Oligonucleotide Analysis

The Custom Oligonucleotide Synthesis Report accompanying each custom oligonucleotide contains extensive thermodynamic and physical data for each oligo. This technical sheet explains each of the values and the method used for arriving at a particular physical data. The figures may differ for the same oligo sequence by using different software applications, the reason being the selection of particular details and constants for arriving at the calculated value. An example is the salt concentration for calculating Tm., Gene Link software uses 100 mM salt concentration which closely approximates the total salt concentration in either PCR or sequencing reactions.

1. Size

The size is the length of the oligo. This is exactly the number of bases in the oligonucleotide sequence. In case of modified bases including 3' and 5' modifications, the software assumes these as bases and is represented in the size of the oligonucleotide. This will be accurate if these were modified bases (e.g. inosine, deoxyuridine, etc.) and *inaccurate* if the modification is not a modified base (e.g. phosphate, amino linker etc.)

2. MW

MW is the molecular weight of the oligonucleotide. The following weights are used for arriving at the calculation.

To enter an oligonucleotide directly, type the oligo in the Oligo window. The maximum number of bases is 100. DNA or RNA may be entered. Inosine, I, is permitted as a base in the sequence. To select an oligonucleotide from a sequence, the sequence must be the currently active window. Highlight the desired oligo sequence then open this analysis. Any highlighted bases in the sequence are automatically entered 5' to 3' in the oligo box up to the limit of 100 bases. Alternatively, set the From: to the 5' base number of the oligo. Set the Length to the length of the oligo. Set the strand to Top or Cmp. The oligo will be displayed in the Oligo window. Note: The 3' bases of the oligo are displayed in the sequence window, use the Home key to display the 5' end on larger oligos.

Mol Wt is the calculated molecular weight of the oligonucleotide. This is calculated from the table of molecular weights for each base found in the appendix A.

Tm is the melting temperature (temperature at which 50% of the oligo is a duplex) of the oligonucleotide calculated using the nearest-neighbor thermodynamic values methods of Breslauer et. al for DNA and Freier et. al for RNA. Note: this is not accurate for long sequences. The formula for the Tm is:

Where H is the enthalpy, S is the entropy, R is 1.987 cal K^{-1} mol^{-1} , and C_T is the total strand concentration. The formulas used for calculating H and S are described below.

Filter Tm is a filter hybridization calculation. It is the thermodynamic Tm - 7.5. Useful as an initial temperature when washing filters.

%GC Tm is based on the %GC and is corrected for salt concentration and % formamide. The formula of Baldino et. al is 81.5 + 16.6 log[Na+] +0.41(%G+%C) - 0.65(% formamide) - 675/length. This formula is more accurate with longer oligos and is useful if hybridizing in formamide or high salt. The salt concentration is set

with the Salt con (mMol) parameter. The default is 1 M. When calculating an oligo for PCR or sequencing, use 50 mMol. The % formamide concentration is adjusted with the % Formamide parameter. The default is 0.

GC + AT Tm is the traditional 2 for each A or T plus 4 for each G or C. This method is considered less accurate and is provided for reference.

nMol/A260 is the concentration of the oligo in nanomoles per O.D. at 260 nm. This is calculated from the table of extinction coefficients and molecular weights for each base in the appendix A.

ug/A260 is the concentration of the oligo in micrograms per O.D. at 260 nm. This is calculated from the table of extinction coefficients in the appendix A.

