Clone Manager Help

Clone Manager is a program to support the Molecular Biologist in planning experiments to create and analyze recombinant molecules. It manages molecules, primers, and sequence alignments and assembly. Provides tools for cloning using classical techniques as well as wizards for advanced cloning methodologies. Numerous viewers provide data views to enable you to explore and visualize the data.

This Clone Manager manual and help system provides comprehensive help topics to assist you in using the software or understanding a results display. You can browse for general information or review a feature you might plan to use.

Clone Manager is implemented as three main modules – Cloning, Primer Designing and Sequence Alignments – with supporting functions for loading, saving, finding, exporting, and processing Molecular Biology data.

Click on a link below to open a list of help topics within the selected category.

- INTRODUCTION
  - Getting Started
  - Molecule file handling
  - Molecule Viewer Window
- DATA LISTS
  - Molecules
  - Primers
  - Enzymes
  - Sequence Phrases
  - DataBook
- CLONING
  - Cloning and enzyme operations
  - Cloning wizards
  - Analytical molecule operations
- PRIMER DESIGNING
  - Design primers
  - Primer Viewer Window

  - Direct and special primer wizards
  - Analyze primers
  - Use, organize primers
  - Primer Collections
- ALIGNING SEQUENCES
  - Multiple sequences
  - Comparing two sequences
  - Alignment Viewer Window
  - Scan for similarities
  - Sequence Assembly
  - Sequence Assembly Viewer
  - NCBI BLAST
- DISCOVERY and TOOLS
  - Sequence Phrases
  - DataBook
  - Find in files
  - Process Molecule
  - Analyze Molecule
  - Export
- MAP DRAWING
  - Molecule graphic maps
  - Enhanced Map
- MAIN MENU REFERENCE
  - File Menu: Open, Save, Preferences, License…
  - View Menu: Molecule, Primer, Enzyme lists…
  - Clone Menu: Cloning operations and wizards.
  - Primer Menu: Design, wizards, analyze.
  - Align Menu: Multi, pair, scan, assembly, Blast…
  - Discover Menu: Find operations…

    - Tools Menu: Translate, analyze, profile, export…
- APPENDIX
    - Installing Clone Manager
    - Molecule File Formats
    - Import Features Preferences
    - REBASE File Updates
    - Primer Criteria
- Table of Contents

Copyright © 1994-2024

Sci Ed Software LLC.

All rights reserved.

Introduction

Clone Manager functions to view and use molecules, and other types of molecular biology data.

The next chapter provides a quick start guide for new users and those upgrading from older versions. The following chapters provide more detailed information about the program and how to use the functions. Click links to quickly jump to more detailed information about a topic.

If you have not already installed the Clone Manager client program, please refer to the Appendix: Installing Clone Manager.

New users and overview:

- Quick Start
- Molecule File Handling
- Molecule Viewer Window

Program Main Menus:

The main menu is divided into the following main categories:

- File Menu: Open, Save, Preferences, Help, License…
- View Menu: Change display view, access data lists…
- Clone Menu: Cloning operations and wizards.
- Primer Menu: Design, wizards, analyze.
- Align Menu: Multi, pair, scan, assembly, Blast.
- Discover Menu: Find operations…
- Tools Menu: Translate, analyze, profile, export…

Other sections:

- Appendix - Supporting documentation
- Table of Contents

Getting Started

This Getting Started guide gives a brief overview of how to use Clone Manager. Later sections will provide an in-depth coverage of these functions.

This getting started guide assumes you have already installed the Clone Manager client program. If not, then please refer to the Appendix: Installing Clone Manager.

Quick Start

- Run Clone Manager – select license type.
- Main window
- Load molecules using the File, Open menu.
- View molecule details.
- Select menu options for File, View, Clone, Primer, Align, Discover or Tools.

Select license type

The first time you run the Clone Manager program, you will see the configuration wizard for selecting how Clone Manager will be licensed. Options are:

- Start a 30-day trial evaluation – select this option if you want to try Clone Manager. The trial is fully functional and has no limitations except for the 30-day time limitation.

- Activate your license on this computer – select this option if you already have a seat license for Clone Manager. On the next screen you will enter your license ID and activation password.

- Join a concurrent license group - select this option if you are in a group that is sharing a concurrent license installed on a network server. On the next screen you can enter any custom settings required by your network administrator. For most users, the default settings will be sufficient, and you can simply click the next button.

Once Clone Manager has been configured, you will see the main window. The next time you start Clone Manager you will come directly to this window.

Main window

At the top of the main window is the main menu bar which gives access to the various categories of functions. Under the menu bar is the main toolbar that gives quick access to the main sections of the program such as the Molecule List that lists the molecules currently loaded into the program. The main part of the window will initially be blank but will contain molecules, cloning wizards and analysis results display as you use the program.

Load a molecule

Use the File, Open menu (or the folder icon at the start of the main toolbar) to show the File-Open dialog box. Use the standard Windows file explorer navigation panes to move to the location of your molecule files on your computer and select the molecule file to load.

Clone Manager can import molecule files in standard GenBank, EMBL and FASTA formats. It can also import plain sequence data. In addition to molecules, Clone Manager can open data files containing primer or alignment data. These are described in more detail in their respective sections.

View the molecule

Once you have opened a molecule file, or imported your sequence data, you can explore details of the molecule.

The Molecule Viewer Window is the main viewer for molecules. Along the bottom of the window are tabs that allow selection of the types of molecule data. Tabs can also be selected using the main menu: View, Molecule Viewer.

- Map – shows a graphic map containing features and sites. Typically, sites will be known restriction enzyme sites and will be listed in the right-hand panel. The map is active, and you can hover over, or click, any item to get more information. The map toolbar gives access to more advanced options which are described in the Map section.
- RMap – shows all possible restriction enzyme sites for the molecule. The toolbar provides more control over the display and allows selecting a different enzyme list or filtering the results. Filtering can be used to identify restriction enzyme(s) that could be used for cloning experiments**.** The toolbar also provides access to REBASE data to get more information such as properties, suppliers, isoschizomers and enzymes with compatible ends.
- Sequence – shows the sequence of the molecule. The toolbar settings icon gives access to the format dialog to enable extensive customization of the display. You can show enzyme sites, features, and translations. Your preferred display style can be saved to allow quick selection when desired. Other toolbar buttons can be used for editing the sequence.
- Features – shows a list of the features of the molecule. The toolbar provides customization of the display, properties, and editing.
- Info – shows basic information about the molecule such as name, description, and notes. Some properties are also shown. The toolbar provides functions for editing and additional properties.

Use the molecule

Use the Clone Manager main menu to select what to do with the loaded molecule.

- File – used to save changes, select preferences, export, and import.
- View – used to change the active display view or view the list of loaded molecules, primers, or other data.
- Clone – provides access to cloning operations. Basic cloning operations are cut with restriction enzyme and ligation. Cloning wizards can be selected, and each wizard will guide you through the steps required.
- Primer – provides access to the primer designing functions. In addition to designing basic primers, or pairs of primers, there are several wizards for advanced primer design. Also found here are functions for analyzing primers and using primers to create products.
- Align – provides access to the sequence alignment functions. Menu options include multiple sequence alignment, pair alignments, scan for similarities, sequence assembly and submitting Blast searches.
- Discover – provides numerous methods for finding. Examples include searching files on disk for some property such as a feature name.
- Tools – provides methods for processing molecules (invert, rotate, translate), analyzing molecules (find open reading frames, find repeats or dyad symmetries), molecule profiles and export functions.

Molecule File Handling

Clone Manager can import molecule files in standard GenBank, EMBL, FASTA or FASTQ formats. It can also import plain sequence data.

In addition to molecules, Clone Manager can open data files containing primer or alignment data. These are described in more detail in their respective sections.

Once you have opened a molecule file, or imported your sequence data, you can explore the sequence and add additional information to complete a more complex molecule file. You can save your molecule data to a disk file, ready for later use.

- Opening molecule files
    - Troubleshoot file loading.
    - Retrieving files from Entrez.
- Saving molecule files
    - Export molecule data
    - Multiple file conversion
- Creating new molecules
- Using Clone Share
- Using the Molecule List

Opening Molecule Files

Use the File, Open menu to open existing molecule files and load them into the program for use. You can have multiple molecule files open at one time, and you can create new molecules, as needed.

You can read molecule files in Clone Manager file formats, standard GenBank, EMBL, FASTA or FASTQ formats, or plain text sequence files. You can control which feature keys are imported when you open standard GenBank or EMBL files. You can also read molecule data directly from the clipboard.

- **To open files containing molecule information:**

1. Click the File Open menu or main toolbar button.
2. Find the file(s) you want in the Windows file open dialog box.
    a. Optionally select more than one file using the control or shift keys.

- **GenBank and EMBL import feature types**

You can set which feature types are imported when loading a databank file. Under main menu File, Preferences the Import Features tab lists the features that will be imported and how they will be displayed. By default, all other feature types are discarded.

- **Open more than one file at the same time**

If 5 or more molecule files are to be loaded, you can elect to keep these files together in a group. If you select a file to load that contains multiple molecules, these molecules will automatically be placed in a group for easy handling.

- **Import from clipboard**

This is a quick way of getting molecules from web browser page displaying a databank file. Select the text with the mouse and copy to the clipboard (Control+C keys or right click). In Clone Manager select menu File, Other Tools, Import from Clipboard.

Problems with File Loading

If the program cannot automatically load a sequence file, it will identify the closest of the recognized formats (GenBank, EMBL, FASTA or ASCII text). It will try to provide you with information about the problem encountered and provide help with file loading.

In the File Help module, the file format and type of molecule are shown at the top left of this information window. The error is described in the top right. The scrolling data window shows the contents of the file. Usually, you will have one of the common file loading errors. You will want to try and correct the error or split the sequence from the rest of the file and then attempt to reload the file.

- **File Split**

  Move the cursor to the start of the sequence and click File Split. The file will be divided at the point you marked so that it will split the sequence from the header. In some cases, the program will insert missing identifiers (like the word "ORIGIN" for a GenBank file).

- **Go To Error**

  Click this button to move the cursor to the point at which the program recognized the error. You can then remove or replace erroneous characters.

- **Retry Load**

  Click this button to repeat the file loading operation for this file (use after editing).

Retrieving Files from Entrez (E-utils)

You can retrieve sequence files from the Entrez system at the NCBI using a module that can directly access the NCBI query system. You need an internet connection to use this function. Once you locate the file you need, you can import the molecule for immediate use in Clone Manager or you can download this file to your hard disk to save the complete GenBank file, including references and annotations.

- **To find the file you need:**

1. Click menu File, Other Tools, Retrieve from Entrez eUtils.
2. In the 'Search In' control, select Nucleotide for DNA molecules or select protein.
3. In the Query box, enter the locus name, accession number or descriptive information.
4. If you need a special sort order for results, set using the Sort control.
5. Click the Search button.
6. When the results are listed, find the file you want and click to select.

Search results are listed in summary format, 20 items per page. Use the Next or Prev buttons to move from one page of items to another. Drag the lower right corner of the window to increase the window size. Drag the join between data columns to change the column width.

- **To display more information about each file**

Use the Display control to switch from the Essentials (concise) display of file contents to the Expanded (all columns) display. The Expanded display includes additional columns of information such as CreateDate, UpdateDate, or TaxId.

- **To view the selected molecule file**

1. Select (click on) the file you want.
2. Click the Show GenBank File button.

- **To import the molecule for immediate use**

1. Select (click on) the file you want.
2. Click the Get Molecule button.
3. Set the checkbox if you want to also save the complete GenBank file to your hard disk.

The molecule will be imported into the Clone Manager program for use now. The sequence data, genes read from the features table, and descriptive information will all be a part of the new molecule. You can AutoScan to add enzyme sites or just use the sequence in an alignment or primer design operation. If you set the checkbox to also

save the complete GenBank file, you will be able to access the references and annotations in this file later.

- **Imported Features**

  Clone Manager files contain essential molecule information (sequence, select features, description), as well as restriction enzyme sites and notes you may add to the file. (Use File, Preferences, Import Features to select which features are imported). GenBank files contain essential molecule information and may also contain references, comments, and additional features that are not imported into Clone Manager.

  If all this information is important to you, you will probably want to have both a Clone Manager and a GenBank file for the molecule. When you create a Clone Manager file from a GenBank file (either by using the import for immediate use operation or by opening a GenBank file previously saved to disk), Clone Manager will automatically store a link between the two files so that you can view the extra information while working in Clone Manager.

Saving Molecule Files

You can save the active molecule to its current name and folder by using the menu File, Save command or the Save main toolbar button. You can also change the filename or location of an existing file before you save it, by using the menu File, Save As… command. If you are saving a file for the first time, you will automatically move to the Save As dialog box so you can name the file and specify its location.

You can save molecule files in the default Clone Manager binary molecule file format (*.cm5), or in an xml-based molecule file format (*.cx5). If you always want to save files in the same format, you can set this as a default preference.

- **Export Molecule Files**

  You can also export the sequence of the molecule or the sequence plus features (in a format like GenBank) using the Export Molecule Sequence option on the Tools, Export menu.

  Another option for exporting molecules is to use Multiple file conversion.

Exporting Molecule Data

Found on the Tools main menu.

You can export molecule data to a disk file or place this data on the Windows clipboard for later use in another program. You can select from the following export modes:

- **Sequence Only**

  Select this option to export only the sequence of the active molecule. The sequence will be exported in plain text format in upper case letters, 50 bases per line, with no base pair numbers, and no header or map information.

- **FASTA**

  Select this option to export molecule data containing the sequence plus one line of descriptive information (molecule name, description) in FASTA format.

- **Compact GenBank**

  Select this option to export essential molecule data in a format that is consistent with the standard GenBank file format. The data will include the molecule name, size, description, features (genes, regions, markers, labels), and sequence data, all with standard GenBank identifiers. This format uses the Windows character set and is a good choice for file sharing with others who may use alternate software applications.

- **Commented GenBank**

  Select this option to export a larger subset of molecule data in a format that is consistent with the standard GenBank file format. The data will include the molecule name, size, description, all features (including information-only features), sequence data, information entered in the Notes field, and the 'draw as' type you assigned for each feature. This format uses the Windows character set, when possible, but will use UTF-8 encoding when Unicode character support is needed. These files can be read back into Clone Manager with either encoding.

- **Converting multiple files**

  If you have a batch of standard Clone Manager files (*.cm5) that you want to convert to another format, perhaps for file sharing or archival storage in a standard format, you can use a special Multiple File Conversion Utility, found on the File, Other Tools menu.

Creating New Molecules

You can create a new molecule file for use in the program and later save this file to disk. The New Molecule Wizard will help you to enter all the information you need for the new molecule. If you just want to use sequence data immediately, you can simply paste a sequence from the clipboard, bypassing the wizard's data entry pages.

- **To create a new molecule**

1. Click menu File, New
2. Select Enter Complete Information and then follow Wizard instructions.
3. Or select Paste Sequence and then copy the sequence to the clipboard for import.

When you are finished creating a new molecule, you should save this file to disk if you want to use this molecule again later. (When the new molecule is open, click File, Save As…).

- **Entering Molecule Information**

You must enter a molecule name and molecule type. You can elect to enter a description or notes about the molecule.

Basic Molecule Information (required) -- Enter a name for the molecule and select the molecule type (DNA or Protein). To update this information later, open the molecule viewer window (select the molecule from the Molecule List if necessary) and click the Info tab.

Molecule Description and Notes (optional) -- Enter a short description of the molecule and/or a lengthier notes section. Some users enter keywords, history of construction, or comments in this field. This information is saved as a part of the proprietary program file format and can be edited or printed later.

- **Entering Sequence Data**

You can enter sequence data (type in or paste from clipboard) or simply enter the size in bps for your new molecule.

To enter sequence data, type in the complete upper strand sequence in the data entry box provided. You can enter valid or ambiguous bases. All sequence entries are in the Insert mode. Use the Delete key to remove erroneous bases, if necessary. You can also paste the sequence from the clipboard if you already have access to this sequence data in another document. Copy the sequence to the clipboard in your other application and then click the Paste button instead of typing in the sequence.

If you do not have sequence for this new molecule, enter the size (in bps) in the space provided.

- **Entering Molecule Format, Linear Ends**

  If your new molecule is DNA, you should indicate if this molecule is circular or linear.

  If you are creating a linear DNA molecule, you can indicate the restriction enzymes used to create each of the linear ends. Enter restriction enzymes names for one or both ends of the molecule. Enzyme names must be recognized by the program -- click the Lookup button for help with enzyme names, if required. You cannot use asymmetric cutters for molecule ends since the program does not know what the end is going to be.

  If you do not enter an enzyme name, the program assumes this to be a blunt end.

- **Entering Features**

  You can enter genes, regions, or markers as features for your new molecule. Enter a short feature name and select the appropriate feature key. Enter a feature description, if desired, and select the Draw As Type and style number. Next, enter start and end base pair positions (markers have only one position) and indicate if on the complementary strand for circular molecules.

  Use the ‘AddNew’ button to get a new blank data entry area for additional features. Use the Prev or Next buttons to move through the features already entered.

Using Clone Share

The Clone Share module is used to share molecule files between Clone Manager users and between users running the Clone Manager Viewer apps on a smart phone or tablet.

Clone Share data stores can be created on supported cloud storage platforms, such as Microsoft OneDrive, as well as on a local web server using the SciEd Web Share application.

- **Access the Clone Share Dialog**

  On the File menu, click Clone Share and select either the Open or Save option to open the Clone Share dialog.

- **Getting Started**

  When you first open the Clone Share dialog you will not have any Clone Shares defined. Clone Manager will automatically open the dialog for adding a Clone Share so that you can enter your first Clone Share connection.

- **Add a Clone Share**

  Open the Clone Share dialog as described above and click the first toolbar button (tooltip Add Clone Share). Select the Share Type to pick where the clone share is located, e.g., Microsoft OneDrive to use Microsoft's cloud data storage service. Enter the required information fields. The Clone Share Name is the name that identifies the data store collection on the remote service (use the lookup hyperlink to show the names of existing shares). Depending on the Share Type, you may need to enter or select a user account.

- **Using SciEd Web Share**

  This service is available if your network administrator has installed this application on a local web server. You will need to enter the URL of the web server that you want to connect to (Examples: https://univerity.edu or http://server.com:8080). Now enter the name of the Clone Share Location that contains the files you want to access. Your network administrator will have set this up for you and will tell you the name to use or you can use the lookup hyperlink. Note that the name is case-sensitive.

- **Using Cloud Storage Services**

  Cloud data services provide data storage allocated to a unique account usually identified by an email address. This account identifier may be either your personal email address or can be an account that you create for your lab. In the latter case, you will be using the account for data storage and do not need to use the email address.

- **Edit a Clone Share Definition**

  Open the Clone Share dialog and use the left pane to select the Clone Share that you want to change. Click the toolbar button (tooltip Edit Clone Share) and make any

changes needed. Note that if you need to select a different share then you will need to add a new Clone Share.

- **Remove a Clone Share**

  Open the Clone Share dialog and select the Clone Share that you want to remove. Click the third button (tooltip Remove Clone Share) and confirm that you want to remove this Clone Share. If you remove a Clone Share definition, you are only removing your access to it. The data contained in the remote Clone Share will not be deleted.

- **Change the order of Clone Shares displayed**

  Open the Clone Share dialog and select the Clone Share that you want to move. Click the fourth or fifth button to move the selected Clone Share up or down.

- **Navigate within a Clone Share**

  Open the Clone Share dialog and click within the left and right panes to open the appropriate folders.

- **Add a Folder**

  Open the Clone Share dialog and right click the node in the left pane where you want to add a new folder.

- **Rename a Folder**

  Open the Clone Share dialog and select the folder in the left pane that you want to change the name. Click a second time to activate edit in place and type in the new name for the folder. Press the Enter key to complete the name change. You can also right click the folder in the right panel and select Rename from the popup menu.

- **Delete a Folder**

  Open the Clone Share dialog and navigate to the appropriate place. Right click the folder in the right panel and select the option Delete from the popup menu. Note that this will also delete any molecule files contained within the folder.

- **Upload Multiple Molecules**

  You can upload molecules stored in folders on your computer in one simple operation using the Multiple File Converter utility located on the File, Tools menu. More information is available in the topic Multiple File Conversion.

- **Move or Delete Multiple Items**

  In the Clone Shares dialog box, navigate to your items. Hold down the Control (or Shift) key to select multiple items in the right list panel. Now mouse right-click one of the selected items and select the appropriate menu option.

- **Maintenance**

  You can perform a validation of the data stored on a Clone Share. Select the Clone Share and then select Edit a Clone Share Definition as described above. Click the Maintenance hyperlink. While the maintenance procedure is running, do not access the data in that share. When maintenance is complete, it will report the results.

- **Problems**

  The left panel location icon has a red X:

  This indicates that the Clone Share Location is not accessible. There are several possibilities:

1. The remote web server is down. Check with your network administrator.
2. Your connection to the web server is down or is being blocked by a security system. Check your network connectivity.
3. The SciEd web server location is not correct or has changed. Check with your network administrator to see what the correct settings for URL and Location are to access this Clone Share. The name of the Clone Share is case sensitive.
4. The user account used by the Cloud storage service is no longer valid. For security reasons, cloud services require re-authentication at intervals or where an account has not been accessed for a while. Click on the share icon and re-enter your account credentials.

Molecule Viewer Window

The Molecule Viewer Window is the main viewer for molecules and shows detailed properties of the molecule. The viewer will be seen when you open or select a molecule.

Along the bottom of the window are tabs that allow selection of the different types of molecule data. Tabs can also be selected using the main menu View, Molecule Viewer.

- Overview – brief description of the viewer window.
- Map – shows a graphic map with annotations for features and sites.
- RMap – shows restriction map data and allows finding enzymes that can be useful for cloning.
- Sequence – shows the sequence of the molecule. Toolbar buttons allow for formatting and editing.
- Features – shows a list of the features of the molecule. The toolbar provides customization of the display, properties, customization, and editing.
- Info – shows basic information about the molecule such as name, description, and notes. Some properties are also shown. The toolbar provides functions for editing and additional properties.
- Zoom – enables viewing data for a sub-region of the molecule in more detail.

Overview

The molecule viewer opens when you load a molecule into the program or when you select a molecule from the Molecule List (and are not in the middle of another operation). Tabs along the bottom of the window let you view the map, restriction map, sequence, features table, or information describing this molecule.

The molecule map is the default view and shows a graphic representation of the molecule and its features (genes or regions of interest).

Click the *RMap* tab to view the detailed restriction map data for this molecule.

You can view the molecule sequence by clicking on the Sequence tab at the bottom of the viewer window. You can display the sequence in a variety of formats, edit the sequence, and use a search function to find places in the sequence that match a search sequence.

The Features tab gives you access to the features table for this molecule. You can add or remove features or view their properties.

The Info tab shows the molecule name, size, description, and notes. You can add or edit this information, as needed.

The Zoom button can be found at the left end of the tab bar. Click this button and select a region of the molecule to view. The map, restriction map, sequence, and features table views will display data for the selected region only.

Molecule map (working map)

To view the molecule map, open the molecule viewer window. If the map is not displayed, click the *Map* tab at the bottom of the window to change to this view. The map will be either circular or linear, depending on the molecule.

The basic molecule map shows a graphic representation of the molecule, including features (genes, regions, and markers). The molecule name and size appear in the window title bar. You can add features (genes, regions, markers) to your molecule map or remove features, if necessary.

The map can also show enzyme sites, primer sites, or feature labels entered for this molecule. You can use a toolbar button to toggle between displaying enzyme sites, primer sites, or labels.

The map is active, and you can hover over, or click, any item to get more information. The map toolbar gives access to more advanced options and is described below.

- Working map display
- Mouse actions
- Molecule Graphic Maps
  - Format Map Print Basics
  - Format Map Features
  - Format Map Sites
  - Format Map Content
- Circular vs Linear Maps
- Working with Map Sites
- Send Map Options

Working Map Display

When a molecule has been loaded into the program or selected from the Molecule List, the molecule viewer window will open, showing a working map of this molecule. If the map is not displayed, click the Map tab at the bottom of the window to change to this view.

You can use the mouse to get more information or initiate actions on this display. You can select a map site to view detailed information about or use an enzyme site, primer site, label site or feature. You can also use other toolbar buttons to find more sites to add to the map or to set options for the working map display or for printing the graphic map.

Features

Features that have been entered to the features table for this molecule will appear as follows:

- Genes appear as arrows below (inside) the map line.
- Regions appear as boxes on the map line.
- Markers appear as short, solid lines crossing the map line.

Click the Features tab of the molecule viewer window to add or modify features. Click the Find button and select ORF search to locate open reading frames and add them to the map as genes, if desired.

Enzyme Sites

Enzyme recognition sites appear above (outside) the map line. By convention, the enzyme site is marked at the left end of the recognition sequence for that enzyme. Enzyme sites that are very close together may be displayed on the same line. ++ symbol next to an enzyme name indicates that one or more enzymes very close to this site could not be displayed without overwriting. Point to the enzyme name to see the bp position for this enzyme and for any suppressed sites at this location. (Use the Map Sites list on the right to select one of these enzymes.)

In highly congested areas, the enzymes sites will be collected in order and moved to a location above or to the side of the map. Multi-cloning sites will usually be displayed in this way. If your map has no enzyme sites, click the Find button and select Find Enzyme Sites AutoScan to find enzyme sites using the specified scanning list. Click Clone, Find Enzyme Sites to locate other enzyme sites and add them to the map.

Primer Sites

Primer sites can be shown above (outside) the map line in much the same way that enzyme sites are displayed. The primer site is marked at the 5' end of the binding site. If your map has no primer sites, you can click Find, Find Primers that Bind to scan your

collection of primer files to find primers that will bind to this molecule. If you are designing primers, you can add these to the molecule map using toolbar buttons in the primer search results list window or in the primer viewer window.

Label Sites

Label sites can be shown above (outside) the map line and can be used to identify short sequence phrase locations. The label site is marked at the start position of the label. If your map has no label sites, you can click Find, Find Sequence Phrases to use your collection of sequence phrases to find those that are present in this molecule.

Click the Change Sites toolbar button to cycle through the display of enzyme sites, primer sites or label sites. Or click on the site list header control (right) to select the list of Enzyme sites, Features, Label sites, or Primer sites to display.

Other links

Working with Map Sites

Format working map display

Circular vs linear molecule maps

Send Map Options

Map toolbar

The map toolbar provides more information on map selections and display options.

- Settings - options for display of map.
- Change sites – selects which type of site is displayed.
- Site Properties
- GoTo Sequence – change to the sequence view to show the sequence around the site.
- Cut at site
- Remove site
- Print List Map – prints molecule information with features and sites.
- Save for Web
- Enhanced view – manage graphics maps for publication.
- Find
    - Find Enzyme Sites (AutoScan)
    - Find ORFs
    - Find Primers that bind
    - Find Sequence Phrases
- Tools

  The options available under tools depends on the type of site selected:

    - Enzyme sites: Suppliers, Isoschizomers, Compatible ends, Wipe all sites, Add site.
    - Features: Export Feature sequence, Create protein, Create fragment, Customize feature style, Shift feature name, Go to start of feature sequence, Go to end of feature sequence.
    - Label sites: Wipe all label sites.
    - Primer sites: Export primer sequence, Create primer, Create product, Change primer name, Wipe all primer sites.

Map Mouse Actions

You can use the mouse to get more information or initiate actions on the working map display.

Features

- Point to a feature name to see feature base pair positions, strand. (Left click with the mouse if hover disabled.)
- Left Click to select a feature. The extent of the feature will be highlighted with a color segment on the map (if enabled).
- Or click on a feature in the scrolling list to select. Click the header control and select Features if the current list shows enzymes or primers.
- Use toolbar buttons to view feature properties, jump to the molecule sequence at the start of this feature, create a protein or fragment from this feature, or export the feature sequence.
- Right Click on a feature and select Customize to change the appearance of the feature on the printed high-resolution graphic map or select Shift Name to move the feature name on the printed map to avoid overwriting.

Enzyme Sites

- Point to an enzyme name to see the base pair position for this site. (Left click with the mouse if hover disabled.)
- Left Click to select an enzyme site. This site turns red and will be highlighted in the Map Sites list on the right.
- Or click on a site in the scrolling list to select the site (now visible in red on the map).
- Use toolbar buttons to view enzyme site properties or information about this enzyme, jump to the molecule sequence at this enzyme site, cut at the selected site, or delete this enzyme site.
- Where enzymes are suppressed (++), point with the mouse to see all enzymes at this location. Use the scrolling list to select a suppressed site.
- Right Click on an enzyme site and select a common function from the menu.

Primer Sites

- Point to a primer name to see the base pair position for the 5' end of the binding site. (Left click with the mouse if hover disabled.)
- Left Click to select a primer site. This site turns red and will be highlighted in the Map Sites list on the right.
- Or click on a site in the scrolling list to select the site (now visible in red on the map).

- Use toolbar buttons to view primer properties, jump to the molecule sequence at this primer site, create a primer or product, export the primer sequence, or delete the primer site.
- Right Click on a primer site and select a common function from the menu.

Label Sites

- Point to a label site to see the base pair position for the start of the label site. (Left click with the mouse if hover disabled.)
- Left Click to select a label site. This site turns red and will be highlighted in the Map Sites list on the right.
- Or click on a site in the scrolling list to select the site (now visible in red on the map).
- Use toolbar buttons to view label site properties, jump to the molecule sequence at this label site, or delete the label site.
- Right Click on a label site and select a common function from the menu.

Circular vs Linear Maps

When a molecule is displayed in the molecule viewer window, the map will be circular or linear, depending on the molecule. You can modify the molecule to change this property, if required.

- **Convert Circular to Linear**

  If you have a circular molecule but want to draw a linear map, cut the molecule after the last base using the cut at base pair option (found on the Clone, Cut menu). This will open the molecule at the origin.

- **Convert Linear to Circular**

  If you have a molecule that has come into the program as a blunt-ended linear molecule, but it should be circular (perhaps the descriptor 'circular' was missing from the file), circularize the molecule using the Process Molecule command found on the Tools menu on the main toolbar.

  If a linear molecule has cohesive ends, you can self-ligate the molecule (if its ends are compatible) using the Ligate option on the Clone menu.

  If you want this change in format (circular vs linear) to be permanent, save the modified molecule to disk.

Working with Map Sites

The Map Sites list is found on the right side of the molecule map display. The scrolling list shows enzyme sites, primer sites, label sites, or features entered for the molecule. Click on the site list header control to select what the list displays. Click on a column header to sort the display. Drag the join between column headings to resize the column width.

Enzyme Sites -- List shows the restriction enzyme name, site position, type of ends produced (5' sticky, 3' sticky, blunt) and an indication if this enzyme is a single cutter (sc), cutting this molecule only once. By convention, the enzyme site is marked at the left end of the recognition sequence.

Primer Sites -- List shows the primer name, site position, and an indication if this primer binds to the complement strand. The primer site is marked at the 5' end of the binding site. Complementary strand primers can be shown in a distinctive color (see format Screen Basics).

Label Sites -- List shows the label name, start, and end base pair positions and an indication if this label is on the complement strand.

Features -- List shows the feature name, start, and end base pair positions, and an indication if this feature is on the complement strand. A color segment can be shown on the map when a feature is selected (see format Screen Basics).

Click to select an enzyme site, primer site, label site, or feature in the list and then click a toolbar button.

View Properties

Click the Site Properties button to see information about the selected enzyme site, primer site, label site, or feature. The enzyme site is displayed in the context of the flanking sequence. The recognition sequence bases are shown in color and the cut position is marked by a vertical or staggered line. Genes in this area are shown as a line below the sequence.

Jump To Sequence

Click the Go To Sequence button to jump to the formatted sequence at the selected enzyme site, primer site, label site, or start of the selected feature. (Use More Actions, below, to jump to the end of the selected feature.)

Enzyme Suppliers

Click Tools with an enzyme site selected and then select Suppliers to view a pop-up display listing the companies that supply the selected enzyme. This supplier information is a part of the REBASE information updated monthly.

Enzyme Isoschizomers

Click Tools, Isoschizomers to view a pop-up display listing the isoschizomers for the selected enzyme. Isoschizomers are defined as enzymes with the same recognition sequence that may cut in the same way (or that may cut differently). This information is a part of the REBASE information updated monthly.

Compatible Ends

Click Tools, Compatible Ends to see the cohesive end (if any) produced by cutting with the selected enzyme at this specific site and a list of enzymes that would produce compatible ends.

Cut at a map site

Click the Cut at Site button to cut at the single selected enzyme site. Additional cut options are available using the Clone, Cut operation.

Delete a map site

Click the Remove Site button to delete the selected enzyme site, primer site, or label site. (Use the Features tab to delete a feature.)

Create...

Select a primer site or feature. Click the Tools button and select option from menu. For primer sites, you can create a primer from the information stored in the molecule file, or you can create an amplified product using the selected primer and another primer from this molecule. For features, you can create a protein by translating the selected feature, or you can create a fragment containing the feature sequence.

Export Sequence

Click Tools, Export Feature Sequence to export the sequence of the selected primer or feature. You can export to a disk file or copy to the Windows clipboard for use in another program.

Additional Functions

Select an enzyme site, primer site, label site, or feature. Then click the Tools button and select option from menu.

Enzyme actions: Select Wipe All Sites to remove all the enzyme sites from the molecule map in one easy operation. Select Add Site to manually add an enzyme site if your molecule does not have sequence data. (If your molecule does have sequence, use the AutoScan or Find Sites options to scan the sequence data for enzyme sites.)

Primer actions: Select Wipe All Primer Sites to remove all of the primer sites from the molecule map in one operation. Select Change Primer Name to change the name of the primer that will appear on this molecule map.

Label site actions: Select Wipe All Label Sites to remove all of the label sites from the molecule map in one operation.

Feature actions: Select Customize Feature Print Style to set custom print options for this feature. Select Shift Feature Name to move the feature name on the printed map to avoid overwriting when needed. Select Go to End of Feature Sequence to jump to the formatted sequence at the end of the feature.

Resizing Site List Display Panel

You can change the width of the map site list, if needed. First, position the mouse at the right edge of the map window and drag the window out to increase its overall width. Next, position the mouse over the line between the map and the map sites list (cursor changes to a 2-headed arrow). Drag this line to the left to increase the width of the site list display.

Adding an Enzyme Site Manually

If your molecule does not have sequence data, you can enter an enzyme site manually. If the molecule has some sequence data at the point where the enzyme site is entered, it will be checked for accuracy before the site is accepted.

For molecules with complete sequence data, it is recommended that you use the AutoScan or Find Enzyme Sites options. The program will scan the sequence and find enzyme sites for you.

- **To manually add an enzyme site:**

1. With the map open on screen, click Tools, Add Site.
2. Enter the enzyme name and the base pair position for the left end of the recognition sequence.

Send Map Options

You can use options to send the map to your printer, preview the map in its printed format, copy the map to the clipboard, export it to a disk file for use in other Windows programs, or save the map in a format that can be used on web pages. For standard Clone Manager graphic maps, you can click Settings to change print options for basic text, colors and styles for features, or the printed size/position and content of the map output. Additional output options are available for enhanced view maps.

- **Print**

  Click the toolbar Print button to send the map to your printer. To change the orientation of the printed page, click File, Other Tools, Print Setup. Click Print Preview to check the appearance of your map before printing, copying to the clipboard, or exporting to a disk file.

- **Clipboard**

  Click the toolbar Send View to Clipboard button to copy the high-resolution graphic map to the Windows clipboard for use in other Windows applications. The styles and colors set for printing will be used for clipboard output.

- **Export**

  Click the Send View to File to export the high-resolution graphic map as a Windows Metafile (*.wmf) or an Enhanced Windows Metafile (*.emf) for use in other Windows applications capable of importing these file formats. The styles and colors set for printing will be used for the exported map. T

- **Save for Web**

  Click the Save for Web button on the map window toolbar to prepare a raster graphic map in the Portable Network Graphics (*.png) format, particularly useful for computer screen display (such as on a web page or in an on-screen presentation). For standard graphic maps, the map size is automatically set to 80%, bp interval numbers are turned off, and the use of italics or small fonts for enzyme names is disabled. (Enhanced views of the graphic map are not automatically modified before export, except for map size.)

  Circular map drawings are prepared in a square drawing area. If your circular map is not an exact circle after pasting or inserting the image, you can adjust the picture height or width to restore the picture to a square, if needed.

  Sending graphic data to a file

  Copying graphic data to the clipboard

Molecule restriction map data

The *RMap* tab shows the restriction map analysis for the active molecule (or the zoomed region of the molecule).

Click the *RMap* tab at the bottom of the window to move to this display. You can print the data in this view window, copy this view to the clipboard, or export the view data to a disk file.

Restriction enzyme recognition site information is listed in table format or displayed in a graphic format. Click the data column header or the Select View toolbar button to change the display style. Click the column headers to sort the data by enzyme name or number of sites. Use the combo box in the toolbar to select a different enzyme list if you want to regenerate the restriction map data using another enzyme list (perhaps a list of enzymes you constructed that are routinely used in your lab).

The toolbar provides more control over the display and allows selecting a different enzyme list or filtering the results. Filtering can be used to identify restriction enzyme(s) that could be used for cloning experiments**.** The toolbar also provides access to REBASE data to get more information such as properties, suppliers, isoschizomers and enzymes with compatible ends.

Click the 'Go To Enzyme' button to enter all or part of an enzyme name to scroll to data for this enzyme, if found.

- **Special Notations**

  Enzymes with many sites (more than 50), have the data display suppressed. Click the small plus sign next to the enzyme name to display or hide this data. Enzyme names shown in color denote possible methylation sensitivity (Click Settings button to enable this feature.)

Methylation Sensitivity Alert

Some enzymes may be sensitive to methylation. If you want to be able to identify these enzymes, click the Settings button and then check the Mark Methylation Sensitivity box. The names of the enzymes with possible methylation sensitivity will be shown in color.

Restriction Map Format

In the restriction map window, click the Settings button to change the enzyme list, data display style, or sort order (see below). You can also make these selections using the toolbar in the restriction map window, but you must use this dialog box to mark enzymes with possible methylation sensitivity or to remember your settings.

Sort Options

You can sort the restriction map data by enzyme name or by the number of sites (or fragments) for each enzyme. Select the sort option in the combo box in the Settings dialog box or click on the column header in the display window.

- **Sort by Enzyme Name**

  Restriction map data is sorted alphabetically by enzyme name.

- **Sort by Number of Sites (or Fragments)**

  Restriction map data is sorted by the number of recognition sites found for each enzyme (or the number of fragments that will result when cut by each enzyme). Enzymes with one recognition site or fragment are listed first. Within each number grouping, enzymes are listed in alphabetical order.

- **Memorize Settings**

  If you always want to use a specific display style, click the Settings button, and then select the display style in the Format Restriction Map dialog box. Click the Memorize button to remember this selection. These settings will be used whenever you initially open a restriction map window.

Restriction Map Display Styles

You can display restriction map data in the styles described below. To select a style, click on the data column header in the restriction map window and select a style from the menu offered or click the Select View toolbar button.

List of Recognition Sites

Shows restriction map data in a table format. The enzyme name is shown at the left and then the number of sites for this enzyme. The base pair position for each recognition site for this enzyme appears in ascending base pair order.

Map of Recognition Sites

Shows a graphic representation of the restriction map. Each enzyme on the list selected that has sites in this molecule is listed. The enzyme sites are marked by diamonds at their approximate location relative to the molecule map, shown in the top of this window.

Fragment Sizes

Shows restriction map data in a table format. The enzyme name and the number of fragments is shown at the left, and then each fragment (in bps) produced by cutting with this enzyme. Fragments for each enzyme are listed in descending size order. (Not available for zoomed molecules.)

Gel View of Fragments

Shows a graphic representation of a gel produced by cutting with each enzyme. The enzyme name and number of fragments is shown at the left and horizontal lanes are marked. A logarithmic scale (bottom of the simulated gel view) relates to fragment size in bps. A vertical bar shows the approximate position at which a band would appear on a gel for a fragment of the size predicted by cutting with the enzyme. (Not available for zoomed molecules.)

No Cutters

List of all the enzymes on the specified enzyme list which do NOT cut the molecule. The list is in alphabetical order, reading from left to right. Filter settings have no effect on the list of non-cutting enzymes.

Restriction Map Toolbar

Use the toolbar buttons to view supplier, isoschizomers or compatible ends information for the selected enzyme, or take action to cut with this enzyme, enter sites found to your molecule map, or build a user list from a filtered data set. Use the Filter button to set options to show you just the information you need.

Enzyme properties and action buttons work on the enzyme that is selected.

- **Settings** – Format restriction map.
- **Select View**
    - List of recognition sites – enzymes are listed on separate lines with number of sites and a list of their recognition site positions. By convention, sites are listed at the position of the left end of the recognition sequence.
    - Map of recognition sites – enzymes are listed on separate lines and the positions of the sites are shown graphically as open dots within the extent of the molecule. An iconic map of the molecule shows genes and regions.
    - Fragment sizes (in bps) – enzymes are listed on separate lines with a list of the sizes of cut fragments.
    - Gel view of fragments – enzymes are listed on separate lines and cut fragments are shown as they would appear on an agarose gel.
    - No cutters – lists enzymes that do not cut the molecule.
- Filter

  Turn on, or off, the filter that allows showing only the enzymes that match your requirements. Use to find enzymes that could be used for a cloning experiment.
- **Enzyme List Selection**

  By default, the Commercial (Main) enzyme list is used to generate the restriction map data, but you can select any other enzyme list and regenerate the data. To select a different enzyme list, use the Enzyme List combo box in the Settings dialog box or the restriction map window toolbar.
- **Go To Enzyme**

  Enter the name of the enzyme to select in the display list.
- **Enzyme Properties**

  The pop-up display shows the recognition sequence, cut position, overhang created, and possible methylation sensitivity (if known).

- **Enzyme Suppliers**

  The display lists the companies that supply this enzyme. You can get up-to-date information for enzyme suppliers from the REBASE web site (http://rebase.neb.com) if needed.

- **Isoschizomers**

  The display shows enzyme isoschizomers. (Isoschizomers are defined as enzymes with the same recognition sequence. They may cut in the same way, or they may cut differently.)

- **Compatible Ends**

  The display shows cohesive ends (if any) produced by cutting with this enzyme and a list of enzymes that would produce compatible ends. For enzyme recognition sites with ambiguous bases, each possible cohesive end and compatible enzymes are shown.

  Enzyme information is a part of the REBASE system and is updated when you update your enzyme lists. Update monthly to make sure you have the latest information about commercially available enzymes.

- **Tools** – Enzyme actions

  You can select an enzyme and cut with this enzyme or enter the site(s) to your molecule map or build a new user enzyme list.

  After adding sites to your molecule map or creating new molecules from cut fragments, remember to save the modified molecule files to disk if needed for further use.

Restriction Map Filter

In the restriction map window, click the Filter button to set filter options to limit the data displayed to just what you need at this time. You can filter based on how enzymes will cut your molecule or based on enzyme characteristics.

When a filter has been applied, the Filter toolbar button will be depressed. Click the depressed Filter button to remove the filter.

Filter by Cut Information

You can set filter options to limit the data display to enzymes that cut a specified number of times, or to show only enzymes that cut inside a specified region and/or enzymes that cut outside a specified region. You can use both cut region filter settings to identify two cut zones.

Cut N times

Check the box for this option and then set the value for N, where N is the number of times the enzyme cuts the molecule. Indicate if the number of cuts should be equal to, less than or equal to, or greater than or equal to N. Example: N equals 3 displays only the enzymes that cut exactly 3 times. N < or = 3 displays enzymes that cut 3 or fewer times.

Cut Outside Region

Check the box for this option and then enter the two base pair position numbers for No cuts between bps (start) and (end). Use this option to designate a region you would like to protect -- enzymes that cut in this region will not be included in the display.

Cut Inside Region

Check the box for this option and then enter the two base pair position numbers for Must cut between bps (start) and (end). Use this option to specify the minimum and maximum base pair positions within which you need an enzyme to cut. For enzymes which cut at multiple sites, if any cuts fall within the region specified, this enzyme and all its recognition sites will be included in the display.

Filter by Enzyme Characteristics

You can set filter options to show only enzymes that produce specific end types (like 5' overhang) or that have a specific recognition element size.

Ends produced by cut

Use the check boxes to indicate that you want the display to include data for enzymes that produce ends with a 5' overhang, ends with a 3' overhang, or blunt ends. You can select more than one option. Leave all the checkboxes blank to indicate that you want to accept any type of end.

Recognition element size

Use the check boxes to specify that the data display should include enzymes with recognition elements that are >6 bases, or 6, 5, or 4 bases in length. You can select more than one option. Leave all the checkboxes blank to accept recognition elements of any size.

Specifying Two Cut Zones

You can use the two restriction map filter options Cut Outside Region and Cut Inside Region to construct two cut zones in your molecule. The filtered restriction map display will show enzymes that cut within the two cut zones specified.

1. Check the Cut Inside Region box.
2. Enter the outermost base pair positions for the left [A] and right [B] ends of the region within which you want both cut zones (see top line of diagram below).
3. Check the Cut Outside Region box.
4. Enter the base pair positions for the left [C] and right [D] ends of the protected region in the center, between the two cut zones (see middle line of diagram below).

You have now identified two regions where enzymes should cut (see bottom line of diagram below).

```
---|A//////////////// cut inside /////////////////B|---
////////////|C-- cut outside (protect) --D|////////////
   |xxxxxxxx|                             |xxxxxxxx|
    cut zone                               cut zone
```

Restriction Map Enzyme Actions

While you are looking at the Restriction Map display, you can select an enzyme and cut with this enzyme or enter the sites to your molecule map. If you are looking at restriction map data that has been filtered to find just the enzymes you need at this time, you can enter enzyme sites to the molecule map for all or some of the enzymes shown, or you can use these enzymes to build a new user enzyme list.

Cut with an Enzyme

Click the Tools button and select Cut with Enzyme to initiate the cut operation using the selected enzyme. A copy of the molecule will be cut at all of the sites found for this enzyme, resulting in one or more fragments. You can enter fragments to the molecule list for immediate or later use.

Enter Sites for One Enzyme to Molecule Map

Click the Enter Sites to Map button to enter recognition sites for the selected enzyme. All the recognition sites found for this enzyme will automatically be added to the molecule map.

Enter Many Sites

Click the Enter Many Sites button to add enzyme sites to the molecule map for multiple enzymes in one operation. Mouse-click to tag enzymes in the list that you want to enter sites for. Click OK to add all the sites to the map for all of the enzymes selected. (Available for filtered data.)

Build User List

Click the Build User List button to create a new user enzyme list using the enzymes shown. Mouse-click to remove enzymes from the list if you do not want to include them in your new user enzyme list. (Available for filtered data, including the list of No-Cutters.)

After adding sites to your molecule map or creating new molecules from cut fragments, remember to save the modified molecule files to disk if needed for further use.

Viewing Sequence Information

To view sequence data, open the molecule viewer window, and click the Sequence tab at the bottom of the window. You can view the complete sequence of your molecule or just a part of the sequence and you can select the format for the sequence display.

You can copy a portion of the sequence to paste into another application or you can edit the sequence, if needed. You can also find places in the sequence that match a search string you enter or jump to a specified base pair position in the sequence (see below).

You can also select (highlight) a range of bases and make this sequence into a molecule feature. If you are using Clone Manager Professional, you can also make the selected sequence into a primer.

Click the Settings button to change the appearance of the sequence data. You can set the display style, select text colors, set print options, or select a partial sequence to view or print.

Click the My Style button to format the sequence in your favorite style. (To set up this feature to remember your individual style preferences, set the style options on the Settings, Style dialog box, and then click the Memorize button.)

- **Selecting sequence**
- **Cut, Copy or Paste**
- **View Sequence Toolbar**
- **Find Sequence**
- **Sequence Display Style**
- **Partial Sequence**
- **Editing Sequence**
- **Printed Output Options**

Selecting sequence

If you are having trouble selecting sequence with the mouse, you can click the Edit button (which will position the cursor on the screen), and then select the sequence using the keyboard and copy the selected sequence to the clipboard. When complete, abandon the edit operation so you do not modify your sequence.

Cut, Copy, or Paste Sequence Data

You can use the standard Windows Cut, Copy and Paste operations on your sequence data. Sequence is stored in the program single-stranded and all operations work on sequence in this way.

- **To copy sequence from the molecule viewer window**

1. Select (highlight) the sequence to be copied using the mouse or use Control+A to select the entire sequence.
2. Look in the toolbar area to see the range of bases selected.
3. Click the Copy button to copy the selected sequence to the clipboard.

- **To paste sequence into the molecule viewer window**

1. Copy single-stranded sequence to the clipboard.
2. Open the molecule viewer window and click the Sequence tab.
3. Click to position the edit cursor where you want to paste the sequence currently on the clipboard.
4. Click the Paste button.

The program will check for valid bases and paste the sequence into the location specified. The sequence will be displayed in the format selected.

Sequence Toolbar

- **Sequence Format** - Change the display style, Colors, Print Options and Partial Sequence.
- **My Style** – Reset display to your preferred style as set in the Sequence Format dialog box.
- **Go To BP Number** – enter the position to quickly scroll to the part of the sequence containing the base pair number specified.
- **Edit** – enables editing of the sequence.
- **Accept changes (Update)** – update the molecule to save any changes to the sequence. If you have other views that depend on the molecule's sequence, these displays will be updated. For analysis and design operations you will be alerted that the results may no longer be valid, and you should redo the operation.
- **Discard changes (Cancel)**
- **Cut** – while editing the sequence, you can use the mouse to select a region of sequence to cut/remove. The sequence bases will be copied to the clipboard and can be inserted into another molecule or external program.
- **Copy** – like Cut but does not remove the sequence.
- **Paste** – inserts text from the clipboard into the sequence at the current cursor position. The sequence can come from an external source and will not allow invalid sequence bases to be inserted.
- **Undo** – undoes the last sequence change.
- **Tools**

  The Tools drop down menu provides options for finding sequences and converting selected sequences into features or primers.

  - Find – find and move to a sequence that may include mismatches or ambiguous bases.
  - Find next – moves to the next position that matches the sequence entered.
  - Enter as feature

    Select (highlight) a range of bases and then select this option to make the selected sequence into a feature (gene, region, or marker) to add to your molecule's feature table.

    When you use the mouse to select the sequence bases, you can check the range selected in the upper area, just to the right of the toolbar buttons.

Enter the feature information. This results in the same action as adding a feature by clicking on the Features tab and then adding a new feature to the features table display.

- Create primer

  Select (highlight) a range of bases and then select this option to make the selected sequence into a primer.

  The primer viewer window will open, and you can view the new primer and its evaluation. You can modify the primer, if needed, and save it to disk or add it to your primer collection. This is the same as creating a new primer using Design Primer, Create from Molecule.

- Create complementary primer – converts the selected sequence into a complementary strand primer.

Sequence Format

Enables selection of multiple display options for Style, Color, Print and selection of partial sequence regions. These options are available in a multi-tab dialog box.

- Style
  - Enzyme Sites – select sites from the Working Map, AutoScan list or the restriction map.
  - Primer Sites – options include Color Bases/Arrows and Two Colors.
  - Sequence – single stranded or double stranded, lower case or upper case, and blocks or 10 or 3.
  - Sequence Labels – select from multiple options for how feature labels are displayed.
  - Translation – displays the translation of the DNA sequence. Select which frame(s) are shown. Select how to display the translation.
  - Annotations – Select whether to only show gene features or to show all map features. Translated genes will show gene features with in-frame translation and you can select how the translation will be displayed.
  - Memorize – saves the current settings which can be recalled using the ‘My Style’ toolbar button to quickly set the display to your preferred style.
  - Set Plain – use to quickly reset all styles to the basic settings for single strand sequence displayed on lower case and blocks of 10.

- Colors

  Allows customization of the colors to use for each type of style option selectable in the preceding tab.

- Print Options

  Allows selection of font size and printing in color.

- Partial Sequence

  Allows displaying detailed sequence formatting for a part of a molecule. This is particularly useful to view details of a part of a large molecule such as genomic sequences.

  Show flanking sequence enables you to show the sequence of the full molecule but only show the enhanced sequence formatting for the area of interest. You might use this to show the translated sequence of an important gene without showing annotations for the rest of the molecule.

Find Sequence Tools

You can find places in the sequence that match or nearly match a specific pattern of bases. You can jump to each occurrence found in the sequence display or you can view a report listing all the sites found on both strands of the entire molecule.

- **To start a Find sequence operation**

1. Open the molecule viewer window and click the Sequence tab.
2. Click the Tools button and select Find or use Ctrl + F.
3. Enter the sequence of bases you want the program to search for (click the small button for help with IUPAC ambiguous bases).
4. Use checkboxes to configure the search to permit limited mismatches or mismatches and gaps or insertions.

- **To jump to matches in the sequence**

1. Set the Search Type to **Go To** Sequence Found.
2. Use checkboxes to limit the search to the upper strand or to begin the search at the start of the sequence.
3. Click Find. If a matching string is found, the string of matching bases will appear as selected text (highlighted) in the sequence view.
4. Click Tools, Find Next or use Ctrl + I to locate additional matches.

- **Or, to view a report of all sites found:**

1. Set the Search Type to **Report All Sites** and click Find.
2. View, print, or copy the list of sites found.

- **Advanced search notes**

When you select the option to allow limited mismatches, the number of mismatches (including gaps or insertions, if applicable) is limited to 1 mismatch for each 10 bases in the search sequence. For short search sequences of less than 10 bases, one mismatch is permitted.

