

# Clone Manager

for Windows

Basic Edition

*Molecule, Cloning Operations*

Version 10

## Getting Started

Copyright © 1994-2019  
Scientific & Educational Software.  
All rights reserved.

The software described in this booklet is furnished under a license agreement and may be used only in accordance with the terms of the agreement.

Clone Manager software is intended for molecular biology applications. For Research Use Only. Not for use in diagnostic procedures.

License types can be for Individual users, Small WorkGroups, or network Concurrent use. Please refer to the Software License form received with your program disks to identify your license type and review the terms of the license agreement.

All rights reserved. No part of this booklet may be reproduced, stored in a retrieval system or transmitted in any form or by any means, without the prior written permission of the publisher.

# Table of Contents

---

## Getting Started

Install Clone Manager	1
User Preferences	2
Getting Started	2

## Molecule File Handling

Opening Molecule Files	3
Molecule File Formats	3
Importing Molecules from the Clipboard	4
Retrieving Files from Entrez	4
Find in Files	4
Saving Molecule Files	4
Sharing Molecule Data Files	5
Exporting Molecule Data	5
Multiple File Conversion	5
Using the Molecule List	6
Creating New Molecules	7

## Molecule Viewer Window

View Map	8
Working with Map Sites	9
Format the Working Map Display	9
View Sequence	9
Sequence Style Options	10
Using the Find Sequence Tools	11
Editing Sequence Data	11
View Features	12
Properties, Filter, Customize	13
View Information	13
Add Author Info, View GB, Base Renumbering	14
Zoom Molecule	14

## Graphic Maps

Standard Graphic Maps	15
Settings   Print Basics, Features	15
Settings   Sites, Content	16
Enhanced Views of Graphic Map	16
Basic Molecule Drawing	17
Enhance Feature Drawing	18
Working with Color, Saving views, Map output	19

## Enzyme Operations

Finding Enzyme Sites	21
Find Sites Results	21
Auto Scan Molecule	22
Enzyme Lists	22
Viewing Enzyme Lists, List Types	23
Updating Enzyme Lists	24
Viewing Restriction Map Data	25
Restriction Map Format	25
Using the Restriction Map Filter	26

## Cloning Operations

Cutting Molecules, Cut Results	27
Modifying Molecule Ends	28
Ligating Molecules	29
Join Sequences	30
Use the Plan Cloning Wizard	30
Use the Gateway Cloning Wizard	31
Use the Topoisomerase Cloning Wizard	33
Use the Assembly Cloning Wizards	34

## Additional Functions

Process Molecule (Invert, Rotate, Circularize, Translate)	36
Analyze Molecule	36
Open Reading Frame Analysis	37
DNA Repeats Analysis, Dyad Symmetries	37
Hydrophilicity (Protein) Analysis	38
Protein Structure Analysis	38
Open Reading Frame Search	38
Compatible Ends	39
Mutagenesis Profile	39
Find 'Almost' Enzyme Sites	40
Change Amino Acid, Show Silent Changes	40
Create Mutant Molecule	40
Export Data	41
Sequence Phrase Collections	41
Find Sequence Phrases	42
Scan for Molecules with Sequence Phrase Sites	42
DataBook Database	43
Databook searches	43
Add to DataBook	43

# Getting Started

---

Version 10 of Clone Manager Basic provides a comprehensive set of tools for sequence manipulation and analysis, sequence alignment, and primer design. You can also use Clone Manager as a quick and easy way to view or edit sequence files, find open reading frames, translate genes, or scan molecule files, searching for genes or text. The molecule viewer window puts all molecule information in one convenient location. Click tabs at the bottom of the window to see the molecule map, restriction map data, sequence, features table, or molecule information.

Clone Manager acts as an expert system to help you with cloning simulation, enzyme operations and graphic map drawing. Clone Manager will find restriction enzyme sites, cut molecules, modify ends, and ligate DNA. Through all of these operations, Clone Manager will track your features and enzyme sites and automatically renumber positions in the new molecules. A very fast and powerful restriction enzyme analysis function will help you to find enzymes to do needed operations. A set of cloning wizards will help you simulate cloning using current cloning techniques.

Molecule graphic maps are drawn quickly and accurately and high-resolution maps can be sent to your printer or to the Windows clipboard for use in another application. Use Enhanced Views of the graphic map to customize special molecule maps for use in a presentation or publication. Extra fonts, colors, styles, line width controls will help you to make some very special maps.

## Install Clone Manager

The seat-licensed version of Clone Manager can be installed on your personal computer. Please refer to information on our website ([www.scienced.com](http://www.scienced.com)) for current information on evaluation copies, purchases, installations and program activation.

Alternatively, you may have your computer set up to access Clone Manager under a network concurrent-user license.

When you use the program for the first time, a home directory will be set up for you. This home directory will be used to hold your personal program configuration settings and user enzyme lists. Subdirectories will be created below this directory to hold your DataBook (database utility) files and your primer and sequence phrase collection files.

## User Preferences

Click File, Preferences to select options to set user configuration options. Click the Basic tab to change your home directory, enter default author information, or select the default file format for saving molecule files. You can save molecule files in the standard Clone Manager molecule format (\*.cm5) or in the Clone Manager xml format (\*.cx5). The \*.cm5 format is preferred and has a file association set.

If you are running the program under a multi-user license (Small WorkGroup or Concurrent User license types), you can share enzyme lists, DataBook database files, or primer collection files, if these resources have been made available for you. You can also store copies of important molecule files in a central location (repository), perhaps in a standard file format. Click the Share Data tab to identify these locations and the repository data type. Click the Colors tab to change the program default settings for special screen display colors, if needed. These colors are used to bring your attention to important information, to distinguish one type of text from another, or to mark displays with ruler lines or highlighting.

Click the Import Features tab identify the feature keys that should be imported when you open a file in standard GenBank or EMBL format. You can select designated feature keys and set the option for handling all other, non-designated features.

