BCDB Qualifying Exam Part I: May 24 & 25, 2016

Instructions:

The following pages contain six questions. You must answer five out of the six questions, on each of the two days. Do not answer extra questions; indicate which one you are skipping on each day. Any one question may ask for more than one response so please read each question carefully for specific instructions pertaining to that question.

All answers are to be typed on the provided computer and saved to a provided USB drive. Hand-written figure pages will be included with your typed answers when they are processed and distributed to graders. SAVE ANSWERS FREQUENTLY, in case of computer issues.

Start typing your answer after the end of its question. Do not make changes to the format, font, color, etc. Do not change the header or footer. For the question you choose not to answer, just type "SKIP". Use the provided pages for your figures.

Please remember as you answer the questions that you have approximately one hour per question. This means that we expect in depth answers unless otherwise noted. Use diagrams to illustrate your answer wherever possible. Although each question is worth the same number of points, not all questions will require the same amount of time. Allocate your time wisely. We recommend that you first make an outline of your answer, rather than just making up your answer as you write. Some questions ask you to propose experiments. Choose the most direct and realistic approaches and explain your experimental rationale as clearly as possible. Be sure to include controls, expected outcomes and possible problems and solutions.

Do NOT put your name on any of the question or answer sheets. To keep the exam anonymous, label your figure pages with the question number (*e.g.,* Question 1) and with a coded name using the code distributed by Susan.

PLEASE MAKE SURE THAT ALL FIGURE PAGES ARE NUMBERED and IDENTIFIED!!!!!

This is a closed book exam. Absolutely no discussion will be allowed between the students while the exam is in progress. Cell phones must be silenced and in a bag – NOT accessible on the desk! You will be held to an honor code by agreeing to not receive or give aid on this exam.

You will have from 8:00 a.m. until 2:00 p.m. to finish the exam. Lunch will be available from (about) 12:00-1:00. No exam materials can leave the room with you during lunchtime or any other time, and you may not refer to outside materials at any time. The exams will be collected no later than 2:00 p.m. sharp.

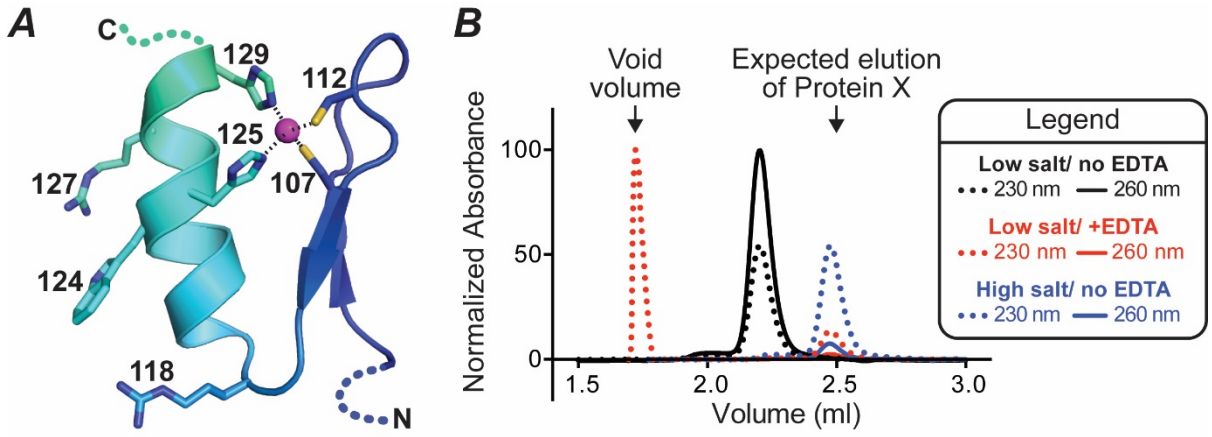
If you have questions during the exam please contact Sho Ono: (404) 727-3916. In case of emergency or no answer, call his cell (770) 940-9272.

If you need supplies, call Susan: 404 727-1594.

Be sure to save your final exam version to your USB drive. Put the drive and all papers into your envelope. If you have figure papers that are NOT to be used, just draw an X on them, fold them, or tear them in half.