When entering the search string, you can use the check box to indicate that you do not want to expand ambiguous bases. You might want to do this to locate an ambiguous base within your actual sequence data.

- **Find Sequence Search Results**

You can search for matching or nearly matching sites in your sequence and set the search type to report all sites found. When the search has been completed, the results will appear in a scrolling list box with a location map display.

The scrolling list shows each site found, with its start position, sequence of bases, and the number of mismatches (*m*) in the sequence shown. Click on column headers to sort by position, sequence, or mismatch number.

For all sites found, the sequence shown is the upper strand sequence. Mismatched bases are shown with upper case (capital) letters, gaps appear as dashes, and inserted bases are shown within parentheses.

The location map display is shown below the list of sites found. Vertical bars above the icon map show the approximate location of the site found and a red arrow marks the site highlighted in the scrolling list.

Add Feature -- click the Add Feature button to enter the highlighted site as a feature for this molecule.

Print List -- print the report of sites found. If you have sorted the list on the screen, the printed list will also be sorted.

Copy List -- copy the report of sites found to the Windows clipboard.

Sequence Style Options

You can set the display style for the sequence data and select the additional information that will appear along with the sequence. (Default colors for sequence text items can be selected on the Colors tab.)

- **To set sequence display style**

1. With the sequence visible on the screen, click the Settings button
2. Click the Style tab if not already selected
3. In the Sequence options area (center), select single- or double-stranded
4. Select block size (10 or 3) and choose upper- or lower-case letters

- **To select additional information**

1. Use check boxes to indicate which items should be displayed with the sequence data
2. For selected items, set related options (see below)
3. Use the Memorize button to remember the settings for future use.

Enzyme Sites -- Enzyme sites are shown above the sequence line. Select the source of the enzyme data: you can use enzyme sites that appear on the working map (Map tab), enzyme sites the program will find using the AutoScan enzyme list, or enzyme sites that are currently shown on the RMap tab.

Primer Sites -- Primer sites are shown above the sequence line. You can select the display style and colors to use for primer sites.

Sequence -- Single-stranded sequence displays the upper strand for the full sequence and the normal or complementary strand (as indicated) for a partial sequence. A double-stranded sequence displays both strands of the molecule, including sticky ends, where appropriate. You can select the block size and case for sequence characters.

Sequence Labels -- Shown on the sequence line itself. You can display sequence labels by coloring the bases in the sequence and/or applying color to the background behind the sequence bases, or you can invert the case of the sequence characters.

Translation -- You can show 3 frames, single frame 1, 2 or 3 (as selected), or all 6 frames of translation below the sequence line. Show amino acid codes in 1-letter, 1-letter with amino acid numbers, or 3-letter styles. Translated display sample.

Annotations -- You can show genes, all map features (genes, regions, markers), or translated genes below the line of sequence. For translated genes, you can select 1-letter, 1-letter with amino acid numbers, or 3-letter amino acid codes.

- **Sequence Text Colors**

  You can select the default color to be used for each type of sequence text that can be displayed and you can set the base color and background color for sequence labels. Colors can help to distinguish one type of text (such as amino acid codes) from another (such as the sequence being translated).

  Colors selected for features (gene, region, marker annotations and sequence labels) will apply to all text of that type. You can use a custom feature setting to distinguish a special feature, if needed.

  To set sequence text colors

  1. With the sequence visible on the screen, click the Settings button
  2. Click the Colors tab
  3. Select the preferred color for each item type

  To set custom feature colors

  1. Click the Features tab to move to the features table
  2. Select (highlight) the feature
  3. Click the Customize Feature button
  4. Select the color for the annotations text for this one feature

Partial Sequences

You can view the complete sequence of your molecule or just a part of the sequence. You can also show a part of the sequence in a distinctive style (perhaps translated or annotated) and the rest of the sequence in a plain format.

- **To view or format part of a sequence**

1. With the sequence visible on the screen, click the Settings button
2. Click the Partial Sequence tab
3. Set the Partial Sequence option button
4. Enter the start and end base pair positions for the part of the sequence you want.
5. Use the small Features button to see where features start and end.
6. Use the checkbox if you want to show flanking sequence

The sequence you identified as the partial sequence will appear in the style you selected on the Style tab. The flanking sequence will appear as plain sequence.

Editing Sequence Data

You can edit the sequence of a molecule by adding or removing bases. You can be in single-stranded, doubled-stranded, or translated mode, but you can only edit the upper strand of the molecule. You can also use the standard Windows Cut, Copy and Paste operations on your sequence data.

- **To edit sequence data**

1. Open the molecule viewer window and click the Sequence tab
2. Click the Edit Sequence button on the viewer toolbar.
3. Update the sequence data (see hints below).
4. Click the Accept Changes (Update) button to complete the changes you made and modify the sequence of the molecule.

You can click the Discard Changes (Cancel) button to abandon your changes and leave the sequence unmodified.

- **Sequence editing**

Use the arrow keys or the mouse to position the cursor where you need to add or remove bases.

  - Enter (type in) bases to be added
  - Use the Delete key or select and then delete to remove bases
  - Use the Undo button on the toolbar to reverse a change you have just made, if needed.

Printed Output Options

You can select the font size to be used for sequence data sent to your printer. The font face will always be Consolas if this is supported in your printer. Otherwise, the font face will be Courier New. (A mono-spaced font is required so that sequence data lines up correctly.)

- **To set printed output options**

1. With the sequence visible on the screen, click the Settings button
2. Click the Print Options tab
3. Select the font size you want to use for your sequence text
4. Use the checkbox if you want to print in black and white only

The approximate number of bases per line for each font size is shown below.

- **For Portrait pages:**

Small (8 pt)	80 bases / line
Medium (10 pt)	60 bases / line
Large (12 pt)	50 bases / line

- **For Landscape pages**

Small (8 pt)	120 bases / line
Medium (10 pt)	90 bases / line
Large (12 pt)	70 bases / line

- **Print in Black and White only**

Translated and Annotated styles will use colors to differentiate between sequence bases and the translation or annotation information. You can use this checkbox to turn off color printing, if desired, forcing all text to black.

Viewing Molecule Features

Found on the *Features* tab in the molecule viewer window. Molecule features appear in the listing with a symbol for the feature type, the feature name, the feature base pair positions, a brief description of the feature, and the feature key, if assigned.

The display consists of a toolbar at the top of the window. An iconic map of the molecule shows the major features. The main part of the display is the list of features. At the bottom are shown the styles for the selected features which show how it will appear on the graphic map and the sequence displays.

Features Toolbar

The toolbar provides functions for controlling the display, adding, editing, and customizing features as well as displaying properties and tools for using the features.

- **Format Features table**
  - Display As – select details or column mode. In column mode you can click the column headers to sort the features table by name, start or end position, descriptive text, or key.
  - Show BP Numbers – in column display mode you can select to show base positions using either Left-Right or Start-End.
  - Show Information-only features – less important features can be set as Information-only so that they will only be shown in the features list when needed.
  - Show simple features map above list – show or hide the iconic map.
  - Show selected feature styles in the lower area – show or hide the styles.
  - Memorize – click to remember the settings so that they will be used the next time you open Clone Manager.
- **Filter**

  Used to restrict the features list to find the most relevant features. Particularly useful for large molecules or those with a highly detailed features list.

  - Feature Type – select to show: Genes, Markers, Regions, and/or Labels.
  - Location – use to select features based on their position in the molecule
  - To show or hide, Information only features, use the Format Features table button.
- **Feature Properties**

  Shows properties of the selected feature appropriate to the type of the feature.

  - Genes – shows name, position, description, size, and translation including molecular weight, isoelectric point, amino acid composition and codon usage.
  - Regions – shows name, position, description, base composition.

- Markers, Labels, and Info-only – shows name, position, description.

- **GB Location/Qualifiers**

  Shows a popup display with more detailed information about feature qualifiers including feature location data (base pair positions) and qualifiers available for this feature. This is particularly useful when molecule files are imported from GenBank, or EMBL. Also shown are base pair positions for join segments, when applicable.

  You can also use a special mouse cursor to click on the feature symbol (to the left of the feature name) to display this pop-up information. Click outside the pop-up to dismiss.

- **Customize Feature**

  Allows customization of the display style for the selected feature.

  Below the features table list, a boxed area is used to show the styles set for the selected (highlighted) feature. If you have not set any custom styles for this feature, the settings will reflect the default values for printed map features or for annotated sequence display. If this box is not shown, click Settings and set to display.

  - Map settings – color and display style for the feature on the graphic map.
  - Sequence settings – change color of the feature when displayed on the annotated sequence view.
  - Select the 'Default' option to remove custom styling.

- **Go To Sequence**

  - Go To Sequence Start – change to the sequence display and move to show the region around the start of the feature. This is the default action for the toolbar button.
  - Go To Sequence End – click the down arrow and then select from the drop-down menu. Changes to the sequence display and moves to show the region around the end of the feature.

- **Edit Feature**

  Allows editing of the properties of the selected feature.

  - Name – will be used to display the feature on the graphic name and on the annotated sequence displays.
  - Feature Key – the GenBank/EMBL key to annotate the feature (optional).
  - Description – one line of text to describe the feature (optional).
  - Draw As Type – sets how the feature will be displayed on the graphic map.

- **Add Feature**

  Allows adding a new feature using the same dialog interface used by Edit Feature.

- **Remove Feature**

  Removes the selected feature. You will be asked to confirm the deletion.

- **Tools**

  Provides several useful functions:

  - Make Protein

    The program will translate the gene feature to create a protein, starting at the base pair position for the start of the gene, on the strand specified, and continue until it reaches the end base pair position (or encounters a stop codon). If the codon start for this feature has been set to a value other than 1, translation will begin at the start codon specified.

    When the protein has been created, you can enter the new molecule to the Molecule List, changing the name or description as needed. You can use this new molecule now or save it to disk for future use.

  - Make Fragment

    The program will prepare a new molecule from the fragment of sequence containing the feature. If the feature is on the complement strand of the original molecule, you can select the orientation of the new molecule.

    You can choose to use the feature orientation (feature will be normal strand after the molecule is inverted) or you can maintain the orientation of the original molecule (feature will be on complement strand). You can also include extra bases at each end of the feature sequence (enter the number of bases you want to add).

    When the fragment has been created, you can enter the new molecule to the Molecule List, changing the name or description as needed. You can use this new molecule now or save it to disk for future use.

  - Export Feature Sequence

    Extracts the sequence of the selected feature. You can copy the result to a file or to the clipboard.

Feature Draw as Types

Feature Draw as Types are used to tell the program how to draw the feature on the molecule map or show the feature on the annotated sequence. Molecule feature types include Genes, Regions, Markers, Labels, and Information-only features. You can set the style and color to be used for each feature type when printing your molecule maps or annotated sequence.

- **Genes**

  Genes usually represent translated sequences, have a start, and end base pair position, and can be on the complementary or normal strand. Genes appear on the molecule map inside/below the mapline and can be displayed in the annotated sequence.

- **Regions**

  Regions are zones of special significance for the user, have a start and end base pair position, and can be on the complementary or normal strand. Regions appear on the molecule map on the mapline and can be displayed in the annotated sequence.

- **Markers**

  Markers have a direction, occupy one base position, and can be on the complementary or normal strand. Markers appear on the molecule map on the mapline and can be displayed on the annotated sequence.

- **Labels**

  Labels are areas within the sequence that are of special interest to the user and have a start and end base pair position. Labels can be shown on the annotated sequence, but do not appear on the molecule map.

- **Info only**

  Information-only features are not shown on the molecule map or on the annotated sequence. (They do not appear on the Features Table list unless you click the Settings toolbar button to set the option to show this feature type.) Use for features where you want to preserve the information about the feature but do not want it to clutter your map or features table.

Segmented features

Genes and regions may consist of joins. In the features table, the symbols for these segmented features appear like the symbols used, but with the symbol shown in two segments. To see the base pair positions for the join segments, move the mouse over the feature symbol (cursor changes to rectangle shape) and click. Click outside the pop-up box to dismiss.

Using the Features Table Filter

If your molecule has a very large number of features, you can click the Filter button to set filter options to transiently limit the data displayed in the features table by location or feature type. (The feature table is displayed on the Features tab in the molecule viewer window.)

When a filter has been applied, the Filter toolbar button will be depressed. Click the depressed Filter button to remove the filter.

Use the toolbar Info Features button to show or hide information-only features.

- **Filter by Feature Type**

  Use the check boxes to indicate the feature types that you want to display. You can select more than one option. Leave all of the checkboxes blank to indicate that you want to list all feature types.

- **Filter by Location**

  Check the box for this option and then enter the two base pair position numbers that define the region where you want the features listed. All features that are found within this region will be shown, even if the feature is not fully contained within the region.

Entering Feature Information

Feature information is found on the Features tab in the molecule viewer window. You can add or edit information about each of the features for your molecule:

- **Feature name** -- Enter a name for the feature. Feature names will usually appear on molecule maps so a short name is preferred. This field may also be left blank.
- **Feature Key** -- Select the appropriate feature key from the drop-down list of keys supported in GenBank/EMBL file formats. The list will initially be limited to your designated feature keys. Click More... to see all of the possible feature keys. (Click File, Preferences, Import Features to modify your list of designated features.)
- **Description** -- Enter additional information about the feature, if desired. This information does not appear on map displays.
- **Codon Start** -- Select the offset (1, 2 or 3) for the first complete codon of the feature. Use for features that can be translated.
- **Draw as Type** -- Select the feature type Gene, Marker, Region, Label, or Info only. The feature type is used to tell the program how to display this feature on the molecule map or annotated sequence.
- **Style** -- Select the style number for the Draw as Type for this feature. You can have 3 different styles for the major feature types (genes, regions, markers).

- **Location** -- Enter the start and end positions for genes, regions, labels, and information-only features. Enter the single base pair position for a marker. If the feature is segmented, you can enter the join segments (see below). If join segments are used, the start and end position of the feature is calculated from the segment positions.
- **Complement checkbox** -- Check this box if the feature in a circular molecule is on the complement strand (counterclockwise orientation). For linear molecules, this information is calculated by using the start and end base pair positions.
- **Join Segments** -- Click the Add button to enter a new join segment. Enter the base pair positions for the segment. Click the Edit or Del buttons to modify or remove the selected segment.
- **Qualifiers** -- Standard qualifiers have been generated based on the Feature Key selected and mandatory qualifiers are identified (if any). Click to select a mandatory qualifier and click Edit to view the list of possible qualifier values or type in text, if appropriate. Do not enter quotation marks. Click the Add button to enter a new qualifier. Select the qualifier name from the list provided and enter the qualifier value. Click the Edit or Del button to modify or remove the selected qualifier.

  After you have added qualifiers to a feature, you can move a qualifier up or down in the qualifier list. Right-click on the qualifier you want to move and select the appropriate action. You can also use options on the right-click menu to copy the text (Value) in the selected qualifier and paste this text into the Feature Name or Description field, replacing the text currently in that field.

Viewing Molecule Information

Details about the molecule are found on the *Info* tab in the molecule viewer window. You can view or edit the molecule name, description, or notes. You can also view molecule properties information. For DNA molecules, you can see base composition. For protein molecules, you can see molecular weight, pI, and amino acid composition.

Click the Edit toolbar button to update the molecule name, description, or notes. Click toolbar buttons in this window to add author information to the notes field, change the starting base pair number for this molecule, change the translation table, or access GenBank annotations (header information) if the file is a GenBank or EMBL format file.

Molecule descriptive information:

- **File Name and Location** – shows where the molecule is stored and the date when the file was last modified. These entries will be blank if the molecule has not been saved to disk.
- **Molecule name** -- appears in the Molecule List, on program screens and in maps and other printed output.
- **Size** -- if the molecule does not contain actual sequence data, you can change the size by typing in the number of bases. If sequence is present, you will have to edit the sequence (add or remove bases) to change the size of the molecule.
- **Description** -- appears on many types of printed output and can optionally be printed on molecule map pages.
- **Translation table** -- assigns the genetic code used to translate this molecule. By default, the Standard Code (translation table = 1) is used unless CDS features for this molecule contain qualifiers indicating an alternate translation table.
- **Notes** -- more detailed information about the molecule. Can be keywords, history of construction, comments. This can also be printed on molecule map pages.

Information Toolbar

- **Edit**

  Allows updating the following informational data for the molecule:

  - Molecule Name – the name shown in the Molecule List and all analysis displays.
  - Size – Can be edited if the molecule does not contain any sequence data.
  - Description – a one-line description of the molecule.
  - Notes – multi line annotation of the molecule.

- **Author Stamp**

You can add an author name, date, and notebook reference to the beginning of the Notes field for your Clone Manager molecule (*.cm5) files. This can help you to identify the source of a file or the location of needed documentation. As part of the Notes field, this information can be printed on the page with your molecule map or written in a DataBook entry.

To avoid having to re-enter your name for each file, click File, Preferences, Basic and enter an author name that should automatically appear in the Author Name field when this option is selected.

- **Base Renumber**

  You can designate a number other than 1 to use as the first base pair number for the sequence of the active molecule. You might want to renumber your sequence if you are using part of a larger molecule and want to maintain consistent base numbers for features or enzyme sites.

  Notes: Negative numbers are permitted, but zero is not a valid base number.

- **Translation Table**

  Select the default translation table to be used for all gene features in this molecule. Individual features can override this setting if appropriate.

- **GenBank Annotations**

  Annotations refer to the header information found in a GenBank or EMBL format file. In addition to reference information, you may find some helpful feature information or comments about the file that you may wish to review.

  If you save a databank file as a Clone Manager file, the annotation information is not saved as a part of this file format, but the program will automatically save a link to the original databank file so that you can later find this information.

- **Tools**

  Used to synchronize changes to shared Clone Share files.

Zoom Molecule

You can select to display only a part of a large molecule. This is often helpful if there are many features, with a very cluttered map, or if the features are very small relative to the size of the molecule and cannot be displayed clearly.

- **To zoom a molecule:**

1. Select the molecule so the molecule viewer window opens.
2. Click the **Zoom** button (left end of tab bar at bottom of window).
3. Select option Zoom to Show Region Specified.
4. Enter the start and end bp positions for the region to display.

The data display is regenerated, and the window title identifies the zoomed region.

Map view -- shows features and enzyme or primer sites for the region specified, always as a linear map. A mapline bar and colored box at the top of the map window show the approximate location of the zoomed region. If features are not fully contained within the zoomed region, a set of dots before or after the feature name indicates that the feature extends further in that direction.

Restriction map view -- shows restriction enzyme sites found within the zoomed region. Gel view of fragments and Fragment size displays are not available for a zoomed molecule.

Sequence view -- includes only the sequence of bases in the zoomed region.

Features table view -- lists features that are found within the zoomed region, even if they are not fully contained within this region.

Displaying Primer Sites

Primer binding sites can be added to your molecule file from the Find Primer Sites Results display, the Link to Molecule results display, the primer design search results display, or from the primer viewer window for a specific primer. Once these sites have been added to the molecule file, they can be displayed on the molecule map or in the formatted sequence.

Map Display

To view primer binding sites on the molecule map:

1. Open the molecule in the molecule viewer window and click the **Map** tab
2. Click the Change Sites toolbar button to switch to primer sites.

Primer binding sites entered for this molecule will appear on the molecule map (in place of enzyme sites). The map sites list (right) shows the primer name and the 5' end of the primer binding site. Click the column headers to sort by primer name or position.

You can select a primer site and then use this primer. Toolbar buttons in the map display window can be used to export the primer sequence, create the primer, use the primer to create an amplified product, change the primer name, or view the primer site properties.

Sequence Display

To display primer binding sites on the formatted sequence:

1. Open the molecule in the molecule viewer window and click the **Sequence** tab
2. Click the Settings button to select the sequence style
3. On the Style tab, use the checkbox to turn on the display of Primer Sites and select the display style.
4. On the Color tab, select the text and background colors for primer sites

Primer sequences will appear above the molecule sequence in their binding location. Lower case letters within the primer sequence indicate a mismatched or unbound base.

Data Lists

Clone Manager maintains lists of active data to enable quick access to data already loaded into the program.

- **Molecules**

  Shows a list of the molecules that have been loaded into the program as well as molecules created using cloning operations. Allows quick selection of molecules for viewing or performing cloning or analysis functions.

- **Primers**

  Shows a list of the primers that have been loaded into the program as well as primers created using primer design functions or cloning operations that use primers. Allows quick selection of primers for viewing, analyzing, or for use in primer operations.

- **Enzymes**

  Shows restriction enzymes available for classic cloning operations or restriction mapping. Allows for custom subsets of enzymes that can be used for special, or repeated, functions. The Commercial Main list contains a list of enzymes that are available commercially and would be a good choice for planning experiments.

- **Sequence Phrases**

  Shows collections of sequence phrases that can be used to search for motifs.

- **DataBooks**

  Is a simple, built-in database you can use to organize and store information about your molecule or data files, including primers or alignment results. You can think of it as a list of the files that you want to keep track of and just enough information about each file to let you find and open the file when you need it.

Molecule List

Access using the Molecule List button on the main toolbar or the main View menu. The molecule list contains molecules loaded into the program or created during the current work session. Newly created or modified molecules that have not been saved are marked with a diamond symbol.

Use this list to identify the molecule to use for the next operation, remove molecules no longer needed, or view additional information about the molecule. You can also look for additional files to add to the molecule list (see below).

You can sort the information in the Molecule List (click on column header), and you can resize the columns as needed to view the data most important to you (drag the join between column headings). If molecules have been loaded in groups, you can temporarily ungroup these files and then regroup them as needed.

- **Selecting a Molecule**

  You can identify the molecule to use for the next operation or you can open the molecule viewer window for any molecule loaded in the program. Click on the name of the molecule of interest and then click Select. You can also select groups of molecules for alignment operations.

- **Loading another Molecule**

  Click Browse to access the File Open box to find and load another molecule file, if needed. The file will be opened, entered to the Molecule List, and selected for use. Click DataBook to access the built-in database to find and open a specific file. You can search DataBook entries, as needed, to find just the file you are looking for.

- **Handling groups of molecule files**

Enter to Molecule List, AutoScan

When you create a new molecule by some action you have taken, you will need to enter this molecule to the Molecule List. You can enter a name and description for the new molecule. Following certain cloning operations (like using the PCR cloning wizard or the Gateway cloning wizard), you can indicate how you want to manage enzyme sites for the new molecule.

Once on the Molecule List, the new molecule will be available for use in the program and can be saved to a disk file for later use.

- **Molecule Name**

  Enter a name for the new molecule. This name appears on the Molecule List and is used to identify the molecule on all data screens and printed output.

- **Molecule Description**

  Enter a short description of the new molecule. This description will appear on some program screens and on printed output.

- **Process Enzyme Sites (AutoScan)**

  Enzyme sites for the new molecule can be auto scanned using your designated AutoScan enzyme list, looking for enzyme sites (single cutters or all on user list, as specified in the AutoScan settings). The enzyme sites will be entered, replacing the existing enzyme sites.

  You can clear the AutoScan checkbox, instructing the program to rebuild the map sites list by rescanning with enzymes currently listed. This is the procedure used following a standard Clone, Ligate operation, but it may result in enzyme sites that are not single cutters for the newly created molecule.

Enzyme Lists

The Clone Manager Program uses enzyme lists to hold groups of restriction enzymes that can be quickly selected for a cloning purpose. Restriction enzymes include the standard REBASE repository of enzyme information and can be extended by user specific custom enzymes.

Information about each enzyme can be viewed. You can also search or print enzyme lists and you can build or modify user enzyme lists to meet your individual needs. You can also update your master enzyme file monthly using REBASE data files so that you are always up to date with new enzyme information.

Left mouse click in the left-hand panel to select an enzyme list to display its content in the right panel.

Right click on an enzyme list in the left-hand panel to show a set of actions on that list – the actions are the same as available on the toolbar 'Folder options' described below.

Viewing Enzyme Lists

Click the Enzyme Lists button on the main toolbar, or use the main View menu, to open the enzyme list resource.

Click on an enzyme list in the left panel to display the contents of this list in the right panel. For each enzyme in the selected list, the name, recognition sequence, cut position, ends produced and overhang bases are shown. Click on column headers to sort by the values in that column.

Select an enzyme by clicking on the enzyme in the list in the right panel. Press a letter key to quickly move to enzymes in the list that begins with that letter.

If the recognition sequence is longer than the space provided and you want to be able to see the entire sequence of bases, double-click on the column header divider (right after the word Recognition). You can now use the horizontal scroll bar in this list box to scroll to the right to see the rest of the recognition sequence.

Enzyme lists are divided into the following categories:

- **Program Lists** – Most commonly used lists, program pre-defined lists.
- **My User Lists** – lists you can create to make collections of enzymes you routinely use.
- **Shared Lists** – lists that are shared among lab, or group, members.

If you have multiple computers running Clone Manager, you can share a current master enzyme file, a set of helpful user lists, or a list of special enzymes. To access shared resources, click File, Preferences, Share Data, and enter the location (complete path and folder) where you can access shared enzyme lists. Shared lists are read-only resources and cannot be modified.

- **Special Enzymes** – Enzymes that you enter yourself, perhaps not commercially available or historic.
- **On The Fly** – quick creation of a list with special characteristics, such as 6 base cutters with 5' sticky ends.

Enzyme List Toolbar

- **Folder options**
  - Add a new (empty) user list.
  - Import an existing list.
  - Rename a list (for user defined lists).
  - Delete a list (for user defined lists).
  - Export a list (for user defined or shared lists).
- **Rebase**

  Enables updating your master enzyme list from the REBASE online repository.
- **Enzyme Info**

  Provides several types of information about the selected enzyme:
  - Enzyme suppliers
  - Isoschizomers – enzymes with the same recognition sequence but may have a different cut position.
  - Compatible ends – enzymes that create the same cohesive end but may not have the same recognition sequence. Useful for classical cloning since the overhangs can be ligated.
  - Enzyme properties – lists name, recognition sequence, cut position and overhang.
- **Copy enzyme**

  Copy the selected enzyme to a user-defined enzyme list. Provides an easy way to build user lists.
- **Print list**
- **Build user list**

A wizard for building a new user list. You can type in the names of a set of enzymes or point-and-click to select enzymes from a list.

- **Edit user list**

  Allow editing an existing user defined enzyme list. Select the list

  - Unlock list – lists are normally locked to prevent inadvertent changes. Use the Unlock List option to enable editing.
  - Add enzyme – type in the name of the enzyme or select from the list.
  - Delete enzyme – deletes the selected enzyme.

  Can also be used to manage the Special Enzymes list which are enzymes that are not yet included in the master REBASE enzyme list.

  - Add enzyme – type in the name of the enzyme, its recognition sequence and cut position.
  - Delete enzyme – deletes the selected enzyme.
  - Edit enzyme – update the selected special enzyme.

- **Ambiguous bases**

  Shows a popup box listing the IUPAC ambiguous base codes.

Enzyme List Locations

When you install Clone Manager, the master enzyme file is copied to the installation directory. When you do a REBASE update, the new enzyme master file will be written to your home directory. On start-up, the program will check your home directory and shared location (if any) for a new master enzyme file and use the latest file found. Your user-specific enzyme information (user lists, special enzymes) will always be stored in files in your home directory.

Program Enzyme Lists

Clone Manager constructs and maintains three general-purpose enzyme lists. Each time you update your master enzyme file using the latest REBASE enzyme file, these lists will be regenerated. You cannot edit these program-defined enzyme lists.

Commercial (Main)

This enzyme list is a subset of the master enzyme list (All Enzymes). Duplicate enzymes with the same recognition sequence and cut position are eliminated. This list is the program default for most enzyme operations. This list contains the major, commercially available enzymes recognized by the Clone Manager Program (approximately 240 enzymes).

All Enzymes

This is the master list containing all the enzymes recognized by the program. There are more than 650 enzymes in this list, including many isoschizomers for the common commercially available enzymes found in the Commercial (Main) List.

Scanner List

This list contains the major, commercially available enzymes with recognition sequences of six or more bases. This list is the program default for the AutoScan operation.

User Enzyme Lists

You can create user-defined enzyme lists to hold the enzymes you routinely use or enzymes you plan to use for a specific task. Clone Manager provides a wizard to help you through the process of building new user enzyme lists.

- **To create a new user enzyme list:**

1. Click menu View, Enzyme Lists button
2. Click the Build User List toolbar button.
3. Follow the wizard instructions.

Once your enzyme list has been created, you can later add or remove enzymes, as necessary, and you can rename or delete a user enzyme list. You can add enzymes that the program recognizes or Special Enzymes (that may not be commercially available) that you define.

While viewing filtered restriction map data, you can build a new user enzyme list using the enzymes shown on the display (including the list of NoCutters). Just click the toolbar button Build User List in the restriction map window to get started.

- **To add an enzyme to an existing list:**

1. Click Enzyme Lists toolbar button
2. In the left panel, select the user-defined enzyme list that you want to add enzymes to
3. Click the Edit User List toolbar button and select Unlock list to allow changes to this list.
4. Next select Add Enzyme and type in the enzyme name in the data entry box provided. An enzyme lookup list box will help you get the name exactly right.
5. Click OK to add this enzyme to the selected user enzyme list.

To remove an enzyme from the unlocked user list, select (highlight) the enzyme and then select the Delete Enzyme option on the Edit User List menu.

- **Rename or Delete a user list**

With the Enzyme Lists open, select the user-defined enzyme list in the left panel. Click the Folder Options toolbar button or just right-click on the list and then select the Rename or Delete option.

- **Export or Import a user list**

You can use the Export option to prepare a user enzyme list file (*.dat) on disk. You might want to do this to provide a list to a colleague or to move your list from one computer to another. To export a user list, click Clone, Enzyme Lists, and right-click on the user list. Select the Export option and provide a file name and location.

The Import option will read in user enzyme list data from a file previously exported from the Clone Manager program. To import a user enzyme list, right-click on My User Lists and select the Import option. Identify the file (*.dat) on disk that you want to import.

Special Enzymes

You can enter detailed information about special enzymes (new, non-commercial, or historical enzymes), if needed. These enzymes can then be used to find recognition sites or cut molecules. You can edit or delete these special enzymes, and you can add special enzymes to your user lists.

- **To view the list of special enzymes:**

1. Click the menu View, Enzyme Lists.
2. Click on Special Enzymes in the left panel.
3. View enzyme information in the right panel.

- **Add a special enzyme**

Open the Special Enzymes list and click the Edit User List toolbar button then select Add Enzyme. Enter the information required. Use upper- or lower-case letters for the enzyme name, as appropriate. Do not enter trailing Ns in the recognition sequence.

- **Edit or remove a special enzyme**

Open the Special Enzymes list and select (highlight) the special enzyme you want to change or remove. Click the Edit User List toolbar button and select Edit Enzyme to modify the information about the special enzyme or select Delete Enzyme to remove this enzyme (and its detailed information) from the special enzyme file.

You must close the Enzyme List, exit, and then re-start the Clone Manager program to use new special enzymes or update the lists containing these enzymes.

On the Fly Enzyme List

You can create a transient enzyme list just when you need it and discard the list at the end of the current operation. Or you can create a transient list in the Enzyme Lists resource, where you can sort the list and review pertinent information about the enzymes. To create and use a transient list:

1. Start the operation you want to perform.
2. When asked to select an enzyme list, pick On The Fly.
3. Specify enzyme characteristics:

   Recognition Element Size -- Select recognition element sizes of 4, 5, 6, or more than six bases. You can select more than one size. Leave all boxes clear to permit any size element.

   Ends Produced -- Select the type of ends required: 5' overhang, 3' overhang, or blunt ends. You can select more than one end type. Leave all boxes clear to permit any type of ends.

   Selection Criteria -- You can specify if only commercially available enzymes should be included in your transient list. If you do not check this box, a small number of enzymes with unique recognition sequences will be included although they are not commercially available. You can also indicate the supplier you would prefer to use. During enzyme selection, if more than one enzyme with the same recognition sequence and cut position is available, the enzyme from your preferred supplier will be used.

- **View transient list data:**
  1. Click Enzyme Lists toolbar button
  2. In the left panel, click On The Fly (bottom).
  3. Specify enzyme characteristics (see below).

Cloning and Enzyme Operations

Clone Manager will find restriction enzyme sites, cut molecules, modify ends, and ligate DNA. Through all these operations, Clone Manager will track your features and enzyme sites and automatically renumber positions in the new molecule.

Or you can use one of the Cloning Wizards to construct a recombinant molecule using Gateway, Topoisomerase, Gibson, In-Fusion, NEBuilder, or Golden Gate cloning procedures.

Classic Cloning Techniques

- Finding restriction enzyme sites
- Cutting molecules
- Modifying molecule ends
- Ligating molecules
- Join sequences

Finding Restriction Enzyme Sites

Found on the Clone menu.

You can find recognition sites for specific restriction enzymes, or you can find sites for lists of enzymes.

Once found, you can see the number and location of these sites, enter all these sites to the molecule map, if desired, or cut the DNA with one enzyme or at one specific site.

You can also rebuild the map sites list for a molecule by rescanning with enzymes currently listed.

If your molecule does not have sequence data, you can add enzyme sites manually.

- **To find recognition sites for enzymes:**

1. Click Clone, Find Enzyme Sites, then click Enzyme(s).
2. Type in the names of the enzymes, separated by a comma, or find the enzymes in the lookup list to the right (see below).
3. Click OK to find sites for these enzymes.

- **Find Sites Results**

When a search for enzyme sites has been completed, the results will be presented in a scrolling list box and location map display. Enzyme recognition sites are listed, showing the enzyme name and the site position. Click on the column headers to sort by enzyme name or position. The vertical bars above the icon map show the approximate location of the enzyme sites and a red arrow marks the highlighted position (see below).

Add Sites -- Add all the enzyme recognition sites found to the molecule map.

Cut Enz -- Cut the molecule at all the sites found for the enzyme highlighted.

Cut Site -- Cut the molecule at the one enzyme site highlighted (not available when finding sites for lists of enzymes).

One or more arrows appear above the vertical bars, depending on the data displayed:

- Find sites for one or several enzymes: arrow marks recognition site currently highlighted in list.
- Find sites for double cutters on list: two arrows mark sites for highlighted enzyme.
- Find all sites for user list: arrows mark all sites for highlighted enzyme.

- **Finding Sites for a List of Enzymes**

You can find recognition sites for all enzymes on an enzyme list in one easy operation. Once found, you can see the number and location of these sites, enter all these sites to the molecule map, if desired, or cut the DNA with an enzyme.

You can find all Single Cutters (enzymes that cut the molecule only once) or all Double Cutters (enzymes that cut the molecule in two places) using any enzyme list. You can find all sites for all enzymes on the list for user-designated enzyme lists.

- **To find recognition sites for lists of enzymes:**

1. Click Clone, Find Enzyme Sites and then click **List of Enzymes**.
2. Click Single Cutters, Double Cutters, or All for User List.
3. Select the enzyme list using the Enzyme List combo box on the right.
4. Click OK to find sites for these enzymes.

To change the enzyme list, use the combo box above the list of enzymes on the right. You can select one of your user enzyme lists, create an On-the-Fly list of enzymes with specific characteristics, or use one of the available enzyme lists.

Rebuild a Molecule’s Enzyme Sites

This option prepares a list of the enzymes used in the current molecule map and then automatically rescans for each enzyme, rebuilding the list of map sites.

This is useful after a ligation of two molecules or after a change to the sequence of a molecule.

- **To rebuild an enzyme site list:**

1. Click Clone, Find Enzyme Sites and then click **Rebuild**.
2. Click OK to begin the operation.

Cutting Molecules

Found on the Clone menu.

You can cut a molecule by selecting an enzyme from an enzyme list and cutting at all sites for this enzyme. (You can also cut with 2 or 3 enzymes at the same time). Or you can cut at one specific enzyme site already marked on your molecule map or at specified base pair positions.

Options:

- Cut at Map Site
- Cut with Enzyme
- Cut at Base Pair position

Cutting at Map Sites

You can cut a molecule by selecting one enzyme site already marked on the molecule map. By convention, the enzyme site is marked at the left end of the recognition sequence for that enzyme. The program will cut at the correct cut position for this enzyme site.

- **To cut a molecule at one specific map site:**

1. Click Clone, Cut
2. Click **Map Site** and select one enzyme site from the list box.
3. Click OK to cut.

Alternatively, you can use the working map display screen to cut at a single map site. Select (click on) an enzyme site on the working map and click the Cut at Site toolbar button. If you want to check a cut position first, select the enzyme and then click the Site Properties toolbar button to see the exact cut position and ends produced.

- **Cut Results**

Following a cut operation, a copy of the original molecule will be cut into one or more fragments. Resulting fragments can be entered to the molecule list, making them available for use within the program. (If needed for further use, you should save these new molecules to disk.)

If a cut action results in only one molecule fragment, you can immediately enter this fragment to the molecule list. When multiple fragments result from the cut action, you can view the resulting fragments and identify which to enter to the molecule list for immediate or later use.

Multiple fragments are listed, showing the start and end base pair positions and the fragment size. Click on the column header to sort on the values in that column. The red horizontal bar above the icon map shows the approximate location of the highlighted fragment. The enzyme producing the left end of the fragment is shown to the left and above the red bar; the enzyme producing the right end is shown to the right and below the bar.

Use Now -- enter the highlighted fragment to the molecule list and then continue to use this as the active molecule. All other fragments will be discarded.

Create -- enter the highlighted fragment to the molecule list and then return to this dialog box to create additional fragments.

Cutting with Enzymes

You can cut a molecule using restriction enzymes. The molecule will be cut at all the sites for an enzyme. (To cut at only one site, use the cut at Map Site option). You can select 1, 2 or 3 enzymes and cut with these enzymes in one operation.

- **To cut a molecule with enzymes:**

1. Click Clone, Cut and then click **Enzyme(s)**.
2. Type in the names of the enzymes, separated by a comma, or find the enzymes in the lookup list to the right (see below).
3. Click OK to cut at all sites for these enzymes.

- **Enzyme List Lookup**

The enzymes on the selected enzyme list are shown along with their recognition sequence. Press a letter key to move to enzymes in the list that begins with that letter. Double-click to paste an enzyme name from the list to the text box.

To view a different collection of enzymes, change the enzyme list, using the combo box above the list of enzymes. You can select one of your user enzyme lists, create an On-the-Fly list of enzymes with specific characteristics, or use one of the program-defined enzyme lists.

If the recognition sequence is longer than the space provided and you want to be able to see the entire sequence of bases, double-click on the column header divider (right after the word Recognition). You can now use the horizontal scroll bar in this list box to scroll to the right to see the rest of the recognition sequence.

Cutting at Base Positions

You can cut a molecule by specifying one or two base pair positions. You might want to do this to cut out a section of the molecule. The positions you specify do not have to be enzyme sites.

- **To cut at base positions:**

1. Click Clone, Cut and then click **Base Position(s)**.
2. Type in 1 or 2 base pair numbers, separated by a comma
3. Click OK to cut after these base pair positions.

The molecule will be cut blunt ended into one or more fragments and you can select the fragment that you wish to keep.

Modifying Molecule Ends

Found on the Clone menu or access using the Modify Ends buttons from the Ligate display window. Use this option to modify the ends of a linear molecule. An icon map of the molecule is shown, with features (if any). If enzymes were used to create the molecule's ends, these enzyme names appear above the respective end. The actual sequence of each end is shown below the map.

Select the type of end modification using the combo box in the lower left to do Klenow Fill-In, 5' ->3' exonuclease, 3' -> 5' exonuclease, or Partial Fill-In operations (see below). The proposed modification will appear in the sequence of the ends if this type of modification has an effect on the molecule.

Select **Modify** to change the molecule ends, as shown, and alter the sequence for this DNA. You can also directly edit the single stranded extensions by selecting Custom (see below).

Standard End Modifications

Klenow Fill-In -- Treatment with DNA polymerase I Klenow fragment will fill in 5-prime overhang extensions of the molecule to create a blunt end. The bases to be added appear as lower-case color letters in the display.

5' -> 3' Exonuclease -- Treatment with 5' -> 3' exonuclease will remove bases from the 5-prime single-strand ends of the molecule (i.e., a 5-prime sticky end will be trimmed down to a blunt end). The extent of exonucleolytic action is limited to the single-strand extensions. The bases to be excised appear as lower-case letters with a cross-out line in color.

3' -> 5' Exonuclease -- Treatment with 3' -> 5' exonuclease will remove bases from the 3-prime single-strand ends of the molecule (i.e., a 3-prime sticky end will be trimmed down to a blunt end). The bases to be excised appear as lower-case gray letters with a cross-out line in color.

Partial Fill-In -- Select this option and then use the check boxes to select the bases A, C, G or T to partly fill in 5-prime overhang extensions. The bases to be added appear as lower-case color letters.

Custom End Modification

Click the Custom button to directly edit the single-strand extensions. Type in bases or remove bases, as needed. The single-strand extension is limited to a maximum of nine bases. To replace a 5' sticky end with a 3' sticky end, remove the 5' end first and then add bases for the new 3' overhang.

Ligating Molecules

Found on the Clone menu. Use this option to join two or more pieces of DNA or to join the two ends of the same molecule (self-ligate). The ligation display window lets you see the molecules as they are being prepared for ligation.

In the upper half of the window, an icon map of the active molecule is shown, with features (if any). If enzymes were used to create the molecule's ends, these enzyme names appear above the respective end. The actual sequence of each end is shown below the map. The molecule in this window can be modified (see below). Use the Fragment Information button to view the features table for this molecule or to find enzymes that could prepare ends that would be compatible to the fragment to the left or right (in lower window).

In the lower half of the window, all the molecules or fragments are shown in boxes. Click the Add Molecule button to select additional molecules to add to this window. You should arrange these molecules in the order in which they should be joined. Click on a molecule and drag to move to the right or left, as needed. Small icons appear between each fragment box to indicate if ligation is possible. If changes are needed to a fragment, click to select the fragment, causing it to appear in the upper box, where modifications can be made.

- **Cutting the DNA**

  Use the Cut Fragment button to access the cut molecule operations for the selected molecule. You can cut a circular molecule to make it linear, cut out a region to be cloned, or cut with an enzyme to make a compatible end.

- **Modifying the Molecule Ends**

  Use the Modify Ends button to access the modify ends operations for the selected molecule. You can use standard or custom end modifications to prepare molecule ends for ligation. This is the same function as available from the main Clone menu: Modifying molecule ends

- **Invert the molecule**

  Use the Invert Fragment button to reverse (invert) the sequence of the selected molecule.

- **PCR Amplify Fragment**

  Use the PCR toolbar button to use the PCR Amplify Fragment wizard to create a PCR amplified product from the active molecule. The wizard can introduce restriction enzyme sites or add necessary bases.

- **Undo Last Operation**

  Use the Undo button to undo the last modification to the selected molecule.

- **Perform ligation**

  Click the Ligate button to join all compatible ends of the fragments in the order shown. If you do not want the ligate operation to circularize the molecule, click on the drop-down arrow to the right on the Ligate button and select Do not circularize. Once ligation is complete, you can enter the molecule to the molecule list for further use or save it to disk.

  Note: if you do not use a mouse, you can select or move a fragment in the lower window using options found on the Move or Select Fragment button menu.

Join Sequences

Found on the Clone menu.

You can use this function to join two sequences to create a larger molecule. Molecule features (genes, regions, etc.) will be retained during this operation and enzyme sites can be auto-scanned anew or rebuilt from existing site lists. After joining, you can enter the new molecule to the Molecule List, entering a molecule's name and description.

- **Sequences**

  Identify the molecules to join. Use the Change... button to access the molecule list to select a molecule to use for Sequence A and then for Sequence B. The sequence selected for Sequence A should be the first part of a new linear molecule or the sequence into which the other sequence (Sequence B) will be inserted.

  Use the Invert checkbox to invert (reverse complement) a molecule sequence, if needed. The positions of all sites and features will be adjusted.

  The join operations work on upper strand sequence. If a molecule has cohesive ends, the upper strand of sequence will be treated as if it were blunt ended. If a molecule is circular, it will be cut after the last base to open the molecule at the origin. For an Insert operation, Sequence A will retain its original format (it can be circular or linear with sticky ends).

- **Operation**

  You can select the method to use to join the two sequences:

  Append -- simply add Sequence B to the end of Sequence A. A blunt-end ligation will attach the 3' end of sequence A to the 5' end of sequence B. Molecule features will be retained and base pair positions recalculated. If you select to circularize the resulting molecule, the program will blunt-end ligate the linear molecule to make it circular.

  Splice -- remove overlap when appending B to A. Prior to ligating the sequences, the program will check the bases at the 3' end of sequence A and at the 5' end of sequence B, looking for exact homology. If at least ten bases of exact matches are found at the ends of the molecules, the matching bases of Sequence B will be trimmed, and the ligation performed. If at least ten bases of exact matches are also found at the ends of the newly spliced linear molecule, then these matching bases will also be trimmed, and the molecule circularized.

  Insert -- insert Sequence B after bp number *n* in Sequence A. After selecting this option, use the edit box provided to enter the base pair number in Sequence A after which the insert should be made. Use the Features... button to look at the features table for the Sequence A molecule.

  Auto-scan for Enzyme Sites -- Enzyme sites for the new molecule can be auto-scanned at the completion of the join operation. The AutoScan procedure uses your designated

AutoScan enzyme list, looking for enzyme sites (single cutters or all on user list, as specified in the AutoScan settings). The enzyme sites found will be entered, replacing the existing enzyme sites.

You can clear the AutoScan checkbox, instructing the program to rebuild the map sites list by rescanning with enzymes currently listed for either molecule. This is the procedure used following a standard Clone, Ligation operation, but it may result in a number of enzyme sites that are not single cutters for the newly joined molecule.

Cloning Wizards

Cloning Wizards provide guidance on implementing specialized cloning protocols. Each wizard will guide you through selecting the input fragments and cloning vectors as well as selecting the appropriate options for cloning.

- Using the Plan Cloning Wizard
- Gateway Cloning Wizard and Results
- Topoisomerase Cloning Wizard and Results
- Assembly Cloning Wizards (Gibson, In-Fusion, NEBuilder) and Results
- Golden Gate Cloning Wizard
- CRISPR
- Ligation Independent Cloning
- PCR Cloning Wizard

Plan Cloning Wizard

You can describe a cloning experiment and have the Plan Cloning Wizard suggest some possible cloning solutions for your review. You can select one of the suggested solutions and view more details about this procedure or create the resultant recombinant with the click of a button.

To get started, click Plan Cloning on the Clone menu, and follow the wizard instructions to identify the insert and vector sequences and indicate your preferences for the cloning experiment.

- **Define What You Want to Clone**

  Identify the molecule that contains the region you want to clone. Use the Change... button to select another molecule if the one shown is not the one you want to work with. Use the Features... button to look at the features table for this molecule and select a gene, if appropriate. The program will enter the base positions for you.

- **Define the Vector You Want to Use**

  Click the Change... button to select a molecule that you want to use as the vector. The Wizard can only use circular vector molecules. If the vector molecule contains a multi-cloning site (shown on a molecule map as a grouped display to the right or left of the map), the base pair positions for this region will automatically be entered for you.

  You can use the Features... button to look at the features table for this molecule. You might want to select a gene (such as a gene for antibiotic resistance) to do recombinant selection by insertional inactivation. If you routinely use a vector molecule and want to remember the region you clone into (assuming it is not a multi-cloning site), you can enter this region into the features table of the molecule.

- **Select an Enzyme List**

  In finding possible solutions for your cloning experiment, the program will consider only those enzymes contained in the list you select. The default list is the one you designated as your AutoScan list. You might want to try more than one enzyme list, depending on how many or how few possible solutions you get. When you are viewing the results, you can easily Redefine your search without entering all the information again.

- **Specify Your Cloning Preferences**

  You can accept the default settings for all preferences, or you can select your own individual preferences. You can indicate whether to use 1 or 2 enzymes to cut the insert, if you will allow the use of blunt-cutting enzymes or Klenow fill-in, if you have a preference for orientation of the insert, or if you want to limit extra insert DNA.

  When you select an option that says "Prefer ..." the program will increase the weight for results of this type and move them closer to the top of the list of possible solutions.

When you select an option that says "Require ..." the program will eliminate results that do not satisfy this requirement.

The preferred enzyme supplier selection is used to collect specific information about each enzyme shown in the details of construction. This information includes buffers, incubation temperature, heat inactivation status.

- **View Plan Results**

  If solutions can be found for the cloning experiment you described, these results will be shown in a table display, sorted with the best results (based on your preferences) at the top.

  The enzymes used to cut the insert are shown in the central columns (in color), with the insert size in the center area. The enzymes used to cut the vector are shown at the right or left of the insert, producing the compatible ends needed to ligate the recombinant. Square brackets indicate a blunt cut, k marks the use of Klenow fill-in.

  The orientation of the insert is shown as > (clockwise), B (both), or < (counterclockwise). Exclamation points at the far right indicate that one or more of the enzymes used was not available from your preferred supplier.

  You can view the details of a suggested solution, construct a recombinant, or redefine the experiment.

- **Details of Construction**

  Select (highlight) the result of interest and click the Details of Construction toolbar button. The Details display shows the enzyme cuts required, the number of fragments resulting, the size and orientation of the insert, and information for each enzyme (buffers, incubation temperatures and heat inactivation).

- **Construct Recombinant**

  Select (highlight) the result of interest and click the Construct Recombinant toolbar button. The program will automatically do the required steps to simulate the cloning experiment and produce the recombinant molecule.

- **Redefine Experiment Plan**

  Click the Redefine toolbar button in the Plan Results display window to redefine (fine-tune) the cloning experiment. You may want to do this if you end up with too many or too few possible solutions. You can change the enzyme list selected, change the preferences, or increase the size of the regions involved, without re-entering all the other information.

Gateway® Cloning Wizard

This wizard will help you to simulate cloning using Gateway® Recombination Cloning techniques (Invitrogen Corp). It will help you select the appropriate components, show you the proposed result, and create the resulting recombinant molecule. If you are using a Multisite Pro kit, the wizard will walk you through constructing multiple fragments with the right attB sites and donor vectors. Please refer to the technical publications received with the Gateway® products or available at www.invitrogen.com for more information about the recombination reactions or possible use restrictions.

Selected donor and destination vector files are provided. These files are copied to your home directory and can be found in the subfolder GWC_Vectors.

To get started:

Select the Gateway Cloning Wizard from the Clone menu. Next, identify the procedure you want to do. You can make an entry clone (BP reaction), an expression clone (LR reaction), prepare a PCR product with attB sites, or simulate a complex multisite procedure.

- **Select Insert and Vector Molecules (BP and LR reactions)**

  To make an Entry Clone, identify the molecule that contains the region you want to clone, flanked by attB sites. Click the Change button to access the Molecule List or Browse. Next, identify the donor vector you want to use. In many cases, a donor vector will be suggested. Click the Change button to select another vector. The donor vector you select must have attP sites that match the attB sites in your insert.

  To make an Expression Clone, identify the insert (entry clone) with attL sites that you want to use. Click the Change button to access the Molecule List or Browse. Next, select the destination vector. Click the Change button to select another vector. The destination vector must have matching attR sites. If you want to use more than one entry clone in this reaction, set the checkbox and click Next to move to a page where you can add multiple entry clones.

- **Define What you Want to Amplify (PCR product generation)**

  Identify the molecule that contains the region you want to amplify. Use the Change button to select another molecule if the one shown is not the one you want to work with. Enter the upper strand coordinates of the region you want to amplify and set the strand (normal or complement) to amplify. Use the Features button to look at the features table for this molecule and select a gene, if appropriate. The program will enter the base positions for you.

  Select the attB extensions you want to add to each primer. For a simple one fragment recombination, attB1 and attB2 are standard. When amplifying a region on the normal strand, the attB extension for Primer A will lead into the expressed sequence. When amplifying a region on the complement strand, the attB extension for Primer B will lead

into the expressed sequence. Use the box below the attB extension to type in other bases that you want to add after the attB extension and before the template sequence. These additional nucleotides may assist in protein expression or maintain proper reading frame. For both Primer A and Primer B, enter the additional bases 5' to 3'.

In Clone Manager Professional, primers can be designed and evaluated. The primer type, criteria, and primer length within the sequence template have been set and should not need to be changed. Set the primer design goal to Restrict to Amplified Region if you do not want any extra sequence before or after the region you selected to amplify. The program will construct the possible primers of different lengths at the two fixed positions and select the best primer pair. If you need higher quality primers and can accept some extra sequence, set the design goal to Better Primers. Specify the GC content range and criteria adjustment that you will allow, if needed, to find better primers. The program will search the region 20-40 bases before and after the bounds of the region to amplify, attempting to find primers that meet all criteria settings and then selecting the best primers.

- **Multisite Pro kits**

  Select the 2, 3 or 4 fragment procedure and indicate the method to use. If inserts have already been prepared, identify the molecule or PCR product that contains the region you want to insert for each fragment, flanked by the needed attB sites. The first line (Make Fragment...) shows the attB sites required for this fragment. Click the Change button to access the Molecule List or Browse. The donor vector with matching attP sites is suggested. Continue to create the additional fragments needed.

  If you need to prepare the PCR products, begin by identifying the molecule, region, and strand to amplify for this fragment. The required attB extensions are preset for primers A and B for each fragment. You can add extra bases between the attB extension and the template sequence, if needed. This PCR product will be generated and used for the subsequent BP reaction. The donor vector with matching attP sites is suggested. Continue this two-step process to create the additional fragments needed.

