Clone Manager

for Windows

Professional Edition

Volume 2
Alignment, Primer Operations

Version 9.5

Getting Started

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

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

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

Table of Contents

Working with Primers

Design Primers

Analyze Primers

Primer Operations

Align Multiple Sequences

You can align multiple sequences against a reference sequence using a global or assembled procedure or do an exhaustive pairwise global alignment of all sequences with a progressive assembly of the alignments (multi-way alignment). If you frequently do the same alignment, you can save the setup file to disk. At another time, you can open the alignment setup file previously saved to disk, automatically loading the sequence files and returning to the setup dialog box. You can make any changes needed and begin the alignment when ready.

Alignment results are displayed in a tabbed window, described in a later chapter (Alignment Results). You can view a table of information summarizing the alignment results, the actual aligned sequences, a set of simple maps showing areas of significant similarity for each sequence, or a dendrogram for multi-way phylogeny, and you can save the alignment results to disk. Some alignment parameters can be adjusted.

Setting up the Alignment

Found on the Align menu or use Toolbar to access. You can use this procedure to set up a Global-Ref, Multi-Way or Assembled alignment of multiple sequences. To set up the multiple sequence alignment:

- Click Align, Align Multiple Sequences
- Click Global-Ref, Multi-Way or Assembled.
- Select to align sequences as DNA, amino acids, or a translation displayed as DNA
- Select the scoring matrix or use the default setting (not used for Assembled alignments).
- Click **Next** to identify the sequences to align.

Global-Ref Alignment -- This procedure will align each of the sequences you specify against a reference sequence (always in position 1). The objective is to maximize the number of matching bases over the full length of each molecule. This type of global alignment works best when sequences are of similar size and reasonably related to the reference molecule.

Multi-Way Alignment -- This procedure will do exhaustive pairwise global alignments of all sequences and progressive assembly of alignments using Neighbor-Joining phylogeny. When the alignment is complete, you will have the option to re-order the sequences based on the phylogenetic data produced.

Assembled Alignment -- The program will compare each sequence against the reference sequence (always in position 1), looking for significant regions of local homology to lock in (using the FastScan - MaxQuality method). The alignment is then completed using a global alignment procedure to align the rest of each sequence against the reference sequence. This type of alignment can be useful when one or more areas of strong similarity are found in otherwise poorly-related sequences.

Alignment Parameters

Alignment parameters have been set to generally give the best results for DNA or protein alignments. You do not have to make any adjustments to these settings unless you want to achieve specific objectives with your alignment. To change the parameters for the current alignment, click the Param button in the alignment setup dialog box. When you make a change to a parameter setting, it will be in effect for the current alignment only. If you want to make the change affect all future alignments of this type, click the Defaults, Set As… button in the lower right. The following parameters can be adjusted for global or multi-way alignments using the standard Linear Scoring Matrix:

Mismatch Penalty -- Penalty assigned to each mismatched position. The default value is 2. An increase in this penalty favors gaps over mismatches.

Open Gap Penalty -- Penalty incurred when a gap is introduced into an alignment. The default value is 4. An increase in this penalty reduces the number of small gaps permitted.

Extend Gap Penalty -- Penalty assigned to each missing residue in a gap. The default value is 1. An increase in this penalty reduces the length of gaps.

Conservative Changes -- For protein alignment, you can indicate whether conservative changes will be permitted. If conservative changes are not permitted, exact matches will be required. The conservative changes recognized by the program are: (Ile, Leu, Val, Met), (Asp, Glu), (Asn, Gln), (Lys, Arg), (Ala, Ser, Thr)

Alignment Item Names

Molecules selected for multiple sequence alignments can display the molecule name or file name, up to a maximum of 16 characters. If the name is too long, the program will truncate the name, using the rule selected. To specify which names to use for your alignments, click the Names button in the alignment setup dialog box.

Select molecule name or file name and select the truncation rule to apply when names are too long. You can direct the program to use the first 16 characters of the name, the last 16 characters (not including file extension) or a combination of the first 7 characters, followed by two dots, then the last 7 characters (omitting characters in the middle, as needed.) When you make changes, the new settings will apply to subsequent molecule selections.

Translating DNA for Alignment

If it is necessary for the program to translate DNA so that it can align the amino acid sequence, it will translate in frame. Translation will start from the basepair position indicated on the setup list (or from the first base of the complete sequence) and continue to the end basepair, going through termination codons, if encountered. To change the translation frame, change the start basepair position by one or two bases.

Scoring Matrix Options

You can select the scoring matrix for global ref or multi-way alignments. User adjustment of alignment parameters is available when using the Standard Linear Scoring Matrix

Standard Linear -- Can be selected for DNA or amino acid sequence alignments. Penalty values are applied when bases do not match or when gaps in the alignment must be opened or extended. These penalty settings can be adjusted by advanced users. For protein alignments, you can indicate whether conservative changes will be permitted. If conservative changes are not permitted, exact matches will be required.

BLOSUM 62 -- Can be selected for amino acid sequence alignments. Uses a substitution matrix (BLOSUM 62) which assigns a score for each amino acid substitution. This substitution matrix was derived from data using the BLOCKS database (Henikoff and Henikoff) in the study of distantly related sequences.

PAM 250 -- Can be selected for amino acid sequence alignments. Uses a substitution matrix (PAM 250) which assigns a score for each amino acid substitution. This substitution matrix was derived using a data set of closely related sequences and is based on the point-accepted-mutation (PAM) model of evolution described by Dayhoff.

Sequences to Align

You can add sequences to the list of sequences to align, specifying the regions for alignment within each sequence. You can move sequences up or down within the list and delete items, if needed. If you are doing a Global-Ref or Assembled alignment, remember that the reference molecule should be in position 1.

To add sequences to the list:

- Click the Molecule List button at the right.
- Select the sequences or group from the molecule list
- To remove a sequence, select (highlight) it and then use the delete button at the right.
- Use the up or down arrowhead buttons to move the highlighted sequence within the list, if needed.

The molecule name, molecule size, and file name of the selected molecule are shown in the area below the list. When all of the sequences have been selected, click Finish to complete the setup and start the alignment.

To align only a part of a molecule sequence:

- Select (highlight) the name of the molecule within the list
- Click the Edit Item button at the right.
- Click Partial Sequence to align less than the full molecule and enter the start and end basepair positions. The start base should always be the first nucleotide to be aligned.
- Use the Complement checkbox if the region to be aligned is on the complement strand.
- Click OK to complete this item.

To change the name of an aligned molecule, select (highlight) the name of the molecule within the list and click the Edit Item button. Enter the Name to Display in the space provided.

Saving and Retrieving Alignment Setup Files

When you have completed the Align Multiple Sequences setup dialog boxes, you can save this alignment setup information to a file on disk. You may want to use this option if the alignment setup is used routinely in your work. Alignment setup files are saved in a program-specific format and use the extension *.am4. The information saved includes molecule names, file names and locations, position in the alignment, start and end basepair position, and the alignment parameters.

To save a setup file, click the Save Setup button (lower left) on the last page of the alignment setup dialog box. Enter a file name and select the Save in… location.

You can later retrieve this setup, adding or inserting new sequences as needed, and repeat the alignment procedure without having to enter all of the molecule names, regions, or alignment parameters again. When you open an alignment setup file previously saved to disk, the program will automatically load the sequence files and return you to the setup dialog box. You can make any changes needed and begin the alignment when ready.

To retrieve a setup file, click Align, Align Multiple Sequences and then click the Retrieve button (lower left). Select an alignment setup file previously saved to disk. Retrieved alignment sets appear on the molecule list as groups for easy handling. You can remove the group when the alignment is complete or select the group to do an additional alignment.

Compare Two Sequences

You can compare two sequences by searching for regions of local homology or by aligning the two sequences using a global or an assembled alignment procedure. Local homology searches can be done using the FastScan method or the Needleman-Wunsch method. The global alignment procedure will maximize the alignment of bases along the full length of both molecules. This procedure works well for sequences that are very similar overall. The assembled alignment can be used effectively for sequences with some similar and some different regions. The assembled alignment is based on finding local homologies and then aligning the remaining sequence using a global strategy.

You can compare sequences as DNA bases or as amino acids, translating a DNA sequence in frame, if needed. You can also translate DNA sequences to determine the best protein alignment and then display the results as DNA. And you can change the parameters for the current alignment as needed (click the small button in the upper right corner of the alignment setup dialog box). From the results screen, you can use the Redefine toolbar button to return to the alignment setup dialog box to change the type of comparison and then restart the search or alignment procedure.

Alignment results are displayed in a tabbed window, described in a later chapter (Alignment Results). For local homology searches, you can view a list of the local homologies found or a set of simple maps locating these local homologies. For each homology block listed, you can view the actual aligned sequences or a more detailed map showing the local homology in each molecule.

Setting up the Comparison

Found on the Align menu or use Toolbar to access. You can use this option to set up a local homology search, a global alignment, or an assembled alignment of two sequences.

To set up the sequence comparison:

- Click Align, Compare Two Sequences and then click Local, Assembled or Global.
- For Local homology searches, select the method (see below) or use the default setting and indicate the maximum number of results to report (default 20).
- Select to compare sequences as DNA or amino acids and use the checkbox to indicate if both DNA strands should be compared.
- For Global alignments, select the scoring matrix or use the default setting.
- Click **Next** to identify the sequences to compare.

Compare with Local Homology Searches -- You can compare two sequences by searching for regions of local homology. Select to use the FastScan or Needleman-Wunsch method. Once the local homology search is complete, you can view a list of the best regions found, inspect the aligned sequences or view a location map.

Compare with Assembled Alignment -- This procedure compares the two sequences that you specify by finding significant regions of local homology (using the FastScan - MaxQuality method) and then completing the alignment using a global alignment procedure. This type of alignment can be useful when one or more areas of strong similarity are found in otherwise poorly-related sequences.

Compare with Global Alignment -- You can compare two sequences that you specify using a global alignment procedure. The objective is to maximize the number of matching bases over the full length of each molecule. This type of global alignment works best when sequences are of similar size and reasonably related to each other.

Local Homology Search Methods

FastScan Method -- The FastScan procedure is based on the rapid search lookup table methods popularized by FAST and BLAST, modified to allow homologies with multiple gaps and frame shifts. The FastScan method is much faster than Needleman-Wunsch and, in general, tends to find shorter regions of higher homology.

Select FastScan - MaxScore to bias the search in favor of high-scoring regions. (Longer regions can increase score, even if homology is lower.) Select FastScan - MaxQual to bias the search in favor of high-quality regions. (Favors finding shorter regions of very high homology.) Based on the length of the search region, the program will automatically select the appropriate ktup (block size) for the homology search to provide a very sensitive search. You cannot adjust the ktup settings for this local alignment operation. You can indicate whether or not to allow conservative changes for protein alignments.

Needleman-Wunsch Method -- The Needleman-Wunsch method uses exhaustive base-by-base comparisons to find long regions of homology, even if they contain numerous mismatches. This method has been optimized to handle large sequences. The local alignment parameters used with this procedure have been set to generally give the best results for DNA or protein sequences. You can adjust these parameters to achieve specific objectives.

Alignment Parameters

User adjustment of alignment parameters is available for FastScan and Needleman-Wunsch alignments and for global alignments when using the Standard Linear Scoring Matrix. Alignment parameters have been set to generally give the rest results for DNA or protein alignments. You do not have to make any adjustments to these settings unless you want to achieve specific objectives with your alignment. To change the parameters for the current alignment, click the small button in the upper right corner of the alignment setup dialog box. When you make a change to a parameter setting, it will be in effect for the current alignment only. If you want to make the change affect all future alignments of this type, click the Defaults, Set As… button in the lower right.

FastScan Parameters -- For local homology searches, parameters are automatically set to give the most sensitive search. For protein alignment, you can indicate whether conservative changes will be permitted. If conservative changes are not permitted, exact matches will be required. The conservative changes recognized by the program are: (Ile, Leu, Val, Met), (Asp, Glu), (Asn, Gln), (Lys, Arg), (Ala, Ser, Thr)

Needleman-Wunsch Parameters -- The following parameters can be adjusted for local homology searches using the Needleman-Wunsch method:

Match Score -- Value assigned when there is a match. The default value is 1. An increase in this score favors finding regions of poorer homology.

Mismatch Penalty -- Penalty assigned to each mismatched position. The default value is 1. An increase in this penalty favors finding only regions of higher homology. In general, the mismatch and gap penalties should be identical.

Gap Penalty -- Penalty assigned to each missing residue in a gap. The default value is 1. An increase in this penalty favors mismatches over gaps.

Conservative Changes -- For protein alignment, you can indicate whether conservative changes will be permitted. If conservative changes are not permitted, exact matches will be required. The conservative changes recognized by the program are: (Ile, Leu, Val, Met), (Asp, Glu), (Asn, Gln), (Lys, Arg), (Ala, Ser, Thr)

Linear Scoring Matrix Parameters -- The following parameters can be adjusted for global or multi-way alignments using the standard Linear Scoring Matrix: Mismatch Penalty, Open Gap Penalty, Extend Gap

Penalty, Conservative Changes. A description of these parameters and default values is provided in the preceding chapter, Align Multiple Sequences.