Question 1

You have been tasked with analyzing the structure and interactions of Protein X, a factor predicted to have a role in control of gene expression. Structural homology to well characterized proteins is limited but you have been able to develop a model for a domain within Protein X and to identify 7 residues (labeled in **Figure A**) potentially critical for its function. You first examine your purified protein using an analytical gel filtration column and find that Protein X elutes at a volume different from that expected (**Figure B**, black chromatograms). You suspect that Protein X may co-purify with an endogenous binding partner and test the effect of a range of different solution conditions on Protein X's elution profile. Two changes have the greatest impact: addition of EDTA (+EDTA; red chromatograms) and addition of 1 M NaCl (High salt; blue chromatograms).



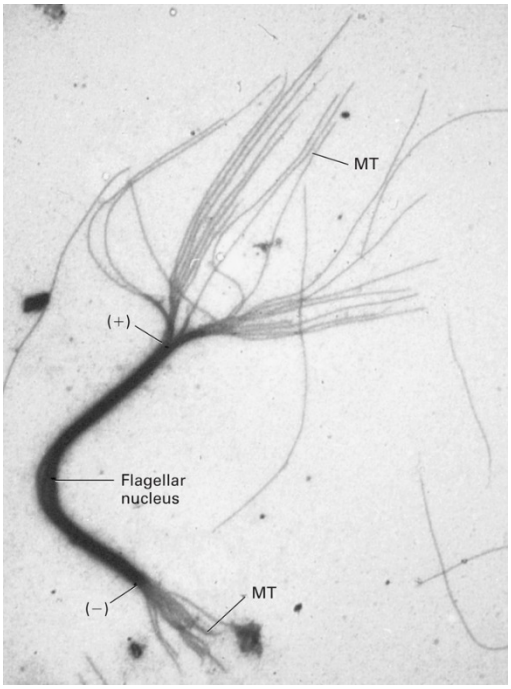
(A) Briefly explaining your reasoning, from the data presented what do you surmise about:
i) the identity of the magenta sphere and its function within this protein? (**2 points**)
ii) the co-purifying macromolecular binding partner of Protein X? (**2 points**)

(B) Describe an experiment with appropriate controls to test your hypothesis about the nature of Protein X's binding partner. (**6 points**)

(C) You are able to purify both Protein X and the binding partner in sufficient quantities for crystallization but the complex fails to crystallize. Describe an alternative approach to map the interaction surface within Protein X, explaining the rationale for your choice. (**4 points**)

(D) Describe the types of interaction you expect each of the following residues: 118, 124, and 127 to make with Protein X's binding partner and explain in detail how you would test their importance for that interaction. Be sure to include appropriate controls to ensure the validity of your conclusions. (**6 points**)

Question 2



This electron micrograph has captured the results of in vitro polymerization of microtubules (MT) onto an isolated ciliary axoneme. The ciliary axoneme is the thick, dark staining structure (labeled as flagellar nucleus).

(A) What are the definitions of the plus and minus ends of the microtubule? What is the relationship of the plus end of the microtubule to the orientation of the alpha and beta tubulin in the tubulin dimer? Illustrate as required. (**10 points**)

(B) The ciliary intraflagellar transport (IFT) motor Kin2 is called a "plus-end" directed motor. What does this mean? (**5 points**)

(C) Propose ONE experiment used to distinguish whether Kin2 is a plus- or minus- end directed motor. Illustrate as required. (**5 points**)

Question 3

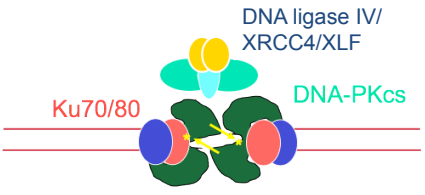


Fig. 1. From a Powerpoint provided as background

(Background.) In Foundations, we discussed the application of single-molecule methods to the study of DNA double-strand break (DSB) repair. Some background: we studied a specific pathway called nonhomologous end joining (NHEJ). This pathway involves six polypeptides, acting as follows (Fig.1):

- Recognition of DNA ends by Ku70/80 dimer
- Binding of a regulatory kinase, DNA-PKcs and subsequent autophosphorylation and displacement.
- Recruitment of DNA ligase IV, XRCC4, and XLF and DNA ligation.

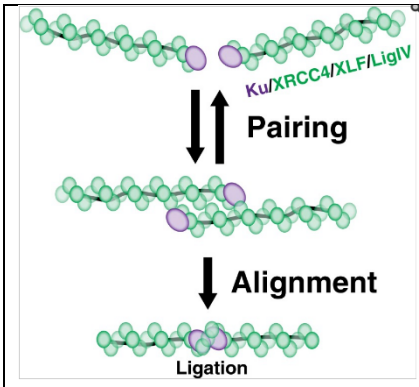


Fig. 2. Model from Reid et al.