  When all fragments have been prepared, they will be collected and automatically added as entry clones for the subsequent LR reaction. Select a destination vector to prepare the expression clone. The destination vector you select must contain attR1 and attR2 sites. More information on Multisite Gateway® Cloning Components

Gateway® Cloning Results

After using the wizard to simulate the construction of an entry or expression clone using Gateway® Recombination Cloning techniques (Invitrogen Corp), you can view the proposed result and create the recombinant molecule. An iconic map of the recombinant molecule is shown, with the insert marked in a lighter color, inserted in the direction indicated by the arrow. If features were present in the vector or insert molecules, these features will be shown above the map line. The area below the map shows detailed information about the recombinant molecule, the insert molecule, and the vector used.

Click the Use Now button to take the molecule shown and use it in the next logical step in Gateway cloning. If you have just done a BP reaction, for example, the Use Now action will start an LR reaction using the entry clone you just created.

Click the Create Molecule button to have the program automatically do the required steps to simulate the cloning experiment, produce the recombinant molecule, and enter to the Molecule List. You can use this new molecule now or save it to disk for later use. Following a multisite procedure, you can also collect the intermediate molecules generated during the procedure (entry clones, PCR products).

If you are using Clone Manager Professional, you can design and evaluate primers to generate PCR products with attB extensions for use in Gateway cloning steps. If you requested more than one solution, you could view the solutions sequentially using the Next and Previous toolbar buttons. For each solution, the iconic map shows the amplified product, flanked by the attB sites. The area below the map includes information about the amplified product and about the primers used to generate this product. If either primer (or the primer pair) does not meet the criteria set, the notation 'Caution -- check primer pair report' appears above the primer descriptions.

Click the Primer Pair Report button to view the evaluation report and related primer analysis screens for the primers associated with this solution.

You can use other toolbar buttons to export primer sequences or add these primers to the primer list, saving the primers to disk or adding them to your primer collections. Following a multisite procedure, the Create Molecule option allows you to add all the primers to a new primer collection or export all primer sequences in one easy step.

- **Verify Results**

  When the simulated recombinant molecule has been generated, you will want to review this new molecule to verify that the results are as expected. You might want to view the Features Table to verify your inserts are present in the correct position and orientation. Select an inserted gene and use the toolbar button Go To Start of Sequence to open the sequence view with the translated gene highlighted in the context of the full molecule sequence. Other helpful tools you might want to use include ORF Search and Analyze Molecule, Open Reading Frames, both found on the Operations Menu.

- **Redefine Cloning Experiment**

  Click the Redefine toolbar button to update the cloning experiment just completed. You may want to do this if you find you inserted a gene in the wrong orientation or selected the wrong vector.

Create Molecules - Multisite Procedure

After using the wizard to simulate a complex multisite procedure, you can have the program automatically do the required steps to simulate the cloning experiment and produce the recombinant molecule and all intermediate molecules in one batch operation. If you are using Clone Manager Professional, you can also create and save the primers used to prepare PCR products and export the primer sequences.

Click the Create Molecule button to start the process.

- **Molecules**

  Use the checkboxes to indicate which molecules you would like the program to create and enter to the Molecule List for you. You can select to create the recombinant shown on the results screen (Expression Clone) and the entry clones prepared in each of the BP reactions. (These entry clones were then used in the final LR reaction to produce the expression clone.) If you prepared PCR products to make the entry clones, you can also elect to create and enter these molecules.

  Molecule names follow the suggested naming conventions: Expression clones and entry clones begin with pEXP or pENTR, as appropriate, followed by a short vector name and then the subcloned gene name. PCR products begin with pcr followed by the two flanking attB sites and the amplified gene name. Once molecules are on the Molecule List, you can open these molecules to review their contents or save them to disk for later use. When saving to disk, it may be helpful to create a new folder and store all the molecules from one recombinant experiment in the same location.

- **Process Molecules**

  Enzyme sites for the new molecules can be auto scanned using your designated AutoScan enzyme list, looking for enzyme sites (single cutters or all on user list, as specified in the AutoScan settings). The enzyme sites will be entered, replacing the existing enzyme sites.

  You can clear the AutoScan checkbox, instructing the program to rebuild the map sites list by rescanning with enzymes currently listed for the molecules being recombined. This is the procedure used following a standard Clone, Ligate operation, but it may result in a number of enzyme sites that are not single cutters for the newly created molecule.

- **Primers**

  If you are using Clone Manager Professional, you can use the checkboxes to enter or use primer data generated in this experiment. You can elect to create each of the primers used to prepare PCR products and save these primers to a new primer collection. The new collection will be named GateWiz_ followed by the subcloned gene name (as found in the expression clone). Primer names will indicate the attB site appended and the gene to be amplified. If you select to export all primer sequences, the list of primers (name and sequence) will be copied to the Windows clipboard. When you see the message

indicating that this data is on the clipboard, you can paste this information into another document. You might want to do this to order oligos or document your work.

Topoisomerase Cloning Wizard

This wizard will help you to simulate cloning using Topo® Cloning techniques (Invitrogen Corp). It will help you select the appropriate components, show you the proposed result, and create the resulting recombinant molecule. Please refer to the technical publications received with the Invitrogen products or available at www.invitrogen.com for more information about the recombination reactions or possible use restrictions.

Selected vector files are provided. These files are copied to your home directory and can be found in the subfolder Topo_Vectors.

To get started:

Select the Topoisomerase Cloning Wizard from the Clone menu. Next, identify the procedure you want to do. You can do TA, directional, or blunt Topo cloning, or you can prepare a PCR product for use in a subsequent cloning procedure.

- **Select Insert and Vector Molecules (Topo cloning)**

  Identify the molecule that contains the region you want to clone. Click the Change button to access the Molecule List or Browse. For TA cloning, the program will automatically add a 3' terminal A base to each strand if not already present. For directional cloning, the program will automatically add 5' terminal CACC to the appropriate strand if not already present.

  Next, identify the vector molecule you want to use. In many cases, a vector will be suggested. Click the Change button to select another vector. The vector you select must have Topo sites that the program can recognize. For TA cloning, the program will automatically add a 3' terminal T base to each strand if not already present. For directional cloning, the program will automatically add a complementary GTGG overhang if not already present.

- **Prepare PCR product**

  Identify the molecule that contains the region you want to amplify. Use the Change button to select another molecule if the one shown is not the one you want to work with. Enter the upper strand coordinates of the region you want to amplify and set the strand (normal or complement) to amplify. Use the Features button to look at the features table for this molecule and select a gene, if appropriate. The program will enter the base positions for you.

  Next select the cloning ends required for use in the subsequent cloning procedure. You can select blunt, TA cloning, or directional cloning ends. For TA cloning, the program will add 3'A overhangs to each end of the amplified product. For directional cloning, 5' CACC will be added to the Lead In primer (see below).

  When amplifying a region on the normal strand, Primer A will lead into the expressed sequence. When amplifying a region on the complement strand, Primer B will lead into

the expressed sequence. Use the boxes below the primer designations to type in other bases that you want to add to the 5' extension before the template sequence. These additional nucleotides may assist in protein expression or maintain proper reading frame. For both Primer A and Primer B, enter the additional bases 5' to 3'.

In Clone Manager Professional, primers can be designed and evaluated. The primer type, criteria, and primer length within the sequence template have been set and should not need to be changed. Set the primer design goal to Restrict to Amplified Region if you do not want any extra sequence before or after the region you selected to amplify. The program will construct the possible primers of different lengths at the two fixed positions and select the best primer pair. If you need higher quality primers and can accept some extra sequence, set the design goal to Better Primers. Specify the GC content range and criteria adjustment that you will allow, if needed, to find better primers. The program will search the region 20-40 bases before and after the bounds of the region to amplify, attempting to find primers that meet all criteria settings and then selecting the best primers.

Topoisomerase Cloning Results

After using the wizard to simulate the construction of a clone using Topo® Cloning techniques (Invitrogen Corp), you can view the proposed result and create the recombinant molecule. An iconic map of the recombinant molecule is shown, with the insert marked in a lighter color, inserted in the direction indicated by the arrow. If features were present in the vector or insert molecules, these features will be shown above the map line. The area below the map shows detailed information about the recombinant molecule, the insert molecule, and the vector used.

Click the Create Molecule button to have the program automatically do the required steps to simulate the cloning experiment, produce the recombinant molecule, and enter to the Molecule List. You can use this new molecule now or save it to disk for later use.

If you are using Clone Manager Professional, you can design and evaluate primers to generate PCR products for use in subsequent topoisomerase cloning procedures. If you requested more than one solution, you could view the solutions sequentially using the Next and Previous toolbar buttons. For each solution, the iconic map shows the amplified product. The area below the map includes information about the amplified product and about the primers used to generate this product. When you select the option to design Better Primers, if either primer (or the primer pair) does not meet the criteria set, the notation 'Caution -- check primer pair report' appears in the primer description area.

Click the Primer Pair Report button to view the evaluation report and related primer analysis screens for the primers associated with this solution.

You can use other toolbar buttons to export primer sequences or add these primers to the primer list, saving the primers to disk or adding them to your primer collections.

If you have just prepared a PCR product, you can click the Use Now button to take the amplified product shown and use it in a TA, directional, or blunt topoisomerase cloning procedure.

Note: In the case of directional cloning, if two or more bases at the 3' end of the insert are homologous to the CACC overhang, you will be alerted that this may result in bidirectional cloning. You can continue or abandon the procedure and perhaps modify the insert molecule or add a cap to one of the primers.

- **Verify Results**

  When the simulated recombinant molecule has been generated, you will want to review this new molecule to verify that the results are as expected. You might want to view the Features Table to verify your inserts are present in the correct position and orientation. Select an inserted gene and use the toolbar button Go To Start of Sequence to open the sequence view with the translated gene highlighted in the context of the full molecule sequence. Other helpful tools you might want to use include ORF Search and Analyze Molecule, Open Reading Frames, both found on the Operations Menu.

- **Redefine Cloning Experiment**

  Click the Redefine toolbar button to update the cloning experiment just completed. You may want to do this if you find you inserted a gene in the wrong orientation or selected the wrong vector.

Assembly Cloning Wizards

You can use assembly cloning wizards to help you to simulate cloning using Gibson Assembly™ (Synthetic Genomics, Inc.), In-Fusion® Cloning (Clontech Laboratories Inc), Ligation-Independent Cloning or related cloning techniques. The wizards will help you select the appropriate components, show you the proposed result, and create the resulting recombinant molecule. Please refer to the technical publications received with any products for more information about the recombination reactions or possible use restrictions.

- **To get started:**

  Select one of the assembly cloning wizards from the Clone menu to get started. Use the Ligation-Independent Cloning wizard to join fragments previously prepared with correct overlaps. Use the Gibson Cloning or In-Fusion Cloning wizards to use PCR to create amplified fragments with overlaps, and then assemble the fragments, using the assembly procedure selected. Set the procedure and primer design options, if applicable.

- **Select Fragments with Overlaps to Assemble**

  Click the Add Fragment button to access the Molecule List or Browse to identify the fragment with overlaps to add to the list of fragments to assemble. Click the Add Vector button to identify a linearized vector fragment with overlaps to add to the list. A vector with overlaps at both ends can result in a circular recombinant molecule.

  Starting with the first fragment in the list, overlaps will be used to determine the order of assembly. The fragment in the top position is used to seed the assembly and set the start and orientation. A vector fragment will automatically be placed in the first position of the list. If you are not using a vector, you can move an important fragment to the first position.

- **PCR Amplify Fragments to Assemble**

  Click the Add Fragment button to add fragments to the list of fragments to assemble. For each fragment, identify the molecule that contains the region you want to amplify. Use the Change button to select another molecule if the one shown is not the one you want to work with. Enter the upper strand coordinates of the region you want to amplify and set the strand (normal or complement) to amplify. Use the Features button to look at the features table for this molecule and select a gene, if appropriate. The program will enter the base positions for you. Set the check box if you need to add extra bases before (lead in) or after (lead out) this fragment. Enter the extra bases as they would appear in the upper strand of the final fragment assembly.

  Click the Add Vector button to identify the vector molecule you want to clone into. For circular vectors, you can enter the base pair positions (upper strand) that immediately flank the site where the assembled insert will be placed. When you select to enter base pair numbers, you have the option to PCR amplify the vector. The PCR primers

generated for the vector will not have insert overlaps. Alternatively, you can select one or two enzymes to cut the vector, creating the insert site. The drop-down controls list the single-cutting enzymes you can use for this molecule. If a single cut is to be used, enter the same enzyme name in both boxes. (Note that 5' overhangs of a restriction site will be removed during the Gibson assembly procedure.) Set the orientation for the assembled insert.

In the list of fragments to assemble, the fragments should be listed in the correct order for assembly. Use the Move Up or Move Down buttons to re-order the fragment list as needed. If a vector is used, the vector fragment will automatically be placed in the first position of the list of fragments to assemble.

In Clone Manager Professional, primers can be designed and evaluated. The primer type and criteria have been set and should not need to be changed. The program will construct the possible primers of different lengths at the two fixed positions, adding overlaps and extra bases (if needed), and select the best primer pair. If the option has been selected, the overlap segments will be split between the two adjacent fragments, when possible. This permits greater flexibility in primer design and may result in primers of shorter length. When the junction is between a fragment and a vector, the entire overlap segment will be a part of the insert primer.

Assembly Cloning Results

After using the wizard to simulate the assembly of a molecule using assembly cloning techniques, you can view the proposed result and create the recombinant molecule. An iconic map of the recombinant molecule is shown, with the inserted fragments drawn in alternating lighter colors. Fragment numbers appear below the map segments. If features were present in the vector or insert molecules, these features will be shown above the map line. If primers were created, these appear above or below the map line at the fragment junctions. The area below the map shows detailed information about the recombinant molecule, the insert fragments, and the vector used.

If you are using Clone Manager Professional and have designed primers to amplify the fragments and create the overlaps, you can view more detailed information about the primers. Use the drop-down list box to move from the Summary view to the view of primer pairs for each fragment. These detailed views give the primer description, primer sequence, possible primer cautions and a three-frame translation of the primer sequence (5' to 3'). Bases in upper case letters are homologous to the template; bases in lower case letters represent the overlap and extra bases added (if any).

Click the Create Molecule button to have the program automatically do the required steps to simulate the cloning experiment, produce the recombinant molecule, and enter to the Molecule List. You can use this new molecule now or save it to disk for later use. When entering to the Molecule List, you can change the molecule's name or description. You can also indicate if enzyme sites for the new molecule should be auto scanned using your designated AutoScan enzyme list, looking for enzyme sites (single cutters or all on user list, as specified in the AutoScan settings). The enzyme sites will be entered, replacing the existing enzyme sites.

If you are using Clone Manager Professional, you can also create and save the primers used to prepare PCR amplified fragments. Use the check boxes to add all the primers to a new primer collection, export all primer sequences, or add primer sites to the recombinant molecule in one easy step. The new primer collection will be named AsmWiz_ followed by the name you entered (or the default Assembly 1). If you select to export all primer sequences, the list of primers (name and sequence) will be copied to the Windows clipboard. When you see the message indicating that this data is on the clipboard, you can paste this information into another document. You might want to do this to order oligos or document your work.

- **Verify Results**

  When the simulated recombinant molecule has been generated, you will want to review this new molecule to verify that the results are as expected. You might want to view the Features Table to verify your inserts are present in the correct position and orientation. Select an inserted gene and use the toolbar button Go To Start of Sequence to open the sequence view with the translated gene highlighted in the context of the full molecule

sequence. Other helpful tools you might want to use include ORF Search and Analyze Molecule, Open Reading Frames, both found on the Operations Menu.

If you are using Clone Manager Professional and have selected to add the primer sites to your recombinant molecule file, you can view these primer sites on the molecule map. Click the Enzyme Sites/Primer Sites toolbar button to toggle between enzyme sites and primer sites. Right click on a primer site and select Go To Sequence to view the molecule sequence at the position of the primer, with primers shown above the sequence bases.

- **Redefine Cloning Experiment**

  Click the Redefine toolbar button to update the cloning experiment just completed. You may want to do this if you find you inserted a gene in the wrong orientation or selected the wrong vector.

Golden Gate Cloning Wizard

This method provides for the directional assembly of two, or many, fragments in a single reaction. Fragments can be joined in a scarless, or seamless, fashion. Recombinant clones can be used to provide a source of functional cassette modules which can be combined into a final recombinant.

The Golden Gate cloning method employs a mix of a type II-S restriction enzyme and ligase in a single reaction tube. Type II-S restriction enzymes have an asymmetric recognition sequence and cut outside of the recognition sequence. By positioning a pair of recognition sites outside the region of interest, and pointing inwards, the insert can be excised without including any extraneous sequence. Thermal cycling between the temperatures optimal for restriction enzyme digestion and ligation allows for high cloning efficiency by preventing the re-ligation of the original clones or vector.

Since type II-S enzymes cut outside of the recognition sequence, the exact nature of the overhang is determined by the adjacent sequence. The typically used type II-S restriction enzyme cuts to produce a four-base overhang. Because the enzyme cuts outside of the recognition sequence, the overhang produced can be controlled by the adjacent sequence or the sequence of the PCR primers used for amplification. This provides 256 sequence combinations that can be used to uniquely define the order of assembly of insert fragments using simple ligation.

- **Clone Manager Wizard**

The Golden Gate wizard can be started from the main Clone menu. The wizard is based on the standard cloning ligation window and provides access to all the standard ligation operations in addition to the specific design options to support the Golden Gate cloning method.

The wizard assumes that any molecule selected either has a suitable pair of type II-S restriction enzyme sites to excise the fragment to be cloned or else will use PCR amplification to introduce the pair of sites.

When the Golden Gate wizard is started, the first dialog page will allow you to enter the default settings for the type II-S restriction enzyme being used and defaults for design of PCR primers used to amplify fragments. Each of these settings can be customized when designing individual PCR primers.

- **Settings**

Select the type II-S restriction enzyme that you want to use. Clone Manager will check that there are no unexpected internal recognition sites in any of the fragments being cloned. Available enzymes can be selected from the drop-down list and include all enzymes with at least a six-base recognition sequence that produces at least a 4 base overhang and cut within 4 bases of the end of the recognition sequence.

Other settings relate to how the PCR primers will be designed. The 5' extension (not homologous to the region being cloned) adds the bases for the recognition sequence of the desired type II-S enzyme together with additional bases that provide the base overhang that will be used to ligate to the adjacent fragment.

The primer bases that precede the restriction enzyme sequence provide a few bases to enable the enzyme to bind and cut efficiently. These bases will not be included in the amplified product.

The primer bases that come after the restriction enzyme sequence allows for introducing additional bases into cloned fragments and can be changed when you are designing an individual fragment to be cloned. On the settings page, it is recommended that you enter one base here and check the box labeled 'Match overhang of adjacent fragment.' The one base is required, for the most used restriction enzymes, to fill the gap between the end of the restriction enzyme recognition sequence and its cut position. This base will not appear in the final recombinant. Checking the 'Match' box enables the wizard to automatically design appropriate overhangs to enable the correct assembly of fragments in the order you specify.

- **Add or Replace Molecule**

  Use the first toolbar button to add a molecule or the drop-down button to replace an existing molecule. Other toolbar buttons allow you to re-organize or modify the fragments and the order in which they will be joined.

  Clone Manager will automatically use a suitably positioned pair of type II-S restriction enzyme sites, if found. If you do not want to use the selected enzyme, you can click the toolbar Undo button.

  If a pair of type II-S sites are not found, then Clone Manager will assume that you will be using PCR amplification, and the PCR wizard will allow you to design the fragment that you want to clone. Should you not want to use PCR amplification, you can click the cancel button to close the PCR wizard.

  The fragment to clone will now appear in the parts list across the bottom of the window. The display will show the overhang ends for the fragment and its relationship to adjacent fragments. Fragments will be assembled into the final recombinant in the order they appear in this window.

  Toolbar buttons are available to review, or change, the last PCR step or undo the last construction step.

- **PCR Wizard**

  This wizard can be started from the toolbar to review, or customize, the design of the PCR primers.

  The wizard defines the region of the molecule to clone. The region is defined using coordinates from the upper strand, but you can select which strand you want to clone.

Selecting to clone the complement strand will effectively invert the selected region to give you the reverse complement sequence. The features button allows quick selection of region and strand.

The PCR design criteria assume that you are subcloning from an existing cloned recombinant and uses rather loose design criteria to evaluate and design primers.

The next two pages of the PCR wizard allow you to design the primers for the left and right ends of the fragment region you are cloning. Each page contains the same primer design fields but the second page for designing the right-hand primer will show the amplification results in the reverse compliment direction to reflect the content of the final amplified product after cutting with the restriction enzyme. Selection of type II-S restriction enzyme and the preceding bases were described under Settings above.

The field for additional bases will be disabled if you are using the default automatic match to allow ligation to the adjacent fragment. You can customize this field by unchecking the match box. This will enable you to enter extra bases into the primer sequence and these bases will appear in the final recombinant between the two adjacent fragments.

- **Assembly design display**

  The display updates to show you an overview of the design of each fragment in this project. Across the bottom of the window will be a list of iconic fragments shown in the order that they will be ligated. The icons will show a brief description of the fragment together with the overhangs that will be generated. Between each fragment will be an icon that confirms that the overhangs between adjacent fragments are compatible for ligation. If the failure icon is shown, then ligation is not possible, and you will need to resolve the incompatibility before continuing.

  Above the ordered list of fragments is a more detailed iconic map of the selected fragment including its size and sequence of the ends.

- **Ligation of product**

  Complete the wizard and generate the final recombinant by clicking the ligate button which is the final button on the toolbar. The default action assumes that you want to circularize the final product if possible. More options are available from the drop-down menu on this button. Also included as an option is to display construction details which will report how each fragment is prepared. This information is also saved in the notes field of the final recombinant.

  You can also design and save suitable primers. Primer design will use the criteria for ‘General Primer Pair’ which uses rather loose design criteria assuming that you are using PCR amplification to subclone from an existing cloned recombinant. Saving the primer sequences will enable you to use them for analysis or other purposes. You can also export primer sequences which may be convenient for ordering the primers.

References

Engler, C., Kandzia, R., & Marillonnet, S. (2008). A one pot, one step, precision cloning method with high throughput capability. PloS one, 3(11), e3647.

Engler C, Marillonnet S. Golden Gate cloning. Methods Mol Biol. 2014; 1116:119-31. doi: 10.1007/978-1-62703-764-8_9. PMID: 24395361.

Weber, E., Engler, C., Gruetzner, R., Werner, S., & Marillonnet, S. (2011) A Modular Cloning System for Standardized Assembly of Multigene Constructs. Plos one Published: February 18, 2011

Cleavage close to the end of DNA Fragments: https://www.neb.com/tools-and-resources/usage-guidelines/cleavage-close-to-the-end-of-dna-fragments

Ligation Independent Cloning

Found on the Clone main menu. The wizard will guide you through the steps required.

- Settings – allows entering the minimum number of bases allowed and the minimum fragment size.
- Table of fragments
  - Add Fragment – select one or more fragments to be assembled
  - Add Vector – this molecule will be used to set the orientation of the resulting recombinant.
  - Move Top – moves the selected fragment to the top of the list of fragments. (Any vector molecule will always be in the first position.)
  - Remove – the selected fragment or vector.
- Final page – reports number of fragments to assemble and the size of the final recombinant.

Results:

The results display shows an iconic map of the assembled recombinant with the positions of prominent features. The locations, and extents, of each assembled fragment are shown as colored segments.

Reported are the size of the recombinant molecule and each contributing fragment.

Results Toolbar:

The following toolbar buttons are active for this cloning method:

- Create Molecule – enter the recombinant to the Molecule list. You can enter a name and description.
- Redefine – shows the wizard and allows making changes.

CRISPR

CRISPR (Clustered Regularly Interspaced Short Palindromic Repeat) is a powerful genetic modification technique adapted from a simple bacteria immune-like system. The system uses a specially designed nucleotide guide sequence to direct action at a specific gene target. The guide sequence works with a Cas enzyme (CRISPR associated protein) to introduce a double strand cut in the target. The cut is then inefficiently repaired, using non homologous end joining, resulting in a deletion, or insertion, of base(s) at the target site. The net effect is the disruption of the translation coding frame resulting in a knockout of protein function.

There are three components required for the design of a CRISPR experiment. A guide sequence that is homologous to the specific target region of the molecule. A Protospacer Adjacent Motif (PAM) sequence adjacent to the guide sequence is required to enable a Cas nuclease to bind the complex and cleave the target.

Design Crispr guide sequence

The CRISPR wizard is designed to assist in guide sequence selection. The guide needs to be in the region of interest, and it needs to have a PAM sequence adjacent. In general, the guide location should be located nearer the amino terminal end of the gene to maximize knockout. This prevents the likelihood that a truncated protein will still be functional. Similarly, if the knockout occurs too early in the translation, there is a possibility that a second translation initiation codon will allow a functional gene product.

The wizard will identify potential guide sequences by searching for PAM sequences within the target region. The guides will then be evaluated for off-target homologies which represent alternate sites that could be acted upon and cause unintended mutations in other genes.

CRISPR enzyme selection

There are an increasing number of enzymes that can be used for CRISPR experiments. Each enzyme has a corresponding PAM sequence, and this will affect the selection of potential guides. If the wizard does not find good potential guides, you can select a different Cas enzyme.

Each Cas enzyme has an optimal guide sequence length. When you select a different Cas enzyme, the wizard will update the design parameters to match your preference for that enzyme. Your settings preferences will be remembered when you complete the wizard. You can also click the ‘Reset’ button to return the settings to default.

CRISPR wizard conditions

Length of guide RNA: The default guide sequence length is set to your preference for the selected Cas enzyme (typically 20-24 bases). You can adjust this value if preferred.

PAM adjacent perfect match bases: This is used when evaluating possible off target sites. Bases adjacent to the PAM site are most important for binding of the Cas complex and mismatches within this close region will be most likely to prevent binding and reduce off target consequences. The default value requires ten adjacent bases to be perfectly matched. Selecting a lower number will provide more sensitivity to finding potential problems, but the search for off-target sites will be slower.

Max mismatches permitted: Off target sites will be ignored if there are more than the specified number of mismatches.

CRISPR wizard Off Target Search

This allows you to specify the genomic sequence that will be used to search for off target sites that may also be affected by Cas cleavage.

If you do not select an off-target sequence, the wizard will use the full molecule used for the guide design. This is most appropriate if you are designing a knockout experiment using a region of a full bacterial genome. For other organisms you can download genome sequences from NCBI (see references).

You can select a file using the 'Browse' button. Your most recently used selections are shown by clicking the 'Recent' button.

Sequence files may be either a single molecule file or a FASTA format multi molecule file. Single molecule files can be complete bacterial genome sequences (mega-base). FASTA format files can include complete eukaryotic genome sequence collections, such as the Human genome (GRCh38, file size 3 giga-bases).

The Clone Manager wizard has been optimized for rapid processing of very large sequences. Time to complete the analysis will typically be a few seconds. However, larger guide design target regions and large off-target sequences will increase the time.

CRISPR wizard results

The results view displays a list of possible guide sequences. For each possible guide, its position and sequence are reported together with the result of the search for off target hits. The number of off target hits are reported together with the sequence of the worst homology. Mismatches are shown as upper-case bases so you can evaluate the significance of the off-target site.

CRISPR wizard results toolbar

The results view toolbar enables you to change the displayed results. Hovering your mouse over a control will show a tooltip explaining the function.

- Position: shows the base position of the PAM sequence which is usually close to the Cas nuclease cut position. The other option is to show the position of the guide sequence.
- Sort: defaults to showing results in position sorted order. The optimal position of a guide will be earlier in the target gene but not too close to the start. The other option is to sort by specificity and will show guides at the top of the list that have no, or fewer, off target homologies. Higher specificity is indicative of fewer unintended off target mutations.
- Binding: defaults to show all off target homologies found. For very large off-target database files there will usually be many possible off-target sites found. This combo box allows you to show only those sites that have strong, or moderate, binding.
- Export guide sequence: clicking this icon will copy the sequence of the selected guide to the clipboard.
- Redefine: icon is located at the right end of the toolbar and enables you to redefine the wizard design settings. You can change the guide design region, evaluation conditions and off target search.

CRISPR wizard results output

The results view can be copied to the clipboard or a file using the menu View, Send view options. Results view can be printed using the menu File, Print.

References:

A guide to the CRISPR process: https://www.addgene.org/guides/crispr/

Download Genome sequences for off target searches:
https://www.ncbi.nlm.nih.gov/genome/

Analytical Molecule Operations

Clone Manager can quickly scan your molecule for enzyme sites, entering all found sites in one operation. You can also search for and analyze open reading frames. Clone Manager can also help you analyze DNA repeats, dyad symmetries, hydrophilicity, and protein structure analyses.

- AutoScan for enzyme sites
- Open Reading Frame (ORF) search
- Open Reading Frame analysis
- Analyze molecule
- Mutagenesis Profile
- Export Data

Auto Scan Molecule

Found on the Discover menu or use toolbar button in map window to access. Also available during import of files that contain no enzyme site information (such as GenBank or EMBL files).

You can use this operation to scan a molecule using the AutoScan-designated enzyme list, looking for enzyme sites. Sites found will automatically be entered for this molecule in one batch operation. If the molecule already has enzyme sites, you can decide if you want to replace the existing site list or simply add the new sites (append).

AutoScan will look for single cutters using the Scanner enzyme list unless you designate an alternate enzyme list. The Scanner list contains the major, commercially available enzymes with recognition sequences of 6 or more bases. To designate a different enzyme list, click Set (List) in the AutoScan dialog box.

For program-defined enzyme lists (Commercial (Main), All Enzymes, Scanner List), you can find all single cutters during the AutoScan operation. If you designate a user-defined enzyme list instead, you can find all sites for enzymes on your list or all single cutters.

- **Use AutoScan in the map window**

  If you load molecules in groups, the AutoScan operation will be bypassed. When you open the molecule viewer for a newly loaded molecule with no enzyme site information, you can click the AutoScan button in the map window toolbar to start this operation.

Open Reading Frame Search

Found on the Discover menu or use toolbar button in map window to access.

Use this option to locate Open Reading Frames for the active molecule. You can indicate if you require that each ORF start at a specific codon. Select ATG, GTG, or type in the base letters for the start codon required. Similarly, you can indicate specific stop codons or accept the default settings. Select the checkbox if you want the program to include open reading frames that start within longer ORFs. You can also enter the minimum size cutoff (in amino acids) and the maximum number of open reading frames to be reported.

The program defaults are:

Start codon	ATG
Stop codons	TAA, TAG, TGA
Minimum size	100 amino acids
Maximum number	20

ORF Search Results

When a search for Open Reading Frames is done, the results will appear in a scrolling list box and location map display. The scrolling list shows each ORF found, with its start position (in bps), frame, and the length (in amino acids). Click on the column header to sort by the value in that column.

The red arrow above the molecule map shows the approximate location of the highlighted ORF. The arrow points right for normal strand, left for complementary strand ORFs.

Enter Gene -- enter the highlighted ORF as a gene in the features table for this molecule, with the option to enter a gene name.

Enter All -- enter every ORF as a gene in the features table, automatically entering generic names (ORF-n) with numbers incrementing in the order shown. Sort the display to change the order, if needed.

You can also set the checkbox to **Show Overview** to see all the listed open reading frames positioned in their approximate locations in the six translation frames. The highlighted ORF is shown in red.

Analyze Molecule

Found on the Tools, Analyze molecule menu.

You can do some basic analyses on your DNA molecules. You can analyze a molecule to find open reading frames, viewing an ORF map of start and stop codons, which help to identify potential protein coding regions. A numerical value (Fickett's TESTCODE) can help you to determine if a region is likely to encode a protein.

You can analyze direct or inverted repeats and dyad symmetries to identify regions that might be associated with biologically active sites such as control regions, transcription terminator regions, or RNA processing sites.

You can also do some basic analyses on protein (or translated DNA) molecules. You can analyze protein structure for alpha-helix, beta-sheet (extended), beta-turn and coil (random) states using the method of Garnier-Osguthorpe-Robson. You can use hydrophilicity analysis with beta turn prediction to assist in identifying antigenic regions.

To get started:

1. Click Tools, Analyze Molecule.
2. Select the type of analysis.
3. Identify which part of the sequence to analyze.

If the selection (highlight) bar color is difficult to see on your display, you can use File, Preferences, Colors to assign your selected Highlighter color for this bar.

Open Reading Frame Analysis

You can analyze open reading frames for a full molecule sequence, or for a portion of the sequence. You can set the minimum size of the open reading frame (in amino acids) that the program should report.

As a part of this analysis, the program uses Fickett's TESTCODE algorithm to generate a score that can help you determine whether an open reading frame is likely to be a protein coding region. This method is based on the observed asymmetry of base utilization within coding regions.

Score	Prediction
>= 0.95	Coding region
0.74 - 0.95	No opinion
< 0.74	Non-coding region

The analysis results are shown in a table format of numerical values and as a picture overview showing all 6 possible reading frames with stop and start codons. Click the Start Codon toolbar buttons to display analysis results with Any, ATG or GTG start

codons. Use the toolbar buttons to select Text or Graphic displays. Use the Redefine button to return to the Analyze Molecule setup dialog box to select another analysis type or another region of the molecule to analyze.

- **Text Display** -- for each open reading frame, the program reports start and end base pair position, length in amino acids, frame, and Fickett's TESTCODE score values. The first amino acids of the open reading frame sequence are displayed. You can sort the list of open reading frames found by frame, length (in descending order), start position, or score (in descending order). You can also use the Add Feature button to enter the selected ORF as a gene in your molecule's features table.
- **Graphics Display** -- an ORF map shows all 6 possible reading frames. Open reading frames that exceed the minimum size specified and that exceed 5% of the analysis region length are shaded in pale blue. Full height black bars mark terminators. Half height-colored bars can be used to mark start codons.

Reference: Fickett, James W. 1982. Recognition of protein coding regions in DNA sequences. Nucleic Acids Res 10:5303-5318.

Analyze Molecule Toolbar

The toolbar is shared by all the analysis methods in this section. Buttons that do not apply to a particular analysis will be grayed out and not active.

- Redefine – return to the analysis dialog box to change the analysis.
- Sort selector – changes how text list results are sorted.
- Text – display a table of results. The columns will depend on the analysis method.
- Graphics – display a graphic image of the results.
- Start codon – select the start codon to use for finding open reading frames.
- Add feature – enter the selected open reading frame as a feature.
- Repeats type – show results for: Direct repeats, Inverted repeats and Palindromes.

DNA Repeats Analysis

You can analyze direct or inverted repeats (and dyad symmetries using a separate option) to identify regions that might be associated with biologically active sites such as control regions, transcription terminator regions, or RNA processing sites.

The repeats found are shown in a table display. Use the toolbar buttons in this window to move between the three repeats displays (direct, inverted or palindromes) or return to the analyze set-up dialog box (Redefine). You can sort the list of repeats by length (in descending order), or by position of the first or second repeat.

If the number of repeats found exceeds 100, the program will eliminate from its list all repeats containing the smallest number of bases (for example, all 4-base repeats may be dropped).

- **Direct Repeats**

  Click the Repeats Type toolbar button in the Analyze Repeats display window and select Direct Repeats.

  The program identifies all repeated sequences of four or more bases that are oriented in the same direction (head-to-tail) in the region to be analyzed. The repeated sequences can be abutted (tandem repeats) or separated by unrelated sequences.

- **Inverted Repeats**

  Click the Repeats Type toolbar button and select Inverted Repeats.

  The program identifies all inverted repeats of four or more bases in the region to be analyzed. (If you want to limit the search to those inverted repeats that would be able to form stable stem-loops or hairpin structures, use the Dyad Symmetries analysis.)

- **Palindromes**

  Click the Repeats Type toolbar button and select Palindromes.

  The program identifies inverted repeats of four or more bases that are not separated by any unrelated sequence. Such DNA sequences read the same on both strands.

Dyad Symmetries

You can analyze dyad symmetries (and direct or inverted repeats using a separate option) to identify regions that might be associated with biologically active sites such as control regions, transcription terminator regions, or RNA processing sites.

This option identifies sequences which could form stable stem-loop or hairpin structures. The program looks for sequences that are repeated in the opposite direction on the same strand and then uses RNA energy parameters to determine if each stem loop would be stable.

The program reports the largest and most stable duplexes which could be formed in descending order (up to a maximum of 100 duplexes with a negative delta-G value). Please note that G:U base pairs can be stable under these circumstances and may be a part of duplexes found.

You can use the Redefine toolbar button to return to the Analyze Molecule set-up dialog box to select another analysis type or another region of the molecule to analyze.

Protein Structure Analysis

You can analyze protein structure for alpha-helix, beta-sheet (extended), beta-turn and coil (random) states. Predictions of secondary structure are calculated using the method of Garnier-Osguthorpe-Robson. You can use hydrophilicity analysis with beta turn prediction to assist in identifying antigenic regions.

The analysis results are shown in a table format of numerical values and as a set of line graphs. Use the toolbar buttons to select Text or Graphics displays. Use the Redefine toolbar button to return to the Analyze Molecule set-up dialog box to select another analysis type or another region of the molecule to analyze.

Text Display -- for each amino acid residue, numerical values show the likelihood for that residue to be in helix, sheet, turn or coil states. The columns on the far-right use a plus symbol (+) to indicate which state is most likely.

Graphics Display -- three-line graphs plot the predicted values for alpha-helix, beta-sheet, and beta-turn configurations. For all three graphs, values above the axis line represent regions which are likely to form helix, sheet or turn structures.

Reference: Garnier, J., D. J. Osguthorpe & B. Robson. 1978. Analysis of the Accuracy and Implications of Simple Methods for Predicting the Secondary Structure of Globular Proteins. J. Mol. Biol. 120:97-120.

Hydrophilicity (Protein) Analysis

You can analyze protein (or translated DNA) molecules for hydrophilicity. Predictions are calculated using methods derived from Kyte & Doolittle and Hopp & Woods, and for surface exposure. The analysis results are shown in a table format of numerical values and as a set of line graphs. Use the toolbar buttons to select Text or Graphics displays. Use the Redefine button to return to the Analyze Molecule setup dialog box.

Text Display -- for each amino acid residue, numerical values are given for Kyte-Doolittle, Hopp-Woods, and Surface-Exposure analyses. All three methods provide an indication of the hydrophilic character of the environment, and the range for each analysis is shown at the top of the column.

Graphics Display -- three-line graphs plot the values calculated for each of the three analyses. For all three graphs, values above the axis line are hydrophilic or predicted to be exposed at the surface of the protein.

- ### Kyte & Doolittle Analysis

  Kyte & Doolittle values fall within a range of +4 to -4, with hydrophilic residues having a negative score. The most hydrophilic reside has a value of -4.5 (arginine). On the graphic display, values above the axis line are hydrophilic; values below the axis line are hydrophobic.

Kyte & Doolittle represents a composite hydrophobicity scale derived from interpretation of free energy changes on a water-vapor phase transition and an analysis of buried side chains. Each value is the average of the values of 5 adjacent residues and is plotted at the middle residue. The range of values is approximately ±4 relative units.

Reference: Kyte, Jack & Russell F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157:105-132.

- **Hopp & Woods Analysis**

Hopp & Woods values fall within a range of -3 to +3, with hydrophilic residues having a positive score. The most hydrophilic residues have a value of +3.0. On the graphic display, values above the axis line are hydrophilic; values below the axis line are hydrophobic.

Hopp & Woods derived hydrophilicity values from a study of antigenicity and adjusted the values to maximize the accuracy of predicting antigenic determinants. Each value is the average of the values of 6 adjacent residues and is plotted at the middle point. The range of values is approximately ±3 relative units.

Reference: Hopp, Thomas P. & Kenneth R. Woods. 1981. Prediction of protein antigenic determinants from amino acid sequences. Proc. Natl. Acad. Sci. USA. 78:3824-3828.

- **Surface Exposure Analysis**

The surface exposure value is presented as a proportion of the residue which is exposed on the surface of the protein. These values fall within a range of 0 to 1.000. The most exposed amino acid has a value of 0.97 (lysine). On the graphic display, peak values which fall above the axis line are predicted to be exposed on the surface of the protein.

Surface exposure analysis uses the data of Janin, et al, which provides values representing the fraction of residues of a given amino acid that have a surface area of greater than 20 angstroms squared. High values therefore represent amino acids that are likely to be exposed on the surface of the protein. Plotted values are the average of 6 residues and are plotted at the middle point.

Reference: Janin, Joel, Shoshanna Wodak, Michael Levitt & Bernard Maigret. 1978. Conformation of Amino Acid Side-chains in Proteins. J. Mol. Biol. 125:357-386.

Mutagenesis Profile

Found on the Tools menu. You can also access this feature using the Tools menu and selecting Open Mutagenesis Profile on the edit tab of the primer viewer window.

You can view sequence of your molecule around the target position specified. To see more bases, make the window wider. Restriction enzyme sites are shown above the sequence and one frame of translation below. You can change one or more bases in the sequence and see if the change affects enzyme sites or the translation. You can also find 'almost' enzyme sites (one-base mismatches), view base changes required to change from one amino acid to another, and view possible silent changes. You can instruct the program to create a mutant primer or molecule incorporating your changes.

The currently active base is marked with a vertical high light bar and can be moved using the keyboard left or right arrow keys, toolbar arrow buttons, or by selecting the base using the mouse.

- **Restriction enzyme sites**

  The sequence will be scanned using restriction enzymes recognized by the program (commercially available, 6 base cutters with no ambiguous bases). Enzyme sites are shown by a horizontal line, indicating the length of the recognition sequence. The cut position is marked by a small pointer and the enzyme name appears over the line.

- **Translation**

  One frame of translation is shown. You can use the toolbar control to change the frame, if necessary. If you identify the target position using the features table, the gene selected is shown in the lower left and the frame is automatically set. Termination codons are shown as ---.

- **Changing bases in the sequence**

  You can change bases (mutate) the sequence shown in this display. Click on the base that you want to change or move the vertical ruler using the toolbar or keyboard left or right arrow keys. The base pair position is reported at the top of the ruler line. With the vertical ruler line on the base that you want to change, click the A, C, G or T letter keys on the keyboard to change the base to the letter key pressed. Or click the Edit Base button and select the base change you want.

  Base changes appear in red (or the assigned alert color) below the sequence lines. When a base change adds an enzyme site, the enzyme name appears in color. When a base change removes an enzyme site, the enzyme name is crossed out with a colored line. When a base change results in a change to the protein translation, the amino acid specified by the mutation appears in color below the original amino acid.

- **Finding 'almost' enzyme sites**

  On the toolbar, click Tools, Find Almost Enzyme Sites to access this view.

  You can use this option to find 'almost' enzyme sites (one-base mismatches to the recognition sequence of a restriction enzyme). The sequence will be scanned using restriction enzymes recognized by the program (commercially available, 6 base cutters with no ambiguous bases) and a list of 'almost' enzyme sites constructed.

  The list shows each single base change that could introduce an enzyme recognition sequence, the position for the change, the name of the enzyme site that would be introduced and the type of end produced by cutting with this enzyme. You can sort this list by position or enzyme name by clicking on the column header.

  Select the change you want to make and click the Change Base button. The sequence on the Mutagenesis Profile display will be changed at the position specified. Newly introduced enzyme sites appear in red (or the assigned alert color) above the sequence.

- **Mutate - Change Amino Acid**

  You can use this option to find out which bases you would need to change in the sequence so that an existing amino acid is replaced with another amino acid.

  Click on the amino acid you want to change or move the vertical ruler line to the amino acid. On the toolbar, click Tools, Change Amino Acid and then use the Change To combo box to select the amino acid you would like to introduce. The list shows the codon positions and the base changes that would need to be made in each position to specify the amino acid shown. You can move the highlight bar in this list to select the row containing the base changes you would like to make and then click the Make Changes button.

  The sequence on the Mutagenesis Profile display will be changed, as instructed, and the change in the translation will appear in red (or the assigned alert color). If the base changes remove an enzyme site, the enzyme name is crossed out with a colored line. If an enzyme site is added, the new enzyme site appears in color.

- **Mutate - Show Silent Changes**

  You can use this option to see which bases in the sequence you could change without changing the translation shown.

  On the toolbar, click Tools, Show Silent Changes button to toggle this display on or off.

  The display shows possible silent base changes directly below the upper strand of sequence. If only one base change can be made at a position without changing the translation, this base is shown as a letter (a, c, g, or t). If more than one base change is possible at this position, an asterisk character (*) is used to indicate the multiple possibilities.

Click on the base that you want to change or move the vertical ruler line to the base. Use the Edit Base toolbar button or press the letter keys A, C, G, or T on the keyboard to make the change to the sequence base. If the base change removes an enzyme site, the enzyme name is crossed out with a colored line. If an enzyme site is added, the new enzyme site appears in color.

Click Tools, Show Silent Changes again to clear this display.

- **Creating a mutant primer or molecule**

  On the toolbar, click Tools, Create Mutant to access this feature.

  After you have updated the bases shown in this display (introduced a mutation), you can use the Create Mutant feature to have the program create a primer or molecule that contains the modified bases.

  If you select the option to create a mutated molecule, the program will enter the molecule to the molecule list and then open the molecule viewer window for this new molecule.

  If you select the option to create a primer with the mutation, you can specify the criteria type you want to use to evaluate the primer (usually PCR). The program will select a primer length that includes your modified bases, and you can indicate if you want the complement strand to be used. The program will analyze all possible primers that would include the mutated bases and select the best choice, opening the primer viewer window for this primer. You can analyze this primer, give it a more meaningful name and description, and save the primer to disk.

Molecule Graphic Maps

The graphic map is a high-resolution map that can be sent to your printer, copied to the clipboard, or exported to a disk file for use in other programs.

In addition, you can create as many as 6 graphic map enhanced views for each molecule file. These special enhanced views might be used to customize a molecule map specifically for use in a presentation or publication.

- Graphic maps - overview
- Format Map Screen Basics
- Format Map Print Basics
- Format Map Features
- Format Map Sites
- Format Map Content

Graphic Maps – Overview

The graphic map is a high-resolution map that can be sent to your printer, copied to the clipboard, or exported to a disk file for use in other programs. Load a molecule into the program or select a molecule from the Molecule List to open the molecule viewer window. The map in this window is a lower resolution, working map of the molecule.

- **Set Display Options**

  Click the Settings button to change the appearance or content of the high-resolution graphic map. You can turn on or off the printing of types of basic text information, select feature styles and colors, select enzyme or primer site styles, or change the size or position of the graphic map or add descriptive text to the map page.

  Changing map print options will not affect the appearance of the working map in the molecule viewer window (except for feature styles and colors), but will affect printed, copied, or exported maps.

- **Print or Copy Graphic Map**

  You can use options on the main File menu to preview or print the high-resolution graphic map. Options on the main View menu to copy the graphic map to the clipboard or to a file for use in another program. The highest resolution graphic maps are exported as vector graphic files (*.wmf or *.emf), which can easily be scaled in their destination application. You can create an enhanced view map and export it as a raster graphic file, if needed.

You can create as many as 6 graphic map enhanced views for each molecule file. These special graphic maps might be used to customize a molecule map specifically for use in a presentation or for a figure for a publication.

Format Map Screen Basics

You can set working map (screen) display options. These settings will be remembered and used for all working map displays. You can also select the list of sites to display (at right) and the site list sort order for the current window.

- **To set working map display options**

1. With the map window open, click the Settings button.
2. Click the **Screen Basics** tab, if needed.
3. Use checkboxes to enable or disable options (see below)
4. Click OK to apply and remember your settings.

- **Selected Features**

Set the checkbox to show a color segment on the map when a feature is selected. (Click on a feature on the map to select or select from the list at right.) The color segment can be helpful in marking the feature extent close to the map line, where enzyme or primer sites are positioned.

- **Complement Strand Primers**

Set the checkbox to show complement strand primers in a distinctive color so that you can distinguish complement from normal strand primers on the map. Normal strand primers appear in the color specified for enzyme or primer names (Site tab). The color used for complement strand primers is the Auxiliary text color, set in File, Preferences, Colors.

- **Mouse Hover**

When this feature is enabled, you can move the mouse to an enzyme or primer site or a molecule feature and see the base pair position for this element without clicking the mouse. If this feature is disabled, you can point and then left click with the mouse to get the same base position information.

- **Line Smoothing**

When enabled, anti-aliasing is applied to feature and map lines to improve the appearance of the items in the on-screen map. This does not affect printing and is not applied to maps copied to the Windows clipboard in vector formats. If disabled, the feature and map lines are drawn without smoothing.

- **Suppress very small Features**

If you are using large molecules (greater than 100,000 bp), you can elect to suppress the drawing of features that are very small relative to the size of the molecule. With genome-sized molecules, you can use this feature to avoid drawing maps that are so crowded with features that they are not legible.

1. Click the Settings button on the map page and then click the **Features** tab.
2. Use the checkbox to suppress the drawing of very small features on the map.

- **Suppression Rules**

  Name suppressed -- if the gene is less than 1/500 (0.2%) of the molecule size, the feature name will not appear.

  Feature and Name suppressed -- for circular molecules, if the gene is less than 1/2500 (0.04%) of the molecule size, neither the feature (arrow, box, pointer, etc.) nor the feature name will appear. For linear molecules, if the gene is less than 1/1000 (0.1%) of the molecule, neither the feature nor the feature name will appear

  You can view less than the complete molecule by using the Zoom feature. When you zoom to focus on just a part of a large molecule (perhaps 10,000 bases), you will get a clearer view of small features.

Format Map Print Basics

You can select the font face and colors used to print text information on your graphic maps and you can control the thickness and color of the mapline:

- **To set basic options for the graphic map**

1. Click the Settings button and then click the **Print Basics** tab.
2. Set one typeface (font) for all text on your map. Arial is a simple, modern, sans serif typeface, while Times Roman is a more classic, slender, serif typeface.
3. Use checkboxes to indicate which text items to print
4. Select a color for each item type.
5. Select a color and weight for the mapline.

Blue and black are the most readable colors for text and would be good for enzyme names. Feature names could be in a bright color to draw attention, or in a basic black to minimize distraction. You might want to try using a lighter color for less important items like base pair interval numbers (especially if your map starts to look a little crowded).

Format Map Features

You can control the style and color for feature types (genes, regions, and markers) to be printed on your graphic map. Within the broad categories of genes, regions, and markers, you can set three distinct styles. The shape/fill pattern and color settings you select will apply to all features of that type.

You can use a custom feature setting to distinguish a special feature, if needed, and you can move a feature name to avoid overwriting if necessary. If you are using very large molecules, you can elect to suppress the drawing of very small features.

Feature positions are preset -- genes will be displayed inside the map line, regions will be displayed on the map line, and markers occupy one base position and will be drawn around this base position on the map line. (If you need more flexibility with feature positions, you can use an Enhanced View map.)

To set feature print styles

1. Click the Settings button and then click the **Features** tab.
2. Use checkboxes to indicate which feature types to print.
3. Select a shape/fill pattern and color for Style 1 of each feature type.
4. Select a shape/fill pattern and color for Style 2 and Style 3 of each feature type, if desired.

To set custom feature settings

1. With the map visible on the screen, click to select a feature and then click Tools, Customize Feature Style.
2. Or right-click on a feature and select **Customize.**
3. Select the custom style and color for this one feature or Default to remove this custom setting.

To shift a feature name

1. With the map visible on the screen, right-click on a feature and select **Shift Name**
2. Or select a feature and then click Tools, Shift Feature Name.
3. Enter the number of increments (in units' of 1/10 inch) to move the text horizontally or vertically on the printed page.

   Solid arrows or boxes can be bold and eye-catching. Use solid styles as a custom setting with a bold color for an important gene or use for all features of one type with lighter colors.

Shaded features are edged in solid color and filled with a muted shade of that color. Vibrant colors often do best with shaded styles, especially if you have a large number of features.

Open styles are edged in color and are not filled. They are understated and will not draw attention away from other features.

Line or Open Box styles are very useful for maps with many features.

Format Map Sites

You can control the size and color for enzyme and primer sites to be printed on your graphic map. You can select one of two preset font sizes for enzyme and primer names, and you can control the color to be used for these names and tick lines. You can also indicate if the program should print base pair numbers for each enzyme or primer site or if it should italicize the first three letters of the enzyme name.

- **To set enzyme print styles**

1. Click the Settings button and then click the **Sites** tab.
2. Set the font size for enzyme or primer names to Standard or Small.
3. Select a color for names and tick lines.
4. Use the checkboxes to print base pair numbers or italicize enzyme names.

The standard font size for enzyme names is 10 points. You can select a smaller font (7 points) for enzyme or primer names if you have a very large number of enzymes and are experiencing problems displaying all the enzyme names.

Format Map Content

You can print stand-alone graphic maps, or you can select a pre-defined map format that adds additional text information to the printed map page. If you do not add additional text, you can move or resize the graphic map before printing or copying to the clipboard.

In almost all cases, the default map size and position will print a complete, centered, high-resolution graphic map in either landscape or portrait orientation. (To change the page orientation, select Print Setup from the File, Other Tools menu).

- **Maps with Additional Text**

You can select a graphic map style for printing that adds additional text to the map page. You can select to include the molecule description and notes and/or a list of features and map sites.

- **To print a map with text annotations**

1. Click the Settings button and then click the Content tab
2. Select **Molecule Map and Text**.
3. Use checkboxes to indicate which text you would like to include.

