

Clone Manager Suite

for Windows

Volume 1
Molecule, Cloning Operations

Version 7.1

Getting Started

Copyright © 1994-2004
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.

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.

Clone Manager Suite, Volume 1

Table of Contents

Getting Started

Installing the Clone Manager Suite	1
Getting Started	2
Output Data	2
User Preferences	3

Molecule File Handling

Opening Molecule Files	4
Sequence File Formats	4
Importing Molecules from the Clipboard	5
Retrieving Files from Entrez	5
Using Work Boxes	5
Saving Molecule Files	6
Sharing, Exporting Molecule Data	6
Using the Molecule List	7
Creating New Molecules	8

Molecule Viewer Window

Zoom Molecule	9
Viewing Molecule Maps	9
Viewing Sequence Information	10
Sequence Style Options	10
Using the Find Sequence Tools	12
Editing Sequence Data	13
Viewing Molecule Features	13
Viewing Molecule Information	15
Add Author Info, Base Renumbering	15

Molecule Maps

Working Map Display	16
Working with Map Sites	16
High-Resolution Graphic Maps	17
Control of Map Basics	18
Control of Map Features	18
Control of Map Sites	19
Control of Map Content	19
Graphic Map Output	20

Enzyme Operations

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

Cloning Operations

Cutting Molecules	27
Cut Results	28
Ligating Molecules	28
Ligation Results	29
Modifying Molecule Ends	29
Use the Plan Cloning Wizard	30
View Plan Results	31

Analysis Functions

Open Reading Frame Search	32
Compatible Ends	32
Mutagenesis Profile	33
Analyze Molecule	34
Open Reading Frame Analysis	34
DNA Repeats Analysis	35
Dyad Symmetries	35
Hydrophilicity (Protein) Analysis	35
Protein Structure Analysis	36

Operations and Utilities

Process Molecule (Invert, Rotate, Circularize, Translate)	37
Mutagenesis Profile	38
Multiple File Conversion Utility	38
DataBook Database	39
Index of DataBooks	39
DataBooks (Tables)	39
DataBook Searches	40
Add to DataBook	40
Export / Import DataBook Contents	40

Clone Manager Suite 7

Getting Started

The Clone Manager Suite 7 contains the latest versions of Clone Manager, Align Plus and Primer Designer. These programs install in the Sci Ed Central program hub and integrate seamlessly to provide a comprehensive set of tools for sequence manipulation and analysis, sequence alignment, and primer design. The central program hub handles file operations, molecule viewing and general utility functions. 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 and molecule information. Use the built-in database to store and organize information about your molecule files.

The Clone Manager program 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. Graphic maps are drawn quickly and accurately and high-resolution maps can be sent to your printer. Or you can copy/paste a graphic map through the Windows clipboard or export as a Windows metafile for use in other applications.

The Align Plus program provides several methods of aligning and comparing DNA or protein sequences. Multiple sequence alignments can be used to determine highly conserved domains within a family or group of related proteins. Rapid scans for similarity can quickly find homologous molecules among large collections of sequences either on your own computer or on your local network. Comparison of two sequences enables the fine details of their similarity to be studied, including identifying frameshift mutations or sequencing errors. A new subclone locator module will help you to locate clones on the primary sequence.

The Primer Designer program provides powerful search functions to find a 'short list' of optimal primers in target regions. Primer Designer will also position a primer in the molecule sequence at a target position, showing you the primer evaluation, the sequence flanking the primer, genes that are nearby, and the location of good neighboring primers. You can also enter a primer at the keyboard, analyze it, and even link it to a molecule. Following primer design, you can save primers to disk files and later use your collection of primers to search for binding sites on any molecule or to create products with the Analyze Mix Wizard.

Install Clone Manager Suite

To install an individual or workgroup copy of the Clone Manager Suite, follow the instructions below. (If you need to install a network copy of the Clone Manager Suite, please refer to the separate instructions accompanying the program disk). Windows should be running on your computer and you should close other open programs before you begin this installation.

- Insert the Clone Manager Suite program CD in your CD-ROM drive
- Allow the disk to autorun or select Start, Settings, Control Panel, and then click Add/Remove Programs.
- Follow the instructions on your screen.

The installation program will ask you to specify the directory where the program files should be copied. Press Next to accept the default directory or press Browse to enter the directory name you prefer. Once the directory has been specified, the program files will be copied. When you use the program for the first time, you will be

asked to designate a home directory. This home directory will be used to hold your personal program configuration settings and user enzyme lists. Two subdirectories will be created below this directory -- one will hold your DataBook (database utility) files and the other will hold your WorkBoxes shortcuts.

Getting Started



Use the Clone Manager Suite icon to start the program. All of the functions of the program are accessed by menus, buttons or shortcut keystrokes. The Toolbar will be an important shortcut to the most common operations. To see a Toolbar button name, rest the mouse cursor on the button for a second or two. For those who prefer the keyboard to the mouse, use the Windows Application key to call up a menu of the toolbar-accessible options for a display window.

Click Help, Help Topics to browse the application help for detailed information about a feature. Use the special Help with View toolbar button (far right) on complex displays to view help directly related to the data display.

Opening Molecule Files



Select Open from the File menu or use the Toolbar button to access. Use this option to load a sequence file for use in the Clone Manager Suite. In the Open dialog box you can use the WorkBoxes button to open the directory containing the shortcuts to your favorite locations and you can add shortcuts using the Add to WorkBoxes button.

If you have placed molecule information on the clipboard (perhaps from an e-mail message), you can import this information using the Paste from Clipboard button. This information must be in one of the standard formats that the program could read from a disk file (like a standard GenBank format or simple ASCII sequence data). For more information about these options and about file formats or help with file loading, please refer to the chapter Molecule File Handling.

When a molecule has been loaded into Clone Manager, it will appear in the molecule viewer window as a working map showing the molecule features and enzyme sites (if any). The molecule will also be added to the Molecule List. Even if you close the molecule viewer window, the molecule will remain on the Molecule List and can be selected for use later in your work session.

Selecting a Molecule for Use



Select Molecule List from the File menu or use the Toolbar button to access. The Molecule List contains a list of the molecules loaded into the program or created during the current work session. Newly-created or modified molecules that have not yet been saved are marked with a diamond symbol.

Use this list to identify the molecule to use for the next operation. Small Molecule List buttons are also found in many dialog boxes, where it might be necessary to change the molecule selected, so that you can change molecules without leaving the dialog box. For more information about the Molecule List, please refer to the chapter Molecule File Handling.

Output Data

Data in the active window can usually be printed, copied to the clipboard, or exported to a disk file. This data can be a molecule map or analysis results or formatted sequence data. Use the toolbar Print button to send the data to your printer.

Send View to Clipboard

Found on the View menu or use toolbar to access. You can copy data from a results window and place this information on the clipboard for use in another Windows program. When you click the Copy View to Clipboard button, the information in the view window is automatically selected and then copied to the clipboard. To place this information in a document in your word processing program, move to the appropriate place in the document and paste the contents of the clipboard into your document (usually Edit, Paste or Edit, Paste Special).

