DRY LAB: EXCEL DATA ANALYSES

Introduction:

Collagen I is the most abundant protein in the connective tissues of mammals, present in fibrils and found in parts of the body such as tendons, ligaments, skin, corneas, cartilage, and bones. It is produced by fibroblast cells, which are responsible for making much of the extracellular matrix that gives structure to animal tissues. In research, because of its physiological relevance, collagen has been used as scaffolds in tissue engineering and as three-dimensional environments for cell studies. In the Kaufman lab, we use collagen gels to study cancer cell migration within the context of a crowded environment (1-3). **Fig. 1** shows a multicellular tumor spheroid implanted into a collagen gel, where



the dark center is the original shape of the spheroid and the cells have clearly invaded radially outward. We have found that the extent of cellular invasion is dependent on the properties of the gels, which is why it is important for our lab to have control over making gels with different structural and mechanical properties. For example, a recent article from our lab (3) reported that breast cancer spheroids were much more invasive when implanted in 1.0 mg/ml collagen gels than in 4.0 mg/ml gels. It is suggested that the limited invasive behavior is due to the physical constraints of the system imposed by the increased number of collagen fibers and resulting dense meshwork.

Experimentally, collagen gels are typically prepared by diluting a commercially available stock solution of collagen monomers (extracted in our case from the tails of rats) in a buffer that provides all the necessary nutrients for cells to survive. The solution is neutralized to pH 7.4, which activates the gelation process, immediately transferred to a sample cell and placed at the appropriate temperature to allow *fibrillogenesis* to occur. During this process, collagen monomers self-assemble into fibrillar structures with network properties that are highly dependent on concentration and temperature. In general, increasing collagen concentrations will produce gels with shorter fibrils, whereas lower temperature gelation produces thicker and more bundled fibrils. See Figure 1 in reference (4) for confocal reflectance images of what fibrils look like when the self-assembly process takes place at different concentration and temperatures.

A better understanding of fibrillogenesis will allow researchers to control for these properties. To date, there have been many studies investigating the mechanism and kinetics of the self-assembly process, which have primarily monitored turbidity. During collagen I gelation, turbidity increases as a nearly transparent solution of monomers develops into fibrils that scatter significant amounts of light, resulting in a sigmoidal curve similar to that seen in **Fig. 3** below. Our studies and others support the hypothesis that collagen fibrillogenesis is a nucleation and growth process, as the sigmoidal curves have distinct lag and growth phases. Imaging techniques can also be employed to reveal the structure of the developing gel. In this lab, you will be using actual data from turbidity measurements and microscopy taken during gelation to reveal information about the collagen self-assembly process. The data and

suggested analyses comes from a recent publication (4) and has been simplified for the purpose of this exercise.

The majority of the publication deals with reconciling information from traditional turbidity measurements and modern imaging techniques like confocal reflectance microscopy by collapsing the reflectance images into a single variable that tracks traditional turbidity fairly well. Along the way, we developed a new way to simultaneously obtain *in situ* turbidity from scanning transmittance images. These measurements helped us conclude that fibril and network formation occurred before substantial turbidity was present, and the majority of increasing turbidity during collagen self-assembly was due to increasing fibril thickness (lateral assembly of monomers and small fibrils) rather than due to branching that increases overall network density (**Fig. 2**).

FIGURE 2







References

(1) Kaufman, L.J., C.P. Brangwynne, ..., D.A. Weitz. 2005. Glioma expansion in collagen I matrices: analyzing collagen concentration-dependent growth and motility patterns. *Biophys. J.* 89:635-650.

(2) Yang, Y.L., S. Motte, and L.J. Kaufman. 2010. Pore size variable type I collagen gels and their interaction with glioma cells. *Biomaterials*. 31:5678-5688.

(3) Guzman, A., M.J. Ziperstein, and L.J. Kaufman. 2014. The effects of fibrillar matrix architecture on tumor cell invasion of physically challenging environments. *Biomaterials*. In press, available online.

(4) Zhu, J. and L.J. Kaufman. 2014. Collagen I self-assembly: Revealing the developing structures that generate turbidity. *Biophys. J.* 106:1822-1831.