**Designated Features** -- Click the Add button to include a new feature key on the import list. Select the Feature Key from the list of all keys currently supported in GenBank or EMBL file formats and select the Draw as Type and style number 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.

**All Other Features** -- This option is used to tell the program how to handle all feature keys that are not listed as Designated Features. You can discard all non-designated features (to keep your features table and file size as small as possible) or import as Information-only features. Information-only features do not appear on the map or annotated sequence but can be viewed on the features table if desired (click Settings button to show.)

## Getting Started



Double-click the Clone Manager icon on your desktop or click Start, Clone Manager to start the Clone Manager Basic program. All of the functions of the program are accessed by menus, toolbar buttons or shortcut keystrokes. For those who prefer the keyboard to the mouse, use the Windows Application key (or Alt + W) to call up a menu of the toolbar-accessible options for a display window.

The Clone Manager Help system provides comprehensive help topics to assist you in using the software or understanding a results display. Click Help on the File menu or use the toolbar Help button to view application help to browse for general information or review a feature you might plan to use. Click the Show button in the Help window toolbar to open the search pane to find all topics that contain specific keywords.

# Molecule File Handling

---

The molecule file handling functions of the Clone Manager program include opening and loading molecules for use, saving molecules to disk, and creating new molecules. These functions also maintain the list of loaded molecules (Molecule List) and handle selection of molecules for use in subsequent operations.

## Opening Molecule Files

You can use the File, Open command 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 molecule files, as needed. You can read molecule files in Clone Manager file formats, standard GenBank, EMBL or FASTA formats, or plain DOS text sequence files. You can control which feature keys are imported when you open standard GenBank or EMBL files. To open files containing molecule information:

- Click the Open toolbar button or click File, Open
- Find the file(s) you want in the Windows file open dialog box
- Click Open.



If 5 or more files are loaded at the same time, 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.

## 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: features, description, notes, enzyme sites, primer sites. 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.

If the program cannot automatically load a sequence file, it will identify the closest of the recognized formats (GenBank, EMBL, FASTA or ASCII text) and then try to provide you with information about the problem encountered and provide help with file loading. Usually, you will have one of the common file loading errors. Try and correct the error or split the sequence from the rest of the file and then attempt to reload the file.

## 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:

- In your e-mail document or other Windows workspace, select (highlight) the molecule data
- Click Edit, Copy to place the molecule information on the Clipboard
- In Clone Manager, 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.

## 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 file 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:

- Click File, Other Tools > Retrieve from Entrez eUtils
- In the Search In control, select Nucleotide for DNA molecules or select protein
- In the Query box, enter the locus name, accession number or descriptive information.
- If you need a special sort order for results, set using the Sort control.
- Click the Search button.
- 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. 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.

To view the selected molecule file, click on the file you want. Click the Show GenBank File button. To import the molecule for immediate use, select the file you want. Click the Get Molecule button. 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 at a later time.

## Find in Files



Found on the Discover menu or use 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, click Find in Files and enter the feature name or the text you want the program to search for. Identify the file or folders of files that should be searched and use the checkboxes to identify the locations in each file where the program should look. When the search has been completed, you can view the list of molecules found to contain the search text. You can print or copy the results list or load a molecule file for use.

## Saving Molecule Files

You can save the active molecule to its current name and folder by using the File, Save command or the Save toolbar button. You can also change the name or location of an existing file before you save it, by using the File, Save As... command. You can save molecule files in the 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 using File, Preferences.

## Sharing Molecule Data Files

In many cases, you will want to share files with other users in your organization. You can use File, Preferences to set options to prompt you to add author information or send a copy of a molecule file to a specified central location (repository) when it is saved for the first time.



**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. Click Tools, 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. Use File, Preferences to store the author name.

**Submit to Repository** -- You can instruct the program to send molecule data to a folder (location) you specify. This location may be a central repository for molecule files shared by a group, or a back-up location where you have decided to keep an extra copy of critical files. You can send a copy of the molecule file or just export the sequence, features, and descriptive information in a file format similar to GenBank. Click the Add Entry toolbar button or click View, Add to > Submit to Repository to send a copy of the active file to this location now. Use File, Preferences to store the repository location and default file format.

## Exporting Molecule Data

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 export the sequence of the molecule or the sequence plus features (in a format like GenBank) using Tools, Export > Export Molecule Data option. You can select from the following export formats:

**Sequence Only** -- Select this option to export only the sequence of the active molecule. The sequence will be exported in ASCII (DOS text) format in upper case letters, 50 bases per line, with no basepair 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.

**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.

## Multiple File Conversion

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. 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.

To start the file converter, click File, Other Tools > Multiple File Conversion. Select the file conversion process (molecule files, primer collections or DataBook tables) and identify the files or folders of files to convert. Click the Add Folder button to select a folder location to add to the list. Subfolders will be included. Click the Add File button to select one or more molecule files to add to the list.

For molecule file conversions, select the file format you want to convert to, using the guidelines provided in the application help, and select the location (folder) where you want the program to save the converted molecule files. For primer collection or databook table construction, enter the name for the new file or folder. The program will place this in the Collections or DataBook folder in your home directory. For databook table construction, use the checkboxes if you want to include primer or alignment files in the new databook table under construction.

## Using the Molecule List



Found on the View menu or use main toolbar button to access. 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 an asterisk 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 select groups of molecules for alignment operations in the Align Plus program.

**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're looking for.

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 of the molecules in the group for an alignment procedure or to remove all of these molecules from the program in one step.

**More Info / Less Info** -- show or hide additional information about the selected molecule. This information appears below the molecule list and includes the number of features, primers, and enhanced maps included for this molecule and the start of the molecule notes.

### Enter to Molecule List

When you create a new molecule by some action you have taken (perhaps by cutting an existing molecule, doing a ligation, or translating a sequence to create a protein), you will need to enter this molecule to the Molecule List. 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. You can enter a name and description for the new molecule.