Send View to File

Found on the View menu when a results window is open. You can export the data from a results window and save this information in a file on disk. When you select this option, the information in the view window is automatically selected and saved to disk in the file format specified. To place this information in a document in your word processing program, move to the appropriate place in the document and insert the contents of the disk file into your document (usually Insert, File or Insert, Picture, from File).

User Preferences

Click File, Preferences to select options to set user configuration options. Click the Basic tab to change your home directory or select the default file format for saving molecule files. The home directory is the location where the program will look for personal configuration data and user-specific files and folders (databooks, workboxes, enzyme lists). You can set a file save format the program should use when saving molecule files for the first time. You can save molecule files in the standard Sci-Ed Central molecule format (*.cm5), in a new xml format (*.cx5), or in the older Clone Manager 4 text format. The *.cm5 format is preferred and has a file association set (you can double-click on a *.cm5 file in Explorer and open Clone Manager with this file loaded).

Click the Share Data tab to enter author information or set up a repository location where you can automatically send copies of newly-created molecule files. You can add an author name, date, and notebook reference to the beginning of the Notes field for your Sci Ed Molecule files. This can help you to identify the source of a file or the location of needed documentation. As part of the Notes field, this information can be printed on the page with your molecule map or written to a DataBook entry.

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 SciEd Molecule file or just export the sequence, features, and descriptive information in a file format similar to GenBank.

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 to identify the feature keys that should be imported when you open a file in standard GenBank or EMBL format, and indicate how each feature key should be handled in Clone Manager. You can also tell the program how to handle all other feature keys that are not listed as designated features -- these can be discarded or automatically imported as information-only features.

Sci Ed Central

Molecule File Handling

The Sci Ed Central program hub controls all file handling, molecule viewing, sequence editing, and general utility functions for the Clone Manager Suite. The molecule file handling functions 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 use Work Boxes to store location information for groups of files you want to use often. 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. You can also read molecule data directly from the clipboard. Special, program-specific non-molecule files can also be loaded, when appropriate. To open files containing molecule information:

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



You can open more than one file at a time, using the mouse and the Control or Shift keys. If 5 or more files are to be loaded, you can elect to keep these files together in a group. If you select a file to load that contains multiple molecules, these molecules will automatically be placed in a group for easy handling.

Sequence File Formats

Molecules in the Sci Ed Central molecule file formats (*.cm5, *.cx5) and the older Clone 4 file format contain sequence data, molecule name and size, and may contain any of the following additional information: restriction enzyme sites, features, description, notes. This format should not be edited with a word processor.

Sequences in other standard file formats are imported as double-stranded and blunt-ended. (If the sequence has been identified as a linear molecule and you want it to be circular, you can use the Circularize option during file loading.) Only one strand of a sequence is needed and IUPAC codes for ambiguous bases can be used. A number of sequence databanks and analysis programs will prepare files in a format that is similar (but not always identical) to GenBank or EMBL formats.

If the program cannot automatically load a sequence file, it will identify the closest of the recognized formats (GenBank, EMBL, FASTA or ASCII text). It will try to provide you with information about the problem encountered and provide help with file loading. In the File Help module, the file format and type of molecule are shown at the top left of this information windows. The error is described in the top right. The scrolling data window shows the contents of the file. Usually, you will have one of the common file loading errors. You will want to try and correct the error or split the sequence from the rest of the file and then attempt to reload the file.

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 this program, click File, 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

You can retrieve sequence files from the Entrez system at the NCBI using a built-in web browser interface. You need an internet connection to use this function. Once you locate the file you need, you can download this file to your hard disk to save the complete GenBank file, including references and annotations, or you can import the molecule for immediate use. To find the file you need:

- Click File, Retrieve from Entrez
- In the search bar, leave the control set to Nucleotide if you are looking for a DNA molecule
- In the Search For box, enter the locus name, accession number or descriptive information.
- Click the Go button.
- When the results are listed, find the file you want and click on the accession number link to view the contents of this file.

If the file contains sequence and molecule information in a standard GenBank format, you can directly use this data. To download the file, click the checkbox to the left of the accession number identifier and then click the Save button on the web page. When the download message box appears, select the option to save the file to disk and enter a filename and select the location on your hard disk. You can open this file for use at any time by clicking File, Open and selecting the file.



To import the molecule for immediate use, click the Get Molecule button on the browser window toolbar (not on the web page). If the active page contained molecule data in the GenBank format, the molecule will be imported into the 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 add enzyme sites or just use the sequence in an alignment or primer design operation.

Using Work Boxes

The implementation of Work Boxes is similar to the concept of "Favorites" introduced with Windows 95. The program creates a folder (directory) called "WorkBoxes" and uses this location to store shortcuts you create for folders that you want to use often.

To add a new folder shortcut to Work Boxes:

- Click File, Open or use the Open toolbar button.
- Navigate through the directory structure until you see the folder you want in the upper box called "Look in:"
- Click the button Add to WorkBoxes



To use the shortcuts in Work Boxes:

- Click the button WorkBoxes
- Double-click to open the folder of interest to select a file.



You can use File, Preferences to set a home directory and to set the default file format for saving molecule files. The home directory is the location where the program will look for personal configuration data and user-specific files and folders (databooks, workboxes, enzyme lists). For individual or workgroup license holders, if you have installed Sci Ed Central under Program Files, you will probably want to assign a home directory that is more easily accessible. (The default home directory is C:\SECentral.) Network concurrent license holders should identify a home directory on each local machine.

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 Sci-Ed Central binary molecule file format (*.cm5), a new xml-based molecule file format (*.cx5), or in the older Clone Manager 4 format. The *.cm5 format is preferred and has a file association set (you can double-click on a *.cm5 file in Explorer and open Clone Manager with this file loaded). The new xml format is a human readable format that can make data portability easier. Although this file format is readable, it should not be edited. If you always want to save files in the same format, you can use File, Preferences to set this as a default preference.

Sharing Molecule Data

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). You can use the checkbox to instruct the program to automatically prompt for this information when files are saved for the first time. To set data sharing preferences:

- Click File, Preferences, then click the Share Data tab.
- Enter an author name that should be used for all files generated from the local machine. (You can also leave the Author Name field blank so that multiple users can type in their names.
- Specify the data type and the repository location where the data should be sent.

Author Stamp -- You can add an author name, date, and notebook reference to the beginning of the Notes field for your Sci Ed Molecule files. This can help you to identify the source of a file or the location of needed documentation. Click File, 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.

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. If you are saving a file for the first time and are prompted for author information, you can use the checkbox to send the file to the repository location in one easy step during the file save operation. Click File, Submit to Repository to send a copy of the active file to this location now.

Exporting Molecule Data

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

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

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

Compact GenBank -- Select this option to export essential molecule data in a format that is consistent with the standard GenBank file format. The data will include the molecule name, size, description, features (genes, regions, markers, labels), and sequence data, all with standard GenBank identifiers.

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.

Using the Molecule List



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

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

Selecting a Molecule -- Identify the molecule to use for the next operation or 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.

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 using the Align Plus program or to remove all of these molecules from the program in one step.