Translating DNA for Alignment

If it is necessary for the program to translate DNA so that it can align or compare the amino acid sequence, it will translate in frame. Translation will start from the basepair position indicated (or from the first base of the complete sequence) and continue to the end basepair, going through termination codons, if encountered. If you need to change the translation frame, you can do this by changing the basepair position for the start of the region to compare. Click Partial Sequence and then change the "Start At" number by one or two bases, as needed, to change the frame.

Scoring Matrix Options

You can select the scoring matrix for global ref or multi-way alignments. User adjustment of alignment parameters is available when using the Standard Linear Scoring Matrix

Standard Linear -- Can be selected for DNA or amino acid sequence alignments. Penalty values are applied when bases do not match or when gaps in the alignment must be opened or extended. These penalty settings can be adjusted by advanced users.

BLOSUM 62 -- Can be selected for amino acid sequence alignments. Uses a substitution matrix (BLOSUM 62) which assigns a score for each amino acid substitution. This substitution matrix was derived from data using the BLOCKS database (Henikoff and Henikoff) in the study of distantly related sequences.

PAM 250 -- Can be selected for amino acid sequence alignments. Uses a substitution matrix (PAM 250) which assigns a score for each amino acid substitution. This substitution matrix was derived using a data set of closely related sequences and is based on the point-accepted-mutation (PAM) model of evolution described by Dayhoff.

Sequences to Compare

You can identify the two sequences to compare, specifying the regions within each sequence, as needed.

To identify a sequence to compare:

- For Sequence box 1, click the Change button at the right
- Select the sequence from the molecule list
- Repeat this procedure for Sequence box 2.
- Click the Flip button to exchange the sequences between the boxes, if needed.

To compare only a part of a molecule sequence:

- Below the molecule name, click Partial sequence
- Enter the start and end basepair positions of the region. The start base should always be the first nucleotide to be compared.
- Use the Complement checkbox to indicate that the region is on the complement strand.
- Use the Features button to access the features table for this molecule. You can select a feature to automatically enter the basepair positions and strand information.

Click **Finish** to start the comparison procedure.

Scan for Similarities

You can scan files or folders on your computer or local network, looking for regions of similarity to all or part of a search molecule. To set up a scan operation, you identify the molecule and range you want to match, and indicate whether or not to translate this molecule (there are specific advantages to doing either DNA or protein searches). Next, you select the files or folders you want to search and set the scan parameters.

The search method used by the scan operation is based on the rapid search lookup table methods popularized by FAST and BLAST, modified to allow homologies with multiple gaps and frame shifts. Once the scan operation is complete, you can view different aspects of the scan results in a tabbed display window, described in a later chapter (Alignment Results). You can view a list of the regions of similarity found, inspect the aligned sequences or view a location map. The Scan Results Wizard will help you to use the molecules with sequence similarities identified by the scan. You can also retrieve scan results previously saved to disk, and continue your analysis of these results.

Setting up a Scan Operation

Found on the Align menu or use Toolbar to access. You can look for regions of similarity to the full sequence of the search molecule or just a part of this molecule (match region) and you can indicate whether or not to translate (process) the molecule.

To set up a scan operation:

- Click Align, Scan for Similarity and then click New Search.
- If the molecule shown is not the one you want to use, click the Change... button to select a different sequence from the molecule list.
- Identify the Match Region, if less than the full molecule (see below)
- Select to process (translate) the molecule or use the no process setting.
- Click **Next** to identify the files or folders to scan.

To identify the match region:

- Click Partial sequence (below molecule name)
- Enter the start and end basepair positions of the region. The start base should always be the first nucleotide to be used for the search.
- Use the Complement checkbox to indicate that the region is on the complement strand.
- Use the small Features button to access the features table for this molecule. You can select a feature to automatically enter the basepair positions and strand information.

Processing the Molecule

You can indicate if the search molecule should be translated or if no processing is required.

None -- Use the DNA or protein molecule as is. For DNA molecules, search for regions of similarity on both strands of the DNA molecules or library sequences. Protein files will not be scanned when you specify a DNA search. For protein molecules, sequences will be compared to each protein molecule or to all 6 frames of translation for each DNA molecule or library sequence.

1 Frame Translated -- The program will translate the DNA molecule from the start basepair position identified (or from the first base of the complete molecule) to the end basepair position. Translation will

proceed through termination codons, if encountered, and the translated protein will be used as the search sequence. This search sequence will be compared to each protein molecule or to all 6 frames of translation for each DNA molecule or library sequence.

6 Frames Translated -- The program will translate the region identified (or the complete molecule) in all 6 frames and then look for homologies to each protein molecule or to all 6 frames of translation for each DNA molecule or library sequence. This option will take approximately 6 times longer than 1 frame translation.

Identifying Search Locations

You can scan molecule files in folders on your computer or local network, and you can scan databank files such as GenBank flat file, Unigene or PIR, if your organization downloads and stores these files for local access. For each location specified, the list box reports the size (in kb) of the data files to be scanned and the number of files at that location. An error log file will report on any problems encountered during the scan. If you routinely search the same locations, you can save the list of files and folders and recall this list at a later time. You can also set the scan parameters (speed and cutoff value) for this search.

Selecting files or folders to scan

Click the Add Folder button to select a folder location to add to the search list. Navigate to the folder you want to add to the list and click OK. The program will attempt to open and read all molecule files in the location specified, including those contained in subfolders.

Click the Add File button to select a library file to add to the list. Navigate to the file you want to add to the list and click OK. The program will attempt to open and read this file during the scan procedure. GenBank databank files and standard files in any of the supported file formats that contain multiple molecules can all be added to this list. Supported file formats include GenBank (flat file), EMBL, FastA and PIR formats.

Click **Finish** to complete the setup and begin the scan procedure.

Scan Parameters

You can set the speed and cutoff value for each scan operation.

Speed -- Select Standard, Rapid or Sensitive. The program will automatically select the appropriate ktup (block size) for the homology search based on the length of the search region and the setting you select. In most cases, the Standard setting will be the most appropriate. Select Sensitive to increase the sensitivity of the scan by decreasing the ktup by one unit, where possible. For Rapid scans, a streamlined procedure and an alternate set of ktup values are used.

Cutoff Values -- The cutoff value is the score required in order to add the homology block to the results list. The Standard cutoff value is 30 for DNA search molecules and 20 for protein search molecules. Select Reduced or Low to lower the cutoff value in order to accept results of lower significance. Reduced is approximately 2/3 the standard value, while Low is 1/2 the standard value.

Saving and Retrieving Scan Results

The results of a completed scan can be saved to disk in a program-specific file format with the extension *.as4. The file format saves the search molecule information and search parameters, as well as the information about the homologies found. To save scan results, click File, Save As.. while the scan results are in the active window on the screen. You can later retrieve this file to continue analyzing the results of the scan operation. To retrieve a scan results file, click Align, Scan for Similarity and then click Retrieve Results. Click the Retrieve Now button at the right and select the scan results file from the file open dialog box.

Using the Scan Results Wizard

At the completion of a scan for similarity, you can select one or more molecules and load the molecule files for use now, extract the molecules from a databank file for use later, or set up a multiple sequence alignment of these molecules, using the scan results as a starting point.

One molecule -- move the highlight bar to the molecule name in the results list and then click the Scan Results Wizard toolbar button.

Multiple molecules -- click the wizard button and then pick the molecules from a list box provided. Mouse-click to tag each molecule you want to select. You can also click the Tag button to quickly tag the first 10 molecules on the list (list is sorted the same way you sorted the results list on the screen display). If you want to tag more or less than 10 molecules, use the spin control above the Tag button to increment or decrement the setting for 'n'. Click **Next** when you have completed molecule selection.

Loading Molecules

You can have the wizard load the selected molecules for use now, placing them on the Molecule List. If only one molecule is selected, the wizard will enter the molecule to the Molecule List and automatically open the molecule viewer window so you can view the molecule sequence or features table or begin another task.

If multiple molecules are selected, these will be entered to the Molecule List. You can then select a molecule from the Molecule List to open the molecule viewer window. If more than 5 molecules are chosen, you will have the option to enter these molecules as a group for easier handling.

Extracting Molecules from Databank Files

You can have the wizard extract the selected molecules from a databank file, writing the new file(s) to disk for later use. If only one molecule is selected, the wizard will extract the molecule and display the Save As.. dialog box so that you can name the file and indicate its location. If multiple molecules are selected, you can decide how to save these files:

Auto-name files and then save to one location -- The program will assign a name for each file, using the molecule name or locus name (or an existing filename, if appropriate). You can specify one location for all of the files extracted in this operation.

Save each file individually -- The program will extract each molecule and display the Save As.. dialog box so that you can name the file and indicate its location.

Combine multiple databank molecules in one output file -- The program will extract each molecule and add it to a new multiple-molecule file. Enter a file name and location in the Save As.. dialog box. All files must be in the same file format (GenBank, EMBL or FastA).

Setting up a Multiple Sequence Alignment

The wizard will set up an Assembled Multiple Sequence alignment. The source sequence and match region used for the scan operation will be specified as the reference sequence for the assembled alignment. Each of the molecules you selected will be aligned against the reference sequence, using the homology region identified in the scan. Once you see the alignment setup dialog boxes, you can change this information, as needed. Click Finish to complete the setup and begin the alignment procedure.

Alignment Results

When an alignment operation has been performed (align multiple sequences, compare two sequences, scan for similarities), the alignment results will be displayed in a comprehensive, tabbed viewing window. Click the tabs to access alternate views of your results. Toolbar buttons in the results window let you sort data, navigate through results lists, format the display style for aligned sequences, or view additional data, where applicable. In some data displays, you can also use your mouse to point to a molecule to get additional information or point to a position in a map or sequence and find out the approximate basepair position. You can print the data in a view window, copy this view to the clipboard, or export the view data to a disk file.

The data displayed will be dependent on the type of alignment operation. When you align multiple sequences using global-ref, multi-way or assembled alignment protocols, the procedure ends with one alignment result. Local alignments and scans for similarity usually end with multiple results, providing information about each local homology found.

Info Tab

Click the Info Tab to view a table of information that gives a concise overview of the multiple sequence alignment results, including the sequences and regions aligned, the number of matches found, the number of nonmatches found, and the calculated percent matching bases. For Assembled alignments of two sequences, you can also see the local homology blocks used to drive the alignment.

When you move the mouse over a molecule name in the multiple sequence tables, the cursor changes to a rectangle shape. Use this special cursor to click on a molecule name to see a pop-up box that will give you more information about this molecule (name, size, file name, description).

You can use the Redefine toolbar button to return to the alignment setup dialog box to change the alignment setup, if needed, and then repeat the alignment.

Use the Edit Name toolbar button in multiple sequence alignments to view or edit the name, description, or notes information that can be saved with the results to define this alignment.

Local Alignments

The Info Tab shows a table that lists the regions of local homology or similarity found in two-sequence comparisons using the Local Homology option and scans for similarity. You can sort this list by score, percent matches, length, or the start position of the region of homology in molecule 1 or molecule 2 (local homologies) or the start position in the search sequence (scan operation). Click the Picture tab to see a graphic representation of this listed information.

You can move the highlight bar to any region of homology found. When you click the Sequence or LMap tabs, you will see aligned sequence data or a detailed map for the selected region of homology. While you are viewing the aligned sequences or detailed maps, you can move to the next or previous result using the navigation buttons on the results window toolbar.

For scan results, when you move the mouse over the list of locus names, the cursor changes to a magnifying glass. Use this cursor to click on a locus name to see more information about this molecule.

Picture Tab

When a multi-way alignment has been performed, you can view a dendrogram which suggests the pattern of relatedness of all of the sequences aligned. Click the Picture tab in the alignment results window to access this view. The dendrogram is constructed with a distance-based tree-building method using the neighbor-joining algorithm. The amount of dissimilarity (the distance) between two aligned sequences is used to derive the tree. The method is very fast and results in a single tree, similar to the phylogenetic guide trees produced using the CLUSTAL W approach (Higgins et al).

Local Alignments

Click the Picture tab in the alignment results window to view a graphic representation of the list of regions of homology found. (Click the Info tab to see this data in table form.) For two-sequence comparisons using the Local Homology option, each display entry shows the region of local homology (shown as a colored bar), between lines representing the two molecules. For Scan for Similarity operations, each display entry shows the approximate location of the region of homology (shown as a colored bar) in the search molecule sequence. You can sort these displays by score, percent matches, length, or the start position of the region of homology.

You can move the highlight bar to any region of homology found. When you click the Sequence or LMap tabs, you will see aligned sequence data or a detailed map for the selected region of local homology.

Sequence Tab

You can view the actual aligned sequences in a variety of formats, emphasizing the differences or similarities between the sequences. Click the Sequence tab in the alignment results window to access this view.

Click the Similarity/Difference button to toggle between the default similarity and default difference formats to get a rapid impression of which view will give you the best display for this data. In the difference formats, dots indicate matches. In protein alignments, * indicates a terminator and is always shown as a mismatch.