The paper we read by Reid et al. examined the last of these steps in detail. The paper presented evidence that DNA ligase IV, XRCC4, and XLF bind in multiple copies to form a filament along each DNA. Single-molecule FRET was then used to examine the interaction of these filaments. These were TIRF experiments where one DNA was tethered to a surface and the other was free in solution. The DNA ends were labeled with a FRET donor-acceptor pair so that the distance between them could be measured. The donor fluorophore was excited and FRET was measured. A typical trace showed an initial, oscillating FRET signal, which was interpreted as alternation between side-by-side pairing of the DNAs (where FRET is weak and variable) and end-to-end alignment (where FRET is high and constant). In the authors' model, the end-to-end complexes then undergo ligation, resulting in a stable high FRET signal (Fig. 2).

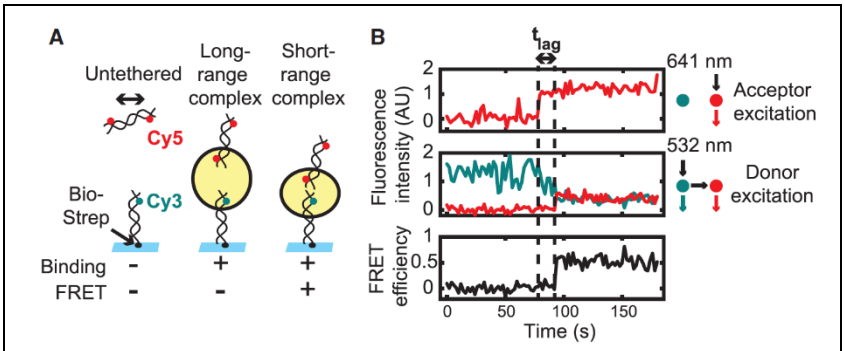


Fig. 3. Single molecule data from new paper.

Another single-molecule study of NHEJ has recently been published. The TIRF setup was similar to the first study. However, rather than incubating with purified proteins, the new study used a crude cellular extract containing all of the core NHEJ factors (not just a subset), and many other proteins as well. The authors of the second study put forward a two-stage model, different than the one in the first study (Fig. 3A). They propose a long-range complex, drawn quite differently than the paired complex in the first study, which transitions to a short range complex, drawn as very similar to the aligned complex in the first study. Fig. 3B shows an example of supporting data. In this single molecule FRET experiment, one DNA substrate was tethered to a surface, and the other was untethered. The sample was subjected to high angle illumination (TIRF) at 532 nm (Cy 3 excitation) and 641 nm (Cy5 excitation) in rapid alternation. Fluorescence was continuously monitored.

(Continue on the next page)

Question 3 (Continued from the previous page)

(Questions:)

(A) Interpret the traces shown in Figure 3B and explain how they support the model presented in Figure 3A. Specifically, explain why the red and green curves rise and fall as they do. Explain how the curve in the bottom panel is derived. Explain how these data differ from the model in the first study (**6 points**).

(B) Under the experimental conditions employed, traces exactly like those shown in Figure 3B were actually quite rare. In most cases the long-range complex simply dissociated without undergoing the transition to a short range complex. To be able to study the transition more efficiently, the authors developed a hack, shown in Figure 4, where a single long DNA was tethered in the middle to the surface. Consequently, the range of motion of the two DNA ends was constrained – they could not diffuse out of the TIRF zone, and they continuously encountered one another.

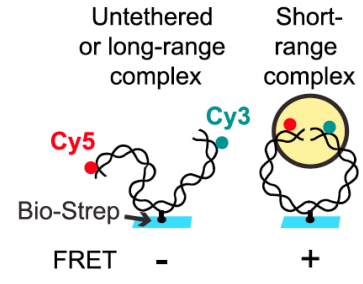


Fig. 4. Circularization substrate.

Using this substrate, the authors were able to measure repeated cycles of formation and dissociation of short-range complexes (there was no covalent end joining under the conditions used). Draw hypothetical traces, showing this repeated cycle for a single molecule. Draw three traces, like those in Figure 3B, showing what happens when Cy5 is excited directly, Cy3 is excited directly, and FRET efficiency is depicted **(6 points)**.

