## Overview of Features

Click the contents button above to see a list of available help topics.

<u>Multi Line Display</u>	Display of up to seven formulas with their molecular weights simultaneously.
Percent Composition	Percent composition of up to seven formulas.
Parentheses Are Allowed	Handles up to 4 layers of embedded parentheses. For example, (CH3)3CH2CH3 is equivalent to CH3CH3CH3CH2CH3.
<u>Hydrates</u> or other appended compo	ounds are allowed. For example, FeCl3-6H2O.
User-definable <u>abbreviations</u>	Default abbreviations are included for common parts of compounds, including amino acids. See <u>full list</u> . For example, PhCI = C6H5Cl and HOAc = CH3COOH.
Smart Case Conversion	The program will automatically convert lowercase letters to uppercase where appropriate for ease of entering a formula. Exact case matching and non-

conversion are also available.

Edit and save abbreviations while program is running.

<u>Isotopes</u> are recognized using the following notation: ^13C is Carbon-13 C6H5^18OH is heavy-oxygen (Oxygen-18) labeled phenol

Feature of <u>weighting</u> parts of a compound relative to the other parts. For example,

[.2Na]Cl would have a weight of 0.2\*22.989768+35.4527=40.0507 NaCl-[.5H2O] would have a weight of 22.989768+35.4527+0.5\*(2\*1.00794+15.9994)=67.4501

<u>Percent Solver</u> mode for finding the value of "x" in a compound that satisfies userspecified percent composition requirements.

Edit and save elemental values while program is running.

<u>Accuracy</u> of the final digit of the molecular wt. and percent composition.

Capability to set optional features at the <u>command line</u> when starting the program.

Capability of <u>saving options as defaults</u> and automatic loading of the saved options upon program start.

Easily <u>Cut, Copy, and Paste</u> information between the Molecular Weight Calculator and other Windows applications.

<u>Mole/Mass Converter</u> for easily translating moles to mass (kg, g, mg, pounds, ounces)

and back. You can also use the mole/mass converter for molarity related calculations.

<u>Formula Finder</u> for finding possible compound empirical formulas for a given molecular weight or for a given set of percent composition data.

Capability of <u>printing</u> results.

Extensive On-Line Help (including context sensitive tool tips) and Error Checking

## Entering Formulas

Once the program starts, you may enter a molecule's formula on the numbered line. Use the LEFT and RIGHT arrow keys, HOME key, and END key to navigate a line. Calculate the weight of a formula by pressing the ENTER key or clicking the Calculate button. In <u>Multi View</u> mode, this will create a new blank line for a new formula, with a maximum of seven formulas (unless the <u>Advance on Calculate</u> option is set to No). This is useful for comparing molecular weights of compounds.

Move among the lines using the UP and DOWN arrow keys or the TAB and SHIFT-TAB keys. Pressing ENTER on line 'n' when line 'n+1' already has a formula will simply move the cursor to line n+1. A formula can consist of elements, element subscripts, <u>parentheses</u>, <u>dashes (for hydrates)</u> (-) followed by numbers to separate parts of a compound (!!NOT!! subtract them), and <u>brackets</u> for weighting parts of a compound.

See the <u>examples</u> for a wide variety of example formulas.

NOTE: For Radioactive elements, the most stable isotope is NOT used; Instead, an average Molecular Weight is used, just like with other elements. See the <u>isotopes</u> <u>section</u> for the method of indicating an exact isotopic mass. Data obtained from the Perma-Chart Science Series Chemical Periodic Table, 1993; ISBN 1-55080-063-9.



Parentheses are used to repeat parts of a compound. Up to four levels of parentheses may be embedded. For example, C6(COOH)6 has 12 carbon atoms, 12 oxygen atoms, and 6 hydrogen atoms while CH(C(CH3)3)3 is a carbon atom with three C(CH3)3 groups attached and a hydrogen atom attached.



A formula may contain brackets to "weight" specific parts of a compound relative to the other parts. For example, [.2Na]Cl would have a weight of: 0.2\*22.989768+35.4527=40.0507and NaCl-[.5H2O] would have a weight of: 22.989768+35.4527+0.5\*(2\*1.00794+15.9994)=67.4501

A colleague of mine requested this feature for the purpose of trying to figure out if a compound he had synthesized and sent off for elemental analysis had a little bit of water in it (hadn't been completely dried), as in C6H6-[.1H2O], to see if the theoretical percent composition data would better match the actual lab results. As you can see, decimal numbers can be used after brackets whereas they cannot be used after closing <u>parentheses</u> or after a <u>dash</u>.

NOTE: The percent solver feature will perform this operation for you.



Dashes are used to add hydrates to a compound. For example, FeCl3-6H2O means ferric chloride with 6 water molecules attached. As shown, an integer can follow the dash, though it's not required (in which case 1 is assumed to follow the dash). In fact, the dash can be used many times to add multiple hydrates onto a compound (if you have the desire). Also, other molecules besides hydrates can be added. For example (fictional compound) C6H6-H2O-2ZnHgMg-U would be benzene + water + 2\*(zinc+mercury+magnesium) + uranium for a total molecular weight of 1152.76. Use <u>brackets</u> in conjunction with the dash to add fractional parts of hydrates or other compounds.



Isotopes may be entered in the form  $^{XY}$  where x is the isotopic mass of the isotope and Y is the name of the element. For example,  $^{13C}$  is carbon-13 and  $^{59Fe}$  is iron-59. The program will check to make sure that the isotopic mass entered is a reasonable value for the particular element that it is associated with (see the <u>isotope math section</u>). The caret symbol ( $^{\circ}$ ) may be entered before any element in a formula, not just at the beginning. For example, C6H5 $^{18OH}$  is oxygen-18 labeled phenol. It has molecular weight 96.114 instead of 94.113

The isotopic mass entered is also used when figuring the percent composition of a compound. Thus, C6H6 is 92.258% carbon and 7.7421% hydrogen while ^13C6H6 is 92.8045% carbon and 7.1955% hydrogen. Isotopic masses may not be entered for abbreviations. For example, ^166Bpy is invalid. I considered allowing this, so that 166 would be used for the weight of Bpy (rather than 156.187), but I decided that the percent composition calculations would then be invalid since the program would not know which elements in the abbreviation were contributing the extra (or less) weight of the abbreviation. Therefore, for carbon-13 labeled Bpy, for example, simply type ^13C10H8N2 instead of ^166Bpy.