## 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 elect to simply paste a sequence from the clipboard, bypassing the wizard's data entry pages. To create a new molecule:

- Click File, New
- Select Enter Complete Information and then follow Wizard instructions.
- 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 at a later time. (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 make changes to 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 more lengthy 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 at a later time.

### Entering Sequence Data

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 entry is 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 and cannot be asymmetric cutters -- click the Lookup button for help with enzyme names, if required. 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 feature type and style number. Next, enter start and end basepair positions (markers have only one position) and indicate if on the complementary strand. 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.

## Molecule Viewer Window

---

The molecule viewer window 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, regions, or markers). 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 string. 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 RMap tab provides detailed restriction map data (described in a later chapter).

### View Map

Click the Map tab at the bottom of the window to display the molecule working map. The map will be either circular or linear, depending on the molecule, and features and enzyme sites will be displayed. 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 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 to add or modify features. Click the Find button and select Find ORFs 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. 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 bp position for this enzyme and for any suppressed sites at this location. 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. If your map has no enzyme sites, click the Find button 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.

**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 the Find button to find Sequence Phrases in your collection that are present in this molecule.

**Primer Sites** -- If you are using Clone Manager Professional, primer sites can be shown above (outside) the map line in much the same way that enzyme sites are displayed. 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.



Click the Change Sites toolbar button to select to display enzyme sites, features, 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.

## 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, primers sites, label sites, or features entered for the molecule. Click on the site list header control to select what the list displays. Click to select an enzyme site, primer site, or feature in the list.

**Enzyme Sites** -- list shows the 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. You can use toolbar buttons to view the site properties, jump to the formatted sequence at the location of the enzyme site, cut at the site, or use the Tools button to view enzyme suppliers, enzyme isoschizomers, or compatible ends.

**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. You can use toolbar buttons to view primer properties, jump to the formatted sequence at the location of the primer site, or use the Tools button to export the primer sequence, create the primer, or start the process to create an amplified product using this primer.

**Label Sites** -- list shows the label name, start and end basepair positions and an indication if this label is on the complement strand. You can use toolbar buttons to view label properties or jump to the formatted sequence at the location of the label site.

**Features** -- list shows the feature name, start and end basepair 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. You can use toolbar buttons to view feature properties or jump to the formatted sequence at the start of the selected feature. Use the Tools button to export the feature sequence, create a protein by translating the feature, create a fragment containing the feature sequence, customize feature print styles or shift a feature name on the map.

## Format the Working Map Display



You can set several working map (screen) display options. These settings will be remembered and used for all working map displays. Click the Settings button and then click the Screen Basics tab. Use checkboxes to enable or disable options.

**Selected Features** -- set the checkbox to show a color segment on the map when a feature is selected. This 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. Normal strand primers appear in the color specified for enzyme or primer sites (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 basepair 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.

## View Sequence

Click the Sequence tab at the bottom of the window to view the molecule sequence. 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 basepair position in the sequence. You can select (highlight) a range of bases and make this sequence into a molecule feature or into a primer (requires Clone Manager Professional).



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

Click the Go To basepair position button and enter a basepair number. The program will scroll to the part of the sequence containing the basepair number specified.

## 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. To set sequence display style:

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



To select additional information:

- Use check boxes to indicate which items should be displayed with the sequence data
- For selected items, set related options (see below)
- 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 (requires Clone Manager Professional).

**Sequence** -- Single-stranded sequence displays the upper strand for the full sequence and the normal or complementary strand (as indicated) for a partial sequence. Double-stranded sequence displays both strands of the molecule, including sticky ends, where appropriate. 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.

**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. To set default sequence text colors, with the sequence visible, click the Settings button and then click the Colors tab. Colors selected for features will apply to all text of that type.

You can use a custom feature setting to distinguish a special feature, if needed. To set custom feature colors, click the Features tab to move to the features table. Select (highlight) the feature and click the Customize Feature button. Select the color for the annotations text for this one feature

## 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 a mono-spaced font (required so that sequence data lines up correctly). To set printed output options click the Print Options tab and select the font size you want to use for your sequence text.

## Partial Sequences

You can view the complete sequence of your molecule or just a part of the sequence. Click the Settings button, then click the Partial Sequence tab. Set the start and end basepair positions for the part of the sequence you want. 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.

## Make Selected Sequence Into...

While you are in the sequence display, you can select a range of bases and then make this selected sequence into a feature (gene, region, or marker) to add to your molecule's feature table. If you are using Clone Manager Professional, you can also make selected sequence into a primer or complementary strand primer. 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.

To make selected sequence into a new feature, select (highlight) the sequence bases and click the Tools button. Select Enter as Feature from the menu and enter the feature information. If you select to make the sequence into a primer, the primer viewer window will open and you can view the new primer and its evaluation. Modify the primer if needed and save it to disk or add it to a primer collection.

## Using the 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 of the sites found on both strands of the entire molecule.

While the sequence is visible on the screen, click the Tools button and then select Find or use Ctrl+F. Enter the sequence of bases you want the program to search for. Use the checkboxes to configure the search to permit limited mismatches or mismatches and gaps or insertions. To jump to matches in the sequence -- set the search type to Go To Sequence Found. If a matching string is found, the string of matching bases will appear as selected text (highlighted). Click Tools, Find Next or use Ctrl+I to locate additional places with the same sequence of bases. To view a list of sites found -- set the search type to Report All Sites. View or print the list of sites found.

## Editing Sequence Data

You can edit the sequence of a molecule by adding or removing bases. You can be in single-stranded, double-stranded, or translated mode, but you can only edit the upper strand of the molecule.



To edit sequence data, click the Edit Sequence button on the toolbar in the sequence view window. Make changes to the sequence data as needed and then click the Accept Changes (Update) button to complete the changes you made and modify the sequence of the molecule. Or click the Discard Changes (Cancel) button to abandon your changes and leave the sequence unmodified.

