Chapter 8

Online Tools for Teaching Large Laboratory Courses: How the GENI Website Facilitates Authentic Research

Benjamin J. McFarland*

Department of Chemistry and Biochemistry, Seattle Pacific University, Seattle, Washington 98119, United States
***E-mail: bjm@spu.edu**

One educational problem that technology can address is how to facilitate authentic research experiences in the classroom for large numbers of students. The collaborative online platform GENI (Guiding Education through Novel Investigation, found at geni-science.org) provides an application that transfers protocols among institutions and into undergraduate teaching laboratories, then collects the data from students for analysis and publication. I have used this tool for several years to conduct bioinformatics research with three or four 16-student lab sections in Biochemistry I and to prepare recombinant proteins with two or three 16-student lab sections in Biochemistry II. These projects also feed into a Survey of Physical Chemistry course. Here I present technical details of how research projects have transitioned into the teaching lab. Others in the GENI consortium have accomplished similar projects in molecular biology and genetics contexts. Education researchers on our team are developing and applying assessments to shape our use of GENI as a tool. Altogether, these show that online lab protocols and data collection can be an effective way to teach students through the creative and exploratory process of asking and answering research questions.

Introduction

Large courses with many lab sections pose particular challenges for the chemistry teacher. One acute challenge is the tension between the value of individual projects in learning science and the limitations of resources in time and funds for accomplishing those projects in the context of a large class composed of students with varied backgrounds and motivations. These authentic research projects, also called Course-Based Undergraduate Research Experiences (CUREs), have been integrated into many different chemistry courses, ranging from general chemistry to analytical chemistry (*1–3*). CUREs require organized management of information in multiple directions, primarily in providing protocols to the students before an experiment and in collecting data from students after an experiment. The general benefits to providing structured relationships to the undergraduate research experience have been noted (*4*). This suggests that a course of study that more closely mimics an authentic research experience will teach students more effectively than a standard set of labs in which the instructor already knows the expected outcomes. Here I describe my use of an online tool that facilitates research collaboration in a central location that can be easily transferred and modified.

The adjectives "large" (as in "large courses") and "authentic" (as in "authentic research") are relative terms that may vary based on the individual institution and instructor. Here "large" is defined relative to student expectations and institutional resources such as teaching load. In this chapter, the tool described is most useful for any situation in which a project is carried out in *multiple lab sections*, whether they are multiple lab sections offered in the same academic term at one institution, multiple sections offered at multiple institutions, or the same project carried out over multiple academic terms. In all of these cases, the online tool described here can provide standardization of input and output that helps to scale up the project and disseminate protocols and data among multiple sections of students.

Many different institutions are implementing authentic research in situations that fall under this definition of large courses. One of the largest collaborations is the Genomics Education Partnership, a consortium of more than 100 institutions that investigated evolution of the Muller F element in undergraduate laboratories, and which developed a central support system of resources tailored to that particular research project (*5*). The tool described in this chapter was developed to serve as similar, more flexible online resource that could be readily adapted by the individual instructor to match individual needs for conducting diverse projects across multiple sections. Recent examples of CUREs conducting authentic research with multiple lab sections either within or among courses include a systems biology project described as "large-scale" at The University of Queensland with yearly cohorts of 33 and 47 students (*6*); a project studying zebrafish development in two undergraduate biology laboratory courses at Indiana University-Purdue University (*7*); and a medicinal chemistry program conducted as an undergraduate capstone experience at multiple small universities (*8*). For the first two projects, the same project is carried out over multiple years, and for

the third, similar projects are carried out at multiple institutions. These are the types of projects that this tool can best facilitate.