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 the New toolbar button
- 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

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

You can enter sequence data (type in or paste) or simply enter the size in bps for your new molecule. To enter sequence data, type in the complete upper strand sequence in the data entry box provided. You can enter valid or ambiguous bases. All sequence 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 Features

You can enter genes, regions, markers, or labels as features for your new molecule. Select the feature type to tell the program how to display this feature on the molecule map or annotated sequence. Enter a short feature name and enter start and end basepair positions (markers have only one position). Use the checkbox to 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.

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.

Molecule Viewer Window

The molecule viewer window provides one location for all of the information about a molecule and is controlled by tabs along the bottom of the window. When you open a molecule, the molecule viewer window will show you a map of the molecule. You can click a tab to quickly move from the molecule map to the restriction map data, sequence, features table, or molecule information (description, notes or GenBank annotations). Use the Help with View button on the window toolbar to access help related to the active view.

The Map tab is the default view and shows a graphic representation of the molecule and its features (genes or regions of interest). The Sequence tab lets you view the molecule 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.

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, resulting in a 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 bp positions for the region to display. The data display is regenerated and the window title identifies the zoomed region.

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

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

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

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

Viewing Molecule Maps

To view the molecule map, open the molecule viewer window. If the map is not displayed, click the Map tab at the bottom of the window to change to this view.

The basic molecule map shows a graphic representation of the molecule, including features (genes or regions). The molecule name and size appear in the window title bar. The map will be either circular or linear, depending on the molecule, and will also show enzyme sites or primer sites, if entered for the molecule. You can use a toolbar button to toggle between displaying enzyme sites and displaying primer sites.

Features

Features that have been entered to the features table for this molecule will appear as follows: Genes appear as pointers below (inside) the mapline; Regions appear as boxes on the mapline; Markers appear as short, solid lines crossing the mapline. Click the Features tab of the molecule viewer window to add or modify features. Click the ORF search button to locate open reading frames and add them to the map as genes, if desired.

Enzyme Sites

Enzyme recognition sites appear above (outside) the mapline. 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 and hold the left mouse button down to see bp position for this enzyme and for any suppressed sites at this location. (Use the Map Sites list on the right to select one of these enzymes.) In highly congested areas, the enzymes sites will be collected in order and moved to a location above or to the side of the map. Multi-cloning sites will usually be displayed in this way. If your map has no enzyme sites, click the AutoScan button to find all single cutters using the specified scanning list. Click Clone, Find Enzyme Sites to locate other enzyme sites and add them to the map.

Primer Sites

Primer sites can be shown above (outside) the mapline in place of enzyme sites (use toolbar button to toggle between enzyme and primer site displays). If your map has no primer sites, you can click Primer, Find Primer Sites to scan your collection of primer files to find primers that will bind to this molecule. If you are designing primers, you can add these to the molecule map using toolbar buttons in the primer search results list window or in the primer viewer window.

Viewing Sequence Information

To view sequence data, open the molecule viewer window, and click the Sequence tab at the bottom of the window. You can view the complete sequence of your molecule or just a part of the sequence and you can select the format for the sequence display. You can copy a portion of the sequence to paste into another application or you can edit the sequence, if needed. You can also find places in the sequence that match a search string you enter or jump to a specified basepair position in the sequence (see below).



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



Go To basepair position -- Click the Go To button and enter the 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 the sequence display style:

- With the sequence visible on the screen, click the Format 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 with My Style.

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 designated AutoScan enzyme list, or enzyme sites that are currently shown on the RMap tab.

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

Sequence -- Single-stranded sequence displays the upper strand for the full sequence and either 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. You can select the block size and case for sequence characters.

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

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

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 or 3-letter amino acid codes.

Sequence Text Colors

You can select the default color to be used for each type of sequence text that can be displayed and you can set the base color and background color for sequence labels. Colors can help to distinguish one type of text (such as amino acid codes) from another (such as the sequence bases). Colors selected for features (gene, region, marker annotations and sequence labels) will apply to all text of that type. You can use a custom feature setting to distinguish a special feature, if needed.

To set default sequence text colors:

- With the sequence visible on the screen, click the Format button
- Click the Colors tab
- Select the preferred color for each item type



To set custom sequence feature colors:

- Click the Features tab to move to the features table
- Select (highlight) the feature
- Click the Customize style 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 Courier New, if this is supported in your printer. Otherwise, the font face will be Courier. (A mono-spaced font is required so that sequence data lines up correctly.) To set printed output options, click the Print Options tab,

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

Partial Sequences

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

- With the sequence visible on the screen, click the Format button
- Click the Partial Sequence tab
- Set the Partial Sequence option button
- Enter the start and end basepair positions for the part of the sequence you want.
- Use the small Features button to see where features start and end.
- 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.

Using the Find Sequence Tools

You can find places in the sequence that 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.

To start a Find sequence operation:

- Open the molecule viewer window and click the Sequence tab.
- Click the Find button
- Enter the sequence of bases you want the program to search for (click the small button for help with IUPAC ambiguous bases).



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 the Find Next button to locate additional places with the same sequence of bases.




Or, to view a list of sites found:

- Set the Search Type to Report All Sites.
- View and/or print the list of sites found.

Advanced search notes: When entering the search string, you can use the check box to indicate that you do not want to expand ambiguous bases. You might want to do this to locate an ambiguous base within your actual sequence data. If you are currently located at any position other than the top of the sequence, you can use a check box option to start your search at the beginning of the sequence.

Editing Sequence Data

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

- Open the molecule viewer window and click the Sequence tab
- Click the Edit Sequence button  on the viewer toolbar.
- Make changes to the sequence data.
- 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.

Cut, Copy, or Paste Sequence Data

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

To copy sequence from the molecule viewer window:

- Select (highlight) the sequence to be copied using the mouse or use Control+A to select the entire sequence.
- Look in the toolbar area to see the range of bases selected.
- 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.
- 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.
- 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.

Viewing Molecule Features

Found on the Features tab in the molecule viewer window. Molecule features appear in the listing with a symbol for the feature type, the feature name, the start and end basepair positions, a brief description of the feature, and the feature key, if assigned. Click the column headers to sort the features table by name, position, descriptive text, or key. Click the Filter button to set options to limit the data displayed by location or feature type. Click the Info Features button to show or hide the display of information-only features. Click toolbar buttons in this window to add, edit or delete features (see below), view feature properties, create a protein from a gene, or set custom style settings for an individual feature. To add a new feature:






- Click the Add Feature button
- Enter the feature information.




To edit feature information, select (highlight) the name of the feature, click the Edit Feature button, and then modify the feature information, as required. To delete a feature, move the highlight to the name of the feature you want to remove and click the Delete Feature button

Feature Types


Feature 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, Markers, Regions, Labels and Information-only features. Most features have a start and end base pair position and can be on the complementary or normal strand.

-  **Genes** -- Genes usually represent translated sequences. Genes appear on the molecule map inside/below the mapline and can be displayed on the annotated sequence.
-  **Regions** -- Regions are zones of special significance for the user. Regions appear on the molecule map on the mapline and can be displayed on the annotated sequence.
-  **Markers** -- Markers have a direction and occupy only one base position. Markers appear on the molecule map on the mapline and can be displayed on the annotated sequence.
-  **Labels** -- Labels are areas within the sequence that are of special interest to the user. Labels can be shown on the annotated sequence, but not on the molecule map.
-  **Info only** -- Information-only features are not shown on the molecule map or on the annotated sequence. They do not appear on the Features Table list unless you click the Info Features toolbar button to temporarily show this feature type. Use for features where you want to preserve the information about the feature but do not want it to clutter your map or features table.

