field.

After a box is created, its field name can be changed with the SETFIELD command.

DBOX and FRAME example

This example uses the DBOX command to create a short, wide data box near the top left of the form, and display the Formula field within it.

It then uses the FRAME command to label the data box with a slightly-larger frame.

```
DBOX 10 20 210 50 FormulaFRAME 5 5 215 55 Formula
```

CAL commands FRAME syntax

Creates a new frame on the active form.

```
FRAME left top right bottom [text]
```

The FRAME command has the following arguments:

Part	Description
left	Integer value representing the distance, in pixels, between the left edge of the frame and the left edge of

Part	Description
	the active form.
top	Integer value representing the distance, in pixels, between the top edge of the frame and the top edge of the active form.
right	Integer value representing the distance, in pixels, between the right edge of the frame and the left edge of the active form.
bottom	Integer value representing the distance, in pixels, between the bottom edge of the frame and the top edge of the active form.
<i>text</i>	String displayed as a label on the top left corner of the frame.

Remarks

When a new frame is created, it adopts the current font settings. These settings can be controlled with the FONT command.

After a frame is created, its label can be changed with the SETTEXT command.

CAL commands TEXT syntax

Creates a new static text box on the active form.

```
TEXT left top right bottom [text]
```

The TEXT command has the following arguments:

Part	Description
left	Integer value representing the distance, in pixels,

Part	Description
	between the left edge of the text box and the left edge of the active form.
top	Integer value representing the distance, in pixels, between the top edge of the text box and the top edge of the active form.
right	Integer value representing the distance, in pixels, between the right edge of the text box and the left edge of the active form.
bottom	Integer value representing the distance, in pixels, between the bottom edge of the text box and the top edge of the active form.
<i>text</i>	String displayed as the static text.

Remarks

When a new static text box is created, it adopts the current font settings. These settings can be controlled with the FONT command.

After a static text box is created, its contents can be changed with the SETTEXT command.

Text will wrap to fit within the dimensions specified. If all text cannot be displayed within the bounds of the text box, only the first part will be displayed. However, all text is still present and will appear if the box is resized to larger dimensions.

FONT and TEXT example

This example uses the FONT command to set the font to be large, bold, and bright (24-point Arial bold, in red), then uses the TEXT command to create a short, wide data box near the top middle of the form.

```
FONT Arial 24 1 255 0 0TEXT 200 40 450 70 My Compound Database
```

CAL commands PICT syntax

Creates a new picture box on the active form.

```
PICT left top right bottom [filename]
```

The PICT command has the following arguments:

Part	Description
left	Integer value representing the distance, in pixels, between the left edge of the box and the left edge of the active form.
top	Integer value representing the distance, in pixels, between the top edge of the box and the top edge of the active form.
right	Integer value representing the distance, in pixels, between the right edge of the box and the left edge of the active form.
bottom	Integer value representing the distance, in pixels, between the bottom edge of the box and the top edge of the active form.
<i>filename</i>	Name of the Windows metafile to display in the picture box.

Remarks

If the Windows metafile is in ChemBioFinder's standard search path, only the actual file name and extension (BUTTERFLY.WMF) is required. Otherwise, the complete file path (c:\some\long\path\BUTTERFLY.WMF) is necessary.

PICT example

This example uses the PICT command to create a large picture field toward the right of the form, and display in it the BUTTRFLY.WMF picture in the ChemBioFinder System directory.

```
PICT 300 100 500 300 butterfly.wmf
```

CAL commands BUTTON syntax

Creates a new button on the active form.

```
BUTTON left top right bottom [filename]
```

The BUTTON command has the following arguments:

Part	Description
left	Integer value representing the distance, in pixels, between the left edge of the button and the left edge of the active form.
top	Integer value representing the distance, in pixels, between the top edge of the button and the top edge of the active form.
right	Integer value representing the distance, in pixels, between the right edge of the button and the left edge of the active form.
bottom	Integer value representing the distance, in pixels, between the bottom edge of the button and the top edge of the active form.
<i>filename</i>	Name of the CAL script file to display in the button

Remarks

If the CAL script file is in ChemBioFinder's standard search path, only the actual file name (BROWSE) is required. Otherwise, the complete file path (c:\some\long\path\BROWSE.CFS) is necessary.

When a new button is created, it adopts the current font settings. These settings can be controlled with the FONT command.

By default, the name of the CAL script file (for example, BROWSE) is shown on the button. This can be changed with the SETTEXT command.

BUTTON example

This example uses the BUTTON command to create a button near the bottom of the form, and assign to it the REGISTER.CFS script in the same directory as the form.

```
BUTTON 50 300 200 330 Register
```

CAL commands SUBFM syntax

Creates a new subform on the active form.

```
SUBFM left top right bottom
```

The SUBFM command has the following arguments:

Part	Description
left	Integer value representing the distance, in pixels, between the left edge of the subform databox and the left edge of the active form.
top	Integer value representing the distance, in pixels, between the top edge of the subform databox and the top edge of the active form.
right	Integer value representing the distance, in pixels, between the right edge of the subform databox and the left edge of the active form.
bottom	Integer value representing the distance, in pixels, between the bottom edge of the subform databox and the top edge of the active form.

Remarks

It is not currently possible to set the data source or link fields of a subform using CAL.

SUBFM example

This example uses the SUBFM command to create a large subform in the middle of the form.

```
SUBFM 50 50 600 300
```

Box manipulation commands

```
FORMEDIT ON|OFF
```

```
SELECT box
```

```
UNSELECT box
```

```
DELBOX box
```

```
SETFIELD box fieldname
```

```
SETTEXT box text
```

```
ACTIVATE subform
```

```
DEACTIVATE
```

```
SCALE
```

```
SCALE_TO_FIT
```

To work with a box or subform on the form, you specify a keyword followed by a box identifier. A box is identified in one of two ways:

- By point—Give the coordinates of a point anywhere within the box. When coordinates fall in more than one box, the most recently created box is used.
- By text—The text you enter to identify a box may be:
 - The field name, if the box contains data from a database.
 - The text label, if the box contains static text, such as frames or text labels.
 - The box name assigned in Box Properties.
 - For subform boxes, the table name.

In all cases, you need not spell out the entire name, just enough to be unique.

Often, more than one box matches the text you enter, for example a box displaying field “molweight” within a frame labelled “MolWeight.” In this case, preference is given to the box connected to the database.

FORMEDIT controls the behavior of the SELECT command. With FORMEDIT ON, SELECT affects boxes. With FORMEDIT OFF, SELECT affects the contents of those boxes.

SELECT and UNSELECT choose boxes to be modified by editing operations. DELBOX is a shortcut for FORMEDIT ON, followed by SELECT, followed by Edit Clear. SETFIELD connects a database field to an existing box; for a button, it attaches a script name (see above regarding script names). SETTEXT attaches a text string to a box or

replaces the one currently attached; see box creation commands (above) for a list of what the text strings mean for various box types.

ACTIVATE and DEACTIVATE are the same as SELECT and UNSELECT, but work on subforms. Following an ACTIVATE command, all subsequent commands apply to the active subform and any boxes it contains. The DEACTIVATE command is required to once again refer to the main form and its contents.

SCALE scales all items in the current form by the percentage factor you specify. Use this to fit your forms to different size screens. See SCALE_TO_FIT below.

SCALE_TO_FIT scales all elements in the form such that the string you specify will fit in the form at the currently selected font size. See FONT below.

The table below shows examples of Box Manipulation Commands

:

Command	Action
SELECT 11 11	select formula frame created above
SELECT formu	select formula data box created above
SETFIELD 21 21 molweight	change field for formula data box to molweight

CAL commands DELBOX syntax

Deletes a data box, frame, static text box, picture box, button, or subform the active form.

DELBOX box

The DELBOX command has the following argument:

Part	Description
box	Box identifier indicating the box to delete.

Remarks

The DELBOX command is a shortcut for FORMEDIT ON, followed by SELECT, followed by the Menu Command Edit Clear.

The DELBOX command deletes only the box, not its actual contents. For example, deleting a button does not delete the CAL script file from your disk.

CAL commands SETFIELD syntax

Connects a database field to an existing data box or a CAL script name to an existing button, or replaces one already connected.

```
SETFIELD box fieldname
```

The SETFIELD command has the following arguments:

Part	Description
box	Box indicating the target.
fieldname	String containing the database field name (if the target is a data box) or CAL script name (if the target is a button) to assign to the target.

CAL commands SETTEXT syntax

Attaches a text string to a frame, text box, or button, or replaces one already attached.

```
SETTEXT box text
```

The SETTEXT command has the following arguments:

Part	Description
box	Box identifier indicating the target.
text	String to attach to the target.

Depending on the box type, the text is attached to different places:

Box type	Text location
data box	Within the text box, wrapped to fit.
frame	On the top border, aligned to the left.
text box	Within the text box, wrapped to fit.
button	Centered within the button.


