OPEN ACCESS Freely available online

CourseSource

Lesson

Using Bioinformatics and Molecular Visualization to Develop Student Hypotheses in a Malate Dehydrogenase Oriented CURE

Kevin P. Callahan[1*], Tamara Mans[2*], Jing Zhang[3], Ellis Bell[4], and Jessica Bell[4]**

[1]St John Fisher College

[2]North Hennepin Community College

[3]University of Nebraska, Lincoln

[4]University of San Diego

*co-first authors

Abstract

Developing student creativity and ability to develop a testable hypothesis represents a significant challenge in most laboratory courses. This lesson demonstrates how students use facets of molecular evolution and bioinformatics approaches involving protein sequence alignments (Clustal Omega, Uniprot) and 3D structure visualization (Pymol, JMol, Chimera), along with an analysis of pertinent background literature, to construct a novel hypothesis and develop a research proposal to explore their hypothesis. We have used this approach in a variety of institutional contexts (community college, research intensive university and primarily undergraduate institutions, PUIs) as the first component in a protein-centric course-embedded undergraduate research experience (CURE) sequence. Built around the enzyme malate dehydrogenase, the sequence illustrates a variety of foundational concepts from the learning framework for Biochemistry and Molecular Biology. The lesson has three specific learning goals: i) find, use and present relevant primary literature, protein sequences, structures, and analyses resulting from the use of bioinformatics tools, ii) understand the various roles that non-covalent interactions may play in the structure and function of an enzyme. and iii) create/develop a testable and falsifiable hypothesis and propose appropriate experiments to interrogate the hypothesis. For each learning goal, we have developed specific assessment rubrics. Depending on the needs of the course, this approach builds to an in-class student presentation and/or a written research proposal. The module can be extended over several lecture and lab periods. Furthermore, the module lends itself to additional assessments including oral presentation, research proposal writing and the validated pre-post Experimental Design Ability Test (EDAT). Although presented in the context of course-based research on malate dehydrogenase, the approach and materials presented are readily adaptable to any protein of interest.

Citation: Callahan KP, Mans T, Zhang J, Bell E, Bell J. 2021. Using bioinformatics and molecular visualization to develop student hypotheses in a malate dehydrogenase oriented CURE. *CourseSource*. https://doi.org/10.24918/cs.2021.43

Editor: Srebrenka Robic, Agnes Scott College

Received: 10/12/2020; **Accepted:** 6/29/2021; **Published:** 1/14/2022

Conflict of Interest and Funding Statement: This work was supported by a grant from the National Science Foundation. NSF-1726932 EHR-IUSE, Principal Investigator Ellis Bell, Co Investigators Jessica Bell and Joseph Provost. None of the authors has a financial, personal, or professional conflict of interest related to this work.

Supporting Materials: Supporting Files S1. Using Bioinformatics – Specific Learning Objectives with Associated Rubrics; S2. Using Bioinformatics – PUI capstone cCURE materials; S3. Using Bioinformatics – PUI junior/senior cCURE materials; S4. Using Bioinformatics – R1 junior/senior mCURE materials; S5. Using Bioinformatics – CC intro level mCURE materials; S6. Using Bioinformatics – CC intro level cCURE materials; and S7. Using Bioinformatics – Hypothesis Development Module.

****Correspondence to:** Department of Chemistry and Biochemistry, University of San Diego, 5995 Alcala Park, San Diego, 92110. Email: jessicabell@sandiego.edu

Learning Goals

When students have completed this lesson, they should be able to:

1. Find, use and present relevant primary literature, protein sequences and structures, and bioinformatics tools.
2. Understand the various roles that non-covalent interactions may play in the structure and function of an enzyme.
3. Create/develop a testable and falsifiable hypothesis and propose appropriate experiments to interrogate the hypothesis.

Our three major learning goals encompass several of the society learning goals (italicized) of both the Bioinformatics and the Biochemistry and Molecular Biology frameworks, and associated learning objectives (indented) as detailed below.

Bioinformatics:

What is the role of computation in hypothesis driven discovery processes within the life sciences?

- Explain the necessity for computation in life sciences research. Explain the role of wet-bench techniques in verifying computational results in life science research. Compare and contrast computer-based research with wet-lab research. Read a scientific article and evaluate how bioinformatics methods were employed by the authors to explore a particular hypothesis. Given a scientific question, develop a hypothesis and define computational approaches that could be used to explore the hypothesis.

Where are data about the proteome found (e.g., amino acid sequence and structure) and how are they stored and accessed?

- Describe the types of metadata that accompany sequence, structure, and function data to make for useful biological interpretation (e.g., biological source, accession number, UniProt number, journal articles, etc.).

How can bioinformatics tools be employed to examine protein structure and function?

- Query a dataset with a specific protein sequence to learn about potential functions (e.g., Pfam, CDD, SwissProt, UniProt, etc.). View and interpret the structure output from Protein Data Bank (e.g., Cn3D, Jmol, etc.). Propose potential functions for a given protein structure.

Biochemistry and Molecular Biology:

How do enzymes catalyze biological reactions?

- Identify the factors contributing to the activation energy of a reaction. Use kinetic parameters to compare enzymes. Interpret the physical meaning of various kinetic parameters and describe the underlaying assumptions and conditions on which different parameters depend

What factors determine structure?

- Compare and contrast the primary, secondary, tertiary and quaternary structure of proteins. Use bioinformatics and computational approaches to compare primary sequences and identify the impact of conservation and/or evolutionary change on structure and function.

How are structure and function related? What is the role of noncovalent intermolecular interactions?

- Predict the biological and chemical effects of mutation on the affinity of binding and design appropriate experiments to test the predictions.

What is the Scientific Process?

