Summer Professional Development Workshop 2021

Authoring a Molecular Case Study

Based on the instructions for engaging students in writing molecular case studies (submitted to BAMBED) and discussions in the Biome Institute 2020 and FMN in Spring 2021.

Table of Contents

Goal of Molecular Case Studies

Make a clear link between the molecular structure (including chemical interactions) and biological function(s) of a protein.

Consider how this case will help students

- Apply foundational knowledge about protein sequence, structure, and function of a specific protein of interest.
- Find and interpret information from the scientific literature and/or other bioinformatics data resources that are related to the protein or system of interest.
- Study examples of how similar biochemical principles are applied to solve a diverse set of biological and biomedical problems.

Consider using the following community standards to guide the case's specific learning objectives:

- Biology: ASBMB Foundational concepts
- Chemistry: Macromolecular, Supramolecular, and Nanoscale (MSN)Systems in the Curriculum
- Bioinformatics: Core Competencies
- Molecular Visualization: BioMolViz Framework
- Vision and Change in Biology (BioCore): BioCore Guide

Other disciplinary community standards can be added here based on participating educator interests.

Steps for authoring a molecular case study:

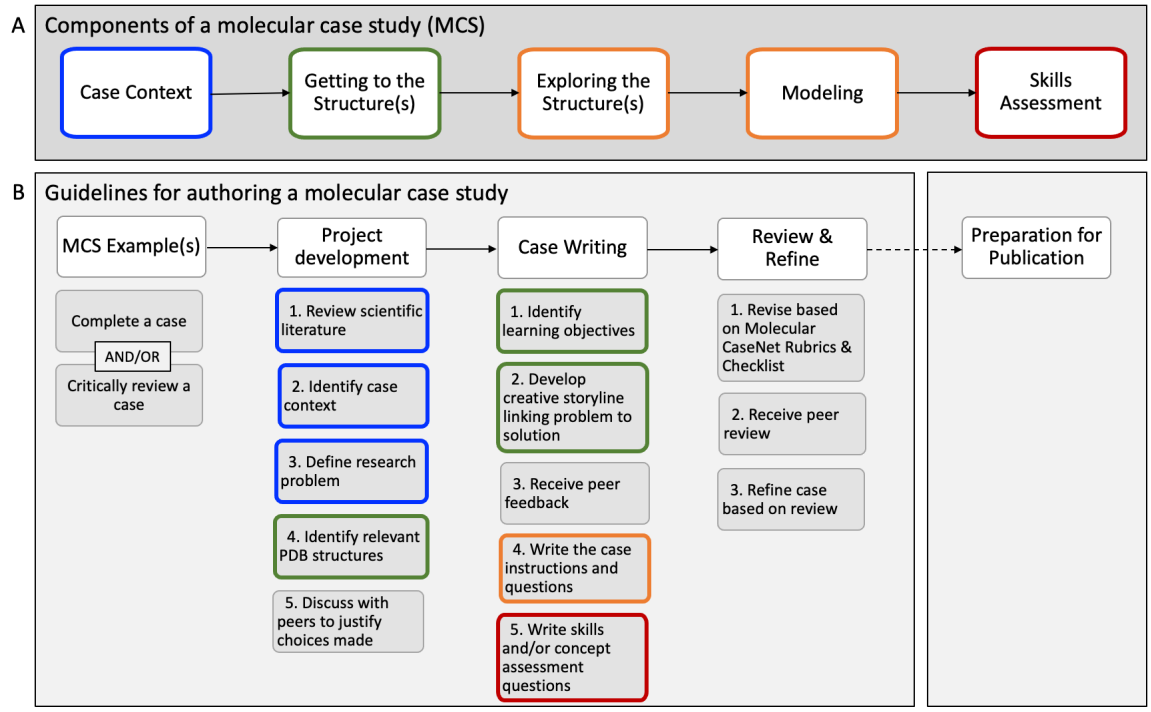

Figure 1: A. The main components of a molecular case study; B. Guidelines for authoring a molecular case study

1. Molecular Case Study Example(s)

Complete a published case and/or critically review either a published or late-stage draft of an example molecular case study (MCS). Reflect on the following:

- What was done well?
- What should be improved? and
- What did you learn from the way these cases are written?