CAL commands ACTIVATE syntax

Marks a subform on the active form to become the new active form. It is equivalent to clicking on the subform with the pointer tool.

`ACTIVATE subform`

The ACTIVATE command has the following argument:

Part	Description
subform	Box identifier indicating the subform to activate.

 **Note:** This command will change the active form.

Remarks

The ACTIVATE command acts only on subforms. The similar command SELECT is used for data boxes, frames, text boxes, buttons, and picture boxes.

CAL commands FORMEDIT syntax

Determines whether the active form is in form edit mode or data edit mode.

`FORMEDIT value`

The FORMEDIT command has the following settings:

Setting	Description
ON	Form is set to form edit mode.
OFF	Form is set to data edit mode.

Remarks

When the form is in form edit mode, the SELECT command affects data boxes themselves. When the form is in data edit mode, the SELECT command affects the data within the data boxes.

CAL commands SELECT syntax

Marks a box on the active form `active_form_definition` for further manipulation. It is equivalent to clicking in the box with the pointer tool.

`SELECT box`

The SELECT command has the following argument:

Part	Description
box	Box identifier indicating the box to select.

Remarks

The SELECT command acts only on data boxes, frames, text boxes, buttons, and picture boxes. The similar command ACTIVATE is used for subforms.

The behavior of the SELECT command is modified by the FORMEDIT command. When the form is in form edit mode, the SELECT command affects data boxes themselves. When the form is in data edit mode, the SELECT command affects the data within the data boxes.

CAL commands UNSELECT syntax

Reverses the action of the SELECT command, making sure that no boxes are marked for further manipulations. It is equivalent to clicking in the form background with the pointer tool.


UNSELECT

The UNSELECT command takes no arguments.

Reverses the action of the ACTIVATE command, making sure that the main form is the active form. It is equivalent to clicking in the main form background with the pointer tool.

UNSELECT

The UNSELECT command takes no arguments.

 **Note:** This command will change the active form.

Program execution commands

`CALL scriptname`

DOS `doscommand`

EXEC `wincommand`

EXEC_BUTTON `boxname`

DDE `ddecommand`

LAUNCH `filename`

OUTPUT_STR `string`

Program execution commands call up external processes and pass data to them. All require an argument, which may need multiple components.

CALL executes a specified CAL script. If the call is made from within another script, when the called script finishes executing it returns control to the caller. CALL must be followed by a scriptname, as described above.

DOS executes a DOS command line. Follow the DOS keyword with any string you might type at a DOS command prompt. If you do not type any arguments after the DOS command, you get an interactive command prompt window. You can use this feature to manipulate files, execute programs or batch files, get directory listings, format disks, etc. During execution of the command, a command prompt window appears on the screen; when finished, the window goes away and control returns to the calling script.

EXEC starts a Windows program and optionally passes it command-line arguments. Follow the EXEC keyword with any string you might use in the Program Manager's File Run command. This command starts a program, but does not return from it. To return to ChemBioFinder, you need to use Task Manager or click in the ChemBioFinder frame window.

EXEC_BUTTON transfers execution to the script connected to a particular button, where boxname is the button's box name in the Box Properties dialog box, usually the same as the text on the button.

DDE sends a Dynamic Data Exchange message to a specified application. Follow the DDE keyword with three arguments:

- The service name—usually the name of the recipient application.
- The topic name—a string recognized by the recipient, identifying the nature of the message. It is typically SYSTEM.
- The command—an instruction to the service indicating what you want it to do.

OUTPUT_STR gives CAL access to the output window used by Python, where string writes a string into the output window, followed by a new line. The command brings the window up on the first call, if necessary. The string need not be quoted, even if it contains multiple words.

Details of these components depend on the service you are addressing.

The table below shows examples of Program Execution Commands

:

Command	Action
CALL myscript	execute script file located in scripts directory
DOS erase junk- file.dat	execute DOS command to delete a file
EXEC note- pad myfile.txt	execute Windows pro- gram
DDE CHEM3D CFWIN "open benz.mol"	send DDE message to qualified service

CAL commands CALL syntax

Executes the specified CAL script

```
CALL scriptname
```

The CALL command has the following argument:

Part	Description
<i>scriptname</i>	String containing the name of the CAL script file to execute.

Remarks

If the CAL script file is in ChemBioFinder's standard search path, only the actual file name (BROWSE) is required. Otherwise, the complete file path (c:\some\long\path\BROWSE.CFS) is necessary.

If the CALL command is executed from within another script (as opposed to from the Command Line), the calling script will continue to execute with its next command after the called script finishes executing.

A CALLED script inherits all user variables from its calling script. The calling script also inherits the user variables back from the called script when it has completed executing.

CAL commands DOS syntax

Executes a standard DOS command line

```
DOS doscommand
```

The DOS command has the following argument:

Part	Description
<i>doscommand</i>	Any string that might be typed at a DOS command prompt.

Remarks

This command may be used to manipulate files, execute programs or batch files, get directory listings, format disks, etc.

During execution, a DOS command prompt window appears on the screen; when finished, the window goes away and control returns to the calling script.

CAL commands EXEC syntax

Starts a Windows program and optionally passes it command-line arguments.

```
EXEC filename [commandargs]
```

The EXEC command has the following arguments:

Part	Description
commandargs	String containing arguments to be passed to that application.
<i>filename</i>	String containing the file name of any Windows application

Remarks

If the file to be launched is in ChemBioFinder's standard search path, only the actual file name (CHEMDRAW) is required. Otherwise, the complete file path (c:\some\long\path\CHEMDRAW.EXE) is necessary.

The EXEC command starts a program, but does not return from it; in order to return to ChemBioFinder, you need to use the Task Manager or click in the ChemBioFinder frame window.

EXEC example

This example uses the EXEC command to start Microsoft Excel.

The first line starts with an asterisk (*) and so is not executed. It serves only to make the example more readable for the user.

```
* RUNEXCEL.CFS -- script to start ExcelEXEC "c:\msoffice\excel\excel.exe"
```

CAL commands DDE syntax

Sends a Dynamic Data Exchange message to a specified application

```
DDE [timeout] servicename topic command
```

The DDE command has the following arguments:

Part	Description
<i>timeout</i>	An optional time in seconds to try establishing the connection
servicename	String containing the identifier of the recipient application; often but not always the application name.
topic	String identifying the nature

Part	Description
	of the message
command	String containing instructions to recipient application indicating what you want it to do.

Remarks

DDE commands are highly specific to each application. Please consult the documentation for the desired recipient application for information on service names, topic names, and commands it supports.

The DDE command is always sent as an EXECUTE instruction.

Communicating with Excel example

This example uses the WRITETEXT command to write the current list to a file, then sends two DDE commands to open that file in Microsoft Excel and resize the columns to a nice size. Finally, it uses the DOS command to delete the temporary file.

The first line starts with an asterisk (*) and so is not executed. It serves only to make the example more readable for the user.

```
* TOEXCEL.CFS -- script to send data to Excel
WRITETEXT C:\DATA.TMPDDE Excel System
[OPEN("C:\DATA.TMP")]DDE Excel System [COLUMN.WIDTH(1,"C1:C4",,3,1)]DOS erase
C:\DATA.TMP
```

General commands

MSG *message*

STATMSG *message*

TIMEDMSG [n] <*msg*>W

FONT *fontname* [size [style r g b]]

SEEPTAB [*decisecs*]

QUIETCLOSE

HIDETABLE

SET

CLEAN [*cleanup, tidy, redraw, denovo*]

HELP

GET [*v*] *section number, itemnumber*

SOUND *filename.wav*

SYSMETRIC [*v*] *index*

MSG displays a message box with the specified text, and waits for the user to click.

STATMSG displays information on the status line.

TIMEDMSG displays a message box similar to the MSG command, but the message disappears automatically after a specified number of seconds. For example: TIMEDMSG 10 "This message will disappear automatically after 10 seconds."

FONT changes the current font. Any boxes created subsequently adopt this font. Follow the keyword with a name from your Windows font list, followed optionally by:

- Font size in points (default is 8).
- Style from the list below (default is 0, or plain text).
- Color; three values for red, green, and blue, each ranging from 0 to 255 (default is 0,0,0, or black).

Font styles are sums of bold (1), italic (2), and underline (4). Thus a style value of 1 means bold, 3 means bold+italic, 6 means italic+underline, and so on.

SEETAB briefly displays the periodic table window, then closes it. The window will be displayed for the amount of time specified (1 decisecond = 0.1 second). If the duration is omitted, it will be displayed for four seconds.

QUITECLOSE closes the active form. This command is similar to FILE CLOSE, but automatically discards any changes instead of prompting the user what to do.

HIDETABLE closes the Table View window for the current form. If a subform is active, it returns the subform to form view.