Some formats apply color if the similarity exceeds a specified value (Color area of high matches or Similarity summary bars formats). Use the Similarity Significance Value combo box to set the level of significance (range 50-100%) you want to apply for these formats.

Click the Format button to access the full set of sequence format options or to set page style options. When the format is just the way you want it, you can print the aligned sequences in this format, or send them to the clipboard or to a disk file to later insert in a document.

In multiple sequence alignments, you can change the name used for one of the molecules aligned, if needed. Right-click on the name displayed and enter the name you would like to use (maximum 16 characters).

Format Display Styles for Aligned Sequences -- You can select whether to show all bases in the aligned sequences (focus on similarity) or to show only non-matching bases (focus on difference). Next, you can select a display style. Check the sample box in the upper right to see how your selected alignment format will appear. Some styles use a background color behind the sequence. You can change to a different background color using the combo box below the sample area.

Page Styles for Aligned Sequences -- You can set style options for the molecule names and basepair numbers that will appear with your aligned sequences. The number of bases of sequence per line will be affected by these selections.

Printed output -- You can select the font size to be used for sequence data sent to your printer. The font size will also affect the number of bases of sequence per line. The font face will always be Courier New if this is supported in your printer. Use the printer set-up box if you want to change page orientation.

Location Map Tab

When a global-style sequence alignment has been performed, you can view a summary of the extent and location of similarity of the sequences and additional information about regions of DNA or protein molecules. Click the LMap tab in the alignment results window to access this view.

Each of the sequences is shown on a separate line. The entire search region is shown for each molecule, scaled to fit in the data box. You can drag out the box to increase the resolution of the map. For global alignments, each sequence is compared to the reference sequence. For multi-way or two-way alignments, each sequence is compared with a virtual composite sequence. You can set the significance value for this display using the Similarity Significance Value combo box in the toolbar.

Regions where similarity is greater than or equal to the significance value (shown in combo box) are marked with colored bars. For assembled alignments, a blue bar indicates that the molecule is being aligned 5' to 3', while a light red bar indicates the molecule is being aligned in the reverse direction (3' to 5'). You can point with the mouse at a map area and depress the left mouse button to see the basepair position at the cursor location. You can also view composition data as an overlay on top of the similarity data.

Compress view -- If you have a large number of sequences, you can compress this view to fit more similarity map data in the display. In the compressed view, the lines are closer together and the molecule names are suppressed. Click the Compress View button to enable or disable this option.

Overlay Composition Data -- Click the Overlay Composition Data button to add composition data to the uncompressed similarity map data in the alignment results window. Click again to remove this overlay. For DNA sequences, AT-rich regions are shown with magenta bars, GC-rich regions are shown with green bars. For protein sequences, Polar-rich regions are shown with magenta bars, Nonpolar-rich regions are shown with green bars

Local Alignments

For each region of local homology or similarity, you can view a map showing the approximate position of the region of homology within each of the molecules. Click the LMap tab in the alignment results window to access this view. Molecule names are shown, along with homology block basepair coordinates. The regions of local homology appear as colored bars. Homologous regions on the normal strand appear in blue, while regions on the complement strand appear in light red.

Click the Next Result or Previous Result toolbar buttons to navigate through the list of local homologies found without leaving this view window.

Saving Alignment Results

When you have completed an alignment, you can save the results to a disk file. You can later re-open this results file and continue to review the results data. To save alignment results, with the results visible on the screen, click File, Save As... and enter a filename when prompted. The results will be saved in program-specific file formats with the following extensions: *.am4 -- used for multiple sequence alignment results data or alignment setup files, *.ac4 -- used for compare two sequences results data,*.as4 -- used for scan for similarity results data, or *.aa4 -- used for sequence assembly results data.

To open a results file, click File, Open and select the file previously saved to disk. You can change the format options or print the aligned sequences, export results data to disk, or copy to the clipboard for use in other applications.

Exporting Alignment Data

You can copy or export the aligned sequences, as shown on the display screen, for use in another program. Click File, Export to use these standard export and copy functions.

Using special alignment data export functions found on the Operations menu, the consensus sequence of an alignment or a special merged/spliced consensus sequence can be exported to a disk file, copied to the clipboard, or loaded as a molecule for use now. The sequence will be exported in ASCII (DOS text) file format in upper case letters, 60 bases per line, with no base pair numbers. Aligned sequences can also be copied or exported in CLUSTAL W multiple sequence alignment file format for later use in another program.

Export Consensus Sequence -- Select this option to export the standard consensus sequence derived from the alignment. Gaps in the aligned sequences are significant and will affect the final consensus sequence.

Export Merged/Spliced Consensus Sequence -- Select this option to export a special merged consensus sequence that can be used for splicing two overlapping regions. In generating this consensus sequence, gaps in the alignment zone are ignored.

Export Aligned Sequences -- Select this option to export the aligned sequences in standard CLUSTAL W multiple sequence alignment file format.

Editing Alignment Data

When viewing aligned sequences, you may feel it is necessary to make small changes to the alignment data. If you believe there is an error in the sequence, you can change one base for another. If the alignment was performed using sequence files from automated sequencing machines and these files contain trace data (ABI or SCF format files), you can view the trace for the questionable position in the aligned sequence. Where there are gaps in the sequence, you can move unaligned bases left or right.

To edit alignment data, open the alignment viewer window and click the Sequence tab to show the aligned sequences. Then follow the steps below.

To replace a base -- Right-click on the base you want to replace and select Edit. Click the Replace Base option in the dialog box and verify that the base you pointed to is shown. Enter the base you want to use as a replacement. You can use the checkbox if you want the program to make the same change in the molecule sequence. Remember to save the updated molecule to disk if you want to retain the modified sequence.

To display a trace -- Right-click on the base in question. Select Trace from the menu shown. The display will show the base in question and about 20 bases to the left and right. The base you specified is marked with a small arrowhead below the base letter. If you feel that the base was not called correctly, you can use the Edit function to change the base in the aligned sequence (see above).

To shift a base left or right -- Right-click on the base you want to shift. Select Left Shift or Right Shift to move the base in the direction indicated. The base will be moved within the sequence gap (marked with ----) until the base lines up with the same base in the reference sequence or until it reaches the end of the gap. (In a multi-way sequence alignment procedure, the virtual composite sequence is used as the reference.)

Assemble Sequences

You can use this module to assemble sequences for simple or complex projects or to locate subclones on a reference sequence. You identify the location of the sequence files or folders of sequence files that you want to assemble and the program handles opening each file and reading the sequence data. Results are presented in a data window with multiple views -- you can review a table of results information, a picture view of alignment locations, or the aligned sequences with optional accessory data in a split pane. You can save the results and later open the file to continue reviewing this data. With the results data in the active window, click File, Save As... and enter a name for the results file. To open this file at a later time, click File, Open and select the file.

Setting up an Assembly Project

Found on the Align menu or use Toolbar to access. You can use this option to assemble sequences for simple or complex projects.

- Click Align, Sequence Assembly
- Click the Simple or Complex sequence assembly option.
- Click Next to specify the location of the sequence files and set the assembly parameters.

Simple sequence assembly project -- you can use this option to assemble a contig from a small set of sequences with limited overlap coverage. You might select this option to assemble the sequences for a gene or region of a sequence that you have cloned.

Complex sequence assembly project -- use this option for larger sequence assembly projects with many sequences and multiple overlap coverage. This is a good option to select when assembling unknown sequences, such as those for a small viral genome. This module is not intended for very large assemblies, like genome-sized projects. Special parameters for the complex sequence assembly module have been set to generally give the best results. You do not have to make any adjustments to these settings unless you want to achieve specific objectives with the assembly. To change the parameters, click the small Param button (at right).

Selecting sequence files to assemble

Click the Add Folder button to select a folder location to add to the search list. Navigate to the folder you want to add to the list and click OK. The program will attempt to open and read all sequence files in the location specified, including those contained in subfolders.

Click the Add File button to select a file to add to the list. Navigate to the sequence file you want to add to the list and click OK. The program will try to open and read this file during the assembly procedure.

Set assembly parameters

Significant Overlap (simple sequence assembly) -- Enter the score that must be achieved in order to consider the overlap between two sequences to be significant (default = 50). Decrease the overlap score to allow assembly of less significant overlaps and enable a finer search.

Expected Coverage (complex sequence assembly)-- Set the option for 6X coverage when you expect to have 6 or more reads along most of the sequence. Set the option for 2X or 3X when coverage is expected to be less than 6X (perhaps for data that is not yet complete).

Shuffle sequences -- Set this checkbox to instruct the program to randomize the sequence order before beginning the assembly process.

Click **Finish** to complete the setup instructions and begin the sequence assembly procedure. When the sequence assembly is complete, one or more contig sequences will be reported. For each contig, the consensus sequence will appear in the first position and each of the other sequences will be aligned against this reference.

Using the Subclone Locator

You can use the Subclones option of the Sequence Assembly module to prepare an alignment that will help you to locate clones on the reference sequence. You identify the reference sequence molecule and then specify the location of the subclone files that you want to scan. As each file is read, the program will compare both strands of each file to the reference sequence, looking for the most significant region of local homology.

- Click Align, Sequence Assembly
- Click the Subclones option.
- If the reference molecule shown is not the one you want to use, click the Change... button to select a different sequence from the molecule list.
- Click Partial Sequence to identify a region less than the full molecule (see below)
- Click Next to specify the location of the sequence files

To identify a partial region, Click Partial sequence and enter the start and end basepair positions of the region. The start base should always be the first nucleotide to be aligned. Use the Complement checkbox to indicate that the region is on the complement strand. Use the Features button to access the features table for this molecule. You can select a feature to automatically enter the basepair positions and strand information.

Selecting sequence files to assemble

Click the Add Folder button to select a folder location to add to the search list. Navigate to the folder you want to add to the list and click OK. The program will attempt to open and read all sequence files in the location specified, including those contained in subfolders. Click the Add File button to select a file to add to the list. Navigate to the sequence file you want to add to the list and click OK. The program will try to open and read this file during the assembly procedure.

Click **Finish** to complete the setup instructions and begin the sequence assembly procedure. The results of the assembly will show the local homologies aligned against the reference sequence (in position 1).

Sequence Assembly Results Viewer

When a sequence assembly operation has been performed, the assembly results will be shown in a data window that has multiple alternate displays. You can print the data in a view window, copy this view to the clipboard, or export the view data to a disk file.

The results viewer has resizable columns and split display areas. Use your mouse to pull on a column divider to make the column wider or pull on a panel divider to make the split panel taller or shorter. Resize the entire data window by pulling out the window from the lower right corner with your mouse.

If the sequence assembly resulted in more than one contig, use the combo box in the toolbar to select the contig data to view. Click the Enter Contig to Molecule List button to load the consensus sequence for this contig as a molecule for later use. Once on the Molecule List, you can select this molecule, add information or analyze the molecule, or save it to disk.

Click toolbar buttons to view info, picture, sequence, sequence and overview, sequence and coverage, or sequence and trace. Click the Summary button to view a concise overview of the assembly results.

View Info

Click the View Info toolbar button to display a table of information, reporting the results of the sequence assembly. The consensus or reference sequence is shown in the first position. Each sequence aligned against this reference is reported, showing the position of the homology on the reference molecule, the position of the homology on the sequence molecule, the number of matching bases, the number of non-matching bases, and the percent error. The letter C after the basepair number indicates Complement strand. Percent Error is defined as the number of non-matching bases divided by the total number of bases, multiplied by 100.

View Picture

Click the View Picture toolbar button to display view a graphic picture of alignment locations. For a simple or complex sequence assembly, the consensus sequence is shown as a map line in the first position. Each of the aligned sequences is shown by a colored bar on a separate line. The position of the bar shows the approximate location of the aligned region in the consensus sequence. Sequence aligned 5' to 3' (blue bar) or aligned in the reverse direction (3' to 5') (light red bar).

For a subclone location operation, the reference sequence is shown as a double-stranded map line in the first position. Black bars above or below the double-stranded map line indicate gaps in coverage on the normal or complement strand. Each of the aligned subclone sequences is shown by a colored bar on a separate line. The position of the bar shows the approximate location of the aligned region in the reference sequence.

View Sequence

Click the View Sequence toolbar button to view the aligned sequences. The consensus or reference sequence is shown in the first position and the other sequences that are aligned against this sequence appear below. The sequence name is shown in the first column (use mouse to widen column if needed) and the basepair number for the first base shown for that sequence appears in the second column. Non-matching bases are shown in red (or with a light orange background). Ambiguous bases in the consensus sequences are shown in light blue (or with a pale blue background). Click the Sequence Format button to toggle between the colored base or colored background style for display of nonmatching or ambiguous bases. Use the horizontal scroll bar below the aligned sequences to move along the contig.

You can select to add a split panel display with accessory information below the aligned sequences. You can view Sequence and Overview (aligned sequence bars and position locator within assembly), Sequence and Coverage (histogram of number of reads, errors within reads), Sequence and Trace (trace data for selected sequence if available). Click the View Sequence button to return to plain sequence, closing the split pane view.