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 select (highlight) the sequence to be copied using the mouse or use Control+A to select the entire sequence. Click the Copy button to copy the selected sequence to the clipboard.

To paste sequence into the molecule viewer window, copy single-stranded sequence to the clipboard. Then open the molecule viewer window and click the Sequence tab. Click to position the edit cursor where you want to paste the sequence currently on the clipboard and 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.

## View Features

Click the Features tab in the molecule viewer window to view the features table. Molecule features appear in the listing with a symbol for the feature type, the feature name, basepair positions, description, and the feature key, if assigned. Click toolbar buttons in this window to add, edit or delete features, view feature properties, create a protein or fragment, export the feature sequence, set custom print style settings for an individual feature, or filter the data displayed. To see the basepair positions for join segments, move the mouse over the feature symbol (cursor changes to rectangle shape) and click. Click outside the pop-up box to dismiss.

Click the Settings button to format the features table display. You can select to show basepair positions start to end or left to right. You can select an option to show Information-only features in the features list. And you can elect to show a simple features map above the list or a box showing the feature style for the selected feature below the list. Click the Memorize button to remember your settings.

## Feature 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** -- usually represent translated sequences, have start and end basepair positions, and can be on the complementary or normal strand. Genes can appear on the molecule map and annotated sequence.



**Regions** -- zones of special significance for the user, have start and end basepair positions, be on the complementary or normal strand. Regions can appear on the molecule map and annotated sequence.



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



**Labels** -- sequence areas of special interest to the user, have start and end basepair position. Labels can be shown on the annotated sequence and appear on the molecule map as label sites.



**Info only** -- not shown on the molecule map or on the annotated sequence. Use for features where you want to preserve the information but do not want it to clutter your map or features table.



Click the Add Feature button to add a new feature, then enter the feature information. To edit feature information, select (highlight) the name of the feature and click the Edit Feature button. To delete a feature, move the highlight to the name of the feature you want to remove and click the Remove Feature button

## Viewing Feature Properties



To view feature properties, select (highlight) the name of the feature and click the Feature Properties button. For Genes, properties include basic information about this gene, its translation, and the amino acid composition, codon usage, molecular weight and pI for the translation product. For Regions, properties include basic information about the region defined and its base composition.

To view GenBank location or qualifiers, select (highlight) the name of the feature. Click the GB Location/Qualifiers button. The information displayed will include feature location data (basepair positions) and qualifiers available for this feature. GenBank format files imported in version 9.2 or later of Clone Manager will read in all feature qualifiers and retain this information when saved as a Clone Manager file.

## 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 molecule viewer window features table by location or feature type. When a filter has been applied, the Filter toolbar button will be depressed. Click the depressed Filter button to remove the filter. Use the Settings toolbar 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 if you want to list all feature types.

**Filter by Location** -- Check the box for this option and then enter the two basepair 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.

## Customize Feature Styles

In the lower area of the features table display, 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. To assign custom styles:

- Select (highlight) the name of the feature and click the Customize Feature button
- Set the custom style and color to be used for this feature for the graphic map.
- Or set the custom text color to be used in the annotated sequence display.
- Set the Default button to return to the default values.



## View Info

Click the Info tab in the molecule viewer window to view molecule information. Click the Edit button to change the molecule name, size, description, or notes. Click toolbar buttons in this window to add author information to the notes field, change the starting basepair number for this molecule, change the translation table, or access GB Annotations (header information) if the file is a GenBank or EMBL format file. 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.

**Molecule name** -- appears in the Molecule List, on program screens and in maps and other printed output.

**Size in bases** -- 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 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 molecule. Can be keywords, history of construction, comments. This can also be printed on molecule map pages.

## Adding Author Information

You can add an author name, date, and notebook reference to the beginning of the Notes field for your Clone Manager molecule (\*.cm5) files. As part of the Notes field, this information can be printed on the page with your molecule map or written to a DataBook entry. To add author information, click the Author Stamp button.

## Viewing GenBank Annotations

Annotations refers to the header information found in a GenBank or EMBL format file. To view GenBank Annotations click the GB Annotations button. After modification to the molecule file (such as adding enzyme sites or performing cloning operations), you will want to save the file. Files are saved in a proprietary format which includes all of the information you entered about this molecule: features, enzyme sites, description, notes. The annotation information is not saved as a part of this file format, but the program will automatically save a link to the original GenBank file so that you can later find this information.

## Base Renumbering

You can designate a number other than 1 to use as the first basepair number for the sequence of the active molecule. To assign a starting basepair number, click the Base Number Start button and enter the starting basepair number for this molecule.

## Zoom Molecule

You can select to display only a part of a large molecule. This is often helpful if there are a large number of 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 just click the Zoom button (left end of tab bar at bottom of window) and enter the start and end basepair positions for the region to display. The data display is regenerated and the window title identifies the zoomed region.

The 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 indicate that the feature extends further in that direction. The 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. The Sequence view includes only the sequence of bases in the zoomed region. The Features table view lists features that are found within the zoomed region, even if they are not fully contained within this region.

To return the map to show the full molecule, click the Zoom tab again and select Show Full Molecule.

# Graphic Maps

---

When a molecule has been loaded into the Clone Manager program or when a molecule is selected from the Molecule List, the molecule viewer window will open, showing a working map of this molecule. The program can also draw a high-resolution graphic map that can be sent to your printer, copied to the Windows Clipboard, or exported as a file for later use by another program. You can also create and save as many as six Enhanced Views of the graphic map 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.

## Standard Graphic Maps

By design, all of the standard graphic maps will look consistent in style. You store your map preferences (like typeface, gene style, region color, etc.) in the 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.

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.

### Settings | 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, click the Settings button and then click the Print Basics tab. 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. Use checkboxes to indicate which text items to print, select a color for each item type and select a color and weight for the mapline.