- When presented with an observation, develop a testable and falsifiable hypothesis, identify appropriate experimental observations and controllable variable.

What skills are needed to access, comprehend and communicate science?

- Identify, locate and use the primary literature. Use databases and bioinformatics tools. Use visual and verbal tools to explain concepts and data.

Learning Objectives

For each of our learning goals, we have specific learning objectives with associated rubrics for assessment (Supporting File S1. Using Bioinformatics – Specific Learning Objectives With Associated Rubrics).

As we adjusted the lesson for different levels of students and different amounts of total time spent, instructors worked to meet the three general learning goals in ways that fit their courses. The specific objectives listed below are a comprehensive list, from which instructors could select the ones which seem appropriate to the course they teach.

Learning Goal 1: Understand the various roles that non-covalent interactions may play in the structure and function of an enzyme.

Students will be able to:

- Compare and contrast the physical basis for coulombic (ionic) interactions and hydrophobic interactions.
- Outline the types of non-covalent interactions you would expect to stabilize secondary structure in a protein.
- Outline the types of non-covalent interactions you would expect to be involved in maintaining a functional tertiary structure in a protein.
- Describe types of non-covalent interactions that might be found across a subunit interface and the functions of the interactions.
- Predict types of non-covalent interactions that would be involved in substrate-enzyme binding and compare their relative strengths.
- Describe how non-covalent interactions in a protein-substrate complex might promote catalysis.
- Predict what types of mutations at nearby residues might alter the pKa of a protonatable group on a protein.

Learning Goal 2: Find, use and present relevant primary literature, protein sequences and structures, and analyses obtained using bioinformatics tools.

Students will be able to:

- Find and use appropriate literature to illustrate the generalizable (big picture) aspects of the work.
- Use appropriate literature to document specific background to the enzyme.
- Use appropriate databases such as Uniprot to obtain sequence information.
- Utilize Clustal Omega and interpret the resultant data.
- Use the Protein Data Bank to obtain 3D coordinates for a protein.
- Use Pymol or other visualization tools to illustrate key features of the protein and their hypothesis.
- Use appropriate bioinformatics tools to design primers for mutagenesis.

Learning Goal 3: Create/develop and present a testable and falsifiable hypothesis and propose appropriate experiments to interrogate the hypothesis.

Students will be able to:

- Describe how the work fits into the field/fills a gap in knowledge.
- Clearly state their hypothesis and the requisite background information that led to the hypothesis.
- Clearly indicate the testable and falsifiable predictions the hypothesis makes.
- Briefly outline the types of experiments that will be used to interrogate the hypothesis.

INTRODUCTION

The "Vision & Change in Undergraduate Biology Education: A Call for Action" initiative and its following report (1), building on prior publications over the past 15-20 years (2-5) recommends 1) integration of core concepts and skills throughout the curriculum, 2) a focus on student-centered learning environments, and 3) validated high impact practices, such as research experiences, as integral components of biology education for all students. Research experiences have major effects on persistence in science (6-9) and positive outcomes in conceptual understanding and skills development, essential for effective workforce development (10-16). The Council on Undergraduate Research (CUR) defines undergraduate research as inquiry or investigation conducted by undergraduates that makes original intellectual or creative contributions to the discipline (16). Such work is a high impact practice that provides robust learning for students, increases retention, enhances student learning though mentorship by faculty and develops a deeper critical thinking ability, as well as intellectual independence. Realization that authentic research was important for student development was evident in the Boyer Report (17), where smaller schools that provide research mentoring disproportionately produce more graduate school students than research intensive institutions where resources and time allocated to research training are not as focused on undergraduates.

Studies in course-based research (18-24) show that, independent of the type of institution, authentic research experience enhances knowledge gains. It is also important to recognize that while reform to include course-embedded undergraduate research experiences (CUREs) in many laboratory courses has been impressive, there remain several important unanswered questions. Assessment has primarily focused on motivation and retention of students (all very positive) but assessment tools linked to student learning outcomes are limited. Further, Brownell and Kloser warn, "despite published articles on CUREs, the impact of these CUREs are still in question" (25). There have been a number of publications discussing the key elements of a CURE as developed by Lopatto, Dolan, Kloser and others (26-30). The NSF-funded CURE network (CUREnet) described the key elements of a CURE (19,31). These elements include 1) the use of scientific practices, 2) discovery, 3) broadly relevant or important work, 4) collaboration, and 5) iteration. Lopatto and others describe seven components of authentic research (11,12): novel questions, student-generated questions, development of a hypothesis, experimental design, data collection, data analysis and presentation or publication of the research. Key parameters for such an experience include minimized role of instructor, an unknown scientific outcome, a project of student design where students are responsible for the design and do most of the work.

The Malate Dehydrogenase CUREs Community (MCC) consists of a diverse community of STEM disciplinary faculty members from institutions that vary across a number of dimensions including type (two-year versus four-year), enrollment size, selectivity, and student population (low-income, first generation). Students enrolled in the CURE use facets of bioinformatics, 3D structure visualization and pertinent background literature to construct a novel hypothesis about the role of a specific amino acid in the activity of malate dehydrogenase. Additional information and resources can be found at the Malate Dehydrogenase CUREs Community (MCC).

In the current lesson we have built on previously published work using a CURE based on malate dehydrogenase (32-34) and incorporated a number of foundational concepts and skills (35-37) that are taken from the ASBMB approved Biochemistry and Molecular Biology Learning Framework (38).

