

USING EXCEL: CUSTOMIZING YOUR SPREADSHEETS AND MAKING PUBLICATION QUALITY GRAPHS

EDIT ON SCREEN!

These instructions will allow you to set up graphs that are publication quality when printed with a laser printer. A general knowledge of EXCEL is assumed. Obtain an EXCEL manual for details about its operation. These instructions are a minimal set to accomplish the goals of customizing your spreadsheets and obtaining excellent quality graphs.

SETTING UP SPREADSHEET:

1. Access EXCEL and put a title and date on your spreadsheet. At this time it would be a good idea to think of a name for saving your spreadsheet on a disc. Limit your name to 8 characters that best describes your data. The extension .xls will indicate the file is an EXCEL spreadsheet. When you create several graphs from data in your spreadsheet, give each graph a name that connects it to the particular data in the spreadsheet. It may be helpful to use a partial date in your title, especially if you collect a large amount of similar data on different days. Put these names on the spreadsheet for future reference.

For the example given on p 5,6:	spreadsheet:	CHY527.xls
	graph:	"GRAPH"
	(only one shown)	

2. Put a title on a column, skip down a line or several lines and enter the **x values** for your graph. Do the same for the **y values** putting them to the *right* of the x values. This placement makes it easier to create the graph from the data and to do linear regression analysis. It is REQUIRED that you use numbers between 0.1 and 10 for entry of data into your columns because of problems displaying exponential notation. The way you enter numbers dramatically affects the appearance of your graph. You must, therefore, keep track of the order of magnitude for the numbers on the spreadsheet, especially for the linear regression analysis, in order to obtain the correct values for the slope and y-intercepts from the data (see sample spreadsheet).

NOTE: Keep the titles for your data points as simple. The width of the column will not be automatically adjusted and a long title will run over into the next column. You can change the width of the columns; however, the default width is adequate for most entries. At this point you customize your entries as follows:

- a. Block the data columns and click on the appropriate icon to **justify** either right, center, or left according to your preference. You can also move items around on your spreadsheet by blocking and moving block with the cursor. You may also move the block by using "EDIT, CUT", and then move cursor to the top cell of where you want to move to, and click on "EDIT, PASTE".
- b. Think carefully about the appropriate **significant figures** for your data. Edit the number of significant figures by using available icons or block your numbers and click on "FORMAT, NUMBERS". DO NOT use scientific notation because the numbers shown this way will crowd the spreadsheet as well as the axes of the graph. You are required to use numbers between 1 and 10 for publication quality graphs.
- c. After you are satisfied with the location and format of the data in your spreadsheet, click on "FILE, SAVE AS" and type in the spreadsheet name you choose in order to **save** your spreadsheet on your disc. If you don't do this, the spreadsheet is automatically given a name (SHEET 1) that is not descriptive, and therefore not useful, and will be written over the next time you create another spreadsheet.

OBTAINING A PRELIMINARY GRAPH FROM DATA:

1. Put x values in the left-most column, followed by data that corresponds to each x -value to the right of these values.
1. Starting at the first x value, highlight (or block) your data set.
2. Click on INSERT/CHART/XY SCATTER plot. Remove the legends and gridlines. Place plot on a separate sheet. A preliminary graph will be created, but you will have to modify it to make it a publication quality graph. Name graph if necessary by double clicking at the bottom line where "chart 1 " appears.

CUSTOMIZING YOUR GRAPH TO OBTAIN PUBLICATION QUALITY FORMAT:

Examine the sample graphs from biochemistry journals. Note the simplicity with which the graphs are displayed. It is easy to achieve the simple format with only a few additional steps in EXCEL.

In general, the editing of your graph will be carried out by clicking (or double-clicking) on various parts of your graph and choosing options from pop-up menus. When you click on a part of your graph, small boxes will appear to indicate that part has been chosen for modification. You can use the FORMAT/SELECTED OBJECT option at the top of the screen to do the editing. Or you can double click on the part you want to edit after which the appropriate editing menus will appear on the screen. Work through the following steps to customize your graph:

TITLE (none): Note that no graphs in journals have titles!