### Settings | Features

You can control the style and color for feature types 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. To set feature print styles click the Settings button and then click the Features tab. Use checkboxes to indicate which feature types to print, then select a shape/fill pattern and color for Style 1 of each feature type. Select a shape/fill pattern and color for Style 2 and Style 3 of each feature type, if desired.

To set custom feature settings, with the map visible on the screen, select a feature and click Tools, Customize. Select the style and color for this one feature or Default to remove this custom setting. To shift a feature name, select a feature and click Tools, Shift Feature Name. Enter the number of increments (in units of 1/10 inch) to move the text horizontally or vertically on the printed page. You can also right-click to access these functions.

## Settings | Sites

You can select one of two preset font sizes for enzyme, primer, and label site 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 basepair numbers for each map site or if it should italicize the first three letters of the enzyme name. To set enzyme print styles, click the Settings button and then click the Sites tab. Set the font size for map site names to Standard or Small and select a color for names and tick lines. Use the checkboxes to print basepair numbers or italicize enzyme names.

## Settings | 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 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, click the Settings button and then click the Content tab. Select Molecule Map and Text and 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 or Position** -- You can change the size of a graphic map before printing. You can also resize a map before copying to the clipboard or exporting. 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. To change the size of the graphic map, click the Settings button and then click the Content tab. Select Molecule Map Only and select 80% or 90% to decrease the map size, or select 110% or 120% to increase the map size. To change the position of the map on the page, click the Settings button and then click the Content tab. Select Molecule Map Only. Move the map left or right by using the horizontal position slider control. 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 Views of Graphic Map

You can use an 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.

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 basepair 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 for use in other programs.

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

- Open the molecule viewer window and click the Map tab.
- Click the Enhanced View button and select New View.



The new enhanced view window will open with your molecule map, drawn using the 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.

**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 text color.

**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 to add basepair numbers to enzyme sites or to italicize the first three letters of the enzyme name. To remove enzyme sites that are not needed, click to select and then click the Remove Selected toolbar 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, label 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 Feature, Add Enzyme Sites, Add Primer Sites, or Add Label Sites, as needed.

## Basic Molecule Drawing

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

**Map Basics** -- 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 basepair 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. (This text can be moved on the map, if needed. Just select the text (click on it) and then drag the text using your mouse.) Clear the check boxes to turn off the display of either the molecule name or the molecule size text.

Select the line width and color to be used for the Map Line.

Basepair intervals can be numbered in bps or kbps and you can set the font for the interval numbers. Clear the checkbox if you do not want to display the basepair interval numbers. Basepair 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 basepair interval tick lines.

**Feature Basics** -- 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. (To 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.)

To move a 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.

To change a 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.

**Enzyme/Primer/Label Site Basics** -- Select the Font Face, Font Style, and Point Size to be used for all enzyme/primer/label (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 site names.

You can use the checkboxes to set the display of basepair numbers for all map sites or to italicize the first three letters of enzyme names.

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 map sites and are using a small point size for map site name text, select a thin tick line. If your names are large or you are using a thick map line, select a thicker tick line for balance.

## 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 a large number of features, while wider styles provide greater impact. Markers have a subset of drawing styles, drawn around the single Marker basepair 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. Sixteen 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.

## 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.

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

You can change the color for the molecule map line, basepair interval tick lines, and enzyme or primer site tick lines only in the Settings dialog boxes. 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.

## 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.

Click the Save or Send toolbar button in the enhanced view display window and select Save View. The enhanced view data will be 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). Remember to save the molecule file to disk before exiting the program.

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 button and select Organize Views. 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.

## Graphic Map Output

You can preview or print the high-resolution graphic map or copy or export the graphic map for use in another program. The highest resolution graphic maps are exported as vector graphic files (\*.wmf or \*.emf), which can easily be scaled. You can create an enhanced view map and export as a raster graphic file, if needed.

**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 File, Other Tools > Print Preview to check the appearance of your map before printing, copying to the clipboard, or exporting to a disk file.

**Clipboard** -- Click the 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. Use Paste to copy the map to your document.

**Export** – Click the Send View to File button 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. Use Insert, Picture to add the map to your document.

**Save for Web** – Click the Save for Web toolbar button in the map window to use this option to prepare a raster graphic map in the Portable Network Graphics (\*.png) format. This format is 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.)



### Additional options for Enhanced View Maps

For enhanced views of the graphic map, click the Save or Send button in the enhanced view display window to copy your enhanced view map to the Windows clipboard or export to a disk file in vector or raster (bitmap) file formats.

Vector file formats (emf, wmf) provide high-resolution images that can be resized accurately. Selected raster file formats (tif, gif, png) provide high quality map images combined with small file size. For raster copy and export options, you can select a low-resolution bitmap (looks similar to the map on the screen) or a high-resolution bitmap (looks similar to the printed map).



# Enzyme Operations

---

The Clone Manager program can find restriction enzyme recognition sites -- for one enzyme at a time, or for all enzymes on an enzyme list meeting specified criteria (like single or double-cutters). Once an enzyme site has been entered, you can view this site in detail (recognition sequence, cut position, cohesive ends produced). The program uses enzyme lists to store the information needed to work with restriction enzymes. You can view, print or update enzyme lists and you can use a helpful wizard to build user-designated enzyme lists. You can keep your master enzyme list up-to-date by doing an update procedure using the latest REBASE enzyme data file (available monthly over the Internet).

Clone Manager has a powerful restriction enzyme analysis module that displays information on restriction enzyme sites as a page in the molecule viewer window (click the RMap tab to get to this display). You can view the restriction enzyme analysis in one of five display styles, including a list of no-cutters. You can also set filter options so that you see only the results you need.

## Finding 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 (all single or double cutters for any enzyme list, or all sites for all enzymes on a user-defined enzyme list). The Rebuild map sites list 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.