The precise meaning of the term "authentic" has been discussed since at least the turn of the 21st century (*9*). Recently, an editorial in *Biochemistry and Molecular Biology Education* (BAMBED) described the current diversity of projects considered to be authentic research in an undergraduate biochemical context (*10*). I will adopt a similar definition of "authentic" here – one that emphasizes that authentic research is both collaborative and publishable in peer-reviewed journals. The editorial's authors note the need for a tool that would promote authentic research collaboration: "In BAMBED and other educational journals, examples of single investigator/institution integration of research into the classroom exist, but these examples function in isolation, lacking the collaboration that promotes long-term authentic research evolution and team-based skills. Again, partnerships between institutions can begin to develop these resources, but an organized community effort would be better positioned to provide the variety and depth of projects necessary to sustain CUREs in the biochemical curriculum" (*10*).

Here I describe such a tool. This tool was initially developed by a consortium and has been used at multiple academic institutions to communicate genomic data. I have used this tool to transfer information among multiple sections in large Biochemistry classes, and among multiple years of the same Physical Chemistry class. The protocols can be passed from institution to institution through HTML-coded webpages and the data is collected in one place for students and instructors to access from any web browser. This tool has been used in multiple contexts and has been adapted to many different types of research projects, so that researchers who teach can adapt their own projects to the classroom and bring an authentic research experience to many students. I have used this tool to bring authentic research protocols from my post-doctoral research used for protein production into the undergraduate biochemistry laboratory since 2012. Its utility has been demonstrated among multiple sections, multiple classes, and multiple projects.

The online tool is titled Guiding Education through Novel Investigation (GENI) and can be found at the URL http://geni-science.org. It was originally developed by a group of biologists from multiple institutions to provide protocols for genome annotation to large groups of students, and was then expanded to include protocols for "wet lab" procedures in molecular biology and genetics courses. GENI's flexible nature allowed me to transfer protein chemistry protocols from my own research projects online for use in three courses: Biochemistry I (four lab sections of up to 16 students each), Biochemistry II (two or three lab sections of 16 students), and Survey of Physical Chemistry (one lab section of 4-8 students). In each of these courses, GENI provided all students with an interactive online framework for carrying out protocols to accomplish novel research projects in the lab, providing new and potentially publishable results in the context of an undergraduate course. Other disciplines inside and outside of chemistry can adapt projects for use on the GENI platform.

The primary purpose of GENI is to give protocols and collect data in the teaching lab, and because it is located online, it can facilitate other types of authentic-research knowledge transfers as well:

1.) The same project at different institutions: The biologists in the GENI consortium have used it to coordinate genome annotation projects at multiple institutions across the U.S.

2.) The same proteins in different courses taught by different professors: At my institution, I have coordinated a bioinformatics GENI project in Biochemistry I with another professor's GENI project in Molecular Biology, so that students can use findings from their homology models obtained in Biochemistry I to plan experiments in Molecular Biology.

3.) The same protein in different courses taught by the same professor: The proteins students purify in Biochemistry II in the winter quarter have been analyzed for binding thermodynamics and kinetics in Survey of Physical Chemistry in the spring quarter.

Because the data collected in these projects are novel, they may ultimately be transferred to the scientific community at large in the form of a peer-reviewed publication. In this way, students contribute to the edifice of scientific knowledge as they learn how to complete protocols, collect data, and make new substances. Through publication, the findings of the new knowledge can be used by the international scientific community.

Using the GENI Website

Once an instructor account is created on the GENI website, a list of available projects becomes accessible, including projects in biochemistry, ecology, functional genomics, and genetics (Figure 1). Many of the tools for the multi-institutional genome annotation projects are collected under the "ACT" heading, being associated with the Integrated Microbial Genomes Annotation Collaboration Toolkit (IMG-ACT) (*11*). Other projects are located under the "Available Projects" heading with brief descriptions. The diversity of available projects demonstrates the broad applicability of the GENI platform.

After an instructor creates a project on GENI and assigns it to a specific class, the website provides a PIN that students can use to associate their account with that specific class and access the class protocols. For each project, five tabs are available to store context for that class: "Background," "Syllabus," "Kit Materials," "Media, Reagent and Chemical List," and "Equipment List." Each of these tabs can be populated with HTML text by the instructor. A sixth tab, "Files," can be used to store PDF documents or image files for student use.