Viewing Feature Properties

 Found on the Features tab in the molecule viewer window. To view feature properties select (highlight) the name of the feature and click the Feature Properties button. The information displayed will depend on the feature type. 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 other feature types, properties include basic information about the feature defined and its base composition, if applicable.

Using the Features Table Filter

 If your molecule has a very large number of features, you can click the Filter button to set filter options to transiently limit the data displayed in the features table by location or feature type. When a filter has been applied, the Filter toolbar button will be depressed. Click the depressed Filter button to remove the filter.

To 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. To 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

Found on the Features tab in the molecule viewer window. Below the features table list, a boxed area is used to show the styles set for the selected (highlighted) feature. If you have not set any custom styles for this feature, the settings will reflect the default values for printed map features or for annotated sequence display. To assign custom styles:

- Select (highlight) the name of the feature.
- Click the Customize Style button
- Set the style and color to be used for this feature for the printed map.



- Or set the text color to be used for this feature in the annotated sequence display.

To turn off the custom style settings, click the Customize Style button and then click the Clear button to set the styles back to your default values.

Viewing Molecule Information

Found on the Info tab in the molecule viewer window. You can view or edit the molecule name, size, description, notes. Click toolbar buttons in this window to add author information to the notes field, change the starting basepair number for this molecule, or access GB Annotations (header information) if the file is a GenBank or EMBL format file. You can also view molecule properties information in the lower right section. For DNA molecules, you can see base composition. For protein molecules, you can see molecular weight, pI, and amino acid composition. Molecule descriptive information includes:

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 (add or remove bases) to change the size of the molecule.

Description -- appears on many types of printed output and can optionally be printed on molecule map pages.

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



Found on the Info tab in the molecule viewer window. You can add an author name, date, and notebook reference to the beginning of the Notes field for your Sci Ed Molecule (*.cm5) files. This can help you to identify the source of a file or the location of needed documentation. To add author information, click the Author Stamp button. Enter your name and a notebook reference, if applicable.

Viewing GenBank Annotations



Annotations refers to the header information found in a GenBank or EMBL format file. In addition to reference information, you may find some helpful feature information or comments about the file that you may wish to review. 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 save a link to the original GenBank file so that you can later find this information.

Base Renumbering



Found on the Info tab in the molecule viewer window. You can designate a number other than 1 to use as the first basepair number for the sequence of the active molecule. You might want to renumber your sequence if you are using part of a larger molecule and want to maintain consistent base numbers for features or enzyme sites. To assign a starting basepair number, click the Base Numbering button and enter the starting basepair number for this molecule. Negative number are permitted, but zero is not a valid base number.

Clone Manager 7

Molecule 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. The Clone menu contains the commands used for graphic map functions and is accessible when a molecule map is in the active window.

Working Map Display

Click the Map tab in the molecule viewer window to view a working map of the active molecule. The map will be 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 an enzyme site, cut at this site, delete this site from the map, or access information about enzyme suppliers or isoschizomers. Other toolbar buttons let you set options for printing the graphic map or start operations that will search for enzyme sites or open reading frames.

Features

Features that have been entered to the features table for this molecule will appear as follows: Genes appear as pointers below (inside) the mapline; Regions appear as boxes on the mapline; Markers appear as short, solid lines crossing the mapline. Click the Features tab of the molecule viewer window to add or modify features. Click the ORF search button to locate open reading frames and add them to the map as genes, if desired. You can use your mouse to get feature information. Point to a feature name and depress the left mouse button to see feature basepair positions, strand.

Enzyme Sites

Enzyme recognition sites appear above (outside) the mapline. 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 and hold the left mouse button down to see bp position for this enzyme and for any suppressed sites at this location. (Use the Map Sites list on the right to select one of these enzymes.)

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

Working with Map Sites

The Map Sites list is found on the right side of the molecule map display and shows a scrolling list containing each enzyme site or primer site entered for the molecule. By convention, the enzyme site is marked at the left end of the recognition sequence for that enzyme, the primer site at the 5' end of the binding site. Click on the column header to sort the display by name or position. You can view the properties for a site, delete the site, or

clear the entire list of sites. Select the site you want to use by clicking on the site on the map display or on the list of map sites at the right. For enzyme sites, you can also cut at the site or add an enzyme site to the list.

Enzyme Sites / Primer Sites -- Click the Enzyme Sites/Primer Sites button to switch from the display of enzyme sites to the display of primer sites. Click the depressed button to return to the display of enzyme sites.

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

Deleting a map site -- Click the Delete Site button to delete the selected map site.

Wiping the list of map sites -- Click the Wipe All Sites button to remove all of the enzyme or primer sites from the molecule map in one easy operation. If enzyme sites are being displayed, Wipe All Sites will remove all of the enzyme sites. If primer sites are being displayed, Wipe All Sites will remove all of the primer sites.

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

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

Enzyme Suppliers -- Click the Enzyme Suppliers button to view a pop-up display listing the companies that supply the selected enzyme. This supplier information is a part of the REBASE information updated monthly.

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

Adding Enzyme Sites -- If you do not have sequence data, you can enter an enzyme site manually. If the molecule has some sequence data at the point where the enzyme site is entered, it will be checked for accuracy before the site is accepted. (For molecules with complete sequence data, it is recommended that you use the AutoScan or Find Sites options. The program will scan the sequence and find enzyme sites for you.)

Circular vs Linear Maps

When a molecule is displayed in the molecule viewer window, the map will be circular or linear, depending on the molecule. You can modify the molecule to change this property, if required. If you have a circular molecule but want to draw a linear map, cut the molecule after the last base using the cut at basepair option (found on the Clone, Cut menu). This will open the molecule at the origin. If you have a molecule that has come into the program as a blunt-ended linear molecule, but it should be circular (perhaps the descriptor 'circular' was missing from the file), circularize the molecule using the Process Molecule command found on the Operations menu. If a linear molecule has cohesive ends, you can self-ligate the molecule (if its ends are compatible) using the Ligate option on the Clone menu.

High-Resolution Graphic Maps

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



Click the Map Print Options (format) button to change the appearance or content of the high-resolution graphic map. You can turn on or off the printing of types of basic text information, select feature styles and colors, select enzyme or primer site styles, or change the size or position of the graphic map or add descriptive text to the map page. Changing map print options will not affect the appearance of the working map in the molecule viewer window (with the exception of feature colors), but will affect printed, copied or exported maps.

You can preview or print the high-resolution graphic map using options on the File menu. You can copy or export the graphic map using options on the View menu (Send View to Clipboard, Send View to File). Copied or exported graphic maps are vector graphic files (*.wmf or *.emf) which can easily be scaled in their destination application. When pasting or inserting these files, select Enhanced Metafile where possible (use Paste, Special). Click Clone, Send Map and then select the destination for the map.

Control of Map 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 Map Print Options button and then click the Map Basics tab.
- Set one typeface (font) for all text on your map. Arial and Tahoma are simple, modern, sans serif typefaces, 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.
- Select a color and weight for the mapline.