Relevant Definitions

Turbidity Cloudiness of a solution caused by individual particles scattering light; can be measured using a spectrophotometer. Limited because measurement only gives a value of how much light has not passed through

Confocal reflectance microscopy (CRM) Imaging technique dependent on contrast emerging from backscattered light; we use the total intensity of these images as a measure of the gelation process. Monomers and small fibrils are diffraction-limited, so "size" is much bigger in images than in reality

In situ turbidity (IST) Measure of forward scattering light calculated from intensity of scanning transmittance microscopy images. Essentially same as regular turbidity, except this has the advantage of being performed simultaneously with CRM

Goals:

Use Excel comfortably to:

- 1. Perform typical calculations using large amounts of data, including background subtraction, and averaging
- 2. Linearize data to extract slope and y-intercept
- 3. Interpret data in the context of a research problem

Experimental Approach:

Traditional turbidity measurements

FIGURE 3 UV spectrophotometer set-up



The collagen solution was made, neutralized with the appropriate amount of NaOH, then immediately transferred to a cuvette and placed into the cuvette holder. Gelation occured rapidly with the addition of NaOH, so this step was completed very quickly. The reference solution consisted of everything that normally goes into the collagen solution except for the collagen itself, which includes cell-friendly medium (even though no cells were used in these experiments, but normally we implant cells into gels), buffers, and distilled water.

For these experiments, standard turbidity measurements were performed using a dual-beam Cary 5000 spectrophotometer, which reported absorbance A

$$A = -\log\frac{I}{I_0} = \alpha p \tag{1}$$

with *I* the intensity of transmitted light, I_0 the intensity of transmitted light in the reference solution (the solution without collagen), α the absorption coefficient (cm⁻¹), and *p* the pathlength in cm. Turbidity τ was calculated via

$$\tau = \frac{A}{p} \ln 10 \tag{2}$$

These measurements were completed at certain wavelengths: 400, 458, 488, 515, 543, and 633 nm. 400nm is the standard wavelength used for collagen turbidity, while the other five wavelengths were chosen because they match with the lasers that we have in lab. Both the sample and reference were held in 1 cm quartz cuvettes, and measurements were taken every 10 s.

In situ *turbidity measurements*

FIGURE 4 Microscopy set-up



For these measurements, collagen solutions were made in the same way as above. Instead of transferring the solution to a cuvette post-neutralization, we transferred a smaller volume of the solution into a homemade cylinder that was epoxied to a glass coverslip bottom. The coverslip bottom is made of thin enough glass so that we can image using a 60x oil immersion objective, which has a small working distance of only ~50 μ m into the sample. In the publication, we imaged collagen gelation using three different channels, but for the data below, we will be focusing only on the scanning transmittance channel. Laser light passes through the sample and is collected using a photomultiplier tube, producing

images that get gradually darker as gelation occurs since less light is able to pass through the sample due to scattering. The total intensity of these images (sum of all pixels) is then extracted and translated into turbidity, also using Eqs. (1) and (2) and with the total intensity of the first image in the series as I_0 .

There is some discussion about what the pathlength should be for this particular set-up. This is important because of the p to use in Eq. 2. Based on the actual volume of the solution used and the dimensions of the cylinder, p would be 0.4 cm. However, due to the presence of a meniscus, the actual pathlength when measured with a ruler is more like 0.3 cm. Furthermore, because we use a confocal microscope with laser, the light is actually focused into a tiny spot so that potentially changes the pathlength as well. We attempt to correct for this in the data analyses.

Data analyses:

Wavelength dependent measurements of turbidity and in situ turbidity for 1.0 mg/ml collagen at 37°C

Purpose Use light scattering equations to extract information about fibrillogenesis.

The data here is mostly presented in the supplementary info. You will effectively be reproducing Figures S3/S4 from the paper.

There exists a known relationship between wavelength of light and turbidity:

$$\tau \lambda^5 = A\mu (\lambda^2 - Bd^2)$$
 with $A = \frac{88}{15} \frac{\pi^3 cn (\frac{dn}{dc})^2}{N_A}$ and $B = \frac{92}{77} \pi^2 n^2$ (3)

with τ turbidity, λ wavelength of light, μ mass per unit length, c mass concentration of collagen, n the refractive index of the solvent, dn/dc the refractive index increment, N_A Avogadro's number, and d the fibril diameter. For collagen in water, n = 1.33 and dn/dc = 0.186 cm³/g. Plotting $\tau \lambda^5$ vs. λ^2 yields a slope of $A\mu$ and a y-intercept of $-A\mu Br^2$.