The graphic map will automatically be reduced to 90% of full size and it will be moved upward from the center to make room for the extra text. (This format is intended for printing in portrait orientation.) The text information will be printed on the lower section of the map page and may go on to a second page, if required.

- **Changing Map Size**

You can change the size of a graphic map before printing. You can also resize a map before copying to the clipboard or exporting.

1. Click the Settings button and then click the Content tab.
2. Select **Molecule Map Only**.
3. Select 80% or 90% to decrease the map size.
4. Or select 110% or 120% to increase the map size.

Making a map slightly larger will give more room for enzyme and feature names since the font sizes used for the text do not change when the map size changes. Making a map slightly smaller may provide a better figure for export to your word processor or drawing program. Less reduction will be required within the importing application and enzyme names will not get as small and map lines and enzyme tick marks will not get as thin as when a larger map is reduced.

- **Changing Map Position**

  You can change the position of a graphic map on the printed page, if necessary. You can shift the map up or down on the page, or you can shift the map slightly to the left or right.

  1. Click the Settings button and then click the Content tab
  2. Select **Molecule Map Only**.
  3. Move the map left or right by using the horizontal position slider control.
  4. Move the map up or down by using the vertical position slider control.

  The default position is centered horizontally and vertically on the printed page. Click the Defaults button to return the map to its original position (centered) and size (100%).

Enhanced Maps

You can use Enhanced View to create a special graphic map, perhaps for a presentation or a publication. You might want to enhance the molecule map to differentiate or emphasize an important set of features, show only the most significant enzyme or primer sites, or control line thickness and color to give the best visual image.

- Overview
- Getting started with Enhanced Views
- Enhanced Map Toolbar
- Enhance feature drawing
- Working with color
- Saving Enhanced Views
- Save or Send Enhanced View Maps

Overview

When you start a new Enhanced View, you begin with the molecule map, formatted for printing using your Clone Manager preferences. Then you can adjust each individual feature's style, color, pattern, or position. You can drag feature names with the mouse, to position them just where you want them. You can directly control the font face, font style, point size and color for enzyme, primer, or feature names, or the molecule name or size text. If needed, you can remove features or enzyme sites from this map, without removing them from your molecule file. You can control the thickness and color of the enzyme and base pair interval tick lines, as well as the map line. You can even switch between circular and linear map formats.

When complete, you can save the Enhanced View to the molecule file for later use. As many as 6 enhanced views can be saved with any molecule file. You can print the enhanced map, copy it to the clipboard or export it to a disk file in standard graphic formats for use in other programs.

- **Standard graphic maps vs enhanced views**

  Clone Manager can generate standard high-resolution graphic maps, also suitable for publication or presentations. By design, all the maps will look consistent in style. You store your map preferences like typeface, gene style, region color, etc. in the Format Map print options settings. Each time you print a map or copy to the clipboard, the program uses the map print option settings to format your map. If the settings are unchanged, each of your maps will look similar. If you do change the settings, all maps that you print or copy thereafter will use the new settings. In Enhanced Views, however, styles, fonts and color settings are saved as a part of the map data and are independent of the standard Clone Manager preferences.

Getting Started with Enhanced Views

To create a special, enhanced map of the active molecule:

1. Open the molecule viewer window and click the Map tab.
2. On the map window toolbar, click Enhanced View, New View.

The new enhanced view window will open with your molecule map, drawn using the Format Map print options settings stored in Clone Manager. Use your mouse to drag on the lower right corner of the window to increase the window size, as needed. Toolbar buttons and controls let you modify the appearance of text, lines, and features drawn on the molecule map. You can click the Undo button to reverse the last action taken, if needed.

Basic molecule drawing

Click the Settings button on the toolbar to modify basic settings for the molecule map, features, or enzyme/primer sites. Basic settings include map format, line widths, line colors, base pair interval markings, special text display settings and font characteristics.

Text

Use the font controls in the toolbar to change the text characteristics for a selected text item. Similar items (enzyme/primer names, bp interval numbers, feature names) will all be formatted to the same font face, style, point size. Click the Color button to change the color of the selected item.

Features

Click on a feature (gene, region, marker) drawn on the map to activate the toolbar buttons that enhance feature drawing. You can change the feature style, position, fill color, fill pattern, and feature name text color. To remove a feature not needed for this map, click to select, and then click the Delete toolbar button.

Enzyme sites

Click Settings, Enzyme Site Basics add base pair numbers for enzyme sites or to italicize the first three letters of enzyme names. To remove enzyme sites that are not needed on this map, click to select, and then click the Remove Selected button.

Adding features or sites

The enhanced view of the molecule map is built from the data in the Clone Manager molecule file. You can add features and enzyme sites or primer sites that are a part of the molecule but not currently displayed on the enhanced view map. Click the Add button and select add features, add enzyme sites, add primer sites, or add label sites.

Copy or export map

Click the Save or Send button to send your map to the Windows clipboard or to a disk file in vector or raster (bitmap) file formats. Click the Print button (main toolbar) to send your map to the printer.

Enhanced Map Toolbar

The enhanced map toolbar provides functions for customizing display options for the components of the enhanced graphic map of the molecule. Some of the functions act on the selected item(s) and will only be available if the appropriate item is selected. For example, changing the display style for a feature requires the feature to be selected by using the mouse to select the feature. Some of the functions, for example color, can act on multiple items. You can select multiple items by holding down the control key as you select items.

- Save or Send – Used to save the view, copy, and export the map
- Settings
    - Map Basics – How the main molecule is drawn.
    - Feature Basics – Customize the display of features.
    - Site Basics – Customize the display of sites.
- Add
    - Features
    - Enzyme, Primer, and Label sites
- Remove – the selected item(s).
- Undo – the last action.
- Edit selected feature.
- Move Feature
    - Out or up depending on whether the map is circular or linear.
    - In or down
- Feature Style – change how the selected feature is drawn.
- Feature Pattern – change the pattern used to fill the selected feature.
- Color – change the color of the selected item(s).
- Zoom In – Magnify the map.
- Zoom Out – reduce the size of the map.
- Font face – change the font used for the selected item. (Arial, Times, Verdana, or Tahoma)
- Font style – change the selected item(s): Regular, Bold, Italic or Bold Italic.
- Font size – change the font point size for the selected item(s).

Enhance Feature Drawing

You can set the drawing style, color, fill pattern, and position for each individual feature in the enhanced view map. Select (click on) the feature and then use the toolbar buttons described below.

Feature Position

Nine pre-set positions can be used. Four positions are inside/below the map line, one position is on the map line, and four positions are outside/above the map line.

Click the Move Feature Out/Up button to move the feature to a position further from the center of a circular map or to a higher position on a linear map.

Click the Move Feature In/Down button to move the feature to a position closer to the center of a circular map or to a lower position on a linear map.

Feature Style

Eight drawing styles are available for genes or regions. Narrow styles are most useful for maps that have many features, while wider styles provide greater impact. Markers have a subset of drawing styles, drawn around the single Marker base pair position.

Click the Feature Style button and select the drawing style for this feature.

Feature Pattern

Eight fill patterns can be used for genes or regions. Select Open (no fill), Solid, Shaded, or patterned with vertical, horizontal, or diagonal lines. Markers are always drawn with the Open style. Fill patterns can be printer- and resolution-dependent.

Click the Feature Pattern button and select the fill pattern for this feature.

Feature Color

You can select the color for the feature drawn on the map and you can select the color for the individual feature name. Twenty standard colors can be used, or you can click More Colors to select a non-standard (custom) color.

Click the Color button and select the color for this feature or feature name text.

Move feature name

On the molecule map, click on the feature name to select. Then drag the name with the mouse to move to a new location, as needed.

Change feature details

You can edit feature names or base pair positions, if needed. Select the feature, then click Edit Selected Feature. Click in the feature name or base pair position fields to edit this information.

Working with Color

Color can help to make your enhanced view maps more interesting and easier to understand. Bold, bright colors catch the viewer's eye and indicate the importance of an object. Too much color, however, can lessen this impact.

You can change the color for individual features or for text elements.

1. Select (click on) the item.
2. Click the Color button.
3. Click on the color you want or click More Colors to select a color not shown.

- **Using more colors**

When you click the Color button, you will see a color selector that shows the 16 standard colors. In some cases, you just need one or two different colors to complete your figure or match a color in use elsewhere. To do this, you can select custom colors, using controls to mix the color components or you can type in numerical values.

Easy steps to a custom color:

1. Click the Color button, then click **More Colors**.
2. Click on a basic color or on an area in the color matrix that is close to what you want.
3. Check the Color|Solid sample box (lower right)
4. Use the slider at the far right to make the color lighter or darker.
5. When the color is right, click **OK** to use this color.

The next time you click the Color button, your custom color will be displayed in the bottom row (above More Colors), so that you can easily use this color again.

- **Line colors**

You can change the color for the molecule map line, base pair interval tick lines, and enzyme or primer site tick lines. Click the Settings button and select Map Basics (map line, bp interval lines) or Site Basics (enzyme/primer lines) and select colors and line widths.

Map Basics

You can select basic options for the molecule map in Enhanced Views of the graphic map.

Map options

Select to draw your molecule as a circular or linear map. If the molecule is defined in its data file as circular and you select to draw it in the linear format, the program will open out the molecule at base pair position 1. If the molecule is defined as linear and you select to draw it in the circular format, the program will draw the molecule map as a circle, where the two molecule ends meet at the top. When a molecule is defined as linear, features cannot go over the top.

You can draw the map smaller or larger than the normal (100%) size. Select 80, 90, 100, 110, or 120% of normal. Adjusting the size of the map can give more room for features or enzyme sites or change the appearance and balance of the map elements.

Click the Font buttons to set font characteristics for the molecule name and molecule size text. Clear the check boxes to turn off the display of either item. If you need to move the molecule name or molecule size text on the map, select the text (click on it) and then drag the text using your mouse.

Select the width and color to be used for the Map Line.

Base pair intervals

Base pair intervals can be numbered in bps or kbps, and you can set the font characteristics for the interval numbers. Clear the checkbox if you do not want to display the base pair interval numbers.

Base pair interval ticks and numbers can be displayed inside the circular map, at the map line, or in an intermediate position. For linear maps, these intervals can be displayed at the map line, on a ruler line below the map (select Standard), or in an intermediate position. Clear the checkbox if you do not want to display interval tick lines or numbers in any position.

Select the width and color to be used for the base pair interval tick lines.

Font settings and text color selections can be made using toolbar controls (select the text first). All other map settings must be made in the Map Basics dialog box (click Settings).

Feature Basics

You can select basic options for all feature names in Enhanced Views of the graphic map.

Select the Font Face, Font Style, and Point Size to be used for all feature names. The most common settings are Arial, Regular, 10 points. If you have a large number of features, select a smaller point size. You can also select a color to be used for all feature names or you can select individual colors for each feature name, if desired (see below).

Move feature name

Select (click on) the feature name and drag to a new position using the mouse. If you expect to change the size of the molecule map or adjust the position of the features, move the feature names after you have made the other changes.

Change feature name

Right-click on the feature name and select Edit Feature. Click in the Name field and make any changes to the feature name -- shorter names are often helpful for a crowded map. (You can also change all style settings for this feature in this dialog box.)

Change color for one feature name

Select the feature name on the map and then use the Color toolbar button to select a color for this one feature name.

Font settings can be made using toolbar controls (select a feature name first). If you want to set one color for all feature names, you must click Settings and select the Feature Basics dialog box.

Site Basics

You can select basic options for all enzyme, primer, and label sites (map sites) in Enhanced Views of the graphic map.

Select the Font Face, Font Style, and Point Size to be used for all map site names. The most common settings are Tahoma, Regular, 10 points. If you have a large number of map sites, select a smaller point size. You can select one color for all enzyme names, one color for all primer names, and one color for all label names.

Site options

You can use the checkboxes to set the display of base pair numbers for all map sites or to italicize the first three letters of enzyme names.

Tick lines

Select the width and color to be used for map site tick lines. (These lines extend from the map line to the enzyme, primer, or label name). When you have many enzyme sites and are using a small point size for enzyme name text, select a thin tick line. If your enzyme names are large or you are using a thick map line, select a thicker tick line for balance.

Font settings and color selections can be made using toolbar controls (select a map site name first). Tick line width and color can only be set in the Site Basics dialog box (click Settings).

Add Features to Enhanced View Map

The enhanced view of the molecule map is built from the data in the Clone Manager molecule file. You can add features, enzyme sites, primer sites, or label sites that are a part of the molecule but not currently displayed on the enhanced view map.

Click the Add button and select Add Feature.

The list will show features on the features table in the molecule file that are not currently shown on the map. If you have information-only features in your molecule file, you can use the checkbox to show this feature type.

Sequence labels can appear in this list since they are not displayed on the map as features. You can select a label feature and have it drawn as a region or gene. It is possible that some labels on this list already appear on the map as map sites (label sites), rather than features.

Once you have selected the feature you want to add to the Enhance view map, click OK. Next, assign the style settings for this feature.

If all the features are already on the map, the list box will be blank.

- **New features**

  If you need to add a feature that is not a part of the current Clone Manager molecule, leave the enhanced view map on the screen and go back to the molecule viewer window (usually the window immediately below the enhanced view). Click the Features tab and enter the new feature to the molecule. Then return to the enhanced view and click the Add button and repeat the steps described above to add this new feature to the enhanced view map.

Add Sites to Enhanced View Map

The enhanced view of the molecule map is built from the data in the Clone Manager molecule file. You can add features, enzyme sites, primer sites, or label sites that are a part of the molecule but not currently displayed on the enhanced view map. To add additional map sites to the enhanced view map:

Click the Add button and select Add Enzyme Sites, Add Primer Sites, or Add Label Sites.

If you selected the option to add enzyme sites, the list would show enzyme sites present in the molecule file that are not currently shown on the map. If you previously deleted enzyme sites from the enhanced view map, you could add them back using this procedure. If you selected the option to add primer sites, the list would show primer sites present in the molecule file that are not currently shown on the map. If primer sites were present in the file, but not visible on the map at the time the enhanced view was created, you would need to use this procedure to add the primer sites to the enhanced view map.

If you selected the option to add label sites, the list would show label sites present in the molecule file that are not currently shown on the map. If label sites were present in the file, but not visible on the map at the time the enhanced view was created, you would need to use this procedure to add the label sites to the enhanced view map. Note that if you added a label as a feature, designating the gene or region display style, this label site would still appear on the list of possible label sites to add to the map.

Mouse-click in the checkbox to tag an enzyme, primer, or label site to add to the map. Click the Tag All button to select all sites on the list. Mouse-click a second time to remove a tag from a site, if needed. Click the Clear button to remove all tags.

If all the sites in the molecule file in this category are already on the map, the list box will be blank.

- **New enzyme sites**

  If you need to add new enzyme sites that are not a part of the current Clone Manager molecule file, click Clone, Find Enzyme Sites to find and then add new enzyme sites to the molecule file. Then click the Add button and repeat the steps described above to add these new enzyme sites to your enhanced view map.

- **New primer sites**

  If you need to add a primer site that is not a part of the current Clone Manager molecule file, you can use Discover, Find Primers that Bind to find and then add new primer sites to the molecule file. Then click the Add button and repeat the steps described above to add the new primer site to your enhanced view map.

- **New label sites**

  If you need to add a label site that is not a part of the current Clone Manager molecule file, you can use Discover, Find Sequence Phrase sites to find and then add new label sites to the molecule file. Then click the Add button and repeat the steps described above to add the new label site to your enhanced view map.

Convert Map Data

You can convert the map data for an enhanced view of the graphic map into a format that can be used in programs that require a raster file format. You can export in TIF, PNG, JPG, GIF or BMP file formats.

To convert enhanced view map data:

1. Open the molecule viewer window and click the Map tab.
2. Click the Enhanced View button and select the view you want to convert
3. Click File, Export to open the export module
4. Set the option to convert to 96 dpi raster, or 300 dpi raster (very large)
5. Set the option to send the output to a disk file or copy to the clipboard
6. If sending to a file, select the file format (TIF, PNG, JPG, GIF or BMP)

- **96 dpi Raster**

  Select this option to prepare a low-resolution raster image of the enhanced view map to export to a disk file in TIF, PNG, JPG, GIF or BMP file formats or copy to the clipboard in standard bitmap format. A 96-dpi raster image looks like the map on the screen and may be suitable for use in on-screen presentations. The image size before reduction in another application is 7 x 7 inches.

- **300 dpi Raster**

  Select this option to prepare a high-resolution raster image of the enhanced view map to export to a disk file in Tif, Png, Jpg, Gif or Bmp file formats or copy to the clipboard in standard bitmap format. A 300-dpi raster image looks similar to the printed map and may be suitable for publication. The image size before reduction in another application is 21 x 21 inches.

- **Maps for web use**

  If your image is for use on a web page, and the Clone standard graphic map is sufficient, try the special Save for Web option found on the Map toolbar. This option will produce a raster graphic map in the Portable Network Graphics (*.png) format using a smaller image drawing size (approximately 5 x 5 inches) and will automatically crop excess white space to minimize file size and download time.

Saving Enhanced Views

When you have completed work on your enhanced view map, you can save this special map data to the molecule file. You can later re-open this enhanced view for further work or print or export the graphic map. You can have a maximum of 6 enhanced views saved in each molecule file. You can organize the enhanced views, renaming the views or deleting ones that are no longer needed. Remember to save the molecule file to disk before exiting the program.

- **Save an enhanced view:**

  Click the Save or Send toolbar button in the enhanced view display window. The enhanced view data is saved as a part of the molecule file. The view name assigned consists of the date and time the view data was saved, but you can change this to a more meaningful name (see below).

- **Rename or delete enhanced views:**

  To organize the enhanced views saved to a molecule file, open the molecule in the molecule viewer window and click the Map tab. Click the Enhanced View toolbar button and select Organize Views. In the Organize Enhanced Views dialog box, select (click on) the view that you want to rename or delete. To change the name of the selected view, click the Rename button and then type in the new name. To permanently remove the selected view, click the Delete button.

- **Keeping track of maps**

  Add a one-line comment about the enhanced view map to the molecule Notes field on the Info tab in the molecule viewer window. (Click the Edit toolbar button to enable text entry). Something like "Enhanced View -- map for seminar slide" will give you a keyword or two that can be searched for using Find in Files. You can also view the first line or two of notes in the More Info section of the molecule list to find the file containing your enhanced map.

  Remember to save the molecule file after you have updated the Notes field or saved enhanced view maps.

Save or Send Enhanced View Maps

You can click the Save or Send button to send your enhanced view map to the Windows clipboard or to a disk file in vector or raster (bitmap) file formats. These options are described below. You can also use options on the Map menu to send the map to your printer, preview the map in its printed format, or save the map for use on a web page.

- **Copy to the clipboard (vector)**

  Copies the map to the clipboard for use in other Windows applications. The vector output can be resized accurately and will print with high resolution. (See note for restoring line anti-aliasing, below)

- **Copy to the clipboard (low-res bitmap)**

  Suitable for pasting into another application for on-screen use. This option prepares a 96-dpi raster image that looks like the map on the screen.

- **Copy to the clipboard (high-res bitmap)**

  Suitable for pasting into another application for subsequent printing or higher resolution on-screen use. This option prepares a 300-dpi raster image that looks like the printed map.

- **Export vector (emf, wmf)**

  Prepares a high-resolution graphic map as a Windows Metafile or an Enhanced Windows Metafile for use in other Windows applications capable of importing these file formats. As vector drawings, the figures can be resized accurately.

- **Export low-res bitmap (tif, gif, png)**

  Prepares a low-resolution (96 dpi) raster image that you can save in a variety of file formats. GIF and PNG files provide high quality map images combined with small file size. A 96-dpi raster image looks like the map on the screen and may be suitable for use in on-screen presentations. The image size before reduction in another application is 7 x7 inches.

- **Export high-res bitmap (tif, gif, png)**

  Prepares a high-resolution (300 dpi) raster image that you can save in a variety of file formats. GIF and PNG files provide high quality map images combined with small file size. A 300-dpi raster image looks like a printed map and may be suitable for publication. The image size before reduction in another application is 21 x 21 inches.

- **Save for Web (PNG)**

  Prepares a raster graphic image in the Portable Network Graphics (*.png) format, particularly useful for computer screen display (such as on a web page or in an on-screen presentation). The map size is automatically set to 80%. No other modifications are automatically made for Enhanced View maps (images are automatically simplified

for standard Clone Manager graphic maps). Although this is a low-resolution image, it looks like the map on the screen and the image size before reduction in another application is 3.5 x 3.5 inches.

Primer Designing

The Primer main menu contains the functions for designing and using primers.

The View main menu contains the functions for viewing the Primer list containing loaded primers and those designed during the active session. Also included is changing the active view when viewing primer search results or viewing the details of a primer.

The following sections provide more information on designing primers, viewing primers and search results, direct entry of a primer's sequence, special wizards for cloning operations using primers, use and organization of primers and the Primer collection which provides quick access to commonly used primers.

- Overview
- Designing primers
- Primer Search Results
- Primer Viewer Window
- Direct and special primer wizards
- Analyze primers
- Primer List
- Use, organize primers
- Primer Collections

Overview of Designing Primers

All design options are found on the Primer menu. You can design primers or primer pairs for PCR, you can find a primer matching an existing PCR primer, or you can design sequencing primers or probes.

You can design primers by searching for primers that meet your criteria, by positioning a primer in the molecule sequence, by typing in the sequence of a known primer, or you can use special primer design wizards for sequencing primers, PCR cloning, RT-PCR or LCR. Following primer design, you can analyze the primer or primer pair, save the primers to disk files for later use, or create amplified products using the primers and their source molecules.

- **Searching for Primers**

  Search the molecule and region you specify, looking for primers that meet the evaluation criteria you set. For PCR primer pairs, you can indicate if you want to amplify a complete target region or just a fragment within a specified size range. Results are listed with basic information about each primer or pair.

- **Create Primer from Molecule**

  Position one primer of the length you specify in the molecule sequence and then adjust its position or length until the primer meets your design criteria. Especially useful when primer position is the most critical element in primer design.

- **Direct Enter Primer Wizard**

  You can type in the sequence and other information needed to create a new primer. This might be a primer that was purchased commercially, received from a colleague, or a primer where you already know the primer sequence but want to save the primer, analyze it, link it to a molecule, or use it with another primer to create a product.

- **PCR cloning wizard**

  Design PCR primers to amplify the region of a molecule you want to clone, use enzymes to prepare the vector or the region for insertion, and then generate recombinant molecules. You can view several solutions and save selected primers, amplified products, and recombinants for later use.

- **Sequencing Primer Wizard**

  Design a set of primers for sequencing a molecule region. The wizard will select three non-overlapping primers per sequencing block, where possible. You can export primer sequences, save to a primer collection, or view a map of sequencing coverage.

- **LCR Oligo Set Wizard**

  Prepare a set of four primers for the Ligase Chain Reaction. The program will select four primers, correctly positioned around the target base position you specify, and adjusted for length to have a Tm as close as possible to the Tm you specify.

- **RT-PCR Primer Design Wizard**

  Search for a pair of primers that can be used for RT-PCR, indicating the region where each primer should be located. For genomic DNA, these areas will usually be exons.

Designing Primers

Clone Manager Professional can help you design, analyze, and use primers. You can search for PCR primers or pairs, sequencing primers, or probes. You can also position a primer in the molecule sequence at a target position to initiate primer design at this position. Or enter a primer sequence and other information for primers already in use or purchased commercially.

The *Primer* main menu contains options for designing, analyzing, and using primers.

- Design Primers
  - Create Primer
  - Search for primers
  - Starting values for primer design
- Primer Evaluation
  - Primer Criteria
  - Primer Ranking and Sorting Weights
- Primer Search Results
- Direct and Special Primer Wizards
- Analyze
  - Overview
  - One Primer Analysis
  - Primer Pair Analysis
  - Mix of Primers and Molecules
- Use Primers
  - Create Product
- Settings
  - Default Primer Criteria
  - Default Weights and Values
  - Primer Design Experimental Conditions
- Primer Collections

Create Primer from Molecule

You can create a PCR primer, sequencing primer, or probe using the sequence of a source molecule. The program will position a primer of the length you specify at the initial position you specify. You can then move the primer along the sequence or adjust its length until the primer evaluation meets the criteria set. This design option is especially useful if you must have a primer very close to a specific location in your molecule.

- **To create a new primer**

1. Select menu Primer, Design
2. Set the option **Create from Molecule** and select the Primer Type.
3. Enter the primer length and the 5' extension (if any).
4. Click **Next** to identify the source molecule and target position.
5. Confirm that the source molecule shown is the one you want to use.
6. Enter a target position for initial placement of the primer and use the checkbox if you want to design a primer using the complement strand.
7. Click **Finish** to create the primer and open the Primer Viewer Window to view the primer and its evaluation.

- **Primer Length**

Enter the number of bases you want to use for your primer (required length).

- **5' Extensions**

Use the checkbox to enable this option and enter the sequence for the 5' extension. Use the small button to the right of the 5' extension field to get help with restriction enzyme recognition sequences. Use the checkbox to indicate if the 5' extension should be added to the primer length entered above. (If you do not check this box, the primer length is assumed to include the 5' extension.)

- **Target Position**

Identify a base position to be used as the initial 5' anchor point for the primer. Use the Features... button to view a list of features (genes or regions) in the source molecule. You can select a feature and automatically enter the base position and strand.

- **Sequencing Primers or Probes**

Because of the wide variation in experimental conditions used for hybridization probes or sequencing primers, the default criteria for these primer types have been set very loose.

It is recommended that you update these settings to reflect your preferences and save these settings for future use. Of particular importance are the annealing temperature and the settings for salt and oligonucleotide concentration.

- **To set the default primer criteria for probes or sequencing primers:**

  1. Click Primer, Default Primer Criteria
  2. Click the tab for sequencing primer or probe
  3. Set basic criteria.
  4. Click the Conditions tab
  5. Enter salt and oligo concentration values.

Searching for Primers

Designing and searching for primers is found on the 'Design' menu under the Primer main menu.

- **Primer Design Details**

The default option of the wizard is to search for primers. The other option is to create a primer from a molecule and is described earlier.

Select the type of primer that you want to design. Several of the options are used to support designing pairs of primers for cloning wizards. For classic PCR you have the option of designing a single primer, a pair of primers, or a primer that matches an existing primer.

Designing primers uses the evaluation criteria and settings specific to the type of primer being designed. You can customize the criteria used by clicking the 'Criteria' button.

- **Setting primer length**

Enter the number of bases you want to use for your primer or for each primer in a primer pair. PCR primers are often 18-22 bases in length. You can indicate if this is the required or optimal length. If you select optimal length and set the acceptable range, primers of lengths within the range indicated will be considered.

If you usually design primers of the same length and acceptable range, you can set the default starting values that the program will use. Click Primer, Default Weights, and Values, then click the Starting Values tab.

- **Setting primer 5' extension**

Use the checkbox to enable this option and enter the sequence for the 5' extension. Use the Lookup button to the right of the 5' extension field to get help with restriction enzyme recognition sequences. Recognition sequences for the most useful restriction enzymes are shown in a scrolling box. Move through the list or type in the enzyme name to quickly move to the enzyme of interest. Select the enzyme with its recognition sequence to paste this sequence into the 5' extension box. Set filters to limit the display to enzymes with recognition sequences of a certain size or enzymes producing certain ends.

Use the checkbox to indicate if the 5' extension should be added to the primer length entered above. (If you do not check this box, the primer length is assumed to include the 5' extension.)

- **Settings for Primer Pair searches**

If you are designing a pair of primers, then you can enter separate settings for the forward and reverse primers.

- **Source Molecule**

  The currently active molecule is automatically selected but you can use the ‘Change’ button if you need to select a different molecule.

- **Primer Target**

  If you are designing a single primer, then enter the base pair position of the optimal position of the primer in the molecule. You can also check the ‘Use complement strand’ to design a reverse strand primer.

  If you are designing a pair of primers, then enter the base pair positions of the target region within the molecule.

  The ‘Features’ button allows quick lookup of the positions of the molecule’s features.

One Primer Search Options

- **Use target position**

  Find single primers that match the criteria starting at the target position entered for the primer search. The program will extend its search to the left of the target position the number of bases specified and to the right of the target position the number of bases specified.

- **Anywhere in the region specified**

  Find single primers that match the criteria whose 5' starting position is located within the region specified as the Primer Target region.

Primer Pair Search Options

- **Amplify complete target region**

  Find primer pairs that will amplify the entire region specified. Enter the start and end base position for the target region. The left primer (Primer A) will optimally be placed at the position you enter as the start of the target region and the program will extend its search to the left of this position by the number of bases you specify. The complementary primer (Primer B) will optimally be placed at the position you enter as the end of the target region and the program will extend the search to the right of this position by the number of bases you specify.

- **Amplify fragment in region**

  Find primer pairs that will amplify a fragment of a specific size anywhere within the target region specified. Enter the start and end base position for the target region and the acceptable product length range (minimum and maximum size in bps).

Source Molecule and Target Positions

After providing information about the primers, you want to design, you can enter source molecule, target, and search option information. Search options and the identification of target positions or target regions will vary depending on whether you are searching for primer pairs or single primers.

1. Confirm that the source molecule shown is the one you want to use.
2. Enter the target region or target position and use the checkbox if you want to find single primers using the complement strand.
3. Select the search option and set related values.
4. Click **Finish** to begin the primer search.

Primer Criteria - Overview

To find or evaluate primers, the program refers to the values specified in the Primer Criteria settings. Separate settings are stored for PCR primers, sequencing primers, probes, or specialized cloning oligos. The program also stores the concentrations of salts and oligos (experimental conditions) and uses these values for calculations.

You can change the criteria settings for an individual primer or search, if needed. You can also change the default settings so that all new primers of that type use the modified criteria settings. If necessary, you can change the type of criteria used to evaluate an existing primer.

- **To change primer criteria during search setup**

1. In the setup dialog box, click the Criteria button.
2. Make criteria changes as required.
3. These criteria settings will be used in finding and evaluating primers during this search.
4. Click the Defaults, Set As... button if you want to use the new settings for all future primers of this type.

- **Default Primer Criteria Settings**

You can also view or edit all the primer criteria or oligonucleotide, or salt concentrations used for all types of primers.

1. Click Primer, Default Primer Criteria
2. Click the tab for the type of primer
3. Review or change criteria settings
4. Click the Conditions tab
5. Review or change oligonucleotide or salt concentrations

Criteria for sequen0ing primers or probes

Primer Ranking and Sorting Weights

Search results consist of a list of primers that have passed an evaluation using criteria you specify (GC range, Tm range, stability, dimers, etc.) You can sort the resulting list of primers by rank so that the best primers will be at the top of the list. Ranks are assigned to primers based on scores calculated by multiplying Sorting Weights by numbers the program generates for certain specific quality-associated conditions.

You can modify the Sorting Weights for a completed search to increase the penalty for a condition and thereby change the position of primers in the list. The program stores default Sorting Weights for PCR primers, sequencing primers, probes and LCR oligos and you can change the default settings for these weights, if desired.

- **To change Sorting Weights for a completed search**

1. Click the Sorting Weights button in the results display.
2. Make any changes to the penalty values required. You can enter penalty values between 1 - 5, with 5 being the most severe penalty.
3. Press OK to return to the primer search results.
4. Select 'Rank Sort' if this has not already been selected.

The search results will re-order based on the new sorting weights penalties applied.

- **Default Sorting Weights**

You can view or edit the sorting weights used for all types of primer and the starting values used by the program as defaults in the Design Primer dialog box.

1. Click Primer, Default Weights, and Values
2. Click the tab for the type of primer
3. Review or change sorting weights
4. Click the Starting Values tab
5. Review or change default starting values

- Primer Criteria vs Sorting Weights

Starting Values for Primer Design

You can change the values the program enters as defaults in the Design Primers dialog box. You can enter default values for the minimum and maximum size range for product length (used in some pair searches), default values to extend search regions, and default values for primer length settings.

When you change the default starting values, the program will always use the revised values when the Design Primers dialog box opens. You can, of course, change any of these values for the current design operation.

- **To change default starting values**

1. Click Primer, Default Weights, and Values
2. Click the Starting Values tab
3. Enter the starting values you want the program to use.

The following starting value settings are stored:

- **Product length range**

  Used in some pair searches. Enter the minimum and maximum size in nucleotide bases to identify the acceptable fragment size range. Preset values: Min length 100, max length 500

- **Extend search (primer pair search)**

  For primer pair searches, enter the number of bases to extend the search to the left of the target position for the primer on the normal strand and the number of bases to extend the search to the right of the target position for the primer on the complementary strand. Preset values: 100 bases to left, 100 bases to right.

- **Extend search (one primer search)**

  Used in single primer searches. Enter the number of bases to extend the search to the left of the target position and to the right of the target position. Preset values: 50 bases to left, 50 bases to right.

- **Primer length**

  Enter the primer length in bases and indicate if this is the required or optimal primer length. If optimal primer length is selected, enter the size range of acceptable primers (usually the optimal primer length ± 1 or 2 bases). Preset value: 20 bases, required length.

Primer Search Results

Results of a search for primers are shown as a list in table format, showing basic information about each primer or primer pair. You can sort this list, view a report for the selected primer, enter the primer to the Primer List for later use, or add the selected (highlighted) primer binding site to the molecule file.

Primers can be sorted by position or quality (rank sort), and you can affect ranking by adjusting the Sorting Weights, if required.

If you find fewer primers than you would like (or none), you can switch to an alternate view to see which evaluation criteria eliminated possible primers. You can use the Auto Adjust Length or Auto Adjust Criteria features to adjust and repeat the search to try to find additional primers.

- Results from Primer Pair searches
- Results from One Primer Searches
- Toolbar buttons
- Enter to Primer List
- Finding More Primers
- Adjust (Relax) Primer Criteria
- Auto Adjust Primer Length
- Auto Adjust Primer Criteria
- Primer Search Statistics Display
- Sorting Primer Search Results

Results from Primer Pair searches

For primers from DNA molecules, pair rank is reported, as well as the base pair position, GC content and Tm for each primer in each pair. Max Dimers reports the worst-case inter or intra-primer dimer found in interactions involving the 3' end or at any position. False Priming reports the Tm of the worst-case false priming sites found for each primer or shows '...' if no false priming found. Product length reports size of amplified product in bps.

For primers from protein molecules, pair rank is shown, as well as the position and start base (offset within codon), and average GC% for each primer in each pair. 3' Degeneracy reports the degree of ambiguity within the 6 bases at the 3' end of each primer. Total Degeneracy reports the degree of ambiguity in the total primer sequence for each primer in the primer pair.

Results from One Primer Searches

For primers from DNA molecules, rank, base pair position, primer length, GC content and Tm are reported for each primer. Dimers shows the worst-case dimer with 3' end involvement and the worst-case dimer at any position. False Priming reports the Tm of the worst-case false priming site found for each primer or shows '...' if no false priming found. Worst-case non-specific binding is reported for probes.

For primers from protein molecules, rank, position (codon), start base (offset within codon), primer length, and average GC % are shown for each primer. Degeneracy reports the degree of ambiguity within the 6 bases at the 3' end of the primer or in the total primer sequence.

Toolbar buttons

- **Primer Report**

  View a report (or analyze) the selected primer or pair. The Primer Viewer Window will open with primer report data, and you can access the more detailed analyses here.

- **Enter to primer list**

  Enter the primer(s) to the primer list. You can enter the name, description, and notes. You can also enter the primers to a primer collection and save to disk.

- **Enter primer site to map**

  Allows you to add the primers as sites on the molecule's graphic map as well as the formatted sequence display.

- **Sort Combo box**

  Allows sorting the results by rank, position, and product length for pair searches.

- **Sorting Weights**

  Allows you to change the weights applied to rank the quality of the search results.

- **Search results/Search stats**

  This toggle button allows you to see where criteria eliminated primers. (If no primers were found, the Search Statistics display is the only display available).

- **Auto adjust length**

  Provides a quick way to explore the effects of allowing variation in the length of the primer(s) used for the search.

- **Auto adjust criteria**

  Provides a quick way to explore the effects of allowing variation in the criteria used to evaluate the primer(s).

- **Redefine**

  Opens the search wizard to allow changes.

Enter to Primer List

Found on the search results window toolbar.

Use this function to formally add the active primer to the list of primers. Once entered to the Primer List, you can select this primer by name for analysis or other operations.

During the Enter to Primer List action, you can enter a name and brief description for the primer. This information will appear in the Primer List and on primer reports. You can also enter additional comments about this primer in the Notes field. All this information will be saved when a primer is saved to disk. You can modify this information later, if necessary, from the Info tab in the primer viewer window. Click the Edit Primer Definition button to view or edit this information.

While you are entering the primer to the Primer List, you can also elect to add the primer to your primer collection or save the primer to a disk file. Use the checkboxes (lower left) to select these options.

Finding More Primers

When a search has been unable to find primers that meet all of the evaluation criteria, you can modify the search conditions (vary primer length or relax evaluation criteria) or you can redefine the search region to increase the size of the search range. Use the Redefine button in the results window to return to the search setup dialog box.

Changing the primer length by one or two bases may influence the number of primers found within the area specified and meeting the criteria selected. Or you may be willing to accept a slightly longer run of bases or an increase in non-3' dimers, without changing primer length.

- **Finding More Primer Pairs**

  When a search has been unable to find either normal strand (Primer A) or complementary strand (Primer B) primers, you will want to view the results for this specific primer and take steps to find such primers (see Find More Primers, above). You can view separate statistics displays for Primer A, Primer B, or the primer pairs. The results displayed are identified on the line above the bar charts (Drop-thru Tests Results: Pairs, for example). Click the Primer A or Primer B buttons in the lower right to change the display.

  When a search has found Primer A candidates and Primer B candidates, but is unable to find A:B pairs meeting your criteria, you can modify the matching criteria to accept larger differences between the two primers, find more single primers that can meet your matching criteria, or redefine the search to change the product length range (if applicable).

- **Finding More Matching Primers**

  Primer A is defined as the primer you indicated that you wanted to match. When a search has been unable to find matching (Primer B) primers, you will want to view the results for Primer B and take steps to find such primers (see Find more single primers).

  When a search has found Primer B candidates but is unable to find A:B pairs meeting your criteria, you can modify the matching criteria to accept larger differences between the two primers, find more single primers that can meet your matching criteria, or redefine the search to change the product length range (if applicable).

Adjust (Relax) Primer Criteria

If the search for PCR primers is not successful, the program will attempt to adjust (relax) the primer evaluation criteria. You can indicate if you will authorize Conservative criteria changes, Aggressive criteria changes, or No changes to the evaluation criteria. These adjustment protocols are the same as those used in standard primer searches when you select the Auto Adjust Primer Criteria option.

The adjustment of primer evaluation criteria will be as follows:

- Runs of bases, Repeats, and Dimers will be adjusted by increasing the setting by 1.
- Stability will be adjusted by decreasing the value by 0.5 kcals.
- GC clamp requirement will be adjusted by decreasing the setting by 1.
- Pairs matching values will be adjusted by increasing the setting by 5.
- **Conservative adjustment permitted:**

  Affects Runs of bases, Repeats, GC clamp, and Dimers (any).
- **Aggressive adjustment permitted:**

  Additionally affects Dimers (3'), Stability, Pairs GC Match, Pairs Tm Match.

Auto Adjust Primer Length

Found on the Search Results/Search Statistics display window toolbar.

You can use this function to start an automated sequence that will repeat the primer search, adjusting the primer length as specified, and then display the new results.

- Select **± 1** to check primers at the original primer length, one base shorter and one base longer than the original length.
- Select **± 2** to check primers within 2 bases of the original length.
- Select **No adjustment** to return to the primer length specified in the search setup dialog box.

Auto Adjust Primer Criteria

Found on the Search Results/Search Statistics display window toolbar.

You can use this function to start an automated sequence that will adjust (relax) primer evaluation criteria, repeat the search, and then display the new results. Select the Search Statistics display to see which criteria eliminated primers.

Adjustment of primer evaluation criteria will be as follows: Runs of bases, Repeats, and Dimers will be adjusted by increasing the setting by 1. Stability will be adjusted by decreasing the value by 0.5 kcals. GC clamp requirement will be adjusted by decreasing the setting by 1. Pairs matching values will be adjusted by increasing the setting by 5.

- Select **Conservative** or **More Aggressive** to adjust (relax) the settings marked with checks.
- Select **Custom** to check the criteria settings that you want to adjust (relax).
- Select **No adjustment** to return to the criteria specified for this primer in the search setup dialog box.

Primer Search Statistics Display

This display is an alternate to the search results display and shows which criteria eliminated primers or pairs during this search. Eliminated primers are shown in red on the bar charts. Click the Search results/Search statistics button to view the list of primers found (if the search was successful).

Use the Auto Adjust Length or Auto Adjust Criteria buttons to make adjustments and repeat the search, if you want to find additional primers, or Use the Redefine button to return to the search set-up dialog box to change the search region or product length requirements.

Generation of Search Statistics

The program initially identifies all primers in your search region. It then begins to eliminate primers that fail to pass the tests shown. The test steps are sequential, and only primers that pass a given test are considered in the following test(s). The number of primers evaluated and eliminated in each step is reported numerically and displayed as a bar chart on the right (eliminated primers shown in red).

For pairs searches, if acceptable single primers are found, the program then identifies all primer pairs using the list of normal strand primers (Primer A) and complementary strand primers (Primer B) found in your search region. It then begins to eliminate primer pairs that fail to pass the tests shown for primer pairs. As with single primers, the test steps are sequential, and only primer pairs that pass a given test are considered in the following test(s). The number of pairs evaluated and eliminated in each step is reported numerically and as a bar chart.

Finding more primers

Sorting Primer Search Results

When a search has been successfully completed, primers found will be listed in a table format showing basic information about each primer or primer pair. You can sort this list by rank, position, or product length (pairs searches).

To sort the list of primers found, click on the type of sort in the combo box found on the toolbar for this window.

- **Sort by Rank**

  Place the best primers, or primer pairs, at the top of the list. Ranks are assigned to each primer or primer pair. Where equal scores are achieved, the same rank number is used, and the equal sign (=) is affixed to the rank number for all similar scores.

- **Sort by Position**

  Lists the primers found in ascending base pair position order. If the results show primer pairs, you can sort by the base pair position of the first primer or by the base pair position of the second primer (Primer B).

- **Sort by Product Length**

  If the results show primer pairs and the two primers can amplify a product, you can sort by the length of the amplified product.

Primer Viewer Window

Use the primer viewer window any time you want to view, modify, or analyze a primer.

To open the primer viewer window, select a primer from the Primer List. The primer viewer window will also open when you load a pre-existing primer into the program, when you click the Report button in a search results window, or when you set up a primer analysis operation.

- **Molecule associated primers**

  For primers that are associated with molecules, you can click tabs at the bottom of the window:

  - Info tab - View or edit detailed primer information.
  - Edit tab - edit the primer sequence or position.
  - Analyze tab - analyze the primer for dimers, runs of bases, hairpins, or false priming.

- **Other primers**

  For primers that are not associated with molecules, you can click tabs at the bottom of the window:

  - Details tab - View or edit primer information
  - Analyze tab - analyze the primer for dimers, runs of bases or hairpins.

From within the primer viewer window, you can access functions to view or edit the primer evaluation criteria, link, or unlink primers with molecules, change the primer name or add descriptive information, or export the primer sequence.

Primer Info tab

This tab provides basic information on a molecule-linked primer. The toolbar allows editing, linking, and exporting.

- **Primer Definition** – reports the name of the primer with any description or notes that you have entered.
- **Sequence** – of the primer.
- **Length** – of the primer.
- **Meets criteria** – reports evaluation of the primer against its primer criteria.
- **Composition** – number of each base.
- **Molar absorbance** – can be used to calculate the amount of the primer for experiments.
- **Linked molecule** – the position in the molecule that the primer aligns with.

Primer Details tab

The Details tab provides basic information about a primer that is not currently associated with a molecule. Detailed information includes the primer sequence, length, composition, and molar absorbance.

The analysis criteria type is reported and the Evaluation Bar at the bottom of the window is used to indicate if the primer meets the evaluation criteria. The toolbar allows editing, linking, and exporting.

In addition, the header area will show the primer name, description and the first part of the note's information. You can modify this information (Primer Definition), if needed. You can also link the primer to a molecule or check the criteria settings.

- **Primer Definition** – reports the name of the primer with any description or notes that you have entered.
- **Sequence** – of the primer.
- **Length** – of the primer.
- **Composition** – number of each base.
- **Molar absorbance** – can be used to calculate the amount of the primer for experiments.
- **Analysis criteria** – the type of criteria used for evaluation.
- **Primer evaluation bar** – reports evaluation: %GC, Tm, Dimers, Stability, GC clamp, Runs, Hairpins. Any property that fails to meet the evaluation criteria will be shown in the alert color. You can click on an item to quickly jump to the Analysis tab for that item.

You can also change the bases in the sequence of a primer, if necessary. For primers that are not linked to a molecule, these changes are straight-forward -- you simply add, remove, or change bases as needed in the primer sequence.

Change the primer sequence:

1. With the primer viewer window open to the Details tab, click the Edit Primer Definition toolbar button
2. Click in the Sequence field and then make changes needed.
3. Use the keys for A, C, G or T to add bases. Use the delete key to remove bases, as needed.
4. Click OK when your changes are complete.
5. Check the Evaluation Bar to see how these base changes have affected the evaluation of the primer.

Primer Info and Details Toolbar

- **Edit** – Change the name of the primer. Optionally, enter a description and notes. If the primer is not associated with a molecule, you can also edit the primer's sequence.
- **Primer Criteria** – view or edit the criteria used to evaluate the primer
- **Link** – find sites that the primer can bind to in another molecule. Also used to unlink from the current molecule.
- **Create product** – not enabled for single primers.
- **Add to collection**
- **Export** – primer sequence

Primer Evaluation Bar

The primer Evaluation Bar appears at the bottom of the primer viewer window on the Edit tab, for molecule-associated primers, or on the Details tab, for primers that are not associated with a molecule.

A boxed display area is used to show the primer evaluation. The primer is evaluated for GC content, melting temperature (Tm), presence of dimers, stability, GC clamp, presence of runs or hairpins, and possible false priming. Degenerate primers (designed using protein molecules) are evaluated for average GC content and degeneracy (3' end and total).

The icon at the left edge of the evaluation bar indicates whether the current primer satisfies all the criteria set.

A check mark in a green circle indicates that the primer meets all the criteria set. Click the Analyze tab to find out more about this primer or print a primer design report. Click the Info tab for molecule-associated information (if applicable). Save the primer to disk for later use, if desired.

A yellow caution sign indicates that the primer does not meet one or more of the criteria set. Values that do not meet criteria will appear in color. You can click on the value in the Evaluation Bar to jump to the appropriate view on the Analyze tab for more details. (Click the Edit or Details tab to return.)

If you are designing a primer using the Create from Molecule option and the primer does not meet criteria, adjust the primer position or length until you find a primer that does meet criteria.

Primer Edit tab

This tab provides the ability to change the position of a primer within its linked molecule. More information is available under Edit Primers.

- **Molecule position**

  The molecule sequence is shown across the center of the screen, with the 5' and 3' ends marked. For DNA molecules, a simple map showing genes or regions appears above the sequence. A horizontal line above the map may show tick marks at some locations. These are neighboring primers that will meet the criteria set. For protein molecules, the amino acid sequence is shown above the degenerate DNA sequence derived from this protein sequence.

  The primer sequence is shown in a boxed area below the molecule sequence. If a 5' extension is present, these bases are shown as non-homologous to the molecule sequence (no vertical match lines).

  Mismatched bases are shown with an asterisk character in the alert color.

- **Primer evaluation bar**

  Reports the primer evaluation: %GC, Tm, Dimers, Stability, GC clamp, Runs, Hairpins. Any property that fails to meet the evaluation criteria will be shown in the alert color. You can click on an item to quickly jump to the Analysis tab for that item.

Primer Edit Toolbar

Allows operations on the primer:

- **Primer Criteria** – view or edit the criteria used to evaluate the primer
- **Redesign primer** – enables changing the position of the primer within the linked molecule.
- **Scan left** – moves the primer position leftwards until the next position satisfying the primer evaluation criteria is found. Positions matching the criteria are shown as tick marks above the molecule line.
- **Move left** – moves the primer position one base to the left. The primer evaluation bar will update to show whether the primer meets criteria.
- **Move right** – moves the primer position one base to the right. The primer evaluation bar will update to show whether the primer meets criteria.
- **Scan right** - moves the primer position rightwards until the next position satisfying the primer evaluation criteria is found.
- **Increase primer length**
- **Decrease primer length**

- **Edit 5’ extension** – allows for entering or modifying extra bases. Such bases can be used to introduce restriction enzyme recognition sequences.
- **Tools**

  Provides additional options:

  - **Link to molecule / find sites** – used to find where the primer will bind to another molecule.
  - **Enter primer site to map** – allows entering the primer as a site on the molecule map.
  - **Edit primer sequence** – allows introducing mismatches that could be used to introduce mutations into the linked molecule.
  - **Open Mutagenesis Profile** – used to design mutational changes that can change amino acid translation or create, or destroy, restriction enzyme sites.

Primer Analyze tab

The analyze tab provides a variety of analysis methods. The toolbar allows selection of the different analyses:

- Primer Report

  Reports the criteria used for evaluation. For each criterion, reports the value, the acceptable range and whether the primer meets the criteria.

- Dimers

  Shows any possible 3′ end dimers or any dimers within the primer

- Cautions

  Reports primer stability (selective primers should have a 5' end pentamer stability exceeding the 3' end pentamer stability), GC clamp, runs, repeats and hairpins that could be problems.

- False priming

  Shows places where the 3’ end of the primer could bind to the linked molecule and cause unwanted priming to occur. The locations and extents of potential false priming sites are shown together with the Tm (melting temperature), and the percent binding at 55ºC, which give an indication of how likely this false priming is to be a problem.

  At the bottom of the display is an iconic map of the molecule showing the possible binding sites. Binding sites with Tm of 20ºC are marked in the alert color

- Any homologies

  Shows any locations and extents of places in the linked molecule where the primer could bind. Also shown is the Tm (melting temperature), and the percent binding at 55ºC, which give an indication of how likely this binding is to be a problem. Homologies that are close to the 3’ end of the primer could be significant since trace exonuclease activity could enable false priming to occur

  At the bottom of the display is an iconic map of the molecule showing the possible binding sites. Binding sites with Tm of 20ºC are marked in the alert color

- Check another molecule

  Enables changing the linked molecule to show false priming and any homology binding to a different molecule.

Save Primers to Disk

You can use the File, Save or Save As... commands to save the active primer to a disk file for later use by the program. This option is available when you have the Primer Viewer Window open on the screen. For primers listed in search results lists, you can enter the primer to the primer list and select the option to save the primer to disk during the Enter to Primer List action.

Primer files are saved in a program-specific format and use the extension *.pd4. Primer files contain detailed information about one primer: primer sequence, including the 5' extension (if any), source molecule information (where appropriate), the primer criteria, and the primer name, description, and notes information. (To create a file containing only the primer sequence in standard ASCII text format, use the Export Primer Sequence option, found on the Operations menu.)

You can also save essential primer information to a primer Collection File. Collection files store the basic information for multiple primers (often a group of related primers), but do not contain any source molecule information or detailed criteria settings.

Editing primer name, description, or notes

Edit Primers

To edit a primer that is associated with a molecule, click the Edit tab in the primer viewer window. The screen display provides information about the primer and its evaluation, and about the molecule in the region around the primer.

The molecule sequence is shown across the center of the screen, with the 5' and 3' ends marked. For DNA molecules, a simple map showing genes or regions appears above the sequence. A horizontal line above the map may show tick marks at some locations. These are neighboring primers that will meet the criteria set. For protein molecules, the amino acid sequence is shown above the degenerate DNA sequence derived from this protein sequence.

The primer sequence is shown in a boxed area below the molecule sequence. If a 5' extension is present, these bases are shown as non-homologous to the molecule sequence (no vertical match lines). The primer evaluation bar appears below the primer.

You can edit a primer by changing its position or length, or you can edit the primer sequence to introduce a mutation. You can also use toolbar buttons in this display window to change the primer evaluation criteria, link or unlink the primer, or add the primer binding site to the molecule map.

To edit a primer that was designed previously, click the Redesign Primer button to activate the toolbar buttons that will allow you to move the primer or change its length.

Adjusting Primer Position

You can move the primer one base pair position at a time, or you can jump to the next position that meets the criteria specified. The primer will be re-evaluated in its new position.

Move one base pair position -- Use the single arrowhead buttons or the keyboard left or right arrow keys to move the primer 1 bp position in the direction of the arrow.

Move to next acceptable primer -- Use the double arrowhead buttons or the Ctrl + arrow keys on the keyboard to move the primer in the direction of the arrow to the next position that meets the criteria set.

Adjusting Primer Length

You can increase or decrease the length of the primer, as needed. Click the Increase Primer Length button to add one base to the 3' end of the primer. Click the Decrease Primer Length button to remove one base from the 3' end of the primer.

Edit Primer Sequence

You can change the bases in the sequence of a primer, if needed. The way in which you change the primer sequence will depend on whether the primer is associated (linked) with a molecule.

For primers that are not linked to a molecule, these changes are straight-forward -- you simply add, remove, or change bases as needed in the primer sequence. Click the Details tab in the primer viewer window to access the function to edit the unlinked primer sequence.