2. Project development

a. **Review scientific literature:** Select a biological system of interest or learn about the biological system you are assigned for your case

b. **Identify case context:** Outline the specific scientific question(s) you would like to answer within this biological system and identify the protein players that are involved in the chosen biological context.

c. **Define research problem:** The case study may focus on

- **Physiological function vs altered function:** Identify a well defined and specific function and a change in that function that is being investigated in the case.
- **Native structure vs altered structure:** There should already be a minimal amount of available structural information in the protein data bank about this protein and its altered form relevant to the functional change investigated. Note that sequence is the primary structure of a protein.

Details for how to include explanations for these are described in the "Describe the research problem" section below.

d. **Identify relevant PDB structures:** Perform preliminary search in the Protein Data Bank (PDB, www.rcsb.org) to determine if adequate structural information is available for the proteins identified in the previous step and determine the main protein structure you want to build your case around.

Consider:

- Which one(s) would help in answering the functional question at hand?
- Are there multiple structures of the protein in the PDB (e.g., bound to ligands/partner proteins, at different temperatures, pH, etc.)? Can these structures be compared/contrasted to learn something related to the case or to make a specific point?
- What do you want to learn from the 3D structure that cannot be learned by looking at its primary structure alone?
- Note the structural method(s) used to elucidate this/these structures? Are there specific considerations due to limitations of the method(s) used?
- What is the focus of the structural experiment - i.e., why did the authors solve the 3D structure? (were they examining the apo structure, inhibitor bound structure, mutant structure etc.)
- What are some of the interesting/unique features and spatial organization of the protein molecule (such as domains, motifs, secondary structures, etc) that are relevant to the function investigated?
- What is the quality of the structure? (Examine validation reports where possible)
- What unique information could be obtained or hypothesized about the protein based on the deposited structural information?
- What aspects of the amino acid sequence, 3D protein structure and biological function are worth exploring in connection to the case?

- **Explore other bioinformatics resources:** Expand your search for the biomolecules relevant to the case into other biological databases and identify tools beyond molecular visualization to supplement your story line. Consider resources that look both at DNA/RNA sequences and protein sequences, structures, functions, and more. For example:
    - Gene Sequences: **NCBI Genbank**: https://www.ncbi.nlm.nih.gov/genbank Protein Sequences: **UniProt**: https://www.uniprot.org/
    - Protein secondary structure prediction **JPred**: https://www.compbio.dundee.ac.uk/jpred
    - Protein family annotations (functions, domains etc.) **Pfam**: http://pfam.xfam.org
    - Metabolic/signaling pathways **KEGG PATHWAY**: https://www.genome.jp/kegg/pathway.html
    - Comparing protein/nucleic acids sequences **BLAST**: https://blast.ncbi.nlm.nih.gov/Blast.cgi
    - Comparing protein sequences **CLUSTAL Omega**: https://www.ebi.ac.uk/Tools/msa/clustalo

Consider:

- What can be learned quickly/easily from the primary and/or secondary structure of proteins, and what cannot?
- How can you use the DNA or RNA sequences to learn more about the wild-type or a mutant protein?
- Are there proteins with related/redundant functions? Is your protein part of a metabolic pathway? Signaling pathway? Multi-protein complex?
- How conserved is the protein across species? Does it make sense to do a multi-species alignment in CLUSTAL OMEGA?

e. Justify choices made: The justification will ensure that the choice of PDB entries and information gathered from various resources are supported by primary literature and/or data supporting experimental data for the claimed link between the structure and function of the central protein. The structure(s) and bioinformatics data selected should help answer the problem or question in the case.

3. Case Writing

a. **Identify learning objectives:** All major projects should begin and end with the objective of the study. Your learning objectives should be appropriate for the specific audience you are targeting. Each of the sections of your case should have 2-3 learning objectives that all tie to one main overall learning goal for the case. Learning objectives should be matched to existing standards whenever possible. Need help writing Learning Objectives? Here are some suggested tips.

b. **Develop a creative/ engaging storyline:** A context for the molecule that is delivered in a non-scientific language aimed to capture the interest of the audience. This "hook," which can be in the form of an image, a short audio/video or written piece (such as a newspaper article or story), will constitute the introduction to the case and convey the biological/ chemical/ ecological etc. question to be addressed in the study.

