Chemistry 365 Biochemistry
Laboratory Unit #1
Tables and Figures with a Spreadsheet

References:

Robyt, J.F.; White, B.J. *Biochemical Techniques Theory and Practice*; Waveland: Prospect Heights, IL, 1990.

The objective of this project is to teach you the general operation of the spreadsheet program that you will be using to construct tables and figures for your laboratory project reports. Scientific data are quickly and easily processed and can be presented professionally with the use of a spreadsheet program. Repetitive calculations, correction of mistakes, and graphing are all conveniently performed. The spreadsheet program we will use is **Microsoft Excel,** available on the campus computer network.

Tables and Figures (graphs):

Experimental data are best presented as tables and/or figures, which offer a concise and rapid method of communicating to the reader the data collected and the experimental results obtained. **Both tables and figures should have explanatory titles and should contain enough information to be intelligible without reference to the text.**

Tables require careful labeling of all data presented. Every row and column should have an appropriate heading or label, and any codes or abbreviations should be explained in a footnote to the table. The units of all measurements should be specified. If the condition is the same for all tabulated data, this is indicated at the top of the column. It is not wise to put too much data into a single table. Use of several tables will emphasize or highlight the important aspects of the data. Do not include more significant figures than are justified by the accuracy of the determinations.

Figures (graphs) are used to illustrate quantitative data and can quickly indicate the relationship between two or more variables and any special trends that might be present. There are several important rules for making good figures.

Rule 1: Always label both the y-axis (ordinate) and the x-axis (abscissa). Place the dependent variable(s) on the ordinate and the independent variable on the abscissa. If possible, enclose the top and right-hand sides of the figure.

Rule 2: Choose sets of coordinates that are easy to read and have major divisions expressed as integers, multiplied by the appropriate power of 10, if necessary (or a mixture of integers and decimals, if more convenient).

Rule 3: Select scales that have uniform intervals and distribute the data over most of the area of the graph rather than having the data compressed into one region of the graph. Use of computer software designed for making tables and graphs, such as Excel, simplifies this task.

Rule 4: Clearly indicate experimental points. Circles can be used to indicate the uncertainty (error) of the measurement in both variables. Vertical lines (error bars) can be used to represent uncertainty in the dependent variable.

Rule 5. Keep the number of lines on a single graph to a minimum. Assign distinct symbols (e.g., circles, open and solid, triangles, open and solid) to each set of data. Symbols should be explained in the legend to the graph. Curves or lines used to show a theoretical "best fit" to the data should be shown as a solid, continuous line.

Rule 6: Prepare a title that describes the graph and gives enough experimental detail to permit the material to stand on its own. **Readers should be able to understand what the data in the graph mean without reference to the text.**

Spreadsheets:

A spreadsheet is simply a matrix of **cells** that are designated by a **column letter** and a **row number**. Cell D5 is in the fifth row down and the fourth column from the left. When working on a spreadsheet, you must always **select** (or **activate**) a cell, row, column or block, and then put something in the area. The things that you put into these areas are either **labels** (text; i.e. words, phrases, sentences, etc.), **values** (numbers; the data you want to calculate with and plot), **formulas** (mathematical expressions with cell addresses as variables), or special built-in **functions** (such as logarithm, natural log, exponents). Type in formulas by starting with an equal sign and then using standard calculating procedure and priority, with **+** for addition, **-** for subtraction, ***** for multiplication, **/** for division and **^** for raise to the power. Functions may be included in formulas. In functions or formulas, cell addresses may be entered in two ways. If, in a calculation, you wish to always be using the same value (i.e. a constant), then enter the cell address as an **absolute reference** with a dollar sign before the letter and the number - such as D5. If in a calculation, a value is a variable, enter the cell address as a **relative reference** without the dollar sign - such as D5. This way, when the formula or function is **copied** and **pasted** to other cells for computing with different variables, these cell addresses will be **automatically** changed to the new values.

When using the spreadsheet, select a cell, row, or column, by using the mouse to set the pointer on the cell, row number or column letter you desire, then **click** the mouse by pressing the upper left button once. To select a block, set the pointer at the cell on one extreme of the block, push the button and hold down, **drag** the pointer to the cell on the other extreme, and then release the button. To select any of the program menu selections, such as **File**, **Edit**, **Format, Tools**, etc. set the pointer over the word and click.

Any cell entry that begins with a letter will automatically be formatted as a label and the alignment will be left justified. To enter numbers as labels, format the cell as numbers,text before typing in the number. Any cell entry that begins with a number will automatically be formatted as a value and will be right justified. Formatting of fonts, numbers, colors, etc. in blocks of cells, rows or columns along with height and width may all be changed by selecting them and then using the **Format** menu. To enter formulas, begin with an equal sign and enter the formula as any mathematical expression. In most cases, cells can be called as a **relative reference**, where only the column number and row letter are given, A5, F2, etc. When the formula is then copied to other cells, these cell references automatically change to reflect the new location. This is normally used when a cell contains a <u>variable</u> in a formula. To designate an **absolute reference**, enter a dollar sign before the column letter and also row number, A5 or F2. This means that in all formulas, even after copying, only one particular cell will be used. This is normally used for <u>constants</u> appearing in a formula. To enter functions, find the function you want from the **Insert** menu and then enter the range of cells inside of the parentheses.

Numerical Curve Fitting: The Method of Least Squares

Many types of data collection result in a linear curve when the dependent variable y is plotted against the independent variable x. The equation for a straight line that results when a set of data (x_1,y_1), (x_2,y_2)…… (x_n,y_n) is obtained is: $y = mx + b$

where m is the slope of the line and b is the y-intercept that results when x = 0. When a set of data points lies almost, but not quite, on a straight line, the experimenter must try to find the line that "best" fits the data points. "Best" can be defined in terms of the so-called **least-squares criterion,** where the line must be such that the sum of the squares of the vertical departures of the points from the best line, commonly called the **residuals,** are a minimum.

Assignment:

You will hand in a table and a figure for each of the two data sets on the back of this page (i.e. Table 1 and Figure 1; Table 2 and Figure 2). In creating the Table first, prepare a title that describes the table and gives **enough experimental detail to permit the material to stand on its own**. Create and label a column for the independent variable data. Create and label a column for the dependent variable data. Perform a regression analysis on the data and use the resulting slope and y-intercept to calculate a "best fit" column on each table using the formula $y = mx + b$. Regression ("best straight line") analysis is found in the **Tools, Data Analysis...** menu. Enter your y and x series and indicate an output range below the data (a regression analysis table will appear). Your intercept and slope (called x variable 1) are shown in the coefficients column of the regression table. The values under standard error, next to the slope and intercept, indicate the uncertainties of the values obtained. You may clear all but the first three columns of the regression table to include on your data table.

Now create a Figure from the data in the Table. To construct a figure (Excel calls it a chart), select chart from the **Insert** menu. Most figures you will make are the XY(scatter) chart type. Next indicate your data range (select your block of data that includes independent variable, dependent variable and “Best fit line”) and that they are entered in columns. Now add your title (that gives **enough experimental detail to permit the material to stand on its own**), axis labels with units, major gridlines and legend position. Finally, make the chart a new sheet in your workbook. You can "clean up" the formatting of your axis labels, numbers and scale, data series, background, and titles by clicking on the desired feature and using the **Format** menu.