(C) Imagine you were a reviewer for the second study, and you criticized the authors for not offering an explanation why their results differed from the first study. Assume that both studies were technically flawless. Propose the most plausible explanation for the difference; note that the Background section of this question provides all the information needed (**3 points**).

(D) Propose an experiment that would let you test this explanation being sure to include proper controls to interpret your results (**3 points**).

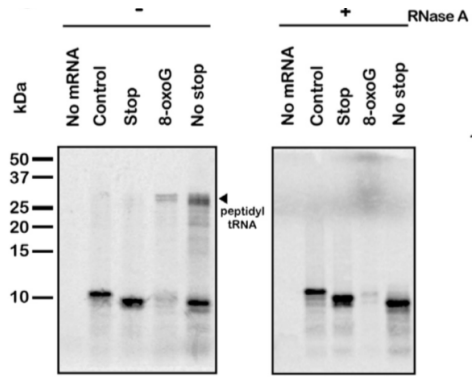
(E) In your opinion, which study is likely to provide more insight into how this process happens in cells? Briefly (a sentence or two) explain the rationale for your response (**2 points**).

Question 4

The identification of modified nucleotides in both DNA and RNA signals an additional layer of regulation only recently appreciated. In particular, the addition of modifications to RNAs post-transcriptional either by environmental damage or as a mechanism to alter gene expression suggests a role to change the rules of protein synthesis. While looking through the published datasets of high throughput sequencing data, you find that the mRNA of the human protein you study contains a number of 8-oxoGuanosines which is a modification to the guanine base that expands its capacity to basepair with adenosine.
Since you know that your protein of interest has a number of different isoforms and mutations that cause a defect in function, your hypothesis is that the 8-oxoG modification may be the reason for these protein expression changes.

(A) You want to test the effect of the 8-oxoG mRNA modification on its ability to be translated in wheat germ S30 extracts. To do this you will add synthetic mRNAs (that have been modified to resemble authentic eukaryotic mRNA) to the extract and test its ability to be translated. Draw your eukaryotic mRNA and identify the four most important parts of the mRNA transcript needed for translation. Remember the mRNA will not need to be processed but it should be designed ready to be translated. (**4 points**)

The results of *in vitro* translation assay of this mRNA are shown below:



(B) What is an S30 extract and why perform the assays in this extract? (**4 points**)

(C) What kind of experiment and gel would you need to run to visualize *only* your protein products? (**2 points**)

(D) In addition to the 8-oxoG mRNA experiments, you also performed three controls called "control", "stop" and "no-stop". Based on the results from the above gel, explain how you think the reactions that resulted in these three control lanes were done. (**4 points**)

(E) On the right gel, RNase A was included after translation and just before samples were loaded on the gel. Explain the rationale behind this control and what the results from the figure tell you. (**2 points**)

(Continue on the next page)

Question 4 (Continued from the previous page)

(F) Your experiments seem to suggest that the 8-oxoG mRNA alters protein synthesis by a stalling mechanism. What is the evidence that suggests this from your gel analysis of the assay? (**2 points**)

(G) One other possibility is that modified mRNAs caused by environmental damage may be degraded before they are translated into erroneous polypeptides. Define and describe one mechanism by which eukaryotic mRNAs are degraded. (**2 points**)

Question 5

As a new graduate student, you have been assigned to characterize a gene in the mouse. You would like to start by examining its expression, but no antibody exists to detect the protein. Your advisor suggests a transgenic approach, where you tag the protein with green fluorescent protein (GFP) and incorporate it randomly into the genome. After successfully isolating five transgenic lines, you begin to analyze patterns of GFP expression. Curiously, three of the lines show clear expression in neurons; however, one line expresses in the heart, kidney and lungs, and the last line has no detectable expression at all.