FONT: For x and y axes numbers and labels: ARIAL BOLD
12 point
Block those characters you want to subscript or superscript and go to FORMAT/SELECTED OBJECT to choose "superscript" or "subscript".

SIG. FIG: Choose significant figures appropriate to your data. DO NOT USE SCIENTIFIC NOTATION.

COLOR: Remove color from graph points and lines for a better printout. Choose the large, filled symbols

CAPTION: Place caption at the bottom of the graph. Wait until later because the items below may alter where you put the caption.

1. If not inserted during creation of the graph, label the x and y axes, including the order of magnitude in parentheses with proper units. See example graph.
2. Double click on an empty space inside the graph so that all four sides are highlighted and the pop-up menu shows. Edit the lines to be the thickest bold lines on all four sides. Click on each x and y axis lines or numbers edit, choosing a thinner line than the one chosen for all four sides so that tick marks will show on these axes. Choosing the TICK MARKS to be INSIDE. If you do not choose a line for the individual x and y axes, the tick marks will not show. DO NOT put a box around the outside of the graph area since this is never done in scientific journals.
3. Edit the symbols used for the data points making them as big as possible, choosing black for both the symbol outline and fill to obtain better printing quality.

OBTAINING A THEORETICAL LINE FOR LINEAR PLOTS:

There are two methods outlined here to obtain a theoretical fit to a set of linear data points. The **FIRST** one involves three steps: (1) using **TOOLS/DATA ANALYSIS/REGRESSION** to obtain linear regression information; (2) using the slope and y-intercept obtained from this analysis to calculate a set of y values; and (3) using the calculated y values to create the theoretical line for the graph. The **SECOND** method uses **INSERT/TRENDLINE** and is faster. However, this latter method will be used only to check that the linear regression analysis has been carried out properly. This will give you confidence that your theoretical y values are correct.

FIRST METHOD:

1. Determine an empty area in your spreadsheet for printing the linear regression analysis. Write down the top cell location for this area.
2. Click on **TOOLS/DATA ANALYSIS/REGRESSION** and click on the y box. Move cursor to the spreadsheet and highlight the y values. Do the same for the x box and x values. Enter the cell for the output and make sure it is in a blank part of your spreadsheet. Click on **OK**.
3. The printout for the regression analysis has more information than you will actually use. You can delete areas of the analysis, as shown on the sample spreadsheet, to simplify your printouts. Consult the spreadsheet to find the slope of the line (X Variable 1), the y-intercept (Intercept), the r^2 value, and the Standard Error for the slope. These values are often used in reporting data in captions of graphs. See example caption on graph.
4. Identify the cells containing the slope and y-intercept values (as "C34 and C33" or whatever they are, respectively). You will use these values to calculate the theoretical y values for each of your x values using $y = mx + b$:
 - a. Move the cursor to the empty cell to the right of your first observed y value. This column will be used to calculate the theoretical y values. If the slope value from the regression analysis is in cell C34, the y-intercept in cell C33, and the observed y values start in cell B10, then your first theoretical y value will be calculated in cell C10 using the equation: $= (A10 * C34) + C33$. This is just $y = mx + b$ using your first x value in A10. Copy this calculation to obtain the remainder of the theoretical y values corresponding to your x values. Later put a title on this column (see sample spreadsheet) to identify these values as the theoretical or calculated y values.
 - b. Go to the graph and click on **INSERT/NEW DATA**. Then click back on the spreadsheet and highlight the theoretical y values you just calculated. Click on **OK** in the **NEW DATA** box and these points will appear with symbols on your graph. At this time, change the symbols to a line only without symbols. Choose the width of the line to be one size thinner than the maximum thickness. This will give the graph a more balanced look. Make the line black for better printing.

OBTAINING A THEORETICAL LINE FOR LINEAR PLOTS, continued:

SECOND METHOD:

This method is only a check that you have done the regression analysis correctly and that you have calculated the theoretical y values for your theoretical line correctly.