If the primer is linked to a molecule, the change of one base for another will produce a mismatch with the molecule sequence at that point. This can be used to introduce a mutation site. Click the Edit tab in the primer viewer window to mutate the primer sequence.

You can also effectively change the sequence of bases in a primer by changing the position of a primer within a molecule or by changing the length of a primer that is linked to a molecule. Click the Edit tab to move or modify the primer within the molecule sequence.

Changing Primer 5' Extension

You can modify the primer 5' extension, if necessary. Click the Edit 5' Extension button and then enter or edit the 5' extension for this primer.

Primer Definition

You can view or edit the name of a primer, or the description and notes information entered for this primer.

To edit the primer definition:

1. Open the primer viewer window and click the Info or Details tab.
2. Click the Edit Primer Definition button
3. Enter or edit the information, as needed.

The primer name will be used to identify the primer on program screens and reports. A brief description of the primer appears on printed reports. The longer notes information can be saved to the primer file or written to a DataBook entry.

If the primer is not associated with a molecule, you can edit the sequence of the primer.

Entering Primer Information

You can view, edit, or enter information about a primer. If some of this information was generated during the primer design process, you can view this information, but it cannot be modified with direct edit.

- **Location**

  For a collection primer, the collection file name is reported. For a new entry, the active collection file open at the left is shown -- this is the location where the new primer will be entered. For a molecule-associated primer, the source molecule and base pair position are reported. The primer location cannot be edited.

- **Primer Name**

  Enter a name for the primer. This name will appear on the primer list and in primer reports. If you plan to enter the primer site to a molecule map, a short primer name is helpful.

- **Primer Type**

  Select the primer type. The primer type sets the criteria used to evaluate the primer. (Primers designed using the program already have a primer type and this cannot be modified here.)

- **Primer Sequence**

  Enter the sequence of bases for your new collection primer or edit the sequence of bases for an existing primer, as needed. (Primers designed using the program already have sequence data and this cannot be modified here.)

- **Primer Description and Notes**

  The description and notes information are saved as a part of the primer file or collection file format. Enter a brief description of the primer. This description will appear on some program screens and printed reports. Notes can contain more detailed information about the primer (source, how used, where stored).

Direct and Special Primer Wizards

Clone Manager Professional can use special primer design wizards for sequencing primers, PCR cloning, RT-PCR or LCR. Or use the direct entry wizard to enter sequence and other information for primers already in use or purchased commercially.

- Direct entry of primer - Enter the sequence of a primer.
- PCR cloning wizard - Use PCR to amplify inserts for cloning.
- Sequencing primer wizard - Design a set of primers for sequencing.
- RT-PCR design wizard - Reverse transcriptase PCR cloning.
- LCR oligo set wizard - Ligase chain reaction method.

Direct Entry of Primer

You can type in the sequence and other information needed to create a new primer. This might be a primer that was purchased commercially, received from a colleague, or previously used but not saved.

- **To direct enter a primer**

1. Click Primer, Direct Entry to start the Direct Enter Primer Wizard.
2. Follow the wizard instructions to enter a primer name, primer type, and sequence (required information).
3. You can also enter a description and notes for this primer (optional).

When the primer is complete, the primer viewer window will open, and you can see the primer evaluation and have access to the analyze options.

- **Entering Basic Primer Information**

Enter the basic, required information for your new primer. You can later update this information, if necessary.

Enter a name for the primer. This name will appear on the primer list and in primer reports.

Select the primer type. The primer type sets the criteria used to evaluate the primer.

- **Entering the Primer Sequence**

Enter the sequence of bases for your new primer. Sequence entry is in the Insert mode. Use the Delete key to remove erroneous bases, if necessary. A base counter shows you how many bases you have entered so far. You can also paste the sequence of the primer from the clipboard if you already have access to this sequence data in another application. Copy the sequence to the clipboard in your other application and then click the Paste button instead of typing in the sequence.

After you have entered the primer sequence, you can invert (reverse complement) this sequence. This may be useful if you have entered a sequence from a publication but need to use the complement strand for your primer. To reverse complement only a part of the sequence, select (highlight) the bases to invert and then click the RevCompl button.

You can enter bases in the primer sequence to code for a specific amino acid, if needed. Position the cursor at the point where you want to insert the bases and then click the Ins AA button and select the amino acid.

- **Entering Primer Description and Notes**

The description and notes information are saved as a part of the primer file format and can be edited later.

Enter a brief description of the primer. This description will appear on some program screens and printed reports.

Notes can contain more detailed information about the primer (source, how used, where stored).

PCR Cloning Wizard

This wizard will help you to design PCR primers to amplify the region of a molecule you want to clone, use enzymes to prepare the vector or the region for insertion, and then generate recombinant molecules. You can view several solutions and save selected primers, amplified products, and recombinants for later use.

- **Getting Started**

  Select the PCR Cloning Wizard from the Primer or Clone menus. Next, identify the type of PCR cloning you want to do and then follow the wizard instructions.

- **Define What you Want to Amplify for Cloning**

  Identify the molecule that contains the region you want to clone. Use the Change... button to select another molecule if the one shown is not the one you want to work with. Enter the upper strand coordinates of the region you want to clone (you can set the orientation of the insert later). Use the Features... button to look at the features table for this molecule and select a gene, if appropriate. The program will enter the base positions for you.

- **Set the Primer Design Options**

  Set the primer type to PCR Primer to set the appropriate primer evaluation criteria. You can leave the criteria set to default conditions, or you can edit the criteria settings for this primer search. Enter the optimal primer length you would like and indicate the acceptable primer size range. Enter the number of bases the search can extend to the left for the forward primer, and to the right for the reverse primer.

- **Specify Preferences**

  If the search for PCR primers is not successful, the program will attempt to extend the acceptable GC range, adjust (relax) the primer evaluation criteria, or increase the search regions to try and find primers you can use. You can accept the default settings or indicate your preferences for each of these categories.

  For Topoisomerase or Classic TA Cloning, you can indicate your preferences about initial residues at the 5' end of your PCR primers that can affect cloning efficiency. These preferences will affect the quality score for each primer and help to sort the best primers to the top of the list.

- **Define the Vector you Want to Use**

  Click the Change... button to select a molecule that you want to use as the vector. The Wizard uses circular vector molecules. If the vector molecule contains a multi-cloning site (shown on a molecule map as a grouped display to the right or left of the map), the base pair positions for this region will automatically be entered for you. For Topoisomerase cloning, the program will find the Topo sites and enter the cut positions

for you. You can change these positions, if necessary, and you can select the orientation of the insert.

- **Select Enzymes**

  For Blunt End cloning and Classic TA cloning, select the enzyme you want to use to cut the vector molecule. The program will show you enzymes from the specified enzyme list that cut only inside the region specified and do not cut the insert molecule. If you want to consider other enzymes, you can select a different enzyme list.

  For Restriction Enzyme cloning, you can select different enzymes to cut the left and right sites of the vector molecule, and you can designate a cap for each enzyme (recommended). The program will show you enzymes that cut only once in the region specified, produce sticky ends, do not have ambiguous bases, and do not cut the insert molecule. The enzyme recognition sequence and cap will be added to the 5' end of the appropriate primer, increasing the length of the primer sequence you specified.

  Select the orientation of insert -- clockwise (cw), or counterclockwise (ccw).

- Adjust (Relax) Primer Criteria

View PCR Cloning Results

If solutions can be found for the PCR cloning experiment you described, the selected number of best solutions can be viewed sequentially.

Use the Next Solution button to view the next solution in the set. Use the Previous Solution button to return to a prior solution in the set.

For each solution, an iconic map of the recombinant molecule is shown, with the insert marked in a lighter color, inserted in the direction indicated by the arrow. If features were present in the vector or insert molecules, these features will be shown above the map line. The area below the map shows detailed information about the recombinant molecule, the insert (amplified product), the primers used for amplification, the vector molecule, enzymes used and position of cuts, and the insert orientation.

Use the Solution Details button to view a complete text summary of this solution. You can print this information or copy it to the clipboard using the standard toolbar Print or Send View to Clipboard buttons.

Use the Create Recombinant button to have the program automatically do the required steps to simulate the cloning experiment, produce the recombinant molecule, and enter to the Molecule List. You can use this new molecule now or save it to disk for later use.

Use the Enter to Primer List button to add the PCR primers to the Primer List, saving the primers to disk or adding them to your primer collection. During this process you can give each primer a name and enter a description.

You can use other toolbar buttons to view the primer pair report and related primer analysis screens, export the primer sequences, or create the initial amplified product using the insert molecule and the two primers specified.

- **Redefine Cloning Experiment**

  Click the Redefine toolbar button to change (fine-tune) the cloning experiment. You may want to do this if you end up with too few possible solutions or solutions that require enzymes that are not readily available. You can change the enzyme list selected, change your preference settings, or change the size of the regions involved, without re-entering all the other information.

Sequencing Primer Wizard

This wizard will help you to design a set of primers for sequencing a molecule region. The wizard will select three non-overlapping primers per sequencing block, where possible. You can re-order the primers in a block to pick another primer as your top choice, where needed. You can export primer sequences, save to a primer collection, or view a map of sequencing coverage.

To get started:

Select the Sequencing Wizard from the Primer menu. Next, follow the wizard instructions to specify the molecule region and strand you want to sequence, enter primer design preferences, and specify which design criteria are most important to you so that primers will be ranked and selected accordingly.

- **Define What you Want to Sequence**

  Identify the molecule that contains the region you want to sequence. Use the Change... button to select another molecule if the one shown is not the one you want to work with. Enter the upper strand coordinates of the region you want to sequence and use the option buttons to indicate if the region to sequence is on the upper/normal strand or on the lower/complement strand.

  You can use the Features... button to look at the features table for this molecule and select a gene, if appropriate. The program will enter the base positions for you.

- **Set the Primer Design Options**

  For automated DNA or cycle sequencing, set the primer type to PCR Primer to set the appropriate primer evaluation criteria. You can leave the criteria set to default conditions, or you can edit the criteria settings for this primer search. If the annealing temperature (shown below the Criteria button) is not correct for this protocol, click the Criteria button and change this value.

  Enter the optimal primer length you would like and indicate the acceptable primer size range. Enter the number of bases preferred for each sequencing block and enter the values that reflect the trace constraints or accept the default values.

- **Specify Weight Settings**

  You can change the weight settings that are used to rank and select the best primers. You can indicate if distance from the optimal target position is more or less important to you, if the penalty for false priming should be increased or ignored, or if avoiding primer dimers is critical to you.

  The standard (default) settings will usually give very good results.

- **Specify Search Preferences**

  If the search for primers is not successful for any sequencing block, the program will attempt to adjust the block positioning and/or adjust the acceptable GC range and repeat the primer search procedure. If primers still cannot be found for a given sequencing block, the program will adjust (relax) the primer evaluation criteria for primers in this block to try and find primers you can use.

  You can accept the default settings or indicate which adjustments you will allow for each of these categories.

  Adjust (Relax) Primer Criteria

View Sequencing Wizard Results

The primers designed for the sequencing project are shown in a text display, grouped by sequencing block. For each primer listed, the data display shows the primer rank in this block (1, 2 or 3), the binding position for the 5' end of the primer to the template molecule, the primer sequence (5' to 3'), the primer length, and the %GC and Tm°C. If the primer was selected using relaxed primer criteria, the symbol !! follows the Tm value.

You can see one primer per sequencing block or three alternate primers per block. Use the drop-down control in the display toolbar to select First Choice or Three Alternates to control the number of primers to display. If difficult areas are encountered and primers cannot be found or if primers are not positioned to provide sequence for the entire block, the notation ** coverage alert ** is added to the block header line.

Sequencing Primer Results Toolbar

Single-primer actions -- With the highlight bar on one of the primers, you can use the first three toolbar buttons to view the primer report for the selected primer, enter this primer to the Primer List for later use, or export the sequence for this primer.

- **Primer Report** – shows the analysis properties of the selected sequencing primer.
- **Enter to Primer List** – used to save the primer.
- **Export Primer Sequence**
- **Drop down selection** for First Choice or Three Alternates.

  You can use this drop-down control in the toolbar to view one primer-associated sequence per block or three alternates per block. Select (highlight) a sequence bar and return to the text display to identify the primer or its binding position. Click the depressed Map Sequencing Coverage button to return to the standard list display.

- **Rotate Alternate Primers**

  Re-order the primers in the block containing the selected (highlighted) primer. This will move an alternate primer to the First-Choice position at the top of the list for this sequencing block. The rank numbers (1, 2, 3) reflect the wizard's rank assignment and do not change with the rotation. (You can use these numbers to return the primers to the original order if needed.)

- **Make new Primer Collection**

  Create a new primer collection containing all the primers displayed -- either 1 primer (first choice) per block or 3 alternate primers per block, depending on the display style selected. You will be asked to enter a short identifier that can be used as part of the name for your collection and its primers.

- **Export all Primer sequences**

  Export primer sequences to a disk file or copy to the Windows clipboard for later use in another program. The wizard will export the primer data displayed -- either 1 primer (first choice) per block or 3 alternate primers per block, depending on the display selected. You can export primer sequences only, primer names and primer sequences, or a tab-delimited data set which also adds the primer description to the data.

- **Map sequencing coverage**

  An alternate display, showing the approximate extent of reliable sequence resulting from each primer-associated sequencing reaction. A map line for the region to be sequenced is shown at the top of the display. Blocks with possible incomplete coverage are marked with red.

  Each of the expected sequences is shown by a colored bar on a separate line. The position of the bar shows the approximate location of the sequence. A blue bar indicates sequence is being aligned 5' to 3', while a light red bar indicates the sequence is being aligned in the reverse direction (3' to 5').

  Clicking this toolbar button a second time returns the display to the textual view of primers.

- **Redefine**

  Returns to the wizard to change (fine-tune) the sequencing primer search. You may want to change the preferred size of the sequencing blocks or modify the primer selection criteria or annealing temperature.

RT-PCR Design Wizard

You can use this option to set up a search for a pair of primers that can be used for RT-PCR. You can indicate the region where each primer should be located. For genomic DNA, these areas will usually be exons. Each primer is designed from a different exon. If the exons are separated by an intron, the area between the exons will not be included in the amplified product, and this difference in product size can help you distinguish cDNA amplification from genomic amplification.

When the search has been completed, the primer pairs found are shown in a list display. In the search results displays, Primer A is located on the normal strand and Primer B is located on the complement strand.

- **Design Details**

  Identify the molecule that you want to use to design the pair of primers for RT-PCR. Use the Change... button to select another molecule, if necessary. Enter the preferred primer length. Both primers will use the same initial length. (Once search results are displayed, you have the option to find more primers or pairs by using the Auto Adjust Primer Length function, if needed). Enter the minimum and maximum values to define the size range that you will accept for the amplified product. The product does not include the area between the two exons. You can also use the Criteria button to view or edit the PCR primer evaluation criteria.

- **Defining Exons 1 and 2**

  Two sequential pages in the wizard are used to enter the regions you want to define as Exons 1 and 2. For each, enter the start and end base pair positions that define the genomic region where you want to locate the primer. Use the Features... button to look at the features table for this molecule to find the start and end position of genes. You will enter a total of four base pair position numbers (two for each exon) on two pages in this wizard. All four base pair position numbers must be in either ascending or descending order. If you are not using genomic DNA, the last base pair position for Exon 1 should be one base pair different than the first base pair position for Exon 2. (Example Exon 1 81-1276, Exon 2 1277-1342). There will be no deletion when the amplified product is created.

- **Search Results**

  Search results will show you the possible primer pairs that could be used for RT-PCR to amplify the region you identified. On the results screen, Primer A is the normal strand primer; Primer B is the complement strand primer. The Product Length reported does not include the area between the two exons you defined. You can move the highlight bar to the primer pair of interest and click Primer Report to open the primer viewer window for this pair. The Info tab will show you a location map. The Analyze tab will show you the primer pair report and provide access to all the detailed analysis screens.

LCR Oligo Set Wizard

You can use this option to prepare a set of four primers for the Ligase Chain Reaction. The program will select four primers, correctly positioned around the target base position you specify, and adjusted for length to have a Tm as close as possible to the Tm you specify.

Primers A and B are located on the normal strand and abut. For optimal specificity, the 3' end position of Primer A covers the target position you specified. Primers C and D are located on the complementary strand and abut. The 3' end position of Primer C covers the target position on the complement strand. When the correct base is present at the target position, amplification will result.

- **Design Details**

  Identify the molecule that you want to use to design the LCR oligo set. Use the Change... button to select another molecule, if necessary. Enter the target position so that the four primers can be constructed around this position. Use the Features... button to look at the features table for this molecule to find the start or end of genes. Enter the preferred melting temperature (Tm). Primers closest to this value will be constructed. You can also use the Criteria button to view or edit the LCR primer evaluation criteria.

- **Primers Selected**

  The four primers selected are shown in the upper area. You can add a 5' extension to primers A and C, if desired. A small graphic shows the relative positions of primers A, B, C and D. At the conclusion of this operation, you can enter the primers to the primer list and then analyze the primers or save the primers to disk.

Analyzing Primers

Use Clone Manager to evaluate primers using GC content, melting temperature, dimers, stability, runs of bases and repeats, hairpins, and false priming, do in-depth analyses, or use the Analyze Mix Wizard to analyze primer binding and products formed in a mix of primers and molecules.

- Analyzing primers - overview
- Primer analyses toolbar
- Primer viewer window
- View primer and evaluation
- Primer evaluation bar
- Using the Analyze Mix wizard
- Primer Values profile

Analyzing Primers - Overview

Found on the Primer menu.

You can analyze single primers, primer pairs, or a mix of primers and molecules. The results of the one primer or pair analyses are shown on the Analyze tab in the primer viewer window. The Primer Pair analysis includes information about the amplified product (if any) which could result from a PCR-type reaction using the two primers and their associated molecule.

- **To set up a primer analysis**

1. Click Primer, Analyze
2. Select the Analysis Type
3. For the one primer or primer pair analyses, click Select Primer to access the primer list to select the primers you want to analyze

To analyze a mix of primers and molecules, the Analyze Mix Wizard will help you to select the primers and molecules, set the analysis conditions, and display the results.

You can also use the Primer Values Profile to calculate and plot primer-related values (Tm, GC content, and stability) to help you find regions for primers in difficult molecules. This option is found on the Tools menu.

Primer Analyses toolbar

Analysis information for a primer or a pair of primers is shown on the Analyze tab of the primer viewer window. The default view is the primer report. Use the toolbar buttons in this window to see other, more specific analyses (see below). These reports can be viewed on screen, sent to the printer, or copied to the Windows clipboard.

You can get to this display by clicking the Primer Report button on the search results screen, by clicking the main menu Primer, Analyze and setting up an analysis, or by clicking the Analyze tab if you are already in the primer viewer window.

The viewer toolbar provides the following analyzes:

- **Primer Reports**

  View general analysis information in the form of a One Primer Report or a Primer Pair Report. The Primer Pair analysis includes a product analysis (found on the Info tab), where appropriate.

- **Analyze Dimers**

  Display homologies between primers (primer dimers).

- **Analysis Cautions**

  View a summary of stability, GC clamp, runs, repeats and hairpins analysis information for single primers or pairs of primers.

- **Analyze False Priming**

  Search the source molecule for 3'-end homologies (>= 6 bases) between the primer and the template.

- **Analyze Any Homologies**

  Search the source molecule for regions of homology (>= 7 bases) between the primer and the template.

- **Change Molecule**

  Select a different molecule to use for False Priming or Any Homology searches.

Primer Reports

You can view general analysis information in the form of a One Primer Report or a Primer Pair Report. The Primer Pair analysis includes a product analysis (found on the Info tab), where appropriate.

Click the Primer Report button on the toolbar in the Analyze tab of the primer viewer window or on the search results display screen. You can also click Primer, Analyze, and select one or two primers from the Primer List for analysis.

One Primer Report

The analysis shows the criteria settings used to evaluate this primer, the actual values for each item, and an indication of whether the criteria were met. For primers against DNA molecules, the criteria include % GC range, Tm range, 3' dimers, any dimers, stability, GC clamp, runs of bases, dinucleotide repeats, hairpins, and worst-case false priming. For degenerate primers, the criteria include the % GC range, 3' end degeneracy and total degeneracy. In addition, degenerate bases are noted.

The printed One Primer Report includes the information on this screen, as well as the more general information contained on the Info tab (primer sequence, length, composition, molecule information and map, if applicable). Primer Notes are added as the last section of the printed report.

Primer Pair Report

The analysis of a pair of primers (usually two primers prepared from the same source molecule) shows a summary of the evaluated values for each primer and comments, where appropriate. Values that do not meet the evaluation criteria are shown in color. For primers against DNA molecules, the values include length, % GC, Tm, 3' dimers, any dimers, stability, GC clamp, runs of bases, dinucleotide repeats, hairpins, and worst-case false priming. For degenerate primers, the values include length, % GC, 3' end degeneracy and total degeneracy. In addition, the degenerate bases for each primer are noted.

The printed Primer Pair Report includes the information on this screen, as well as the more general information contained on the Info tab (primer sequences, length, molecule information and map, if applicable, and product information if an amplified product would result) as well as primer evaluation criteria settings.

If the primer contains a 5' extension, two Tm values are reported. The first (superscript 1) reports the Tm for the first round of amplification and represents the association of only the region of primer homologous to the template (i.e., excluding the 5' extension). The second (superscript 2) reports the Tm for the second round of amplification where the full length of the primer (including the 5' extension) is homologous to the amplified product.

Analyze Dimers

Click the Dimers button on the Analyze tab screen in the primer viewer window. Use this option to display homologies between primers (primer dimers).

The presence of dimers can reduce the efficiency of priming and can cause artifacts such as primer-dimer bands. Dimers involving the 3' end of a primer are particularly important to avoid.

For single primer analyses, the display shows the largest region of homology involving the 3' end and the largest region of homology in any position on the primer.

For two primer analyses, the display shows a chart of interactions between the primers. The program reports the number of adjacent bases in the largest duplex found in interactions involving the 3' ends of primers or in interactions at any position in the primers. The primer interaction shown in the lower area corresponds to the highlighted chart value.

Analysis Cautions

Click the Cautions button on the Analyze tab screen in the primer viewer window. Use this option to view a summary of stability, GC clamp, runs, repeats and hairpins analysis information for single primers or pairs of primers.

- **Stability**

  The program reports the differential stability between the 5' and 3' domains of the primer. If this value does not meet the criteria setting, it is shown in color. Also reported are the values of the 3' end pentamer and the maximally stable non-3' pentamer.

- **GC Clamp**

  If enabled, the program reports the number of G's or C's at the 3' end of each primer. If the criteria value is set to 0, the 3' end of the primer is not checked for G's or C's.

- **Runs**

  The program reports the longest run of identical bases for each primer. If the number of bases in the run exceeds the criteria value, these bases will be shown in color.

- **Repeats**

  The program reports the highest number of dinucleotide repeats for each primer. If this number exceeds the criteria value, the repeats will be shown in color.

- **Hairpins**

  The program checks for possible secondary structure (hairpins or stem loops) within the primer. If found, the program will show the structure of the hairpin and its energy in kcals at the temperature specified in the criteria settings.

Analyze False Priming

Click the False Priming button on the Analyze tab screen in the primer viewer window. Use this option to search the source molecule for 3'-end homologies (>= 6 bases) between the primer and the template. Mismatches or gaps can be included in these areas of homology.

If any regions of 3'-end homology are found, the display screen will show a list of these regions, sorted by Tm. Homologous bases are shown 5' to 3' and the base pair position and Tm for each region of homology are given, as well as the percent binding at the annealing temperature you set. An ellipsis (small dots) at the 5' end of the homologous base display indicates that the primer sequence was longer than the space allowed.

The map in the lower area shows the actual primer (thicker, dark green arrow) and false priming sites (slender arrows). Arrows point right for normal strand, left for complementary strand and color arrows mark sites where the Tm is >= 20°C.

Click the Any Homologies button to search for perfect homologies not restricted to the 3' end of the primer.

Check for False Priming against another molecule

Click the Check Other Molecule button and follow the instructions to check for false priming against a molecule other than the source molecule. When you close this window, this transient connection will be discarded.

Analyze Any Homologies

Click the Any Homologies button on the Analyze tab screen in the primer viewer window. Use this option to search the source molecule for regions of perfect homology (>= 7 bases) between the primer and the template.

If any regions of homology are found, the display screen will show a list of these regions, sorted by Tm. Homologous bases are shown 5' to 3' and the base pair position and Tm for each region of homology are given, as well as the percent binding at the annealing temperature you set. Homologous bases are aligned with the corresponding primer sequence, when possible. An ellipsis (small dots) at either end of the homologous base display indicates that the primer sequence was longer than the space allowed, preventing alignment of bases.

The map in the lower area shows the actual primer (thicker, dark arrow) and regions of homology (slender arrows). Arrows point right for normal strand, left for complementary strand and color arrows mark sites where the Tm is >=20°C.

Click the False Priming button to search for 3'-end homologies, which can include gaps or mismatches.

Check for any homologies against another molecule

Click the Check Other Molecule button and follow the instructions to check for any homologies against a molecule other than the source molecule. When you close this window, this transient connection will be discarded.

Analyze Mix Wizard

Found on the Primer menu. You can use this option to analyze primer interactions, binding, and products formed from a mixture of primers and molecules. This wizard will help you to identify the primers and molecules to add to the mix and set parameters for the analysis.

When the analysis has been completed, a list display will show you the possible products produced, binding sites, and any cautions noted (primer dimers that might affect the results, molecules, or primers that are not used).

You can also use this wizard to analyze interactions from a group of primers where no molecules are present in the mix and no products are produced.

- **Identify Primers**

  Click the Add Primer button to access the Primer List to select the primers to add to the mix. You can select more than one primer at the same time by using the Shift or Control keys in conjunction with the mouse. If primers are linked to a molecule, you do not necessarily have to reload the original source molecules to use the primers in this analysis.

- **Identify Molecules**

  Click the Add Molecule button to access the Molecule List to select molecules to add to the mix. You can select more than one molecule at the same time, and you can use the Browse button to find molecules not yet loaded. If you do not add any molecules to the mix, the program will simply report primer interactions.

- **Set Mix Parameters**

  Select the primer type. This information is used to preset annealing temperature and dimer criteria. You can change these settings, if needed. You can also set the cutoff for percent priming needed to add a result to the list or you can accept the default setting. Use the checkboxes to indicate which results categories you would like to view/print.

Analyze Mix Results

When the analysis has been completed, the results will be displayed in a scrolling window.

- **Products**

If products can be created, they will be listed in the first section, grouped by molecules. The product length, binding positions, and the two primer names are shown for each product. You can select (highlight) a product and create this product or see a detailed report.

Click the Create Product button on the toolbar to create this molecule. Set the option to enter the molecule to the Molecule List for use now, save to a disk file for later use, or export the product sequence.

Click the Primer Report button to view a report that gives detailed information about the two primers involved in creating the product, as well as additional information about the amplified product.

- **Binding Sites**

Binding sites are reported in the next section, grouped by primer. The molecule name, binding position, number of homologous bases, number of mismatches, and Tm are shown for each binding site. You can select (highlight) a primer and see a detailed report about this primer.

- **Analysis Cautions**

Analysis cautions are shown in the last section. 3' end primer dimers or any primer dimers that exceed the warning levels set will be noted here. The primers involved and the homologous bases are shown for each interaction. Molecules or primers that are not used are listed at the end of this section.

Two bookmark buttons let you move quickly to the **Cautions Section** or return to the top of the **List** of products.

Primer Values Profile

Found on the Tools menu.

You can profile regions of your molecule and plot primer-related values to help you find regions for primers in difficult molecules. The program will calculate Tm, GC content and stability for a virtual primer of specified length. (For protein molecules, values are calculated for average GC content, 3' degeneracy and total degeneracy.)

The graphs cover approximately 200 bases of molecule sequence, centered around the target position specified. Primer values are plotted at the 5' position for each primer. Use the Redefine button to select another target position.

- **Identifying acceptable values**

  Values that meet the criteria specified for PCR primers will appear in the region of each graph marked with a green bar at the right. Tm and GC range limits are represented by horizontal rules, as is the stability cutoff value (DNA) or the degeneracy cutoff values (protein). Click on an area of the graph where all three values are plotted in the green ranges. The vertical ruler line will move to this position, and you can read out the base pair position number at the top of the line.

- **Using the vertical ruler line**

  The vertical reference line initially marks the target position you specified. You can move this ruler line in 10-base increments using the toolbar or keyboard left or right arrow keys. You can also click with the mouse on an area of interest and the ruler line will move to this new location and report its base pair position at the top of the ruler line.

- **Changing the primer length**

  The length specified for the virtual primer is reported in the toolbar area. You can use the toolbar Increase Primer Length or Decrease Primer Length buttons to change the length of this virtual primer. The program will recalculate the values and replot the graphs.

Primer List

Click the Primer List button on the main toolbar or use the main View menu. The Primer List contains folders for the active and loaded primers (described below) and for the primer collections found in your home location (My Collections) or in a shared resource location (Shared Collections).

You can sort the information in the Primer List, and you can resize the columns as needed to view the information most important to you. Click on a column heading to sort by the values in that column. Click again to reverse the sort order. Move the mouse to the join between two column headings, depress the mouse button and drag to resize the column. Use the mouse to drag out the dialog box borders to resize the Primer List.

- **Active/Loaded Primers**

  The folder Active/Loaded primers contain newly created primers, primers loaded into the program from disk, and primers highlighted in search results windows (in parentheses). You can use this list to identify the active or loaded primer to use for the next operation, open the primer viewer window to see additional information about the primer, or remove primers no longer needed on this list. Use the Add To button to add one of these primers to a primer collection.

- **Selecting a Primer**

  You can identify the primer to use for the next operation or you can open the primer viewer window for any primer listed. Click on the name of the primer of interest and then click Select.

- **Removing a Primer**

  You can remove a primer from the Active/Loaded folder. This does not delete a disk file (if the primer had previously been saved to disk). Click on the name of the primer you want to discard and then click Remove. If the primer you selected for discard is highlighted in a search window, the discard action will pertain to the search window, and you will be asked if it is okay to close the search in order to remove the primer specified.

- **Loading another Primer**

  Click Browse to access the File Open box to find and load a primer file, if needed. The file will be opened, the primer entered to the Primer List and then selected for use. Click DataBook to access the built-in database to find and open a specific primer file. You can search DataBook entries, as needed, to find just the file you are looking for.

- **Primer collections**

  Organize collections of primers for quick access and search operations.

Use, Organize Primers

Clone Manager uses the Primer List to hold primers loaded, or created, during the current session. You can also store primers in collection files for streamlined storage and access. The Primer Viewer Window puts all information about a primer in one convenient location. And primer sites can appear on your molecule map and be used to export the primer sequence, create the primers, or start to create an amplified product.

- Primer List
- Primer Collections
- Primer viewer
- Create amplified product
- Find primers that bind
- Use primer sites on map
- Scan for molecules with sites
- Link a primer to a molecule
- Internet BLAST primer searches

Create Product

Found on the Primer menu. You can use this option to create a new DNA molecule that is the amplified product of a PCR-type reaction using two primers and a DNA molecule that you specify. If an amplified product can be produced, the results screen will show the product size, GC content, melting temperature, and annealing temperature. You can create a molecule for use now or later, or you can export the product sequence to an ASCII text file.

If an amplified product cannot be produced using standard PCR reaction conditions, you will have the option to relax these conditions and try again. For this operation, Standard PCR reaction conditions use the default PCR annealing temperature (55°C if default unchanged) and a percent binding cutoff of 50%. Relaxed conditions lower the annealing temperature by 10°C and the percent binding cutoff to 20%.

If more than one product is possible, you will have the option to open the Analyze Mix Wizard to view the more detailed results. In the mix wizard results, you can select (highlight) a product and click the Create Product toolbar button to complete this operation.

- **Identify Components**

  1. Click the **Select Molecule** button to access the Molecule List to select a DNA molecule for the procedure.
  2. Click the upper **Select Primer** button to access the Primer List and select the primer for the upper strand (forward primer).
  3. Click the lower **Select Primer** button to access the Primer List and select the primer for the lower strand (reverse primer).

  Click **OK** to create the amplified product.

- **Create product result**

  If an amplified product can be produced, the results screen will show the product size, GC content, melting temperature, and annealing temperature. You can create a molecule for use now or later, or you can export the product sequence to an ASCII text file.

- **Enter Product Molecule to Molecule List**

  The amplified product will be created as a molecule file containing sequence, features (if any), name and description. This molecule will be entered into the Molecule List and the molecule viewer window opens, displaying this molecule for use. You can view the sequence, features, or restriction map information. The file will not be saved to disk unless you specifically use the File, Save function.

- **Save Product Molecule to Disk File**

  The amplified product will be created as a molecule file containing sequence, features (if any), name and description. This file will be saved to disk, using the filename you specify. The molecule will not be entered to the Molecule List and will not be available for immediate use. You can, of course, load the file using File, Open.

- **Export Product Sequence**

  The sequence of the amplified product will be exported in ASCII text format to a disk file, using the filename you specify.

Finding Primers that Bind

Found on the Discover menu or use the Tools button to access in the map window.

You can use this function to find sites where primers you have could bind effectively to the active molecule. You can scan primer files in folders on your computer or local network. Primer files (*.pd2 or *.pd4 files) can be read by the program, as well as primer Collection files (*.px5). The program will open each primer file, read each primer sequence, and analyze its binding potential.

You can set the binding cutoff, to adjust the stringency of the search, and you can set the binding temperature. (Standard conditions are at least 50% binding at a binding temperature of 55°C.) When the search has been completed, you can view the search results. Primer binding sites found can be added to the molecule file and can be displayed on the molecule map or sequence display.

- **Selecting primer files or folders to search**

Click the Add Folder button to select a folder location to add to the search list.

Navigate to the folder you want to add to the list and click OK. The program will attempt to open and read all primer files in the location specified, including those contained in subfolders.

Click the Add File button to select a primer file to add to the list.

Navigate to the primer file you want to add to the list and click OK. The program will attempt to open and read this file during the search procedure.

Click the Add Collection button to access your primer Collections and select a file to add to the search list.

Select the collection file you want to add to the list and click OK. The program will attempt to open the file and read all the primer data contained in the file. If you have selected a folder of collection files, the program will attempt to open each file within the folder.

When you have selected the files and folders to search, click **OK** to begin the search for primer binding sites.

Primers created against protein molecules or primers with degenerate bases cannot be analyzed. These files will be skipped.

If you routinely use the same primer file locations, you can save the list of files and folders and recall this list later.

- **Find Primer Sites Results**

When a search for primer binding sites has been completed, the results will be presented in a scrolling list box and location map display. You can print or copy this list

to the clipboard. You can add some or all these sites to your molecule file for display on the map or formatted sequence.

Results dialog box:

- Table – shows a list of primer binding sites with the primer name and the base pair position of the 5' end of the binding site. Also reported are the length of the binding site (LenB) and the length of the primer (LenP). If there are mismatches within the binding site, the number of mismatched bases is shown in the column *m*. Click on the column headers to sort by primer name, position, or length of the binding site.
- Iconic Map – shows the location(s) of all the primers on the molecule. The vertical bars above the icon map show the approximate location of the primer binding sites. A red arrow and red vertical bar mark the highlighted primer binding site.
- To tag primer sites you want to add, mouse-click on the checkbox in front of the primer name. To un-tag a site, mouse-click again.
- Tag All – button will select all the found primers so they can be added to the molecule using the Add Sites button.
- Clear – button removes the check mark against all primers.
- AddSites –adds all of the tagged primer binding sites to the molecule file and closes the results window. The primer sites can be viewed by selecting primer sites on the main molecule viewer window.
- Print – button will print the list of found primer sites.
- Copy – button will copy the list of found primer sites to the clipboard.

Using Primer Sites

To view primer binding sites on the molecule map, open the molecule in the molecule viewer window and click the **Map** tab. Then click the Change Sites toolbar button to switch to primer sites or you can use the combo box selector above the list of sites.

The location of primer sites will be displayed on the map. You can select a primer site by selecting it from the sites list or clicking on its name on the map. Use the Sites properties toolbar button to view the primer site properties. The tools toolbar button enables exporting the primer sequence, creating the primer, using the primer to create an amplified product, or changing the primer name. (If you do not have any primer sites on your molecule map, you can add primer sites to a molecule file in several ways.)

View Properties

Click the Properties button to view an information pop-up that shows the primer name, length, base pair positions that mark the binding site on the molecule, sequence, and description.

Jump to Sequence

Click the Go To Sequence button to jump to the formatted sequence at the selected primer site.

Create (primer or product)

Click the Tools button and select option from menu. You can create a primer from the information stored in the molecule file or you can create a new DNA molecule that is an amplified product of the active molecule, using the primer selected and another primer from the molecule map that you can select during the operation

Export Sequence

Click the Tools button and select Export Primer Sequence to export the sequence of the primer to a disk file or copy to the Windows clipboard for use in another program.

More Actions

Click the Tools button and select option from menu. You can select Wipe All Primer Sites to remove all the primer sites from the molecule map in one operation, or you can select Change Primer Name to change the name of the primer that will appear on this molecule map.

Scan for Molecules with Sites

Found on the Discover menu. You can use this function to scan molecule files to look for sites where the specified primer could bind effectively. You can scan molecule files in folders on your computer or local network. The program will open each molecule file, read the sequence, and analyze primer binding potential.

You can set the binding cutoff, to adjust the stringency of the search, and you can set the binding temperature. (Standard conditions are at least 50% binding at a binding temperature of 55°C.) When the search has been completed, you can view the search results. You can print or copy the list of molecules and binding sites found.

- **Selecting molecule files or folders to search**

Click the Add Folder button to select a folder location to add to the search list.

Navigate to the folder you want to add to the list and click OK. The program will attempt to open and read all molecule files in the location specified, including those contained in subfolders.

Click the Add File button to select a molecule file to add to the list.

Navigate to the molecule file you want to add to the list and click OK. The program will attempt to open and read this file during the search procedure.

If you routinely use the same molecule file locations, you can save the list of files and folders and recall this list at a later time.

Click **OK** to begin the scan for molecule sites.

- **Notes:**

Primers created against protein molecules or primers with degenerate bases cannot be used to search for molecule sites.

Link to Molecule

Found on the Info or Details tab in the primer viewer window or on the Tools menu. This option is available when viewing a single defined primer (not part of a search in progress).

You can use this option to search the molecule you specify, attempting to find possible sites where the primer could initiate priming. If priming sites are found, you can add a primer binding site to your molecule map or you can link a copy of this primer to the molecule at the site you select. You can also create an unlinked primer by making a copy of this primer that is not linked to any molecule.

- **Link a Primer to a Molecule**

  To link a primer copy to a molecule:

  1. Open the primer viewer window for the primer you want to link to and click the Info or Details tab.
  2. Click the Link to Molecule button and set the Link option.
  3. Select the molecule you want to link to and click OK to begin the search for possible priming sites.
  4. Select the position for the primer-to-molecule link and press **Link**.

  A copy of the primer will be linked at the position specified. Non-homologous bases at the 5' end will be treated as a 5' extension. As the primer is entered to the Primer List, you can modify the primer name or description, as needed. You can use this new primer now or you can save the primer to disk for future use.

- **Priming Sites Found**

  If the Link to Molecule operation has found possible sites where the primer can initiate priming, each of the possible priming sites are shown in a scrolling list. If you attempt to link a primer to its own source molecule, the search will show you possible priming sites other than the active linked site. The potential binding site can contain mismatched bases (reported in the column headed *m*). Non-homologous bases at the 5' end will be treated as a 5' extension. You can select one of these positions and press Link to create a copy of the active primer linked to the molecule at the position specified.

- **Add Site to Map**

  Click the Add Site button to add the selected (highlighted) primer binding site to the molecule file. This site can appear on the molecule map or the formatted sequence.

Unlink a Primer

To unlink a primer copy from the current molecule:

1. Open the primer viewer window for the primer you want to unlink and click the Info tab.
2. Click the Link to Molecule button
3. Set the option **Unlink** and click OK.

A copy of the primer with no molecule association will be made. As the primer is entered to the Primer List, you can modify the primer name or description, as needed.

Mutate Linked Primer Sequence

You can change one or more bases in the sequence of a primer that is linked to a molecule to create a mismatch or mutation site. When making changes, you can see the primer bases in a scrolling list, along with the molecule bases at the associated position.

To mutate the primer sequence:

1. Open the primer viewer window for the primer (select primer from Primer List) and click the Edit tab.
2. Click Tools, Edit Primer Sequence
3. Select (highlight) the base you want to change. Press A, C, G or T to overwrite the primer base letter with the new base letter.

   Mismatched bases are marked in the scrolling list with *m* and on the edit screen display with a red asterisk (*) character. Mismatches will affect the evaluation of the primer.

   Click the Open Mutagenesis Profile button to view a display that provides information about restriction enzyme sites and sequence translation, if needed.

Internet BLAST Primer Searches

Found on the Discover menu. You can set up a remote search using the NCBI BLAST server. You need an internet connection to use this function. This search is the equivalent of using the BLAST Program Search for Short, Nearly Exact Matches, recommended to find primer binding sites. The NCBI documentation suggests a common use is to check the specificity of primers used in PCR reactions or hybridization.

The program will help by handling query sequence formatting, setting the expect value and default word size parameters to values appropriate for short primer-length sequences, and will send your request to the server over the internet. The built-in web browser will show results and allow you to access molecules of interest.

- **Setting up a BLAST Search**

  To initiate a BLAST search, you will want to identify the query sequence (primer), and then select the BLAST database to search.

  1. Click Discover, **BLAST** primer (NCBI).
  2. Click the Change... button to select a different primer from the primer list, if needed.
  3. Use the combo box to select the BLAST database you want to search. A description of the selected database is provided below the selection.
  4. Click the7 **OK** button to prepare your query sequence and send your request over the internet to the NCBI BLAST server.

- **Notes:**

  - If you are using this operation for the first time, you can click the Options button to view or modify your internet access configuration settings.
  - For BLAST searches, the query sequence should contain no ambiguous bases.

Primer Collections

Primer collections provide a means to hold groups of primers that can be viewed or used for primer operations throughout Clone Manager.

- Constructing a Primer Collection
- Use a Primer Collection
- Exporting and Importing Primer Collections
- Searching a Primer Collection

Overview

Primer collections are stored in special primer collection files (*.px5). These collection files contain essential information about multiple primers, often a group of primers that are related in some way. Storing primer information in collection files streamlines storage and access.

The information stored for each primer includes the primer name, type, sequence, description, and notes. Information about a source molecule or a binding site within that molecule is not stored in a collection file entry. If this information is important to you, use File, Save As... to save the primer and its associated molecule information to a disk file in standard primer file (*.pd4) format.

You can construct a primer collection using existing primer files or you can add new primers directly to a collection file. You can create new collection files, as needed, and create folders to organize your collection files. All your collection files are stored within the folder Collections, located in your home directory.

When using Clone Manager Professional, just click on Primer List to access and use your primer collections. At this location, you can view, edit, or organize the primers in your collection or select a primer for use.

Shared Primer Collections

Primer collections can be shared between users or computers. To access the shared resources, click File, Preferences, Share Data, and enter the location (complete path and folder) where you can access shared enzyme lists and primer collection files. When you open the primer list, shared collection files will appear in a folder below your primer collections. Shared collection files will be opened as read-only resources.

Setting up Shared Resources

Constructing a Primer Collection

You can enter new primers directly to a collection file, you can add active (open) primers to a collection, you can convert batches of existing primer files to collection files in one easy operation, or you can import primer collection data.

- **Getting started**

  Click on the Primer List toolbar button to open the Primer List.

  Click on the Collections folder to open folders or files in this location. When you are just getting started, you will have one empty file ready for your first primer to be added to this collection file. You can click the first toolbar button Folder Options to add new folders or new empty collection files, or to rename or delete any of the files or folders you have created.

Adding to your collection

- **Enter new primer** -- If you have the sequence for a new primer, you can create the primer in place in the collection file. Open the Primer List and click on a collection file to open this collection. Then click the toolbar button Add New Entry. Fill in the information about the new primer.
- **Add one primer** -- When a primer is open in the primer viewer window, you can add this primer to one of your collection files. Just click the Add Entry button on the main toolbar and select Add to Primer Collection and select the collection file. This primer can be one you created previously and stored to disk, or it can be a primer you just created.
- **Convert batches of existing files** -- You can use the Multiple File Conversion utility to prepare collection files from your existing *.pd4 or *.pd2 files. You can put all your primers in one collection file or, if you have many primers, you can prepare a set of files to organize them by type, source, or intended use. Just click File, Other Tools, Multiple File Conversion to use.
- **Import primer collection data** -- you can use an Import function to prepare a primer collection from a tab-delimited text file that contains primer information prepared in another application.

Using your Primer Collections

Once you have primers in your collection files, you can view, edit, organize, or select these primers for use. You can also search your collection files to find a specific primer.

Open the Primer List to use your primer collections. Click on the My Collections folder to open folders or files in this location. Use the Tab key or Shift + Tab to move between the left and right panels or click with the mouse.

Click on the Folder Tools toolbar button to add new folders or new empty collection files, to rename or delete any of the files or folders you have created, or to export or import primer collections.

Click on a collection file in the left panel to display the contents of this file in the right panel. For each primer in this collection, the primer name, description, sequence, length, and primer type are shown.

Click on column headers to sort by the values in that column. Move the mouse to the join between two column headings, depress the mouse button and drag to resize the column. Use the mouse to drag out the dialog box borders to resize the Primer List window.

View or Edit Primer Information

Select a primer by clicking on the primer in the list in the right panel. Use the toolbar buttons to edit the information about this primer or view Notes information, delete this primer entry, copy the entry to paste into another collection file, or export the primer sequence. When changes have been made to your collection files, these changes are automatically saved for you.

Select a Primer for use

Click on a primer and then click the Select button at the lower right to use this primer in an on-going operation or to open the primer viewer window for this primer. If more than one primer can be selected for an operation, use the Shift or Control keys in conjunction with the mouse to select additional primers.

Document collection file contents

You can use toolbar buttons to print the contents of the active collection file or send this list of primers to the clipboard for use in another application. Or you can export the data in the active collection file in tab-delimited format to import into another application -- click the Folder Tools button and select Export.

Export / Import Primer Collections

You can use the Export option to prepare a tab-delimited text file containing the information in a primer collection file. You can also export a subset of primer data in a multiple sequence FASTA file format. This subset includes the primer name, description, and primer sequence.

Export a primer collection:

1. Open the Primer List and select (highlight) the primer collection you want to export.
2. Click the Folder Tools button and select Export.
3. In the Export Primer Collection dialog box, select which data to export.
4. Use the checkbox to indicate if you want the exported file to contain headers at the top of each column of data.
5. Select to export to a disk file or copy to the clipboard.

You can use the Import option to read in primer data contained in a tab-delimited file prepared in another application to create a new primer collection.

Import a primer collection:

1. Open the Primer List and click My Collections.
2. Click the Folder Tools button and select Import.
3. Select the tab-delimited text file containing your primer data.
4. In the Import Primer Collection dialog box, specify the first row that contains primer data to import. (If your file has column headers, the data usually starts at row 2.)
5. Use the small drop-down arrow keys above each column to pick the type of data in that column or indicate that the data should be skipped (not imported).
6. Enter the name for the new primer collection and click OK.

Search in Primer Collections

You can search in your primer collection files to locate a specific primer or a group of primers that have a common text string in the primer entry. You can search using a primer name or any text that could be found in the name, description, or notes fields.

- **To start a search**

1. Open the Primer List
2. Click on the collection file or folder in the left panel that you want to search
3. In the Find edit box (lower left), type in the text to search for
4. Click the Go button

If primers are found that meet the search criteria, these primers will be listed in a new folder called Search Results, found at the bottom of the folder list in the left panel. (Click on the Search Results folder to open this folder if it is not already open.)

You can select a primer for use from this list or you can print the list of primers or copy it to the clipboard. When you close the Primer List, the Search Results folder will be discarded.

Aligning Sequences

Clone Manager Professional can help you align many sequences against a reference sequence or against each other. You can also compare two sequences to find regions of local similarity, or scan sequence files in folders on your computer searching for similarities to a specified sequence.

Clone Manager can also help with laboratory-sized sequence assembly projects or the location of subclones on a reference sequence.

- Getting started with alignments
- Aligning multiple sequences
- Comparing two sequences
- Scanning for similarities
- Using the scan results wizard
- Sequence assembly
- Internet BLAST searches
- Alignment Parameters and Settings
- Saving Alignment Setup Files
- Alignment results viewer

Getting Started with Alignment Operations

You can use the alignment features in Clone Manager Professional to do several tasks effectively. For new users in particular, it can be difficult to figure out just where to start.

This section will try to point you in the right direction and get you started with looking at a new sequence, showing aspects of relatedness, checking for vector contamination, and organizing your alignment setup files.

What do you want to learn more about?

- Looking at a new sequence
- Showing aspects of relatedness
- Checking for vector contamination
- Organizing alignment setup files

Looking at a New Sequence

When you have just acquired a new sequence, you probably want to find out as much as possible about this new sequence. Local and remote searches and sequence comparisons are good places to start. You can find places where your new sequence is like a known sequence, you can find other similar sequences on your own computer or local network system or in public databases, or you can find proteins that may have similar functions.

- **Find similar regions in a known sequence.**

  Click Align, Compare Two Sequences and set up a Local Homology search. Compare the new sequence to the known sequence, entering the molecule names in the boxes provided. Select the FastScan method to quickly find common regions of high homology. Select the Needleman-Wunsch method to find longer regions of lower homology. Find similar sequences - local search.

  Click Align, Scan for Similarity to set up a search on your personal computer or local network. Identify your new sequence as the search molecule and click Next to pick the search locations. Click the Add Folders button to select folders on your computer that contain sequence files you want to scan. Click the Add Files button to select databank library files (like EMBL or GenBank) that you have local access to.

- **Find similar sequences - remote search.**

  Click Align, BLAST Search (NCBI REST) to set up a search request to send to the NCBI BLAST server. Identify your new sequence as the search molecule and click Next to set BLAST options. Do a nucleotide search to see if your new sequence has possible errors or frame shifts. Do a protein (or translated DNA) search to find similarity with sequences from different organisms. Do all 6 frames of translation to find your alignment frame.

Showing Aspects of Relatedness

Sometimes you want to capture the best visualization of where sequences are similar to each other, or specifically how they differ. The program has some very special views. You can show an overall view of areas of similarity in many sequences, a phylogeny-style of diagram showing a pattern of relatedness, or specific areas where differences are hard to identify.

- **Show overall view of areas of similarity in many sequences.**

  Set up a Global-Ref multiple sequence alignment, comparing each of the sequences against the first sequence (the reference sequence). Click the LMap tab to see a graphic showing areas of similarity.

- **Show a phylogeny-style diagram.**

  Set up a Multi-Way multiple sequence alignment, comparing each of the sequences against each other. Click the Picture tab to see the diagram.

- **Show specifically where sequences differ.**

  Set up a Global-Ref or Multi-Way multiple sequence alignment, comparing each of the sequences against a reference sequence or against each other. Click the Sequence tab and set the format to Color Behind Non-matching Bases.

Checking for Vector Contamination

If you are concerned that some vector sequence may have been incorporated into molecules you have cloned, you can set up a quick and easy way to scan new sequences against a vector library.

- **Set up vector library.**

  Make a copy of the sequences of each of the vectors in use in your laboratory. Store these copies in a folder (directory) called VecCheck.

- **Check new sequence.**

  Do a local scan to check your new sequence for similarity against your vectors.

  Click Align, Scan for Similarities. Identify the new sequence as the search molecule. Click the Add Folder button and select the folder containing your vector sequences (VecCheck) as the search location.

- **Follow up on hits.**

  If you get a significant hit, you can make a more complete comparison.

  Click Align, Compare Two Sequences and compare the new sequence against the vector sequence that gave the significant hit.

Organizing Alignment Setup Files

If you perform recurring multiple sequence alignments, you probably want to save your alignment setup files so you do not have to reconstruct the alignments each time they are used.

- **Save alignment setup file.**

  When you have set up a multiple sequence alignment just the way you want it, with all the sequences and regions identified, with the sequences in the order you want, and with alignment parameter changes made (if any), click the **Save Setup** button (lower left). Enter the file name and location for this setup file.

- **Enter to DataBook.**

  Record your alignment setup files in the DataBook database. Create a new DataBook specifically for this purpose. Enter records in this new DataBook with the file name of each alignment setup file and add comments about the alignment. You can sort the records, print the list, or search the comments field to find the alignment setup file you need.

- **Retrieve alignment setup file.**

  To run an alignment again, click Align, Align Multiple Sequences, and then click the **Retrieve** button (lower left). Select the alignment setup file to retrieve. Each of the sequence files will be loaded automatically. You can add a new sequence or update the setup, as needed, and begin the alignment procedure.

Aligning Multiple Sequences

Found on the Align menu. You can align multiple sequences against a reference sequence using a global or assembled procedure or do an exhaustive pairwise global alignment of all sequences with a progressive assembly of the alignments (multi-way alignment). You can also retrieve an alignment setup file previously saved to disk so that you can repeat the alignment without re-entering all the information.

Alignment Wizard

The alignment setup wizard allows you to set the alignment strategy, set parameters, and select sequences to align. You can also save alignment setups so that you can repeat the alignment later without having to re-enter all the selections. Saved alignment setup files can be retrieved within the wizard.

- **Type of alignment:**
  - Global reference alignment

    Aligns each of the sequences you specify against a reference sequence. The objective is to maximize the number of matching bases over the full length of each molecule.