In many CURE type activities, the emphasis is on data collection and analysis. Here we place the emphasis on hypothesis development and proposal construction, areas that are frequently difficult to incorporate into a CURE or coursework in a meaningful way (39-42). We have also placed a focus on collaboration and teamwork since both contribute to the high impact nature of the lesson (43-45). Many CUREs use molecular biology techniques and answer other types of questions; here we place emphasis on protein structure-function relationships, and in particular the role that non-covalent interactions (a gateway concept in biochemistry) (46) play in enzyme function.

We have chosen to focus on malate dehydrogenase (MDH) because the genetic, species, and organismal diversity of MDH make it an ideal protein to explore foundational concepts of protein structure and function. The interaction between MDH and its substrate allow students to explore hydrogen bonding, steric effects, charge-charge interactions, and the hydrophobic effect. However, the approach, materials, and rubrics that we have developed are broadly applicable to almost any protein of interest and can be easily tailored to a faculty person's given research interests.

Intended Audience

This lesson is intended for a broad audience and has been used with a variety of students ranging from incoming first year students in a primarily undergraduate institution (PUI), first and second year community college students, and junior/senior biochemistry majors in both PUI and research intensive university settings. It has also been used with third and fourth year non-science majors.

Required Learning Time

The lesson has been implemented in a variety of formats using up to a total of eight teaching sessions (for example four lecture periods plus four laboratory sessions). Students have multiple homework assignments associated with the lesson and are expected to spend about 1-2 hours out of class working on the homework assignments and preparing for class; written proposals and presentations will typically require more than 2 hours to prepare.

Prerequisite Student Knowledge

The lesson requires a basic knowledge of the central dogma, an introductory knowledge of the levels of protein structure (primary, secondary, tertiary & quaternary) and function (catalytic mechanism and enzyme kinetics) and an understanding of various types of non-covalent interactions.

Prerequisite Teacher Knowledge

Instructors should have a basic understanding of protein structure and function, basic enzymology, and bioinformatics, as well as fundamental knowledge of evolution and molecular biology.

SCIENTIFIC TEACHING THEMES

Active Learning

The various components of the module were developed with student-centered learning as their focus. Student work is framed by a discussion of "biological literacy" as defined by the Vision and Change Final Report (1), focusing on how science is done in a real-world context and how the practice of science has changed. Students are provided with background reading and assignments for out of class work, allowing them to connect the concepts and learning objectives of the module with previous knowledge and skills from other courses. Many lesson periods include think-pair-share and other student-centered activities. Students are encouraged to discuss their ideas with one another as well as with the faculty involved. Students work in groups to learn background information and new skills, and these groups become research teams as they develop their hypotheses.

Assessment

Many formative assessments were used in this lesson, although these varied as appropriate for the different course levels (from introductory to capstone) and with the amount of time that the students were expected to spend on the lesson. For example, introductory students used an instructional worksheet with explanations of the structure of a scientific paper along with prompts to find specific information from assigned papers. Senior capstone students were asked to search the literature and make a mind map of the information they gleaned from the papers they found. Likewise, though all students used rendering software to examine protein structure, introductory and non-major students were given a very specific set of questions and prompts to answer as they learned how to use the program, while senior students were given more general guidelines of what type of pictures they might wish to generate, to fit the broader range of scientific questions their hypothesis might address. See the Supporting Files for details of five different deployments of the lesson (Supporting Files S2. Using Bioinformatics – PUI capstone cCURE materials, S3. Using Bioinformatics – PUI junior/senior cCURE materials, S4. Using Bioinformatics – R1 junior/senior mCURE materials, S5. Using Bioinformatics – CC intro level mCURE materials, and S6. Using Bioinformatics – CC intro level cCURE materials).

In all variants of the lesson, a summative assessment was given as a culminating assignment, in which students wrote and/or orally presented a hypothesis statement and prediction. In a modular setting (in which a shorter time of the semester was used for the CURE), students were guided to write a paragraph with their hypothesis and explanatory information or assigned to deliver a short oral presentation of their hypothesis and reasoning. In complete CURE (full semester) settings, students wrote formal scientific proposals. Opportunities for peer review, instructor feedback, and revision were given so that students could improve their work.

Inclusive Teaching

We have been pleased to see this lesson successfully incorporated in a variety of institution types and at all levels, including courses for non-majors. We see this flexibility as part of its strength, as instructors can see many different depths and lengths that the lesson can reach. The resources used in the various parts of the lesson are all freely available through the web and students learn to access and assess these materials while creating their own knowledge and understanding, thereby promoting ownership of the hypotheses they develop, all important aspects of inclusive teaching (47). Papers selected for class discussion are chosen to highlight a diversity of ideas and author backgrounds, demonstrating how diversity in its many forms contributes to better problem solving (48). To further illustrate and model a scientific community, the authors of the discussion paper frequently attend (virtually) the hypothesis presentations of the other authors' students, leading to a robust discussion that enhances the students' sense of ownership in their project. In several iterations of the project, students have collaborated across institutions, so that they may brainstorm with each other, and mentor and learn from students at other levels of study. In some cases, project ideas were "handed off:" hypotheses developed by one group of students were tested by other groups, allowing for shared resources and time, as well as ideas.

LESSON PLAN

We have implemented this lesson in a variety of settings, ranging from a PUI upper-level capstone course to a community college lab associated with an introductory course (course descriptions and syllabi are listed in part A of each of the following: Supporting Files S2. Using Bioinformatics – PUI capstone cCURE materials, S3. Using Bioinformatics – PUI junior/senior cCURE materials, S4. Using Bioinformatics – R1 junior/senior mCURE materials, S5. Using Bioinformatics – CC intro level mCURE materials, and S6. Using Bioinformatics – CC intro level cCURE materials.). Our Malate Dehydrogenase CURE Community has differentiated courses in which the complete semester is spent in CURE-related activity as cCUREs (c for complete) and those in which CURE-related activity takes a shorter part of the semester as mCUREs (m for modular). We have used between two and eight class meetings for the lesson. Table 1 concisely lays out the varied settings and time allowances of five different deployments of the lesson. Below, we explain our general approach, with references to the collections of notes, assignments, and rubrics from our varied course formats all shared in the supporting materials.