1. Go to plot and click once on any data point to highlight the set. Do not click on the theoretical line. Click on INSERT/TRENDLINE/LINEAR/OPTIONS/DISPLAY EQUATION ON CHART/DISPLAY R-SQUARED ON CHART/OK. The best line through your points will be calculated automatically and the equation for the line with the r^2 value will be printed on the graph. If this line exactly coincides with your theoretical line, you may not be able to see it. However, the TRENDLINE remains highlighted and you can change its color (to red) and its appearance (to - - - -) to see it and to verify your calculations are correct.
2. Write down the equation for the line (with m and b values) and the r^2 value and verify that the parameters are identical to the ones in your spreadsheet for the linear regression analysis.
3. Since publication quality graphs do not have equations written on them, click on the equation and DELETE it. You should also DELETE this TRENDLINE since it is only a check that you have carried out the regression analysis and calculation of the theoretical line correctly. Check with the instructor to use the SET INTERCEPT TO ZERO mode if your data should go through zero and does not.

PRINTING YOUR SPREADSHEET AND GRAPH: **EDIT ON SCREEN BEFORE PRINTING**

EDIT EXTENSIVELY ON SCREEN BEFORE PRINTING. DON NOT WASTE PAPER!

Examine your spreadsheet and graph under PRINT PREVIEW under the FILE option menu and change the placement of the spreadsheet and your graph on the page. Note that you may have to modify margins. A good placement of the graph is shown in the example graph. After the print preview comes up, you may also choose the SETUP mode to remove page titles and to insert page numbers, if appropriate. Remove HEADERS and FOOTERS (except page numbers) as they are not found on publication quality graphs.

NOTE: It is recommended that you frequently save your file during the time that you are making the changes in order not to lose any of these changes. Use a "FILE, SAVE AS" mode, to verify the name. Examine "WINDOW" and close, without saving, any spreadsheets that you inadvertently created that are of no use to you. BEWARE: EXCEL can generate unwanted spreadsheets (and charts) especially when the inexperienced user is at the keyboard.

SAMPLE
SPREADSHEET

EDIT ON SCREEN BEFORE PRINTING

DETERMINATION OF k_{cat} FOR CHYMOTRYPSIN

5-27-01
CHY527.XLS

Conditions: 0.067 M phosphate at pH 7.0

Sample for EXCEL handout
remove this note before
printing!

x	y	calcd $y = mx+b$
ENZYME	RATE	RATE
E-04	E-06	E-06
M	M min^{-1}	M min^{-1}
0.00	0.00	0.17
1.10	2.20	2.25
1.90	4.20	3.77
3.40	6.40	6.61
3.90	7.60	7.56
5.10	9.80	9.84

SEE "GRAPH"

SUMMARY OUTPUT

<i>Regression Statistics</i>	
Multiple R	0.99797105
R Square	0.99594623 = r^2
Adjusted R Square	0.99493278
Standard Error	0.25711922
Observations	6

NOTE: Parts of the regressions analysis have been removed for simplicity and the remainder has been rearranged for easier viewing.

	<i>Coefficients</i>	<i>Standard Error</i>
b = Y-INTERCEPT = Intercept	0.166986352	0.18739199
m = SLOPE = X Variable 1	1.895979343	0.060480505

Standard Error for X Variable 1 = 1.90 +/- 0.06

From "INSERT/TRENDLINE":	$y = 1.896x + 0.167$ $r^2 = 0.9959$
--------------------------	--

Slope must be multiplied by E-06/E-04 = E-02
to get proper magnitude:

Show proper units:
Slope of graph = k_{cat} in min^{-1} =
 $1.90 \pm 0.06 \times 10^{-2} \text{ min}^{-1}$
(use proper number of sig. figs.)