### Finding Sites for Enzymes

You can search a molecule looking for recognition sites for one or more restriction enzymes. Once found, you can see the number and location of these sites, enter all of these sites to the molecule map, if desired, or cut the DNA with one enzyme or at one specific site. To find recognition sites for enzymes:

- Click Clone, Find Enzyme Sites and then click Enzyme(s).
- Type in the names of the enzymes, separated by a comma, or find the enzymes in the lookup list to the right (see below).
- Click OK to find 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 begin 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.

### 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 of 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:

- Click Clone, Find Enzyme Sites and then click List of Enzymes.

- Select Single Cutters, Double Cutters or All Sites User List.
- Select the enzyme list using the Enzyme List combo box on the right.
- 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 program-defined enzyme lists.

## 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. Vertical bars above the icon map show the approximate location of the enzyme sites and a red arrow marks the highlighted position. You can add all of the enzyme recognition sites found to the molecule map, cut the molecule at all the sites found for the enzyme highlighted, or cut the molecule at the one enzyme site highlighted.

## Auto Scan Molecule

Found on the Discover menu or use Find 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.

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 Find button in the map window toolbar and select Find Enzyme Sites (Autoscan) to start this operation.

## Enzyme Lists

The Clone Manager program uses enzyme lists to hold the information about each enzyme that is necessary for its use. You can view, 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. Types of enzyme lists include program lists, user lists, shared lists, special enzymes, and the on-the-fly list.

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 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.

## Viewing Enzyme Lists



Click the Enzyme Lists button on the main toolbar to open the enzyme list resource. Click on Program Lists or My User Lists to see enzyme lists in these categories. Click on the Folder Options toolbar button to add a new empty enzyme list, or to rename, delete, import or export a user enzyme list.

Click on an enzyme list in the left panel to display the contents of this list in the right panel. For each enzyme in this 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. Type in the initial letters of an enzyme name to quickly move to that enzyme in the list. Use the Enzyme Info toolbar button to view enzyme suppliers, isoschizomers, compatible ends, or enzyme properties.

## Program Enzyme Lists

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

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

**All Enzymes** -- This is the master list containing all of the enzymes recognized by the program. There are more than 640 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 6 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 enzyme list:

- Click the Enzyme Lists toolbar button
- Click the Build User List button.
- Follow the wizard instructions.



Once your enzyme list has been created, you can later add or remove enzymes, as necessary. You can add enzymes that the program recognizes or Special Enzymes (perhaps not commercially available) that you define. To add an enzyme to an existing list:

- Click the Enzyme Lists toolbar button
- In the left panel, select the user-defined enzyme list that you want to add enzymes to
- Click the Edit User List button and select Unlock List to allow changes to this list.
- 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.
- Click OK to add this enzyme to the selected user enzyme list.



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

You can export or import user enzyme lists, as needed, to provide a list to a colleague or to move your list from one computer to another. You can also rename or delete a user enzyme list, if required. Select the list and then click the Folder Options button.

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 No Cutters). Just click the Tools button and select Build User List to get started.

## 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 enzyme list:

- Start the operation you want to perform or click the Enzyme Lists button on the main toolbar.
- Select the enzyme list On-The-Fly (left panel).
- Specify enzyme characteristics (see below).

**Recognition Element Size** -- Select recognition element sizes of 4, 5, 6, or more than 6 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.

## Updating Enzyme Lists

You can update the Clone Manager master enzyme file monthly using REBASE data files so that you are always up-to-date with new enzyme information. 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. To do a REBASE update:

- Click the Enzyme Lists toolbar button
- Click the REBASE Update button
- If you already have the latest REBASE file (via email or from the web), set the option **Have the File** and provide the complete filename and path.
- Set the option **Get the File Now** to have the program get the file for you using anonymous FTP.
- Click **Update**



The program 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 in order to use the new enzyme lists. To get the latest REBASE data file, connect to the REBASE web site <http://rebase.neb.com>, click on Data Files, click on file format #35, Clone Manager, and then click Get File. Save the data file to your computer. When you do an update procedure, you will need to identify this file.

## Viewing Restriction Map Data

Restriction map data for the active molecule (or the zoomed region of the molecule) is displayed in the molecule viewer window. Click the RMap tab at the bottom of the window to move to this display. Restriction enzyme recognition site information is listed in table format or displayed in a graphic format. You can print the data in this view window, copy this view to the clipboard, or export the view data to a disk file.



Click the Select View button or click the data column header 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).

**Special Notations** -- Enzymes with a large number of 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.)

**Enzyme Information** -- If you find an enzyme on the Restriction Map display that looks like it will cut in just the right place, but you need more information about this enzyme, you can get enzyme properties, suppliers, isoschizomers, or compatible ends information. Select the enzyme and then click the Enzyme Info button.

**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. Click the Tools button to access these options.

### Restriction Map Settings

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.

**Memorize Settings** -- Once you have selected the enzyme list, display style and sort order, click the Memorize button to instruct the program to remember these settings and use them whenever you initially open a restriction map window.

**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 restriction map window toolbar or in the Settings dialog box.

**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.

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

**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.

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 use the Select Views button. 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.

**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 basepair position for each recognition site for this enzyme appears in ascending basepair 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 affect on the list of non-cutting enzymes.

## Using the 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.

Use the Cut Outside Region option to designate a region you would like to protect -- enzymes that cut in this region will not be included in the display. Use the Cut Inside Region option to specify the minimum and maximum basepair 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 of 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. Use the check boxes to select the characteristics required. Leave all of the check boxes blank to accept any type of ends or to accept any size recognition elements.

# Cloning Operations

---

Clone Manager can cut DNA at enzyme cut sites or at specified basepair positions. Molecule features and enzyme sites are automatically recalculated after a cut operation. The program can ligate DNA, checking for compatible ends. You can modify molecule ends, if needed for ligation, and you can select the cloning orientation. For fast simulated cloning, you can cut, ligate, and modify ends from the same display window showing sticky ends and molecule icon maps.

The Plan Cloning Wizard lets you describe your objectives, review the list of possible solutions, and instruct the Wizard to construct the recombinant of your choice. The Gateway, Topoisomerase, Ligation-Independent, Gibson Assembly and In-Fusion cloning wizards simulate cloning using these techniques. The wizards will help you select the appropriate components, show you the proposed result, and create final recombinants.