SET allows you to modify basic ChemBioFinder settings, including most of those found in the Preferences dialog box and several that cannot be accessed in any other way. See the online help for a complete listing of parameters that can be SET.

CLEAN will attempt to standardize the bond lengths and angles of the current molecule. It is the same as opening the molecule in ChemDraw, and selecting the Clean Up Structure command. The arguments are optional.

HELP displays a list of all valid CAL commands. The valid parameter types are listed for each command, as is a brief description.

SET/GET functions store data in the Windows registry under HKEY_CURRENT_USER\Software\CambridgeSoft\ChemFinder\14.0\General\Options

SOUND plays a .wav file you specify.

SYSMETRIC retrieves the specified system metric into a variable.

The table below shows examples of General Commands:

Command	Action
MSG "Click OK to con- tinue"	display message, wait for click
FONT Times 12 1 255 0 0	change font to red 12- point Times bold
GET V1,4,5	place the value of item number 5 in section 4 into the variable V1
SOUND beep- .wav	play the sound in the files beep.wav

CAL commands MSG syntax

Displays a message in a dialog box, and waits for the user to dismiss the dialog box.

MSG message

The MSG command has the following argument:

Part	Description
message	String displayed as the mes- sage in the dialog box.

CAL commands FONT syntax

Changes the current font, size, style, and color used for all subsequent box creation.

FONT fontname [size [style [r g b]]]

The FONT command has the following arguments:

Part	Description
fontname	String containing the name of a font from your Windows font list
size	String containing the size of the font, in points.
style	String containing the style in

Part	Description
	which to display the font, as described in Settings.
r	String containing the red component of the font color, in the range from 0 to 255.
g	String containing the green component of the font color, in the range from 0 to 255.
b	String containing the blue component of the font color, in the range from 0 to 255.

Settings

The settings for style are:

Setting	Description
0	Plain text.
1	Bold
2	Italic
3	Bold and italic.
4	Underline
5	Bold and underline.
6	Italic and underline.
7	Bold and italic and underline

CAL commands SEEPTAB syntax

Displays the Periodic Table window, then closes it.

```
SEEPTAB [deciseCs]
```

The SEEPTAB command has the following argument:

Part	Description
decisecs	Amount of time to display the Periodic Table window before closing it.

Remarks

10 decisec = 1 second.

If the decisecs argument is omitted, the Periodic Table window will be displayed for 4 seconds.

CAL commands QUIETCLOSE syntax

Closes the topmost form immediately, without checking to see if it has been modified.

```
QUIETCLOSE
```

The QUIETCLOSE command takes no arguments.

Remarks

This command is similar to the Menu Command File Close. However, if the user has modified the form in any way, the File Close command will present an alert asking if the user wants to save changes, while the QUIETCLOSE command will not.

CAL commands HIDETABLE syntax

Closes the Table View window for the active form.`active_form_definition`

```
HIDETABLE
```

The HIDETABLE command takes no arguments.

Remarks

The Table View for the active form can be displayed with the Menu Command View Data Table.

If the active form is a subform, the HIDETABLE command will return it to form view.

CAL commands SET syntax

Changes the values of various preferences that control the overall behavior of ChemBioFinder.

```
SET section preference setting
```

The SET command has these arguments:

Part	Description
section	String describing the general category of preference.
preference	String naming the preference to change.
setting	String describing new value of preference.

Values

The values for section, preference, and value are:

Section	Preference	Description	Values
DISPLAY	ATOM_LABEL_COLOR	The color in which to display atom labels that cannot be interpreted by ChemBioFinder. The default is Purple.	color
DISPLAY	ATOM_NUMS	Whether or not to display atom numbers on each atom. The default is 0 (do not display)	0 or 1
DISPLAY	ATOM_SPECIAL_COLOR	The color of query atoms and other atom types not listed in the Periodic table. The default is Purple.	color
DISPLAY	BOND_NUMS	Whether or not to display bond numbers on each bond. The default is 0 (do not display)	0 or 1
DISPLAY	BOND_SPACING	Spacing between double bonds, as a percentage of bond length. The default is 12 (%).	0 ... 100
DISPLAY	FONT_ADJUSTMENT	A correction to the font size of atom labels based on a structure's average bond length. More specifically, the font size of atom labels	-50 ... 80

Section	Preference	Description	Values
		is: (nominal value + (FONT_ADJUSTMENT/100) * average bond length). The default is 20 (%).	
DISPLAY	FORM_COLOR	The background color of the form. This parameter can be set manually in the Color tab of the Preferences dialog box. The default is Grey.	color
DISPLAY	FRAME_PICTURES	Whether or not to display borders around pictures. This parameter can be set manually in the Display tab of the Preferences dialog box. The default is 1 (display).	0 or 1
DISPLAY	GRID_SPACING	The spacing, in pixels, between hash marks on the grid. This parameter can be set manually in the Display tab of the Preferences dialog box. The default is 10 (pixels).	0 ... 100
DISPLAY	MAX_BND_LEN	The scaling factor by which the maximum bond length is modified. This value is only applicable when the Fit to box option is deselected. The default is 100 (%).	10 ... 500

Section	Preference	Description	Values
DISPLAY	MOL_HIGHLIGHT_COLOR	The color in which sub-structure search matches are displayed. This parameter can be set manually in the Color tab of the Preferences dialog box. The default is Red.	color
DISPLAY	NET_OUTSIDE	Whether or not to move bracketed text to the outside, where possible, in the display of net reaction changes in molecular formulas and molecular weights. The default is 1 (move to outside).	0 or 1
DISPLAY	NET_RXN_CHANGE	Whether or not to display net reaction changes in brackets in molecular formulas and molecular weights. The default is 1 (display changes).	0 or 1
DISPLAY	PICTURE_BACKGROUND	The background color of Picture boxes. This parameter can be set manually in the Color tab of the Preferences dialog box. The default is White.	color
DISPLAY	QUERY_BOND_COLOR	The color in which bonds with query properties should	color

Section	Preference	Description	Values
		<p>be displayed. This parameter can be set manually in the Color tab of the Preferences dialog box.</p> <p>The default is Green.</p>	
DISPLAY	RXN_CTR_COLOR	<p>The color in which reaction centers should be highlighted, if highlighting is turned on. This parameter can be set manually in the Color tab of the Preferences dialog box.</p> <p>The default is 33023 (purplish).</p>	color
DISPLAY	SHOW_AAMAP	<p>Whether or not to highlight matching mapped atoms in a reaction in color. This parameter can be set manually in the Display tab of the Preferences dialog box.</p> <p>The default is 0 (do not highlight).</p>	0 or 1
DISPLAY	SHOW_H_CHARGED_ATOMS	<p>Whether or not to display implicit hydrogens on all atoms bearing a charge or radical.</p> <p>The default is 1 (display).</p>	0 or 1
DISPLAY	SHOW_H_ON_HET	<p>Whether or not to display implicit hydrogens on all heteroatoms. This parameter can be set manually in the</p>	0 or 1

Section	Preference	Description	Values
		<p>Display tab of the Preferences dialog box.</p> <p>The default is 1 (display).</p>	
DISPLAY	SHOW_H_TERM_C	<p>Whether or not to display implicit hydrogens on all terminal carbons. This parameter can be set manually in the Display tab of the Preferences dialog box.</p> <p>The default is 0 (do not display).</p>	0 or 1
DISPLAY	SHOW_RXN_CENTER	<p>Whether or not to highlight reaction centers. This parameter can be set manually in the Display tab of the Preferences dialog box.</p> <p>The default is 0 (do not highlight).</p>	0 or 1
OPTIONS	CFW_METAFILES	<p>Whether to generate pictures using ChemBioFinder or using ChemDraw when copying a structure to the clipboard.</p> <p>The default is 0 (use ChemDraw to generate).</p>	0 or 1
OPTIONS	DDE_TIMEOUT	<p>Number of seconds to wait for a response to a DDE transaction.</p> <p>The default is 30 (seconds).</p>	0 ... 60
OPTIONS	EXPLICIT_METAL_H	Whether or not to require	0 or 1

Section	Preference	Description	Values
		<p>that metals have hydrogens drawn implicitly. If this parameter is off, ChemBioFinder will automatically add hydrogens to fill each metal's minimum valence. This parameter can be set manually in the General tab of the Preferences dialog box.</p> <p>The default is 1 (require explicit drawing).</p>	
OPTIONS	EXPORT_HEADER	<p>ASCII value of the delimiter to use when exporting delimited text files. This parameter can be set manually in the General tab of the Preferences dialog box.</p> <p>The default is 9 (tab).</p>	0 ... 255
OPTIONS	LOAD_LAST_FORM_ON_STARTUP	<p>Whether or not to load the last-active form file the next time ChemBioFinder is launched.</p> <p>The default is 0 (do not load).</p>	0 or 1
OPTIONS	NON_INTERACTIVE	<p>Whether or not to suppress non-critical dialogs. Principally useful when ChemBioFinder is being used unattended, such as over the WWW. It is not recommended to have this option on when using</p>	0 or 1