  - Assembled alignment

    Aligns all sequences against a reference sequence, but it will progressively assemble the alignment using local similarities.

  - Multi-way alignment

    Does an exhaustive pairwise global alignment of all sequences and progressive assembly of alignments using Neighbor-Joining phylogeny. When complete, you can re-order the sequences based on the phylogenetic data.

- **Align sequences as:**
  - DNA bases – default is to align DNA sequences as is.
  - Amino acids or translate in frame – the default for aligning protein sequences. DNA sequences will be translated in the frame set by the region selected.
  - Translation, display as DNA – allows for the alignment of gene sequences from related organisms where differences in codon usage would provide a poor alignment at the level of DNA bases.

- **Scoring matrix**
  - Standard Linear – uses a simple linear penalty for mismatches and gaps. Use 'Settings' to set the penalties.
  - BLOSUM 62 – used for amino acid sequence alignments.
  - PAM 250 – used for amino acid sequence alignments.
- **Settings**
  - Param – sets the linear scoring matrix properties for mismatches and gaps. Also allows permitting conservative amino acid changes. Under Defaults you can click the 'Set As' button to remember your settings and use them for future alignments. If you make a change to the settings and want to revert to the default settings, you can click the 'Restore' button.
  - Names – sets how you want the name to be shown in the alignment results. Options are molecule name or file name. Long names can also be truncated
- **Retrieve**

  Allows you to load a previously saved alignment setup file.

  If you frequently do the same alignment, you can save the setup file to disk. At another time, you can open the alignment setup file previously saved to disk, automatically loading the sequence files, and returning to the setup dialog box. You can make any changes needed and begin the alignment when ready. You can also save the alignment results (alignment setup and aligned sequences) after an alignment has been completed.
- **Select sequences**

  The next page of the wizard enables you to select the molecules that you want to align. The list shows the currently selected molecules with the region selected to align.

  The toolbar on the right allows managing molecule selection:
  - Add molecule – select molecule(s) from the Molecule List.
  - Identify region – allows selecting a portion of the molecule to align.
  - Remove molecule.
  - Move up – allows changing the order of the molecule. The molecule at the top is labeled as 'Ref' and all other molecules will be aligned with that molecule.
  - Move down.
- **Selecting multiple sequences**

  To streamline alignment procedures, you can load, select for alignment, and later remove groups of molecule files. Just load more than 5 files at the same time, assign a

group name, and then select this group for alignment instead of working with individual files.

- **Save Setup**

  Allows you to save the current alignment setup for later reuse.

Global-Ref Alignment

This procedure will align each of the sequences you specify against a reference sequence (always in position 1). The objective is to maximize the number of matching bases over the full length of each molecule.

This type of global alignment works best when sequences are of similar size and reasonably related to the reference molecule.

To get started:

1. Click Align, Align Multiple Sequences and then click **Global-Ref**.
2. Select to align sequences as DNA, amino acids, or a translation displayed as DNA
3. Select the scoring matrix or use the default setting.
4. Click **Next** to identify the sequences to align.

- **Alignment parameters**

  Alignment parameters have been set to generally give the best results for DNA or protein alignments. You do not have to make any adjustments to these settings unless you want to achieve specific objectives with your alignment. To change the parameters, click the Param button in the alignment setup dialog box.

- **Name Settings**

  Molecules selected for alignment can display the molecule name or file name, up to a maximum of 16 characters. Click the Names button in the alignment setup dialog box to view or modify these settings.

- **Alignment results**

  View the results of an alignment.

Assembled Alignment

The program will compare each sequence against the reference sequence (always in position 1), looking for significant regions of local homology to lock in (using the FastScan - MaxQuality method). The alignment is then completed using a global alignment procedure to align the rest of each sequence against the reference sequence.

This type of alignment can be useful when one or more areas of strong similarity are found in otherwise poorly related sequences.

To get started:

1. Click Align, Align Multiple Sequences and then click **Assembled.**
2. Select to align sequences as DNA, amino acids, or a translation displayed as DNA.
3. Click **Next** to identify the sequences to align.

- **Alignment parameters**

  Alignment parameters have been set to generally give the best results for DNA or protein alignments. You do not have to make any adjustments to these settings unless you want to achieve specific objectives with your alignment. To change the parameters, click the Param button in the alignment setup dialog box.

- **Name Settings**

  Molecules selected for alignment can display the molecule name or file name, up to a maximum of 16 characters. Click the Names button in the alignment setup dialog box to view or modify these settings.

- **Alignment results**

  Show and explore the alignment results.

Multi-Way Alignment

This procedure will do exhaustive pairwise global alignments of all sequences and progressive assembly of alignments using Neighbor-Joining phylogeny. When the alignment is complete, you will have the option to re-order the sequences based on the phylogenetic data produced.

To get started:

1. Click Align, Align Multiple Sequences and then click **Multi-Way**
2. Select to align sequences as DNA, amino acids, or a translation displayed as DNA
3. Select the scoring matrix or use the default setting.
4. Click **Next** to identify the sequences to align.

- **Alignment parameters**

  Alignment parameters have been set to generally give the best results for DNA or protein alignments. You do not have to make any adjustments to these settings unless you want to achieve specific objectives with your alignment. To change the parameters, click the Param button in the alignment setup dialog box.

- **Name Settings**

  Molecules selected for alignment can display the molecule name or file name, up to a maximum of 16 characters. Click the Names button in the alignment setup dialog box to view or modify these settings.

- **Alignment results**

  Show different views of the alignment.

Sequences to Align

You can add sequences to the list of sequences to align, specifying the regions for alignment within each sequence. You can move sequences up or down within the list and delete items, if needed. If you are doing a Global-Ref or Assembled alignment, remember that the reference molecule should be in position 1.

To get started:

1. Click the Add Molecule toolbar button on the right.
2. Select the sequences or group from the molecule list
3. To remove a sequence, select (highlight) it and then use the Remove Molecule button at the right.
4. Use the up or down buttons to move the highlighted sequence within the list, if needed.

The molecule name, molecule size, and file name of the selected molecule are shown in the area below the list. When all the sequences have been selected, click **Finish** to complete the setup and start the alignment.

- **To align only a part of a molecule sequence:**

1. Select (highlight) the name of the molecule within the list
2. Click the Identify Region button on the right.
3. Click **Partial Sequence** to align less than the full molecule and enter the start and end base pair positions. The start base should always be the first nucleotide to be aligned.
4. Use the Complement checkbox if the region to be aligned is on the complement strand.
5. Click OK to complete this item.

- **To change the name of an aligned molecule:**

1. Select (highlight) the name of the molecule within the list
2. Click the Identify Region button on the right.
3. Enter the **Name to Display** in the space provided.
4. Click OK to complete this item.

- **Translating DNA for Alignment**

If it is necessary for the program to translate DNA so that it can align the amino acid sequence, it will translate in frame. Translation will start from the base pair position indicated on the setup list (or from the first base of the complete sequence) and continue to the end base pair, going through termination codons, if encountered. To change the translation frame, change the start base pair position by one or two bases.

Alignment Parameters and Settings

User adjustment of alignment parameters is available for global and multi-way alignments when using the Standard Linear Scoring Matrix, and for FastScan and Needleman-Wunsch alignments. Alignment parameters have been set to generally give the best results for DNA or protein alignments. You do not have to make any adjustments to these settings unless you want to achieve specific objectives with your alignment.

To change the parameters for the current alignment: click the Param button in the alignment setup dialog box. (If this button is not active, you cannot change the alignment parameters for this alignment matrix type.) When you make a change to a parameter setting, it will be in effect for the current alignment only. If you want to make the change affect all future alignments of this type, click the Defaults, Set As… button in the lower right.

- **Linear Scoring Matrix Parameters**

  The following parameters can be adjusted for global or multi-way alignments using the standard Linear Scoring Matrix:

  Mismatch Penalty -- Penalty assigned to each mismatched position. The default value is 2. An increase in this penalty favors gaps over mismatches.

  Open Gap Penalty -- Penalty incurred when a gap is introduced into an alignment. The default value is 4. An increase in this penalty reduces the number of small gaps permitted.

  Extend Gap Penalty -- Penalty assigned to each missing residue in a gap. The default value is 1. An increase in this penalty reduces the length of gaps.

  Conservative Changes -- For protein alignment, you can indicate whether conservative changes will be permitted. If conservative changes are not permitted, exact matches will be required. The conservative changes recognized by the program are: (Ile, Leu, Val, Met), (Asp, Glu), (Asn, Gln), (Lys, Arg), (Ala, Ser, Thr)

- **FastScan Parameters**

  For local homology searches, parameters are automatically set to give the most sensitive search.

  Conservative Changes -- For protein alignment, you can indicate whether conservative changes will be permitted. If conservative changes are not permitted, exact matches will be required. The conservative changes recognized by the program are: (Ile, Leu, Val, Met), (Asp, Glu), (Asn, Gln), (Lys, Arg), (Ala, Ser, Thr)

- **Needleman-Wunsch Parameters**

  The following parameters can be adjusted for local homology searches using the Needleman-Wunsch method:

Match Score -- Value assigned when there is a match. The default value is 1. An increase in this score favors finding regions of poorer homology.

Mismatch Penalty -- Penalty assigned to each mismatched position. The default value is 1. An increase in this penalty favors finding only regions of higher homology. In general, the mismatch and gap penalties should be identical.

Gap Penalty -- Penalty assigned to each missing residue in a gap. The default value is 1. An increase in this penalty favors mismatches over gaps.

Conservative Changes -- For protein alignment, you can indicate whether conservative changes will be permitted. If conservative changes are not permitted, exact matches will be required. The conservative changes recognized by the program are: (Ile, Leu, Val, Met), (Asp, Glu), (Asn, Gln), (Lys, Arg), (Ala, Ser, Thr)

- **Name Settings**

  Molecules selected for multiple sequence alignments can display the molecule name or file name, up to a maximum of 16 characters. If the name is too long, the program will truncate the name, using the rule selected. New settings apply to subsequent molecule selections.

  Name to Enter -- Select molecule name or file name (file extension will not be included).

  Truncation Rules -- If the name selected is too long, you can direct the program to use the first 16 characters of the name, the last 16 characters (not including file extension) or a combination of the first 7 characters, followed by two dots, then the last 7 characters (omitting characters in the middle, as needed).

Saving Alignment Setup Files

When you have completed the Align Multiple Sequences setup dialog boxes, you can save this alignment setup information to a file on disk. You may want to use this option if the alignment setup is used routinely in your work.

Alignment setup files are saved in a program-specific format and use the extension *.am4. The information saved includes molecule names, file names and locations, position in the alignment, start and end base pair position, and the alignment parameters.

To save a setup file:

1. From the last page of the Align Multiple Sequences dialog box click the **Save Setup** button (lower left).
2. Enter a file name and select the Save in… location.

   You can later retrieve this setup, adding or inserting new sequences as needed, and repeat the alignment procedure without having to enter all the molecule names, regions, or alignment parameters again.

To save alignment results:

You can also save the alignment results (alignment setup and aligned sequences) after an alignment has been completed. With the alignment results visible on the screen, click File, Save As... and enter a filename when prompted. Click File, Open to re-open this results file and continue to review the results data.

Retrieving alignment setup files

Comparing Two Sequences

Found on the Align menu. Used to compare two sequences using four different algorithm methods.

- **Alignment Comparison type**

  Select the alignment method to use. You can change the parameters for the current alignment as needed (click the small button in the upper right corner of the alignment setup dialog box).

  - **Local Homology**

    Compares two sequences by searching for regions of local homology.

  - **Assembled alignment**

    Is best for sequences with some similar and some different regions. The assembled alignment is based on finding regions of high local homology and then aligning the remaining sequence using a global strategy.

  - **Global alignment**

    Procedure will maximize the alignment of bases along the full length of both molecules. This procedure works well for sequences that are very similar overall.

  - **Layout (genomes)**

    Aligns two genome sequences allowing for shifts and inversions of highly conserved regions. Useful to identify genomic rearrangements and insertions of mobile elements.

- **Compare Sequences as**

  You can compare sequences as DNA bases or as amino acids, translating a DNA sequence in frame, if needed. You can also translate DNA sequences to determine the best protein alignment and then display the results as DNA.

- **Scoring Matrix**

  Used for global alignments to set the method used for penalizing mismatches.

- **Sequences to Compare**

  Select the two sequences to compare, Select to compare the full sequence or a partial region.

- **Save results**

  To save the compare two sequences results data, with the sequence comparison results visible on the screen, click File, Save As... and enter a filename when prompted. Click File, Open to re-open this results file and continue to review the results data.

Sequences to Compare

You can identify the two sequences to compare, specifying the regions within each sequence, as needed.

1. For Sequence box 1, click the Change button at the right
2. Select the sequence from the molecule list
3. Repeat this procedure for Sequence box 2.
4. Click the Flip button to exchange the sequences between the boxes, if needed.

- **To compare only a part of a molecule sequence:**

1. Below the molecule name, click **Partial** sequence
2. Enter the start and end base pair positions of the region. The start base should always be the first nucleotide to be compared.
3. Use the Complement checkbox to indicate that the region is on the complement strand.
4. Use the small Features button to access the features table for this molecule. You can select a feature to automatically enter the base pair positions and strand information.

Click **Finish** to start the comparison procedure.

- **Translating DNA for alignment**

Allows aligning two DNA sequences using their amino acid translations. The sequences will be translated in frame starting from the base pair position indicated (or from the first base of the complete sequence) and continue to the end base pair, going through termination codons, if encountered.

If you need to change the translation frame, you can do this by changing the base pair position for the start of the region to compare. Click Partial Sequence and then change the "Start At" number by 1 or 2 bases, as needed, to change the frame.

Compare with Local Homology Searches

You can compare two sequences by searching for regions of local homology. You can select to use the FastScan or Needleman-Wunsch method. Once the local homology search is complete, you can view a list of the best regions found, inspect the aligned sequences, or view a location map.

To get started:

1. Click Align, Compare Two Sequences and then click **Local**.
2. Select the method (see below) or use the default setting.
3. Indicate the maximum number of results to report (default 20).
4. Select to compare sequences as DNA or amino acids and use the checkbox to indicate if both DNA strands should be compared.
5. Click **Next** to identify the sequences to compare.

- **FastScan Method**

  The FastScan procedure is based on the rapid search lookup table methods popularized by FAST and BLAST, modified to allow homologies with multiple gaps and frame shifts. The FastScan method is much faster than Needleman-Wunsch and, in general, tends to find shorter regions of higher homology.

  Select FastScan - MaxScore to bias the search in favor of high-scoring regions. (Longer regions can increase score, even if homology is lower.) Select FastScan - MaxQual to bias the search in favor of high-quality regions. (Favors finding shorter regions of very high homology.)

  Based on the length of the search region, the program will automatically select the appropriate ktup (block size) for the homology search to provide a very sensitive search. You cannot adjust the ktup settings for this local alignment operation. You can indicate whether to allow conservative changes for protein alignments.

- **Needleman-Wunsch Method**

  The Needleman-Wunsch method uses exhaustive base-by-base comparisons to find long regions of homology, even if they contain numerous mismatches. This method has been optimized to manage large sequences.

  The local alignment parameters used with this procedure have been set to generally give the best results for DNA or protein sequences. You can adjust these parameters to achieve specific objectives.

- **Alignment results**

  Show and explore the results of the alignment.

Compare with Assembled Alignment

This procedure compares the two sequences that you specify by finding significant regions of local homology (using the FastScan - MaxQuality method) and then completing the alignment using a global alignment procedure.

This type of alignment can be useful when one or more areas of strong similarity are found in otherwise poorly related sequences.

To get started:

1. Click Align, Compare Two Sequences and then click **Assembled**.
2. Select to compare sequences as DNA or amino acids and use the checkbox to indicate if both DNA strands should be compared.
3. Click **Next** to identify the sequences to compare.

Alignment results

Compare with Global Alignment

You can compare two sequences that you specify using a global alignment procedure. The objective is to maximize the number of matching bases over the full length of each molecule.

This type of global alignment works best when sequences are of similar size and reasonably related to each other.

To get started:

1. Click Align, Compare Two Sequences and then click **Global**.
2. Select to compare sequences as DNA or amino acids and use the checkbox to indicate if both DNA strands should be compared.
3. Select the scoring matrix or use the default setting.
4. Click **Next** to identify the sequences to compare.
5. Click **Finish** to view the Alignment results

Compare with Layout (Genomes)

You can compare two sequences that you specify using a layout alignment procedure that is optimized for comparing genomes. The objective is to maximize the number of matching bases over the full length of each molecule. With genomic sequences, there are often substantial rearrangements of segments of the molecules including inversions, and shifts in the positions, of segments. Such rearrangements can be seen pictorially and used to identify, and investigate, the cause and significance of evolutionary events.

To get started:

1. Click Align, Compare Two Sequences and then click Layout (Genomes).
2. Select to compare sequences as DNA or amino acids and use the checkbox to indicate that both DNA strands should be compared.
3. Click **Next** to identify the sequences to compare.
4. Click **Finish** to view the Alignment results

Alignment Results Viewer

When an alignment has been completed, Clone Manager Professional will display alignment results in a tabbed window. The data displayed will depend on whether a global or local alignment procedure was done. Data in a view window can be copied to the clipboard, printed, or saved to a disk file.

Global-type alignment results display:

- Info tab -- shows a table of information, summarizing the alignment results.
- Picture tab -- phylogeny dendrogram (multi-way alignments).
- Sequence tab -- view the actual aligned sequences in a variety of formats.
- LMap tab -- shows a graphic view identifying areas of significant similarity for each sequence.

Local-type alignment results display:

Local alignments usually end with multiple results, providing information about each local homology found. You can display different aspects of the local alignment results by clicking the tabs at the bottom of the window.

- Info tab -- shows a sortable list of the local homologies found
- Picture tab -- view a set of simple maps locating these local homologies
- Sequence tab -- shows the aligned sequences for the selected homology block
- LMap tab -- more detailed map showing the selected homology block in each molecule.

Other Operations

- Editing Alignment Data
- Saving Alignment Results
- Exporting Alignment Data
- Retrieving Alignment Setup Files

Info Tab - global alignments

Multiple sequence alignments using the Global-Ref, Multi-Way or Assembled options and two-sequence comparisons using the Global or Assembled options display results using a similar tabbed data window.

The **Info Tab** shows a table of information that gives a concise overview of the sequence alignment, including the sequences and regions aligned, the number of matches found, the number of nonmatches found, and the calculated percent matching bases. For Assembled alignments of two sequences, you can also see the local homology blocks used to drive the alignment.

When you move the mouse over a molecule name in the multiple sequence tables, the cursor changes to a rectangle shape. Use this special cursor to click on a molecule name to see a pop-up box that will give you more information about this molecule (molecule name, size, file name, description).

You can use Tools, Redefine to return to the alignment setup dialog box to change the alignment setup, if needed, and then repeat the alignment. Use Tools, Edit Name in multiple sequence alignments to view or edit the name, description, or notes information that can be saved with the results to define this alignment.

Calculations:

The percent matching bases is calculated as the number of matching bases divided by the length of the alignment, with the result multiplied by 100 to convert to a percent figure. For global alignments, the length of the alignment is determined for each individual sequence as it is aligned with the reference sequence, including any gaps needed for this two-way alignment. For multi-way alignments, the length of the alignment is derived from the virtual composite sequence length, plus gaps, if present.

Info Tab - local alignments

Two-sequence comparisons using the Local Homology option and scans for similarity display results in a similar tabbed data window.

The **Info Tab** shows a table that lists the regions of local homology or similarity found. You can sort this list by score, percent matches, length, or the start position of the region of homology in molecule 1 or molecule 2 (local homologies) or the start position in the search sequence (scan operation). Click the Picture tab to see a graphic representation of this listed information.

You can move the highlight bar to any region of homology found. When you click the Sequence or LMap tabs, you will see aligned sequence data or a detailed map for the selected region of homology. While you are viewing the aligned sequences or detailed maps, you can move to the next or previous result using the navigation buttons on the results window toolbar.

For scan results, when you move the mouse over the list of locus names, the cursor changes to a magnifying glass. Use this special cursor to click on a locus name to see more information about this molecule.

Use Tools, Edit Name in the scan results window to view or edit the name, description, or notes information for the alignment. For two sequence comparisons, you can use Tools, Redefine to return to the alignment setup dialog box to change the alignment setup, if needed, and then repeat the alignment.

Calculations:

Alignment score is calculated using the method of Needleman-Wunsch (Needleman, S.B. and C.D. Wunsch. 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. Mol Biol 48:443-453) or calculated for FastScan as follows: +1 for each homologous base, -1 for each mismatch, terminator, or gap.

Picture Tab - global alignments

When a multi-way alignment has been performed, you can view a dendrogram which suggests the pattern of relatedness of all of the sequences aligned. Click the **Picture** tab in the alignment results window to access this view.

The dendrogram is constructed with a distance-based tree-building method using the neighbor-joining algorithm. The amount of dissimilarity (the distance) between two aligned sequences is used to derive the tree. The method is very fast and results in a single tree, like the phylogenetic guide trees produced using the CLUSTAL W approach (Higgins, D.G. and P.M. Sharp. 1988. CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene 73:237-244).

Picture Tab - local alignments

You can view a graphic representation of the list of regions of homology found. Click the **Picture** tab in the alignment results window to access this view. (Click the Info tab to see this data in table form.)

For two-sequence comparisons using the Local Homology option, each display entry shows the region of local homology (represented by a colored bar), between lines representing the two molecules. You can sort this display by score, percent matches, length, or the start position of the region of homology in molecule 1 or molecule 2.

For Scan for Similarity operations, each display entry shows the approximate location of the region of homology (represented by a colored bar) in the search molecule sequence. If the search molecule is currently loaded, a simple molecule map appears across the top of the data window. You can sort the display by score, percent matches, length, or the start position of the region of homology in the search molecule sequence.

You can move the highlight bar to any region of homology found. When you click the Sequence or LMap tabs, you will see aligned sequence data or a detailed map for the selected region of local homology.

All molecule maps are drawn by converting to nucleotide units in order to support mixed DNA and protein displays.

Sequence Tab - alignment results

You can view the actual aligned sequences in a variety of formats, emphasizing the differences or similarities between the sequences. Click the **Sequence** tab in the alignment results window to access this view.

Click Show Similarity or Difference button to toggle between the default similarity and default difference formats to get a rapid impression of which view will give you the best display for this data.

In the difference formats, dots indicate matches. In protein alignments, * indicates a terminator and is always shown as a mismatch.

Some formats apply color if the similarity exceeds a specified value (Color area of high matches or Similarity summary bars formats). Use the Similarity Significance Value combo box to set the level of significance (range 50-100%) you want to apply for these formats.

Click the Settings button to access the full set of sequence format options or to set page style options.

When the format is just the way you want it, you can print the aligned sequences in this format, or send them to the clipboard or to a disk file to later insert in a document. You can also use a feature on the Tools main menu to export alignment data in a special way. Using this feature, you can export the consensus sequence of an alignment, or you can export the aligned sequences in CLUSTAL W format for later use in another program.

You can also make small changes (change one base for another or move unaligned bases left or right) to edit the alignment data, if needed.

In multiple sequence alignments, you can change the name used for one of the molecules aligned, if needed. Right-click on the name displayed and enter the name you would like to use (maximum 16 characters).

Format Display Styles for Aligned Sequences

You can select whether to show all bases in the aligned sequences (similarity formats) or to show only non-matching bases (difference formats). Next, you can select a display style (see descriptions below). Check the sample box in the upper right to see how your selected alignment format will appear. You can also set page style and printer font options as well as feature annotations and memorize preferred settings.

- **Plain Text**

  Bases are shown (or printed) in black. This format will print the fastest since individual character formatting is not required. For the difference format, dots indicate matches (matching bases are not shown).

- **Color Non-matching Bases**

  Mismatches are shown in the alert text color. Conservative amino acid changes are shown in the auxiliary text color. You can change these colors, if needed, using File, Preferences, Colors. For the difference format, dots indicate matches (matching bases are not shown).

- **Color Behind Non-matches**

  The background area behind non-matching bases is colored light red by default. You can change this color, if needed, using the combo box below the sample area on the Alignment Format dialog box. For the difference format, dots indicate matches (matching bases are not shown).

- **Color Areas of High Matches**

  The background area behind bases matching at the same base position will be colored when the extent of similarity is greater than or equal to the Similarity Significance Value. You can set this value using the combo box in the window toolbar. The area is colored light blue by default. You can change this color, if needed, using the combo box below the sample area on the Alignment Format dialog box. This style is available when all bases are shown.

- **Color by Composition**

  For DNA sequences, bases A and T are shown in blue, bases C and G are shown in orange. For protein sequences, polar amino acids (Arg, Asn, Asp, Gln, Glu, His, Lys, Ser, Thr) are shown in blue, nonpolar amino acids (Ala, Cys, Gly, Ile, Leu, Met, Phe, Pro, Trp, Tyr, Val) are shown in orange. This style is available when all bases are shown.

- **Similarity Summary Bars**

  Bases are shown as plain text (black letters). Blocks of color appear below the aligned sequences when the extent of similarity at the base pair position is greater than or equal to the Similarity Significance Value. You can set this value using the combo box in the window toolbar. This style is available when all bases are shown.

- **Consensus**

  Shows base letter for the most common base in each position. For each position, calculate the frequency of the most common base and compare it to the Similarity Significance Value. (You can set this value using the combo box in the window toolbar.) Show an upper-case letter when the frequency is greater than or equal to the Similarity Significance Value, and a lower-case letter for frequency 20% below the significance value shown. The consensus sequence is shown in the auxiliary text color. You can change this color, if needed, using File, Preferences, Colors. This style is available when all bases are shown.

- **Notes for clipboard use:**
    - Characters in color will retain this color when sent to the clipboard or to a disk file in html or rtf formats.
    - Colors used for background areas behind bases can be put on the clipboard or sent to a disk file in html format.
    - Similarity summary bars will be replaced by character representations when sent to the clipboard or to a disk file.

Feature Annotations

Features can be shown above the sequence for the reference molecule. This can be used to quickly identify where sequence differences occur. There are three options for controlling annotations.

- None – no annotations will be shown.
- Genes only.
- Genes and Regions (default).

Overlay Composition Data

Click toolbar button: Tools, Overlay Composition Data to add composition data to the similarity map data in the alignment results window. Click again to remove this overlay.

- **For DNA sequences:**

AT-rich regions are shown with blue bars
GC-rich regions are shown with orange bars

- **For protein sequences:**

Polar-rich regions are shown with blue bars
Nonpolar-rich regions are shown with orange bars

Polar amino acids recognized by the program are: Arg, Asn, Asp, Gln, Glu, His, Lys, Ser, Thr.
Nonpolar amino acids recognized by the program: Ala, Cys, Gly, Ile, Leu, Met, Phe, Pro, Trp, Tyr, Val.

To see the base pair position for an area of interest, point with the mouse and depress the left mouse button. To see composition by color for each individual base, click Sequence and select the Color by Composition display style.

Page Styles for Aligned Sequences

You can set style options for the molecule names and base pair numbers that will appear with your aligned sequences. The number of bases of sequence per line will be affected by these selections. You can also set options for printed output.

- **Molecule name**

  Show at left or do not display.

- **Sequence bases**

  Select lower- or upper-case letters for all bases in the sequences.

- **Base pair numbers**

  Show base pair interval numbers at the left or right edges, across the top of each block of aligned sequences, or do not display.

- **Printed output**

  You can select the font size to be used for sequence data sent to your printer. The font size will also affect the number of bases of sequence per line. The font face will always be Courier New if this is supported in your printer. Use the printer set-up box if you want to change page orientation.

Memorize Settings

Click the Memorize link to remember your preferred settings. They will be used each time you start Clone Manager. Click the Reset link to clear your remembered settings.

Location Map Tab - global alignments

When a global-style sequence alignment has been performed, you can view a summary of the extent and location of similarity of the sequences and additional information about regions of DNA or protein molecules. When the alignment is against a reference molecule and this molecule is currently loaded, a simple molecule map appears across the top of the data window. Click the **LMap** tab in the alignment results window to access this view.

Each of the sequences is shown on a separate line. The entire search region is shown for each molecule, scaled to fit in the data box. You can drag out the box to increase the resolution of the map. For global alignments, each sequence is compared to the reference sequence. For multi-way or two-way alignments, each sequence is compared with a virtual composite sequence. You can set the significance value for this display using the Similarity Significance Value combo box in the toolbar.

Regions where similarity is greater than or equal to the significance value (shown in combo box) are marked with colored bars. For assembled alignments, a blue bar indicates that the molecule is being aligned 5' to 3', while a light red bar indicates the molecule is being aligned in the reverse direction (3' to 5').

You can point with the mouse at a map area and press the left mouse button to see the base pair position at the cursor location. You can also view composition data as an overlay on top of the similarity data.

If you have many sequences, you can compress this view to fit more similarity map data in the display. In the compressed view, the lines are closer together and the molecule names are suppressed. Click Tools, Compress View to enable or disable this option.

To see the molecule name and base pair position for an area of interest, point with the mouse and depress the left mouse button.

Location Map Tab - local alignments

For each region of local homology or similarity, you can view a map showing the approximate position of the region of homology within each of the molecules. Click the **LMap** tab in the alignment results window to access this view.

Molecule names are shown, along with homology block base pair coordinates. The regions of local homology appear as colored bars. Homologous regions on the normal strand appear in blue, while regions on the complement strand appear in light red.

Click the Next Result or Previous Result toolbar buttons to navigate through the list of local homologies found without leaving this view window.

Editing Alignment Data

When viewing aligned sequences, you may feel it is necessary to make small changes to the alignment data. If you believe there is an error in the sequence, you can change one base for another. If the alignment was performed using sequence files from automated sequencing machines and these files contain trace data (ABI or SCF format files), you can view the trace for the questionable position in the aligned sequence. Where there are gaps in the sequence, you can move unaligned bases left or right.

To get started:

With alignment data in the alignment viewer window, click the Sequence tab to show the aligned sequences and follow the steps below. Click the Edit Alignment toolbar button to initiate these actions. You can click on a base in the aligned sequence with the mouse, or you can type in the molecule name and base position and then select the edit action required.

- **To replace a base**

1. Enter (or verify) the base position and molecule to edit
2. Click the Replace Base option
3. Verify that the base you pointed to is shown
4. Enter the base you want to use as a replacement

Use the checkbox if you want the program to make the same change in the molecule sequence or trace file on disk.

- **To display a trace**

1. Enter (or verify) the base position in question
2. Select **Display Trace** from the menu shown

The display will show the base in question and about 20 bases to the left and right. The base you specified is marked with a small arrowhead below the base letter. If you feel that the base was not called correctly, you can use the Edit function to change the base in the aligned sequence.

- **To shift a base left or right**

1. Enter (or verify) the base position to edit
2. Select **Left Shift** or **Right Shift** to move the base in the direction indicated

The base will be moved within the sequence gap (marked with ----) until the base lines up with the same base in the reference sequence or until it reaches the end of the gap. (In a multi-way sequence alignment procedure, the virtual composite sequence is used as the reference.)

- **Update molecule sequence**

  If the alignment has just been done and the molecules are still loaded, you can use the checkbox in the Edit Alignment dialog box to indicate that you want the program to make the same change in the molecule sequence. If the molecules are not presently loaded (the alignment results were opened from a disk file), you can use the Redefine toolbar button to repeat the alignment, reloading the molecules, and then make the base change and update the molecule sequence.

  Sequences are updated for standard molecule files (*.cm5, GenBank, etc.) in program memory. You should save the updated molecule to disk before exiting the program if you want to retain the modified sequence. Sequences in trace files on disk (*.abi, *.scf) are updated and saved to disk at the time the base change is requested.

- **Alignment Name, Description**

  You can enter descriptive information about an alignment or assembly. When you save the results of the alignment to disk, this information will also be saved. If you use the Add to DataBook function, the descriptive information will automatically be added to the DataBook record.

  - **Alignment/Assembly Name**

    Enter a name for the alignment or assembly. This name will appear in the results window title bar and as part of the header information in printed reports.

  - **Description**

    Enter a description of the alignment or assembly. The first part of this information (about 70 characters) will appear in the header in printed alignment results.

  - **Notes**

    Can contain more detailed information about the alignment or assembly.

    You are automatically prompted to enter this information the first time you save the results of an alignment or assembly.

    Click the Tools, Edit Name in the alignment or assembly results window to view or edit alignment name, description, or notes information.

Saving Alignment Results

When you have completed a multiple sequence alignment, a two-sequence comparison, a scan for similarity, or a sequence assembly operation, you can save the results to a disk file. You can later re-open this results file and continue to review the results data.

- **Saving alignment results**

  With the alignment results visible on the screen, click File, Save As... and enter a filename when prompted. The results will be saved in program-specific file formats with the following extensions:

  *.am4 -- used for multiple sequence alignment results data or alignment setup files.

  *.ac4 -- used for comparing two sequences results data.

  *.as4 -- used for scanning for similarity results data.

  *.aa4 -- used for sequence assembly results data.

  During the save operation, you will have a chance to enter a name, description, and notes for the alignment.

- **Open results file**

  Click File, Open and select the file previously saved to disk. You can change the format options or print the aligned sequences, export results data to disk, or copy to the clipboard for use in other applications. You can also use special alignment data export functions to export the consensus sequence or the aligned sequences in standard CLUSTAL W multiple sequence alignment format.

  Export results data

Exporting Alignment Data

Special alignment data export functions can be found on the Tools main menu or as an option in the Export function (found on the File menu).

Sequence data for these export operations will be the aligned sequences shown on the Sequence tab in the Alignment Viewer Window. These may be the result of a multiple sequence alignment operation or one of many results from a local alignment operation.

The consensus sequence of an alignment or a special merged/spliced consensus sequence can be exported to a disk file, copied to the clipboard, or loaded as a molecule for use now. The sequence will be exported in ASCII (DOS text) file format in upper case letters, 60 bases per line, with no base pair numbers. Aligned sequences can also be copied or exported in CLUSTAL W multiple sequence alignment file format for later use in another program.

Export Consensus Sequence

Select this option to export the standard consensus sequence derived from the alignment. Gaps in the aligned sequences are significant and will affect the final consensus sequence.

Following a multiple sequence alignment, you can select the Consensus display style to see the consensus sequence that will be exported. In this view, upper case letters are used to show the most common base in each position when the frequency is greater than or equal to the Similarity Significance Value shown in the combo box. You can set this value to control the significance of the consensus sequence generated. Only the upper-case letters that appear in the consensus line will be used for the exported consensus sequence. Where lower case letters or gaps appear, the consensus sequence will show N for DNA sequences or X for protein sequences.

Export Merged/Spliced Consensus Sequence

Select this option to export a special merged consensus sequence that can be used for splicing two overlapping regions. In generating this consensus sequence, gaps in the alignment zone are ignored.

Export Aligned Sequences

Select this option to export the aligned sequences in standard CLUSTAL W multiple sequence alignment file format.

Standard use of alignment data

You can copy or export the aligned sequences, as shown on the display screen (not in CLUSTAL W format), for use in another program. Click View, Send View to Clipboard or Send View to File to capture the alignment data. Or select the Data in View option in the Export function and then select a Send To destination.

Retrieving Alignment Setup Files

You can open an alignment setup file previously saved to disk. After you select a file to open, the program will automatically load the sequence files and return you to the setup dialog box. You can make any changes needed and begin the alignment when ready.

To get started:

1. Click Align, Align Multiple Sequences and then click **Retrieve** (lower left).
2. Select the setup file from the Retrieve Alignment Setup dialog box.
3. Click **OK** to accept the file and continue the retrieve process. Each of the sequence files will be loaded at this time.
4. Replace or delete sequences from the list, if necessary. Click Back to modify the alignment parameters, if necessary.

Click **Finish** to complete the set up and start the alignment.

Notes:

Retrieved alignment sets appear on the molecule list as groups for easy handling. You can remove the group when the alignment is complete or select the group to do an additional alignment.

Saving alignment setup files

Scanning for Similarities

Found on the Align menu.

You can scan files or folders on your computer or local network, looking for regions of similarity to all or part of a search molecule. To set up a scan operation, you identify the molecule and region you want to match and indicate whether to translate this molecule. Next, you select the files or folders you want to search for and set the scan parameters.

Once the scan operation is complete, you can view different aspects of the scan results in a tabbed display window. You can view a list of the regions of similarity found, inspect the aligned sequences, or view a location map. The Scan Results Wizard0 will help you to use the molecules with sequence similarities identified by the scan. You can also retrieve scan results previously saved to disk and continue your analysis of these results.

The search method used by the scan operation is based on the rapid search lookup table methods popularized by FAST and BLAST, modified to allow homologies with multiple gaps and frame shifts.

Scoring is +1 for each homologous DNA base and +2 for each homologous protein base. Mismatches and gaps are penalized by 1. For perfect DNA homology, the score will equal the length of the homologous region. High scores therefore represent long regions of high homology.

Scan Wizard

- Click Align, Scan for Similarity and then click **New Search**.
- If the molecule shown is not the one you want to use, click the Change... button to select a different sequence from the molecule list.
- Identify the Match Region, if less than the full molecule:
  - Click **Partial sequence** (below molecule name)
  - Enter the start and end base pair positions of the region. The start base should always be the first nucleotide to be used for the search.
  - Use the Complement checkbox to indicate that the region is on the complement strand.
  - Use the small '?' button to access the features table for this molecule. You can select a feature to automatically enter the base pair positions and strand information.

Processing the Molecule

You can indicate if the search molecule should be translated or if no processing is required.

None -- Use the DNA or protein molecule as is. For DNA molecules, search for regions of similarity on both strands of the DNA molecules or library sequences. Protein files will not be scanned when you specify a DNA search. For protein molecules, sequences will be compared to each protein molecule or to all 6 frames of translation for each DNA molecule or library sequence.

1 Frame Translated -- The program will translate the DNA molecule from the start base pair position identified (or from the first base of the complete molecule) to the end base pair position. Translation will proceed through termination codons, if encountered, and the translated protein will be used as the search sequence. This search sequence will be compared to each protein molecule or to all 6 frames of translation for each DNA molecule or library sequence.

6 Frames Translated -- The program will translate the region identified (or the complete molecule) in all 6 frames and then look for homologies to each protein molecule or to all 6 frames of translation for each DNA molecule or library sequence. This option will take approximately 6 times longer than 1 frame translation.

DNA or Protein Scans

There are specific advantages to doing either DNA or protein searches when using the Scan for Similarity operation.

- **DNA Scans**

  Searching for DNA homologies is most useful when you are looking for non-protein homologies (such as ribosomal RNA genes or transcriptional control zones) or searching for homologies without wanting to bias the search by requiring a continuous in-frame translation product. This can be especially important in cases where your sequence may contain frame-shift mutations or errors. A DNA search is able to cope with multiple single-base frameshifts, whereas a protein search would only find the longest common in-frame portion.

- **Protein Scans**

  Searching for protein homologies is most useful when you are interested in finding similar gene products either for studying gene families or for finding proteins that may provide insight into the biological mechanisms of action of a newly acquired sequence.

  An advantage of selecting the option to translate the search sequence in all 6 frames is that you can find which reading frames contain identifiable genes. Such a search, obviously, takes much longer than translating a known gene in one frame.

Identifying Search Locations

You can scan molecule files in folders on your computer or local network, and you can scan databank files such as GenBank flat file or PIR, if your organization downloads and stores these files for local access. For each location specified, the list box reports the size (in kb) of the data files to be scanned and the number of files at that location. An error log file will report on any problems encountered during the scan.

If you routinely search the same locations, you can save the list of files and folders and recall this list later. You can also set the scan parameters (speed and cutoff value) for this search.

- **Selecting files or folders to scan**

Use the toolbar buttons to add or manage files and folders to scan.

Add Folder: Click to select a folder location to add to the search list.

> Navigate to the folder you want to add to the list and click OK. The program will attempt to open and read all molecule files in the location specified, including those contained in subfolders.

Add File: Click to select a library file to add to the list.

> Navigate to the file you want to add to the list and click OK. The program will attempt to open and read this file during the scan procedure. GenBank databank files and standard files in any of the supported file formats that contain multiple molecules can all be added to this list. Supported file formats include GenBank (flat file), EMBL, FastA and PIR formats.

Remove: Click to remove the selected item from the list.

- ### Using Search Lists

If you routinely do operations that search through files and folders (such, as Find Primer Sites or Scan for Similarity), you can save the list of locations to search and recall this list at a later time.

Save List

When you have completed the identification of search-in locations, click the Save List button (toolbar at right), and enter a list name when prompted.

Use List

To use a list previously saved, click the Open List button (toolbar at right) and select the name of the list you want to use. The listed locations will be added to the list box. You can add or remove locations, as needed, before you begin the search operation.

Setting Scan Parameters

You can set the speed and cutoff value for each scan operation.

Speed -- Select Standard, Rapid or Sensitive.

The program will automatically select the appropriate ktup (block size) for the homology search based on the length of the search region and the setting you select.

In most cases, the Standard setting will be the most appropriate. Select Sensitive to increase the sensitivity of the scan by decreasing the ktup by one unit, where possible. For Rapid scans, a streamlined procedure, and an alternate set of ktup values are used.

Cutoff Values – Provides a filter for the required homology.

The cutoff value is the score required to add the homology block to the results list. The Standard cutoff value is 30 for DNA search molecules and 20 for protein search molecules.

Select Reduced or Low to lower the cutoff value to 0accept results of lower significance. Reduced is approximately 2/3 the standard value, while Low is 1/2 the standard value.

Scan Results Viewer

When a scan for similarity is complete, the results display will include information about each homology or similarity found. You can display different aspects of the local homology results by clicking the tabs at the bottom of the window.

- **Info tab** – shows a list of similar molecules in table format:
    - Score: the higher the score the more extensive the similarity is.
    - Locus: name of the molecule with homology
    - Description: the molecule's description
- **Picture tab** – shows a picture list of where all the similarities are within the search sequence
- **Sequence tab** – shows the sequence of the selected aligned similarity. Selected similarity can be changed using the toolbar or the Info or Picture tab display. The header reports the percent match, score, and length of the homology.
- **LMap tab** – shows a graphic image of the location of the selected aligned similarity. This also reports details of the location of the homology within the pair of sequences and shows if the alignment is inverted and uses the complementary strand.
- Scan Results Toolbar
- Scan Results Wizard

  Use to load molecules of interest, extract molecule files for use later, or set up a multiple sequence alignment using the scan results as a starting point.
- Saving scan results
- Export Alignment Data

Scan Results Toolbar

The toolbar allows for operations on the results display:

- **Sort:** the first combo box allows for selection of how the results should be sorted. Options are unsorted, score, percent (homology), length (of homology) and Start 1 (position on the reference sequence).
- **Left and Right arrow:** Changes the selected sequence homology. Most useful when browsing the sequence of LMap tab displays.
- **Settings:** enabled on the sequence tab. Provides extensive settings for the sequence display format.
    - Display aligned sequences: options allow to focus on similarity or differences.

- Display style: Plain text, Color non-matching bases, or Color behind non-matches.
- Page style: allows for turning off the display of the molecule names, showing sequence bases in lower or upper case and location of the base pair numbers.
- Printed output: allows selection of the font size

- **Show Similarity or Difference** - provides a quick toggle action for Settings, Display style
- **Percent homology combo box:** Not used here - primary use is for multiple molecule alignments.
- **Edit alignment:**
  - Aligned molecule: Select either the reference or the aligned molecule.
  - Base position to edit: Enter the base position for the base to edit. You can also click on the base to edit before clicking this toolbar button to pre-select the correct molecule and base.
  - Edit action: Allows for replacing the base and optionally updating the original sequence (as opposed to simply updating the alignment). You can also shift the base left or right to the other end of a gap. If trace information is available in the molecule, you can view the trace to review the confidence of updating the sequence.
- **Tools:**
  - Edit name: you can enter a name, description, and notes to describe this sequence similarity search. This is useful if you are saving, or printing, the results.
  - Scan results wizard - allows several options for using the scan results.

Scan Results Wizard

You can select one or more molecules and load the molecule files for use now, extract the molecules from a databank file for use later, or set up a multiple sequence alignment of these molecules, using the scan results as a starting point.

The Scan Results Wizard can be selected from the drop-down menu by clicking the Tools toolbar button.

One molecule

Move the highlight bar to the molecule name in the results list and then select the Scan Results Wizard from the Tools toolbar button.

- Load molecule for use now - enters the molecule to the Molecule List and automatically opens the molecule viewer window so you can view the molecule sequence or features table or begin another task.
- Extract molecule and save to disk – saves the molecule to disk. Useful if you are searching for a multiple molecule database file.

Multiple molecules – pick molecules from a list

From the Search Results Wizard, select this option to pick which molecules you want to include.

Mouse-click to tag each molecule you want to select. You can also click the Tag button to quickly tag the first 10 molecules on the list (list is sorted the same way you sorted the results list on the screen display). If you want to tag more or less than 10 molecules, use the spin control above the Tag button to increment or decrement the setting for 'n'.

Then select how you want to act on the selected molecules.

- Load Molecules to Molecule List
- Extract molecules and save to disk
- Set up a multiple sequence alignment

Extract molecules and save to disk

The Scan Results Wizard will auto-name files and then save to one location -- The program will assign a name for each file, using the molecule name or locus name (or an existing filename, if appropriate). You can specify one location for all the files extracted in this operation.

Save each file individually -- The program will extract each molecule and display the Save As dialog box so that you can name the file and indicate its location.

Combine multiple databank molecules in one output file -- The program will extract each molecule and add it to a new multiple-molecule file. Enter a filename and location in the Save As dialog box. All files must be in the same file format (GenBank, EMBL or FastA).

Setting up a Multiple Sequence Alignment

The Scan Results Wizard will set up an Assembled Multiple Sequence alignment. The source sequence and match region used for the scan operation will be specified as the reference sequence for the assembled alignment. Each of the molecules you selected will be aligned against the reference sequence, using the homology region identified in the scan.

Once you see the alignment setup dialog boxes, you can change this information, as needed. Click **Finish** to complete the setup and begin the alignment procedure.

Retrieving Scan Results

You can open a scan results file previously saved to disk.

1. Click Align, Scan for Similarity and then click **Retrieve Results**.
2. Click the Retrieve Now button on the right.
3. Select the scan results file from the file open dialog box.

Scan results files are saved in a program-specific format and use the extension *.as4

Once the scan results have been retrieved, you can continue to analyze regions of homology found. You can view aligned sequences, location maps, and extract information of interest.

Saving Scan Results

The results of a completed scan can be saved to disk in a program-specific file format with the extension *.as4. The file format saves the search molecule information and search parameters, as well as the information about the homologies found.

You can later retrieve this file to continue analyzing the results of the scan operation.

- **To save scan results:**

Click File, Save As. while the scan results are in the active window on the screen.

Sequence Assembly

Found on the Align menu. You can use this module to assemble sequences for simple or complex projects or to locate subclones of a reference sequence. Most commonly used for assembling sequence trace files into contig(s) but can also be used to assemble sequencing files against a known reference sequence to check for errors or mutations.

In each case, you identify the location of the sequence files or folders of sequencing files that you want to assemble. The program handles opening each file and reading the sequence data. Results are presented in a data window with multiple views -- you can review a table of results information, a picture view of alignment locations, or the aligned sequences with optional accessory data in a split pane.

- **Sequence Assembly Wizard**
- **Sequence Assembly Viewer**

Sequence Assembly Wizard

The wizard consists of 2 pages for setting up the purpose of the assembly and selecting the files to assemble. If the files contain sequence trace information (ABI or SCF) then you will be able to view the trace data in the resulting assembly.

Assembly mode and objective:

- **Simple** – Assemble a contig from a small sequence set with limited overlap coverage. Use to assemble the sequences for a gene or region that you have cloned.
- **Complex** – Larger sequence assembly project with many sequences and multiple overlap coverage. Use when assembling unknown sequences.
    - Param button – enables changing the assembly parameters controlling how sequences are merged into contig(s). The Hints box provides an explanation.
- **Subclones** – Locate subclone sequences onto a known reference sequence. Use to confirm the original sequence or rapidly identify SNPs.
    - Enter the reference molecule and optionally select the part of the sequence you want to use. Use the Change button to select a different molecule if needed. You can use the Features button to quickly look up the position of genes and regions.
- **Retrieve** – Load a previously saved sequence assembly project**.**
- **Names** – Change how the name of the sequences will be displayed. You can show the molecule name or the filename. For long names you can select to show the first 16 characters, the last 16 characters or a split of the first and last characters.

Sequences to assemble:

The toolbar on the right provides methods for adding, or removing, files and folders to use for the assembly project.

- **Add Folder** – uses the 'Browse For Folder' dialog box to select a folder. All the files in the folder will be used for the assembly. It is most convenient if the files for each sequencing project are stored in separate folders on your computer.

- **Add File** – uses the 'Select Sequence Files' dialog box to select one, or more, files to be used for the assembly. Multiple files can be selected using the standard Control-click or Shift-click method.

- **Remove** – removes the selected item in the list of sequences.

- **Open List** – allows quick selection of files and folders that you often use.

- **Save List** – allows saving the current list of files and folders so that they can be used in the future without having to re-enter manually.

Assembly Parameters:

- **Expected Coverage** – used for complex assembly projects. The default is to expect at least 6 sequences for each contig base. You can reduce this expectation if appropriate.
- **Overlap score** – used for simple assembly projects. Score is the number of matching bases in the overlap with a penalty for minor mismatches. Overlaps between 2 sequences that do not meet this requirement will not be combined into an assembly contig.
- **Shuffle sequences** – used for simple and complex assembly. Because sequence assembly proceeds in a linear fashion, it is possible that the order in which sequences are assembled could affect the resulting assembly contigs. Checking this box will shuffle, or re-order, the sequences and can be used to verify that the details of the assembled contigs remain valid. The best use is to run an assembly then re-run using shuffle to verify that the contig remains valid.

Sequence Assembly Viewer

When a sequence assembly operation is complete, the assembly results will be shown in a data window that has multiple alternate displays.

You can view a table of results information, a picture view of alignment locations, or the aligned sequences with optional accessory data in a split pane. You can print the data in a view window, copy this view to the clipboard, or export the view data to a disk file.

If the sequence assembly resulted in more than one contig, you can use the combo box in the toolbar to select the contig data to view.

Click the Enter Contig to Molecule List button to load the consensus sequence for this contig as a molecule for later use. Once on the Molecule List, you can select this molecule, add information, analyze the molecule, or save it to disk.

Click the Select View toolbar button to view info, picture, sequence, sequence and overview, sequence and coverage, or sequence and trace.

Click the Edit Name toolbar button to view or edit the name, description, notes for the sequence assembly results data, or the Edit Consensus button to edit the consensus sequence, if required.

- **Save sequence assembly results**

  When the sequence assembly operation has been completed, you can save the results and later open the results file to continue reviewing this data. With the results data in the active window, click File, Save As... and enter a name for the results file. To re-open this file later, click File, Open, and select this file.

Sequence Assembly Toolbar

- **Redefine** – modify and rerun the assembly project.
- **Summary** – a concise overview of the assembly results.
- **Edit Name** - view or edit the name, description, notes for the sequence assembly results data.
- **Contig** – Used if the sequences can be assembled into more than one contig. The drop-down list allows selection of which contig to view.
- **Enter contig to molecule list** - load the consensus sequence for the selected contig as a molecule for later use. Once on the Molecule List, you can select this molecule, add information, analyze the molecule, or save it to disk.
- **Select view** – provides different ways to view the assembly results:
    - View info
    - View picture
    - View sequence
    - Sequence and Overview
    - Sequence and Coverage
    - Sequence and Trace
- **Color Base or Background** – select how the sequence is shown.
- **Go To ambiguous** – moves the active position to the next ambiguous base.
- **Edit consensus sequence** – allows making corrections to the assembled sequences.

View Info

When a sequence assembly operation has been performed, you can view a table of information, reporting the results of the sequence assembly.

Click the Select View toolbar button and then click View Info.

The consensus or reference sequence is shown in the first position.

Each sequence aligned against this reference is reported, showing the position of the homology on the reference molecule, the position of the homology on the sequence molecule, the number of matching bases, the number of non-matching bases, and the percent error. The letter C after the base pair number indicates Complement strand.

Percent Error is defined as the number of non-matching bases divided by the total number of matching and non-matching bases, multiplied by 100.

View Picture

When a sequence assembly operation has been performed, you can view a graphic picture of alignment locations.

Click the Select View toolbar button and then click View Picture.

For a simple or complex sequence assembly:

The consensus sequence is shown as a map line in the first position. Each of the aligned sequences is shown by a colored bar on a separate line. The position of the bar shows the approximate location of the aligned region in the consensus sequence. A blue bar indicates that the sequence is being aligned 5' to 3', while a light red bar indicates the sequence is being aligned in the reverse direction (3' to 5').

For a subclone location operation:

The reference sequence is shown as a double-stranded map line in the first position. Black bars above or below the double-stranded map line indicate gaps in coverage on the normal or complement strand. Each of the aligned subclone sequences is shown by a colored bar on a separate line. The position of the bar shows the approximate location of the aligned region in the reference sequence. A blue bar indicates that the sequence is being aligned 5' to 3', while a light red bar indicates the sequence is being aligned in the reverse direction (3' to 5').

View Sequence

When a sequence assembly operation has been performed, you can view the aligned sequences or aligned sequences with optional accessory data in a split pane.

Click the Select View toolbar button and then click View Sequence.