NOTE: It is suggested that D be used for deuterium rather than ^2H since most deuterium used in the laboratory is of weight 2.0159 rather than exactly 2, which is the weight ^2H would calculate as.

NOTE: For Radioactive elements, the most stable isotope is NOT used; Instead, an average Molecular Weight is used, just like with other elements. Data obtained from the Perma-Chart Science Series Chemical Periodic Table, 1993; ISBN 1-55080-063-9.



# Examples of Formulas

The following are examples of formulas that can be entered on the formula line:

Formula	Description
C6H6	Benzene. A basic formula.
nacl	Sodium Chloride. Program will capitalize nacl to NaCl for
	you.
FeCl3-6H2O	Iron Chloride with 6 hydrates attached (shows use of
	dashes for <u>hydrates</u> )
Al(OH)3	Aluminum and 3 hydroxide groups (shows use of
	<u>parentheses</u> )
HC(C(CH3)3)3	Carbon with three C(CH3)3 groups attached (shows use of
	embedded <u>parentheses</u> )
Et2O	Diethyl Oxide (shows use of an <u>abbreviation</u> )
^13C6H6	Carbon-13 labeled benzene (shows use of isotopes)
[.2Na]Cl	Only one fifth of a sodium atom exists for each chlorine
	atom (shows use of <u>brackets</u> )
C6H6-[xH2O]	Benzene with an unknown amount of water, denoted by
	the x. Use with <u>percent solver</u> calculations.



The New Formula Button (Alt-N) creates a new blank formula line provided the current formula is not blank. This is especially useful in <u>Single View</u> mode since the ENTER key does not create a new line.



The file MWT\_ELEM.DAT contains a listing of all the elements and their uncertainties (errors in their weights). This file may be edited via a text editor or by selecting Edit Elements Table under the File menu. Only elemental weights and uncertainties may be changed, not element names. New elements can only be added by entering them as abbreviations. See the <u>abbreviations section</u> for a description of how to edit abbreviations. Select Reset to Defaults in the Editing Elements window to reset the elements to their default average atomic weight values.

By default, the elemental weights are based on the elements' average atomic weights in nature. By clicking on the Use Exact Weights button, you can set the elements' weights to their exact weights, which is typically only used in analysis of mass spectrometry results. This change will also affect the percent composition results to reflect the elements' exact weights appropriately. Typically, you will want to use the elements' average weights, which can be set by either clicking Use Average Weights or Reset To Defaults.



To use an abbreviation, simply type it as you would type any element. Use the F3 key (or the Edit/Abbreviation Mode menu) to cycle through the three abbreviation recognition modes: Normal recognition, Amino Acid (Extended) recognition, and No Recognition. The MWT\_ABBR.DAT file contains the <u>default abbreviations</u> included with the program. Besides using the default abbreviations, you can <u>edit the abbreviations</u> to add your own.



### **Using a Text Editor**

You can add your own abbreviations directly to the MWT\_ABBR file via a text editor, up to 100 normal and 50 extended abbreviations. The following conventions must be followed:

- 1) Comments may be added by preceding the comment with an apostrophe.
- 2) Two group headings must exist: [AMINO ACIDS] and [ABBREVIATIONS]
- 3) Abbreviations consist of the abbreviation and molecular formula under the appropriate column in either section.
- 4) Only the first letter of an abbreviation is allowed to be capitalized.
- 5) Elements MUST be properly capitalized; eg. na is incorrect, use Na
- 6) Parentheses and dashes are NOT allowed; eg. CH3(CH2)2 is incorrect, use C3H7 or CH3CH2CH2
- 7) The abbreviations in the Amino Acid section are treated as extended level abbreviations -- they are only recognized when extended abbreviations are on.
- 8) The maximum length of any abbreviation is 6 characters.
- 9) Abbreviation names may contain only letters, while abbreviation formulas may contain letters and numbers only.

#### **Using the Built-in Editor**

You may also add or delete abbreviations by selecting Edit Abbreviations under the File Menu. The above conventions apply to any new or modified abbreviations. While editing the abbreviations, use the arrow keys to move around, then press Enter or click on the value to change it. Erasing the value or selecting cancel in the Change Value window WILL ERASE the abbreviation or value. Any lines containing one or more blank columns will be erased from the MWT\_ABBR.DAT file. Pressing Cancel in the Editing Abbreviations window will cancel the most recent changes by simply reloading the MWT\_ABBR.DAT file. Select Reset to Defaults in the Editing Abbreviations screen to reset the abbreviations to their default names and formulas.

If the MWT\_ABBR.DAT file becomes corrupted, the MWT program will inform the user and ignore incorrect lines. If the file becomes deleted, the MWT program will create a new file with the default abbreviations. Thus, you can erase the .DAT files or select Restore to Defaults to restore the default abbreviations.

One quirky aspect of abbreviation recognition is that the first abbreviation of a user entered formula that is matched to the available abbreviations (during program operation) is the abbreviation that will be used -- furthermore, extended abbreviations are matched first if they are turned on.

For example, say the user types in the formula phen2-meph. If only normal abbreviations are turned on, the program will match phen to Phen and not to Ph and En since Phen appears first in the MWT\_ABBR.DAT file. To force the Ph and En abbreviations, type PhEn2-MePh. However, if extended abbreviations have been activated (by pressing F3), the formula phen2-meph will be translated to PheN2-MePh and nothing else, since the program will match Phe to the amino acid phenylalanine, leaving the n to be interpreted as nitrogen. Numerous methods exist to avoid this specific problem:

- 1) Properly capitalize formulas while typing them in
- 2) Don't use extended abbreviations
- 3) Erase the Phe amino acid

4) Place the Phen abbreviation in the amino acid section ABOVE the Phe abbreviation, so that the program tries to match Phen before Phe.

This example should serve as a model for how to handle possible future abbreviation conflicts. This example should also illustrate that if the Phen abbreviation is placed anywhere below the Ph or En abbreviations in the MWT\_ABBR.DAT file, it will be IMPOSSIBLE for the program to recognize it, since it will match Ph or En or both before even encountering Phen.