Section	Preference	Description	Values
		ChemBioFinder directly. The default is 0 (interactive mode).	
OPTIONS	VALIDATE_VALENCE	Whether or not to provide a warning when attempting to enter a substance containing an element in a non-standard valence. This parameter can be set manually in the General tab of the Preferences dialog box. The default is 0 (do not warn).	0 or 1
PRINT_OPTIONS	PRINT_BORDER	Whether or not to display a border around forms when printing The default is 1 (print a border).	0 or 1
PRINT_OPTIONS	PRINT_FOOTER	The contents of the footer to be printed at the bottom of all pages containing forms. Any text is valid, as well as the following special codes: %D current date %F name of the CFX file %P page number %T current time The default is “ChemBioFinder %D %T Record %P”	string

Section	Preference	Description	Values
PRINT_OPTIONS	PRINT_HEADER	<p>The contents of the header to be printed at the top of all pages containing forms. Any text is valid, as well as the same special codes recognized for the footer.</p> <p>The default is "%F" (file name)</p>	string
SEARCHING	EXACT	<p>Whether or not to allow a substructure match when searching. This parameter can be set manually in the Search tab of the Preferences dialog box.</p> <p>The default is 0 (substructure).</p>	0 or 1
SEARCHING	FIND_CHARGED_CARBON	<p>Whether or not to allow uncharged carbon atoms to match charged ones when searching. This parameter can be set manually in the Search tab of the Preferences dialog box.</p> <p>The default is 1 (allow the match).</p>	0 or 1
SEARCHING	FIND_CHARGED_HET	<p>Whether or not to allow uncharged heteroatoms to match charged ones when searching. This parameter can be set manually in the Search tab of the Preferences dialog box.</p>	0 or 1

Section	Preference	Description	Values
		The default is 1 (allow the match).	
SEARCHING	MATCH_STEREO_DB	Whether or not to match cis/trans stereochemistry around double bonds. This parameter can be set manually in the General tab of the Preferences dialog box. The default is 1 (match stereo).	0 or 1
SEARCHING	MATCH_STEREO_TET	Whether or not to match tetrahedral stereochemistry. This parameter can be set manually in the Search tab of the Preferences dialog box. The default is 1 (match stereo).	0 or 1
SEARCHING	MAX_HITS	Maximum number of hits to be returned by a search. The default is 2500 (hits).	1...
SEARCHING	SIMILARITY	Whether or not to allow a similarity match when searching. This parameter can be set manually in the Search tab of the Preferences dialog box. The default is 0 (require exact match).	0 or 1
SEARCHING	SRCH_TANIMOTO_PCT	Cutoff percentage below which two compounds are	0 ... 100

Section	Preference	Description	Values
		not considered similar. This parameter can be set manually in the Search tab of the Preferences dialog box. The default is 90 (%).	
SEARCHING	SRCH_USE_RXN_CTRS	Whether or not to require that query substructures have at least one atom or bond in the reaction center of the target. This parameter can be set manually in the Search tab of the Preferences dialog box. The default is 1 (require in center).	0 or 1
SEARCHING	STOP_AFTER_MAX_HITS	Whether or not to stop searching after the maximum number of hits have been found. If this preference is turned off, ChemBioFinder will report the total number of hits, but will still display only the maximum number. The default is 1 (stop).	0 or 1

CAL commands CLEAN syntax

Attempts to standardize the bond lengths and angles of the current molecule.

```
CLEAN [param]
```

The CLEAN command takes the following optional arguments:

Part	Description
cleanup	(default) Progressively stronger drawing modes, starting with Tidy, until something changes.
tidy	Minimum redrawing, if any.
redraw	License to redesign rings
denovo	Complete redesign from scratch.

Remarks

This command is equivalent to opening the structure in ChemDraw, selecting the Clean Up Structure command, and returning to ChemBioFinder.

CAL commands HELP syntax

Displays a list of all valid CAL commands. The valid parameter types are listed for each command, as is a brief description.

HELP

Remarks

This command is not of much use during the execution of a CAL script, but can be a quick syntax aid when entered into the ChemBioFinder Command Line.

File commands

READMOL *filename*

READPICT *boxname filename*

WRITEMOL *filename*

WRITEPICT *boxname filename*

WRITETEXT *filename*

READMOL and WRITEMOL operate on the current molecule. These commands will work only if there is at least one structure-related box (structure, formula, or molweight) on the form, and you are positioned to a valid entry in the database. READMOL reads a specified structure file and replaces the current molecule in the form. WRITEMOL saves the current molecule to a specified structure file; if the file doesn't exist, it is created, otherwise it is overwritten. For

both READMOL and WRITEMOL, the format of the file is determined by its extension; for example, benzene.cdx is a ChemDraw file.

READPICT and WRITEPICT transfer pictures into or out of picture boxes, where boxname is the name of the picture box on the form and filename is a pathname to a graphics file, with extension .bmp, .jpg, .gif, .png, .emf, or .wmf.

WRITETEXT exports the current list as delimited text with whatever delimiter is specified in the Preferences dialog box.

The table below shows examples of File Commands

Command	Action
READMOL benz.mol	read specified file to become current molecule
WRITEMOL saved.mol	write current molecule to specified file
WRITETEXT hits.txt	export current list to a text file

OPENDB [R/E/RE] dbname

CRETABLE tablename

DELTABLE tablename

SELTABLE tablename

CREFIELD fieldname

DELFIELD fieldname

SORT [D] fieldname

Oracle database commands:

SQL <SQLStatement>

SQLSELECT [v] <SQLSelectStatement>

OPENDB opens a standard molecule database. Specify a pathname to the MDB file, or a name from the ODBC Data sources list. When connected to an Oracle database, OPENDB takes a table name instead of a database name. You can specify the mode of database opening:

- R—read-only
- E—exclusive
- RE—read-only and exclusive

CRETABLE creates tables in the current database.

DELTABLE deletes tables in the current database.

SELTABLE selects a table from the current database and makes it the current working table for subsequent field actions.

CREFIELD creates fields in the current table in the current database. You can specify the field type and width in the CREFIELD command. The default text field is 50 characters wide.

DELFIELD deletes fields in the current table in the current database.

SORT sorts the database contents on the specified field. The default sorts the database in ascending order. You can specify descending order with D.

SQL allows SQL to be passed directly from CAL. *SQLStatement* is a SQL statement that returns no results. It is therefore useful for data manipulation operations. Available only when connected to an Oracle database.

SQLSELECT returns a one-row recordset. It is useful for calculations or lookups. *v* is the name or number of a variable to receive the results of a Select statement; *SelectStatement* is a SQL Select statement designed to retrieve a single result. If *v* is omitted, results go in V1. Available only when connected to an Oracle database.

If you omit the password from the connection string, and you will be prompted for it when running the command.

The table below shows examples of Database Commands

Command	Action
CRETABLE MyTable	creates a new table
DELTABLE MyTable	deletes a new table

CAL commands OPENDB syntax

Opens a database and assigns it as the Data Source of the active form.

OPENDB dbname

The OPENDB command has the following argument:

Part	Description
dbname	String containing the file name of a database, which could be a *.MDB file or a name from the ODBC Data Sources list.

Remarks

If the file to be launched is in ChemBioFinder's standard search path, only the actual file name (CS_DATA) is required. Otherwise, the complete file path (c:\some\long\path\CS_DATA.MDB) is necessary.

OPENDB example

This example uses the OPENDB command to assign a new data source to the current form.

```
OPENDB c:\data\mydb.mdb
```

Variable commands

```
SETVAL [v] text
```

```
READVAL [v] filename
```

```
WRITEVAL [v] filename
```

```
INPUT [v] [prompt]
```

```
PASSWORD [v] [default]
```

```
GETDATA [v] box
```

```
PUTDATA box text
```

```
FROM_SMILES [str]
```

```
TO_SMILES [v]
```

```
APPEND ON|OFF
```

```
APPENDVAL [v] [text]
```

```
INCREMENT [v]
```

```
DECREMENT [v]
```

```
LET v v op v
```

```
SUBSTR [v] <str> <from> [<to>]
```

CAL has nine variables called V1, V2, ..., V9. These are temporary storage locations you can use to move information between form boxes and other data sources.

Variable names can be substituted in any CAL command. Any item shown in italics in a command description may be replaced by the name of a variable, prefixed with a dollar sign.

For example:

MSG message may be either:

```
MSG "An explicit message"
```

—a message explicitly coded into the script—or

```
MSG $V2
```

—a message taken from variable 2 at run time

You may concatenate variables, and variables and literals together in CAL commands. You must add both a leading and trailing dollar sign to the variable name.