Below the tabs is a list of expandable headings for protocols and data collection, each of which contains four parts: "Introduction," "Protocol," "Upload Results," and "View Results." The first two of these are HTML text documents, including, for example, a step-by-step list of numbered items for the students to follow in the lab (Figure 2). At my institution, we commonly make laptops available in the lab for protocol access and allow students to bring their own laptops into the lab. If this is not possible, the text in each window can be printed

by the students so they can bring paper protocols into the lab. (Assessment tools and surveys may be embedded under appropriate headings at the beginning and/or end of the project.) The "Upload Results" heading contains up to 20 fields for students to enter text or upload a file as prompted by the instructor (Figure 3). File sizes up to ~1MB can be accommodated at present. Once the information is uploaded, all users can view it under the "View Results" heading. This feature makes collaboration among student groups possible, because they can download and view the results of all other groups in the class.

Figure 1. Screenshot of biochemistry projects available on GENI site.
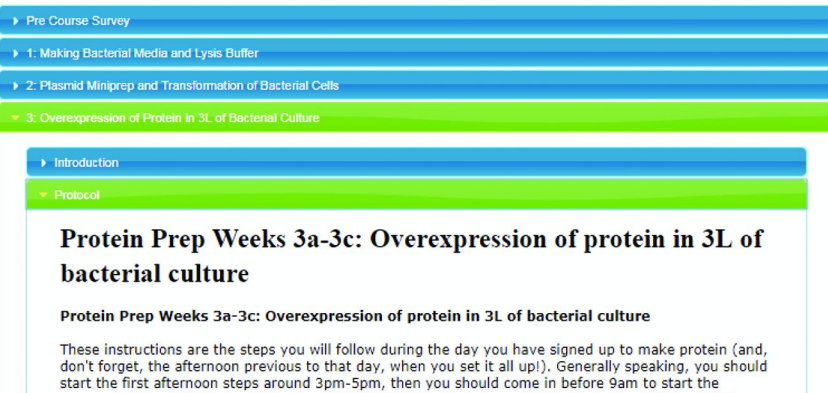*Reproduced with permission from Kathryn Houmiel, GENI Program Manager.*
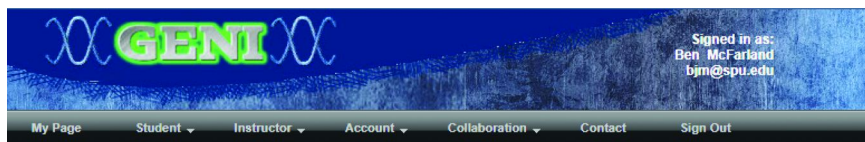
Figure 2. Screenshot of a biochemistry protocol. Reproduced with permission from Kathryn Houmiel, GENI Program Manager.
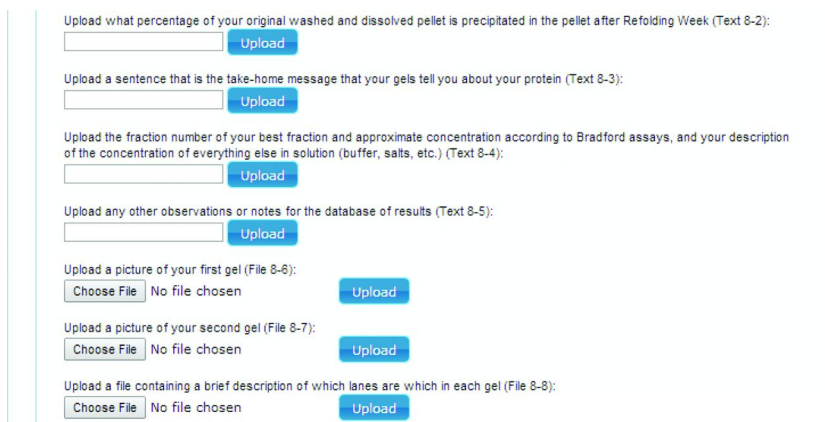

Figure 3. Screenshot of the "Upload Results" tab for the eighth week of protein production. Reproduced with permission from Kathryn Houmiel, GENI Program Manager.