The following is a list of the default abbreviations included with this program. Be aware that BPY is Boron-Phosphorus-Yttrium while Bpy is bipyridine:

**Default Abbreviations for Normal Recognition** (guasi-alphabetically, same as the order in the MWT ABBR.DAT file)

(	it is strongly suggested t	hat longer abbrevia	tions be placed first)
Abbreviation	Group Name	Formula	Molecular Weight
Phen	1,10-Phenanthroline	C12H8N2	180.209
Urea	Urea	H2NCONH2	60.05564
Вру	2,2'-Bipyridine (See Note 1)	C10H8N2	156.187
Oac	Acetate	CH3COO	59.04462
Otf	Triflate	OSOOCF3	149.0704
Трр	Tetraphenylporphyrin	(See Note 2)	612.7333
Bu	Butyl	CH3CH2CH2	57.11546
En	Ethylenediamine	H2NCH2CH2NH2	60.099
Et	Ethyl	CH3CH2	29.0617
Ме	Methyl	CH3	15.03482
Ms	Mesyl	CH3SOO	79.09962
Ox	Oxalate	C2O4	88.0196
Ph	Phenyl	C6H5	77.1057
Ру	Pyridine	C5H5N	79.10143
Ts	Tosyl	CH3C6H4SOO	155.1974
D	Deuterium	Heavy hydrogen	2.01588
(See Note 3)			

Note 1: 2,2'-Bipyridine is also known as 2,2'-Dipyridyl

- Note 2: Tpp is (C4H2N(C6H5C)C4H2N(C6H5C))2 and is assumed to be bound to a metal. Thus, solid, pure Tpp would be notated H2Tpp to add back the two hydrogens that leave upon complex formation. Tpp is also known as 5,10,15,20-Tetraphenyl-21H,23H-porphine.
- Note 3: It is suggested that D be used for deuterium rather than ^2H since most deuterium used in the laboratory is of weight 2.0159 rather than exactly 2, which is the weight ^2H would calculate as.

### **Default Abbreviations for Amino Acids**

(alphabetically, same as the order in the MWT\_ABBR.DAT file)

Abbreviation	Amino Acid	Formula	Molecular Weight
Ala	Alanine	C3H5NO	71.07884
Arg	Arginine (unprotonated NH2)	C6H12N4O	156.1876
Asn	Asparagine	C4H6N2O2	114.1039
Asp	Aspartic acid (undissociated COOH)	C4H5NO3	115.0886
Cys	Cysteine	C3H5NOS	103.1448

	(no disulfide link)		
Gln	Glutamine	C5H8N2O2	128.1308
Glu	Glutamic acid	C5H7NO3	129.1155
	(undissociated COOH)		
Gly	Glycine	C2H3NO	57.05196
His	Histidine	C6H7N3O	137.1412
	(unprotonated NH)		
Нур	Hydroxyproline	C5H7NO2	113.1161
lle or Leu	Isoleucine or Leucine	C6H11NO	113.1595
Lys	Lysine	C6H12N2O	128.1742
	(unprotonated NH2)		
Met	Methionine	C5H9NOS	131.1986
Phe	Phenylalanine	C9H9NO	147.1766
Pro	Proline	C5H7NO	97.11671
Ser	Serine	C3H5NO2	87.07824
Thr	Threonine	C4H7NO2	101.1051
Trp	Tryptophan	C11H10N2O	186.2133
Tyr	Tyrosine	C9H9NO2	163.176
Val	Valine	C5H9NO	99.13259

Note: The amino acids are listed as (NH)CH(CO)R (missing an H and an OH) as if they had formed peptide bonds and were in a protein. Be sure to add an H to the first amino acid and an OH to the last amino acid when you type in a protein (for the neutral form). For example: H-Gly-Ala-Phe-OH is a neutral protein with three amino acids. The zwitterionic (dipolar) form of the protein would be represented as H2-Gly-Ala-Phe-O though the molecular weight would be the same.



Multi View Mode is useful for comparing up to seven formulas and their molecular weights simultaneously. See <u>Entering Formulas</u> for information on formula conventions. You must switch to <u>Single View</u> mode to see percent composition data.

## Single View Mode (percent composition)

You can see the percent composition of the compounds by pressing the F8 key. Although only one formula is shown at a time, the other formulas are kept in memory and can be cycled through using the UP and DOWN keys without switching back to <u>Multi</u> <u>View</u> Mode. In Single View Mode the formulas can still be completely modified. The ENTER key or Calculate Button will calculate a formula's molecular weight and percent composition. However, in Single View Mode, pressing the ENTER key or clicking the Calculate Button will keep the same formula visible; only the UP and DOWN arrow keys will move to a new formula in Single View Mode. Use the <u>new formula button</u> to create a new formula.



The F11 key, or the PAGE UP and PAGE DOWN keys set the program in percent solver mode. Use the PAGE UP, PAGE DOWN, and arrow keys to move among the percent compositions shown. Press Enter or click on a percent to type in the target values for the solver to aim for. Erase the target value or select cancel to un-target the percent.

Before the solver can work, you need to place an x in the formula, directly after a bracket. For example, "C6H6-[xH2O]" is useful for calculating the amount of water in a benzene sample, provided elemental analysis data is available. The formula must be in <u>bracketed form</u> for the solver to work. If you have the formula "NaClx", where you wish to find the number of chlorine atoms based on data, simply retype the formula in the form "Na[xCl]" for the percent solver to be able to interpret the it. When at least one target percent is entered and an x is present after a bracket in the formula, you may navigate the cursor to the formula line then press ENTER or click on the Calculate Button to begin solving.

The computer will show you the progress of its calculations. Press any key or click the mouse during the calculations to stop them before they are completed. The computer should stop calculating once it gets the same answer four times in a row, which should be to approximately four or five decimal places.

After the calculations are complete, a window will appear which shows the obtained percentages, target percentages, and differences between these numbers. This is useful for seeing how well the x value works for matching the target percentages. Use the Copy button to copy the data to the clipboard. Closing the box with the close button returns you to the main program. Press the F11 key again to exit percent solver mode.

NOTE: If the program is not in Percent Solver mode, but an x is present in the formula after a bracket, the program will calculate the molecular weight based on x equaling one.