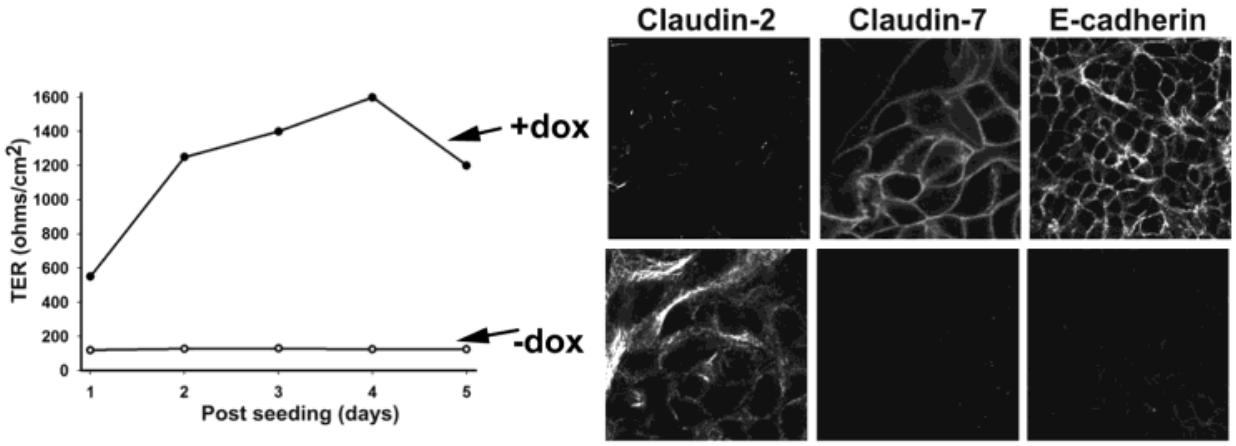
(A) Using your knowledge of chromosome biology, explain what has happened in each of your transgenic lines. (**3 points each: total 9 points**)

(B) For either the line that expresses in the heart, kidney and lungs, or the line that has no detectable expression at all, explain molecularly what has happened to the transgenic construct (**3 points**) and describe one assay that you can use to test your hypothesis. Be sure to include controls (**5 points**)

(C) If you want to create a transgene with a more consistent expression pattern between different lines, what could you include in your transgenic construct? Explain why your modification is expected to result in consistent expression. (**3 points**)

Question 6

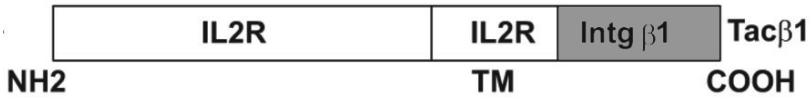
You are interested in how integrins regulate kidney epithelial barrier function and have obtained proximal tubule (PT) cells from tetracycline inducible Cre recombinase (Tet-iCre); β1 integrin flox/flox mice. In one experiment, you take the PT cells cultured on Transwell permeable supports and incubate them for 3 days in either the presence or absence of doxycycline (a derivative of tet). Then, you analyze them for transepithelial resistance (TER) and expression of claudin-2, claudin-7 and E-cadherin by immunofluorescence. Your results are the following:



(A) Under which circumstances (+ vs – doxyclycline, dox) is the PT cell barrier tighter? You measure paracellular flux with a small fluorescent dye (calcein). Based on the TER data, would you predict the paracellular flux of calcein to be higher for untreated cells or for cells treated with doxycycline? (**4 points**)

(B) Propose a molecular mechanism for how the changes in junction protein expression affect PT cell barrier function. (**4 points**)

You then transfect β1 integrin deficient PT cells with an IL2-receptor/β1 integrin chimera where the extracellular and transmembrane domains are from the IL2-receptor and the cytoplasmic tail is from β1 integrin:



(Continue on the next page)

Question 6 (Continued from the previous page)

When you measure barrier function and expression/localization of claudin-2, claudin-7 and E-cadherin, you find that this construct is able to replace the function of full length β1 integrin.

(C) Why is the result with the chimera surprising? (**4 points**)

(D) Propose a hypothesis for how β1 integrin alters claudin and cadherin expression to alter barrier function. Provide two independent ways to test your hypothesis that do NOT involve CRISPR/Cas9. (**8 points**)

Question 7

(A) Tandem Mass Spectrometry (LC-MS/MS) was used in efforts to identify sites of post-translational modification in a 100 kDa protein that was expressed in HeLa cells as a (C-terminal) FLAG-tagged protein. The protein was then purified by immunoprecipitation using the FLAG antibody and appeared as a single band in SDS-polyacrylamide gel electrophoresis, migrating consistently for a 100 kDa protein. Three peptides were found by LC-MS/MS to be 80 Da larger than predicted and thus were thought to be phosphopeptides. However, sulfation is known to add the same mass as phosphorylation and can occur on the side chains of the same amino acids. List **TWO** different methods you could use to distinguish between sulfation and phosphorylation. (**2 points each- total of 4 points possible**)