Foundational Knowledge and Introduction to the Project

The overall organization of the lesson is illustrated in Figure 1 and summarized in Table 2. We introduce or review the foundational knowledge and project plan with slideshow- or whiteboard-supplemented lectures (part B in each of the following Supporting Files: S2. Using Bioinformatics – PUI capstone cCURE materials, S3. Using Bioinformatics – PUI junior/senior cCURE materials, S4. Using Bioinformatics – R1 junior/senior mCURE materials, S5. Using Bioinformatics – CC intro level mCURE materials, and S6. Using Bioinformatics – CC intro level cCURE materials; some of these refer to materials in Supporting File S7. Using Bioinformatics – Hypothesis Development Module.). The detail and depth of the presentations vary, depending on the level and prior knowledge of students. Foundational knowledge for our project includes protein structure and noncovalent interactions, function and action of enzymes, and how the evidence of evolutionary conservation and change may be reflected in biomolecules. These concepts are illustrated when possible with specifics of malate dehydrogenase structure and function, so that students are prepared for future investigation of this particular enzyme.

Teams of 2-4 students are formed. The project plan is laid out, including the basic schedule and expectations, the philosophy behind using a CURE, plans for outside collaboration which will occur during the project, and the rationale and process for various pedagogical tools, such as think-pair-share and mind-mapping.

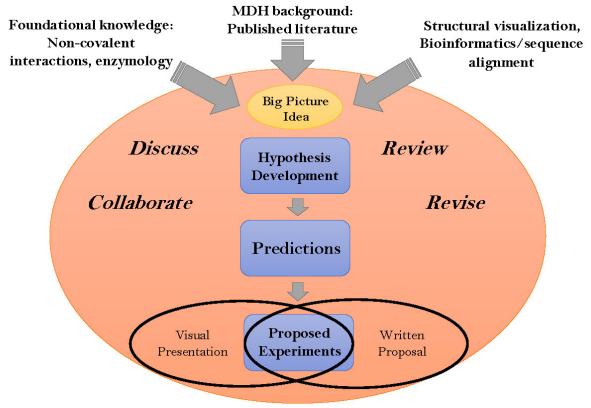

Figure 1. Mind map of the organization and philosophy of the module.

Malate Dehydrogenase Background in Survey of Primary Literature

We may begin with a review quiz or mini lecture to remind students of key points from the foundational knowledge portion of the lesson. As we are preparing students to make unique and authentic contributions to a field that is unfamiliar to them, we teach them that scientists become familiar with a new area of research by reviewing the existing literature in the field. We introduce primary literature, define it, and discuss scientific peer review and publishing norms. We demonstrate how to find scientific literature in general and papers about malate dehydrogenase in particular, and may provide sample papers for the students highlighting approaches or structures of focus during the project. We outline the structure of scientific papers and present approaches for dissecting and analyzing them, through discussion and/or a guided worksheet focusing on a pre-chosen paper. Students are assigned to find and read one to ten papers, and summarize in a form appropriate for their capability. Assignments are in part C in each of the following Supporting Files: S2. Using Bioinformatics – PUI capstone cCURE materials, S3. Using Bioinformatics – PUI junior/senior cCURE materials, S4. Using Bioinformatics – R1 junior/senior mCURE materials, S5. Using Bioinformatics – CC intro level mCURE materials, and S6. Using Bioinformatics – CC intro level cCURE materials.

Bioinformatics, Structural Visualization, and Alignments

We introduce students to databases and bioinformatics tools which can be used to collect information, explore, and develop hypotheses about malate dehydrogenase. These include RCSB Protein Data Bank and UniProt, and molecular rendering and visualization tools such as PyMol, Jmol, and Chimera. We show how multiple sequence alignments (Clustal Omega) can be generated and used to inform our hypotheses about structure and function of the protein. Students are guided via worksheets, pre-recorded videos available on the internet, and instructor coaching when needed. Students complete one or more assignments to summarize data and structures that can inform their hypotheses. The varied programs used in our differing settings are illustrated in Table 3. Assignments are in part D in each of the following Supporting Files: S2. Using Bioinformatics – PUI capstone cCURE materials, S3. Using Bioinformatics – PUI junior/senior cCURE materials, S4. Using Bioinformatics – R1 junior/senior mCURE materials, S5. Using Bioinformatics – CC intro level mCURE materials, and S6. Using Bioinformatics – CC intro level cCURE materials.

Hypothesis Development, Predictions and Experimental Design

In the closing phase of the lesson, the students are guided to develop a hypothesis. Student teams are prompted to synthesize the information they have gathered about malate dehydrogenase from the literature, from the databases, molecular visualizations, and sequence alignments. We remind them that hypotheses make testable predictions; to facilitate the connection to experimentation, at this point or earlier we discuss or practice biochemical assays and consider the sort of data gathered from these. Teams discuss their hypothesis and predictions informally with their partner(s) and instructor, and in some cases, present these more formally before the other teams, guest faculty, or students from collaborating classes. Through the discussions and questions that come out during presentations, the teams refine their hypotheses. After chances for reflection and revision, students are asked to write out their hypothesis with supporting evidence, in appropriate formats such a formal proposal in a capstone course, or a simple paragraph for introductory students participating in an abbreviated version of the project. The varied approaches we used for hypothesis development in different settings are laid out in Table 4. Assignments are in part E in each of the following Supporting Files: S2. Using Bioinformatics – PUI capstone cCURE materials, S3. Using Bioinformatics – PUI junior/senior cCURE materials, S4. Using Bioinformatics – R1 junior/senior mCURE materials, S5. Using Bioinformatics – CC intro level mCURE materials, and S6. Using Bioinformatics – CC intro level cCURE materials.