Sequence and Overview

Click the Sequence and Overview toolbar button to display the aligned sequences with an overview of the sequence assembly shown below the sequences in a split pane. The complete consensus or reference sequence is represented by a gray map line drawn full-width at the top of the split pane. The sequences that have been aligned against the reference appear in their approximate locations relative to the reference sequence. Sequence aligned 5' to 3' (blue bar) or aligned in the reverse direction (3' to 5') (light red bar).

A pair of brackets at the top and bottom of the split pane indicates the position and extent of the sequence you are viewing (in the top panel) within the assembly. Click with the mouse at any point within the overview image to move the brackets to a new location and jump to this position in the aligned sequences. You can change the size of the split pane, if needed to display a larger set of overview data. Use your mouse to pull on the divider between the upper and lower panes to increase or decrease the height of the lower pane.

Sequence and Coverage

Click the Sequence and Coverage toolbar button to view the aligned sequences with a coverage histogram shown below the sequences in a split pane. A histogram of coverage for the complete assembly is shown full-width in the split pane. At each point in the assembly, the number of aligned sequences contributing to the data is plotted on the histogram using a gray color. The height of the gray bar at a given position is proportional to the number of reads for the consensus sequence at that point. A scale is shown at the left and right edges of the histogram.

If non-matching bases are present in an aligned sequence, a red color is plotted at the position of the mismatch. The height and extent of red areas within the histogram can be used to identify possible problem areas. For subclone locator results, a solid red baseline indicates no coverage for this part of the reference sequence.

A pair of brackets at the top and bottom of the split pane indicates the position and extent of the sequence you are viewing (in the top panel) within the assembly. Click with the mouse at any point within the coverage histogram to move the brackets to a new location and jump to this position in the aligned sequences. You can change the size of the split pane, if needed to view the coverage histogram. When the height of the histogram pane is increased, the histogram scales to use the area available. Use your mouse to pull on the divider between the upper and lower panes to increase or decrease the height of the lower pane.

Sequence and Trace

If the assembly was performed using sequence files from automated sequencing machines and these files contain trace data (ABI or SCF format files), you can view the aligned sequences with a trace chromatogram shown below the sequences in a split pane.

Click the Sequence and Trace toolbar button to display one trace chromatogram. Click on a base in one of the aligned sequences to view the trace data for this sequence at and around this base position. Vertical and horizontal highlighters mark the sequence and base in the upper panel that you have selected. The lower panel displays the trace data with the selected base marked with a vertical highlighter and centered in the display, if possible. Click on another base to view trace data at another position.

Click the Sequence and Multiple Traces toolbar button to display trace chromatograms for all sequences shown. Click on a base in any line of aligned sequence data to view trace data for all sequences at this position in the alignment. A vertical highlighter marks the column of sequence bases you selected. The lower panel displays the trace data with the appropriate base marked in each sequence frame.

Use the mouse to pull on the assembly viewer window frame to make the entire data window larger, if needed. Use the mouse to move the split bar between the upper and lower panes to change the proportion of space available for the sequence or trace displays.

You can use the horizontal scroll controls in the upper sequence panel or the keyboard arrow keys to move the contig to the left or right. As long as the selected base stays within the visible sequence, the vertical highlighter remains on this base and the trace data in the lower panel continues to show the trace for the selected base, centered within the display. If, however, the vertical highlighter in the upper panel hits either the left or right window frame, the highlighter will begin to shift from one sequence base to the next, to remain within the visible display. At this point, the vertical highlighter in the lower panel will also begin to shift, synchronizing the trace data to the sequence data scrolling in the upper display. To move the vertical highlighter to the left or right within the sequence, use Alt + arrow key or just click on another base with the mouse.

Remote Searches

Clone Manager can help you to initiate remote searches of public domain databases. You need an internet connection to use these functions. The program can send search requests to the NCBI BLAST server and to several other selected web sites that provide public access to protein pattern databases in the public domain. Search requests are sent to the server over the internet and results are displayed in the built-in web browser. At the present time, you can send search requests to the Pfam, PROSITE and PRINTS protein pattern database web sites.

> Disclaimer: These web sites presently provide public access but their availability is subject to change without notice. In addition, Scientific & Educational Software is not responsible for analysis results or for any change in licensing to permit site access.

Internet BLAST Searches

Found on the Align menu. You can set up a remote search using the NCBI BLAST server. You need an internet connection to use this function. The program will help by handling query sequence formatting, in-frame translation, or excision of a subsequence, and will send your request to the server over the internet. The built-in web browser will show results and allow you to access molecules of interest.

To initiate a BLAST search, you will want to identify the query sequence (molecule and region), and then select the BLAST program to use and the BLAST database to search.

- Click Align, BLAST Internet Search (NCBI).
- Click the Change... button to select a different sequence from the molecule list, if needed.
- Identify the Match Region, if less than the full molecule
- Select to process (translate) the molecule, if needed. Note that DNA or protein query sequences will be appropriate for different BLAST programs and databases.
- Click Next to select the BLAST options for your search.

Note: If you are using this operation for the first time, you can click the Options button to view or modify your internet access configuration settings.

Processing the Molecule

You can indicate if the search molecule should be translated or if no processing is required.

None -- Use the DNA or protein molecule as is. Note that specific BLAST programs will handle query sequence translation as part of the search program.

1 Frame Translated -- The program will translate the DNA molecule from the start basepair position identified (or from the first base of the complete molecule) to the end basepair position. Translation will proceed through termination codons, if encountered, and the translated protein will be used as the search sequence.

Selecting BLAST Options

Once you have identified the query sequence, you can select the BLAST program to use and the database to search. Use the upper combo box to select the BLAST program you want to use. A description of the selected program is provided below the selection. The programs available depend on whether your query sequence is

DNA or protein. Use the lower combo box to select the BLAST database you want to search. A description of the selected database is provided below the selection. The databases available depend on the program selected (above).

Click the Finish button to prepare your query sequence and send your request over the internet to the NCBI BLAST server.

BLAST Programs -- BLAST programs blastn, blastx, tblastx, and MegaBlast are available for DNA query sequences. blastn compares against a nucleotide sequence database, blastx translates and then compares against a protein sequence database, tblastx translates both the query sequence and the sequences in the nucleotide database. MegaBlast is specifically designed to efficiently find long alignments between very similar sequences and find identical matches to your query sequence. BLAST programs blastp and tblastn are available for protein query sequences. blastp compares against a protein sequence database and tblastn compares against a translated nucleotide sequence database.

BLAST Databases -- A variety of protein or nucleotide sequence databases are available for your search. Descriptions are provided below the selection combo box. The database combo box will limit your selection to those databases appropriate for the BLAST program you selected for use. In addition, please note that the tblastx program should not be used with the nr database.

Viewing BLAST Search Results

Once the BLAST search request has been sent to the NCBI BLAST server, the web browser will open, displaying the NCBI Format Request page. To view results, click the View Report button on the web page.

When viewing results, you can jump to molecules of interest. You can save molecules to a disk file on your computer and later read these files into Clone Manager or other software programs. Select the GenBank display format before saving to a file.

When you are viewing a molecule file, if the molecule data is in the GenBank format, you can import the molecule directly 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. Click the Get Molecule button on the browser window toolbar (not on the web page) to import the molecule for immediate use.

Other Internet Searches

Found on the Align menu. You can send search requests to other selected web sites that provide public access to protein pattern databases in the public domain. You need an internet connection to use this function.

The program will help by handling query sequence formatting, in-frame translation, or excision of a subsequence, and will send your request to the server over the internet. The built-in web browser will show results and allow you to access more information about the protein domains or patterns. To initiate a protein pattern search, you will want to identify the query sequence (molecule and region), and then select the public domain database to search.

- Click Align, Other Internet Search.
- Click the Change... button to select a different sequence from the molecule list, if needed
- Identify the Match Region, if less than the full molecule
- Molecule processing is automatic. If you select a DNA sequence, it will be translated in frame.
- Click Next to select the public database for your search.

Selecting Public Database Options

Once you have identified the query sequence, you can select the public database to search. Use the upper combo box to select the database to access. A short description of the selected database is provided below the selection. The lower combo box provides the address (URL) for the public domain database web site. You can change this address, if needed.

Click the Finish button to prepare your query sequence and send your request over the internet to the database web server.

At the present time, you can send search requests to the protein pattern database web sites described below.

Pfam Database Information -- Pfam is a comprehensive database of protein domain family alignments and HMMs. The protein family alignments were constructed semi-automatically using hidden Markov models (HMMs). Pfam families contain functional annotation and cross-references to other databases. The Pfam database is accessed using the Pfam server at the Washington University in St Louis. Interactive access URL http://www.sanger.ac.uk/Software/Pfam

PROSITE Database Information -- The PROSITE database is a collection of protein sites and patterns. Biologically significant protein domains were cataloged and thoroughly annotated by experts. Each domain is summarized by a pattern, which becomes a part of the profile library. This database is accessed with the ExPASy proteomics server at the Swiss Institute of Bioinformatics (SIB), hosted in the US by the North Carolina Supercomputing Center (NCSC). Interactive access URL http://us.expasy.org/prosite

PRINTS Database Information -- The PRINTS database is a collection of protein fingerprints (groups of conserved motifs used to characterize a protein family). This database is accessed with the FingerPrints web server at the University of Manchester. Interactive access URL http://umber.sbs.man.ac.uk/dbbrowser/PRINTS/

Viewing Internet Search Results

Once the internet search request has been sent to the appropriate web server, the web browser will open, displaying the search results. You can usually print or save this page for later review. In some cases, you can also link to more information about the protein domain or pattern.

You may want to visit these sites directly, using the interactive access URLs provided above. Each site contains links to helpful information and references.

Working with Primers

You can use the primer operations of Clone Manager to design and analyze primers. For streamlined storage and access, you can save essential primer information to primer collections that contain information about many primers, in very compact form. These primer collections are always available for quick access within the Primer List display.

You can also use File, Save or Save As... to save an active primer to a separate disk file for later use by the program. Primer files are saved in a program-specific format and use the extension *.pd4. Primer files contain the primer sequence, including the 5' extension (if any), source molecule information (where appropriate), the primer criteria, and the primer name, description and notes information.

You can use File, Open to retrieve primer files prepared in a previous work session and saved to disk. When a primer file has been identified, if the primer was linked to a molecule, the program will attempt to load the molecule file and relink the primer. If the original molecule file cannot be found or opened, you can try to find the sequence file or you can work with the primer without its source molecule.

Primer List

Click the Primer List button to access. The Primer List contains folders for the active and loaded primers (described below) and for the primer collections found in your home location (My Collections) or in a shared resource location (Shared Collections).

You can sort the information in the Primer List and you can resize the columns as needed to view the information most important to you. Click on a column heading to sort by the values in that column. Click again to reverse the sort order. Move the mouse to the join between two column headings, depress the mouse button and drag to resize the column. Use the mouse to drag out the dialog box borders to resize the Primer List.

Active / Loaded Primers

The folder Active/Loaded primers contains newly-created primers, primers loaded into the program from disk, and primers highlighted in search results windows (in parentheses). You can use this list to identify the active or loaded primer to use for the next operation, open the primer viewer window to see additional information about the primer, or remove primers no longer needed on this list.

Selecting a Primer -- You can identify the primer to use for the next operation or you can open the primer viewer window for any primer listed. Click on the name of the primer of interest and then click Select.

Removing a Primer -- You can remove a primer from the Active/Loaded folder. This does not delete a disk file (if the primer had previously been saved to disk). Click on the name of the primer you want to discard and then click Remove.

Loading another Primer -- Click Browse to access the File Open box to find and load a primer file, if needed. The file will be opened, the primer entered to the Primer List and then selected for use. Click DataBook to access the built-in database to find and open a specific primer file.

While viewing search results, you can use the Enter to Primer List function to formally add the active primer to the list of primers, entering a name and short description for the primer. Once entered to the Primer List, you can select this primer by name for analysis operations or save the primer to disk for later use. You can also enter additional comments about this primer in the Notes field. All of this information will be saved when a primer is saved to disk.

Primer Collections

Primer collections are stored in special primer collection files (*.px5). These collection files contain essential information about multiple primers, often a group of primers that are related in some way. Storing primer information in collection files streamlines storage and access. Just click on Primer List to access and use your primer collections. At this location, you can view, edit, or organize the primers in your collection or select a primer for use.

The information stored for each primer includes the primer name, type, sequence, description, and notes. Information about a source molecule or a binding site within that molecule is not stored in a collection file entry. If this information is important to you, use File, Save As... to save the primer and its associated molecule information to a disk file in standard primer file (*.pd4) format.

You can construct a primer collection using existing primer files or you can add new primers directly to a collection file. You can create new collection files, as needed, and create folders to organize your collection files. All your collection files are stored within the folder Collections, located in your home directory.

Shared Primer Collections -- If you are running the program under a multi-user license (Small WorkGroup or Concurrent User license types), you can share primer collection files, if these files have been made available for you. To access the shared resources, click File, Preferences, Share Data, and enter the location (complete path and folder) where you can access shared enzyme lists and primer collection files. When you open the primer list, shared collection files will appear in a folder below your primer collections. Shared collection files will be opened as read-only resources.