EDIT ON SCREEN BEFORE PRINTING

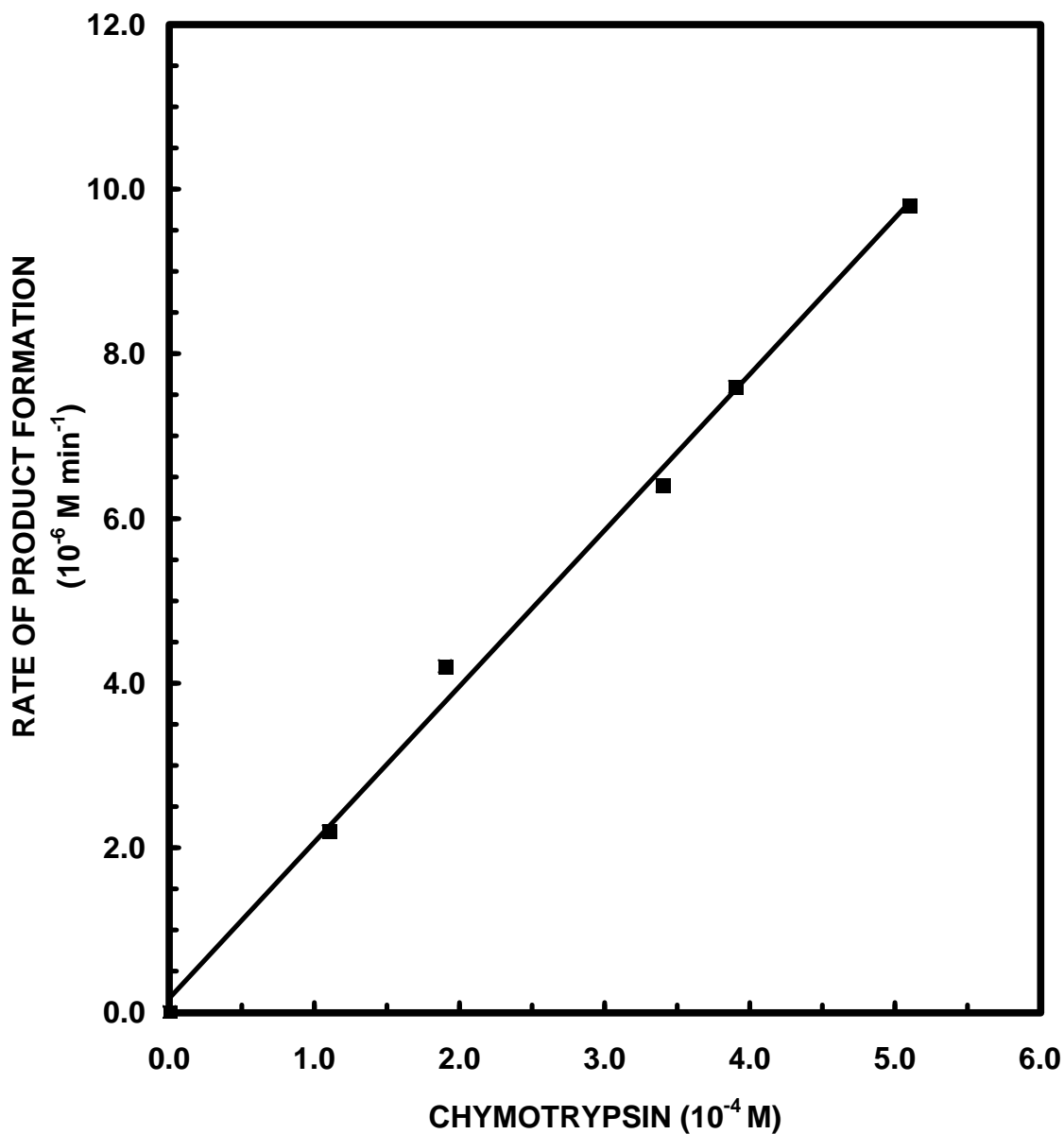


Figure 1. Dependence of the rate of hydrolysis of p-nitrophenylacetate on the concentration of chymotrypsin at pH 7.0 in 0.036 M phosphate.