For example:

```
MSG$V2$$V1$
```

Displays a message box with the value of V2 immediately followed by the value of V1.

Most commands for manipulating variables take an optional variable number as the first argument. If the number is omitted, it is assumed to be 1.

SETVAL puts the specified text into a variable. READVAL reads the contents of a text file into a variable. WRITEVAL copies the contents of a variable to a file.

INPUT displays a text input dialog box, accepts data from the user and stores it in a variable. An optional prompt string is displayed in the box when it appears.

PASSWORD is the same as INPUT, but displays all characters as asterisks.

GETDATA retrieves the contents of a form box (as long as it is not a structure) into a variable; see above for how to identify a box. PUTDATA copies specified text into a form box. Form names are case-sensitive. Be sure to use the same case as the form box name into which you want to store data.

TO_SMILES retrieves the contents of the current structure as a smiles string. FROM_SMILES generates a structure from a smiles string

APPEND determines whether data in a storage location is kept if other data is being moved there. This setting is also consulted when you do a File Export to SD or delimited text file.

APPEND ON causes any subsequent data movements to append new data to old; APPEND OFF causes old data to be overwritten. This applies to all variable commands except PUTDATA. The default at program startup is APPEND OFF.

APPENDVAL is a shortcut designed to make it easier to build string values. It is the same as setting APPEND ON, doing a SETVAL, and then restoring APPEND to its original state.

Tip: *The difference between APPEND ON and APPENDVAL is that APPENDVAL does not add a carriage return to the string being appended, whereas APPEND ON does. If you want to insert a carriage return at the end of the string, use APPEND ON.*

INCREMENT adds 1 to the value of a variable. Use INCREMENT in loops.

DECREMENT subtracts 1 from the value of a variable. Use DECREMENT in loops.

LET allows you to perform mathematical operations on variables that contain integer or real-number values. Only one operator per line is supported. It recognizes the following operators:

- + for addition
- - for subtraction
- * for multiplication
- / for division

SUBSTR extracts a substring from a variable. The extracted substring extends from a specified character to the end of the string, or from one specified character to another.

The table below shows examples of Variable Commands

Command	Action
SETVAL "some text"	store text in V1
SETVAL 2 "other text"	store text in V2
READVAL 9 tmpdata.txt	read text from file into V9
APPEND ON	turn on append mode
WRITEVAL output.txt	append contents of V1 to file
SETVAL 2 "append me"	append text to contents of V2
APPEND OFF	turn off append mode
INPUT 2 "enter new data"	prompt and get user input to V2
GETDATA 3 molname	copy data from molname box to V3
PUTDATA mol- name "ben- zene"	store given string in mol- name box
LET 3 \$V1 + \$V4	sets V3 equal to the value stored in V1 plus the value stored in V4
SUBSTR 2 abcdef 3 5	stores cde in V2

CAL commands SETVAL syntax

Assigns the specified string to a user variable.

```
SETVAL [variable] text
```

The SETVAL command has the following arguments:

Part	Description
variable	User variable to which the text is assigned.
text	String that is assigned to the variable

Remarks

If variable is omitted, the text will be assigned to V1.

CAL commands READVAL syntax

Reads the contents of a text file into a user variable. `user_variable_definition`

```
READVAL [v] filename
```

The READVAL command has the following arguments:

Part	Description
variable	User variable to which the text is assigned.
filename	String containing the file name from which to read the text.

Remarks

The complete file path (c:\some\long\path\TEMP.TXT) is necessary when reading files

If variable is omitted, the text will be assigned to V1.

If the file is greater than 32K in size, only the first 32K is read.

READVAL example

This example uses the READVAL command to assign the contents of a file to a variable. This might be the first step, for example, in writing the data to a database.

```
READVAL d:\reports\status.txt
```

CAL commands WRITEVAL syntax

Copies the contents of a user variable to a text file.

```
WRITEVAL [v] filename
```

The WRITEVAL command has the following arguments:

Part	Description
variable	User variable from which the text is obtained.
filename	String containing the file name in which to store the text.

Remarks

The complete file path (c:\some\long\path\TEMP.TXT) is necessary when saving files. If the specified file does not exist, it will be created; otherwise it will be overwritten.

WRITEVAL example

This example uses the WRITEVAL command to write the contents of a variable to a text file.

```
WRITEVAL 2 f:\formulas\mydata.txt
```

CAL commands INPUT syntax

Displays a text input dialog box, accepts data from the user, and stores it in a variable

```
INPUT [variable] [input]
```

The INPUT command has the following arguments:

Part	Description
variable	User variable to which the user's data is assigned.
input	String presented as the default input.

Remarks

If variable is omitted, the text will be assigned to V1.

CAL commands PASSWORD syntax

Displays a text input dialog box, accepts data from the user, and stores it in a variable. Each character entered by the user is displayed as an asterisk to conceal its identity.

```
PASSWORD [v] [default]
```

The PASSWORD command has the following arguments:

Part	Description
variable	User variable to which the user's data is assigned.
default	String presented as the default password.

Remarks

If variable is omitted, the text will be assigned to V1.

Password entry example

This example uses the MSG command to prompt the user for a password, then the PASSWORD command to retrieve it. After using the IF command to test whether the password is correct, it either EXITS or continues with the rest of the script.

```
msg Please enter the secret password...password lif $v1 <> "the secret password"
exit* ...etc.
```

CAL commands GETDATA syntax

Reads the contents of a data box into a user variable.

```
GETDATA [v] box
```

The GETDATA command has the following arguments:

Part	Description
variable	User variable to which the text is assigned.
box	Box identifier indicating the source of the text.

Remarks

GETDATA accepts only boxes that contain text or data that can be represented as text (including numbers, formulas, or dates). You cannot GETDATA from a data box containing a structure or a picture.

If variable is omitted, the text will be assigned to V1.

CAL commands PUTDATA syntax

Copies the contents of a user variable to a data box.

```
PUTDATA box text
```

The PUTDATA command has the following arguments:

Part	Description
box	Box identifier indicating the destination of the text.
text	String containing the text to store.

Remarks

If PUTDATA attempts to store text in a data box that contains a non-text data type (such as a number, date, or formula), the text will automatically be converted to the correct type.

CAL commands APPEND syntax

Determines whether data in a user variable `user_variable_definition` is kept if other data is being moved there.

`APPEND value`

The APPEND command has the following argument:

Part	Description
value	Value describing whether new data should be added to the end of existing data or whether it should overwrite the existing data.

Settings

The settings for value are:

Setting	Description
ON	New data is added to the end of existing data.
OFF	New data replaces existing data.

Different form example

This example shows how you can use a script to open a second form with some of the same fields as the original form but, for example, a different layout. This same sort of procedure can be used to display related records if you do not want to use an explicit subform.

The APPEND, SETVAL, and GETDATA commands build a search string based on the contents of the link_field field on the original form. The Menu Command File Open opens the second form, or switches to it if it is already open. Finally, the Menu Command Search Enter Query enters search mode, the PUTDATA command enters the search string generated earlier in the script, and the Menu Command Search Find executes the search.

```
append offsetval 1 "="append ongetdata 1 link_fielddappend offfile open d:\-
chemfndr\otherform.cfx search enter queryputdata link_field $v1search find
```

CAL commands APPENDVAL syntax

Concatenates a new string to the end of an existing user variable.

```
APPENDVAL [v] text
```

The APPENDVAL command has the following arguments:

Part	Description
variable	User variable to which the text is concatenated.
text	String that is concatenated to the variable.

Remarks

APPENDVAL is a shortcut designed to make it easier to build string values. It is the same as setting APPEND ON, doing a SETVAL, and then restoring APPEND to its original state.

If variable is omitted, the text will be assigned to V1.

Saving compounds example

This example uses the LOOP...ENDLOOP commands and the Menu Command Record Next to step through each record in the current list. For each record, it builds a unique file path based on the current \$INDEX using SETVAL and APPENDVAL commands. Finally, it saves each compound to disk with the WRITEMOL command. Since the file path ends in .cdx, this script saves all of the compounds in ChemDraw format.

```
append offloopsetval 1 c:\mydir\appendval 1 $indexappendval 1 .cdxwritemol $v1rec
nextendloop
```

Environment variables

\$INDEX

\$RECNO

\$NUM_RECS

\$DTBA_NAME_LONG

\$DTBA_NAME_SHORT


\$FORM_NAME_LONG

\$FORM_NAME_SHORT

\$CFW_DIR

\$FORM_DIR

\$SYSTEM_DIR

 **Note:** To view a complete list of variables, click Help on the Enter Cal Command dialog box.

ChemBioFinder provides several variables to allow you to retrieve information about the current ChemBioFinder environment. They cannot be set directly, but are modified as a consequence of other commands.

For example, the menu command:

```
Record Next
```

increments the current RECNO by 1.