Consider:

- Who is your audience and what is the take home message you would like them to have after completing this case study?
- Are there YouTube videos (or home-made videos) or articles to get a reader hooked into your story?
- Are you going to create a fictional story line or can you use a real life example from a newspaper story or clinical publication?
- Is your storyline entirely scientifically plausible?
- Are the images, videos, stories used for the case appropriately attributed? If using any copyrighted materials all necessary permissions should be obtained.

c. Write case instructions and questions: Each MCS needs to include the following five sections: Preparation (case context), getting to the structure(s), exploring the structure(s), modeling, and skill assessment. You can have multiple subdivisions or modules under each of these sections

- **Decide the number of subsections/modules your case study will have** (under each of the five mandatory MCS sections).
- **List the literature data, bioinformatic tools, and PDB structures** you will use in each section to meet your stated objectives. You must explore molecular interactions within the protein as well as between the protein and a binding partner for the structure/function relation sections of the case.
- **Write the instructions for how to obtain information** related to the case:

Consider:

- Where and how will readers find information? How can you provide or help them obtain information?
- If some of your information should be described in embedded text, figures, or tables—keep the text short and sweet and custom tailor your figures/tables.
- You can use some already-written instructions or a YouTube tutorial for finding specific details like a hydrogen bond or coloring an individual residue, etc.?
- Is there a logical order for exploring various aspects of the case?

- **Write your case questions:** Each learning objective should be measured with 1-2 questions (assessments) to determine if students have met the objective.

Consider:

- After the student obtains the structure, what will they do to gain skills in understanding how to interpret the 3D image?
- How do your questions require the reader to demonstrate the link between molecular structure and function?
- Do you have variety in the kinds of questions you are asking and the ways in which the student can respond (e.g. by pasting an image, a M.C. question, a brief essay, etc.)?
- Do your questions require the students to engage with the figures or structure, or can they just answer the questions without them?
- Are the questions written with the target audiences' prior exposure to the subject matter and using inclusive and accessible language?
- Need help writing multiple choice questions? Check out the following resources
  - Writing Good Multiple Test Questions (Vanderbilt University)
  - Best Practices for Designing Exams (University of Michigan)
  - Designing Effective Multiple Choice Questions (McGill University)
  - Best Practices for Designing and Grading Exams (University of Michigan)
  - Guide to Writing Multiple Choice Questions (from National Board of Medical Examiners, NBME)

- **While writing the case you may have to describe the research problem:**
  - **Physiological function vs altered function:** Identify a well defined and specific function and a change in that function that is being investigated in the case.

Within the introduction section of the case, there should be a well articulated question that clearly ties the hook (introductory video, image, story etc.) to the investigated functional question.

There should be a dedicated section in the case that delivers adequate biological and chemical background to explain the biological system at hand and help appreciate the importance of the investigated question.

Furthermore this section should clearly describe all the important cellular players within the biological system, how they are related functionally with respect to each other, and how they individually or collectively impact the investigated function. The case should include several specific questions (and answers) about this section that assess the audience's understanding of the functional information related to the case.

- **Native structure vs altered structure:** There should already be a minimal amount of available structural information in the protein data bank about this protein and its altered form relevant to the functional change investigated. Note that sequence is the primary structure of a protein. Case discussions should include sequence and structure explorations.

  There should be a dedicated section in the case that delivers adequate background information as well as specific structural information at a molecular level about the central protein investigated. Furthermore this section should clearly describe the structural relationship between the central protein and its interacting partners within the investigated biological system. The case should include several specific questions (and answers) about this section that assess the audience's understanding of the structural information related to the case.

Consider using literature or online bioinformatics tool to investigate:

- What biological phenomenon is affiliated with your protein under normal and disease contexts?
- How/why was your protein named as it is?
- What are its substrates, ligands, reaction mechanisms, or binding partners?
- What organism does it come from? How conserved is the protein across species?
- What are the physical stats of the protein (e.g. MW, pI, amino acid length)?
- How is your protein regulated in the cell? What are the post-translational modifications on the protein (if any), and what role do they play in its function?

Create a key. Even if multiple authors are writing different portions of the case, they need to collaboratively create a single comprehensive key.

- Where possible provide a rubric for grading the responses. Use the following as a guideline in developing the key.
  - What is the main idea in the answer? Is it clear and correct?
  - What is the evidence for the answer? Is there a molecular structural figure to support the claims made.
  - Is the language used in the answer grammatically correct, understandable (without too much technicolor)

d. **Write one additional assessment question (and the associated answer key)** that relates to the case study but requires a student to apply knowledge gained from working through the molecular case study to a slightly different scenario that requires molecular visualization to answer the question. This will be used as a post-test to measure student learning of both the concepts and skills learned in the case.