At this point, I have published eight protein chemistry projects on the GENI website and used them in three courses over six years of teaching, so that 425 students have completed 176 projects: constructing and testing 132 homology models, purifying 30 different recombinant proteins, and measuring thermodynamics and/or kinetics of 12 different protein-protein interactions (Table 1). The general limitations to adapting authentic research experiments to the teaching laboratory are the time constraints of the assigned laboratory period and the resource constraints of the institution. Bioinformatics projects have an advantage in both of these areas, but benchtop experiments are possible as well, if carefully planned and organized with the help of the GENI website.

Another factor to consider when planning experiments is the constraint of the academic calendar and the institution's class schedule. Because fewer classes are scheduled during the summer, independent research projects can be assigned to students to investigate interesting results from the authentic research projects carried out in teaching labs during the previous academic year. Then those summer projects can help shape the questions asked and proteins assigned to students during the next academic year. In the GENI consortium (a group of academics who developed and use GENI), we call this self-reinforcing interplay between independent research and research in the teaching lab the "Research Cycle" (Figure 4). The key is to adapt research projects to the class time and schedule given by one's particular institution.

The primary technical characteristic of GENI that sets it apart from cloud drive storage is that it is an independent, central, and persistent web platform dedicated to CUREs and based on the HTML web programming language. Every web browser therefore has the potential to access it, and project registration and authentication is handled on the GENI site, rather than by an external third party. Adapting a research project to GENI is assisted by basic conversance with HTML tags. The results entered into GENI are stored in a central database accessible to all project users, so accessibility and authentication issues associated with student user accounts on third-party cloud drives are minimized. These issues can interfere with collaborations among groups of students in multiple lab sections, so GENI was designed to facilitate authentic research by facilitating this type of information transfer.

Because GENI has been developed by a relatively small consortium, most new features are added upon request from users. For example, at some institutions laptops posed safety concerns in the lab, so a feature was added allowing students to print protocols in one step directly from the website before lab. GENI is optimized for users who want to carry out projects among multiple lab sections or at multiple institutions at a scale at which cloud drives become less efficient or transferable, but who do not have the resources to develop their own central support system.

Table 1. Biochemistry projects completed with GENI.

| Quarter	Course	Project Title	Students	Results
Winter 2012	Biochem II	Preparative Protein Production from Inclusion Bodies…	19 in 9 groups	5 MICA & 2 NKG2D proteins
Fall 2012	Biochem I	Bioinformatics… of Genes Related to Three Agrobacterium Paralogs	47 in 23 groups	23 ProC homology models
Winter 2013	Biochem II	Preparative Protein Production from Inclusion Bodies…	27 in 13 groups	2 MICA & 6 NKG2D proteins
Fall 2013	Biochem I	Comparative Homology Modeling of ArgE Paralogs and Orthologs	62 in 29 groups	29 ArgE homology models
Winter 2014	Biochem II	Preparative Protein Production from Inclusion Bodies…	28 in 13 groups	3 MICA & 5 NKG2D proteins
Spring 2014	P. Chem Survey	Protein-Protein Binding by Surface Plasmon Resonance	11 in 5 groups	5 MICA-NKG2D interactions
Fall 2014	Biochem I	Comparative Homology Modeling of ProC Paralogs and Orthologs	53 in 28 groups	28 ProC homology models
Fall 2015	Biochem I	Predicting Structure and Function of ProC Paralogs	47 in 22 groups	22 ProC homology models
Winter 2016	Biochem II	Preparative Protein Production from Inclusion Bodies and Crystallization	31 in 15 groups	1 MICA & 3 NKG2D proteins
Spring 2016	P. Chem Survey	Protein-Protein Binding by Surface Plasmon Resonance	4 individuals	4 MICA-NKG2D interactions
Fall 2016	Biochem I	Predicting Structure and Function of Siderocalin Orthologs	60 in 30 groups	30 lipocalin homology models
Winter 2017	Biochem II	Preparative Protein Production from Periplasmic Expression	42 in 21 groups	5 lipocalin proteins
Spring 2017	P. Chem Survey	Protein-Protein Binding by Surface Plasmon Resonance	4 individuals	3 antibody-antigen interactions
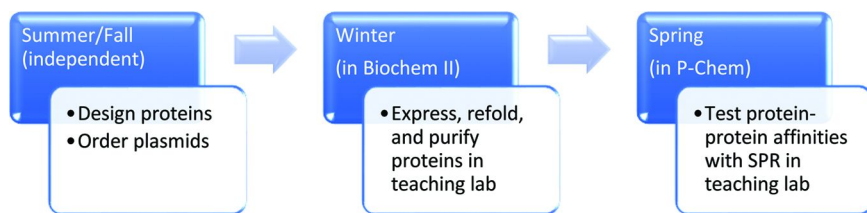
Figure 4. Flowchart of Research Cycle connecting Biochem II and Physical Chemistry Survey in the academic year.