Control of Map Features

You can control the style and color for feature types (genes, regions, and markers) to be printed on your graphic map. These style and color settings 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. 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 Map Print Options button and then click the Features tab.
- Use checkboxes to indicate which feature types to print.
- Select a style and color for each feature type.

To set custom feature settings:

- With the map visible on the screen, right-click on a feature and select Customize
- Select the style and color for this one feature or Clear to remove this custom setting.

To shift a feature name:

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

Suppress very small features -- If you are using large molecules (greater than 100,000 bp), you can elect to suppress the drawing of features that are very small relative to the size of the molecule. Use the checkbox on the Features tab to enable this option.

When enabled, the program will suppress drawing the feature or the feature name using the following rules: If the gene is less than 1/500 (0.2%) of the molecule size, the feature name will not appear. For circular molecules, if the gene is less than 1/2500 (0.04%) of the molecule size, neither the feature (arrow, box, pointer, etc) nor the feature name will appear. For linear molecules, if the gene is less than 1/1000 (0.1%) of the molecule size, neither the feature nor the feature name will appear.

Control of Map Sites

You can select one of two preset font sizes for enzyme and primer names and you can control the color to be used for these names and tick lines. You can also indicate if the program should print basepair numbers for each enzyme or primer site or if it should italicize the first three letters of the enzyme name. To set enzyme print styles:

- Click the Map Print Options button and then click the Sites tab.
- Set the font size for enzyme or primer names to Standard or Small.
- Select a color for names and tick lines.
- Use the checkboxes to print basepair numbers or italicize enzyme names.

Control of Map Content

You can print stand-alone graphic maps, or you can select a pre-defined map format that adds additional text information to the printed map page. If you do not add additional text, you can move or resize the graphic map before printing or copying to the clipboard. In almost all cases, the default map size and position will print a complete, centered, high-resolution graphic map in either landscape or portrait orientation. (To change the page orientation, select Print Setup from the File 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 Map Print Options button and then click the Content tab
- Select Molecule Map and Text.
- Use checkboxes to indicate which text you would like to include.

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

Changing Map Size -- You can change the size of a graphic map before printing. You can also resize a map before copying to the clipboard or exporting. To change the size of the graphic map:

- Click the Map Print Options button and then click the Content tab.
- Select Molecule Map Only.
- Select 80% or 90% to decrease the map size.
- Or select 110% or 120% to increase the map size.

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

application and enzyme names will not get as small and map lines and enzyme tickmarks will not get as thin as when a larger map is reduced.

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

- Click the Map Print Options 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%).

Graphic Map Output

You can use options on the Clone, Send Map menu to preview the map in its printed format, send the map to your printer, copy the map to the clipboard, export it to a disk file for use in other Windows programs, or save the map in a format that can be used on web pages. You can also print a list of features or enzyme sites found on the map.

Preview -- Click Preview to check the appearance of your map before printing, copying to the clipboard, or exporting to a disk file. To change the orientation of the printed page, Click File, Print SetUp.

Print -- Click Print or use the toolbar Print button to send the map to your printer. You can click Map Print Options to change the printed size and position of the map, colors and styles for features, and basic text options.

Clipboard -- Click Clipboard or use the toolbar Send View to Clipboard button to copy the high-resolution graphic map to the Windows clipboard for use in other Windows applications. The styles and colors set for printing will be used for clipboard output. To paste your map in a Word document, open your word processing program and move to the place where you want your map and then click Edit, Paste Special and select Enhanced Metafile.

Export -- Use this option to export the high-resolution graphic map as an Enhanced Windows Metafile (*.emf) or a Windows Metafile (*.wmf) 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. To use your map metafile in a Word document, open your word processing program and move to the place where you want to insert your map and then click Insert, Picture.

Print List Map -- Use this option to print a list of the features and enzyme sites which appear on your graphic map. Features are listed in a table format in order by basepair position of the start of the feature. Enzyme sites are listed by site basepair position in ascending order (reading from left to right). You can also print this information on the same page as the graphic map (control map content).

Save for Web -- Use this option to prepare a special raster graphic map in the Portable Network Graphics (*.png) format, particularly useful for computer screen display (such as on a web page or in an on-screen presentation). For this format, the map size is set to 80%, bp interval numbers are turned off, and you cannot set italics or small fonts for enzyme names. Solid feature styles and bold colors work best, as does the Arial type face.

Clone Manager 7

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 or use toolbar button to access. 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). You can also rebuild the map sites lists for a molecule by rescanning with enzymes currently listed.

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 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 Sites and then click List of Enzymes.
- Click Single Cutters, Double Cutters or All for 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.

The Rebuild 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.

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 Operations menu or use toolbar button in map window to access. Also available during import of files that contain no enzyme site information (such as GenBank or EMBL files). You can use this operation to scan a molecule using the AutoScan-designated enzyme list, looking for single cutters. Enzyme 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 add the new sites.

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

Enzyme Lists

Clone Manager 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 new lists to meet your individual needs. You can also update the program-defined enzyme lists monthly using REBASE enzyme files so that you are always up-to-date with new enzyme information. Three types of enzyme lists are used by the Clone Manager Program: program-defined lists, transient lists, and user-defined lists.

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

Program-defined Enzyme Lists

Clone Manager 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 enzyme list is a subset of the master enzyme list (All Enzymes). Duplicate enzymes with the same recognition sequence and cut position are eliminated. This list is the program default for most enzyme operations. This list contains the major, commercially-available enzymes recognized by the Clone Manager Program (approximately 200 enzymes).

All Enzymes -- This is the master list containing all of the enzymes recognized by the program. There are more than 500 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.

Transient Enzyme Lists

You can create a transient enzyme list just when you need it and discard the list at the end of the current operation. You can specify the characteristics for the enzymes you want to include and indicate a supplier preference. To create a transient enzyme list:

- Start the operation you want to perform.
- When asked to select an enzyme list, pick On-The-Fly.
- 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.

User-Defined Enzyme Lists

User-defined lists are enzyme lists you create to make collections of enzymes you routinely use. Once your enzyme list has been created, you can later add or remove enzymes, as needed, and you can share these lists with others using an export/import procedure. The program provides a wizard to help you through the process of building new enzyme lists. To create a new enzyme list:

- Click Clone, Enzyme Lists
- On the Enz Lists tab, click Build New
- Follow the wizard instructions.

Share User Enzyme Lists -- You can export or import user enzyme lists, as needed, to share a list with a colleague or to move your list from one computer to another. You can also rename a user enzyme list, if required. To export, import, or rename a user enzyme list, click Clone, Enzyme Lists. On the Enz Lists tab, click Share and then select the operation you want to perform.

Viewing Enzyme Lists

Select this option to look at or print any of the enzyme lists or to see information about particular enzymes. You can also add to or remove enzymes from your user enzyme lists. To edit a user-defined enzyme list, clear the Read Only check box and click Add Enz or Del Enz, as needed. To view an enzyme list:

- Click Clone, Enzyme Lists.
- On the View List tab, select the enzyme list in the List box.