TEACHING DISCUSSION

The lesson plans presented here were developed as part of the Malate Dehydrogenase CURE Community (MCC). MCC CUREs were instituted with two goals: 1) to provide high impact teaching practices for students, and 2) to create new knowledge on structure function relationships in malate dehydrogenase that can be submitted for publication in peer-reviewed journals.

While all MCC CUREs involve students in each stage of the scientific process (Scientific Background, Hypothesis Development, Proposal, Experiments/Teamwork to test hypothesis, Data Analysis and Conclusions, and Presentation), one of the distinguishing features of MCC CUREs is the emphasis on student-generated hypotheses. As outlined in Figure 1 students use foundational knowledge, primary literature, bioinformatics and structural analysis to develop their hypothesis. Using this information, all student-generated hypotheses should lead to a proposal containing the following elements:

- Big Picture- why is this important to science, to society
- Specific Hypothesis
- Predictions arising from the hypothesis
- Proposed experiments to test the predictions, specifically what types of information do you need to falsify/support the hypothesis

Finally, as outlined in Figure 1, this process should culminate in some type of presentation of the hypothesis and proposed research.

As MCC resources grow, we have available a number of mutants that have validated sequences, and in many cases pre-existing data that has been validated by subsequent experiments after the culmination of the course. These mutants and their validated data can be made available for use in MCC CURE Classes.

We propose that this information be available to students as they develop their own hypotheses in an MCC CURE Class. For example, they will be able to see what mutants have been made previously and any validated data on those mutants. The database of this information does NOT contain any information as to what hypotheses were developed with respect to the mutation, nor are any suggestions as to what hypotheses were being tested given to students.

There are a couple ways this information can be used in a class:

1. Students could be given a selection of these mutants with the validated background information and asked to develop a hypothesis as to what function they think the wildtype residue played in the protein, and what they think the mutation will do to the structure-function relationships of the enzyme and devise and conduct experiments to explore their hypothesis.
2. Students could be given a "big picture" question. For example, what governs folding of malate dehydrogenase? As part of the accessible background information about MDH, the list of validated mutants and associated effects would be provided to the students. They would use this information in much the same way as they would use published information and would develop their own hypothesis. At the instructor's discretion, they could be restricted to using one of the pre-existing mutations.

The goal of this module is to guide students to develop a testable hypothesis about malate dehydrogenase. Two major challenges were encountered by the students taking this course: 1) they only have superficial understanding of malate dehydrogenase and 2) are not familiar with bioinformatic and structural visualization software. To overcome the first challenge, as simple as it may sound, it is necessary to talk about the various aspects of malate dehydrogenase (structure, reaction mechanism, substrate binding, etc.) every single class. It is important that the students feel as though they are becoming experts on MDH. The visualization of the structure of MDH and the specific roles amino acid side chains play in its functionality in conjunction with the evolutionary conservation of these amino acids is critical to the student's ability to formulate their hypothesis. To faculty with limited experience with structural visualization tools, we would direct you to the instruction videos in section D in each of the following Supporting Files: S2. Using Bioinformatics – PUI capstone cCURE materials, S3. Using Bioinformatics – PUI junior/senior cCURE materials, S4. Using Bioinformatics – R1 junior/senior mCURE materials, S5. Using Bioinformatics – CC intro level mCURE materials, and S6. Using Bioinformatics – CC intro level cCURE materials.

Finally, one aspect of the lesson that we have all found extremely effective is the impact of sitting in on one another's student presentations of their hypothesis. We have used Skype or Zoom to virtually sit in on presentations, and students are energized by having someone other than their own instructor involved. This aspect has in some cases led to student collaborations between institutions. It seems clear that student ownership of their project is enhanced by these interactions. Through the Malate Dehydrogenase CUREs Community we are in the process of assessing this aspect in detail as part of our current IUSE grant (49).

SUPPORTING MATERIALS

- Supporting File S1. Using Bioinformatics – Specific Learning Objectives with Associated Rubrics
- Supporting File S2. Using Bioinformatics – PUI capstone cCURE materials
- Supporting File S3. Using Bioinformatics – PUI junior/senior cCURE materials
- Supporting File S4. Using Bioinformatics – R1 junior/senior mCURE materials
- Supporting File S5. Using Bioinformatics – CC intro level mCURE materials
- Supporting File S6. Using Bioinformatics – CC intro level cCURE materials
- Supporting File S7. Using Bioinformatics – Hypothesis Development Module

ACKNOWLEDGMENTS

The materials presented here have in large part been developed as part of an NSF funded IUSE grant: NSF-1726932 EHR-IUSE "MDH CUREs Community, a Protein Centered Approach to CUREs," $598,666, 9/1/2017 to 8/20/2020, Principal Investigator Ellis Bell, Co Investigators Jessica Bell and Joseph Provost.