The consensus or reference sequence is shown in the first position and the other sequences that are aligned against this sequence appear below. The sequence name is shown in the first column (use mouse to widen column if needed) and the base pair number for the first base shown for that sequence appears in the second column.

Non-matching bases are shown in red (or with a light orange background). Ambiguous bases in the consensus sequences are shown in light blue (or with a pale blue background). Click the Color Base or Background button to toggle between the colored base or colored background style for display of nonmatching or ambiguous bases.

Use the horizontal scroll bar below the aligned sequences or the keyboard arrow keys to move along the contig.

Feature Annotations

Features are shown above the sequence for the reference molecule. This can be used to quickly identify where sequence differences occur.

Accessory Information:

You can select to add a split panel display below the aligned sequences. This split panel is used to show additional information about the sequence assembly. You can view:

- Sequence and Overview -- aligned sequence bars and position locator within assembly
- Sequence and Coverage -- histogram of number of reads, errors within reads for assembly
- Sequence and Trace -- display trace data for one or more sequences (if files contain this data)

Sequence and Overview

You can view the aligned sequences with an overview of the sequence assembly shown below the sequences in a split pane.

Click the Select View toolbar button and select Seq and Overview.

The complete consensus or reference sequence is represented by a gray map line drawn full width at the top of the split pane. The sequences that have been aligned against the reference appear in their approximate locations relative to the reference sequence. A blue bar indicates a sequence aligned 5' to 3', while a light red bar indicates a sequence aligned in the reverse direction (3' to 5').

A pair of brackets at the top and bottom of the split pane indicates the position and extent of the sequence you are viewing (in the top panel) within the assembly. Click with the mouse at any point within the overview image to move the brackets to a new location and jump to this position in the aligned sequences.

You can change the size of the split pane, if needed to display a larger set of overview data. Use your mouse to pull on the divider between the upper and lower panes to increase or decrease the height of the lower pane.

Sequence and Coverage

You can view the aligned sequences with a coverage histogram shown below the sequences in a split pane.

Click the Select View toolbar button and select Seq and Coverage.

A histogram of coverage for the complete assembly is shown full width in the split pane. At each point in the assembly, the number of aligned sequences contributing to the data is plotted on the histogram using a gray color. The height of the gray bar at a given position is proportional to the number of reads for the consensus sequence at that point. A scale is shown at the left and right edges of the histogram.

If non-matching bases are present in an aligned sequence, a red color is plotted at the position of the mismatch. The height and extent of red areas within the histogram can be used to identify possible problem areas. For subclone locator results, a solid red baseline indicates no coverage for this part of the reference sequence.

A pair of brackets at the top and bottom of the split pane indicates the position and extent of the sequence you are viewing (in the top panel) within the assembly. Click with the mouse at any point within the coverage histogram to move the brackets to a new location and jump to this position in the aligned sequences.

You can change the size of the split pane, if needed to view the coverage histogram. When the height of the histogram pane is increased, the histogram scales to use the area available. Use your mouse to pull on the divider between the upper and lower panes to increase or decrease the height of the lower pane.

Sequence and Trace

If the assembly was performed using sequence files from automated sequencing machines and these files contain trace data (ABI or SCF format files), you can view the aligned sequences with one or more trace chromatograms shown below the sequences in a split pane.

Click the Select View toolbar button and select Seq and Trace to display one trace chromatogram.

Click on a base in one of the aligned sequences to view the trace data for this sequence at and around this base position. Vertical and horizontal highlighters mark the sequence and base in the upper panel that you have selected. The lower panel displays the trace data with the selected base marked with a vertical highlighter and centered in the display, if possible. Click on another base to view trace data at another position.

Click the Select View toolbar button and select Seq Multi Trace to display trace chromatograms for all sequences shown.

Click on a base in any line of aligned sequence data to view trace data for all sequences at this position in the alignment. A vertical highlighter marks the column of sequence bases you selected. The lower panel displays the trace data with the appropriate base marked in each sequence frame.

Use the mouse to pull on the assembly viewer window frame to make the entire data window larger, if needed. Use the mouse to move the split bar between the upper and lower panes to change the proportion of space available for the sequence or trace displays.

- **Navigation**

  You can use the horizontal scroll controls in the upper sequence panel, or the keyboard arrow keys to move the contig to the left or right. If the selected base stays within the visible sequence, the vertical highlighter remains on this base and the trace data in the lower panel continues to show the trace for the selected base, centered within the display.

  If, however, the vertical highlighter in the upper panel hits either the left or right window frame, the highlighter will begin to shift from one sequence base to the next, to remain within the visible display. At this point, the vertical highlighter in the lower panel will also begin to shift, synchronizing the trace data to the sequence data scrolling in the upper display.

  To move the vertical highlighter to the left or right within the sequence, use Alt + arrow key or just click on another base with the mouse.

- **Multi-Trace**

  Shows trace profiles for all sequences displayed.

Edit Consensus Sequence

When viewing sequence assembly results, you can edit the consensus sequence if you believe there is an error in the base shown. You can replace one base with another or remove a base from the consensus sequence, if required. And you can quickly find ambiguous bases in the consensus sequence, where problems with base calling may be most likely.

You can edit the consensus sequence from any display screen that shows sequence data. You might find the Sequence and Trace display the most helpful in identifying miscalled bases. Click the Go To Ambiguous toolbar button to move to the next position in the consensus sequence that has an ambiguous base.

To edit the consensus sequence:

1. Click on the base in the consensus sequence that you want to change.
2. Click the Edit Consensus Sequence toolbar button.
3. Verify that the correct base is shown, and the edit pointer (black arrowhead) marks the correct base in the sequence.
4. To replace a base, click the Change Base option and enter the base you want to use as a replacement.
5. To remove a base, click the Remove Base option.
6. Click OK to make the requested change to the consensus sequence.

When replacing a base using this method, you can enter (type in) any valid base, including ambiguous bases. When removing a base, a gap will be shown in the consensus sequence on the screen. If this sequence is entered to the molecule list and later used or saved to disk, the gap will be closed.

Remember changes:

If you have already saved the sequence assembly results to a disk file, you should save to disk again so that the changes made to the consensus sequence will be written to file. To save just the consensus sequence to a disk file, click the Enter Contig to Molecule List toolbar button and then click File, Save when the molecule viewer window shows the newly entered molecule.

Recall base in trace file

You can right-click on a base in one of the aligned trace file molecules and select Recall Base to change the base called at this position. You can use the checkbox to indicate that you want to also update the sequence in the trace file and overwrite this file on disk.

NCBI BLAST

BLAST (Basic Local Alignment Search Tool) is a tool administered by NCBI (National Center for **Biotechnology** Information) to search sequence databases to find sequences that show statistical homology to a search sequence. (Altschul, S.F., W. Gish, W. Miller, E.W. Myers and D.J. Lipman. 1990. Basic Local Alignment Search Tool. J Mol Biol 215:403-410.)

Submitting a search request to the NCBI BLAST server involves identifying the search sequence and optionally identifying the region to use for the search. You can then select options that identify which type of BLAST search and which database to use. Clone Manager will submit the search request to the online NCBI BLAST server (internet connection required) and wait for the response. Depending on how busy the NCBI servers are, the request may take some time to process. You will see a task management window displaying that the job is waiting for a response.

When Clone Manager receives the results of your BLAST search, it will display the results in a tabular format and the toolbar can be used to act on the results.

Starting a BLAST Search

Start the BLAST setup wizard by using the main menu Align, BLAST Search (NCBI REST).

- BLAST URL – is the URL used to submit requests to the NCBI servers. NCBI reserves the right to change this URL at any time. Should they do so, you can enter the updated URL here.
- Options – Internet access configuration. Used in case your computer needs special configuration to access the internet. Examples are using a proxy server. Please consult with your IT support department because only they will know how to configure your internet connection.

Selecting Search Molecule for BLAST

- Molecule - When the wizard starts, it will automatically select the currently active molecule. You can select a different molecule by clicking the ‘Change’ button.
- Match region – allows selecting a partial region to use for the search. This is useful if you want to search for matches to a gene or region of a large molecule.
- Process – Allows for translating a DNA sequence and submitting the search as an amino acid sequence.

Selecting BLAST Options

Once you have identified the query sequence, you can select the BLAST program to use and the database to search.

1. Use the **upper** combo box to select the BLAST program you want to use. A description of the selected program is provided below the selection. The programs available depend on whether your query sequence is DNA or protein.
2. Use the **lower** combo box to select the BLAST database you want to search. A description of the selected database is provided below the selection. The databases available depend on the program selected (above).
3. Click the **Finish** button to prepare your query sequence and send your request over the internet to the NCBI BLAST server.

BLAST Programs

BLAST programs blastn, blastx, tblastx, and MegaBlast are available for DNA query sequences. blastn compares against a nucleotide sequence database, blastx translates and then compares against a protein sequence database, tblastx translates both the query sequence and the sequences in the nucleotide database. MegaBlast is specifically designed to efficiently find long alignments between very similar sequences and find identical matches to your query sequence.

BLAST programs blastp and tblastn are available for protein query sequences. blastp compares against a protein sequence database and tblastn compares against a translated nucleotide sequence database.

BLAST Databases

A variety of protein or nucleotide sequence databases are available for your search. Descriptions are provided below the selection combo box.

The database combo box will limit your selection to those databases appropriate for the BLAST program you selected for use. In addition, please note that the tblastx program should not be used with the nr database.

Viewing BLAST Search Results

The BLAST search request is sent to the NCBI BLAST server. Results are displayed in a table format. Select the homology you are interested in and use toolbar buttons to view the GenBank file or export the results table.

Displayed Results:

Shows columns for: Accession, Score, Length, Gi, Description for each homology found. High scores equate to more significant or extensive matches.

Columns can be resized using the mouse to drag the dividers in the column header.

Use the toolbar to act on the selected homology.

Results Toolbar:

- Show GenBank File – retrieves and displays the homologous molecule. This may take some time if the homology is genome sized (Length column) or if the Internet connection to NCBI is slow. You can copy the file to the clipboard for importing into another program or print the file.
- Get Molecule – retrieves the homologous molecule from the NCBI server for use in now in Clone Manager. You also have the option to save a copy of the original GenBank file to disk. The sequence data, genes read from the features table, and descriptive information will all be a part of the new molecule and can be configured in Preferences.
- Export Results Table – copies the list of BLAST hits to the clipboard in tab-delimited format. You can then paste the results into another program such as a spreadsheet.
- RID – opens the search results at the NCBI web site. You can then explore the results in more detail using their analysis options. Please contact NCBI for more information.
- Program – reports the BLAST program used for this search
- Database – reports the database searched.

Sequence Phrases

Sequence Phrases are collections of motifs that can be used to search and annotate molecules.

Clone Manager will store sequence phrases in collection files that you can organize. You can inspect a molecule to find sites that match sequence phrases in your collection files. Or you can scan for molecules that contain a given sequence phrase.

- Sequence Phrase Collections
- Construct and Use Sequence Phrase Collections
- Export / Import Sequence Phrase Collections
- Find Sequence Phrases
- Scan for Molecules with Sequence Phrase Sites

Sequence Phrase Collections

Collections of sequence phrases are stored in special sequence phrase collection files (*.sx5). These collection files contain essential information about your sequence phrases, often stored in a group of phrases that are related in some way. You can create new collection files, as needed, and create folders to organize your collection files. All your collection files are stored within the folder SeqPhrases, located in your home directory.

The information stored for each sequence phrase includes the sequence phrase name, description, sequence, and type (DNA or Protein). For each sequence phrase, you also select the match mode to be used when searching molecules for this phrase and the way you want to add found sites to your molecule file. Match modes include Exact match required, Allow limited mismatches, or Allow gaps and limited mismatches. Found sites can be entered as genes, regions, or labels.

You can open your sequence phrase collections using the Sequence Phrases button on the main toolbar or by using the View main menu.

Shared Sequence Phrase Collections

You can share sequence phrase collection files if these files have been made available for you.

To access the shared resources, click File, Preferences, Share Data, and enter the location (complete path and folder) where you can access shared enzyme lists, primer collection files, and sequence phrase collection files. When you open the sequence phrase collection window, shared collection files will appear in a folder below your sequence phrase collections. Shared collection files will be opened as read-only resources.

More information

Construct and Use Sequence Phrase Collections

Viewing Sequence Phrase Information

Sequence Phrase collections can be viewed using the main menu View, Sequence Phrases option.

The Sequence Phrases dialog box consists of three panels:

- Toolbar with options for managing sequence phrases.
- Left panel with a list of the Sequence Phrase collections available for selection.
- Right panel shows the content of the selected collection in the left panel.

Sequence Phrase Toolbar

Provides the ability to create and organize collections and their contained sequence phrase entries.

- Folder tools – manages the left panel list of folders and collections. You can also right mouse click the selected item to get the same options.
    - New Folder – create a new folder which will contain individual collection files. Allows organizing collections into groups of similar purpose.
    - New (empty) file – create a new collection file to which you can add new sequence phrases.
    - Import – Import a collection file that was previously saved or provided to you by a colleague.
    - Rename – change the name of the selected collection file or folder.
    - Delete – remove the selected collection file or folder. If the selected item is not empty, you will be asked to confirm
    - Export – Export the selected collection. You can export to a disk file to share the collection with a colleague or to use within a spreadsheet program. You can also copy to the Windows clipboard.
- Print – print the selected collection file.
- Copy – copy the selected collection file to the Windows clipboard.
- Edit – edit an existing phrase.
- Add – add a new sequence phrase.
- Remove – remove the selected sequence phrase.
- Copy entry – copy the phrase to another collection.

Entering Sequence Phrase Information

You can view, edit, or enter information about a sequence phrase. Access from the main menu View, Sequence Phrases. Select the collection to use and use the toolbar to add or edit a sequence phrase. The View/Edit Sequence Phrase dialog box has the following data entry fields:

- **Collection**

  The collection file name is reported. This location cannot be edited. For a new entry, the active collection file open in the left panel is shown -- this is the location where the new sequence phrase will be entered.

- **Name**

  Enter a name for the sequence phrase. If you plan to enter the sequence phrase site as a label site on a molecule map, a short name is helpful.

- **Description**

  The description is saved as a part of the collection file format. Enter a short description of the sequence phrase. This description will appear on some program screens and printed reports.

- **Type**

  Set to DNA or Protein, matching the type of bases you will enter for the sequence phrase.

- **Sequence**

  Enter the sequence of bases for your new sequence phrase or copy from the clipboard.

- **Match mode**

  Select the match mode to be used when searching molecules for this phrase. You can select Exact match required, Allow limited mismatches, or Allow gaps and limited mismatches. When searching for sequence phrases you can use the Automatic option to use the match mode you set or you can override for a specific search, as needed.

- **Enter as**

  Specify how you want found sites added to your molecule. You can enter found sites for this sequence phrase as a gene or region, setting a style for the feature type. This option works well for longer sequence phrases that will show up well as features in your molecule. You can also enter found sites as a label. This option works well for short sequence phrases. You can show these on your molecule map in place of enzyme sites or primer sites. When you complete a search for sequence phrases, found sites will always be entered as specified.

Construct and Use Sequence Phrase Collections

You can enter new sequence phrases directly to a collection file or you can import sequence phrase collection data. Once you have sequence phrases in your collection files, you can view, edit, organize, or select these sequence phrases for use. You can also search your collection files to find a specific sequence phrase.

Getting started

Click on the Sequence Phrases toolbar button to open the sequence phrase collections.

Click on the My SeqPhrases folder to open folders or files in this location. When you are just getting started, you will have one empty file ready for your first sequence phrase to be added to this collection file. You can click the first toolbar button Folder Tools to add new folders or new empty collection files, or to rename or delete any of the files or folders you have created.

- **Enter a new sequence phrase** -- click on a collection file to open this collection. Then click the toolbar button Add New Entry. Fill in the information about the new sequence phrase.
- **Import sequence phrase collection data** -- you can use an import function to prepare a sequence phrase collection from a tab-delimited text file that contains sequence phrase information prepared in another application.

Using your Sequence Phrase collections

Click on the Sequence Phrases toolbar button to open the collection. Click on the My SeqPhrases folder to open folders or files in this location.

Click on the Folder Tools toolbar button to add new folders or new empty collection files, to rename or delete any of the files or folders you have created or to export or import sequence phrase collections.

Click on a collection file in the left panel to display the contents of this file in the right panel. For each sequence phrase in this collection, the name, description, length, enter as and match mode are shown.

Click on column headers to sort by the values in that column. Move the mouse to the join between two column headings, depress the mouse button and drag to resize the column. Use the mouse to drag out the dialog box borders to resize the window.

- **View or Edit sequence phrase information** -- select a sequence phrase by clicking on the entry in the list in the right panel. Use the toolbar buttons to edit the information about this phrase, delete this entry, or copy the entry to paste into another collection file. When changes have been made to your collection files, these changes are automatically saved for you.
- **Document collection file contents** -- You can use toolbar buttons to print the contents of the active collection file or send this list of sequence phrases to the

clipboard for use in another application. Or you can export the data in the active collection file in tab-delimited format to import into another application -- click the Folder Options button and select Export.

Export / Import Sequence Phrase Collections

You can use the Export option to prepare a tab-delimited text file containing the information in a sequence phrase collection file. Programs like Microsoft Excel or Microsoft Access can import this data allowing you to use this information in other applications.

Export a sequence phrase collection

1. Click the Sequence Phrases button and select (highlight) the sequence phrase collection you want to export.
2. Click the Folder Tools button and select Export.
3. In the Export Collection dialog box, select which data to export.
4. Use the checkbox to indicate if you want the exported file to contain headers at the top of each column of data.
5. Select to export to a disk file or copy to the clipboard.

You can use the Import option to read in data contained in a tab-delimited file prepared in another application in order to create a new sequence phrase collection.

Import a sequence phrase collection

1. Click the Sequence Phrases button and click My SeqPhrases.
2. Click the Folder Tools button and select Import.
3. Select the tab-delimited text file containing your sequence phrase data.
4. In the Import Collection dialog box, specify the first row that contains sequence phrase data to import. (If your file has column headers, the data usually starts at row 2.)
5. Use the small drop-down arrow keys above each column to pick the type of data in that column or indicate that the data should be skipped (not imported).
6. If your data does not contain match mode or enter as information, you can select one match mode and enter as type to be used for all sequence phrases in the new collection.
7. Enter the name for the new sequence phrase collection and click OK.

Find Sequence Phrases

Found on the Discover menu or use Tools button to access from map window.

You can use this function to find sites where sequence phrases are found in the active molecule. You can scan for sequence phrases that are stored in your Sequence Phrases collection files (*.sx5). The program will open each sequence phrase collection, read each phrase sequence, and attempt to match the sequence in the active molecule.

You can set the match mode to Automatic to use the match mode settings associated with each stored sequence phrase. Or you can override this setting and select a match mode to use for all sequence phrases during this search. When the search has been completed, you can view the search results. Sequence phrase sites found can be added to the molecule file and can be displayed on the molecule map or sequence display.

Selecting sequence phrase files or folders to search

Click the Add Collection button to access your Sequence Phrase collections and select a file or folder to add to the search list.

Select the collection file you want to add to the list and click OK. During the search, the program will attempt to open the file and read all the sequence phrase data contained in the file. If you have selected a folder of collection files, the program will attempt to open each file within the folder.

Set the Match Mode

Use the dropdown control to select the match mode to use for this search. Select Automatic to use the match mode setting associated with each stored sequence phrase. To override these stored settings and use one specified match mode for all sequence phrases in this search, select one of the following settings:

Exact match required -- all bases in the sequence phrase must match the bases in the molecule sequence.

Allow limited mismatches -- the number of differences is limited to 1 mismatch for each 10 bases in the search sequence. For short search sequences of less than 10 bases, one mismatch is permitted.

Allow gaps and mismatches -- the number of differences is limited to 1 gap or mismatch for each of 10 bases in the search sequence. For short search sequences of less than 10 bases, one difference (gap or mismatch) is permitted.

When you have selected the files and folders to search and set the match mode, click **OK** to begin the search for sequence phrase sites.

Note: If your sequence phrase is long, you should think about allowing gaps so that an unexpected gap or insertion (such as a sequencing error) will not cause a failure to find the phrase site.

Scan for Molecules with Sequence Phrase Sites

Found on the Discover menu. You can use this function to scan molecule files to look for sites where the specified sequence phrase can be found. You can scan molecule files in folders on your computer or local network. The program will open each molecule file, read the sequence and search for matching sites.

You can use the match mode stored for this sequence phrase (select Automatic match mode) or override using a different match mode for this search. When the search has been completed, you can view the search results. You can print or copy the list of molecules and binding sites found.

Selecting molecule files or folders to search

Click the Add Folder button to select a folder location to add to the search list.

Navigate to the folder you want to add to the list and click OK. The program will attempt to open and read all molecule files in the location specified, including those contained in subfolders.

Click the Add File button to select a molecule file to add to the list.

Navigate to the molecule file you want to add to the list and click OK. The program will attempt to open and read this file during the search procedure.

If you routinely use the same molecule file locations, you can save the list of files and folders and recall this list at a later time.

Click **OK** to begin the scan for molecule sites.

Find Molecule Sites Results

When a search for molecule sites for the specified sequence phrase has been completed, the results will be presented in a scrolling list box.

In the results display, the molecule name is listed, as well as the base pair position of the 5' end of the matching site. Also reported is the number of differences (mismatches or gaps, if applicable). Click on the column headers to sort by molecule name or file name.

You can print or copy the results list or load a molecule file for use.

Find Sequence Phrase Results

When a search for sequence phrase sites has been completed, the results will be presented in a scrolling list box and location map display. You can print or copy this list to the clipboard. You can add some or all these sites to your molecule file for display on the map or formatted sequence.

In the results display, the found sequence phrase sites are listed, showing the phrase name and the base pair position of the 5' end of the sequence phrase site. Also reported are the number of differences (mismatches or gaps, if permitted), the length of the phrase and the sequence found. Mismatches are shown using upper case letters and gaps appear as dashes. Click on the column headers to sort by name or position of the sequence phrase site.

The vertical bars above the icon map show the approximate location of the sequence phrase sites. A red arrow and red vertical bar mark the highlighted site.

To tag sequence phrase sites you want to add, mouse-click on the checkbox in front of the phrase name. Click the Tag All button to automatically tag all the sites in the list. To untag a site, mouse-click again or click the Clear button to untag all sites in the list.

Click the **Add Sites** button to add all the tagged sequence phrase sites to the molecule file and close the results window.

Note: The dialog display of found sites shows only the first 250 bases of the matching sequences. If your sequence phrase is longer, you can use the Print or Copy functions to show the complete sequences, with all mismatches, gaps, and insertions.

Search in Sequence Phrase Collections

You can search in your sequence phrase collection files to locate a specific phrase or a group of sequence phrases that have a common text string in the entry. You can search using a sequence phrase name or any text that could be found in the name, description, or sequence fields.

To start a search

- Open the Sequence Phrase collections
- Click on the collection file or folder in the left panel that you want to search
- In the Find edit box (lower area), type in the text to search for
- Click the Go button

If sequence phrases are found that meet the search criteria, these will be listed in a new folder called Search Results, found at the bottom of the folder list in the left panel. (Click on the Search Results folder to open this folder if it is not already open.) You can select a sequence phrase for use from this list or you can print the list or copy it to the clipboard. When you close the sequence phrase collections module, the Search Results folder will be discarded.

You can click the small question mark button to the right of the Go button to select the fields you want to search. You can also search within the Search Results folder, to further refine your search, as needed.

DataBook

Access from the View menu or use the main toolbar button to access.

DataBook is a simple, built-in database you can use to organize and store information about your molecule or data files, including primers or alignment results. You can think of it as a list of the files that you want to keep track of and just enough information about each file to let you find and open the file when you need it.

You can construct a DataBook from existing molecule or data files using a batch procedure in the Multiple File Conversion utility. Or you can add molecules or data files directly to an existing DataBook with the click of the Add Entry, Add to DataBook toolbar button. In either case, the program will create new records and pre-enter most of the important information for you.

You can create new DataBook files, as needed, and create folders to organize your DataBook files. All your DataBook files are stored within the folder DataBook, located in your home directory. These files have the extension *.dx5.

- Using the DataBook Database
- Add to DataBook
- View000Edit DataBook Record
- DataBook Searches
- Dat0Book Tools
- Multiple File Conversion

Using the DataBook Database

Once you have records in your DataBook files, you can view, edit, or organize the records within the DataBooks. You can also search the DataBooks to find a specific molecule or data file and open the file for use.

With the DataBook module open, click on the My DataBooks folder to open folders or files in this location.

Click the Folder Tools toolbar button to add new folders or new empty DataBook files, rename or delete any of the files or folders you have created, or export or import DataBook records.

Click on a DataBook file in the left panel to display the contents of this file in the right panel. For each DataBook record, the information in the Data Info, Description and File Name fields will be shown.

Click on column headers to sort by the values in that column. Move the mouse to the join between two column headings, depress the mouse button and drag to resize the column. Use the mouse to drag out the dialog box borders to resize the DataBook window.

View or Edit DataBook Record

Select a record by clicking on the entry in the list in the right panel. Use the toolbar buttons to edit the information in this record or view additional information, delete this entry, copy the entry to paste into another DataBook file, or view the properties of the actual molecule or data file. When changes have been made to a DataBook record, these changes are automatically saved for you.

Load a File for Use

Click on a record and then click the Load File button at the lower right to use this molecule in an on-going operation or to open this file for use now. If more than one file can be selected for an operation, use the Shift or Control keys in conjunction with the mouse to select additional files.

Constructing a DataBook using the Multiple File Conversion utility

Adding molecules or data files to an existing DataBook

Add to DataBook

You can add a molecule, primer, or alignment results file to the built-in DataBook database in a very simple, efficient procedure. The program will pre-enter most of the important information for you. You can leave the entry as is, or you can change or add additional information that will appear in the DataBook only (your molecule, primer, or data file will not be changed).

Add new molecule or data file

1. Open the molecule, primer, or alignment results file.
2. Click the Add Entry button (main toolbar) and select Add to DataBook
3. Identify the DataBook where you want to place this new record.
4. Use the checkbox (lower right) if you want to view the record as soon as it is created.
5. Add or modify information in the data entry fields, as needed, and click OK to save.

You can also enter a new record manually. Just open the DataBook database, click the Add New Entry toolbar button, and fill in the fields with the information about your file.

Copy records already in a DataBook

When viewing records in your DataBook files, you can select a record you want to copy and then click the Copy Entry button. Select the DataBook file you want to add this record copy to and click OK.

View/Edit Databook Record

DataBook Table Construction

One new DataBook table (*.dx5) -- Writes a record for each appropriate file and saves the records to one new DataBook table (in your Home Directory, DataBook). The information entered in each record is dependent on the type of file -- for molecule files, the record will contain the full file name and path, the molecule name and size, feature names and descriptions, and the molecule description and notes (if any).

Set of Files (*.dx5) in a new Folder -- Makes a separate DataBook table (file) for each folder of files found and saves these files in a new folder in the DataBook location (Home directory). Each DataBook table will contain a record for each file it can read, storing essential information about the file (as described above). Each DataBook table will have a name derived from the original folder name. Nested folders will have names that include the parent folder designation to maintain clues to data organization.

Sharing DataBooks

You can share DataBook files if these files have been made available for you.

To access the shared resources, click File, Preferences, Share Data, and enter the location (complete path and folder) where you can access shared enzyme lists, primer collection files and DataBook files. When you open the DataBook, shared DataBook files will appear in a folder below your DataBooks. Shared DataBook files will be opened as read-only resources.

More information

Setting up Shared Resources

View/Edit DataBook Record

You can enter or edit information about your molecule, primer, or data files. If you use the Add to DataBook option, the program will enter information into most of the fields. Any additional information you enter or changes you make will reside in this database file and will not affect your original molecule, primer, or data files. Records in shared DataBooks cannot be edited.

- **DataBook**

  The location of this record.

- **File Name**

  The full file name and path for the molecule, primer, or data file. You can use the Browse button to navigate to the file name and update this information, if needed (see note, below).

- **Data Info**

  A brief identification of the data contained in this file. Molecule name and size, primer name and size, or alignment type are entered in the Add to DataBook procedure.

- **Features/Details**

  Add to DataBook will enter feature names and descriptions (molecule file), primer sequence and linked molecule information (primer file), or alignment results summary data (alignment results file).

- **Description**

  Add to DataBook will enter the molecule, primer, or alignment description found in the file. If you enter extra information in this DataBook field, you may want to put it first so that it will be visible at the start of the field in the standard DataBook list display.

- **Comments/History**

  Add to DataBook will enter the notes found in a molecule or primer file. You can use this additional space for more descriptive information, as needed.

  Note: If you need to change the path for many file names in your DataBook, you can use DataBook Tools, found on the Tools menu, to do a find and replace operation for the file path.

DataBook Searches

You can search in your DataBook files to locate a specific molecule, entries with certain genes, or records that have a common text string. You can search using a molecule name or any text that could be found in the data info, features/details, description, or comments/history fields.

To start a search

- Open the DataBook module
- Click on the DataBook file or folder in the left panel that you want to search
- In the Find edit box (lower area), type in the text to search for
- Click the Go button

If records are found that meet the search criteria, these records will be listed in a new folder called Search Results, found at the bottom of the folder list in the left panel. (Click on the Search Results folder to open this folder if it is not already open.) You can select a file to open for use from this list or you can print the list of entries or copy it to the clipboard. When you close the DataBook module, the Search Results folder will be discarded.

If you do not want to search all four of the default fields, click the small question mark button to the right of the Go button. Click to uncheck fields that you do not want to search. You can also search within the Search Results folder, to further refine your search, as needed.

DataBook Tools

Found on the Tools menu.

You can use DataBook Tools to change the file paths for groups of files in your DataBooks and you can convert old DataBooks to the new format, if you did not do so previously.

Change File Paths

If you have moved your molecule files from one location to another or have changed a mapped drive letter or renamed folders, you can quickly and easily do a find and replace operation on the file name field in a selected DataBook.

1. Click Tools, DataBook Tools
2. Set the **Change File Paths** option.
3. Complete the data entry fields, using the suggestions below.
4. Click OK.

Change from: Enter the complete file path up to, and including, the text that needs to be replaced. This path should begin with the drive letter. It should not include the actual individual file name.

Change to: Enter the complete replacement for the text entered in the "Change from" field (above). This path should begin with the drive letter. You can click the Browse button to navigate to the new location and paste the folder path information.

DataBook: Click the Browse button to select the DataBook that contains the records where you need to change the file path.

The program will check the file name field for each record in the DataBook and make the text replacement, where appropriate.

If you have old DataBook files containing files that have been moved or you are not confident that the old DataBook files are very useful, you might find it easier to create new DataBooks de novo using the Multiple File Converter.)

Export / Import DataBook Contents

You can use the Export option to prepare a tab-delimited text file containing the information in a DataBook file. Programs like Microsoft Excel or Microsoft Access can import this data allowing you to use this information in other applications.

Export a DataBook file:

1. Open the DataBook module and select (highlight) the DataBook file you want to export.
2. Click the Folder Tools button and select Export.
3. In the Export DataBook dialog box, select to export to a disk file or copy to the clipboard.
4. Use the checkbox to indicate if you want the exported file to contain headers at the top of each column of data.

You can use the Import option to read in descriptive data contained in a tab-delimited file prepared in another application in order to create a new DataBook file. You can also use this option to read in the data in a file previously exported using the export DataBook option (above).

Import a DataBook file:

1. Open the DataBook module and click My DataBooks.
2. Click the Folder Tools button and select Import.
3. Select the tab-delimited text file containing your descriptive data.
4. In the Import DataBook dialog box, specify the first row that contains data to import. (If your file has column headers, the data usually starts at row 2.)
5. Use the small drop-down arrow keys above each column to pick the type of data in that column or indicate that the data should be skipped (not imported).
6. Enter the name for the new DataBook file and click OK.

Note: If you want to create a new DataBook using your existing molecule or data files, you can use the Multiple File Conversion utility to construct the new DataBook table.

Multiple File Conversion

The Multiple File Converter is found on the main menu File, Other Tools and is a utility function that can convert or build using files and folders of data on your computer.

You can convert molecule files in standard Clone Manager file formats (*.cm5, *.cx5) to another file format for sharing with others, archival storage, web applications, or use by other programs. You can also construct primer collections using your individual primer files in standard primer file formats (*.pd4, *.pd2) or construct DataBook tables from your molecule or data files in one easy batch operation or upload multiple molecules to the Clone Share companion application.

The file conversion utility will attempt to open and convert the individual files you specify, and all appropriate files found in the folders you specify, including those contained in subfolders.

- Select the file conversion process:
    - Molecule files – convert from one file format to another.
    - Molecule files – upload to Clone Share.
    - Primer Collections – construct from individual primer files.
    - DataBook tables - construct from molecule or data files.
- Use the toolbar to identify the files or folders of files to process:
    - Add Folder - Click to select a folder location to add to the list. Subfolders will be included. The list will show the number of files and their total size.
    - Add File – select one or more files to add to the list. The list will show the size of each file selected.
    - Remove – the selected item from the list.

Molecules convert format

Converts standard Clone Manager molecule files into another format.

- Compact GenBank format (*.gb) - Files will contain essential molecule data in a format that is consistent with the standard GenBank file format. The data will include the molecule name, size, description, features (genes, regions, markers, labels), and sequence data, all with standard GenBank identifiers. This file format is appropriate for sharing with others who do not use the Clone Manager program.
- Commented GenBank format (*.gbk) -- Files will contain a larger subset of molecule data in a format that is consistent with the standard GenBank file

format. The data will include the molecule name, size, description, all features (including information-only features), sequence data, information entered in the Notes field, and the 'draw as' type you assigned for each feature. This might be the format of choice for constructing a repository of your important files.

- FASTA format (*.fa) -- Files will contain the sequence plus one line of descriptive information (molecule name, description) in FASTA format. Many web applications require sequence data in this format.
- Clone Manager XML format (*.cx5) -- Files will contain all the information found in the standard Clone Manager file format (*.cm5), but in a structured, text-based format that can be read by humans or computers. This information can include enzyme sites, primer sites, a user base offset, or links to trace files, as well as standard molecule information. Use this file format for sharing data across applications.

Use the browse button to specify the folder where the converted files will be written to. Each converted file will appear as a separate file.

Upload to Clone Share

Uploads standard Clone Manager molecule files to the Clone Share store.

Click the browse button to select where you want to upload the selected list of files and folders.

Create Primer Collections

Builds one, or more, primer collections from standard Clone Manager primer files.

- One new collection file:

  Collects all primers found and saves in one new primer collection file (in your Home directory, Collections). The collection file will contain essential information about each primer, including the primer name, description, sequence, and primer type. If notes information is available for a primer, this information will also be saved in the collection file.

- Set of files in a new folder:

  Makes a separate collection file for each folder of primers found and saves these files in a new folder in the Collections location (Home directory). Each collection file will contain essential information about each primer, including the primer name, description, sequence, primer type, and notes. Each collection file will have a name derived from the original folder name. Nested folders will have names that include the parent folder designation to maintain clues to data organization.

- Destination – enter the name of the new collection file or folder.

Create DataBook tables

Builds one, or more, DataBook files. All readable molecule files will be included. Use the checkboxes if you also want to add primer or alignment files.

- One new DataBook table – creates a new DataBook with a record for each file selected.
- Set of files in a new folder – if your input selection contains folders, each folder of primer files will be used to create a separate collection of primers.
- Destination – enter the name of the new DataBook file or folder.

Preferences

Found on the File main menu. Allows setting user preferences for program functions.

The tabbed dialog box is divided into functional categories:

- Basic Preferences
  - Home Directory – used to store user settings, files, and collections.
  - Author Stamp
  - Default Molecule File Format
  - Exit prompt – prompt when exiting program with unsaved changes to molecules.
- Options
  - Toolbar size
  - Assembly Cloning Wizard style
- Import Features – used to set which features will be imported when loading a GenBank or EMBL databank file.
- Colors
- Custom Toolbar
- Share Data

Basic Preferences

You can use menu File, Preferences to set a home directory, enter default author information, and set the default file format for saving molecule files.

To set basic preferences:

1. Click File, Preferences, then click the Basic tab.
2. Specify your home directory
3. Enter an author name that should be used for all files generated from the local computer
4. Select the default file format, using the guidelines below.

Home Directory

The home directory is the location where the program will look for personal configuration data and user-specific files and folders (DataBooks, enzyme lists, primer collections). The default home directory is C:\Users\[name]\Documents\CMHome.

Author Stamp

You can add an author name, date, and notebook reference to the beginning of the Notes field for your Clone Manager molecule files. Adding author information can help you to identify the source of a file or the location of needed documentation. Set the Prompt checkbox if you want to be reminded to Author Stamp when you save a file for the first time. Click Operations, Author Stamp to add author information to the current molecule file now or use the Author Stamp toolbar button when editing molecule information on the Info tab of the molecule viewer window.

Default Molecule File Formats

You can set a file save format the program should use when saving molecule files for the first time. You can save molecule files in the standard Clone Manager molecule format (*.cm5) or the Clone Manager xml format (*.cx5). The *.cm5 format is preferred and has a file association set (you can double-click on a *.cm5 file in Explorer and open Clone Manager with this file loaded).

The *.cx5 xml format is a human readable format that can make data portability easier. Although this file format is readable, it should not be edited. *.cx5 files prepared with Clone Manager 9.2 and later are encoded UTF-8 and cannot be used with earlier versions of Clone Manager.

Options Preferences

Toolbar Size

Used to select the size of toolbars in the main and viewer windows.

- Small – best for compact lo-res displays
- Default – best for most computer displays
- Large - best for touch enabled displays

Assembly Cloning Wizards

Used to enable the old cloning wizards used in version prior to version 12.

Import Features Preferences

You can use menu File, Preferences to identify the feature keys that should be imported when you open a file in standard GenBank or EMBL format.

To set feature import preferences:

1. Click File, Preferences, then click the Import Features tab.
2. Select designated feature keys, using the guidelines below.
3. Set the option for managing all other, non-designated features.

Designated Features

Click the Add button to include a new feature key on the import list. When the dialog box appears, select the Feature Key from the list of all keys currently supported in GenBank or EMBL file formats. Select ‘Draw as Type’ to tell the program how to display this feature in Clone Manager. When you next open a GenBank or EMBL file, the program will try to read in features with this feature key.

Draw as Types Gene and Region supports three alternate drawing styles so that you can divide features within these broad categories. Genes are drawn inside/below the mapline and can be translated. Regions are drawn on the mapline. (You can use the Format Map option to assign feature shapes and colors for styles 1, 2, and 3.)

Click the Edit button to modify the information for the selected (highlighted) feature key. Click the Remove button to delete the feature key from the import list

All Other Features

This option is used to tell the program how to manage all feature keys that are not listed as Designed Features (as described above).

Discard all non-designated features -- keep your features table and file size as small as possible.

Import as Information-only features -- these features do not appear on the map or annotated sequence but can be viewed on the features table if you press the Info Features toolbar button.

Feature key selection list

When you are adding a new feature to a molecule and you are selecting a feature key, the program will show you the short list containing your designated feature keys. (You can click More... to see the entire list of possible feature keys.) If you find you often use a feature key that is not on the short list, you can add it to the list of designated features. This will also ensure that such features will be imported when encountered in GenBank files.

Color Preferences

Allows setting your color display preferences.

- Alert (warning) text color – used to direct your attention to important text.
- Auxiliary text color - used to distinguish less import or subsidiary text.
- Highlighter color – Bars of bright color used as highlighters.
  - Use this color to mark selected items in results displays.
- Rule and box color – used to separate data into lines or box text.
- Reset – used to restore the original program settings.

Custom Toolbar

Allows customizing the main program toolbar. You can add, remove, and reorder the buttons you see to make it easier to do operations that you do frequently.

To update the main toolbar, move action icons between the left and right panels. The 'Separator' icon adds a spacer between action icons and is useful for creating groups of actions.

When you have made the changes that you want, click the dialog OK button and the main program toolbar will be updated to reflect the changes you have made. Each time you start Clone Manager, the toolbar will show your preference.

Toolbar list

The list on the left shows the actions available on the main toolbar. Use the buttons on the central control toolbar to move actions.

- Right facing arrow – remove the icon action from the active toolbar list.
- Up facing arrow – move the icon action up in the list.
- Down facing arrow – move the icon action down in the list.
- Info button – get more information about the selected action.

Available actions list

The list on the right shows actions that are not currently on the active toolbar list.

- Left facing arrow – add the action to the active toolbar list. The action will now appear on the main menu.

Clear

Removes all of the action icons from the main toolbar. This provides a quick way to reset the toolbar and enter the icons you want in the order you want.

Reset

Restores the active toolbar to the icons originally set in Clone Manager. These settings are the program default toolbar.

Setting up Shared Resources

You can share enzyme lists, DataBook database files, and sequence phrase collection files with others in the group. If you are using the Clone Manager Professional suite, you can also share primer collection files.

Files that can be shared

Enzyme master file, file name SEEnz32.enz
Set of user lists, file name SEUserList.dat
List of special enzymes, file name SEUserDef.enz
GenBank import filter, file name SEKeyList.dat

DataBook database files, *.dx5
Sequence phrase collection files, *.sx5
Primer collection files, *.px5

Shared resources are placed in one location on a network file server where everyone will be able to access these files. Each user then identifies this location in File, Preferences, Share Data in their program settings. When shared enzyme lists, shared DataBook files, shared sequence phrase collection files, or shared primer collection files are opened from this location, they will be opened as read-only resources. Personal user enzyme lists, DataBook files, sequence phrase collection files, and primer collection files can be used, in addition to the shared resources.

- **Setting up shared resources**

A folder should be created for shared resources in an accessible location. To share the enzyme master file, copy the file SEEnz32.enz to this folder. To share a set of user lists, copy the file SEUserList.dat to this folder. To share DataBook database files, copy the DataBook files (*.dx5) to a subfolder DataBook, in the same shared location. To share sequence phrase collection files, copy the collection files (*.sx5) to a subfolder SeqPhrases, in the same shared location. To share primer collection files, copy the collection files (*.px5) to a subfolder Collections, in the same shared location (see sample, below). In this example, the shared location is the folder Shared_Resources with its path information.

```
Shared_Resources
   SEEnz32.enz
   SEUserList.dat
   SEUserDef.enz
   SEKeyList.dat
   Collections
      AListPrimers.px5
      Project104.px5
   DataBook
      CloningVectors.dx5
```

TeachingSamples.dx5
SeqPhrases
DNA-binding.sx5
HomingEndo.sx5

- **Notes for shared DataBook database files**

  DataBook files contain links to the complete molecule, primer, or alignment data files so that the files can be opened for use. You might want to consider using the following procedure to set up DataBook database files to share:

  Place a copy of molecule files, primer files, or alignment results files in folders on your server or other accessible location. You might want to prevent others from overwriting the files in this location.

  Use the Multiple File Converter module to create DataBook files using the easy batch procedure. These DataBook files should all have file paths that will allow the files to be opened from computers within the group.

  Copy the DataBook files (*.dx5) you just created and place them in the shared location, as described above. The DataBook files in this location will be read-only.

  If you want to change any of the data in the shared DataBook files, you can do so at the computer where you created the DataBook files (since they will be part of your personal DataBooks at this location). Then copy the modified DataBook files to the shared location. If you need to change the path (location) of many of the molecule, primer, or data files, you can use Tools, DataBook Tools to do a find and replace operation for the file path.

- **Repository Information**

  You can set up a shared location to be used as a repository to be used for submitting important molecules.

File Menu

The *File* main menu contains options for opening, saving, and exporting Clone Manager files. Also included are Preferences and License.

- New: Create a new molecule.
- Open: load a Clone Manager or molecule file from disk.
- Save and SaveAs: save changes to the active data.
- Clone Share

  Access from the main File menu using the 'Clone Share' menu item
  - Open – load a data file.
  - Save – save a data file.
- Export
- Print
- Other Tools
  - Import from Clipboard
  - Retrieve from Entrez eUtils
  - Multiple File Conversion
  - Print Preview – shows a preview of what will be printed.
  - Print Setup – configure your printer.
- Preferences
- Help **–** access help topics.
- License Clone Manager
- Recent Files – allows quick selection of files that you recently opened.
- Exit – close the program.

Opening Files

Selecting the File, Open menu will open the windows dialog box where you can select the file that you want to open. Clone Manager will load the file and determine the type of information contained. More information is available for each type of data:

- Molecule files – loads Clone Manager molecules and imports files in standard formats.
- Primer files – loads primer sequence and any molecule association.
- Alignment setup files - loads previously saved alignment setup configuration and any results.

Saving Files

Saving molecules, or other molecular biology data, enables the data to be used in a later session.

When saving the results of an analysis, you will often have the option to enter information that describes the data. For example, if saving the results of an alignment, you can enter a name, description, and notes. Such information will be displayed in printed reports.

- **Save**

The File, Save menu option is enabled when the active window has been modified or not yet been saved. If the data had previously been saved, this option will automatically overwrite the existing data file.

- **SaveAs**

Use the File, *SaveAs* menu option if you want to save to a different location.

Export Data

Found on the File menu. Use this option to export data that is in the active window at the time you select the Export option.

You can export the complete set of data in the view or capture a pixel image of just the data shown in the window.

You can also export basic molecule, primer, or alignment data, depending on the contents of the active window. You can export the data to a disk file or copy it to the Windows clipboard for later use in another program.

- **Header -** reports the source of the data, whether it is text or graphics, and options for exporting the primary data.
- **Data in View**

  Select this option to export all results information in the view. The complete set of information is automatically selected and then sent to a disk file or copied to the clipboard. If the active window shows a molecule map, the high-resolution (printed format) graphic map will be exported, using the styles and colors set for printing.
- **Window Snapshot**

  Select this option to do a screen capture of just the data that is currently shown in the active window, exactly as it appears. You can send this pixel image to a disk file or copy it to the Windows clipboard. Files are saved in the Portable Network Graphics (*.png) format, particularly useful for computer screen displays such as on a web page or in an on-screen presentation.
- **Primary Data (where applicable)**

  Primary data depends on the nature of the active window.

  - Molecule – select the output format to use. Options include sequence only, FASTA, Compact GenBank (essential information only) and Commented GenBank (including extra Clone Manager-specific information about the molecule).
  - Primer – export the primer sequence as plain text for FASTA format.
  - Alignment – export consensus sequence, merge/splice or aligned data.
  - Sequence Assembly – export the consensus sequence.

- **Converted Map Data (where applicable)**

  If the active window shows an enhanced view of the graphic map, you can convert the map data into raster images for use in other Windows programs that require a raster file format. You can export in TIF, PNG, JPG, GIF or BMP file formats.

- **Preparing for screen capture**

  Before you do a Window Snapshot, you may want to optimize the appearance of the screen display. You can resize windows, scroll to an area of interest, or change screen colors for some specialized displays. For example, you can use the Format button in the Alignment Viewer Window to change the color behind aligned sequences, or use File, Preferences, Colors to darken the lines in the Restriction Map, Map of Sites display.

Importing Molecules from the Clipboard

You can read molecule information directly from the Windows clipboard if this information is in one of the standard formats that the program could read from a disk file (like a standard GenBank format). This could be very helpful if you get files using your e-mail system.

- **To read in molecule information:**

1. In your e-mail document or other Windows workspace, select (highlight) the molecule data.
2. Click Edit, Copy to place the molecule information on the Clipboard.
3. In the Clone Manager program, click File, Other Tools, Import from Clipboard.

If your e-mail document contains more than one molecule file, the program will read in the information for each molecule in turn and place this collection of molecules on the Molecule List as a group. Select Ungroup to see the individual molecules.

License Clone Manager

This option provides information on the current license state and provides the methods for managing your license.

The dialog box reports the following information:

- Version of Clone Manager
- License type
    - Trial Evaluation – reports how many days are remaining
    - Activated – reports the License ID used to activate this installation. Also reports if the license is registered.
    - Concurrent – identifies that you are part of a concurrent license.
    - Transferred – the current computer is not licensed as the license was transferred to another computer.
- Registration– shows the registered user's name and company. This is optional but recommended.

Manage License

Managing a license depends on the current license type:

- Trial Evaluation
    - Continue the trial evaluation.
    - Activate your license on this computer.
    - Join a concurrent group.
- Activated
    - Remove your license - allows you to transfer the license to a new computer.
    - Update your license – our technical support may ask you to select this option to refresh your license.
- Concurrent
    - Change your concurrent license settings – use to select a different license server or if your network administrator moves the licensing service to a different server.
    - Leave your concurrent license group – stop using concurrent licensing. You would do this if you were changing to a seat license.

- Transferred
  - Activate your license on this computer.
  - Join a concurrent group.

View Menu

The *View* main menu contains options for changing the selected data view. Also includes access to the lists of loaded data file.

- Send view to Clipboard
- Send view to File
- Molecule List
- Primer List
- Enzyme Lists
- Sequence Phrases
- DataBook
- Add to
  - DataBook
  - Primer Collection
  - Clone Share
  - Submit to Repository
- Molecule Viewer – select the tab to view molecule details.
- Primer Viewer – select the tab to view primer details.
- Alignment Viewer - select the tab to view alignment details.
- Other Windows – select the window to make visible.

Clone Menu

The *Clone* main menu contains options for classic cloning operations and cloning wizards.

- Classic Cloning Operations
  - Find Enzyme Sites
  - Cut
  - Modify Ends
  - Ligate
  - Join Sequences
- Cloning Wizards
  - Plan Cloning
  - Gateway
  - Topoisomerase
  - Assembly Cloning: Gibson, In-Fusion, NEBuilder
  - Golden Gate
  - CRISPR Design
  - Ligation Independent
  - PCR Cloning

Primer Menu

The *Primer* main menu contains options for designing, analyzing, and using primers.

- Design
- Direct Entry
- PCR Cloning wizard
- Sequencing wizard
- LCR oligo set
- RT-PCR
- Analyze
- Analyze Mix wizard
- Create Product
- Default Primer Criteria
- Default Weights and Values

Align Menu

The *Align* main menu contains options for searching for sequence similarities.

- Align Multiple Sequences
- Compare Two Sequences
- Scan for similarity
- Sequence Assembly
- BLAST search
- Parameters and Settings
- Align Idea Wizard

Align Idea Wizard

This wizard will try to give you some ideas about using the alignment methods effectively.

- Looking at a new sequence.
- Showing aspects of relatedness.
- Checking for vector contamination.
- Organizing alignment setup files.

Discover Menu

The *Discover* main menu contains options for discovering commonalities.

- Find 0in Files
- Find Enzyme sites (AutoScan)
- Find Open Reading Frames
- Find Sequence Phrases
- Sc0n for Molecules with Sequence Phrase
- Find Primers that Bind
- Scan for Molecules with Primer sites
- BLAST primer

Find in Files

Found on the Discover menu or use main toolbar button to access. You can use this function to scan molecule files to look for certain genes or specific text in the molecule name or description. For Clone Manager files (*.cm5), you can also look for specific text in the Notes field.

To find text in a molecule file:

1. Click the Find in Files button on the main toolbar or select the option on the Discover main menu.
2. Enter the feature name or the text you want the program to search for.
3. Identify the file or folders of files that should be searched (see below).
4. Use the checkboxes to identify the locations in each file where the program should look.

Selecting molecule files or folders to search

Click the Add Folder button to select a folder location to add to the search list.

Navigate to the folder you want to add to the list and click OK. The program will attempt to open and read all molecule files in the location specified, including those contained in subfolders.

Click the Add File button to select a molecule file to add to the list.

Navigate to the molecule file you want to add to the list and click OK. The program will attempt to open and read this file during the search procedure.

If you routinely use the same molecule file locations, you can save the list of files and folders and recall this list later.

Click **OK** to begin the Find in Files operation. The program will open each molecule file and search all the fields specified, looking for a perfect match for the characters you typed in. Upper case or lower-case letters will not affect the search.

Save search lists

Select the 'Save List' toolbar button to save the list of files and folders. You can use the 'Open List' toolbar button to load the list of files and folders at a later time.

Search Results

When a search for molecule sites for the specified primer has been completed, the results will be presented in a scrolling list box.

In the results display, the molecule name is listed, as well as the base pair position of the 5' end of the binding site. Also reported are the length of the binding site (LenB) and the length of the primer (LenP). If there are mismatches within the binding site, the number

of mismatched bases is shown in the column *m*. Click on the column headers to sort by molecule name, position, or length of the binding site.

Click on column headers to sort the list by molecule name, file name, date, or description. Drag on the join between columns to resize. Double-click on a join to make the column size fit the data and then scroll to view longer molecule descriptions, if needed.

You can print or copy the results list or load a molecule file for use.

Tools Menu

The *Tools* main menu contains options for.

- Process Molecule: Invert, Rotate, Circularize, Translate
- Analyze Molecule
- DataBook Tools
- Author Stamp
- Mutagenesis Profile
- Primer Values Profile
- Link Primer to Molecule
- Export
    - Export Molecule Data
    - Export Primer Sequence
    - Export Alignment Data

Process Molecule

Found on the Tools menu. You can invert the active DNA molecule or rotate a circular DNA molecule. If a DNA molecule was brought into the program as a linear molecule (perhaps because the "circular" descriptor was missing), you can circularize the molecule. You can also translate all or a part of a DNA molecule.

When the process has been completed, you will want to enter the new molecule to the Molecule List, changing the name or description, as needed. The inverted, rotated, circularized, or translated molecule exists only within the current program and should be saved to disk if needed for further use.