You can also enter detailed information about special enzymes (new, non-commercial or historical enzymes), if needed. These enzymes can then be used to find recognition sites or cut molecules. You can also view or delete these special enzymes. To view the list of special enzymes, click Clone, Enzyme Lists and then click the Special tab. To add a special enzyme, click Add New and enter the information required.

Updating Enzyme Lists

You can update the program-defined enzyme lists monthly using REBASE enzyme files so that you are always up-to-date with new enzyme information. Each time you do an update, the All Enzymes, Commercial (Main) and Scanner enzyme lists will be regenerated. The update procedure will not discard your user-defined enzyme lists or any special enzymes that you may have entered. To update your program-defined enzyme lists using the latest REBASE data file:

- Click Clone, Enzyme Lists, and then click on the Update tab.
- Click on the option for the REBASE file source and complete the required information.
- Click Update Now.

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. You can get the latest REBASE data files directly from the REBASE web site, or you can subscribe to the email service. In either case, you will have to identify the disk file and its location during an update procedure. Alternatively, you can let the program get the file for you by dialing out over the Internet using anonymous ftp. This file is accessed directly during the update procedure.

REBASE Files on the Web -- To get the latest REBASE data file from the World Wide Web, connect to the REBASE web site: <http://rebase.neb.com>, click on the REBASE Files icon, click on file format #35, Clone Manager and then click Get File. Save the data file to your computer.

Getting REBASE Files via FTP -- Set this option on the Update tab of the Enzyme List dialog box during an update procedure. The REBASE data file will be used during the update and discarded. It is not saved to your hard disk. Note: FTP access uses TCP/IP internet access protocols. You may need to install remote-access-services and/or dial-up networking.

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. You can print the data in this view window, copy this view to the clipboard, or export the view data to a disk file.

Restriction enzyme recognition site information is listed in table format or displayed in a graphic format. Click the data column header 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). Use toolbar buttons to get enzyme properties, suppliers, isoschizomers, or compatible ends, initiate a cut operation or enter the sites to your molecule map. Use the Filter button to set options to show you just the information you need.



Click the No Cutters button to see a list of enzymes that do not cut within the active molecule (or that do not cut within the zoomed region of the molecule). Click the depressed No Cutters button to return to the prior data display.

Properties -- The pop-up display shows the recognition sequence, cut position, overhang created, and the type of methylation sensitivity (if applicable).

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

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

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

Cut with Enzyme -- A copy of the molecule will be cut at all of the sites found for this enzyme, resulting in one or more fragments. You can enter fragments to the molecule list for immediate or later use.

Enter Sites to Map -- All of the recognition sites found for this enzyme will automatically be added to the molecule map.

Restriction Map Format Options



In the restriction map window, click the Format button to change the enzyme list, data display style, or sort order. 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 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. You can use the toolbar buttons to change the display as needed when viewing data.

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 format options dialog box or the restriction map window toolbar.

Methylation Sensitivity Alert -- Some enzymes are sensitive to methylation. If you want to be made aware of these enzymes, click the Format button and then check the Mark Methylation Sensitivity box. The names of the methylation-sensitive enzymes will be shown in color.

Restriction Map Display Styles

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

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.

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

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.

Cut N times -- Check the box for this option and then set the value for N, where N is the number of times the enzyme cuts the molecule. Indicate if the number of cuts should be equal to, less than or equal to, or greater than or equal to N.

Cut Outside Region --- Check the box for this option and then enter the two basepair position numbers for No cuts between bps (start) and (end). Use this option to designate a region you would like to protect.

Cut Inside Region -- Check the box for this option and then enter the two basepair position numbers for Must cut between bps (start) and (end). Use this 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.

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

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

Clone Manager 7

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.

Cutting Molecules



Found on the Clone menu or use button on main toolbar to access. You can cut a molecule by selecting an enzyme from an enzyme list and cutting at all sites for this enzyme. (You can also cut with 2 or 3 enzymes at the same time). Or you can cut at one specific enzyme site already marked on your molecule map or at specified 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. If the recognition sequence is longer than the space provided and you want to be able to see the entire sequence of bases, double-click on the column header divider (right after the word Recognition). You can now use the horizontal scroll bar in this list box to scroll to the right to see the rest of the recognition sequence.

Cutting at 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 an enzyme site on the working map and click the Cut at Site toolbar button. If you want to check a cut position first, select the enzyme and then click the Site Properties toolbar button to see the exact cut position and ends produced.

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 colored horizontal bar above the icon map shows the approximate location of the highlighted fragment. The enzyme producing the left end of the fragment is shown to the left and above the 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.

Ligating Molecules



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

Use the buttons to the right of a molecule window to cut a molecule or modify its ends, as needed, to make the ends compatible for ligation. Use the Features or CompEnds buttons (lower left) to view the features tables for the molecules or to find enzymes that could prepare ends that would be compatible to the molecule end you select. When ready, click the Ligate button to join all compatible ends or the Partial button to ligate only the two molecule ends you designate.

Changing molecules -- Use the molecule list button to the right of the molecule window to select a molecule for this window or to change the molecule currently shown. Note: To create a new molecule during a ligate operation (perhaps a short linker), minimize the ligate window, then select File, New and enter the information and sequence for the new DNA.

You will want to put the vector in the upper window -- after ligation, a circular recombinant will be rotated to restore the original reference frame of the vector (in the upper molecule window). If needed, you can flip the contents of the two windows. Use the double arrowhead button to the right, between the two molecule windows, to put the lower molecule in the upper window and the upper molecule in the lower window.

Cutting the DNA -- Use the cut button on the right to access the Cut Molecule operations for this 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 on the right to access the Modify Ends operations for this molecule. You can use standard or custom end modifications to prepare molecule ends for ligation.

Ligation Results

When you click the Ligate button in the Ligation working window, a ligation will be performed if the ends of the two molecules of DNA are compatible. All possible ends will be ligated.

One possible result -- If only one ligation orientation is possible, you will be ready to enter the new molecule to the Molecule List.

Two possible results -- If you added a piece of DNA that could be inserted in more than one orientation, you can select the orientation you prefer. The insert is marked in color below the molecule icon map, with an arrow indicating direction. Use the Enter1 or Enter2 buttons (lower left) to enter the product with the desired orientation to the molecule list. If you want to save both products, enter each to the molecule list sequentially.

Partial Ligation

When you click the Partial button in the Ligation working window, a partial ligation will join only the two molecule ends you select. The resulting product will be linear. Set the option that joins the ends in the orientation you want, using the drawings to visualize the ligation event. After you add the new molecule to the molecule list, you can return to the Ligate display to select additional ligation products.

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).

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 gray 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.

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 the Plan Cloning toolbar button
- 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 small molecule list button to select another molecule if the one shown is not the one you want to work with. Use the small 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 small molecule list button to select a molecule that you want to use as the vector. The Wizard can only use circular vector molecules. If the vector molecule contains a multi-cloning site (shown on a molecule map as a grouped display to the right or left of the map), the basepair positions for this region will automatically be entered for you.

You can use the small 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 you routinely use a vector molecule and want to remember the region you clone into (assuming it's not a multi-cloning site), you can enter this region into the features table of the molecule.