REFERENCES

1. Vision and Change in Undergraduate Biology Education: A View for the 21st Century,. 2009, AAAS, https://live-visionandchange.pantheonsite.io/wp-content/uploads/2011/03/Revised-Vision-and-Change-Final-Report.pdf, accessed October 20, 2020.
2. Benvenuto M. 2002. Educational reform: why the academy doesn't change. Thought and Action. 18:63-74
3. Bell E. 2001. The future of education in the molecular life sciences. Nat Rev Mol Cell Biol. 2:221-5. doi: 10.1038/35056610.
4. National Research Council. 2003. BIO2010: Transforming Undergraduate Education for Future Research Biologists. Washington, DC: The National Academies Press. doi: 10.17226/10497
5. Wood WB. 2003. Inquiry-based undergraduate teaching in the life sciences at large research universities: a perspective on the Boyer Commission Report. Cell Biol Educ. 2:112-6
6. Graham MJ, Frederick J, Byars-Winston A, Hunter AB, Handelsman J. 2013. Science education. Increasing persistence of college students in STEM. Science. 347:1455-6. doi: 10.1126/science.1240487.
7. Russell SH, Hancock MP, McCullough J. 2007. The pipeline. Benefits of undergraduate research experiences. Science. 316:548-9. doi: 10.1126/science.1140384.
8. Seymour, E., Hunter, A.□B., Laursen, S.L. and DeAntoni, T. 2004. Establishing the benefits of research experiences for undergraduates in the sciences: First findings from a three□year study. Sci. Ed. 88:493-534. doi:10.1002/sce.10131
9. Eagan MK, Hurtado S, Chang MJ, Garcia GA, Herrera FA, Garibay JC. 2013. Making a Difference in Science Education: The Impact of Undergraduate Research Programs. Am Educ Res J. 50:683-713. doi:10.3102/0002831213482038
10. Crowe M, Brakke D. 2008. Assessing the impact of undergraduate research experiences on students: an overview of current literature. CUR Quart. 28:43-50