- **Considering the bigger picture:** Write a quesΩstion/discussion prompt as a suggestion for describing the case's relationship to issues of diversity/inclusion/equity, science and society, and/or social justice.
- **Format References:** The literature cited and listed at the end of the case should be formatted e.g., using the APA (7th edition) style.
  Example:
  In-text: (Coleman, 2001)
  In the References section (Bibliography): Coleman, J. (2001). Nitric oxide in immunity and inflammation. *International Immunopharmacology*, *1*(8), 1397-1406. https://doi.org/10.1016/s1567-5769(01)00086-8
  Resource: Cite this for me for formatting (https://www.citethisforme.com/)
- Any specific data or images should also be cited using the same style.
- Note: a rigorously researched case study will have 5-10 citations.

4. Review and Refine (the case study):

- Consult Molecular CaseNet Rubric and Checklist and revise as needed
- Find a peer (collaborator or colleague) to review the case you have written and ask them to provide feedback (see Peer Review Criteria list)
- Make sure that you address all issues in the feedback and consult the Molecular CaseNet case review rubrics before submitting the case for publication.

Summer Professional Development Workshop 2021

Peer Review Criteria

Molecular Case Study being reviewed
Title:

Author(s):

The Basics

1. Does the title succinctly convey the main focus of the case study without giving too much away?
2. Are there any major issues related to the formatting or structure of the case study? List them and provide suggestions to improve them.
3. List any errors in grammar, spelling that you notice.
4. Comment on the quality of the stated learning objectives.
    a. Are they stated clearly and succinctly?
    b. Do they represent a variety of levels in Bloom's Taxonomy?
    c. How well do the assigned questions align with the stated learning objectives?
5. Does the storyline serve first and foremost to support the experimental question? Is sufficient introduction provided to help you understand the experimental question?
6. Look at all of the headings. Are they specific, descriptive, succinct, and logical? Provide examples here for improvement.
7. Look up a paper in the journal *Cell*, and look specifically at the citation style at the end of it. Now look at the paper. Are there issues of formatting (indentations, spacing), capitalization, or other inconsistencies of style, either in the citation section or in the in-text citations?

The Writing

1. Does this MCS sustain a single coherent point of view throughout? Are there ways to improve this?
2. How could the readability, clarity, or style of this MCS be improved? Comment briefly here on themes you noticed, but also provide extensive editorial notes on the paper directly as needed.
3. Provide one example of a section where the author's writing is far too "fluffy". Quote the section here, and on the paper, make extensive marks on how to condense or cut the writing into a more technical form without losing the meaning or content.
4. While the case needs to be accessible to a broad range of audiences, a very casual or colloquial language may make it difficult for users from different disciplinary and/or cultural backgrounds to read and comprehend. Identify all of the following: colloquialisms, informality, clichés, redundancy, and wordiness. Words to avoid include reference to "research," "researchers," "scientist," "studies have shown..." etc. List examples of these problems that you noticed here.

The Scientific Content

1. What is the strength of this MCS's experimental question and hypotheses explored? Response to this question can provide ideas for promoting the case (and including a line/phrase in the abstract, keywords etc.)
2. How does the case explore both the structure and function of the wild type protein and altered function (mutation, ligand binding, etc.) Consider language, originality/creativity, scientific legitimacy/logic, justification, and placement within the storyline.
3. Does each part of the case study progress logically or are there transitional leaps that are confusing?

4. Does the case study integrate data and evidence for claims made from the primary literature and/or bioinformatics data resources into the storyline?
5. Are there any parts of the case that presume prior knowledge about concepts or skills. For example,
    a. logic of experimental approaches described in key case-related papers
    b. choice of PDB entry for exploration
    c. what to look for in the PDB structure(s) and related bioinformatic resources.

    Please include suggestions for what might be done to correct this?

Unique Elements of the Molecular Case Study

1. What is the quality, use, creativity, and complexity of the figure(s) and legend(s) included in the case. Is there any place that would have benefitted from an additional figure or table?
2. Try to answer the questions provided in the case study.
    a. Do you agree with the authors' key and is it complete?
    b. Are the questions clearly meeting the objectives stated, or do they feel like random questions?
    c. Do the questions being asked lead you forward in the storyline, or are they just latent observations?
3. Are the experimental approaches and bioinformatics explorations appropriately used? List sections in the case where the author should better describe the experimental approach taken.
4. Review the molecular visualization part of the case. Were you able to obtain images of the structure as you believe the authors intended? Were the instructions clear? Explain.
5. Who do you see as the target audience for this molecular case study? Could it be completed by first-semester biology students? 300-level biochemists? Advanced students? Explain.
6. Does the additional assessment assignment logically connect to the initial MCS and assess skills a student should be learning in the course of completing the MCS? Explain.