- 0. Your first exercise will be to figure out the units for all of the components that go into Eq. 3. HINTS:
 - From Eq. 2, absorbance has no units, but τ has units of cm⁻¹.
 - *c*, concentration, has units of g/ml.

ANSWER:

```
[\tau] = cm^{-1}
[\lambda] = cm
[A] = cm^{3} x mol/g
[c] = g/ml = g/cm^{3}
[n] = unitless
[dn/dc] = cm^{3}/g
[N_{A}] = mol^{-1}
[\mu] = g/mol x cm^{-1} = Da/cm
[B] = unitless
[d] = cm
```

Quantitative information on fiber dimensions may be assessed through wavelength dependent turbidity measurements. Standard turbidity assays of these gels were performed at 400, 458, 488, 515, 543, and

633 nm. From these measurements, wavelength scans were constructed at each time point (every 10 s for 10 min). We will use this data in this section to tell us about the mass/length ratio and fibril diameter throughout gelation.

1. We'll start with the absorbance data from the UV spectrophotometer. Make a new sheet, and label it 'Step 1'. (It is recommended that you create a new sheet for every step in this lab.) Notice that at time \approx 0, slightly different absorbances are obtained across the trials. This might be due to not zeroing the instrument after every trial. Let's call this background and subtract the initial value for each trial.

EXCEL TIP #1: When you copy over the columns from the original data, do NOT just copy and paste the columns. The way to reference a cell from another sheet is by starting a formula by typing '=', clicking on the sheet tab, clicking on the cell you wish to reference, and hitting 'ENTER'. This should produce a formula of the form ='Sheet X'A1 (in this example "A1" is the cell on 'Sheet X', where "A" is the column and "1" is the row). Reference the original data so that if changes are made to this original data, these changes are applied throughout the workbook. The video below will show you how to do this. Once you link the first cell to the corresponding cell in the original data you can fill in the remaining cells in your new sheet. By placing your cursor on the bottom right corner of the cell you should notice that the cursor changes to a black cross, now drag down to the last row of data. To apply this formula to every column, highlight the column of data you just linked and drag across.

EXCEL TIP #2: The format =A\$1 may be useful, the "\$" tells Excel to keep the row number and constant. Once you enter your formula for background subtraction you can apply this calculation to an entire column. Now drag down to fill in the background subtracted data for the rest of the cells. To fill in the background subtracted data for the remaining columns, highlight the column and drag across.

Need help? Here is a link to a video tutorial specific to this step: Step 1 Excel lab: <u>https://www.youtube.com/watch?v=hC0Twv6kB0I</u>

2. Make a new sheet. Link over the columns for the time and the background-subtract absorbance values. For all the trials, calculate the turbidity from this absorbance using Eq. 2.

EXCEL TIP #3: Set-up your turbitity calculations in a new space, to the right of your absorbance data). (This will make step 5 *much* easier.)

Need help? Here is a link to a video tutorial specific to this step: Referencing Data: https://www.youtube.com/watch?v=UjWhNg68yhM

3. Change the time from s to minutes by inserting a new column to the right of your data in seconds, then enter an equation that will convert your seconds to minutes.

4. Although measurements were taken every 10 s, remember that some time elapses after the solution is neutralized and transferred to a cuvette before measurements are started. You are given the time corrections in s – convert the time delays to minutes, add them up for each trial, then take the average and standard deviation. Now divide the standard deviation by the average to estimate the difference in the shifts.

EXCEL TIP #4: Excel has many built in formulas, ie for calculating averages and standard deviations. To calculate the average, type in "=AVERAGE()" and select the range of cells; to calculate the standard deviation, type in "=STDEV()" and select the range of cells.

Because the shifts differ by less than 2% it is likely ok to ignore this correction to the time of measurements.