- **Invert Molecule**

  Use this option to invert a copy of the active molecule, changing the positions of all sites and features. The following information will appear as default entries: Molecule name: i-(original molecule name); Description: inverted (original molecule name).

- **Rotate Molecule**

  Use this option to rotate a circular molecule to bring a different point to the top of the map. The program will rotate a copy of the molecule, changing the positions of all sites and features. To rotate a molecule, enter the base pair position which should appear at the top of the map (the new origin of the molecule). The following information will appear as default entries: Molecule name: r-(original molecule name); Description: (original molecule name) rotated to position (n).

- **Circularize Molecule**

  Use this option to blunt-end ligate a linear molecule to make it circular. (If the molecule has cohesive ends, the upper strand of sequence will be treated as if it were blunt ended.)

- **Translate Molecule**

  Use this option to translate the active DNA molecule to create a protein. Translation will proceed from the starting point specified, on the strand specified, and continue until a stop codon or the end of the molecule is encountered. Translation will use the Standard Code (translation table = 1) unless an alternate translation table is assigned for this molecule. To translate a molecule, enter the starting base pair position for translation. Use the check box to indicate if you want to translate the complement strand. The following information will appear as default entries: Molecule name: protein; Description: translation of (original molecule name) starting at bp (n).

Appendix

The following links provide more information on the listed topics:

- Installing Clone Manager
- Molecule File Formats
- Import Features Preferences
- Groups of Molecule Files
- Common File Loading Errors
- REBASE File Updates
- Translated Display
- Send View to File
- Send View to Clipboard
- Using Search Lists
    - Alignment Scan
    - Find in files
- Basic Preferences
- Working with Map Sites

Installing Clone Manager

The following instructions show how to install the Clone Manager client program on your computer.

If you have not yet purchased a license for Clone Manager, you can use a free trial license which is valid for 30 days.

When you purchase a license for Clone Manager you can activate your license to remove the time limitation of the trial license. Licenses can be either seat (for your use on your computer) or concurrent (install a license manager on your network server for shared access).

- Seat License

Personal license to use Clone Manager on your computer. If your computer is portable, you can use Clone Manager wherever you are.

- Concurrent License

Shared use on your local network. Your network administrator will activate a license management service on your network server. When you start the Clone Manager client program, it will contact the license management service and request a user token. If a token is available, the client program will start, and you can use all the functions of Clone Manager. If all tokens are in use, you will get a message to that effect.

Concurrent licenses work very well for large groups who are willing to share access to Clone Manager. However, you must always remain connected to the license management service. If you need to use Clone Manager off-site, you will need a seat license. (Or you can ask your network administrator to set up a VPN to permit remote access to your internal network.)

- Installing Client Program
- Client computer requirements
- Activate seat license
- Transfer a seat license to a new computer

Installing Client Program

The Clone Manager client program runs on your personal computer and provides the functions described in this help document for handling, viewing, and analyzing molecular biology data.

These instructions show how to install the client program on your personal computer.

- Download the client installer

Open your web browser to our website https://scied.com and select the downloads section. This will open to a page where you can download the current version.

Click the button CMxxSetup.msi to download the latest installer. Other links allow downloading instructions and other helpful information.

If you have an older version, you can click a link to download the version you have.

- Installing Clone Manager

  Run the installer on your computer (administrator privileges required).

  - The installation program will copy the program files to the specified installation directory on your computer, install a menu item for the program on your Start menu, and install a program short-cut on your desktop.
  - File associations are configured to allow you to open a data file in Clone Manager by double-clicking the file in Windows Explorer.
  - User documentation is installed in the Docs folder under the program installation folder.

- Run the client program

  Double click the desktop, or start menu, shortcut.

  The first time you run the program, you will need to select your license type.

  - Start a 30-day trial evaluation.

    You can use all the functions of Clone Manager to evaluate all functions using your own molecules. The trial is fully functional but will expire in 30 days.

  - Activate your license on this computer.

    Use this option if you already have a seat license. On the next page, you will enter your license keys.

  - Join a concurrent license group.

    Use this option if your organization has a concurrent license. In most cases, joining a concurrent license does not need any special configuration settings and you can just complete the licensing wizard. If your network requires special configurations, your network administrator will tell you what to enter on the next page of the wizard.

  Home directory: The first time your run Clone Manager, a home directory will be created for you. The home directory is the location where the program will look for personal configuration data and user-specific files and folders. In Clone Manager, you can look in File, Preferences to identify this location and change, if needed.

  If you have been using an earlier version of Clone Manager, the new installation will use the same home directory (CMHome), so that you will have immediate access to your primer collections, DataBook database files, special enzyme lists, GenBank feature import settings and other custom settings.

Client computer requirements

- Recommended: Windows 10 or 11.
- 64-bit Windows (x64)

N.B. Clone Manager may run on older versions of x64 Windows but are not supported as Microsoft no longer supports these older operating systems. If you have an older 32-bit (x86) version of Windows, please contact us for downgrade rights so you can install Clone Manager 11 which supports older versions of Windows.

Activate seat license

Activating a seat license is used to remove the time limitation of the trial evaluation license.

Run the Clone Manager program and select the main menu Fil, License Clone Manager option. Click the 'Manage' button and select the option to 'Activate your license'. On the next screen, enter your license ID and activation password. We also recommend entering an 'Installation Name' which indicates where the license was installed. This information is available to those managing licenses and can assist in locating where a seat license was installed.

Once activated, you can use Clone Manager on this computer for as long as it is useful to you. At intervals, the license manager will automatically update your license to keep it current.

Transfer a seat license to a new computer

When you get a new computer, you can move your Clone Manager seat license to the new computer.

Please refer to the FAQs https://scied.com/faqs.htm#MoveLicenseSeat

You will need the license number and password that you received in the email confirmation when you purchased your license for Clone Manager. On your old computer, run Clone Manager and click File, License Clone Manager. Click Manage and select the option to transfer your license to a new computer. Confirm this action and note that your license has been successfully removed. Close Clone Manager and uninstall if desired.

On your new computer, download and install Clone Manager. Run Clone Manager and select the option to activate your license. If you are already running the trial license on your new computer, click File, License Clone Manager. Click Manage to select the option to Activate and enter your license details.

You can move your license up to three times. (If you need to frequently move your license, you may want to consider an alternate license type, such as a network concurrent license.)

Molecule File Formats

Molecules in the Clone Manager molecule file formats (*.cm5, *.cx5) contain sequence data, molecule name and size, and may contain any of the following additional information: restriction enzyme sites, features, description, notes. This format should not be edited with a word processor.

Sequences in other standard file formats are imported as double-stranded and blunt-ended. (If the sequence has been identified as a linear molecule and you want it to be circular, you can use the Circularize option during file loading.) Only one strand of a sequence is needed and IUPAC codes for ambiguous bases can be used. A number of sequence databanks and analysis programs will prepare files in a format that is similar (but not always identical) to GenBank or EMBL formats.

GenBank File Format

Sequence and header information. The program looks for "LOCUS" at the start of line 1, marking the beginning of the header information, and "ORIGIN" marking the end of the header and the beginning of the sequence. All information between Locus and Origin can be read as Annotations in the program.

If a features table is present in the file, by default features labeled as "CDS" or "exon" will appear in the molecule features table as genes. You can modify which feature keys are imported.

EMBL File Format

Sequence and header information. The program looks for "ID" plus 3 spaces at the start of line 1, marking the beginning of the header information, and "SQ" plus 3 spaces, marking the end of the header and the beginning of the sequence. All information between ID and SQ can be read as Annotations in the program.

If a features table is present in the file, by default features labeled as "CDS" or "exon" will appear in the molecule features table as genes. You can modify which feature keys are imported.

FASTA File Format

Sequence and a definition line. The program looks for ">", marking the beginning of the file and the definition line. The following lines contain sequence data, usually with 60 characters per line.

DOS Text File Format

Sequence data not in GenBank, EMBL or FASTA formats. Sequences can be single- or double-spaced and can be typed in upper- or lower-case letters. Base pair numbers will be ignored. Spaces between letters (blocks of bases) will not affect the program's ability

to read the sequence. Non-sequence lines (comments) must start with a semi-colon or delimiters that mimic the GenBank or EMBL formats (ID and SQ, for example).

Groups of Molecule Files

When you load 5 or more molecule files at the same time, you can indicate that these molecules should be kept together as a group for easier handling. In the molecule list display, you can temporarily ungroup these files and then regroup them as needed:

- **Ungroup**

  Breaks group structuring and lists all molecules as individual entities. You might want to ungroup to select a single molecule for a specific operation or to open the molecule viewer window to view this molecule's sequence or features table.

- **Regroup**

  Restructures groups previously ungrouped (see above). You might want to regroup to select all the molecules in the group for an alignment procedure or to remove all these molecules from the program in one step.

Common File Loading Errors

There are several common file loading errors. Your file can contain unrecognized characters, the program may be unable to find the start or end of the sequence, or you can have invalid base positions for one of the features in the feature table. Each type of error is described, and suggestions are given for resolving the problems encountered.

Unrecognized Character in File

This error message will appear if the program encounters: any character other than A, C, G, T, or a valid IUPAC ambiguous base when trying to load a DNA sequence file; or any character other than a valid amino acid code when trying to load a protein sequence file.

This error will also be triggered by standard text that is not within the confines of a databank header (between "ORIGIN" and "LOCUS" in a GenBank file or between "ID" and "SQ" in an EMBL file). Non-sequence text in an ASCII file can also cause this error if the text is not commented (comment lines begin with a semi-colon).

Resolution: Repair the file by removing the invalid character(s) or adding semi-colons to comment lines in an ASCII file and try to reload.

Could not Find Start of Sequence

This error can be triggered if a sequence is present, but the program cannot find the end-of-header identifier in a GenBank or EMBL file ("ORIGIN" or "SQ", respectively, missing from file).

Resolution: Repair the file by adding the missing identifier in its correct location and try to load it again. Or split the header from the sequence by marking the beginning of the sequence with the cursor and select File Split.

File Too Large

This file is too large to fit into the File Loading Help System. It may not be too large to be loaded into the main application, if you can identify the error and modify the file in your word processor. This error can also be triggered when the program is unable to find the end of the sequence.

Resolution: Look for invalid bases, comment lines without semi-colons, or missing header identifiers, as discussed in the other file loading error descriptions. Remember to save as a DOS text file (ASCII text) and try to load again.

Invalid Base Position in Features Table

One of the features in the feature table of a databank file has an invalid base pair position.

Resolution: You can correct the features table entry, if you know the correct base pair position, or you can instruct the program to skip this feature.

REBASE File Updates

You can update the Clone Manager master enzyme file monthly using REBASE enzyme files so that you are always up to date with new enzyme information.

You can get the latest REBASE data files directly from the REBASE web site, or you can subscribe to the email service. In either case, you will have to identify the disk file and its location during an update procedure.

Alternatively, you can let the program get the file for you over the Internet using anonymous ftp. This file is accessed directly during the update procedure.

Each time you do an update, the program-defined enzyme lists (Commercial (Main), All Enzymes and Scanner) will be regenerated. The update procedure will not discard your user-defined enzyme lists or any special enzymes that you may have entered.

Citation: Roberts, R.J., Vincze, T., Posfai, J., Macelis, D., REBASE: a database for DNA restriction and modification: enzymes, genes, and genomes. Nucleic Acids Res. 51: D629-D630 (2023).

Update REBASE

1. Open the Enzyme Lists dialog using the View, Enzyme Lists menu.
2. Click the toolbar button for REBASE Update.
3. Select the source:
    a. Have a file – select the file that you downloaded from REBASE
    b. Get the file now – uses FTP to directly download the latest file. Update the URL if needed and click the 'Configure' button if you need to change your internet settings.
4. Click 'Update'
5. Clone Manager will locate and process the REBASE data file and report the changes found. Once the enzyme list update is complete, you must exit the program and then re-start it to use the new enzyme lists.

REBASE Files on the Web

To get the latest REBASE data file:

1. Connect to the REBASE web site: http://rebase.neb.com
2. Click on Data Files.
3. Click on file format #35, Clone Manager.
4. Click **Get File**.

Save the data file to your computer. When you do an update procedure, you will need to identify this file.

Subscribing to REBASE

To subscribe to the REBASE email service:

1. Connect to the REBASE web site: http://rebase.neb.com
2. Click on the **Subscribe**.
3. Complete the required contact information
4. Select data format #35, Clone Manager.

When files are received in your email system, you will want to save these files to your hard disk as text files so that you can use them for the update procedure. When you do an update procedure, you will need to identify this file.

Getting REBASE Files via FTP

Set the option Get the File Now in the REBASE Update dialog box during an update procedure. The default settings are:

ftp address	ftp.neb.com
Directory	/pub/rebase
File	clonemgr.*

You can modify these address components, if necessary. The REBASE data file will be used during the update and discarded. It is not saved to your hard disk.

Note: FTP access uses TCP/IP internet access protocols. Please consult with your network administrator or internet service provider to ensure you have the required setup.

Restoring a prior Master Enzyme List

A backup copy of your master enzyme file is made by the program any time you begin an update procedure. If, for any reason, it is necessary to go back to your prior version of the master enzyme file, follow these steps: Using the Windows File Explorer open the folder for your Home Directory. Rename your current master enzyme file from SEEnz32.enz to another file name. You can also delete this file if you suspect it contains errors. Rename your backup master enzyme file from SEEnz32.bak to SEEnz32.enz.

Translated Display

The translated display for sequence views uses amino acid codes consisting of 1 letter or 3 letters. Both these codes are right aligned below the rightmost base of the codon (even if the codon is read right-to-left for complementary strands):

1 letter code

```
atg aaa
  |   |
  M   K
```

3 letter code

```
atg aaa
  |   |
Met Lys
```

1 Letter Amino Acid Codes

A	Ala	M	Met
C	Cys	N	Asn
D	Asp	P	Pro
E	Glu	Q	Gln
F	Phe	R	Arg
G	Gly	S	Ser
H	His	T	Thr
I	Ile	V	Val
K	Lys	W	Trp
L	Leu	Y	Tyr

Genetic Codes

1 Standard Code

2 Vertebrate Mitochondrial Code

3 Yeast Mitochondrial Code

4 Mold Mitochondrial; Protozoan Mitochondrial; Coelenterate Mitochondrial; Mycoplasma; Spiroplasma Code

5 Invertebrate Mitochondrial Code

6 Ciliate, Dasycladacean and Hexamita Nuclear Code

9 Echinoderm and Flatworm Mitochondrial Code

10 Euplotid Nuclear Code

11 Bacterial, Archaeal and Plant Plastid Code

12 Alternative Yeast Nuclear Code

13 Ascidian Mitochondrial Code

14 Alternative Flatworm Mitochondrial Code

16 Chlorophycean Mitochondrial Code

21 Trematode Mitochondrial Code

22 Scenedesmus obliquus Mitochondrial Code

23 Thraustochytrium mitochondrial Code

24 Pterobranchia Mitochondrial Code

25 Candidate Division SR1 and Gracilibacteria Code

26 Pachysolen tannophilus Nuclear Code

27 Karyorelict Nuclear Code

28 Condylostoma Nuclear Code

29 Mesodinium Nuclear Code

30 Peritrich Nuclear Code

31 Blastocrithidia Nuclear Code

33 Cephalodiscidae Mitochondrial UAA-Tyr Code

Export Data

There are multiple options for exporting data depending on the type of data.

- File menu, Export provides many methods of exporting the data in the view.
- View menu, provides sending the current view to the clipboard or a file:
  - Send view to Clipboard
  - Send view to File
- Tools menu, Export the current data
- Analyses – many methods will have an option to export the data including the option to use the view menu sent options.

Send View to File

Found on the View menu or use toolbar to access. Also found as an option in the Export Data function (found on the File menu).

You can export the data from a results window and save this information in a file on disk. When you select this option, the information in the view window is automatically selected and saved to disk in the file format specified. Click Send View to File or in the Export dialog box, select Data in View, then Export to a Disk File.

To place this information in a document in your word processing program, move to the appropriate place in the document and insert the contents of the disk file into your document. You can also copy this view information to the Windows clipboard.

Different formats are used to export text data and graphic data to the clipboard. Follow these links for more specific information and some hints about using this data.

If you are exporting formatted sequence data, the program will use the format on the screen, including the number of bases per line. If you are working with a very wide window to view sequence data, you will probably want to resize the window so that you see about 50 or 60 bases per line before exporting the data to file. This will avoid line wrapping problems when the data is opened or inserted into documents in your other application.

Send View to Clipboard

Found on the View menu or use toolbar to access. Also found as an option in the Export Data function (found on the File menu).

You can copy data from a results window and place this information on the clipboard for use in another Windows program. When you click the Send View to Clipboard button, the information in the view window is automatically selected and then copied to the clipboard. Or in the Export dialog box, select Data in View, then Copy to the Clipboard.

To place this information in a document in your word processing program, move to the appropriate place in the document and paste the contents of the clipboard into your document (usually Edit, Paste or Edit, Paste Special). You can also capture this information in a file for later use (see below).

Different formats are used to copy text data and graphic data to the clipboard. Follow these links for more specific information and some hints about using this data.

- **Sequence data and window size**

  If you are copying formatted sequence data, the program will copy the format on the screen, including the number of bases per line. If you are working with a very wide window to view sequence data, you will probably want to resize the window so that you see about 50 or 60 bases per line before copying to the clipboard. This will avoid line wrapping problems when the data is pasted into your other application.

Copying Text Data to the Clipboard

Text information is placed on the clipboard in HTML format (htm), Rich Text Format (rtf), and plain text format (txt). Files in rtf and htm formats will include font information (font face Consolas or Courier New, point size 10 pt). Files in htm format also include background color information, when appropriate.

You can paste the contents of the clipboard into your application using Edit, Paste or Paste Special. When you use Edit, Paste (or the toolbar Paste button), the application will select the format it wants from the clipboard. You can often use Edit, Paste Special and then select the format you want to override the default selection, if needed.

After you have pasted text data into your application, you may find the page is too narrow for the data you have pasted, causing the lines to wrap. Sometimes the text on the screen looks okay but the print preview shows line wrapping. To fit the text on the page, you can change the page margins to make the page wider (File, Page Setup) or you can make the point size of the text smaller.

You can change the font or point size after you have pasted the data into your application but remember to select a fixed pitch or monospaced font (so that sequences will line up correctly).

Copying Graphic Data to the Clipboard

Graphic information is placed on the clipboard in Enhanced Metafile format (emf) and Windows Metafile Format (wmf). The Windows clipboard then prepares a generic, picture version of these graphic images.

You can paste the contents of the clipboard into your application using Edit, Paste or Paste Special. When you use Edit, Paste (or the toolbar Paste button), the application will select the format it wants from the clipboard. You can sometimes use Edit, Paste Special and then select the format you want to override the default selection, if needed. In general, Enhanced Metafile format images will give the best results.

When resizing images in your applications, you will want to maintain the original aspect ratio. In many applications, holding down the Shift key while you are dragging out an image will maintain this relationship.

Some older word processing and presentation graphic programs do not fully support the Enhanced Metafile format. Some use filters to convert these files to the older format, losing some information about scaling and image placement. If you are having difficulty pasting graphic data into your application, you may want to try sending this data to a disk file instead (send view to file). Try selecting the older *.wmf file format, which may work better with your older application.

Sending Text Data to a File

Text information can be saved to a disk file in several formats:

- **HTML (*.htm)**

  Default format for import into most word processors. Includes font information (sequence data requires a monospaced font so that bases line up correctly) and background and foreground text color information, when appropriate.

- **Rich Text Format (*.rtf)**

  Format previously used by most word processors. Includes font information and foreground text color information.

- **Plain text (*.txt)**

  Can be read by even the simplest utilities, but does not contain font face, point size, or color information.

Sending Graphic Data to a File

Graphic information can be saved to a disk file in two formats.

- **Enhanced Metafile (*.emf)**

  Default format for current Windows programs. Has improved scaling, image placement and color information.

- **Windows Metafile (*.wmf)**

  Format previously used by all Windows programs.

  If you have had trouble using the clipboard to copy and paste graphic data into your word processing program, you may want to try saving it to a file in the older *.wmf format.

Multisite Gateway® Cloning Components

Many researchers are currently using kits from Invitrogen Corp to perform 2, 3 and 4-fragment recombination. The following components and sequences are used to simulate the preparation of recombinant molecules using Multisite Gateway® Pro kits.

2-Fragment Recombination

Element	Insert Flanking sites	Donor Vector	Entry Clone sites
Fragment 1	attB1, attB5r	pDONR221 P1-P5r	attL1, attR5
Fragment 2	attB5, attB2	pDONR221 P5-P2	attL5, attL2

3-Fragment Recombination

Element	Insert Flanking sites	Donor Vector	Entry Clone sites
Fragment 1	attB1, attB4	pDONR221 P1-P4	attL1, attL4
Fragment 2	attB4r, attB3r	pDONR221 P4r-P3r	attR4, attR3
Fragment 3	attB3, attB2	pDONR221 P3-P2	attL3, attL2

4-Fragment Recombination

Element	Insert Flanking sites	Donor Vector	Entry Clone sites
Fragment 1	attB1, attB5r	pDONR221 P1-P5r	attL1, attR5
Fragment 2	attB5, attB4	pDONR221 P5-P4	attL5, attL4
Fragment 3	attB4r, attB3r	pDONR221 P4r-P3r	attR4, attR3
Fragment 4	attB3, attB2	pDONR221 P3-P2	attL3, attL2

Destination Vectors

For 2, 3 and 4-fragment recombination, any destination vector containing attR1 and attR2 sites can be used.

Primer Sequence att Sites (5'-3')

attB1	GGGG	ACA AGT TTG TAC AAA AAA GCA GGC T
attB2	GGGG	AC CAC TTT GTA CAA GAA AGC TGG GTA
attB3	GGGG	ACA ACT TTG TAT AAT AAA GTT G
attB3r	GGGG	AC AAC TTT ATT ATA CAA AGT TGT
attB4	GGGG	AC AAC TTT GTA TAG AAA AGT TGG GTG
attB4r	GGGG	ACA ACT TTT CTA TAC AAA GTT G

attB5	GGGG	ACA ACT TTG TAT ACA AAA GTT G
attB5r	GGGG	AC AAC TTT TGT ATA CAA AGT TGT

Primer Thermodynamics and Secondary structure

Determination of Tm

Melting temperature is calculated using the method based on nearest-neighbor thermodynamic parameters. The formula for calculating Tm uses the expression from Freier et al (Proc Natl Acad Sci USA 1986, 83, 9373-9377) for non-self-complementary sequences where one component (the primer) is in molar excess.

Thermodynamics

Thermodynamic parameters for enthalpy and entropy are taken from SantaLucia (Proc Natl Acad Sci USA 1998, 95, 1460-1465) or Breslauer et al (Proc Natl Acad Sci USA 1986, 83, 3746-3750). Salt correction uses the values from Wetmur (Crit Rev Biochem & Mol Biol 1991 26, 227-259) which allow correction for both potassium (or sodium) and magnesium concentrations. The concentrations of salt, magnesium, and oligonucleotide are user editable. Separate values are maintained for each type of design: PCR, probe, sequencing, and LCR. The destabilization effect of mismatched bases is estimated using the value for RNA mismatch loops from Turner et al (Cold Spring Harbor Symp. Quant. Biol. 1987, 52, 123-133).

Determination of Optimal Annealing Temperature

PCR optimal annealing temperatures are calculated based on the method of Rychlik et al (Nucleic Acids Research 1990, 18, 6409-6412). Product-template melting temperature is calculated based on the method of Baldino et al (Methods in Enzymol 1989, 168, 761-777).

Determination of Secondary Structure

Secondary structure is determined by identifying dyad symmetries that are energetically stable at the annealing temperature. Thermodynamic stability is determined by the nearest-neighbor method. In the absence of data for DNA hairpin loop formation, we have used the energy parameters from Freier et al (Proc Natl Acad Sci USA 1986, 83, 9373-9377) which refer to RNA:RNA secondary structure. The program uses the annealing temperature to determine whether secondary structures will be stable under experimental conditions. The program's indication of secondary structure therefore should be interpreted as showing the possibility of structure. In general, Primer Designer will show more secondary structure than is likely under normal PCR conditions.

Primer Criteria Settings

These settings make up the values referred to as Primer Criteria. You can change these settings for individual primers or specific searches, or you can change the program default settings, if desired.

- **GC Content**

GC % Range for primers: Preset values (PCR): min 50%, max 60%. Enter the minimum and maximum acceptable percent GC.

Match pairs for GC: Preset value: match GC of primers in pairs to within 5% GC. Enter the percent value to match within. Leave entry blank to disable.

- **Melting Temperature (Tm)**

Tm ^{o}C Range for primers: Preset values (PCR): min 55°C, max 80°C. Enter the minimum and maximum acceptable melting temperature in centigrade degrees.

Match pairs for Tm: Preset value: match Tm of primers in pair to within 10°C. Enter the Tm value to match within. Leave entry blank to disable.

- **Annealing Temperature**

Preset value (PCR): 55ºC. Enter the annealing temperature that you typically use in your PCR reactions. Along with stability, controls formation of hairpin loops and false priming.

- **Stability 5' vs 3'**

Preset value: 1.2 kcals (SantaLucia energy parameters) or 2.0 kcals (Breslauer energy parameters). Enter the value (in kcals) by which the 5' end pentamer stability must exceed the 3' end pentamer stability. Not applicable for probes.

- **Dimers**

Preset values (PCR): Less than 3 matching bases at 3' end, less than 7 adjacent homologous bases. Enter the number of matches at the 3' end or the number of adjacent homologous bases which would cause primer rejection. 3' end dimers not applicable for probes.

- **Runs**

Preset values: Less than 4 base runs (any bases), less than 3 dinucleotide pair repeats. Enter the number of bases (any) or the number of dinucleotide pairs which would cause primer rejection.

- **Protein degeneracy**

  Preset value: less than 8 within 6 bp of 3' end. Enter the degeneracy value for the 6 bases at the 3' end which would cause primer rejection. Not applicable for probes.

- **GC Clamp**

  Preset value: Require 1 G or C at the 3' end of the primer. Enter the minimum number of G's or C's required at the 3' end. Enter 0 to disable. Not applicable for probes.

- **Oligonucleotide and Salt concentrations**
- **Criteria for sequencing primers or probes**

Change Type of Criteria

It is possible to change the type of criteria used to evaluate an existing primer (designed using an alternate criteria type). For example, you may want to change between PCR and sequencing primer criteria.

To change the criteria type for a primer:

1. Open the primer viewer window (select the primer from the primer list, if necessary)
2. Click the Edit tab.
3. Click the Criteria button on the window toolbar
4. Click the small redefine button (near bottom, center) on the criteria page
5. Select the criteria type you want to use to evaluate this primer.

Primer Design Experimental Conditions

You can select the energy parameters you want the program to use to calculate Tm. A new unified set of nearest neighbor parameters (SantaLucia 1998) can be selected to replace the more classic nearest neighbor parameters of Breslauer (1986). When the energy model has been changed, the program will automatically reset the stability value to work with the energy parameters selected.

You can also set the concentrations of oligonucleotides or salts you use in your experimental conditions. These concentrations are used in Tm calculations.

- **To change energy parameters or concentration values:**

1. Click Primer, Default Primer Criteria
2. Click the Conditions tab
3. Select the energy parameters model
4. Enter the concentration values you want the program to use.

The following values are preset by the program.

- **PCR Primers**

Oligonucleotide 100.0 nanomolar

Salt (K/Na) 50.0 millimolar

Free magnesium 1.5 millimolar

- **Sequencing Primers**

Oligonucleotide 20.0 nanomolar

Salt (K/Na) 50.0 millimolar

Free magnesium 1.5 millimolar

- **Probes**

Oligonucleotide 0.4 nanomolar

Salt (K/Na) 1000.0 millimolar

Free magnesium 0.0 millimolar

- **LCR Oligos**

Oligonucleotide 4.0 nanomolar

Salt (K/Na) 100.0 millimolar

Free magnesium 0.0 millimolar

- **Notes:**

  These concentrations are used to calculate Tm (melting temperature). Higher oligonucleotide concentrations result in higher Tm values. We have set the default concentration for PCR to be the typical concentration of primer added to the reaction and, therefore, the Tm predicts the melting temperature for the critical first round of amplification. Some prefer to use sub-nanomolar concentrations to derive an "average" Tm throughout the reaction.

  The effect of magnesium on Tm has been discussed by Wetmur (Crit Rev Biochem & Mol Biol 1991 26, 227-259). Note that deoxynucleotides in the reaction will chelate magnesium and hence reduce the free magnesium concentration.

Primer Sorting Weights

The program will look at each of the items listed below (if appropriate) and generate a numerical score for each item. The score will then be multiplied by the Sorting Weight assigned. These scores will be summed to produce a total score for the primer or pair. The total scores will then be used to rank the primers or pairs in order of preference so that the list can be rank sorted with the best primers at the top.

You can change the sorting weight settings for specific searches, or you can change the program default settings, if desired.

- **3' Dimer Scoring**

  Count the number of homologous bases in the largest duplex found involving the 3' end. Multiply by the 3' Dimer Penalty assigned in Sorting Weights. Preset penalty value: 1. (Not applicable for probes).

- **Any Dimer Scoring**

  Count the number of homologous bases in the largest duplex found. Multiply by the Any Dimer Penalty assigned in Sorting Weights. Preset penalty value: 1.

- **False Priming Scoring**

  Calculate the Tm for the worst-case false priming site found (single primer) or the sum of Tm for the worst-case false priming sites found for primers A and B (pair). Multiply by the False Priming Penalty assigned in Sorting Weights. Preset penalty value: 1. (For probes, calculate Tm for Any Homologies in place of False Priming.)

- **3 ' Degeneracy Scoring**

  Used for protein source molecules. Calculate the difference between the 3' degeneracy of this primer and the primer with the least 3' end degeneracy. Multiply by the 3' Degeneracy Penalty assigned in Sorting Weights. Preset penalty value: 1. (Not applicable for probes)

- **Total Degeneracy Scoring**

  Used for protein source molecules. Calculate a number that is 1 less than the total degeneracy of this primer divided by the minimum total degeneracy. Multiply by the Total Degeneracy Penalty assigned in Sorting Weights. Preset penalty value: 1

- **GC Mismatch Scoring**

  Used for pair search results. Calculate the difference in % GC between primers A and B. Multiply by the GC Mismatch Penalty assigned in Sorting Weights. Preset value: 1.

- **Tm Mismatch Scoring**

  Used for pair search results. Calculate the difference in ºC between primers A and B. Multiply by the Tm Mismatch Penalty assigned in Sorting Weights. Preset value: 1.

Primer Criteria vs Sorting Weights

You can change the primer criteria, or the sorting weights used in the program. These will have different effects, as described below.

- **Primer Criteria**

  You can change Primer Criteria for a primer search. For example, you can increase the 3' dimer cutoff from 3 to 4 bases. This means that you will accept primers with less than 4 matches at the 3' end. The criteria change means that you will probably find more primers (but they may be of lower quality).

  Changes to criteria may affect the number of primers found.

- **Sorting Weights**

  You can change Sorting Weights after a search has been completed. For example, you can increase the penalty for 3' dimers from 1 to 2. This means that primers with more matches at the 3' end will be more heavily penalized and these primers will receive a poorer score for ranking.

  Changes to sorting weights will not affect the number of primers found, but only their position in the list when sorted by rank.

Adding Primer Sites to the Molecule

Primer binding sites can be added to your molecule file from the Find Primer Sites Results display, the Link to Molecule results display, the primer design search results display, or from the primer viewer window (Edit tab) for a specific primer. Once these sites have been added to the molecule file, they can be displayed on the molecule map or in the formatted sequence.

- **Find Primer Sites Results**

  Click the Add Sites button to add all the listed primer binding sites to the molecule file.

- **Link to Molecule**

  Select the molecule to link the primer to.

- **Primer Design Search results**

  Click the Enter Primer Site to Map button to add the binding site for the selected (highlighted) primer or primer pair to the molecule file.

- **Primer Viewer Window, Edit tab**

  Click Tools, Enter Primer Site to Map to add the binding site for this primer to the molecule file.

Primer Identification

You can enter or edit the name of the primer and its description. This information becomes a part of the molecule file and can be viewed using the Site Properties button on the map display. The primer name will appear on the map and on the sequence display, so a short but descriptive name is helpful.

Displaying Primer Sites

Alignment methods

Altschul, S.F., W. Gish, W. Miller, E.W. Myers and D.J. Lipman. 1990. Basic Local Alignment Search Tool. J Mol Biol 215:403-410.

Chao, K-M, W.R. Pearson, and W. Miller. 1992. Aligning two sequences within a specified diagonal band. CABIOS 8:5, 481-487.

Dayhoff, M.O., R.M. Schwartz and B.C. Orcutt. 1978. A model of evolutionary change in proteins. In Dayhoff, M.O. (Ed), Atlas of Protein Sequence and Structure (National Biomedical Research Foundation, Washington DC) vol 5 suppl 3:345-358.

Henikoff, S. and J.G. Henikoff. 1992. Amino acid substitution matrices from protein blocks. Proc Natl Acad Sci USA 89(22):10915-10919.

Higgins, D.G. and P.M. Sharp. 1988. CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene 73:237-244.

Hirschberg, D.S. 1975. A linear space algorithm for computing longest common subsequences. Commun Assoc Comput Mach 18:341-343.

Myers, E.W., and W. Miller. 1988. Optimal alignments in linear space. CABIOS 4:1, 11-17.

Needleman, S.B. and C.D. Wunsch. 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. Mol Biol 48:443-453.

Pearson, W.R. and D.J. Lipman. 1988. Improved tools for biological sequence comparison. Proc Natl Acad Sci USA 85:2444-2448.

Scoring Matrix Options

You can select the scoring matrix for global ref or multi-way alignments. User adjustment of alignment parameters is available when using the Standard Linear Scoring Matrix

- **Standard Linear**

  Can be selected for DNA or amino acid sequence alignments.

  Penalty values are applied when bases do not match or when gaps in the alignment must be opened or extended. Advanced users can adjust these penalty settings.

  For protein alignments, you can indicate whether conservative changes will be permitted. If conservative changes are not permitted, exact matches will be required. The conservative changes recognized by the program are: (Ile, Leu, Val, Met), (Asp, Glu), (Asn, Gln), (Lys, Arg), (Ala, Ser, Thr)

- **BLOSUM 62**

  Can be selected for amino acid sequence alignments.

  Uses a substitution matrix (BLOSUM 62) which assigns a score for each amino acid substitution. This substitution matrix was derived from data using the BLOCKS database (Henikoff and Henikoff, Proc Natl Acad Sci U S A. 1992 Nov 15; 89(22): 10915–10919) in the study of distantly related sequences.

- **PAM 250**

  Can be selected for amino acid sequence alignments.

  Uses a substitution matrix (PAM 250) which assigns a score for each amino acid substitution. This substitution matrix was derived using a data set of closely related sequences and is based on the point-accepted-mutation (PAM) model of evolution described by Dayhoff (Dayhoff, M. O.; Schwartz, R. M.; Orcutt, B. C. (1978). "A model of evolutionary change in proteins". Atlas of Protein Sequence and Structure. 5 (3): 345–352).

FastScan Local Alignment ktup Settings

ktup refers to the number of bases that must exactly match. Different settings will be used for DNA and protein molecules. Translated DNA will use the protein molecule settings.

The ktup selected by the program based on search region length is as follows.

Local ktup Settings

DNA molecules:

<3,000 bases ktup = 4
<10,000 bases ktup = 5
<30,000 bases ktup = 6
<100,000 bases ktup = 7
<300,000 bases ktup = 8
<1,000,000 bases ktup = 9
<3,000,000 bases ktup = 10
>3,000,000 bases ktup = 11

Protein molecules:

All proteins, ktup = 2

Scan ktup Settings

The following settings are used for scan operations when the Standard speed option is selected. The ktup selected by the program based on search region length is as follows:

DNA molecules:

<300 bases ktup = 5
<1,000 bases ktup = 6
<3,000 bases ktup = 7
<10,000 bases ktup = 8
<30,000 bases ktup = 9
<100,000 bases ktup = 10
<300,000 bases ktup = 11
>300,000 bases ktup = 12

Protein molecules:

<1200 bases ktup = 3
>1200 bases ktup = 4

Rapid ktup Settings

The following settings are used for scan operations when the Rapid speed option is selected. The ktup selected by the program based on search region length is as follows:

DNA molecules:

<10,000 bases ktup = 8
>10,000 bases ktup = 12

Protein molecules:

<1200 bases ktup = 4
>1200 bases ktup = 5

For DNA scans, the search procedure is additionally accelerated by increasing the library molecule shift to 4 bases to minimize overlapping.

Table of Contents

Introduction 4
  New users and overview: 4
  Program Main Menus: 4
Getting Started 5
    Quick Start 5
    Select license type 5
    Main window 5
    Load a molecule 6
    View the molecule 6
    Use the molecule 7
Molecule File Handling 8
  Opening Molecule Files 9
  Problems with File Loading 10
  Retrieving Files from Entrez (E-utils) 11
  Saving Molecule Files 13
  Exporting Molecule Data 14
- Sequence Only 14
- FASTA 14
- Compact GenBank 14
- Commented GenBank 14
- Converting multiple files 14

  Creating New Molecules 15
  Using Clone Share 17
Molecule Viewer Window 20
  Overview 21
  Molecule map (working map) 22
  Working Map Display 23
    Features 23
    Enzyme Sites 23

Primer Sites ... 23
Label Sites ... 24
Map toolbar ... 25
Map Mouse Actions ... 26
Features ... 26
Enzyme Sites ... 26
Primer Sites ... 26
Label Sites ... 27
Circular vs Linear Maps ... 28
Working with Map Sites ... 29
View Properties ... 29
Jump To Sequence ... 29
Enzyme Suppliers ... 29
Enzyme Isoschizomers ... 30
Compatible Ends ... 30
Cut at a map site ... 30
Delete a map site ... 30
Create... ... 30
Export Sequence ... 30
Additional Functions ... 30
Resizing Site List Display Panel ... 31
Adding an Enzyme Site Manually ... 31
Send Map Options ... 32
- Save for Web ... 32
Molecule restriction map data ... 33
Methylation Sensitivity Alert ... 34
Restriction Map Format ... 34
Sort Options ... 34
Restriction Map Display Styles ... 35
List of Recognition Sites ... 35
Map of Recognition Sites ... 35

Fragment Sizes 35
Gel View of Fragments 35
No Cutters 35
Restriction Map Toolbar 36
Restriction Map Filter 38
Filter by Cut Information 38
Cut N times 38
Cut Outside Region 38
Cut Inside Region 38
Filter by Enzyme Characteristics 38
Ends produced by cut 38
Recognition element size 39
Specifying Two Cut Zones 39
Restriction Map Enzyme Actions 40
Cut with an Enzyme 40
Enter Sites for One Enzyme to Molecule Map 40
Enter Many Sites 40
Build User List 40
Viewing Sequence Information 41
Selecting sequence 42
Cut, Copy, or Paste Sequence Data 42
Sequence Toolbar 43
Sequence Format 44
Find Sequence Tools 46
Sequence Style Options 48
Partial Sequences 50
Editing Sequence Data 51
Printed Output Options 52
Viewing Molecule Features 53
Features Toolbar 53
Feature Draw as Types 56

Segmented features 56
Using the Features Table Filter 57
Entering Feature Information 57
Viewing Molecule Information 59
Information Toolbar 59
Zoom Molecule 61
Displaying Primer Sites 62
Map Display 62
Sequence Display 62
Data Lists 63
- Molecules 63
- Primers 63
- Enzymes 63
- Sequence Phrases 63
- DataBooks 63

Molecule List 64
- Selecting a Molecule 64
- Loading another Molecule 64

Enter to Molecule List, AutoScan 65
Enzyme Lists 66
Viewing Enzyme Lists 66
Enzyme List Toolbar 67
Enzyme List Locations 68
Program Enzyme Lists 69
Commercial (Main) 69
All Enzymes 69
Scanner List 69
User Enzyme Lists 70
Special Enzymes 72
On the Fly Enzyme List 73
Cloning and Enzyme Operations 74

Classic Cloning Techniques 74
Finding Restriction Enzyme Sites 75
Rebuild a Molecule's Enzyme Sites 77
Cutting Molecules 78
Cutting at Map Sites 79
Cutting with Enzymes 80
Cutting at Base Positions 81
Modifying Molecule Ends 82
Standard End Modifications 82
Custom End Modification 82
Ligating Molecules 83
- Cutting the DNA 83
- Modifying the Molecule Ends 83
- Invert the molecule 83
- PCR Amplify Fragment 83
- Undo Last Operation 84
- Perform ligation 84

Join Sequences 85
Cloning Wizards 87
Plan Cloning Wizard 88
Gateway® Cloning Wizard 90
Gateway® Cloning Results 92
Create Molecules - Multisite Procedure 94
Topoisomerase Cloning Wizard 96
Topoisomerase Cloning Results 98
Assembly Cloning Wizards 100
Assembly Cloning Results 102
Golden Gate Cloning Wizard 104
Ligation Independent Cloning 108
CRISPR 109
Design Crispr guide sequence 109

CRISPR enzyme selection ... 109
CRISPR wizard conditions ... 110
CRISPR wizard Off Target Search ... 110
CRISPR wizard results ... 110
CRISPR wizard results toolbar ... 111
CRISPR wizard results output ... 111
Analytical Molecule Operations ... 112
Auto Scan Molecule ... 113
Open Reading Frame Search ... 114
ORF Search Results ... 114
Analyze Molecule ... 115
Open Reading Frame Analysis ... 115
Analyze Molecule Toolbar ... 116
DNA Repeats Analysis ... 116
- Direct Repeats ... 117
- Inverted Repeats ... 117
- Palindromes ... 117

Dyad Symmetries ... 117
Protein Structure Analysis ... 118
Hydrophilicity (Protein) Analysis ... 118
- Kyte & Doolittle Analysis ... 118
- Hopp & Woods Analysis ... 119
- Surface Exposure Analysis ... 119

Mutagenesis Profile ... 120
- Restriction enzyme sites ... 120
- Translation ... 120
- Changing bases in the sequence ... 120
- Finding 'almost' enzyme sites ... 121
- Mutate - Change Amino Acid ... 121
- Mutate - Show Silent Changes ... 121
- Creating a mutant primer or molecule ... 122

Molecule Graphic Maps 123

Graphic Maps – Overview 123

Format Map Screen Basics 125

- Suppress very small Features 125

Format Map Print Basics 127

Format Map Features 128

To set feature print styles 128

To set custom feature settings 128

To shift a feature name 128

Format Map Sites 130

Format Map Content 131

Enhanced Maps 133

Overview 133

Getting Started with Enhanced Views 134

Basic molecule drawing 134

Text 134

Features 134

Enzyme sites 134

Adding features or sites 134

Copy or export map 135

Enhanced Map Toolbar 136

Enhance Feature Drawing 137

Feature Position 137

Feature Style 137

Feature Pattern 137

Feature Color 137

Move feature name 137

Change feature details 137

Working with Color 138

Map Basics 139

Map options 139

Base pair intervals 139
Feature Basics 140
Move feature name 140
Change feature name 140
Change color for one feature name 140
Site Basics 141
Site options 141
Tick lines 141
Add Features to Enhanced View Map 142
Add Sites to Enhanced View Map 143
Convert Map Data 145
Saving Enhanced Views 146
Save or Send Enhanced View Maps 147
Primer Designing 149
Overview of Designing Primers 150
Designing Primers 152
Create Primer from Molecule 153
Searching for Primers 155
One Primer Search Options 156
Primer Pair Search Options 156
Source Molecule and Target Positions 157
Primer Criteria - Overview 158
Primer Ranking and Sorting Weights 159
Starting Values for Primer Design 160
Primer Search Results 161
Results from Primer Pair searches 162
Results from One Primer Searches 163
Toolbar buttons 164
Enter to Primer List 165
Finding More Primers 166
Adjust (Relax) Primer Criteria 167

Auto Adjust Primer Length 168
Auto Adjust Primer Criteria 168
Primer Search Statistics Display 169
Sorting Primer Search Results 170
Primer Viewer Window 171
Primer Info tab 172
Primer Details tab 173
Change the primer sequence: 173
Primer Info and Details Toolbar 174
Primer Evaluation Bar 175
Primer Edit tab 176
Primer Edit Toolbar 176
Primer Analyze tab 178
Save Primers to Disk 179
Edit Primers 180
Adjusting Primer Position 180
Adjusting Primer Length 180
Edit Primer Sequence 181
Changing Primer 5' Extension 181
Primer Definition 182
Entering Primer Information 183
Direct and Special Primer Wizards 184
Direct Entry of Primer 185
PCR Cloning Wizard 187
View PCR Cloning Results 188
Sequencing Primer Wizard 190
View Sequencing Wizard Results 191
Sequencing Primer Results Toolbar 191
RT-PCR Design Wizard 193
LCR Oligo Set Wizard 194
Analyzing Primers 195

Analyzing Primers - Overview 195
Primer Analyses toolbar 196
Primer Reports 197
One Primer Report 197
Primer Pair Report 197
Analyze Dimers 198
Analysis Cautions 199
- Stability 199
- GC Clamp 199
- Runs 199
- Repeats 199
- Hairpins 199
Analyze False Priming 200
Check for False Priming against another molecule 200
Analyze Any Homologies 201
Check for any homologies against another molecule 201
Analyze Mix Wizard 202
Analyze Mix Results 203
Primer Values Profile 204
Primer List 205
Use, Organize Primers 207
Create Product 208
Finding Primers that Bind 210
Using Primer Sites 212
View Properties 212
Jump to Sequence 212
Create (primer or product) 212
Export Sequence 212
More Actions 212
Scan for Molecules with Sites 213
Link to Molecule 214

Unlink a Primer 215
Mutate Linked Primer Sequence 216
Internet BLAST Primer Searches 217
Primer Collections 218
Shared Primer Collections 218
Constructing a Primer Collection 219
Adding to your collection 219
Using your Primer Collections 220
View or Edit Primer Information 220
Select a Primer for use 220
Document collection file contents 220
Export / Import Primer Collections 221
Export a primer collection: 221
Import a primer collection: 221
Search in Primer Collections 222
Aligning Sequences 223
Getting Started with Alignment Operations 224
Looking at a New Sequence 224
- Find similar regions in a known sequence. 224
- Find similar sequences - remote search. 224

Showing Aspects of Relatedness 225
Checking for Vector Contamination 225
Organizing Alignment Setup Files 226
- Save alignment setup file. 226
- Enter to DataBook. 226
- Retrieve alignment setup file. 226

Aligning Multiple Sequences 227
Global-Ref Alignment 230
Assembled Alignment 231
Multi-Way Alignment 232
Sequences to Align 233

Alignment Parameters and Settings 234
Saving Alignment Setup Files 236
Comparing Two Sequences 237
Sequences to Compare 239
Compare with Local Homology Searches 240
Compare with Assembled Alignment 241
Compare with Global Alignment 242
Compare with Layout (Genomes) 243
Alignment Results Viewer 244
Info Tab - global alignments 245
Info Tab - local alignments 246
Picture Tab - global alignments 247
Picture Tab - local alignments 248
Sequence Tab - alignment results 249
Format Display Styles for Aligned Sequences 249
Feature Annotations 251
Overlay Composition Data 251
Page Styles for Aligned Sequences 251
Memorize Settings 252
Location Map Tab - global alignments 253
Location Map Tab - local alignments 254
Editing Alignment Data 255
- To replace a base 255
- To display a trace 255
- To shift a base left or right 255
- Update molecule sequence 256
- Alignment Name, Description 256

Saving Alignment Results 257
- Open results file 257

Exporting Alignment Data 258
Export Consensus Sequence 258

Export Merged/Spliced Consensus Sequence 258
Export Aligned Sequences 258
Standard use of alignment data 258
Retrieving Alignment Setup Files 259
Scanning for Similarities 260
Scan Wizard 261
DNA or Protein Scans 261
Identifying Search Locations 262
- Using Search Lists 263
Setting Scan Parameters 263
Scan Results Viewer 264
Scan Results Toolbar 264
Scan Results Wizard 265
Setting up a Multiple Sequence Alignment 266
Retrieving Scan Results 267
Saving Scan Results 267
Sequence Assembly 268
Sequence Assembly Wizard 269
Assembly mode and objective: 269
Sequences to assemble: 269
Assembly Parameters: 270
Sequence Assembly Viewer 271
- Save sequence assembly results 271
Sequence Assembly Toolbar 272
View Info 273
View Picture 273
View Sequence 274
Sequence and Overview 275
Sequence and Coverage 276
Sequence and Trace 277
Edit Consensus Sequence 278

NCBI BLAST 279
Starting a BLAST Search 279
Selecting Search Molecule for BLAST 279
Selecting BLAST Options 279
BLAST Programs 280
BLAST Databases 280
Viewing BLAST Search Results 281
Sequence Phrases 282
Sequence Phrase Collections 283
Shared Sequence Phrase Collections 283
Viewing Sequence Phrase Information 284
Sequence Phrase Toolbar 284
Entering Sequence Phrase Information 285
Construct and Use Sequence Phrase Collections 286
Export / Import Sequence Phrase Collections 288
Find Sequence Phrases 289
Selecting sequence phrase files or folders to search 289
Scan for Molecules with Sequence Phrase Sites 290
Selecting molecule files or folders to search 290
Find Molecule Sites Results 290
Find Sequence Phrase Results 291
Search in Sequence Phrase Collections 292
DataBook 293
Using the DataBook Database 294
View or Edit DataBook Record 294
Load a File for Use 294
Add to DataBook 295
Add new molecule or data file 295
Copy records already in a DataBook 295
DataBook Table Construction 295
Sharing DataBooks 296

More information ... 296
View/Edit DataBook Record ... 297
DataBook Searches ... 298
DataBook Tools ... 299
Export / Import DataBook Contents ... 300
Multiple File Conversion ... 301
Molecules convert format ... 301
Upload to Clone Share ... 302
Create Primer Collections ... 302
Create DataBook tables ... 303
Preferences ... 304
Basic Preferences ... 305
Home Directory ... 305
Author Stamp ... 305
Default Molecule File Formats ... 305
Options Preferences ... 306
Toolbar Size ... 306
Assembly Cloning Wizards ... 306
Import Features Preferences ... 307
Designated Features ... 307
All Other Features ... 307
Feature key selection list ... 307
Color Preferences ... 308
Custom Toolbar ... 309
Toolbar list ... 309
Available actions list ... 309
Clear ... 309
Reset ... 309
Setting up Shared Resources ... 310
Files that can be shared ... 310
File Menu ... 312

Opening Files 313
Saving Files 314
Export Data 315
Importing Molecules from the Clipboard 317
License Clone Manager 318
Manage License 318
View Menu 320
Clone Menu 321
Primer Menu 322
Align Menu 323
Align Idea Wizard 323
Discover Menu 324
Find in Files 325
To find text in a molecule file: 325
Selecting molecule files or folders to search 325
Save search lists 325
Search Results 325
Tools Menu 327
Process Molecule 328
- Invert Molecule 328
- Rotate Molecule 328
- Circularize Molecule 328
- Translate Molecule 328
Appendix 329
Installing Clone Manager 330
Installing Client Program 330
Client computer requirements 332
Activate seat license 332
Transfer a seat license to a new computer 332
Molecule File Formats 333
GenBank File Format 333

EMBL File Format ... 333
FASTA File Format ... 333
DOS Text File Format ... 333
Groups of Molecule Files ... 335
Common File Loading Errors ... 336
Unrecognized Character in File ... 336
Could not Find Start of Sequence ... 336
File Too Large ... 336
Invalid Base Position in Features Table ... 336
REBASE File Updates ... 338
Update REBASE ... 338
REBASE Files on the Web ... 338
Subscribing to REBASE ... 339
Getting REBASE Files via FTP ... 339
Restoring a prior Master Enzyme List ... 339
Translated Display ... 340
Genetic Codes ... 341
Export Data ... 342
Send View to File ... 343
Send View to Clipboard ... 344
Copying Text Data to the Clipboard ... 345
Copying Graphic Data to the Clipboard ... 346
Sending Text Data to a File ... 347
- HTML (*.htm) ... 347
- Rich Text Format (*.rtf) ... 347
- Plain text (*.txt) ... 347
Sending Graphic Data to a File ... 348
- Enhanced Metafile (*.emf) ... 348
- Windows Metafile (*.wmf) ... 348
Multisite Gateway® Cloning Components ... 349
Primer Thermodynamics and Secondary structure ... 351

Determination of Tm ... 351
Thermodynamics ... 351
Determination of Optimal Annealing Temperature ... 351
Determination of Secondary Structure ... 351
Primer Criteria Settings ... 352
- GC Content ... 352
- Melting Temperature (Tm) ... 352
- Annealing Temperature ... 352
- Stability 5' vs 3' ... 352
- Dimers ... 352
- Runs ... 352
- Protein degeneracy ... 353
- GC Clamp ... 353
Change Type of Criteria ... 354
Primer Design Experimental Conditions ... 355
Primer Sorting Weights ... 357
Primer Criteria vs Sorting Weights ... 358
Adding Primer Sites to the Molecule ... 359
Alignment methods ... 360
Scoring Matrix Options ... 361
- Standard Linear ... 361
- BLOSUM 62 ... 361
- PAM 250 ... 361
FastScan Local Alignment ktup Settings ... 362
Local ktup Settings ... 362
Scan ktup Settings ... 362
Rapid ktup Settings ... 362