(B) Co-immunoprecipitation (co-IP) is a method that can be used to identify candidates for novel binding partners for a protein of interest (protein X). Protein proximity interactomes (PPIs) is a newer technology that uses a covalent modification, e.g., BirA is a biotin ligase that non-specifically biotinylates nearby side chains, to identify proteins in close proximity (30-40 nm) to your protein X-BirA fusion protein.

i) Describe a situation (e.g., a question you want to answer) in which each of these would be the preferred approach over the other and why it would be the better choice. (**5 points**)

ii) Describe the controls you would include for each approach you propose in part (i) to increase your confidence that the datasets resulting would represent true binding or proximity partners. (**5 points**)

iii) Describe **TWO** follow-up experiments that you would perform to determine if protein Y, found by either PPI or co-IP, interacts directly with protein X. (**1.5 points each- total of 3 points possible**)

iv) Describe an approach to determine if the interaction of protein X and protein Y occurs in a biologically relevant context. (**3 points**)

Question 8

The muscle-specific protein that you are studying, muscle-specific protein 1 (MSP1), is readily detectable by immunoblotting of muscle lysate but absent from any other tissues you examine. You find that you can recapitulate this expression as MSP1 is readily detected in a C2C12 mouse muscle cell line but is not detectable in lysates from a mouse fibroblast cell line. Analysis of RNA from the same cells indicates that the mRNA transcript encoding MSP1 is also undetectable in the fibroblast cell line but is readily detected in the C2C12 cell line.

(A) Design an experiment to determine whether the absence of MSP1 expression in the fibroblast cell line is likely to be due to transcriptional or post-transcriptional regulation. (**4 points**)

(B) From a series of experiments, you determine that *MSP1* transcription is, in fact, decreased in the fibroblast cell line relative to the C2C12 cells. The transcription factor, Mtx, is the only known trans-activator for *MSP1* and you show by immunoblot analysis that Mtx protein is expressed at equal levels in both muscle and fibroblast cell lines. There is a single strong Mtx consensus binding site located upstream of the *MSP1* promoter. You hypothesize that this site is a binding site for Mtx. Describe a biochemical experiment (cell free) to determine whether the consensus site you have identified does bind to the Mtx transcription factor. Be sure to include appropriate controls to ensure the specificity of binding. (**4 points**)

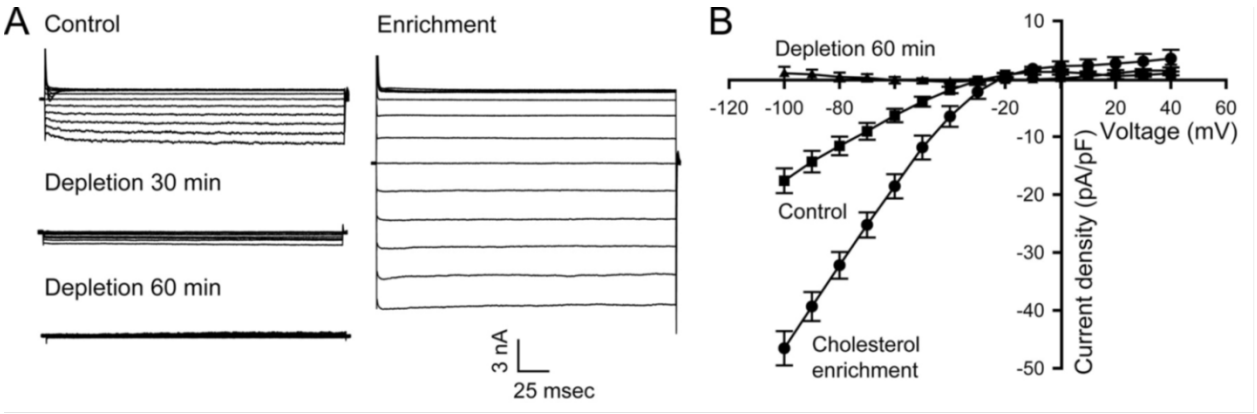
(C) Assuming that the Mtx protein can bind to the Mtx binding site in your biochemical experiment, now design an experiment to test whether this binding is regulated in different cell types. Your data predict that Mtx should bind upstream and activate transcription of the *MSP1* gene in the muscle cell line but NOT in the fibroblast cell line. Describe an experimental approach to test this prediction. Be sure to include appropriate controls and provide an explanation for how the results support or refute (or help you to modify!) your prediction. (**4 points**)