- 5. Obtain average turbidities (and the standard deviation) for the 2 or 3 trials at each wavelength (at each time by selecting across). Use Excel Tip #4 for the first formula and then drag down to the last row of data to repeat calculations at every time point.
- 6. Make a scatterplot of these six curves. The graph should be on a separate sheet, and should include: title, axes labels, appropriate scales for axes, and a legend. Additionally, use y-axis error bars to indicate the standard deviation of the turbidity values.

EXCEL TIP #4: Insert empty scatterplot. Create a new sheet, right click on any cell, and go to Select Data. Click "Add", then select the x and y columns. Do this for each curve separately, and label the data according to the wavelength. To add the error bars, click on a curve, go to Design \rightarrow Add Chart Element \rightarrow Error Bars \rightarrow Custom \rightarrow Specify Value and select the standard deviation for both positive and negative error values. Make sure that these are vertical error bars; if horizontal error bars show up, just delete them. Repeat for each curve.

Need help? Here is a link to a video tutorial specific to this step: Graphs and Error Bars: <u>https://www.youtube.com/watch?v=c7oFwRi4iH0</u> S1500X General Chemistry Laboratory Summer 2015 Writing Assignment #1 (10 points)

Assignment:

Use the Excel data and directions posted on Courseworks to produce the indicated graph (this was started after Exp #3). Then, read the Zhu and Kaufman collagen article posted on Courseworks from *Biophysical Journal*. Using the Excel activity, the journal article, and questions/answers from the author, write a commentary. Additional details are provided below.

Your assignment is to write a short **commentary** on the article and Excel activity. Include a short discussion of the article itself, as well as your final results and graph from the Excel activity in light of the article and feedback from the author.

The following are some questions that can guide your commentary about the article: Did you like the article or not? Does it relate to a topic that we discussed this semester in lab so far? What did you understand or not understand about the article? Did the article help to increase your knowledge of the topic?

Some questions to consider when discussing the Excel activity: What does the graph represent? What are the limitations or assumptions behind the data treatment that resulted in this graph? What did you learn from the Excel activity regarding data analysis?

The paper should be <u>at least 300 words</u> but no more than 500 words in length, 12 point Times New Roman font, double-spaced with 1" margins all around. You can use first person for this assignment. You will not be graded on scientific accuracy since some of the article is highly technical and may be difficult to understand and interpret. We are more interested in what you get from the article and your reaction to it, as well as the data analysis in the Excel activity.

In addition to the paper, you need to include the graph you generated in the Excel activity. The graph should be sized to one full page, scaled to use the graph space effectively. Make sure to label the axes and include a title. The title should be descriptive and more than just 'x versus y'. The graph does not count toward the paper length.

Deadlines:

The Excel activity file must be uploaded to the Dropbox tab on Courseworks for S1500X (as an .xls or .xlsx file) no later than <u>Wednesday</u>, July 8th.

The Writing Assignment must be uploaded to the TurnItIn site AND your Courseworks Dropbox (as .doc or .docx file) no later than <u>Friday</u>, July 10th.

- Approximately how many hours did you spend in class working on the Excel activity? *
- Approximately how many hours did you spend outside of class working on the Excel activity? *
- Please rank your experience and knowledge of Excel BEFORE starting the Excel activity (0 = no experience, 5 = expert). *
- Please rank your proficiency in using Excel AFTER starting the Excel activity (0 = no experience, 5 = expert). *
- 5. Did you find the video tutorials useful? (0 = didn't use them, 5 = extremely helpful) *
- 6. Did you find the in-class computer lab sessions useful? (0 = didn't use them, 5 = extremely helpful) *
- How comfortable are you using Excel to carry out your future data analysis on your own? (0 = not comfortable at all, 5 = extremely comfortable) *
- 8. If you learned a new technique with Excel please list at least one thing you learned doing this activity:
- 9. If you found the lab frustrating please list at least one thing you found frustrating about this activity:
- 10. Do you recommend we include this lab in future semesters? *(Yes, No, or Yes if the following issues were addressed.)
- 11. Issues (if any) from question 10
- 12. Did you find the background material (provided with the Excel activity) helpful?*
- 13. Have you read the original article yet? *
- 14. Did you find the Excel activity helped you better understand the original journal article? (0 = no, not at all, 5 = yes, the Excel activity was very helpful) *
- 15. Do you have any other feedback or considerations as we consider using this activity in future semesters?

*Required Questions