## 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. Or you can cut at one specific enzyme site already marked on your molecule map or at specified basepair positions.

### Cutting with Enzymes

You can cut a molecule using restriction enzymes. The molecule will be cut at all of 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:

- Click Clone, Cut and then click Enzyme(s).
- Type in the names of the enzymes, separated by a comma, or find the enzymes in the lookup list to the right (see below).
- 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 begin 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.

### 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:

- Click Clone, Cut
- Click Map Site and select one enzyme site from the list box.
- 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 Properties toolbar button to see the exact cut position and ends produced.

## Cutting at Base Positions

You can cut a molecule by specifying one or two basepair 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:

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

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

## 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 basepair 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. If cutting with enzymes, 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.

## 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 9 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 together two or more pieces of DNA or to join the two ends of the same molecule (self-ligate). The ligate 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 of 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.

**Invert the Sequence** -- use the Invert Fragment button to reverse (invert) the sequence of the selected molecule.

**Undo Last Operation** -- use the Undo button to undo the last modification to the selected molecule.



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 name and description.

First, 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 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). You can select the method to use to join the two sequences:

**Append** -- 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 basepair 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 10 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 10 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 basepair number in Sequence A after which the insert should be made.

Enzyme sites for the new molecule can be auto-scanned at the completion of the join operation. 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 (but this may result in some enzyme sites that are not single-cutters for the newly-joined molecule).

## Use the 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 use the Plan Cloning Wizard, click Clone, Plan Cloning and follow the wizard instructions. The wizard will help you 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 needed. 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. 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) in order to do recombinant selection by insertional inactivation. If the vector molecule contains a multi-cloning site, the basepair positions for this region will automatically be entered for you.

**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 of 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 possible 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. Select (highlight) the result of interest and use toolbar buttons to view the details of a suggested solution, construct a recombinant, or redefine the experiment, if needed.

## Use the 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](http://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 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'. If you are using Clone Manager Basic, the region to be amplified will be extracted and the attB extensions added to simulate PCR product generation.

If you are using 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.

## View Gateway® Cloning Results

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. Click the Create Molecule button to have the program automatically do the required steps to simulate the

cloning experiment and produce the recombinant molecule. You can enter this new molecule to the molecule list to view now or save to disk for later use. Following a multisite procedure, you can also collect the intermediate molecules generated during the procedure (entry clones, PCR products).

## Use the 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](http://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. 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'. If you are using Clone Manager Basic, the region to be amplified will be extracted and extensions added to simulate PCR product generation.

If you are using 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, selecting the best.

## View Topoisomerase Cloning Results

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 to disk for later use. 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 another topoisomerase cloning procedure.

## Use the Assembly Cloning Wizards

You can use assembly cloning wizards to help you 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.

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 sets 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. 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 basepair positions (upper strand) that immediately flank the site where the assembled insert will be placed. When you select to enter basepair 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.

If you are using Clone Manager Basic, the region to be amplified will be extracted and the overlap extensions added to simulate PCR product generation. If you are using 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.

## View 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 to disk for later use. When entering to the Molecule List, you can change the molecule 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.

If you are using Clone Manager Professional, you can also create and save the primers used to prepare PCR amplified fragments during the Create Molecule operation. 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.

## Additional Functions

---

Clone Manager can search for Open Reading Frames and easily enter found ORFs as genes in the features table for the molecule. It can provide information about cohesive ends to help you select enzymes for use. You can use Clone Manager to analyze DNA molecules to find direct or inverted repeats and dyad symmetries and you can analyze protein (or translated DNA) molecules for hydrophilicity and look at protein structure (helix, sheet, and turn configurations). A more detailed Open Reading Frames analysis is a part of the Analyze Molecule option.

Clone Manager includes operations and utilities to rotate, invert, circularize or translate a molecule, export results data, find text in molecule files and show a mutagenesis profile that provides a complex view of a sequence reacting to base changes. A simple database utility, DataBook, can be used to organize and store information about your molecule or data files. A new module helps you to manage and use sequence phrases.

### 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, 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** -- invert a copy of the active molecule, changing the positions of all sites and features.

**Rotate Molecule** -- 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 basepair position to appear at the top of the map (the new origin of the molecule).

**Circularize Molecule** -- 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** -- 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 basepair position for translation. Use the check box to indicate if you want to translate the complement strand.

### Analyze Molecule

You can analyze a DNA 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 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 analyze your molecule: Click Tools, Analyze Molecule, select the type of analysis, and identify which part of the sequence to analyze.



## 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 or not 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. 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 toolbar buttons to display Text or Graphic displays. Click the Start Codon button to select analysis results with Any, ATG or GTG start codons.

**Text Display** -- for each open reading frame, the program reports start and end basepair position, length in amino acids, frame, and Fickett's TESTCODE score values. The first amino acids of the open reading frame sequence are displayed.

**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 color. Full height black bars mark terminators. Half height colored bars can be used to mark start codons.

## 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 Repeats Type button in this window to move between the three repeats displays (direct, inverted or palindromes) or return to the analyze set-up dialog box (Redefine). 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** -- 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 sequence.

**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** -- 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 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 basepairs can be stable under these circumstances and may be a part of duplexes found.

## 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 Graphic displays. In the text display, numerical values are given for Kyte-Doolittle, Hopp-Woods, and Surface-Exposure analyses for each amino acid residue. 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. In the 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 residue 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. 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  $\pm 4$  relative units.

**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. 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  $\pm 3$  relative units.

**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, values above the axis line are predicted to be exposed on the surface of the protein. High values 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.

## 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 Graphic displays.

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

## Open Reading Frame Search

Found on the Discover menu or use the Find 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.