The Mole/Mass converter is useful for performing common chemical calculations concerning the interconversion of mass, moles, and molarity. To use the converter:

- Type a valid molecular formula on one of the formula lines
- Select Mole/Mass converter from the view menu

In the converter window, the first two lines are used for simple conversion between mass and moles or one unit of mass and other units of mass.

- Select Convert Amounts from the dropdown combo box
- Type the starting number
- Select the correct units for the starting and computed numbers
- The program will perform the conversion for you.

When volume units are selected, a density box will appear so that the program can correctly convert mL of a liquid to moles of the liquid (for example). Note that pounds and ounces conversions are based on the avoirdupois system, commonly used in the U.S. (1 pound = 16 ounces = 0.454 kg).

When Find Molarity is chosen in the dropdown combo box, two other boxes appear which can be used to perform three common molarity-related operations:

- By typing a weight or volume in the top box and entering a volume in the second box, you can select the Find Molarity button to calculate the resultant molarity of a solution of the corresponding volume containing the specific amount of the compound.
- Conversely, after entering a molarity and a volume, you can select the Find Amount button to find the amount of substance required to make a solution of the corresponding volume and molarity.
- Finally, after entering an amount and a molarity, you can find the volume of solvent required to satisfy the amount and molarity requirements by selecting the Find Volume button.

Note: I have chosen to only allow the g/mL unit for density; if anyone can suggest other common units I will implement them in the next version of the program.



## **Formula Finder**

The Formula Finder can be used to find empirical formulas of compounds that have a specific molecular weight or specific percent composition data. This is useful for finding possible compounds based on data, simply the reverse of what the main program does.

For finding compounds based on molecular weight:

- Select the Match Molecular Weight radio button
- Check desired elements
- Type a weight in the Molecular Weight of Target Compound box.
- Set the weight tolerance to an acceptable value.
- Select calculate.

The tolerance defines what range of compounds to match. For example, typing 78 in the weight box and .1 in the tolerance box will find compounds whose weights lie between 77.9 and 78.1. You specify which elements a compound contains by clicking the boxes next to the element names.

The Min and Max boxes appear once an element is selected. These boxes specify the minimum and maximum number of times the element may appear in the empirical formula of the compounds. The Custom1\_, Custom2\_, etc. boxes are used to enter custom elements or abbreviations besides the ones listed. Clicking one of the Custom boxes brings up the typical Min and Max boxes, plus a box where you may type a weight for the custom element, an elemental symbol for one of the valid elements (eg. Fe, Pb, Na), or a valid abbreviation (eg. Me, Ph, Bpy).

Selecting calculate will instruct the program to find all of the compounds that match the specifications. If you used a custom element and typed a weight for it, it will be notated in the results list by C1\_, C2\_, C3\_, and C4\_. This is so that C2\_5 can represent five custom 2 atoms, while C2 represents two carbon atoms. If an elemental symbol such as Fe or Na was entered for a custom element, the program will use that symbol in the results list.

You can limit the maximum number of compounds found by changing the number in the Max Hits box. Double clicking or pressing Enter on any of the lines in the results box will show the entire line. This is useful if the formula is too long to display in the box. You can stop the Formula Finder during calculations by pressing a key or clicking the mouse.

For finding compounds based on percent compositions, select the Match Percent Compositions radio button. Now, selecting specific elements will show a Percent box, along with the Min and Max boxes. In this box you should type the percent of the compound that is the element. Type the percentages as numbers between 0 and 100 and do not type a percent sign (the program won't let you). If you use an abbreviation in a custom element box while Matching Percent Compositions, realize that the percentage you type in for the abbreviation will be for the percent of the compound that is that abbreviation.

Similar to the molecular weight tolerance, the percent tolerance specifies the percent range that the element may have, so that setting carbon's percent to 56 with a tolerance of 2 will match compounds that are between 54% and 58%. Again, to see the entire line of results that don't fit in the list box, just double click on the line or press Enter.

You can minimize the Formula Finder window (by clicking on the underscore just left of the X button in the upper right of the window) while it performs its calculations and it will continue to calculate in the background. This is useful in case you have a very large calculation to perform and you want to use other windows applications or other parts of the Molecular Weight Calculator program while the calculations are performed.



The Molecular Weight Calculator for Windows 95 contains five menus:

<u>File Menu</u> <u>Edit Menu</u> <u>View Menu</u> <u>Options Menu</u> <u>Help Menu</u>



Edit Elements Table Edit Abbreviations Print Results Exit



The results can be printed by selecting Print Results from the File Menu or clicking the Print button in the Formula Finder. In <u>Multi View</u> Mode, the formulas and their molecular weights will be printed. In <u>Single View</u> Mode, the current formula and its percent composition results will be printed. In the <u>Formula Finder</u>, the formulas in the results box will be printed.



In the main part of the program, press the ESC key or Alt-F4 and then press the y or Y key to exit the program. If using the <u>Mole/Mass Converter</u>, the <u>Formula Finder</u>, or the <u>Formula Font</u> selector, exit them first before attempting to exit the program.



Cut, copy, paste, and delete Copy current Molecular Weight Copy current Percent Composition data Duplicate Current Formula Erase All Formulas Erase Current Formula Expand Abbreviations

# Cut, copy, paste, and delete

The Cut, Copy, Paste, and Delete options may be used to copy formulas to other formulas in the program or to other programs. These options will only work if the cursor is on a text input box. Furthermore, text must be selected for cut, copy, or delete to work.

# Copy current weight or percent composition

Copy Current Molecular Weight copies the molecular weight to the clipboard so that it may be pasted into other programs. Similarly, Copy Percent Composition data will copy the data to the clipboard (only available during <u>Single View</u> Mode).



Press the F2 key to copy a formula to a blank line (valid in both <u>Multi</u> and <u>Single View</u> modes). The program will search for a blank line on all seven lines, and put the copy on that line. If no line is blank, it will copy the formula to a new line, unless seven formulas already exist. Also note that the program will not copy a blank line to a new blank line.



Press the F5 key to erase all the formulas. The program will prompt for confirmation. NOTE: All of the Y,N prompts are case insensitive; n is the same as N.