(D) Your experiments from (B,C) indicate that while Mtx can bind the *MSP1* promoter in cell-free/biochemical experiments and the C2C12 cells, no binding can be detected in the fibroblast cell line. Wanting to understand how this regulation occurs, you sequence the *Mtx* gene in both the muscle and fibroblast cell lines and find no polymorphisms or sequence changes that could explain your results. A postdoc in the lab suggests that you test whether treatment of the fibroblast cell line with Trichostatin A (TSA), a histone deacetylase (HDAC) inhibitor, affects MSP1 expression. Indeed, you observe a significant increase in MSP1 protein levels when the fibroblast cell line is treated with TSA. Based on all your results, propose a model for the absence of MSP1 expression in the fibroblast cell line and design an experiment to fully test your model, including appropriate controls and a description of the results that would support your proposed model. (**8 points**)

Question 9

You are interested in whether a cardiac arrhythmia in patients with high cholesterol is explained by an effect of cholesterol on acetylcholine (Ach)-stimulated G protein-regulated inwardly rectifying potassium channels (Kir3). You isolate mouse atrial cardiomyocytes and culture them either in methyl-β-cyclodextrin (MβCD) to deplete cholesterol in the plasma membrane (labelled "depletion" in figures below) or in MβCD+cholesterol solution to enrich cholesterol ("enrichment" in the figures below).

You then generate current/voltage curves from these cells using the voltage protocol shown in Panel C. The experiment is performed in the presence of Ach with and without the Kir3 channel inhibitor tertiapin. The currents in the presence of tertiapin are subtracted from the currents in the absence of tertiapin to obtain the tertiapin-sensitive currents shown in Panels A, B, D. The voltage clamp experiment is performed on control cells and cells with depleted or enriched cholesterol (Panels A,B). You then repeat the experiment for control and cholesterol-depleted cells in the presence of GTPγS in the cytoplasmic solution (Panels C, D).



(Continue on the next page)

Question 9 (Continued from the previous page)

(A) Explain how the traces in panel A were obtained and what they represent.

i) What is the recording configuration? (**2 points**)

ii) What does current density (pA/pF) mean? How was this determined? (**4 points**)

iii) What does the difference in outward vs. inward current tell you about how these channels function? (**2 points**)

iv) What information can you obtain from the reversal potential? In what direction are ions moving? (**2 points**)

(B) The effects of cholesterol on current density could be due to changes in total number of channels (N), channel open probability (P_o) and/or single channel conductance.

i) How do these parameters affect current density? (**2 points**)

ii) Propose a method to distinguish which of these parameters is changed in the presence or absence of cholesterol. Provide as much detail as possible by drawing the expected results of your experiment. (**2 points**)

(C) The effects of GTPγS on cholesterol sensitivity of Kir3 currents shown in panels C,D, suggest a role for heterotrimeric G proteins in mediating or regulating the effects of cholesterol on Kir3.

i) Propose a hypothesis for how G proteins and cholesterol regulate Kir3. How would you test this hypothesis? (**6 point**)

Question 10

You are studying the mechanism by which stereocilia in inner ear hair cells bend and transduce a signal in response to sound waves. From a proteomics analysis of these cells you discover a protein that has some similarities to the giant muscle protein titin. At its N-terminus, this protein, called "vibrostretchin", has 5 consecutive immunoglobulin-like (Ig) domains followed by a 900 residue sequence with no known homologies or domains, followed by a 500 residue sequence that is predicted to form a coiled-coil, and 5 consecutive calponin-like repeats, for a total predicted molecular weight of ~250 kDa. You raise antibodies to the protein, and immunofluorescence staining reveals that it is located at the base of stereocilia. The 900 residue region contains multiple types of short repeats, each 20-100 residues, many in tandem copies, overall low sequence complexity, and no predicted secondary structure elements.