Constructing a Primer Collection

You can enter new primers directly to a collection file, you can add active (open) primers to a collection, or you can convert batches of existing primer files to collection files in one easy operation.

Click on the Primer List toolbar button to open the Primer List. Then click on the Collections folder to open folders or files in this location. When you are just getting started, you will have one empty file ready for your first primer to be added to this collection file. You can click on the first toolbar button Folder Options to add new folders or new empty collection files, or to rename or delete any of the files or folders you have created.

Adding to your collection:

Enter new primer -- If you have the sequence for a new primer, you can create the primer in place in the collection file. Open the Primer List and click on a collection file to open this collection. Then click the toolbar button Add Entry. Fill in the information about the new primer.

Add one primer -- When a primer is open in the primer viewer window, you can add this primer to one of your collection files. Just click File, Add to Collection and select the collection file. This primer can be one you created previously and stored to disk or it can be a primer you just created.

Convert batches of existing files -- You can use the Multiple File Conversion utility to prepare collection files from your existing *.pd4 or *.pd2 files. You can put all your primers in one collection file or, if you have a large number of primers, you can prepare a set of files to organize them by type, source, or intended use.

Create a new collection file -- If you want to place the primer in a new collection file, not yet in your collection file list, create the new file first. Open the Primer List, click My Collections, and then click the toolbar button Folder Options. Select New (empty) File and then give the new file a meaningful name. Now you can add your primer to this collection file.

Search in Primer Collections

You can search in your primer collection files to locate a specific primer or a group of primers that have a common text string in the primer entry. You can search using a primer name or any text that could be found in the name, description, or notes fields. To start a search

- Open the Primer List
- Click on the collection file or folder in the left panel that you want to search
- In the Find edit box (lower left), type in the text to search for
- Click the Go button

If primers are found that meet the search criteria, these primers will be listed in a new folder called Search Results, found at the bottom of the folder list in the left panel. (Click on the Search Results folder to open this folder, if it is not already open.) You can select a primer for use from this list or you can print the list of primers or copy it to the clipboard. When you close the Primer List, the Search Results folder will be discarded.

Using your Primer Collections

Once you have primers in your collection files, you can view, edit, organize, or select these primers for use. You can also search your collection files to find a specific primer.

Open the Primer List to use your primer collections. Click on the My Collections folder to open folders or files in this location. Use the Tab key or Shift + Tab to move between the left and right panels or click with the mouse.

Click on the Folder Options toolbar button to add new folders or new empty collection files, or to rename or delete any of the files or folders you have created.

Click on a collection file in the left panel to display the contents of this file in the right panel. For each primer in this collection, the primer name, description, sequence, length, and primer type are shown.

Click on column headers to sort by the values in that column. Move the mouse to the join between two column headings, depress the mouse button and drag to resize the column. Use the mouse to drag out the dialog box borders to resize the Primer List window.

View or Edit Primer Information -- Select a primer by clicking on the primer in the list in the right panel. Use the toolbar buttons to edit the information about this primer or view Notes information, delete this primer entry, copy the entry to paste into another collection file, or export the primer sequence. When changes have been made to your collection files, these changes are automatically saved for you.

Select a Primer for use -- Click on a primer and then click the Select button at the lower right to use this primer in an on-going operation or to open the primer viewer window for this primer. If more than one primer can be selected for an operation, use the Shift or Control keys in conjunction with the mouse to select additional primers.

Document collection file contents -- You can use toolbar buttons to print the contents of the active collection file or send this list of primers to the clipboard for use in another application. Or you can export the data in the active collection file in tab-delimited format to import into another application -- click the Folder Options button and select Export.

Design Primers

All design options are found on the Primer menu. You can design primers or primer pairs for PCR, you can find a primer matching an existing PCR primer, or you can design sequencing primers or probes. You can design primers by searching for primers that meet your criteria, by positioning a primer in the molecule sequence, or by typing in the sequence of a known primer. You can also use a special primer design wizard to design a set of sequencing primers or wizards for PCR Cloning, RT-PCR or LCR. Following primer design, you can analyze the primer or primer pair, save the primers to disk files for later use, or create amplified products using the primers and their source molecules.

Search for Primers

The program will search the molecule and region you specify, looking for primers that meet the evaluation criteria you set. You can use default criteria settings or modify these criteria to meet your own experimental goals. Primers found will be listed in a table format, showing basic information about each primer or primer pair. You can sort this list, send it to your printer, or analyze or save any of the primers found. Primers can be sorted by quality (rank sort) and you can affect ranking by adjusting the Sorting Weights, if required.

If you find fewer primers than you would like (or none at all), you can switch to an alternate view to see which evaluation criteria eliminated possible primers. You can use the Auto Adjust Length or Auto Adjust Criteria features to make adjustments and repeat the search to try to find additional primers.

Setting up a Primer Search

To find potential primers using an automated search procedure, you will need to provide information about the primer or primer pair you want to design and the source molecule and target region.

- Click the Design Primer button (main toolbar)
- Set the option Search for Primers and select the Primer Type.
- View or edit primer evaluation criteria, if necessary.
- Enter the primer length.
- Enter the 5' extension (if any).
- Click Next to identify the source molecule and target region.

Primer Length -- Enter the number of bases you want to use for your primer or for each primer in a primer pair. PCR primers are often 18-22 bases in length. You can indicate if this is the required or optimal length. If you select optimal length and set the acceptable range, primers of lengths within the range will be considered.

5' Extensions -- Use the checkbox to enable this option and enter the sequence for the 5' extension. Use the small button to the right of the 5' extension field to get help with restriction enzyme recognition sequences. Use the checkbox to indicate if the 5' extension should be added to the primer length entered above. (If you do not check this box, the primer length is assumed to include the 5' extension.)

After providing information about the primers you want to design, you can enter source molecule, target, and design goal information. Search options and the identification of target positions or target regions will vary depending on whether you are searching for primer pairs or single primers.

- Confirm that the source molecule shown is the one you want to use.
- Enter the target region or target position and use the checkbox if you want to find single primers using the complement strand.

- Select the search option and set related values.
- Click Finish to begin the primer search.

Amplify complete target region -- Find primer pairs that will amplify the entire region specified. Enter the start and end base position for the target region. The left primer (Primer A) will optimally be placed at the position you enter as the start of the target region and the program will extend its search to the left of this position by the number of bases you specify. The complementary primer (Primer B) will optimally be placed at the position you enter as the end of the target region and the program will extend the search to the right of this position by the number of bases you specify.

Amplify fragment in target region -- Find primer pairs that will amplify a fragment of a specific size anywhere within the target region specified. Enter the start and end base position for the target region and the acceptable product length range (minimum and maximum size in bps).

Use target position -- Find single primers close to a specified position. Identify a base position to be used as the starting point for the primer search. The program will extend its search to the left of the target position the number of bases specified and to the right of the target position the number of bases specified.

Anywhere in region specified -- Find single primers whose 5' starting position is located within the region specified. Enter the start and end base position to identify the target region.

Primer Search Results

When a search has been completed, the results will be displayed in the Search Results or Search Statistics windows. If the search was successful, primers found will be listed in a table format showing basic information about each primer or primer pair. You can sort the list of primers or pairs by rank or position, view a report for the selected primer, enter the primer to the Primer List for later use, or add the selected (highlighted) primer binding site to the molecule file.

Click the Primer Report button on the toolbar to view a report (or analyze) the selected primer. The primer viewer window will open with primer report data and access to more detailed analyses.

Click the Search results/Search statistics button to see where criteria eliminated primers. (If no primers were found, the Search Statistics display is the only display available).

For primers from DNA molecules, rank, basepair position, primer length, GC content and Tm are reported for each primer. Dimers shows the worst-case dimer with 3' end involvement and the worst-case dimer at any position. False Priming reports the Tm of the worst-case false priming site found for each primer, or shows '...' if no false priming found. Worst-case non-specific binding is reported for probes. For primers from protein molecules, rank, position (codon), start base (offset within codon), primer length, and average GC % are shown for each primer. Degeneracy reports the degree of ambiguity within the 6 bases at the 3' end of the primer or in the total primer sequence.

Sort by Rank places the best primers or primer pairs at the top of the list. Ranks are assigned to each primer or primer pair. Where equal scores are achieved, the same rank number is used and the equal sign (=) is affixed to the rank number for all similar scores. Sort by Position lists the primers found in ascending basepair position order. If the results show primer pairs, you can sort by the basepair position of the first primer or by the basepair position of the second primer (Primer B). If the results show primer pairs and the two primers can amplify a product, you can sort by the length of the amplified product.

The Primer Search Statistics Display is an alternate to the search results display and shows which criteria eliminated primers or pairs during this search. The program initially identifies all possible primers in your search region. It then begins to eliminate primers that fail to pass the tests shown. The test steps are sequential,

and only primers that pass a given test are considered in the following test(s). The number of primers tested and eliminated in each step is reported numerically and displayed as a bar chart on the right (eliminated primers shown in red).

Finding More Primers

When a search has been unable to find primers that meet all of the evaluation criteria, you can modify the search conditions (vary primer length or relax evaluation criteria) or you can redefine the search region to increase the size of the search range. Use the Redefine button in the results window to return to the search setup dialog box. Changing the primer length by one or two bases may have an effect on the number of primers found within the area specified and meeting the criteria selected. Or you may be willing to accept a slightly longer run of bases or an increase in non-3' dimers, without changing primer length.

Auto Adjust Primer Length -- Found on the Search Results/Search Statistics display window toolbar. You can use this function to start an automated sequence that will repeat the primer search, adjusting the primer length as specified, and then display the new results.

Auto Adjust Primer Criteria -- Found on the Search Results/Search Statistics display window toolbar. You can use this function to start an automated sequence that will adjust (relax) primer evaluation criteria, repeat the search, and then display the new results. Select the Search Statistics display to see which criteria eliminated primers. Select Conservative or More Aggressive to adjust (relax) the settings marked with checks. Select Custom to check the criteria settings that you want to adjust (relax).

Finding More Primer Pairs -- When a search has been unable to find either normal strand (Primer A) or complementary strand (Primer B) primers, you will want to view the results for this specific primer and take steps to find such primers (see Find More Primers, above). You can view separate statistics displays for Primer A, Primer B or the primer pairs. The results displayed are identified on the line above the bar charts (Drop-thru Tests Results: Pairs, for example). Click the Primer A or Primer B buttons in the lower right to change the display. When a search has found Primer A candidates and Primer B candidates, but is unable to find A:B pairs meeting your criteria, you can modify the matching criteria to accept larger differences between the two primers, find more single primers that can meet your matching criteria, or redefine the search to change the product length range (if applicable).

Create Primer from Molecule

You can create a PCR primer, sequencing primer, or probe using the sequence of a source molecule. The program will position a primer of the length you specify at the initial position you specify. You can then move the primer along the sequence or adjust its length until the primer evaluation meets the criteria set. This design option is especially useful if you must have a primer very close to a specific location in your molecule. To create a new primer:

- Click the Design Primer button on the toolbar
- Set the option Create from Molecule and select the Primer Type.
- Enter the primer length and the 5' extension (if any).
- Click Next to identify the source molecule and target position.
- Confirm that the source molecule shown is the one you want to use.
- Enter a target position for initial placement of the primer and use the checkbox if you want to design a primer using the complement strand.
- Click Finish to create the primer and open the Primer Viewer Window to view the primer and its evaluation.

Primer Length -- Enter the number of bases you want to use for your primer (required length).

5' Extensions -- Use the checkbox to enable this option and enter the sequence for the 5' extension. Use the small button to the right of the 5' extension field to get help with restriction enzyme recognition sequences. Use the checkbox to indicate if the 5' extension should be added to the primer length entered above. (If you do not check this box, the primer length is assumed to include the 5' extension.)

Target Position -- Identify a base position to be used as the initial 5' anchor point for the primer. Use the Features button to view a list of features (genes or regions) in the source molecule. You can select a feature and automatically enter the base position and strand.

View Primer and Evaluation

After identifying the position and length of a primer using the Create from Molecule primer design option, the program will create the primer and ask you for the primer name and description. The screen display of the new primer provides information about the primer and its evaluation, and about the molecule in the region around the primer.

The molecule sequence is shown across the center of the screen, with the 5' and 3' ends marked. For DNA molecules, a simple map showing genes or regions appears above the sequence. A horizontal line above the map may show tick marks at some locations. These are neighboring primers that will meet the criteria set. For protein molecules, the amino acid sequence is shown above the degenerate DNA sequence derived from this protein sequence. The primer sequence is shown in a boxed area below the molecule sequence. If a 5' extension is present, these bases are shown as non-homologous to the molecule sequence (no vertical match lines). The primer evaluation bar appears below the primer.

You can change the primer position or length, change the primer evaluation criteria, or edit the linked primer sequence to introduce a mutation. Click the Analyze tab to view analytical information about the primer. Use the toolbar buttons in the Analyze display window to access the more detailed analysis options, if desired. Click the Edit tab to return.