Press the F6 key to erase the current formula. The program will prompt you for confirmation. In <u>Single View</u> mode this will leave you with a blank line. Again, the other formulas (if present) can be viewed using the UP and DOWN arrow keys.

NOTE: All of the Y,N prompts are case insensitive; n is the same as N.



Choose expand abbreviations from the edit menu to expand the abbreviations in the current formula to their elemental equivalents. For example, expanding PhOH gives C6H5OH. Parentheses will be inserted for clarity if necessary, for example FeBpy3 expands to Fe(C10H8N2)3 since FeC10H8N23 would be incorrect.



<u>Multi View</u> <u>Single View (percent composition)</u> <u>Percent Solver</u> <u>Mole/Mass Converter</u> <u>Formula Finder</u>



Abbreviation Mode Case Recognition Modes Caution Mode Advance on Calculate Standard Deviation Mode Exit Program Options Formula Font Saving Options as Defaults Restoring Default Options



Use the F4 key to toggle the case recognition mode for the analysis of a formula. In Exact Case Mode elements must be properly capitalized for the program to recognize them. Only the first letter of abbreviations are capitalized. In Smart Case Mode elements can be typed in lower case or upper case and the program will make an educated guess as to what the elements are. For example, nacl would be interpreted as sodium chloride. However, the line will be left displayed the way it was typed in. In Convert Case Up Mode the program will convert a line to proper notation after it has analyzed the line. In the above example, the program would redisplay the formula as NaCl. Another example, c6h6cl, would become C6H5Cl for chlorobenzene.



When caution mode is set to on, the program will warn the user when a formula is possibly mistyped, typically by warning when a character combination could be interpreted as either an inorganic element or two organic elements. For example, Co means cobalt, while CO means carbon monoxide.



The Advance on Calculate feature sets whether or not a new formula is created in <u>Multi View</u> mode when the Calculate button or the Enter key is pressed.



## **Standard Deviation Modes**

The F12 key cycles through the four standard deviation modes:

For H2O with molecular weight 18.01528 and std. dev. of 0.0003159114, Short mode displays the result as 18.0153(3) Scientific mode displays it as 18.0153 (3.159114E-04) Decimal mode displays it as 18.01528 (.0003) (no rounding) Off mode displays it as 18.01528, without standard deviation or rounding.

The standard deviation (accuracy) of calculated molecular weights and percent compositions is calculated according to <u>mathematical formulas</u>.



The Exit Program Options submenu allows you to choose whether or not the escape key (Esc) causes the program to end and whether or not exiting the program brings up the exit confirmation dialog box.



The Change Formula Font option allows you to set the font used to display the formulas and calculated results. The font is saved only when <u>Save Options as Defaults</u> is selected.



The options menu includes an option to save the current options as defaults, allowing the current options to be automatically set when the program is run again. The Restore Default Options feature will set the all the options to their respective defaults, but will not replace the options in the MWTWIN.INI file. Select Save Options as Defaults to update the .INI file.



With the help menu you can see an <u>overview</u> of the program, search the index for a specific topic, or view an about box listing information <u>about the program</u> and the <u>author</u>.

Additionally, the Show Tool Tips command can be used to turn off or on the display of tool tips on various controls throughout the program. Tool tips are pop-up tips that appear when you point at a control with the mouse pointer without moving (or clicking) for one second. I have placed these tips throughout the program at places where I think they will be useful in helping users learn about the various features of the program. For example, if you hold the mouse over the status line at the bottom of the main screen, a tool tip will appear telling you that you can "Double click the status line to expand it" and thus see the entire line in case part of it gets cut off. You can remove the check mark from beside Show Tool Tips by selecting the menu item, causing the tool tips to no longer be displayed throughout the program. By choosing <u>Save Options as Defaults</u> in the options menu, you can save this option in the MWTWIN.INI file.



## Summary of Keystrokes

F1 displays help F2 <u>copies the current formula</u> to a blank line F3 toggles the <u>abbreviation mode</u> F4 toggles the <u>case recognition</u> mode F5 <u>erases all the formulas</u> F6 <u>erases the current formula</u> F7 toggles the <u>caution mode</u> on and off F8 toggles the view mode between <u>Multi</u> and <u>Single</u> F9 toggles whether or not a <u>new formula is created after calculating</u> F10 highlights the menu bar (Windows-defined Key) F11 toggles the <u>percent solver</u> on and off F12 toggles the <u>standard deviation</u> display mode

Enter selects the calculate button or sets the selected <u>percent</u> The up and down arrow keys are used to move among the <u>formulas</u> Page Up, Page Down, and Tab may be used to navigate to the <u>percents</u> ESC and Alt-F4 <u>Exit the Program</u>



The program will display error messages in red on the status line when an incorrect entry is inputted. The error message will disappear when a key is pressed. Any actual program errors (errors with the code) should be trapped by an error trapping routine, though I have tried to beta-test the program as much as possible with every situation I can think of. Please <u>send me a note</u> if you encounter any unhandled error.

The program will also display numerous <u>caution messages</u> for possibly erroneous entries. This feature may be toggled on and off with F7.



If you downloaded mwt4\_1s.exe, the SETUP program will correctly install the files and create a program group and icon for the program. Otherwise, you may create a new shortcut to the program by referencing the MWTWIN.EXE file, moving MSFLXGRD.OCX and COMDLG32.OCX to the \Windows \System directory, and making sure MSVBVM50.DLL is present in the \Windows\System directory. If you do not have the MSVBVM50.DLL file, you can download it from various sites on the WWW, including:

http://www.simtel.net/simtel.net/win95/dll.html under vb5run.zip http://www.sharepaper.com/apps/reviews/q\_ufad.html under Visual Basic 5 Runtimes http://www.softseek.com/Utilities/VBRUN\_Files/ under VB5 Runtime Files http://www.apexsc.com/vb/davem/vbttdl.html under VB 5 Runtime Files

You can also ftp the file from various sites, including: ftp.simtel.net/pub/simtelnet/win95/dll/msvbvm50.zip OAK.Oakland.edu/pub/simtelnet/win95/dll/msvbvm50.zip ftp.uoknor.edu/mirrors/simtelnet/win95/dll/msvbvm50.zip ftp.coast.net/SimTel/win95/dll/msvbvm50.zip

<u>Command line options</u> may be added when starting the program. Typing MWTWIN /? at the command line or in the Run Programs box under the Start menu shows a list of options for setting internal program options before starting the program. All of the options may also be toggled while the program is running and saved during program operation.