There are many challenges inherent to bringing authentic research projects into the teaching lab, first and foremost being the fact that even the instructor does not know exactly what will happen. Straightforward projects with a high probability of success and with a low likelihood of surprise are more appropriate for undergraduate research in general and for authentic research projects in large lab courses in particular. In my experience, such projects are possible and can produce useful data. Below I will discuss a few of my personal experiences with conducting authentic research in the teaching lab to show that these challenges are always present but also can be met, especially if the primary goal is the education of the student rather than the collection of results or construction of new proteins. Novel results will usually be collected, but the students will always be educated with this approach.

Biochemistry I: Homology Modeling and Functional Annotation

GENI organizes a three-week computational protein chemistry exercise in the Biochemistry I (BIO/CHM 4361) laboratory scheduled in the fall quarter at Seattle Pacific University. Fifty to seventy students take this course in three or four lab sections, each staffed with a lab instructor and a teaching assistant. At this point in the course, the students have learned enough about the basics of protein structure that they can start to use PyMol, the industry standard tool for protein structure visualization. Instead of giving them a standard protein structure to visualize, they use online tools to build a homology model of a novel DNA sequence and interpret the structure in light of specific, authentic research questions. GENI provides an online protocol through a web browser with embedded links that take students directly to the bioinformatics tools they are told to use. After data are collected, GENI provides online database fields in the same browser window for students to input the data so that the process is more fully integrated on the students' laptops.

During the first week, students follow a step-by-step tutorial protocol in GENI to examine the structure of cytochrome c in PyMol (originally written by Peter C. Kahn). At the beginning of the second week's lab period, they are introduced to the research question being asked, are given a gene with unknown structure and function, and are shown several online tools and databases for gene analysis, homology modeling, and protein structure analysis. During the lab period, they work in groups on laptops to access online tools that analyze DNA homology and primary protein structure. By the end of the three-hour lab period, they submit their

gene sequence to homology modeling on the I-TASSER server, which typically takes 48-72 hours to complete.

The third week of the laboratory is scheduled as a workshop in which students work independently, examining the quality of their homology model and comparing it to experimentally determined crystal structures using PyMol and online tools for tertiary structure analysis. PyMol has several functions useful in this regard, including the ability to align multiple structures and map the distribution of charges on a protein surface. Students gather information and synthesize it into a written lab report to answer a question about the unknown function of their novel DNA sequence. At the end of the lab report, students design a simple experiment to test one of their functional hypotheses. Some students have used this experience to develop experiments in their later Molecular Biology course, and it could also form the basis of an independent summer research project. GENI's independent, central location online means that students' data remain available to them on the GENI web platform for later use. The ability to update HTML protocols in GENI allows for easy annual reformulation of the research question and updating of online tools used as web addresses change from year to year.