%GC is the percentage of G and C in the oligo. The calculation is the total number of G and C divided by the total number of bases in the oligo. It is used for the Tm calculations.

dG is the G (free energy) for the oligo as calculated by the nearest neighbor method of Breslauer et. al for DNA and Freier et. al for RNA. The G is calculated by the formula G = H-TS. Where H is the enthalpy, S is the entropy and T is the temperature set by the dG Temp. The values used to calculate the G for an oligo are listed in the appendix A. The G is a measure of stability, the greater the negative value, the more stable the duplex formed by the oligo. For example, G of the oligo ATGCTT is the G for AT + G for TG + G for GC + G for CT + G for TT. The value is corrected for temperature if the dG Temp is not 25. The initiation value of +5 for DNA and + 3.6 for RNA is added to the G.

dH is the enthalpy of the oligo as calculated by the nearest neighbor method of Breslauer et. al for DNA and Freier et. al for RNA. The values to calculate the delta H (H) for an oligo are listed in the appendix A. For example, H of the oligo ATGCTT is the H for AT +H for TG + H for GC + H for CT + H for TT.

dS is the entropy of the oligo as calculated by the nearest neighbor method of Breslauer et. al for DNA and Freier et. al for RNA. The values to calculate the delta S (S) for an oligo are listed in the appendix A. For example, delta S of the oligo ATGCTT is the S for AT + S for TG + S for GC + S for CT + S for TT. The initiation value of 10.8 is added to the calculation.

3'- End dG is the nearest-neighbor G for the last 7 (default value) bases of the 3' end. The number of bases may be changed with the 3'-end length parameter. The G is calculated by the formula G = H-TS. Where H is the enthalpy, S is the entropy and T is the temperature set by the dG Temp. The initiation value of +5 for DNA and + 3.6 for RNA is added to the G. Lower numbers are more stable. Values which are very negative may lead to false priming, values which are positive may reduce primer binding efficiency.

<Sense oligo> or < Antisense oligo> label displays the strand the oligo is located on. The Switch oligos button toggles between the two. Note: Sense and antisense are relative, depending upon which oligo was entered first and what the oligo window is set to.

From: is the number of the 5' base of the oligo in the sequence. It is set to the cursor position if a sequence is open and active. When the strand and length parameters are set, the oligo is displayed in the Oligo window. Any contiguous string of bases in the sequence less than 100 bases may be checked for use as an oligo with this parameter. When a sequence is highlighted in the sequence window, the value is the number of the base at the 5' end of the highlighted sequence.

Len: is the length of the oligonucleotide. If the oligo is entered directly in the Oligo window, the length is automatically calculated. Len: parameter also determines the length of the oligo if From is used to extract an oligo from a sequence. When a highlighted sequence is used, the Len: is the length of the highlight up to the limit of 100 bases.

Strand determines the strand the oligo is selected from.

Top and Cmp are used in conjunction with the From: parameter. The default is the sense (top) strand. If the complementary strand is selected the bases are still entered 5' to 3'. When a sequence is highlighted, the strand depends on the strand highlighted.

Molecule selects DNA or RNA as the molecule.

DNA and RNA change the molecule type. All parameters dependent on the type of molecule are recalculated. Thermodynamic parameters, molecular weight and concentration values are calculated for RNA using the RNA values listed in the appendix A. When entering an RNA sequence, U is required in place of T. If RNA values are not available, DNA values are used.

Edit sense or Edit antisense button displays a dialog box in which a second oligo may be entered for comparison to the currently displayed oligo. If the current oligo is a sense oligo then the Edit antisense is displayed. Conversely if the current oligo is antisense then the Edit sense button is displayed. The two oligos are then checked for secondary structures and physical characteristics. This permits checking existing primer pairs for complementarity and Tms, etc.

Switch oligos button switches the analyzed oligo from sense to antisense and back again.

dG Temp: is required for the calculation of the free energy (dG). This affects all G calculations (3' end dG, hairpin loops dG, and dimer dG) as well as hairpin loop Tm. The default value is 25.

Probe con (pico Mol) is required for the Tm calculation. The default is 0.6, but values of 250 or greater are used by some researchers for PCR applications. The default value is accurate for most filter hybridizations. Setting the value to 100 along with 330 mM salt gives accurate thermodynamic Tms when washing filters in 2X SSC.

Salt con (milli Mol) is required for the Tm and %GC Tm. The default is 1000mM (1M). For filter hybridizations, the appropriate salt concentration should be entered. The most common hybridization buffers use 2X SSC which is 330 mM.