Adjusting Primer Position -- You can move the primer one basepair position at a time or you can jump to the next position that meets the criteria specified. The primer will be re-evaluated in its new position. Use the single arrowhead buttons or the keyboard left or right arrow keys to move the primer 1 bp position in the direction of the arrow. Use the double arrowhead buttons or the Ctrl + arrow keys on the keyboard to move the primer in the direction of the arrow to the next position that meets the criteria set.

Adjusting Primer Length -- You can increase or decrease the length of the primer, as needed. Click the Increase Primer Length button to add one base to the 3' end of the primer. Click the Decrease Primer Length button to remove one base from the 3' end of the primer.

Changing Primer 5' Extension -- You can modify the primer 5' extension, if necessary. Click the Edit 5' Extension button and then enter or edit the 5' extension for this primer.

Direct Entry of Primer

You can type in the sequence and other information needed to create a new primer. This might be a primer that was purchased commercially, received from a colleague, or previously used but not saved. To direct enter a primer:

- Click Primer, Direct Entry to start the Direct Enter Primer Wizard.
- Follow the wizard instructions to enter a primer name, primer type, and sequence (required information).

- You can also enter a description and notes for this primer (optional).

When the primer is complete, the primer viewer window will open and you can see the primer evaluation and have access to the analyze options.

Entering Basic Primer Information -- Enter the basic, required information for your new primer. You can later make changes to this information, if necessary. Enter a name for the primer. This name will appear on the primer list and in primer reports. Select the primer type. The primer type sets the criteria used to evaluate the primer.

Entering the Primer Sequence -- Enter the sequence of bases for your new primer. Sequence entry is in the Insert mode. Use the Delete key to remove erroneous bases, if necessary. A base counter shows you how many bases you have entered so far. You can also paste the sequence of the primer from the clipboard if you already have access to this sequence data in another application. Copy the sequence to the clipboard in your other application and then click the Paste button instead of typing in the sequence.

Entering Primer Description and Notes -- The description and notes information is saved as a part of the primer file format and can be edited at a later time. Enter a short description of the primer. This description will appear on some program screens and printed reports. Notes can contain more detailed information about the primer (source, how used, where stored).

Enter New Primer Directly in Collection

If you are using primer collections and you have the sequence for a new primer, you can create the primer in place in the collection file. To create a new primer in a collection file:

- Open the Primer List and click on a collection file to open this collection.
- Click the toolbar button Add New Entry.
- Fill in the information about the new primer.

Location -- For a collection primer, the collection file name is reported. For a new entry, the active collection file open at the left is shown -- this is the location where the new primer will be entered. The primer location cannot be edited.

Primer Name -- Enter a name for the primer. This name will appear on the primer list and in primer reports. If you plan to enter the primer site to a molecule map, a short primer name is helpful.

Primer Type -- Select the primer type. The primer type sets the criteria used to evaluate the primer.

Primer Sequence -- Enter the sequence of bases for your new collection primer or edit the sequence of bases for an existing primer, as needed.

Primer Description and Notes -- The description and notes information is saved as a part of the primer file or collection file format. Enter a short description of the primer. This description will appear on some program screens and printed reports. Notes can contain more detailed information about the primer (source, how used, where stored).

Use the Sequencing Primer Wizard

This wizard will help you to design a set of primers for sequencing a molecule region. The wizard will select three non-overlapping primers per sequencing block, where possible. You can re-order the primers in a block to pick another primer as your top choice, where needed. You can export primer sequences, save to a primer collection or view a map of sequencing coverage. Select the Sequencing Wizard from the Primer menu. Next, follow the wizard instructions to specify the molecule region and strand you want to sequence, enter primer

design preferences and specify which design criteria are most important to you so that primers will be ranked and selected accordingly.

Define What you Want to Sequence -- Identify the molecule that contains the region you want to sequence. Use the Change... button to select another molecule if the one shown is not the one you want to work with. Enter the upper strand coordinates of the region you want to sequence and use the option buttons to indicate if the region to sequence is on the upper/normal strand or on the lower/complement strand. You can use the Features... button to look at the features table for this molecule and select a gene, if appropriate. The program will enter the base positions for you.

Set the Primer Design Options -- For automated DNA or cycle sequencing, set the primer type to PCR Primer to set the appropriate primer evaluation criteria. You can leave the criteria set to default conditions, or you can edit the criteria settings for this primer search. If the annealing temperature (shown below the Criteria button) is not correct for this protocol, click the Criteria button and change this value. Enter the optimal primer length you would like and indicate the acceptable primer size range. Enter the number of bases preferred for each sequencing block and enter the values that reflect the trace constraints, or accept the default values.

Specify Weight Settings -- You can change the weight settings that are used to rank and select the best primers. You can indicate if distance from the optimal target position is more or less important to you, if the penalty for false priming should be increased or ignored, or if avoiding primer dimers is more or less critical to you. The standard (default) settings will usually give very good results.

Specify Search Preferences -- If the search for primers is not successful for any sequencing block, the program will attempt to adjust the block positioning and/or adjust the acceptable GC range and repeat the primer search procedure. If primers still cannot be found for a given sequencing block, the program will adjust (relax) the primer evaluation criteria for primers in this block to try and find primers you can use. You can accept the default settings or indicate which adjustments you will allow for each of these categories.

View Sequencing Wizard Results

The primers designed for the sequencing project are shown in a text display, grouped by sequencing block. For each primer listed, the data display shows the primer rank in this block (1, 2 or 3), the binding position for the 5' end of the primer to the template molecule, the primer sequence (5' to 3'), the primer length, and the %GC and Tm °C. If the primer was selected using relaxed primer criteria, the symbol !! follows the Tm value.

You can see one primer per sequencing block or three alternate primers per block. Use the drop-down control in the display toolbar to select First Choice or Three Alternates to control the number of primers to display. If difficult areas are encountered and primers cannot be found or if primers are not positioned to provide sequence for the entire block, the notation ** coverage alert ** is added to the block header line.

Use the Map Sequencing Coverage button to view an alternate display, showing the approximate extent of reliable sequence resulting from each primer-associated sequencing reaction. A map line for the region to be sequenced is shown at the top of the display. Blocks with possible incomplete coverage are marked red.

Each of the expected sequences is shown by a colored bar on a separate line. The position of the bar shows the approximate location of the sequence. A blue bar indicates sequence is being aligned 5' to 3', while a light red bar indicates the sequence is being aligned in the reverse direction (3' to 5'). You can use the drop-down control in the toolbar to view one primer-associated sequence per block or three alternates per block. Select (highlight) a sequence bar and return to the text display to identify the primer or its binding position. Click the depressed Map Sequencing Coverage button to return to the standard list display.

Use the Rotate Alternate Primers button to rotate (re-order) the primers in the block containing the selected (highlighted) primer. This will move an alternate primer to the First Choice position at the top of

the list for this sequencing block. The rank numbers (1, 2, 3) reflect the wizard's rank assignment and do not change with the rotation. (You can use these numbers to return the primers to the original order if needed.)

Use the Make New Primer Collection button to create a new primer collection containing all of the primers displayed -- either 1 primer (first choice) per block or 3 alternate primers per block, depending on the display style selected. You will be asked to enter a short identifier that can be used as part of the name for your collection and its primers.

Use the Export All Primer Sequences button to export primer sequences to a disk file or copy to the Windows clipboard for later use in another program. The wizard will export the primer data displayed -- either 1 primer (first choice) per block or 3 alternate primers per block, depending on the display selected. You can export primer sequences only, primer names and primer sequences, or a tab-delimited data set which also adds the primer description to the data.

Single-primer actions -- With the highlight bar on one of the primers, you can use the first three toolbar buttons to view the primer report for the selected primer, enter this primer to the Primer List for later use, or export the sequence for this primer.

Click the Redefine toolbar button to change (fine-tune) the sequencing primer search. You may want to change the preferred size of the sequencing blocks or modify the primer selection criteria or annealing temperature.

Export Primer Sequences

When you use the Export All Primer Sequences option in the sequencing primer wizard results, you can export primer sequences only, primer names and primer sequences, or a tab-delimited data set which also adds the primer description to the data.

Sequences only -- Select this option to export only the sequences of the primers, one primer per line, 5' to 3'. The order of the primers exported will be the same as the order in the results display.

Primer name and primer sequence -- This option will construct a name for each primer, using the 5-character identifier you entered, and the basepair position of the binding site in the molecule sequence. The default identifier is the first 5 characters of the molecule name.

Tab-delimited data set -- Select this option to export the following information, separated by tab characters. This format can be imported into other applications, like an Excel spreadsheet.

Name -- primer name consisting of the 5-character identifier you entered plus the basepair position of the binding site in the molecule sequence.

Sequence -- the primer sequence 5' to 3'

Characteristics -- the primer length (in parentheses), followed by the primer %GC value and the Tm value in degrees C.

Description -- includes the number of the sequencing block this primer was designed for, the rank number assigned for this primer in the group of 3 alternates, and the name of the molecule that contains the region to be sequenced.

Analyze Primers

You can analyze single primers, primer pairs, or a mix of primers and molecules. The results of the one primer or pair analyses are shown on the Analyze tab in the primer viewer window. The Primer Pair analysis includes information about the amplified product (if any) which could result from a PCR-type reaction using the two primers and their associated molecule. To set up a primer analysis:

- Click Primer, Analyze
- Select the Analysis Type
- For the one primer or primer pair analyses, click Select Primer to access the primer list to select the primers you want to analyze

Primers are also analyzed or evaluated during primer design using the Create from Molecule design option. The primer evaluation bar appears below the primer in the design window and can be used to jump to the appropriate analysis page using the special magnifying glass cursor.

To analyze a mix of primers and molecules, the Analyze Mix Wizard will help you to select the primers and molecules, set the analysis conditions, and display the results. You can also use the Primer Values Profile to calculate and plot primer-related values (Tm, GC content, and stability) to help you find regions for primers in difficult molecules or the Mutagenesis Profile to help you with primer or molecule mutations.

Primer Evaluation Bar

The primer Evaluation Bar appears at the bottom of the primer viewer window on the Edit tab, for molecule-associated primers, or on the Details tab, for primers that are not associated with a molecule. A boxed display area is used to show the primer evaluation. The primer is evaluated for GC content, melting temperature (Tm), presence of dimers, stability, GC clamp, presence of runs or hairpins, and possible false priming. Degenerate primers (designed using protein molecules) are evaluated for average GC content and degeneracy (3' end and total).

The icon at the left edge of the evaluation bar indicates whether or not the current primer satisfies all of the criteria set.

The primer meets all of the criteria set. Click the Analyze tab to find out more about this primer or print a primer design report. Click the Info tab for molecule-associated information (if applicable). Save the primer to disk for later use, if desired.

The primer does not meet one or more of the criteria set. Values that do not meet criteria will appear in color. You can click on the value in the Evaluation Bar to jump to the appropriate view on the Analyze tab for more details. (Click the Edit or Details tab to return.)

If you are designing a primer using the Create from Molecule option and the primer does not meet criteria, adjust the primer position or length until you find a primer that does meet criteria.

View Primer Analyses

Use the primer viewer window any time you want to view, modify or analyze a primer. To open the primer viewer window, select a primer from the Primer List. The primer viewer window will also open when you load a pre-existing primer into the program, when you click the Report button in a search results window, or when you set up a primer analysis operation. From within the primer viewer window, you can access functions to

view or edit the primer evaluation criteria, link or unlink primers with molecules, change the primer name or add descriptive information, or export the primer sequence.

For primers that are associated with molecules, you can click tabs at the bottom of the window to view or edit detailed primer information (Info), edit the primer sequence or position (Edit), or analyze the primer (Analyze) for possible dimers, runs of bases, hairpins, or false priming. For primers that are not associated with molecules, you can click tabs at the bottom of the window to view or edit primer information (Details), or analyze the primer (Analyze) for possible dimers, runs of bases or hairpins.

Primer Reports

You can view general analysis information in the form of a One Primer Report or a Primer Pair Report. The Primer Pair analysis includes a product analysis (found on Info tab), where appropriate. Click the Primer Report button on the toolbar in the Analyze tab of the primer viewer window or on the search results display screen. You can also click Primer, Analyze and select one or two primers from the Primer List for analysis.

One Primer Report -- The analysis shows the criteria settings used to evaluate this primer, the actual values for each item, and an indication of whether the criteria were met. For primers against DNA molecules, the criteria include % GC range, Tm range, 3' dimers, any dimers, stability, GC clamp, runs of bases, dinucleotide repeats, hairpins and worst-case false priming. For degenerate primers, the criteria include the % GC range, 3' end degeneracy and total degeneracy. In addition, the degenerate bases are noted.

The printed One Primer Report includes the information on this screen, as well as the more general information contained on the Info tab (primer sequence, length, composition, molecule information and map, if applicable). Primer Notes are added as the last section of the printed report.