Syntax: MWTWIN [/M:n] [/C:n] [/A:n] [/S:n] [/U:n] [/X] [/?]

Available options are:

/M:0 or /M:1 for <u>Multi</u> or <u>Single View</u> mode.
/C:0, /C:1, or /C:2 for Convert Up, Exact, or Smart <u>Case Recognition</u> mode.
/A:0, /A:1, or /A:2 for Normal, Extended, or No recognition of <u>abbreviations</u>.
/S:3 or /S:n for <u>Std Deviation</u> Mode: 3 is off and n can be 0, 1, or 2.
/U:0 or /U:1 for <u>Caution message</u> display on or off.
/X to prevent MWTWIN from attempting to <u>read/write abbreviations/elements</u>.
/? or /Help to display this screen.

See MWTWIN.HLP for more help. Program by <u>Matthew Monroe</u>. Send E-Mail to monroem@unc.edu WWW - http://www.unc.edu/~monroem/

This program is Freeware; distribute freely.



Calculation Note:

For Radioactive elements, the most stable isotope is NOT used; Instead, an average Molecular Weight is used, just like with other elements. See the <u>isotopes section</u> for the method of indicating an exact isotopic mass. Data obtained from the Perma-Chart Science Series Chemical Periodic Table, 1993; ISBN 1-55080-063-9.

Uncertainties come from CRC Handbook of Chemistry and Physics, except for Radioactive elements, where uncertainty was estimated to be .n5 where n represents the number of digits after the decimal point but before the last number of the molecular weight. For example, for Nobelium,  $MW = 259.1009 (\pm 0.0005)$ 

Molecular Weight, Percent Composition, and Std. Deviation Calculations

Percent Solver Calculations

**Isotopic Extremes Calculations** 

## Percent Composition and Standard Deviation calculations

### The Mathematics behind the Molecular Weight, Percent Composition and Standard Deviation Calculations

(Concept from Molar Mass program by Christopher King at Eastern Oregon State College).

NOTE: s\_f means s subscript f and is one term; likewise s\_mw means s subscript mw and is also simply one term.

For Addition:

In general: For f = x + y - z, the standard deviation of f, s\_f, is equal to Sqrt[(s\_x)<sup>2</sup> + (s\_y)<sup>2</sup> + (s\_z)<sup>2</sup>], which is in words, the std. dev. of f is equal to the square root of the sum of the squares of the std. deviations of each term in the original sum.

Applied to the Molecular Weight of a molecule:

 $MW = ax + by + cz + ..., where a is the number of atoms of weight x, b is the number of atoms of weight y, etc. The Std. Dev. of MW, s_mw = Sqrt[a(s_x)^2 + b(s_y)^2 + ...]$ 

NOTE: The reason the equation uses  $a(s_x)^2$  is that if the terms were added up consecutively 'a' times for the same element, we would have  $(s_x)^2 + (s_x)^2 + (s_x)^2$  etc. up to 'a' times, which combines to be  $a(s_x)^2$ 

For example, for chlorobenzene, C6H5CI:

 $\begin{array}{l} \mathsf{MW} = 6*12.011 + 5*1.00794 + 35.4527 = 112.5584 \\ \mathsf{s\_mw} = \mathsf{Sqrt}[6*(.001)^2 + 5*(.00007)^2 + (.0009)^2] = .002614 \\ \mathsf{Resulting} \ \mathsf{MW} \ \mathsf{is} \ 112.558 \ (\pm 0.003) = 112.558(\pm 3) \end{array}$ 

For Multiplication:

In general: For  $f = x^*y/z$ , the standard deviation of f, s\_f is equal to  $f^*Sqrt[(s_x / x)^2 + (s_y / y)^2 + (s_z / z)^2]$ , which is in words, the std. dev. of f is equal to f times the square root of the sum of squares of the quotients of the std. deviations of each term divided by each term itself.

Applied to the Percent Composition of a molecule:

%Comp = (a\*x/MW)\*100, where a is the number of atoms of weight x and MW is the total molecular weight of the molecule. The Std. Dev. of %Comp, s\_%c = %c\*Sqrt[(s\_a/a)<sup>2</sup> + (s\_x/x)<sup>2</sup> + (s\_mw/mw)<sup>2</sup>]. But, since a is a finite number, its s\_a = 0, so the formula becomes: s %c = %c\*Sqrt[(s x / x)<sup>2</sup> + (s mw / mw)<sup>2</sup>].

For example, for chlorobenzene, C6H5Cl:

%Comp of Carbon = 6\*12.011/112.5584\*100 = 64.02543%s\_%c = 64.02543\*Sqrt[(.001/12.011)<sup>2</sup> + (.002614/112.5584)<sup>2</sup>] = .00553 Resulting %Comp of Carbon = 64.025% (±0.006) =  $64.025(\pm 6)\%$ 

%Comp of Hydrogen = 5\*1.00794/112.5584\*100 = 4.47741%s\_%c = 4.47741\*Sqrt[(.00007/1.00794)<sup>2</sup> + (.002614/112.5584)<sup>2</sup>] = .0003279 Resulting %Comp of Hydrogen = 4.4774% (±0.0003) =  $4.4774(\pm 3)\%$ 

%Comp of Chlorine = 35.4527/112.5584\*100 = 31.49716%s %c = 31.49716\*Sqrt[(.0009/35.4527)<sup>2</sup> + (.002614/112.5584)<sup>2</sup>] = .0010837 Resulting %Comp of Hydrogen =  $31.497\% (\pm 0.001) = 31.497(\pm 1)\%$ 



## The Mathematics behind the Percent Solver routine

For determining the closest fit between multiple target values and actually calculated data, the value of the following formula is minimized: diff = (actual%value\_i - target %value\_i)<sup>2</sup>, which in words is the summation of the ith calculated percentage value minus the ith target value. The minimum value of this formula should be obtained when all actual values are closest to their target values.