When a search for Open Reading Frames has been completed, 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). The colored 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. You can also set the checkbox to Show Overview to see all of the listed open reading frames positioned in their approximate locations in the six translation frames. The highlighted ORF is shown in red.

## Compatible Ends

Clone Manager can show you the cohesive end that would be produced by cutting with a specified restriction enzyme and a list of enzymes that could be used to produce compatible cohesive ends. This function can be accessed from the molecule map or restriction map displays, or while viewing enzyme lists or working in the ligate window. The list of enzymes that could produce compatible ends is limited to prototype enzymes and neoschizomers. (If more than one enzyme could produce exactly the same compatible end, only one enzyme is shown (the prototype for this group)). If another enzyme has the same recognition sequence but cuts in a different position (a neoschizomer), this enzyme will also be listed if it will produce a compatible end.

**Molecule Map** -- Click on an enzyme site and then click Tools, Compatible Ends. The display will show the cohesive end produced by cutting with this enzyme at this site, as well as enzymes that could produce compatible ends. Even if the enzyme has a degenerate recognition sequence (it can recognize more than one sequence), the sequence of bases at this specific enzyme site is known.

**Restriction Map or Enzyme Lists** – In the RMap window, click on an enzyme and then click the Compatible Ends button. If the enzyme has a degenerate recognition sequence, ambiguous bases in the overhang will be expanded and each of the possible overhangs shown. For each overhang shown, enzymes that could produce compatible ends are listed (if any). In the enzyme list display, click Enzyme Info, Compatible Ends.

**Ligate Window** -- When you are in the process of performing a ligation, you can cut or modify molecule ends, as needed. If you need help with finding an enzyme that can produce cohesive ends compatible to one of the ends in the ligation window, you can click the Fragment Information button and then select the end you need to match. Site-specific compatible ends information will be displayed, as described for the Molecule Map (above).

## Mutagenesis Profile

Found on the Tools menu. You can view 50 bases of sequence in your molecule around the target position specified. 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 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. One frame of translation is shown. You can use the toolbar control to change the frame, if necessary. If you identified 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 basepair position is reported at the top of the ruler line. With the vertical ruler line on the base that you want to change, press the letter key on the keyboard to make the change to the sequence base or click on the Edit Base toolbar 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.

**Find 'Almost' Enzyme Sites** -- Click Tools, Find 'almost' Enzyme Sites to locate 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. Once you identify the change you want to make, 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.

**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. 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.

**Show Silent Changes** -- Click Tools, Show Silent Changes to see which bases in the sequence you could change without changing the translation shown. Possible silent base changes are shown 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. Press the letter key on the keyboard to make the change to the sequence base or click Edit Base and select the change you want. 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 Show Silent Changes again to clear this display.

**Creating a Mutant Primer or Molecule** -- After you have made changes to the bases shown in this display (introduced a mutation), you can click Tools, Create Mutant 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.

## 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.

**Data in View** -- Select this option to export all of the 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 display such as on a web page or in an on-screen presentation.

**Primary Data** (where applicable) -- If the active window is a part of the Molecule Viewer Window or the Alignment Viewer Window, you can export primary molecule data or primary alignment data. If the active window is a part of the Primer Viewer Window, you can export the sequence of the primer in plain text format, or the primer name and sequence in FASTA format.

**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

## Sequence Phrase Collections



Found on the View menu or use main toolbar button to access. 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 groups 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.

### Enter Sequence Phrase Information

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 or export the collection file data in tab-delimited format to import into another application.

To enter a new sequence phrase, click on a collection file to open this collection. Then click the toolbar button Add Entry. Fill in the information about the new sequence phrase. 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 Options to add new folders or new empty collection files, or to rename or delete any of the files or folders you have created.

To view or edit sequence phrase information, select a sequence phrase by clicking on the entry in the list in the right panel of the collections window. Use toolbar buttons to edit the information, delete the 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.

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. In the left panel, click on the collection file or folder you want to search. Type in the text to search for in the Find edit box (lower area) and click the Go button. If sequence phrases are found that meet your 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. You can select a sequence phrase for use from the 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.

## Find Sequence Phrases

Found on the Discover menu or use toolbar Find 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. The program will open each sequence phrase collection you select, 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 particular search.

**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 10 bases in the search sequence. For short search sequences of less than 10 bases, one difference (gap or mismatch) is permitted.

When the search has been completed, the results will be presented in a scrolling list box and location map display. Found sequence sites are listed, showing the phrase name and the basepair position of the 5' end of the 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. 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 to your molecule file, mouse-click on the checkbox in front of the phrase name. Click the Tag All button to automatically tag all of the sites in the list. Click the Add Sites button to add all of the tagged sequence phrase sites to the molecule file and close the results window.

## 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 specific 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 a search 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 basepair position of the 5' end of the matching site. Also reported is the number of differences (mismatches or gaps, if applicable). You can print or copy the results list or load a molecule file for use.

## Export / Import Sequence Phrase Collections

To export a sequence phrase collection, open the Sequence Phrases collection module and select the collection you want to export. Click the Folder Options button and select Export. In the Export Collection dialog box, select which data to export.

To import a sequence phrase collection, open the Sequence Phrases module and click My Collections. Click the Folder Options button and select Import. Select the tab-delimited text file containing your sequence phrase data. In the Import Collection dialog box, specify the first row that contains sequence phrase data to import. 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. Select a match mode and enter a name for the new sequence phrase collection.

## The DataBook Database



Found on the View menu or use 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. Click the Add Entry button on the main toolbar, then select Add to DataBook. 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.

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. Use the Tab key or Shift + Tab to move between the left and right panels or click with the mouse. Click the Folder Options 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.

## 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 and 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 and 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.

## 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).

To add a new molecule or data file, open the molecule, primer or alignment results file and click the Add Entry main toolbar button, then select Add to DataBook. Identify the databook where you want to place this new record. Use the checkbox (lower right) if you want to view the record as soon as it is created. 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. You can also copy a record already in a databook -- when viewing records in your databook files, 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.