Primer Pair Report -- The analysis of a pair of primers (usually two primers prepared from the same source molecule) shows a summary of the evaluated values for each primer and comments, where appropriate. Values that do not meet the evaluation criteria are shown in color. For primers against DNA molecules, the values include length, % GC, Tm, 3' dimers, any dimers, stability, GC clamp, runs of bases, dinucleotide repeats, hairpins and worst-case false priming. For degenerate primers, the values include length, % GC, 3' end degeneracy and total degeneracy. In addition, the degenerate bases for each primer are noted.

The printed Primer Pair Report includes the information on this screen, as well as the more general information contained on the Info tab (primer sequences, length, molecule information and map, if applicable, and product information if an amplified product would result) as well as primer evaluation criteria settings.

Analyze Dimers

Click the Dimers button on the Analyze tab screen in the primer viewer window. Use this option to display homologies between primers (primer dimers). The presence of dimers can reduce the efficiency of priming and can cause artifacts such as primer-dimer bands. Dimers involving the 3' end of a primer are particularly important to avoid.

For single primer analyses, the display shows the largest region of homology involving the 3' end and the largest region of homology in any position on the primer. For two primer analyses, the display shows a chart of interactions between the primers. The program reports the number of adjacent bases in the largest duplex found in interactions involving the 3' ends of primers or in interactions at any position in the primers. The primer interaction shown in the lower area corresponds to the highlighted chart value.

Analysis Cautions

Click the Cautions button on the Analyze tab screen in the primer viewer window. Use this option to view a summary of stability, GC clamp, runs, repeats and hairpins analysis information for single primers or pairs of primers.

Stability -- The program reports the differential stability between the 5' and 3' domains of the primer. If this value does not meet the criteria setting, it is shown in color. Also reported are the values of the 3' end pentamer and the maximally stable non-3' pentamer.

GC Clamp -- If enabled, the program reports the number of G's or C's at the 3' end of each primer. If the criteria value is set to 0, the 3' end of the primer is not checked for G's or C's.

Runs -- The program reports the longest run of identical bases for each primer. If the number of bases in the run exceeds the criteria value, these bases will be shown in color.

Repeats -- The program reports the highest number of dinucleotide repeats for each primer. If this number exceeds the criteria value, the repeats will be shown in color.

Hairpins -- The program checks for possible secondary structure (hairpins or stem loops) within the primer. If found, the program will show the structure of the hairpin and its energy in kcals at the temperature specified in the criteria settings.

Analyze False Priming

Click the False Priming button on the Analyze tab screen in the primer viewer window. Use this option to search the source molecule for 3'-end homologies (>= 6 bases) between the primer and the template. Mismatches or gaps can be included in these areas of homology.

If any regions of 3'-end homology are found, the display screen will show a list of these regions, sorted by Tm. Homologous bases are shown 5' to 3' and the basepair position and Tm for each region of homology are given, as well as the percent binding at the annealing temperature you set. An ellipsis (small dots) at the 5' end of the homologous base display indicates that the primer sequence was longer than the space allowed. The map in the lower area shows the actual primer (thicker, dark green arrow) and false priming sites (slender arrows). Arrows point right for normal strand, left for complementary strand and color arrows mark sites where the Tm is >=20°C.

Check for False Priming against another molecule -- Click the Check Other Molecule button and follow the instructions to check for false priming against a molecule other than the source molecule. When results are shown, an icon in the upper area will remind you that you are temporarily connected to another molecule. When you close this window, this transient connection will be discarded.

Analyze Any Homologies

Click the Any Homologies button on the Analyze tab screen in the primer viewer window. Use this option to search the source molecule for regions of perfect homology (>= 7 bases) between the primer and the template.

If any regions of homology are found, the display screen will show a list of these regions, sorted by Tm. Homologous bases are shown 5' to 3' and the basepair position and Tm for each region of homology are given, as well as the percent binding at the annealing temperature you set. Homologous bases are aligned with the corresponding primer sequence, when possible. An ellipsis (small dots) at either end of the homologous base display indicates that the primer sequence was longer than the space allowed, preventing alignment of bases.

The map in the lower area shows the actual primer (thicker, dark green arrow) and regions of homology (slender arrows). Arrows point right for normal strand, left for complementary strand and color arrows mark sites where the Tm is >=20°C.

Analyze Mix Wizard

Found on the Primer menu. You can use this option to analyze primer interactions, binding, and products formed from a mixture of primers and molecules. This wizard will help you to identify the primers and molecules to add to the mix and set parameters for the analysis.

When the analysis has been completed, a list display will show you the possible products produced, binding sites, and any cautions noted (primer dimers that might affect the results or molecules or primers that are not used). You can also use this wizard to analyze possible interactions from a group of primers where no molecules are present in the mix and no products are produced.

Identify Primers -- Click the Add Primer button to access the Primer List to select the primers to add to the mix. You can select more than one primer at the same time by using the Shift or Control keys in conjunction with the mouse. If primers are linked to a molecule, you do not necessarily have to reload the original source molecules to use the primers in this analysis.

Identify Molecules -- Click the Add Molecule button to access the Molecule List to select molecules to add to the mix. You can select more than one molecule at the same time and you can use the Browse button to find molecules not yet loaded. If you do not add any molecules to the mix, the program will simply report possible primer interactions.

Set Mix Parameters -- Select the primer type. This information is used to preset annealing temperature and dimer criteria. You can change these settings, if needed. You can also set the cutoff for percent priming needed to add a result to the list or you can accept the default setting. Use the checkboxes to indicate which results categories you would like to view/print.

Analyze Mix Results

When the analysis has been completed, the results will be displayed in a scrolling window.

Products -- Possible products will be listed in the first section, grouped by molecule. The product length, binding positions, and the two primer names are shown for each product. You can select (highlight) a product and create this product or see a detailed report.

Click the Create Product button on the toolbar to create this molecule. Set the option to enter the molecule to the Molecule List for use now, save to a disk file for later use, or export the product sequence.

Click the Primer Report button to view a report that gives detailed information about the two primers involved in creating the product, as well as additional information about the amplified product.

Binding Sites -- Binding sites are reported in the next section, grouped by primer. The molecule name, binding position, number of homologous bases, number of mismatches, and Tm are shown for each binding site. You can select (highlight) a primer and see a detailed report about this primer.

Analysis Cautions -- Analysis cautions are shown in the last section. 3' end primer dimers or any primer dimers that exceed the warning levels set will be noted here. The primers involved and the homologous bases are shown for each interaction. Molecules or primers that are not used are listed at the end of this section.

Primer Values Profile

Found on the Operations menu. You can profile regions of your molecule and plot primer-related values to help you find regions for primers in difficult molecules. The program will calculate Tm, GC content and stability for a virtual primer of specified length. (For protein molecules, values are calculated for average GC content, 3' degeneracy and total degeneracy.) The graphs cover approximately 200 bases of molecule sequence, centered around the target position specified. Primer values are plotted at the 5' position for each primer. Use the Redefine button to select another target position.

Identifying acceptable values -- Values that meet the criteria specified for PCR primers will appear in the region of each graph marked with a green bar at the right. Tm and GC range limits are represented by horizontal rules, as is the stability cutoff value (DNA) or the degeneracy cutoff values (protein). Click on an area of the graph where all three values are plotted in the green ranges. The vertical ruler line will move to this position and you can read out the basepair position number at the top of the line.

Using the vertical ruler line -- The vertical reference line initially marks the target position you specified. You can move this ruler line in 10-base increments using the toolbar or keyboard left or right arrow keys. You can also click with the mouse on an area of interest and the ruler line will move to this new location and report its basepair position at the top of the ruler line.

Changing the primer length -- The length specified for the virtual primer is reported in the toolbar area. You can use the toolbar Increase Primer Length or Decrease Primer Length buttons to change the length of this virtual primer. The program will recalculate the values and replot the graphs.

User Settings

Primer Criteria -- To find or evaluate primers, the program refers to the values specified in the Primer Criteria settings for GC Content, Melting Temperature (Tm), Annealing Temperature, Stability 5' vs 3', Dimers, Runs, Protein Degeneracy and GC Clamp. Separate criteria settings are stored for PCR primers, sequencing primers, probes or LCR oligos. You can change the criteria settings for an individual primer or search, if needed. You can also change the default settings so that all new primers of that type use the modified criteria settings. If necessary, you can change the type of criteria used to evaluate an existing primer.

Changes to primer criteria may affect the number of primers found in a primer search. Click the Criteria button to change primer criteria for one primer or a primer search. Click Primer, Default Primer Criteria to change default settings for all new primers.

Experimental Conditions -- The program also stores the concentrations of salts and oligos and uses these values for calculations. You can also select the energy parameters the program will use to calculate Tm. Click Primer, Default Primer Criteria and then click the Conditions tab to change these settings.

Sorting Weights (Primer Ranking) -- Search results consist of a list of primers that have passed an evaluation using the Primer Criteria settings. You can sort the resulting list of primers by rank so that the best primers will be at the top of the list. Ranks are assigned to primers based on scores calculated by multiplying Sorting Weights by numbers the program generates for certain specific quality-associated conditions. You can modify the Sorting Weights for a completed search to increase the penalty for a condition and thereby change the position of primers in the list. The program stores default Sorting Weights for PCR primers, sequencing primers, probes and LCR oligos and you can change the default settings for these weights, if desired.

Changes to sorting weights will not affect the number of primers found in a search, but only their position in the list when sorted by rank. Click the Sorting Weights button in a search results display to change penalty values for these results. Click Primer, Default Weights and Values to change default settings.

Primer Operations

You can view or edit information about a primer (name, description, notes), or you can edit a primer sequence, if necessary, or export or copy this sequence to the clipboard. You can also link a primer to a molecule or create an amplified product using a pair of primers. You can find sites where primers in your collection could bind effectively to the active molecule, or search for molecules that have binding sites for a primer of interest. Primer sites can be added to molecule files for display on the molecule map or sequence and primer sites on the molecule map can be used for a number of actions.

View, Edit Primer Information

You can view information about a primer by opening the primer in the primer viewer. Click the Info or Details tab to view detailed information. The Info tab provides information about primers that are associated with a molecule. This includes the primer sequence, length, whether it meets the evaluation criteria, bp position information, and a molecule map. For single primers, composition and molar absorbance are reported, where appropriate. For primers that are not part of a search in progress, the header area will show the primer name, description and the first part of the notes information. You can modify this information, if needed. You can also check the criteria settings.

For primer pairs, product information is included if an amplified product would result. Amplified product size, GC content, melting temperature, and annealing temperature are reported. Note that PCR optimal annealing temperatures are calculated based on the method of Rychlik et al (see References) and use salt and oligonucleotide concentrations that may differ from those you have set as defaults for your experimental conditions. You can click the Create Product toolbar button to create a new DNA molecule that is this amplified product.

The Details tab provides information about primers that are not associated with a molecule. Detailed information includes the primer sequence, length, composition and molar absorbance. The analysis criteria type is reported and the Evaluation Bar at the bottom of the window is used to indicate if the primer meets the evaluation criteria. In addition, the header area will show the primer name, description and the first part of the notes information. You can modify this information, if needed. You can also link the primer to a molecule or check the criteria settings.

Edit Primer Definition

You can view or edit the name of a primer, or the description and notes information entered for this primer. To edit the primer definition:

- Open the primer viewer window and click the Info or Details tab.
- Click the Edit Primer Definition button
- Enter or edit the information, as needed.

The primer name will be used to identify the primer on program screens and reports. The short description of the primer appears on printed reports. The longer notes information can be saved to the primer file or written to a DataBook entry.

Edit Primer Sequence

You can change the bases in the sequence of a primer, if needed. The way in which you change the primer sequence will depend on whether the primer is associated (linked) with a molecule. For primers that are not

linked to a molecule, these changes are straight-forward -- you simply add, remove or change bases as needed in the primer sequence. To change the unlinked primer sequence:

- With the primer viewer window open to the Details tab, click the Edit Primer Definition toolbar button
- Click in the Sequence field and then make changes needed.
- Use the keys for A, C, G or T to add bases. Use the delete key to remove bases, as needed.
- Click OK when your changes are complete.
- Check the Evaluation Bar to see how these base changes have affected the evaluation of the primer.

If the primer is linked to a molecule, the change of one base for another will produce a mismatch with the molecule sequence at that point. This can be used to introduce a mutation site. You can also effectively change the sequence of bases in a primer by changing the position of a primer within a molecule or by changing the length of a primer that is linked to a molecule.

Click the Edit tab to move or modify the primer within the molecule sequence. The screen display provides information about the primer and its evaluation, and about the molecule in the region around the primer. The molecule sequence is shown across the center of the screen, with the 5' and 3' ends marked. For DNA molecules, a simple map showing genes or regions appears above the sequence. A horizontal line above the map may show tick marks at some locations. These are neighboring primers that will meet the criteria set. For protein molecules, the amino acid sequence is shown above the degenerate DNA sequence derived from this protein sequence.

The primer sequence is shown in a boxed area below the molecule sequence. If a 5' extension is present, these bases are shown as non-homologous to the molecule sequence (no vertical match lines). The primer evaluation bar appears below the primer. To edit a primer that was designed previously, click the Redesign Primer button to activate the toolbar buttons that will allow you to move the primer or change its length.