11. Lopatto, D. 2006. Undergraduate research as a catalyst for liberal learning. Peer Rev. 8:22-25.
12. Lopatto, D. 2009. Science in Solution: The Impact of Undergraduate Research on Student Learning. Research Corporation for Science Advancement: Tucson Arizona.
13. Luckie DB, Maleszewski JJ, Loznak SD, Krha, M. 2004. Infusion of collaborative inquiry throughout a biology curriculum increases student learning: a four-year study of "Teams and Streams" Advances in Physiology Education. 28:199-209. doi.org/10.1152/advan.00025.2004
14. Myers MJ, Burgess BB. 2003. Inquiry-based laboratory course improves students' ability to design experiments and interpret data. Advances in Physiology Education. 27:26-33. doi: 10.1152/advan.00028.2002
15. Rissing S. Cogan J. 2009. Can an inquiry approach improve college student learning in an introductory laboratory? CBE life sciences education. 8:55-61. doi: 10.1187/cbe.08-05-0023.
16. Hansel N. (Ed). 2012. Characteristics of Excellence in Undergraduate Research (COEUR). Council on Undergraduate Research (CUR), ISBN: 978-0-941933-049-0
17. Reinventing Undergraduate Education: A Blueprint for America's Research Universities. Boyer Commission on Educating Undergraduates in the Research University State University of New York at Stony Brook for the Carnegie Foundation for the Advancement of Teaching: Stony Brook, NY, 1998.
18. Lopatto D, Hauser C, Jones CJ, Paetkau D, Chandrasekaran V, Dunbar D, MacKinnon C, Stamm J, Alvarez C, Barnard D, Bedard JE, Bednarski AE, Bhalla S, Braverman JM, Burg M, Chung HM, DeJong RJ, DiAngelo JR, Du C, Eckdahl TT, Emerson J, Frary A, Frohlich D, Goodman AL, Gosser Y, Govind S, Haberman A, Hark AT, Hoogewerf A, Johnson D, Kadlec L, Kaehler M, Key SC, Kokan NP, Kopp OR, Kuleck GA, Lopilato J, Martinez-Cruzado JC, McNeil G, Mel S, Nagengast A, Overvoorde PJ, Parrish S, Preuss ML, Reed LD, Regisford EG, Revie D, Robic S, Roecklien-Canfield JA, Rosenwald AG, Rubin MR, Saville K, Schroeder S, Sharif KA, Shaw M, Skuse G, Smith CD, Smith M, Smith ST, Spana EP, Spratt M, Sreenivasan A, Thompson JS, Wawersik M, Wolyniak MJ, Youngblom J, Zhou L, Buhler J, Mardis E, Leung W, Shaffer CD, Threlfall J, Elgin SC. 2014. A central support system can facilitate implementation and sustainability of a Classroom-based Undergraduate Research Experience (CURE) in Genomics. CBE Life Sci Educ. 13:711-23. doi: 10.1187/cbe.13-10-0200.
19. Auchincloss LC, Laursen SL, Branchaw JL, Eagan K, Graham M, Hanauer DI, Lawrie G, McLinn CM, Pelaez N, Rowland S, Towns M, Trautmann NM, Varma-Nelson P, Weston TJ, Dolan EL. 2014. Assessment of course-based undergraduate research experiences: a meeting report. CBE Life Sci Educ. 13:29-40. doi: 10.1187/cbe.14-01-0004.
20. Corwin LA, Graham MJ, Dolan EL.2015. Modeling course-based undergraduate research experiences: an agenda for future research and evaluation. CBE Life Sci Educ. 14:es1. doi:10.1187/cbe.14-10-0167
21. Denofrio LA, Russell B, Lopatto D, Lu Y. 2007.Mentoring. Linking student interests to science curricula. Science. 318:1872-3. doi: 10.1126/science.1150788.
22. Lopatto D. 2007 Undergraduate research experiences support science career decisions and active learning. CBE Life Sci Educ. 6:297-306. doi: 10.1187/cbe.07-06-0039.
23. Dolan EL, Collins JP. 2015. We must teach more effectively: here are four ways to get started. Mol Biol Cell. 26:2151-5. doi: 10.1091/mbc.E13-11-0675.
24. Lopatto D. 2004. Survey of Undergraduate Research Experiences (SURE): first findings. Cell Biol Educ. 3:270-7. doi: 10.1187/cbe.04-07-0045.
25. Kloser MJ, Brownell SE, Chiariello NR, Fukami T. 2011. Integrating teaching and research in undergraduate biology laboratory education. PLoS Biol. 9:e1001174. doi: 10.1371/journal.pbio.1001174.
26. Lopatto D, Alvarez C, Barnard D, Chandrasekaran C, Chung HM, Du C, Eckdahl T, Goodman AL, Hauser C, Jones CJ, Kopp OR, Kuleck GA, McNeil G, Morris R, Myka JL, Nagengast A, Overvoorde PJ, Poet JL, Reed K, Regisford G, Revie D, Rosenwald A, Saville K, Shaw M, Skuse GR, Smith C, Smith M, Spratt M, Stamm J, Thompson JS, Wilson BA, Witkowski C, Youngblom J, Leung W, Shaffer CD, Buhler J, Mardis E, Elgin SC. 2008. Undergraduate research. Genomics Education Partnership. Science. 322:684-5. doi: 10.1126/science.1165351.
27. Lopatto, D. 2016 CURE Survey. Available from: http://www.grinnell.edu/academics/areas/psychology/assessments/cure-survey.
28. Lopatto D. 2003. The essential features of undergraduate research. CUR Quart. 24:139-142.
29. Hanauer DI, Dolan EL. 2014. The project ownership survey: measuring differences in scientific inquiry experiences. CBE Life Sci Educ. 13:149-58. doi: 10.1187/cbe.13-06-0123.
30. Rahm, J., Miller, H.C., Hartley, L. and Moore, J.C. 2003. The value of an emergent notion of authenticity: Examples from two student/teacher–scientist partnership programs. J. Res. Sci. Teach. 40:737-756. doi:10.1002/tea.10109
31. CUREnet, https://serc.carleton.edu/curenet/index.html, accessed October 20, 2020
32. Bell E. 2011. Using research to teach an "introduction to biological thinking". Biochem Mol Biol Educ. 39:10-6. doi: 10.1002/bmb.20441.
33. Bell JK, Eckdahl TT, Hecht DA, Killion PJ, Latzer J, Mans TL, Provost JJ, Rakus JF, Siebrasse EA, Ellis Bell J. 2017. CUREs in biochemistry-where we are and where we should go. Biochem Mol Biol Educ. 45:7-12. doi: 10.1002/bmb.20989.
34. The Malate Dehydrogenase Cures Community. https://mdh-cures-community.squarespace.com (accessed October 20, 2020).
35. Tansey JT, Baird T Jr, Cox MM, Fox KM, Knight J, Sears D, Bell E. 2013. Foundational concepts and underlying theories for majors in "biochemistry and molecular biology". Biochem Mol Biol Educ. 41:289-96. doi: 10.1002/bmb.20727
36. White HB, Benore MA, Sumter TF, Caldwell BD, Bell E. 2013. What skills should students of undergraduate biochemistry and molecular biology programs have upon graduation? Biochem Mol Biol Educ. 41:297-301. doi: 10.1002/bmb.20729.
37. Wright A, Provost J, Roecklein-Canfield JA, Bell E.2013. Essential concepts and underlying theories from physics, chemistry, and mathematics for "biochemistry and molecular biology" majors. Biochem Mol Biol Educ. 41:302-8. doi: 10.1002/bmb.20728.
38. CourseSource Biochemistry and Molecular Biology Learning Framework. https://www.coursesource.org/courses/biochemistry-and-molecular-biology, (accessed October 20, 2020)
39. Hadzigeorgiou, Y, Fokialis P, Kabouropoulou, M. 2012. Thinking about Creativity in Science Education.. Creative Education, 3:603-611. doi: 10.4236/ce.2012.35089.
40. Strode PK. 2015. Hypothesis Generation in Biology: A Science Teaching Challenge & Potential Solution. The American Biology Teacher, 77:500-506.
41. Teacher Guide. 2017. Using Science News to Develop a Research Question and Hypothesis Science News Educator Guide. https://www.sciencenews.org/wp-content/uploads/2015/09/SNHS_guide_activity_research_in_brief_04_29_17.pdf, (accessed October 20th 2020)
42. Irby SM, Pelaez NJ, Anderson TR. 2018. How to Identify the Research Abilities That Instructors Anticipate Students Will Develop in a Biochemistry Course-Based Undergraduate Research Experience (CURE). CBE Life Sci Educ.17:es4. doi:10.1187/cbe.17-12-
43. Ballen CJ, Wieman C, Salehi S, Searle JB, Zamudio KR. 2017. Enhancing Diversity in Undergraduate Science: Self-Efficacy Drives Performance Gains with Active Learning. CBE Life Sci Educ. 16:ar56. doi: 10.1187/cbe.16-12-0344.
44. Premo J, Cavagnetto A, Davis WB, Brickman P.2018. Promoting Collaborative Classrooms: The Impacts of Interdependent Cooperative Learning on Undergraduate Interactions and Achievement. CBE Life Sci Educ. 17:ar32. doi: 10.1187/cbe.17-08-0176.
45. Stone EM. 2014. Guiding students to develop an understanding of scientific inquiry: a science skills approach to instruction and assessment. CBE Life Sci Educ. 13:90-101. doi: 10.1187/cbe-12-11-0198.
46. Loertscher J, Lewis JE, Mercer AM, Minderhout V. 2018. Development and use of a construct map framework to support teaching and assessment of noncovalent interactions in a biochemical context. Chem. Educ. Res. Pract., 19:1151-65
47. Tanner KD. 2013. Structure matters: twenty-one teaching strategies to promote student engagement and cultivate classroom equity. CBE Life Sci Educ. 12:322-31. doi: 10.1187/cbe.13-06-0115.
48. Phillips KW. 2014. How Diversity Makes Us Smarter. Scientific American. https://www.scientificamerican.com/article/how-diversity-makes-us-smarter/, (accessed October 20, 2020)
49. NSF-1726932 EHR-IUSE

Table 1. Course schedule and structure of the course across all courses and institution types. Please note cCUREs (c for complete) denotes courses in which the complete semester is spent in CURE-related activity; mCURE (m for modular) denotes courses in which CURE-related activity accounts for only a portion of the semester.