## The Mathematics behind the Isotopic Extremes error handler

When a user types in an <u>isotopic mass</u> (eg. ^13C for carbon-13), the program checks to see if the isotope is a reasonable value. Obviously, ^5C (carbon-5) would be unreasonable because any compound with 6 protons (i.e. carbon) must have a minimum weight of 6. However, upon examining the Table of Isotopes in the CRC Handbook of Chemistry and Physics, I found that very few compounds have zero neutrons (and thus an atomic weight equal to the number of protons in the element). Therefore, I examined the table and used MS Excel to create a graph of the highest and lowest isotopic masses for every element (shown below). I then determined two fit-lines that encompassed all the data points -- one for the upper extreme and one for the lower extreme.



Since most isotopic masses occur somewhere around the value twice the atomic number (Z) of the element, I chose to work with the value:

 $x = isotope_value - Z*2$ 

where isotope\_value is the value present after the caret (^) in the formula.

The curve that describes the upper limit of these x values is:  $max_x = INTEGER(0.63*Z + 6)$ where max x is the maximum allowable x value before a warning is issued.

The curve that describes the lower limit of these x values is:

 $min_x = INTEGER(0.008*Z^2 - 0.4*Z - 6)$ 

where min\_x is the minimum allowable x value before a warning is issued. The results were converted to integers to aid in program design.

Finally, because the min\_x value is too low for some of the lighter elements, I included a test to make sure that isotope\_value was at least as large as Z, or else a warning was again issued to the user. The method is an approximate test on the isotope value, and is designed to catch the outrageous isotopic masses. It does not catch the borderline ones since that would have increased the program size and memory usage by too much.

Examples: For Carbon: ^13C has z = 6, x = 1, max\_x = 9 and min\_x = -9 Because x < 9 and x > -9 and x > 6, no warning is issued. (The average atomic weight of carbon is 12.011)

^25C has z = 6, x = 10, max\_x = 9 and min\_x = -9 Because x > 9, a warning that 25 is probably too large is given. In reality, the heaviest known carbon isotope is carbon-16.

^5C has z = 6, x = -7, max\_x = 9 and min\_x = -9Although x < 9 and x > -9, x is less than 6 so a warning is again issued. In reality, the lightest known carbon isotope is carbon-9.

For Silver:

^100Ag has z = 47, x = 6, max\_x = 35 and min\_x = -8 Because x < 35 and x > -8 and x > 47, no warning is issued. (The average atomic weight of silver is 107.87)

^130Ag has z = 47, x = 36, max\_x = 35 and min\_x = -8Because x > 35, a warning that 130 is probably too large is given. In reality, the heaviest known silver isotope is silver-122.

^85Ag has z = 47, x = -9, max\_x = 35 and min\_x = -8 Because x < -8 a warning that 85 is probably too small is given. In reality, the lightest known silver isotope is carbon-96.

The above calculations show, as mentioned above, that this method catches the outrageous isotopic masses but does not catch the borderline cases. Furthermore, it should also be evident that this method is more accurate for the heavier elements, a consequence of the unusual isotopic behavior of the lighter elements.



## Program Availability

### Windows 95 Version (MWTWin v4.1)

The program is available via the web at http://www.unc.edu/~monroem/ or at http://www.simtel.net/pub/simtelnet/win95/chem/ in the mwt4 1.zip file.

The program is also available via Anonymous FTP at ftp.simtel.net under pub/simtelnet/win95/chem/ in the mwt4 1.zip file, and ftp.osc.edu under the pub/chemistry/software/MS-WINDOWS95/Molecular Weight/ directory in the mwt4 1.exe file.

To save space, the mwt4 1.zip and mwt4 1.exe files do not contain the MSVBVM50.DLL file and do not contain an install program; instead, directions for installation are included in the Readme.Txt file and in this help file; MSVBVM50.DLL must be obtained elsewhere. The mwt4 1s.exe file contains MSVBVM50.DLL and a full installation program.

#### Windows 3.x Version (MWTWin v3.2)

The program is available via the web at http://www.unc.edu/~monroem/ or at http://www.coast.net/SimTel/win3/chem.html and http://www.cdrom.com/simtel.net/win3/chem.html and http://www.simtel.net/pub/simtelnet/win3/chem/ in the mwt3 2.zip file.

The program is also available via Anonymous FTP at ftp.coast.net under the SimTel/win3/chem/ directory in the mwt3 2.zip file, at ftp.simtel.net under pub/simtelnet/win3/chem/ in the mwt3 2.zip file, and ftp.osc.edu under the pub/chemistry/software/MS-WINDOWS/Molecular Weight/ directory in the mwt3 2.exe or mwt3 2s.exe files.

#### Dos Version (MWT v2.87)

The program is available via the web at http://www.unc.edu/~monroem/ or at http://www.coast.net/SimTel/msdos/chemstry.html and http://www.cdrom.com/simtel.net/msdos/chemstry.html and http://www.simtel.net/pub/simtelnet/msdos/chemstry/ in the mwt2 87.zip file.

The program is also available via Anonymous FTP at ftp.coast.net under the SimTel/msdos/chemstry/ directory in the mwt2 87.zip file, at ftp.simtel.net under pub/simtelnet/msdos/chemstry/ in the mwt2 87.zip file, and ftp.osc.edu under the pub/chemistry/software/MS-DOS/Molecular-Weight-Calculator/ directory in the mwt2 87.exe file.



Revision History for the Molecular Weight Calculator program by Matthew Monroe		
Revision 1	DOS version Completed August 6, 1995 Contains basic features including: formula weight calculation percent composition calculation can handle parentheses, up to 2 sets of embedding can handle hydrates and other additional compounds can handle numbers up to 2 digits long two types of case conversion help screen	
Revision 2.0	Completed August 8, 1995 Contains the following new features: can handle numbers up to 6 digits long can handle up to 4 sets of embedded parentheses ability to handle brackets and partial weights three types of case conversion expanded help screens	
Revision 2.1	Completed August 9, 1995 Contains the following new feature: display of time	
Revision 2.2	Completed August 10, 1995 Contains the following new feature: calculator	
Revision 2.21	Completed August 17, 1995 Contains minor bug fixes.	
Revision 2.3	Completed September 2, 1995 Contains the following new features: added command line switches to control internal settings added standard deviation option added recognition Bu for butyl group	
Revision 2.4	Completed December 1, 1995 Contains the following new features: Percent solver program. Combined F11 and F12 std. deviation options into just F12 Contains Screen update bug fixes.	
Revision 2.5	Completed December 7, 1995 Improved speed of the % solver algorithm. Contains the following new features: added recognition of Urea, Ox, Py, En, Tpp, Otf.	
Revision 2.55	Completed December 17, 1995 Added error trapping routines. Improved handling of exponentially large and small numbers.	