Adjusting Primer Position -- You can move the primer one basepair position at a time or you can jump to the next position that meets the criteria specified. The primer will be re-evaluated in its new position. Use the single arrowhead buttons or the keyboard left or right arrow keys to move the primer 1 bp position in the direction of the arrow. Use the double arrowhead buttons or the Ctrl + arrow keys on the keyboard to move the primer in the direction of the arrow to the next position that meets the criteria set.

Adjusting Primer Length -- You can increase or decrease the length of the primer, as needed. Click the Increase Primer Length button to add one base to the 3' end of the primer. Click the Decrease Primer Length button to remove one base from the 3' end of the primer.

Changing Primer 5' Extension -- You can modify the primer 5' extension, if necessary. Click the Edit 5' Extension button and then enter or edit the 5' extension for this primer.

Mutate Linked Primer Sequence -- You can change one or more bases in the sequence of a primer that is linked to a molecule in order to create a mismatch or mutation site. When making changes, you can see the primer bases in a scrolling list, along with the molecule bases at the associated position. To mutate the primer sequence:

- Click the Edit Primer Sequence toolbar button
- Select (highlight) the base you want to change. Press A, C, G or T to overwrite the primer base letter with the new base letter.

Mismatched bases are marked in the scrolling list with *m* and on the edit screen display with a red asterisk (*) character. Mismatches will affect the evaluation of the primer.

Link to Molecule

Found on the Info or Details tab in the primer viewer window or on the Operations menu. This option is available when viewing a single defined primer (not part of a search in progress). You can use this option to search the molecule you specify, attempting to find possible sites where the primer could initiate priming. If priming sites are found, you can add a primer binding site to your molecule map or you can link a copy of this primer to the molecule at the site you select. You can also create an unlinked primer by making a copy of this primer that is not linked to any molecule. To link a primer copy to a molecule:

- Open the primer viewer window for the primer you want to link and click the Info or Details tab.
- Click the Link to Molecule button and set the Link option.
- Select the molecule you want to link to and click OK to begin the search for possible priming sites.
- Select the position for the primer-to-molecule link and press Link.

A copy of the primer will be linked at the position specified. Non-homologous bases at the 5' end will be treated as a 5' extension. As the primer is entered to the Primer List, you can modify the primer name or description, as needed. You can use this new primer now or you can save the primer to disk for future use.

Priming Sites Found -- If the Link to Molecule operation has found possible sites where the primer can initiate priming, each of the possible priming sites are shown in a scrolling list. If you attempt to link a primer to its own source molecule, the search will show you possible priming sites other than the active linked site. The potential binding site can contain mismatched bases (reported in the column headed *m*). Non-homologous bases at the 5' end will be treated as a 5' extension. You can select one of these positions and press Link to create a copy of the active primer linked to the molecule at the position specified.

Add Site to Map -- Click the Add Site button to add the selected (highlighted) primer binding site to the molecule file. This site can appear on the molecule map or the formatted sequence.

To unlink a primer copy from the current molecule:

- Open the primer viewer window for the primer you want to unlink and click the Info tab.
- Click the Link to Molecule button
- Set the option Unlink and click OK.

A copy of the primer with no molecule association will be made. As the primer is entered to the Primer List, you can modify the primer name or description, as needed.

Create Product

Found on the Primer menu. You can use this option to create a new DNA molecule that is the amplified product of a PCR-type reaction using two primers and a DNA molecule that you specify. If an amplified product can be produced, the results screen will show the product size, GC content, melting temperature, and annealing temperature and you can decide if you want to create this product.

If an amplified product cannot be produced using standard PCR reaction conditions, you will have the option to relax these conditions and try again. For this operation, Standard PCR reaction conditions use the default PCR annealing temperature (55°C if default unchanged) and a percent binding cutoff of 50%. Relaxed conditions lower the annealing temperature by 10C° and the percent binding cutoff to 20%. If more than one product is possible, you will have the option to open the Analyze Mix Wizard to view the more detailed results. In the mix wizard results, you can select (highlight) a product and click the Create Product toolbar button to complete this operation.

Identify components to create product:

- Click the Select Molecule button to access the Molecule List to select a DNA molecule for the procedure.
- Click the upper Select Primer button to access the Primer List and select the primer for the upper strand (forward primer).
- Click the lower Select Primer button to access the Primer List and select the primer for the lower strand (reverse primer).
- Click OK to create the amplified product.

With the two primers and DNA molecule specified, this procedure will attempt to create a new DNA molecule that is the amplified product of a PCR-type reaction. If an amplified product can be produced, the results screen will show the product size, GC content, melting temperature, and annealing temperature. You can create a molecule for use now or later, or you can export the product sequence to an ASCII text file.

Enter Product Molecule to Molecule List -- The amplified product will be created as a molecule file containing sequence, features (if any), name and description. This molecule will be entered to the Molecule List and the molecule viewer window opened, displaying this molecule for use. You can view the sequence, features or restriction map information. The file will not be saved to disk unless you specifically use the File, Save function.

Save Product Molecule to Disk File -- The amplified product will be created as a molecule file containing sequence, features (if any), name and description. This file will be saved to disk, using the filename you specify. The molecule will not be entered to the Molecule List and will not be available for immediate use. You can, of course, load the file using File, Open.

Export Product Sequence -- The sequence of the amplified product will be exported in ASCII text format to a disk file, using the filename you specify.

Find Primers that Bind

Found on the Primer menu or use toolbar button to access. You can use this function to find sites where primers you have collected could bind effectively to the active molecule. You can scan primer files in folders on your computer or local network. Primer files created in Primer Designer ver 3, 4 or 5 (*.pd2 or *.pd4 files) can be read by the program, as well as primer Collection files (*.px5). The program will open each primer file, read each primer sequence and analyze its binding potential.

You can set the binding cutoff, to adjust the stringency of the search, and you can set the binding temperature. (Standard conditions are at least 50% binding at a binding temperature of 55°C.) When the search has been completed, you can view the search results. Primer binding sites found can be added to the molecule file and can be displayed on the molecule map or sequence display.

Selecting primer files or folders to search

Click the Add Folder button to select a folder location to add to the search list. Navigate to the folder you want to add to the list and click OK. The program will attempt to open and read all primer files in the location specified, including those contained in subfolders.

Click the Add File button to select a primer file to add to the list. Navigate to the file you want to add to the list and click OK. The program will attempt to open and read this file during the search procedure.

Click the Add Collection button to access your primer Collections and select a file to add to the search list. Select the collection file you want to add to the list and click OK. The program will attempt to open

the file and read all the primer data contained in the file. If you have selected a folder of collection files, the program will attempt to open each file within the folder.

When you have selected the files and folders to search, click OK to begin the search for primer binding sites.

Find Primer Sites Results

When a search for primer binding sites has been completed, the results will be presented in a scrolling list box and location map display. You can add some or all of these sites to your molecule file for display on the map or formatted sequence. In the results display, primer binding sites are listed, showing the primer name and the basepair position of the 5' end of the binding site. Also reported are the length of the binding site (LenB) and the length of the primer (LenP). If there are mismatches within the binding site, the number of mismatched bases is shown in the column *m*. Click on the column headers to sort by primer name, position, or length of the binding site. Vertical bars above the icon map show the approximate location of the primer binding sites. A red arrow and red vertical bar mark the highlighted primer binding site.

To tag primer sites you want to add, mouse-click on the checkbox in front of the primer name. Click the Tag All button to automatically tag all of the primer sites in the list. To untag a site, mouse-click again or click the Clear button to untag all sites in the list. Click the Add Sites button to add all of the tagged primer binding sites to the molecule file and close the results window.

Display Primer Sites

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

Map Display -- To view primer binding sites on the molecule map:

- Open the molecule in the molecule viewer window and click the Map tab
- Click the Enzyme Sites/Primer Sites toolbar button to switch to primer sites.

Primer binding sites entered for this molecule will appear on the molecule map (in place of enzyme sites). The map sites list (right) shows the primer name and the 5' end of the primer binding site. Click the column headers to sort by primer name or position. You can select a primer site and then use this primer. Toolbar buttons in the map display window can be used to export the primer sequence, create the primer, use the primer to create an amplified product, change the primer name, or view the primer site properties.

Sequence Display -- To display primer binding sites on the formatted sequence:

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

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

Adding Primer Sites to the Molecule

Primer binding sites can be added to your molecule file from the Find Primer Sites Results display, the Link to Molecule results display, the primer design search results display, or from the primer viewer window (Edit tab)

for a specific primer. Once these sites have been added to the molecule file, they can be displayed on the molecule map or in the formatted sequence.

Find Primer Sites Results display -- Click the Add Sites button to add all of the listed primer binding sites to the molecule file.

Link to Molecule results display -- Click the Add Site button to add the selected (highlighted) primer binding site to the molecule file.

Primer Design Search results display -- Click the Add Primer Site to Map button to add the binding site for the selected (highlighted) primer or primer pair to the molecule file.

Primer Viewer Window, Edit tab -- Click the Add Primer Site to Map button to add the binding site for this primer to the molecule file.

The name of the primer and its description becomes a part of the molecule file and can be viewed using the Properties button on the map display. The primer name will appear on the map and on the sequence display, so a short but descriptive name is helpful.

Using Primer Sites

While you are looking at the primer sites on your molecule map, you can select a primer site and view the primer site properties, export the primer sequence, create the primer, use the primer to create an amplified product, or change the primer name. (If you do not have any primer sites on your molecule map, you can add primer sites to a molecule file in several ways.)

Switch to Primer Sites -- To view primer binding sites on the molecule map, open the molecule in the molecule viewer window and click the Map tab. Then click the Enzyme Sites/Primer Sites toolbar button to switch to primer sites.

View Properties -- Click the Properties button to view an information pop-up that shows the primer name, length, basepair positions that mark binding site on the molecule, sequence, and description.

Jump to Sequence -- Click the GoTo Sequence button to jump to the formatted sequence at the selected primer site.

Create (primer or product) -- Click the Create button and select option from menu. You can create a primer from the information stored in the molecule file or you can create a new DNA molecule that is an amplified product of the active molecule, using the primer selected and another primer from the molecule map that you can select during the operation.

Export Sequence -- Click the Export Sequence button to export the sequence of the primer to a disk file or copy to the Windows clipboard for use in another program.

More Actions -- Click the More Actions button and select option from menu. You can select Wipe All Primer Sites to remove all of the primer sites from the molecule map in one operation, or you can select Change Selected Primer Name to change the name of the primer that will appear on this molecule map.

Internet BLAST Primer Searches

Found on the Primer menu. You can set up a remote search using the NCBI BLAST server. You need an internet connection to use this function. This search is the equivalent of using the BLAST Program Search for Short, Nearly Exact Matches, recommended to find primer binding sites. The NCBI documentation suggests a common use is to check the specificity of primers used in PCR reactions or hybridization.

The program will help by handling query sequence formatting, setting the expect value and default word size parameters to values appropriate for short primer-length sequences, and will send your request to the server over the internet. The built-in web browser will show results and allow you to access molecules of interest.

To initiate a BLAST search, you will want to identify the query sequence (primer), and then select the BLAST database to search.

- Click Primer, BLAST primer (NCBI).
- Click the Change... button to select a different primer from the primer list, if needed.
- Use the combo box to select the BLAST database you want to search. A description of the selected database is provided below the selection.
- Click the OK button to prepare your query sequence and send your request over the internet to the NCBI BLAST server.

If you are using this operation for the first time, you can click the Options button to view or modify your internet access configuration settings. For BLAST searches, the query sequence should contain no ambiguous bases.

Scan for Molecules with Sites

Found on the Primer menu. You can use this function to scan molecule files to look for sites where the specified primer could bind effectively. You can scan molecule files in folders on your computer or local network. The program will open each molecule file, read the sequence and analyze primer binding potential.

You can set the binding cutoff, to adjust the stringency of the search, and you can set the binding temperature. (Standard conditions are at least 50% binding at a binding temperature of 55°C.) When the search has been completed, you can view the search results. You can print or copy the list of molecules and binding sites found.

Selecting molecule files or folders to search

Click the Add Folder button to select a folder location to add to the search list. Navigate to the folder you want to add to the list and click OK. The program will attempt to open and read all molecule files in the location specified, including those contained in subfolders.

Click the Add File button to select a molecule file to add to the list. Navigate to the file you want to add to the list and click OK. The program will attempt to open and read this file during the search procedure.

If you routinely use the same molecule file locations, you can save the list of files and folders and recall this list at a later time. Click OK to begin the scan for molecule sites.

Find Molecule Sites Results

When a search for molecule sites for the specified primer has been completed, the results will be presented in a scrolling list box. In the results display, the molecule name is listed, as well as the basepair position of the 5' end of the binding site. Also reported are the length of the binding site (LenB) and the length of the primer (LenP). If there are mismatches within the binding site, the number of mismatched bases is shown in the column *m*. Click on the column headers to sort by molecule name, position, or length of the binding site. You can print or copy the results list or load a molecule file for use.