Institution type	Student population	Lab sessions/ week	Duration of each lab	Total sections/ semester	Students/ section	CURE type	Total lab sessions on hypothesis development
PUI Capstone	Juniors and Seniors	2	4h	2-3	8-10	cCURE	4
PUI	Juniors and Seniors	1	4h	2-3	10	cCURE	4
R1	Juniors and Seniors	1	4h	2	24	mCURE	2.5
CC	First Year	1	3h	1	12-24	mCURE	2
CC		1	3h	1	12-24	cCURE	5

Table 2. Lesson Plan Timeline across all courses and institution types.

Activity	Description	Estimated Time	Notes
Foundational Knowledge and Introduction to the Project			
Lecture and guided discussion	1. Review protein structure and noncovalent interactions. 2. Introduce or review function and action of enzymes. 3. Discuss how biomolecules reflect evidence of evolutionary conservation and change . 4. Form teams of 2-4 students. 5. Preview project schedule and expectations.	75 minutes for PUI capstone. Somewhat shorter for other settings, approximately 30 minutes.	• Detail required depends on your students' background and your expectations for depth and mastery. • Materials are found in part B of each of the following: S2. PUI capstone cCURE materials, S3. PUI junior/senior cCURE materials, S4. R1 junior/senior mCURE materials, S5. CC intro level mCURE materials, and S6. CC intro level cCURE materials; some refer to materials in S7. Hypothesis Development Module.
Malate Dehydrogenase Background in Survey of Primary Literature			
Mini-lecture followed by activity (think pair share and mind-mapping) or guiding worksheet	1. Remind students of the malate dehydrogenase enzyme that will be the focus of their hypothesis. 2. Discuss how scientists prepare for new areas of study by reviewing the literature. 3. Define primary literature, describe peer review process. 4. Demonstrate how to search PubMed, Google Scholar, etc. to find papers. 5. Guide students through analysis of paper(s) by activity or worksheet.	~2-3 hours	• Materials are found in part C of each of the following: S2. PUI capstone cCURE materials, S3. PUI junior/senior cCURE materials, S4. R1 junior/senior mCURE materials, S5. CC intro level mCURE materials, and S6. CC intro level cCURE materials.
Bioinformatics, Structural Visualization, and Alignments			
Mini-lecture followed by computer-based assignments	1. Introduce databases with protein information (RCSB Protein Data Bank or UniProt) and direct students to download sequence for Malate Dehydrogenase. 2. Open and briefly demonstrate structural visualization tools. 3. Students work through a set of exercises guided by video tutorials and instructor coaching. 4. Introduce multiple sequence alignment and interpretation. 5. Students work through an alignment exercise.	3 hours in class for structural visualization, plus some student teams need additional time to complete assignments. For some courses, additionally up to 4 hours may be used to construct a multiple sequence alignment. Less time (1.5 hours) is required if alignment is provided.	• Students or teams of students will need computers for this work. • Some instructors opt to not include alignment interpretation if time or course topics are limited. • Materials are found in part D of each of the following: S2. PUI capstone cCURE materials, S3. PUI junior/senior cCURE materials, S4. R1 junior/senior mCURE materials, S5. CC intro level mCURE materials, and S6. CC intro level cCURE materials.
Hypothesis Development, Predictions and Experimental Design			
Students informally discuss hypotheses and predictions with team and instructor. As time allows, students make a formal presentation of their hypothesis. Students write out their hypothesis and predictions and support them with evidence.	1. Discuss or practice experimental methods (this has been done in various ways ranging from lectures to actual wet lab work). 2. Prompt students to formulate hypotheses and predictions. Allow time for discussion in teams and with instructor. 3. In some cases, students make a more formal oral presentation. 4. Students are assigned to write out their hypothesis and predictions with evidence in a format appropriate to their level.	Up to 3 hours may be used for overview or practice of experimental methods. Informal and formal discussions of hypotheses take variable time, depending on the number of student groups, at least 15 minutes per team.	• Materials are found in part E of each of the following: S2. PUI capstone cCURE materials, S3. PUI junior/senior cCURE materials, S4. R1 junior/senior mCURE materials, S5. CC intro level mCURE materials, and S6. CC intro level cCURE materials.

Table 3. Comparison of bioinformatic and molecular visualization components across all courses and institution types.

	Bioinformatics, Structural Visualizations, and Alignments				
	PUI capstone cCURE	**PUI cCURE**	**R1 mCURE**	**CC cCURE**	**CC mCURE**
Clustal	☑				
Uniprot		☑			
Alignment provided				☑	☑
PYMOL			☑	☑	☑
JMOL	☑	☑			

Table 4. Comparison of hypothesis development, prediction and experimental design components across all courses and institution types.

	Hypothesis, Predictions and Experimental Design				
	PUI capstone cCURE	**PUI cCURE**	**R1 mCURE**	**CC cCURE**	**CC mCURE**
Mutants provided			☑	☑	☑
Students generate mutants	☑	☑			
Written proposal	☑	☑	☑	☑	
Oral presentation	☑	☑	☑	☑	