Select an Enzyme List

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

The enzymes used to cut the insert are shown in the central columns (in color), with the insert size in the center area. The enzymes used to cut the vector are shown at the right or left of the insert, producing the compatible ends needed to ligate the recombinant. Square brackets indicate a blunt cut, k marks the use of Klenow fill-in. The orientation of the insert is shown as > (clockwise), B (both), or < (counter-clockwise). An exclamation point at the far right indicates that one or more of the enzymes used was not available from your preferred supplier.

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

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

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

Clone Manager 7

Analysis Functions

Clone Manager can search for Open Reading Frames and easily enter found ORFs as genes in the features table for the molecule, and it can provide information about cohesive ends to help you select enzymes for use. A mutagenesis profile is now available in Clone Manager -- you can use this display to show silent changes, find 'almost' enzyme sites, or change one amino acid for another. A set of analysis functions allows you to find direct or inverted repeats and dyad symmetries for DNA molecules. You can also 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 available as part of the Analyze Molecule option.

Open Reading Frame Search



Found on the Operations menu or use toolbar button in map window to access. Use this option to locate Open Reading Frames for the active molecule. You can indicate if you require that each ORF start at an ATG or GTG codon, and you can 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). Click on the column header to sort by the value in that column. The red arrow above the molecule map shows the approximate location of the highlighted ORF. The arrow points right for normal strand, left for complementary strand ORFs. Use the Enter Gene option to enter the highlighted ORF as a gene in the features table for this molecule. 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.

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 display, the restriction map data display, while viewing enzyme lists, or while working in the ligate window.

The list of enzymes that could produce compatible ends is limited to prototype enzymes and neoschizomers. That is, 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 the Compatible Ends button. 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 -- 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).

Enzyme Lists -- When you are viewing enzyme lists (Clone, Enzyme Lists, View List), you can click on an enzyme and then click the Enz Info button. Select Compatible Ends to view enzyme-specific compatible ends information, as described for Restriction Map data (above).

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 Comp Ends 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 Operations 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. (Please see the Operations chapter (last in this guide) for a description of the mutagenesis profile display screen.)

Show Silent Changes



You can use this option to see which bases in the sequence you could change without changing the translation shown. When you click the Show Silent Changes button, 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. Make a base change. If the change removes an enzyme site, the name is crossed out with a colored line. If an enzyme site is added, the new enzyme site appears in color. Click the depressed Show Silent Changes button to clear this display.

Find 'almost' enzyme sites



You can use this option to find 'almost' enzyme sites (one-base mismatches to the recognition sequence of a restriction enzyme). The sequence will be scanned using restriction enzymes recognized by the program (commercially-available, 6 base cutters with no ambiguous bases) and a list of 'almost' enzyme sites constructed. The list shows each single base change that could introduce an enzyme recognition sequence, the position for the change, the name of the enzyme site that would be introduced and the type of end produced by cutting with this enzyme. You can sort this list by position or enzyme name by clicking on the column header. 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 the Change Amino Acid button 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.

After you have made changes to the bases shown in this display (introduced a mutation), you can use the Create Mutant toolbar button 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.

Analyze Molecule

You can do some basic analyses on your DNA molecules. You can analyze a molecule to find open reading frames, viewing an ORF map of start and stop codons, which help to identify potential protein coding regions. A numerical value (Fickett's TESTCODE) can help you to determine if a region is likely to encode a protein. You can analyze direct or inverted repeats and dyad symmetries to identify regions that might be associated with biologically active sites such as control regions, transcription terminator regions, or RNA processing sites.

You can also do some basic analyses on protein (or translated DNA) molecules. You can analyze protein structure for alpha-helix, beta-sheet (extended), beta-turn and coil (random) states using the method of Garnier-Osguthorpe-Robson. You can use hydrophilicity analysis with beta turn prediction to assist in identifying antigenic regions. When viewing analysis results, use the Help with View toolbar button to view help related to the analysis displayed. To analyze your molecule:

- Click Operations, Analyze Molecule
- Select the type of analysis
- 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 analysis results with Any, ATG or GTG start codons. Use the toolbar buttons to select Text or Graphic displays. Use the Redefine button to return to the Analyze Molecule setup dialog box to select another analysis type or another region of the molecule to analyze.

Text Display -- for each open reading frame, the program reports start and end 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. You can sort the list of open reading frames found by frame, length (in descending order), start position, or score (in descending order). You can also use the Add Feature button to enter the selected ORF as a gene in your molecule's features table.

Graphics Display -- an ORF map shows all 6 possible reading frames. Open reading frames that exceed the minimum size specified and that exceed 5% of the analysis region length are shaded in pale blue. Full height black bars mark terminators. Half height colored bars can be used to mark start codons.

DNA Repeats Analysis

You can analyze direct or inverted repeats (and dyad symmetries using a separate option) to identify regions that might be associated with biologically active sites such as control regions, transcription terminator regions, or RNA processing sites. The repeats found are shown in a table display. Use the toolbar buttons in this window to move between the three repeats displays (direct, inverted or palindromes) or return to the analyze set-up dialog box (Redefine). You can sort the list of repeats by length (in descending order), or by position of the first or second repeat. If the number of repeats found exceeds 100, the program will eliminate from its list all repeats containing the smallest number of bases (for example, all 4-base repeats may be dropped).

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

Palindromes -- Click the Palindromes toolbar button in the Analyze Repeats display window to view. The program identifies inverted repeats of four or more bases that are not separated by any unrelated sequence. Such DNA sequences read the same on both strands.

Dyad Symmetries

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

This option identifies sequences which could form stable stem-loop or hairpin structures. The program looks for sequences that are repeated in the opposite direction on the same strand and then uses RNA energy parameters to determine if each stem loop would be stable. The program reports the largest and most stable duplexes which could be formed in descending order (up to a maximum of 100 duplexes with a negative delta-G value). Please note that G:U basepairs can be stable under these circumstances and may be a part of duplexes found. You can use the Redefine toolbar button to return to the Analyze Molecule set-up dialog box to select another analysis type or another region of the molecule to analyze.

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. Use the Redefine button to return to the Analyze Molecule setup dialog box.

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

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

Kyte & Doolittle Analysis

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

Hopp & Woods Analysis

Hopp & Woods values fall within a range of -3 to +3, with hydrophilic residues having a positive score. The most hydrophilic residues have a value of +3.0. On the graphic display, values above the axis line are hydrophilic; values below the axis line are hydrophobic. Hopp & Woods derived hydrophilicity values from a study of antigenicity and adjusted the values to maximize the accuracy of predicting antigenic determinants. Each value is the average of the values of 6 adjacent residues and is plotted at the middle point. The range of values is approximately ± 3 relative units.

Surface Exposure Analysis

The surface exposure value is presented as a proportion of the residue which is exposed on the surface of the protein. These values fall within a range of 0 to 1.000. The most exposed amino acid has a value of 0.97 (lysine). On the graphic display, peak values which fall above the axis line are predicted to be exposed on the surface of the protein. Surface exposure analysis uses the data of Janin, et al, which provides values representing the fraction of residues of a given amino acid that have a surface area of greater than 20 angstroms squared. High values therefore represent amino acids that are likely to be exposed on the surface of the protein. Plotted values are the average of 6 residues and are plotted at the middle point.