% Formamide is required for the %GC Tm . This value is defaulted to zero. When hybridizing in formamide, adjust this to the appropriate concentration.

3'-end len: is the number of bases used for the 3'- End dG . This value is defaulted to 7. A value of less than five or greater than nine, may not provide an accurate indication of the 3' end stability.

Base run >= parameter sets the minimum length for a run of like bases to be listed in the Base runs / Palindromes box. Base runs may permit false priming.

Pal len >= parameter sets the minimum length for a palindromic sequence to be listed in the Base runs / Palindromes box. The default value is 6. This feature is useful for identifying restriction sites in an oligo.

Stem length is the minimum number of base pairs in a hairpin stem required to be listed in the hairpin box. This value is used by the duplex analysis filters as the minimum stacking length for a duplex.

Base runs / Palindromes is the run/palindrome list. Any base runs or palindromes that exceed the run and palindrome length parameters are listed in this box. If more than one is found ,they may be viewed by clicking on the down arrow. The box lists P for a palindrome, and R for a run of bases. The start base and length of the run/palindrome are listed before the sequence.

Secondary structures are displayed in the box at the bottom of the window. The Hairpin loops, Dimers, Bulge loops and Internal loops buttons toggles between displaying the indicated secondary structures. Bulge

loops and internal loops are variations of dimers. Gene Runner considers only uninterrupted complementarity between sequences as dimers. If the complementarity is separated by an equal number of bases on both strands then it is an internal loop, if it is separated by an unequal number of bases it is a bulge loop.

The current and total number of the selected secondary structure is displayed in the boxes in the upper left. The structures are sorted according to the characteristics selected with the Sort function. Only one structure is displayed at a time. If Hairpin loops is selected, the hairpin loop is displayed along with its stem length, position and loop size. Vertical bars indicate the base pairs in the hairpin. If Dimers, Bulge loops or Internal loops is selected, the dimer/bulge loop is displayed along with the location and size of the largest stack. A vertical bar indicates bases in the largest contiguous match, a '+' indicates other base pairs. The display initially is set to displaying the secondary structures of the highlighted oligo. To see secondary structures involving both primers click on the Other button. When viewing intermolecular (two oligo) structures the Other button is changed to the Single button which returns the analysis to intramolecular. The oligos displayed in the window are labeled Sense and Antisense.

<< and >> buttons change which structure is displayed in the window. The number of the desired structure may be typed directly in the first box.

Other or Single button toggles between intermolecular and intramolecular secondary structure analysis. If Other is displayed the analysis is set to intramolercular (single oligo), clicking the button will change the analysis to intermolecular (two oligo) and display the Single button. Clicking the Single button returns the display to intramolecular. The button is blank if the analysis is set to hairpin loops as they are by definition intramolecular.

Tm is the thermodynamic melting temperature for the structure displayed in the window. For dimers it is calculated using the nearest-neighbor values in the appendix A. For hairpins it is the temperature where the G value is 0, the temperature above which the hairpin is no longer stable. These values only indicate the relative stability of the secondary structures and are not exact Tms. They should only be used to compare the relative stability of the structures.

dG is the free energy of the dimer or hairpin . dG for a dimer is calculated by the formula G = H-TS Where H is the enthalpy, S is the entropy and T is the temperature set by the dG Temp. The G for a hairpin is determined by the free energy values of Freier at al. The values are listed in the appendix A.

Sort permits sorting of secondary structures. A standard sort box is displayed.

Print button displays the print dialog box. Selections in the dialog determine the content of the printout.

OK recalculates the analysis.

Cancel cancels the analysis and closes the window.

Help displays the help information directly related to the current window.

Defaults displays a dialog box enabling the user to save the currently selected methods and parameters as the new defaults or returns them to the original Gene Runner defaults. See the help section on default analysis and sorting parameters.