(A) Given the nature of the 900-residue sequence, how do you think it would respond to a pulling force? Design an experiment to test your hypothesis, including positive and negative controls. Show graphs of the results of your experiments. (**5 points**)

(B) What is a coiled-coil and how can we explain its assembly? What does having a coiled-coil region in vibrostretchin predict for the structure of vibrostretchin? Describe two experiments that would allow you to test your prediction. (**5 points**)

(C) You suspect that the calponin-like repeats bind to filamentous (F-) actin. Describe an experimental approach to determine whether the calponin-like repeat region of vibrostretchin: (1) binds to F-actin, and (2) bundles F-actin. (**5 points**)

(D) Assuming that you have a cell that can be induced to form stereocilia, how would you investigate at which stage of this assembly process vibrostretchin is involved? You have succeeded in demonstrating that the region containing calponin-like repeats do bind to F-actin. Describe two experiments to test whether this F-actin binding is important for the localization of vibrostretchin. (**5 points**)

Question 11

You are studying the regulated expression of a water transporter, AQUA1, in kidney tissue. A transcription factor, DBF1, is suspected to be involved. There are human mutations in the gene encoding this transcription factor that all cause defective water retention and dehydration. One class of mutations results in an amino acid substitution within the predicted DNA binding domain of DBF1; the other class of mutations causes amino acid changes within the domain that is predicted to form a coiled-coil domain structure that is a known protein:protein interaction domain.

Both classes of mutations result in decreased expression of the *AQUA1* transcript.

Read all parts before answering

Part I (12 points)

(A) Propose a detailed hypothesis for why amino acid substitutions in the predicted DNA binding domain of DBF1 result in decreased expression of the AQUA1 transcript. (**2 points**)

(B) Design and describe an experiment to test your specific hypothesis. Describe what result will support your hypothesis and also describe a result that would not support your hypothesis. Be certain to describe what controls will be included in your experiment, and why these controls will be essential for interpreting your results. (**4 points**)

(C) You identify a candidate DNA binding motif for DBF1 within the *AQUA1* promoter. Describe how you would test whether DBF1 directly binds to this 6-base pair motif with high affinity and specificity. Use appropriate methods and controls that would allow you to draw robust conclusions about your results. (**4 points**)

(D) During your analyses, you identify one amino acid, arginine 471, within the DNA binding domain of DBF1 that is altered in multiple patients. In one patient a mutation results in a change from arginine to lysine (R->K). In a different patient, the same arginine is altered to aspartic acid (R->D). In which patient do you predict the most severe impact on the DNA binding activity of DBF1? Why? (**2 points**)

Part II (8 points)

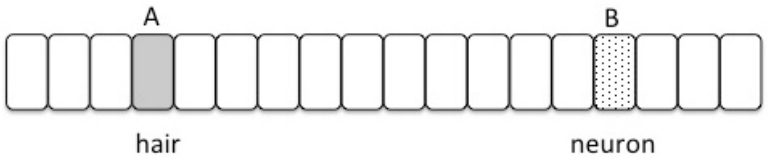
(A) Propose a detailed hypothesis for why amino acid substitutions within the coiled-coil domain of DBF1 affect *AQUA1* expression. (**2 points**)

(B) Design and describe an experiment to test your hypothesis. Describe what result will support your hypothesis and also describe a result that would not support your hypothesis. Be certain to describe what controls will be included in your experiment, and why these controls will be essential for interpreting your results. (**4 points**)

(C) Given the presence of the coiled-coil domain within the DBF1 protein, propose a potential sequence for the 8-base pair DBF1 DNA binding site that you analyzed in **Part I (C)**. What is the rationale for the sequence of the binding site that you proposed? (**2 points**)

Question 12

Within a sheet of developing epithelial cells shown below, two cells will invariably develop into a hair cell (A; grey fill) or a neuron (B; dotted fill) in the mature tissue. Laser ablation (i.e. killing) of the presumptive hair cell leads to an epithelial sheet that proceeds to develop without a hair cell; in this case, development of the neuron is unaffected by killing the hair cell. By contrast laser ablation of the presumptive neuron leads to loss of the neuron in the mature tissue, but it also causes a failure of the presumptive hair cell to adopt a hair cell fate as development proceeds.



(A) Provide a hypothesis as to how these data can be explained by the action of genes and proteins. (**5 points**)

(B) Design an experimental approach that would test each aspect of your hypothesis. Assume that a full-range of genetic tools are available to express or remove genes/proteins at will in any of the cells, even in single cells or in chosen subsets of cells in the sheet.

i) This includes an experiment to test **expression patterns** of the factors involved, and tools you would use to document this pattern(s). You may invoke the existence of any key reagents/probes to carry out these studies. (**5 points**)

ii) This also includes an experimental approach that simultaneously **tests your model and rules out a competing model** of how the presumptive neuron can affect hair cell fate. (**5 points**)

iii) This also includes a discussion of how you would **interpret the data** generated by your proposed experiments to support or refute your model. (**5 points**)