The table below shows examples of Environment Variables

:

Command	Action
SETVAL 1 \$RECNO	V1 now equals the number of the current record
SETVAL V1 \$NUM_RECS	V1 now equals total number of records in the current list
SETVAL V1 \$DTBA_ NAME_ LONG	V1 now equals, for example "c:\ChemBioFinder- \mydb.mdb"
SETVAL V1 \$FORM_ NAME_ SHORT	V1 now equals, for example "mydb"
SETVAL V1 \$SYSTEM_ DIR	V1 now equals, for example, "c:\- chemfndr\system"

Script-only commands

**text*

:label

label:

GOTO *label*

IF *v1 op v2 stmt*

IF *v1 op v2*

<*stmt (s)*>

[ELSE

<*stmt (s)*>]

ENDIF

LOOP [*count*]

ENDLOOP

PAUSE [*dseconds*]

EXIT

INTERACTIVE *ON/OFF*

STEP *ON/OFF*

Script-only commands are useful only in script files, not interactive CAL.

An asterisk (*) is used to mark a line as a comment and not a command that should be executed.


Labels and GOTO give a mechanism for jumping backwards or forwards in a script. A label may be any text string starting or ending with a colon. It may appear anywhere in the script. (Prior to ChemOffice 2005, it had to appear prior to any GOTO that referenced it.)

IF is a standard conditional statement. It compares two values according to an operator, and then executes its final statement if the result is true. In the IF/ELSE format, it executes the ELSE statement when the IF statement is false.

Valid operators are:

=	<>	<	>
---	----	---	---

LOOP begins a section of code that will execute repeatedly, until ENDLOOP is reached. By default, the number of times the code executes is equal to the number of records in the current list, but you can provide a specific count in the LOOP statement.

 **Tip:** ENDLOOP can be used anywhere within a loop clause to mean “continue to the next iteration”. Example:
LOOP

```

RECORD NEXT
IF $HAS_MOL = 0
ENDLOOP
ENDIF
CLEAN
ENDLOOP

```



In this example, if a record has no structure, the program goes to the next record and skips the CLEAN.

PAUSE temporarily stops the script from executing for a specified number of tenths of a second. If no number is given, the duration of the pause is 2 seconds.

EXIT ends the script immediately.

INTERACTIVE turns interactive mode on or off.

STEP toggles single-step (debugging) mode on and off.

The table below shows examples of Script-Only Command.

Command	Action
start:	a label
IF \$V1 = "" MSG Empty value!	provide an error message if the variable V1 is empty
PAUSE 100	pause 10 seconds
GOTO start	go to specified label
LOOP 10	begin loop to repeat ten times
PAUSE	pause the default time: 2 seconds
ENDLOOP	return to LOOP statement until done

 **Note:** Script commands are subject to change. See the Readme for up-to-date information.

CAL commands * syntax

Marks a line as a comment, and not a command to be executed.

*[text]

The * command has the following argument

Part	Description
text	String containing the comment.

Remarks

The * character must be the first character in a line to mark it as a comment.

This command is useful only in a script file, and has no effect when entered on the Command Line.`command_line_definition`

CAL commands: syntax

Labels a line with the preceding or following text.

:label or label:

The **command has the following argument:

Part	Description
label	Text to be used as a label.

Remarks

This command is useful only in a script file, and has no effect when entered on the Command Line.

CAL commands GOTO syntax

Unconditionally transfers execution to a specified line in the current script.

GOTO label

The GOTO command has the following argument:

Part	Description
label	String indicating the destination line.

Remarks

The label must be correctly specified with a ":CAL_:".

This command is useful only in a script file, and has no effect when entered on the Command Line.

GOTO example

This example allows the user to batch-enter several new compounds and input one field for each. It starts with a MSG command to tell the user what to do. Then it has a line starting with a colon (:) that serves as a label to start the loop.

The INPUT command gets the name of the new entry from the user. The IF command tests to see if the user entered

a blank string, and EXITS the script if so. After the Menu Command Record Add creates a new record, the PUTDATA command enters the user's string into the molname field.

Closing the loop, the GOTO command jumps execution back to the start label.

```
msg Enter new compound names. Enter a blank name when complete
:startinput lif $v1 = "" exitrecord addputdata molname $v1goto start
```

CAL commands IF syntax

Conditionally executes a command depending on the value of an expression

```
IF variable operator value command
```

or

```
IF variable operator value
command[ELSE command]ENDIF
```

The GETDATA command has the following arguments:

Part	Description
variable	User variable that is the basis of the comparison.
operator	Comparison function.
value	String that is to be compared to the variable.
command	Any valid CAL command.

Settings

The settings for operator are:

Setting	Description
<	Command will be executed if value is less than variable.
>	Command will be executed if value is greater than variable.
=	Command will be executed if value is equal to variable.
!=	Command will be executed

Setting	Description
	if value is not equal to variable.
<>	Command will be executed if value is not equal to variable.

Remarks

Only one CAL command may be executed by each IF statement. However, that one command may be a GOTO that skips over many other commands.

CAL commands LOOP...ENDLOOP syntax

Executes a block of statements repeatedly.

```
LOOP [count]
[statements]ENDLOOP
```

The LOOP...ENDLOOP commands have the following arguments:

Part	Description
count	Number of times the statements are to be repeated.
statements	Commands that are repeated a number of times.

Remarks

If count is omitted, the statements will be repeated a number of times equal to the number of records in the current list. This command is useful only in a script file, and has no effect when entered on the Command Line.

LOOP example

This example uses the LOOP command and the Menu Command Record Next to display all of the records in the current list.

The PAUSE command keeps each record on screen long enough for people to see it.

```
looprec nextpause 2endloop
```

CAL commands PAUSE syntax

Temporarily stops the script from executing.

```
PAUSE [deciseconds]
```

The PAUSE command has the following argument:

Part	Description
decisecs	Amount of time to suspend script execution.

Remarks

10 decisec = 1 second.

If the decisecs argument is omitted, the script will be suspended for 2 seconds.

CAL commands EXIT syntax

Immediately halts execution of a CAL script.

EXIT

Remarks

This command is useful only in a script file, and has no effect when entered on the Command Line.

ChemScript

ChemScript is the cheminformatics Software Development Kit (SDK). It contains the programming algorithms that are common throughout PerkinElmer products. As a software developer, you can apply ChemScript to create your own scripts to use with ChemBioOffice applications.

All example script files in the ChemScript library are available in Python and C#.NET. If you are familiar with either of these languages, you will find these scripts easy to understand. However, if you are new to either Python or C#.NET, we suggest that you refer to the Web sites and books listed at the end of this chapter.

Although ChemScript is available in Python and C#.NET, we will use Python in this guide to explain ChemScript. Python is a non-proprietary and widely used programming language.

Why use ChemScript?

ChemScript adds considerable versatility to how you manage your chemical data. Using ChemScript, you can modify, view, and transfer your data from one place to another using your own custom rules.¹

Here are just a few common uses:

Salt splitting and stripping

Identify and remove salt fragments from a drawn structure and register the pure compound.

Canonical codes

Generate canonical codes for a set of structures and use the codes to find duplicates in your data.

File format conversion

Convert structure or reaction files from one format to another.

Generate properties

Execute Struct=Name or generate physical property features found in ChemBioDraw.

Common scaffold orientation

Enforce standard orientations of structures based on the established orientation of a common substructure.

2D Structure Diagram Generation (SDG) and Cleanup

Generate new 2D structures from connection tables without coordinates and clean up existing 2D structure using the ChemBioDraw algorithms.

¹ChemScript lets you convert up to 10,000 data records per day. For greater capacity, you will need ChemScript Ultra. For information, contact PerkinElmer.

How ChemScript works

The most fundamental use for ChemScript is to read data from one source, modify the data using a script, and write the modified data to another location. You can retrieve data from or write to almost any database, file(s), or application.

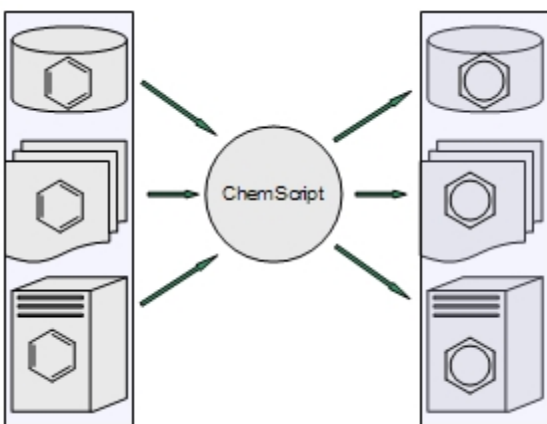


Figure 21.1: A ChemScript script can retrieve data from one source, modify the data, and write it to another location.

How the data is modified is determined entirely by the script. The script can delete data, calculate or add new data, or edit existing data. The data can be either text, structures, or both.