	Added warning statements for possibly ambiguous abbreviations such as Bpy and Py.
Revision 2.6	Completed January 16, 1996 Contains the following new features: Recognizes Amino Acid abbreviations. Recognizes Deuterium, D abbreviation. Percent Solver values can now be set for all elements of a formula. Displays differences between actual and theoretical values at the conclusion of the percent solver routine. Added CTRL-P option to print results. Added more error messages/explanations for incorrect user entries.
Revision 2.7	<ul> <li>Completed March 1, 1996</li> <li>Contains the following new features: <ul> <li>Able to handle isotopes; use the caret (^) symbol.</li> <li>Able to calculate powers of numbers in the calculator.</li> <li>Added more error messages/explanations for incorrect entries.</li> </ul> </li> <li>Included a moderately accurate error handler to check the validity of entered isotopic masses.</li> <li>Added numerous "caution" statements for possibly erroneous entries: eg. Co vs. CO and Sc vs. SC</li> <li>This feature may be toggled on and off with CTRL-C</li> <li>Improved speed of the % solver algorithm; watch it blaze!</li> <li>Removed CTRL-P option use PRINT-SCREEN or SHIFT-PRINT-SCREEN</li> <li>Changed the order and layout of the on-line help screens.</li> <li>Fixed typos in help screens and added two more help screens.</li> </ul>
Revision 2.8	<ul> <li>Completed May 1, 1996</li> <li>Contains the following new features:</li> <li>Able to handle user-defined abbreviations listed in the MWT_ABBR.DAT file.</li> <li>Error checking to prevent duplicate abbreviations or abbreviations duplicating elements.</li> <li>Automatic re-creation of the abbreviation file if it is deleted/missing.</li> <li>Ability to run without abbreviations if program is run on a write-protected drive.</li> <li>Command-line switch (/X) to stop any attempts of reading or creating an abbreviation file (useful on a write-protected drive, like a network drive).</li> </ul>
Revision 2.85	Completed September 1, 1996 Contains the following new features: Now recognizes the MWTPATH environment variable which specifies the location of the MWT_ABBR.DAT file
Revision 2.87	Completed January 1, 1997 Updated the help file to reflect the latest Windows version of the program.
Revision 3.0	Windows 3.x version Completed June 10, 1996 New Windows interface. Contains the following new features:

	Able to modify elemental values stored in the MWT_ELEM.DAT
	Editing and saving of Abbreviations and Elemental values during program operation. Ability to save current options as defaults. Ability to Copy and Paste information to/from other Windows programs. Unlimited formula length formulas can even contain all 103 elements. Ability to print results.
Revision 3.1	Completed June 28, 1996 Contains the following new features: Ability to change the font used to display formulas and calculated results. Mole/Mass Converter Formula Finder Full hypertext help
Revision 3.2	<ul> <li>Completed January 1, 1997</li> <li>Contains the following new features:</li> <li>Allows use of abbreviations as custom elements in formula finder.</li> <li>For Percent Solver, added display of differences between target percentages and obtained percentages.</li> <li>Changed comment delimiter to be a semicolon in the MWTWIN.INI file and the .DAT files.</li> <li>Added detection of a mouse click to stop the percent solver and formula finder.</li> <li>Updated help file with MWT Dos version information.</li> <li>Corrected bug in display of standard deviation values to display the correct number of decimal places.</li> <li>Improved editing of abbreviations and elemental values.</li> <li>Improved formula finder interface.</li> </ul>
Revision 4.0	<ul> <li>Windows 95 version Completed August 25, 1997</li> <li>32 bit Windows 95 code</li> <li>Contains the following new features: <ul> <li>Allows user to turn off confirmation of exiting the program.</li> <li>Added possibility of using exact elemental weights, in addition to average elemental weights. Useful for mass spectrometry.</li> <li>Added context sensitive tool tips to many of the program's controls and added a setting in the help menu to turn them on or off (the settings can be saved in the MWTWIN.INI file)</li> <li>Redesigned program help to Windows 95 standards.</li> <li>Improved dialog box displayed when changing items in grids to include OK, Cancel, and Reset buttons.</li> </ul> </li> <li>Fixed bug that had smart case mode and exact case mode switched around in the Calculate function.</li> <li>Changed all forms and controls to the standard VB 5.0 3-D style.</li> <li>Changed program icon and made the edges transparent.</li> <li>Updated help file with MWT Dos version information.</li> <li>Updated my E-mail and WWW addresses to UNC.</li> </ul>

Revision 4.1

Completed September 2, 1997

Contains the following new features:

Added feature to expand abbreviations to their elemental equivalents.

Fixed bug that wrongly calculated weights and uncertainties for weighting of compounds when abbreviations were involved (i.e. use of brackets, eg. [.1Bpy]).

Fixed bug that did not take abbreviations or isotopes into account when hydrates (dashes) were used (eg. Fe-3Bpy or Fe-6^13CH3)

Enabled error handlers to catch overflow values in various procedures.



Contacting the Author

You can contact me by E-mail at Monroem@UNC.Edu.

### About the Author

I am a graduate student in chemistry pursuing my Ph.D. at the University of North Carolina in Chapel Hill, NC. I received my B.S. in chemistry at the University of Wyoming in Laramie, WY in May, 1997. I taught myself to program in BASIC on an Apple //c (with 128 Kb of Ram and no hard disk) in 1986 during 6th grade. Since then, I have updated to GW-Basic, then QuickBasic v4.5, QuickBasic v7.1 for DOS, Visual Basic 3.0 for Windows, and now Visual Basic 5.0 for Windows 95. I also program in C and C++, but I stick with the various forms of Basic since I am much more comfortable with the language and it is very simple to create powerful Windows applications with Visual Basic.