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

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

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

Sci Ed Central

Operations and Utilities

The general utility functions of the Sci Ed Central program hub include molecule operations like rotating, inverting, circularizing or translating a molecule, a mutagenesis profile, and a multiple file conversion utility.

Sci Ed Central also includes a simple, built-in database called DataBook that you can use to enter or edit information about your molecule or primer files. You can add molecules or primers to the database without additional data entry and later search the entries to find the file you need.

Process Molecule

Found on the Operations menu. You can invert the active DNA molecule, or rotate a circular DNA molecule. If a DNA molecule was brought into the program as a linear molecule (perhaps because the "circular" descriptor was missing), you can circularize the molecule. You can also translate all or a part of a DNA molecule. When the process has been completed, you will want to enter the new molecule to the Molecule List, changing the name or description, as needed. The inverted, rotated, circularized, or translated molecule exists only within the current program and should be saved to disk if needed for further use.

Invert Molecule

Use this option to invert a copy of the active molecule, changing the positions of all sites and features. The following information will appear as default entries: Molecule name: i-(original molecule name); Description: inverted (original molecule name).

Rotate Molecule

Use this option to rotate a circular molecule to bring a different point to the top of the map. The program will rotate a copy of the molecule, changing the positions of all sites and features. To rotate a molecule, enter the basepair position which should appear at the top of the map (the new origin of the molecule). The following information will appear as default entries: Molecule name: r-(original molecule name); Description: (original molecule name) rotated to position (n).

Circularize Molecule

Use this option to blunt-end ligate a linear molecule to make it circular. (If the molecule has cohesive ends, the upper strand of sequence will be treated as if it were blunt-ended.)

Translate Molecule

Use this option to translate the active DNA molecule to create a protein. Translation will proceed from the starting point specified, on the strand specified, and continue until a stop codon or the end of the molecule is encountered. 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. The following information will appear as default entries: Molecule name: protein; Description: translation of (original molecule name) starting at bp (n).

Mutagenesis Profile

Found on the Operations 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, 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.

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

Translation -- One frame of translation is shown. You can use the toolbar control to change the frame, if necessary. If you 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, click on the toolbar A, C, G or T buttons or press these letter keys on the keyboard to change the base to the letter key pressed.

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.

Multiple File Conversion Utility

You can convert molecule files in standard Clone Manager file format (*.cm5) to another file format for sharing with others, archival storage, web applications, or use by other programs. The file conversion utility will attempt to open and convert the individual files you specify and all Clone Manager files found in the folders you specify, including those contained in subfolders. To convert multiple files to a new format:

- Click File, Multiple File Conversion.
- Select the file format you want to convert to.
- Click Next to 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.
- Select the location (folder) where you want the program to save the converted files.
- Click Finish to complete the setup and begin the file conversion process.

You can save files in compact GenBank format (appropriate for sharing with others), commented GenBank format (good choice for constructing a repository of your important files), FASTA format (for use in web applications) or in Clone Manager XML format (for sharing data across applications).

The DataBook Database



Found on the File menu or use toolbar button to access. DataBook is a simple, built-in database you can use to organize and store information about your molecules or primers. It consists of databooks (tables) you create for your projects and an index to these databooks. Use the DataBook Index to open existing databooks or to create a new databook. From the index page, you can export all of the information in one of your databooks in order to move data to another computer or to your favorite database application.

Adding information to a databook can be really easy -- when you open a molecule file or primer file and then click the Add to DataBook button, 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.

While you have a databook open, you can search most of the fields to see which files contain a gene or a keyword in the description or comments, or you can print lists of the molecules in the database. And when you want to find a specific file for use, you can go directly to the DataBook, where you have all the information you need to identify just the file you're looking for.

Index of DataBooks

The DataBook Index shows a listing of all your databooks (tables). You may want to create new databooks at the beginning of new projects so that you will have a convenient place to collect information about molecules or primers in use or newly created. To open one of your databooks:

- Select (highlight) the databook
- Click Open DataBook



To create a new databook, click the Add New Index Entry button and then enter databook name, brief description and your name or initials. You can also change the information about the selected databook (click Edit Index Entry), remove the databook listing (click Delete Index Entry), or export or import the contents of the databook (click Export / Import).

Databooks (Tables)

Each databook is a collection of entries describing a group of molecules or primers. These molecules may be related to each other, or stored in the same location, or in use by the same investigator. You can add entries to a databook using a manual or automated procedure. When you open a databook, the list display screen shows a part of the entry information in a one-line, abbreviated listing. Click on column headers to sort. You can view, edit, copy or delete an entry, or you can open a file for use (see below). You can also search a databook to find entries containing specific text.



Use the List / Form button to toggle between the abbreviated list view and the editable form view of the database entries. In the Form view, use the navigation buttons (arrows) to move to next or previous entries, as needed.

View or Edit Databook Entries -- Select (highlight) the entry you want to view or change in the list view and click the Edit Entry button. (In the form view, just move to this entry using the navigation buttons.) You can modify any information in the database entry without changing your original file in any way. Move to another entry or return to the list view (click the List / Form button) to save these changes. Press the Escape key to abandon the changes you have made.

Copy or Delete Databook Entries -- Select (highlight) the entry you want to copy and click the Copy Entry button. The entry will be copied and you will be prompted to indicate which databook you want to paste the

copy of the entry into. Select (highlight) the entry you want to delete and click the Delete Entry button. The entry will be removed but the original molecule file will not be affected.

Open a File for Use -- Select (highlight) the name of the file you want to open in the list view or move to the entry in the form view and then click the Open File for Use button. For molecule files, the molecule will be loaded and you will automatically move to the molecule viewer window, ready to work. For primer files, the primer will be loaded and you will move to the primer viewer window. The database utility will close when you open a file for use.

DataBook Searches

You may want to search databook entries to locate files with certain genes or with specific text in the Comments or Description fields. To search the entries in a databook:

- Open the databook you think contains the entries.
- Click the Search button
- Type in the text you want the program to search for
- Mark the fields you want searched.



The search function will search all the fields specified, looking for a perfect match for the characters you typed in. Uppercase or lowercase letters will not affect the search. Entries containing the search string will be marked with a symbol in the second column (after File Name). These entries will appear at the top of the entries list (the program automatically sorted on this column for you). To change the sort, click on another column header.

Add to DataBook

You can add a molecule or primer to the built-in database (DataBook) 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 file or primer file will not be changed). To use Add to DataBook:

- Close the DataBook module (if open).
- Open the molecule file.
- Click the Add to DataBook button (main toolbar)
- Identify the databook you want
- Add or modify information in the other data entry fields, as needed.



To enter a new record manually, start the DataBook module and open the appropriate databook. Click the Add New Entry button and fill in the information about your molecule or primer.

Export / Import DataBook Contents



You can export or import complete databooks, as needed, to offload data for another database application or to move your data from one computer to another. This option can be accessed from the DataBook Index screen.

Select the Export option to prepare a tab-delimited text file of all the entries in the selected databook. Programs like Microsoft Access or Microsoft Excel can open this file, bringing your data into their formats and allowing you to prepare custom reports or perform sophisticated data operations. Select the Import option to read in the data in a file previously exported using the Export DataBook option (above).