Since ChemScript scripts are like any other Python scripts or .NET program, you can execute the scripts using either the Windows command line or any development environment.

Getting Started

By default, ChemScript and Python 3.2 are installed on your local computer when you install ChemBioOffice. After the installation, we suggest that you follow the simple exercises in this section to familiarize yourself with ChemScript.

For editing ChemScript files, we recommend that you use one of the many programming tools that are available. One of these tools, IDLE, is included with the Python installation. To learn more about IDLE, see the online Help in the IDLE main menu.

Starting IDLE

From the Start menu, go to **All Programs>Python 3.2>IDLE (Python GUI)**. After Python starts, a Python shell opens in IDLE and the header indicates a version of Python (Ex. 3.2), and a version of IDLE (Ex. 1.2). A prompt will also appear: ">>>"

Loading the ChemScript Library

At the command prompt, type the line below and press <Enter>:

```
import ChemScript14 *
```

 **Note:** The command is case-sensitive.

A “Welcome to CS ChemScript” message appears, followed by a command prompt.

ChemScript Help

You can read a description of any ChemScript class within IDLE. For example, enter the line below to return Help for the ChemScript `Atom` class.

```
help (Atom)
```

The help will begin with a message such as:

“Help on class Atom in module ChemScript14:”

Enter SMILES Data

You can also test to verify that ChemScript is working correctly. One common use for ChemScript is to apply functions to SMILES strings. For example, you can assign a SMILES string to a variable. Type the following line and press <Enter>:

```
myMol = Mol.LoadData('C1CCCCC1C')
```

This message appears:

```
Open molecule successfully: chemical/x-smiles
```

Reporting a Chemical Name

You can now use the variable to report the chemical name. For example, at the command prompt type the line below and press **Enter**:

```
myMol.chemicalName ()
```

The line `methylcyclohexane` is returned.

Counting Atoms

To count the atoms in the string, type the following line and press <Enter>:

```
myMol.CountAtoms ()
```

The command returns the atom count for the structure defined with the `chemical/x-smiles` format for `'C1CCCCC1C'`, which is 7.

Exiting IDLE

To exit, type this line and press <Enter>:

```
exit ()
```

Confirm any prompts to complete the exit command. Python IDLE exits.

Getting Started Guide

ChemScript includes a Getting Started guide to help you begin developing and using your own scripts. To open the guide, go to **Start>All Programs>ChemBioOffice2014 >ChemScript 14.0>Getting Started**.

The document includes notes on the ChemScript objects and functionality, Python, and an overview of examples installed.

Editing Scripts

Using IDLE or another development environment, you can either edit the scripts provided with ChemBioOffice or create your own. Regardless of how you develop a script, it must include these commands:


```
from sys import *
from os import *
from os.path import *
from ChemScript14 import *
```

The first command imports the python system. The second and third commands import the operating system modules. The last command imports all the ChemScript functions. After you include these command lines, how you develop the rest of your script is entirely up to you.

Introducing the ChemScript API

ChemScript includes a ChemScript API reference guide. You can find the guide at **Start>All Programs>ChemBioOffice 2014>ChemScript 14.0>API Reference**. It provides links and information for the ChemScript classes.

The ChemScript object model comprises two fundamental levels of functionality, described below.

ChemScript Classes

At the top level, the API consists of four classes:

Atom

Chemical element, charge, bonds to neighboring atoms, drawing coordinates, 3D coordinates (if available), stereochemistry, etc.

Bond

Bonded atoms, bond order, etc.

Molecule (Mol)

A chemical connection table, which can represent one or more molecular fragments. This class also includes file I/O capabilities and other advanced chemistry functionality such as stereochemistry.

Reaction (Rxn)

A chemical reaction with one or more steps.

Functions and Algorithms

The secondary level consists of the core set of high-level features that you can modify to meet your specific business needs. Some examples are described below.

Template Based Normalization

Enforce standard representations of functional group structures in chemical data.

Template Based Product Generation

Automatic generation of products from a set of reactants and a generically defined reaction. For example, reactions like those between amines and carboxylates.

Substructure Identification and Mapping

Atom-by-atom comparison of a molecule with a substructure. Positive matching provides an atom-by-atom map of the substructure atoms to those in the molecule.

Salt Stripping

Remove salts from a reaction based on a pre-defined list of salt fragments.

Structure Orientation

Enforce standard orientation of structures based on the established orientation of a common scaffold.

2D Structure Generation and Cleanup

Use ChemBioDraw-based algorithms to generate structure from scratch or after modifying chemical data using a program.

Canonical Codes

Generate unique identifying codes from a chemical structure.

File Format Conversion

Read and write file data using all PerkinElmer supported file formats (CDX, CDXML, MOL, CHM, SKC, SMILES, etc.).

Chemical Name and Structure Conversion¹

Use the ChemBioDraw Struct=Name feature to generate structures from chemical names and names from their structures.

Molecular Mechanics

Optimize molecular structures using the MM2 force-field.

The ChemScript API online

PerkinElmer also provides the documentation of the API online. You can find the API at sdk.cambridgesoft.com.

Tutorials

We provide several sample scripts to illustrate how you can develop your own custom code to meet your business needs. Many of the scripts we use are in the ChemScript samples directory. By default, this directory is where ChemBioOffice 2014 auxiliary files are installed:

On Windows 7 and 8:

C:\ProgramData\CambridgeSoft\ChemOffice2014\ChemScript\Examples\

¹Premium functionality that may be licensed from PerkinElmer.

For the sake of brevity, we won't repeat the scripts in this manual or try to teach Python. However, we briefly describe what you can do with the code examples so that you can modify and expand upon them for your own use. As you read the tutorials, you are encouraged to view the code in IDLE and edit it as desired to see how each example works. For more on IDLE, see ["Getting Started" on page 296](#).

Example 1: Automated Structure Clean up

This sample script cleans up the structures in multiple ChemBioDraw files all at the same time. It uses the same cleanup function that is in ChemBioDraw. The script is in the directory `Example.001/script.py`. The script reads the CDX structure files from a source directory, applies the cleanup feature to each structure, and write the modified files to an output directory. The original files remain unchanged.

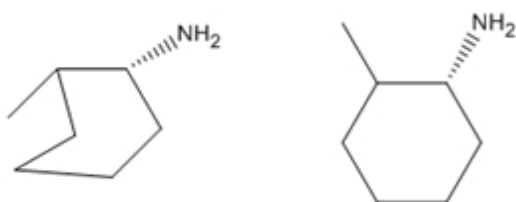


Figure 21.1: The structure cleanup script reads a structure file (left) and creates a new, cleaned up structure file (right).

This example uses the CDXML format. Other formats, such as MDL MOL, may also be applied. You can also force ChemScript to use specific file formats for reading and writing data.

Example 2: Create an SD file

This example illustrates how you can create an SD file from existing CDXML files. You can find the script at `Example.002/script.py`. We begin with a list of CDXML files that each contain a chemical structure. The list of files is hard-coded into the ChemScript script. When executed, the script uses the `SDFFileWriter` method to create an SD file that includes all the structures.

Example 3: Create a list of CDXML files

This example illustrates how to read an SD file and write a list of CDXML files. You can find the source file at `Example.003/script.py`.

Example 4: Filter an SD file

This example uses the `atomByAtomSearch` method to demonstrate a simple application of the "atom-by-atom" substructure search in ChemScript. The program reads an SD file and filters structures into one of two output SD files, structures that contain a phenyl group and structures that don't. It also illustrates how you can read chemical data formatted as a SMILES string. See `Example.004/script.py`.

Example 5: Computing Canonical Codes

This example script checks whether any structures appear in both of two SD files based on the structures' canonical codes. The output is a new SD file with the duplicate structures excluded. See `Example.005/script.py`.

This example first computes the canonical code for each structure. Since the canonical code does not vary with different representations of the same chemical structure, you can use it to determine whether two structures are chemically equivalent.


This example also introduces the Python Dictionary, which is an associative array. The dictionary maps a key to a value. The dictionary is used to determine whether a canonical code has been previously encountered.

This example uses an alternate looping construct to read an SD file.

Canonical codes should never be permanently stored because their representation can change among different versions of ChemScript.

Example 6: Structure Overlay

This script introduces the ChemScript structure overlay feature. It uses a scaffold structure file to superimpose two chemically similar structures. The script first examines the structures in an SD file that contain a common scaffold substructure. It then aligns these structures so that they have the same orientation with respect to the scaffold. See `Example.006/script.py`.

 **Note:** The overlay functionality can also be used to align three dimensional structures.

Example 7: Reaction Transformation

This example demonstrates reaction transformation. This means that you can draw a reaction that defines a transformation of a molecule and then apply that transformation to a set of structure files.

All the files necessary for this tutorial are in the `Example.007` directory. The `transform1.cdxml` file provides the reaction that defines the transformation. The `input.sdf` file contains all the structures that will be transformed.



Figure 21.2: The transform file defines how the transformation is applied to the source structures.

The script searches the input file for structures that contain a nitro group, shown as a reactant in the transformation file. If a structure is found, the script transforms the nitro group to the form shown in the product and copies the entire structure to a new file. The figure below shows one example.

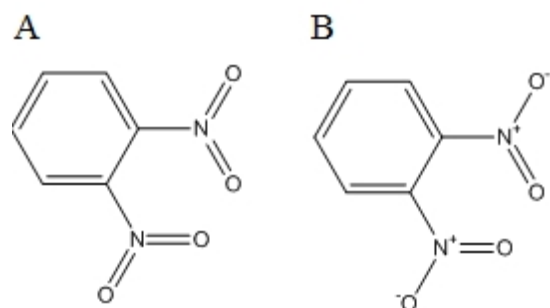



Figure 21.3: A) before the transformation is applied; B) after transformation.

Structures that don't meet the search criteria are ignored.

Example 8: Simple salt stripping

The program reads an SD File, identifies and removes salt components (if any are present), and outputs two SD files. The output structure file contains the original structures without the salt component, and the output salt file contains the salt components that were stripped, along with a reference to the original structure. See

`Example.008/script.py`.

 **Note:** This example uses a default set of salts that PerkinElmer provides. However, you can also define a customized salt table that enables you to designate which chemicals are considered salts.

Useful References

There are numerous resources available for learning Python and C#.NET. Just a few of the many books and Web sites are listed below.

Books

Python

- *Beginning Python: From Novice to Professional* by Magnus Lie Hetland.
- *Dive into Python* by Mark Pilgrim.
- *Learning Python* by Mark Lutz & David Ascher. This is a beginner/intermediate learning manual and reference.
- *Python in a Nutshell* by Alex Martelli. This book is a brief introduction and good reference to Python.

C#.NET

- *C# in a Nutshell* by Peter Drayton, Ben Albahari, and Ted Neward.
- *Pro C# with .NET 3.0* by Andrew Troelsen.
- *C# Essentials* by Ben Albahari, Peter Drayton, and Brad Merrill.
- *C# 3.0 Cookbook* by Jay Hilyard and Stephen Teilhet.

Web Sites

Python

You can find more information on Python at: <http://www.python.org>. This is the official Python programming language site.

.NET

For information on .NET, see <http://msdn.microsoft.com>


CS Oracle Cartridge

Prerequisites for connecting ChemBioFinder/Oracle to an Oracle database:

- ChemBioFinder Ultra
- Oracle Client
- Oracle OLEDB Provider

Oracle Client: We recommend version 11.2 or newer. This is available as a free download from the [Oracle Web site](#).

1. Download and install the Oracle Client software (version 11.2). This is available free from Oracle. Note that there are different versions for 32 bit, 64 bit, and Itanium systems.

 **Note:** *The minimum installation (called “Runtime”) is not adequate to provide support for ChemBioFinder/Oracle. In addition, you must subsequently choose the “Custom” install, and select “Oracle Windows Interfaces.” If you do not have the requisite components installed, an attempt to open an Oracle database will generate the error: “Provider cannot be found. It may not be properly installed.”*

2. Have the Oracle administrator install CS Oracle Cartridge (13.1) . You can use ChemBioFinder/Oracle to access any Oracle database, but one without the CS Oracle Cartridge cannot be used to store or search chemical structures.
3. Have the administrator set up a user account for you. ChemBioFinder/Oracle requires that you log into a tablespace for which you have write privileges. For access to corporate files, the administrator needs to grant you read-only privileges to tables and views of interest outside your tablespace.
4. On your client computer, use Oracle Net Configuration Assistant to connect to each Oracle database you intend to use.

In principle, ChemBioFinder/Oracle can access any Oracle database. In practice, some preparations are recommended:

- You can access data in Oracle databases that do not have the CS Oracle Cartridge installed — for example, biological databases — but you cannot search or sort by structure or structure-related properties. ChemBioFinder/Oracle has been developed with CS Oracle Cartridge version 13.1, and is not likely to work with earlier versions. (To display the Cartridge version, look in the **Database** tab of the **Properties** dialog box. The version is listed, or “None” if unavailable.)
- Any table containing structures should have a unique primary key. Otherwise, a warning message will appear. You can set up a primary key in ChemBioFinder (see ["Indexing" on page 186](#).)
- As a ChemBioFinder/Oracle user, you must have sufficient privileges to create tables in your own tablespaces.

Fast-move caching scheme

Normally, when you open a database or do a search, ChemFinder issues a select statement. The server then processes the statement, prepares a set of records, and provides a “cursor” for navigating through them to retrieve the data. Unfortunately, when the cursor is asked to move to a position N records from the current one, it requires all N to be downloaded to the client. In the worst case, you open a large table and use **Move Last** to see the last record, and the entire table is sent down from the server.

To get around this problem, an alternative mechanism is available in ChemFinder/Oracle. In this scheme, instead of requesting complete forms-full of data over a list or query, ChemBioFinder/Oracle retrieves only the ID or primary key field, and stores (“caches”) the results in an array in memory. It then need not rely on the normal recordset cursor. When a request is made to move to record N, the key for that record is looked up in the array, and a new select statement is issued to retrieve only the corresponding record. Thus, **Move Last** takes no more time than **Move Next**. There are tradeoffs between the two schemes:

- The new scheme takes time to retrieve the set of ID's during a database open, or after a search or sort. Normally, this is quite fast even for a large table, and steps are taken to avoid repeating the operation unless necessary. A message on the status bar indicates when this caching process is taking place.
- In the new scheme, moving from one record to the next is somewhat slower, since each move involves a (fast, one-hit) search. This is not very noticeable during list browsing, but slows down multi-retrieve operations such as filling table view.

Because of these drawbacks, the new caching scheme is OFF, by default. To turn it ON for an Oracle-connected form:

1. Go to **File>Database**, or otherwise bring up the Properties dialog box.
2. Go to the **Oracle** tab of the dialog box.
3. Check the **Cache ID's for faster moves** box.
4. Click **OK**.
You will get an alert instructing you to:
5. Save the form, close it, and reopen it. This is currently necessary in order to reinitialize the database connection with the caching scheme.

Configuring via CF_SETTINGS table

When you select a table in an Oracle database, ChemBioFinder determines and displays the types of columns in the table. Here's how it decides whether a particular column contains structures:

- The column must be of type CLOB or BLOB.
- If the column has been indexed by the CS Oracle Cartridge, then it is a structure column.
- If there is no index, then ChemBioFinder/Oracle looks for a table of configuration info called CF_SETTINGS. If the column is listed in that table as being of type STRUCTURE, then it is a structure column.

The CF_SETTINGS table is also used to determine whether a particular column of type BLOB contains ChemFinder pictures (Windows metafiles).

To designate an unindexed column as a structure column:


1. In the **Field** tab of the **Properties** dialog box, select a column of type BLOB or CLOB.
2. Click **Set As Structure**. If the CF_SETTINGS table has not yet been created, an alert indicates that it is about to be created.
3. Click **OK** to proceed and create the table, or **Cancel** to abort without creating. If the table already exists and there is already an entry for the selected column, you are given an alert and a chance to cancel.
4. Click **OK** to proceed. A new entry is created in the table, marking the selected column as a structure column.

To designate a BLOB column as a picture column:


1. If the CF_SETTINGS table does not exist, create it using steps 1-3 of the above procedure.
2. Edit the table one of two ways:
 - with an Oracle client tool such as SQL*Plus, OR...
 - by opening it in ChemFinder and creating a form.
3. Add a row containing the table name, column name, and a column of type PICTURE.

Searching

Here's what happens when you present a query to ChemBioFinder/Oracle:

1. If the query contains a structure, it is converted to a text representation and copied to a temporary table.
 -  **Note:** *The table is called temp_queries. It is created in the CSCartridge tablespace, and removed as soon as the search is finished or interrupted. ChemFinder does not yet handle the case of multiple structure boxes where more than one contains a query.*
2. The query is converted to a SQL select statement. Query components in form boxes are ANDed together (just as in ChemFinder), where the structural parts are calls into the CS Oracle Cartridge structure search functions.
3. The hits table is created, if it does not already exist.
4. A unique ID is assigned to the new list which will result from the search. The select statement is wrapped in a larger SQL statement which will cause the results to be deposited directly into the hits table.
5. The SQL is executed. When the search is complete, the results are new rows in the hits table. Each row contains the new list ID alongside the ID of a record from the searched table.
6. The final list is prepared by a join, selecting rows from the main table which have record ID's matching those of the new list in the hits table.

The resulting list is ready to browse, save, export, etc.

-  **Note:** *Text searches in Oracle are case sensitive. You will get different hits from the query "benz*" than from "Benz*."*