A nearly unlimited supply of research questions can be found in publicly available genomic databases. The relative simplicity and wide diversity of species represented by microbial genomes make them especially amenable to this approach. The first four projects I completed in this class originated in a collaboration with a colleague in the Department of Biology who conducts experiments on the microbe *Agrobacterium tumefaciens* and was involved in the original sequencing of its genome (*12*). Several amino acid metabolism genes in the genome had paralogs with unknown function, including the proline biosynthesis gene ProC and the arginine biosynthesis gene ArgE. These paralogs were used in various BLAST searches to find similar orthologs with unknown function, compared to the *E. coli* genes and other related genes with known functions. Each year the class examined 20-30 versions of ProC or ArgE genes and asked if physical characteristics of the protein model indicated preservation or divergence of function and binding specificity. The centralized online nature of GENI made modifications easier and allowed students to access other student groups' data for comparison.

The online Enzyme Similarity Tool from the Enzyme Function Initiative (EFI-EST) was used in the most recent projects to organize large numbers of DNA sequences into groups that may correlate with functional subgroups. EFI-EST can collect 10,000 DNA sequences from the UniProtKB database and organize them into a similarity network (*13*). For the projects described here, EFI-EST produces a manageable number of subgroups (4-6) so that representative sequences from each group can be assigned to different students for investigation with the working hypothesis that sequences from the same subgroup may have similar function. Sequences from the subgroup with Protein Data Bank structures or published functional data can be provided to the students as hints for structure and function of that subgroup.

The larger research question of functional annotation of these subgroups has been divided into smaller research questions in different ways. Students

have used direct links in GENI to access the National Center for Biotechnology Information Basic Local Alignment Search Tool (BLAST) and the Joint Genome Initiative Integrated Microbial Genomes (JGI-IMG) (*14*) database in order to compare their sequence to other sequences across genomes or within their source organism's genome. Because all student data on GENI is visible to other students, they can compare results to those obtained by other students with other organisms. One year, the project asked students to identify ribosome binding sites upstream from their assigned sequence because computational identification of these sites is still problematic (*15*), and students may be able to identify divergent, context-dependent sites better than some algorithms. Another year, the project linked to a number of online primary sequence analysis tools and asked students to use them to predict protein characteristics like pI and secondary structure from primary structure, and then to compare these characteristics to those of reference proteins with known function.

For homology modeling, the I-TASSER protein structure and function prediction server was chosen because of its ease of use, high capacity, and good performance in comparison studies (*16*). I-TASSER output includes model validation and functional prediction, and students are directed to particular parts of its output for their own interpretation. Students also perform external model validation using other online tools. In one case, a model validation service stopped working midway through the project. Because GENI is online, the protocol was modified to point the students to another validation service as the students performed their research.

The homology model is downloaded in the .pdb format so that students can compare it to other structures in PyMol and can submit it to other online tools that calculate characteristics like structural conservation, charge distribution, and aggregation propensity. Many web-based tools for structural analysis are published each year; the instructor can write default settings for these into GENI that will work with most students, and then work individually with groups that need different settings. Students are commonly uncomfortable with command-line entry, but most interfaces are menu-driven and can be used by students with minimal computer-science skills.

In the most recent iteration of this project, we explored a new family of proteins representing a new direction developing in my research. My personal interest in siderocalin function, and the fact that I observed several well-defined subgroups in the EFI-EST network, led me to investigate this family of proteins with this process, despite my personal inexperience. We explored this novel area of research in class and built 30 lipocalin homology models. A Linux terminal was set up in the lab to run a binding-site identication program (*17*) and students accessed it during the workshop week. This gave all students time to run the program, which took a few hours, and collect the results to integrate with the others for the lab report.

With the final lab report, students also submit a spreadsheet to GENI containing specific information from each phase of the project. I compile these and organize them by phylogenetic similarity, and can detect patterns within similar subgroups that may be related to function at the overall level. For example, we have developed working hypotheses for the binding specificities of

the different subgroups of lipocalins, and now these provide independent research projects for students.

GENI's online nature allows for fast response before the final due date for the project. I set an earlier deadline for data entry than for the final written report and review student data before the report is due. When some data from a particular tool are inconsistent on the spreadsheet, I notify the groups that they have used a particular tool incorrectly, and they can return to that tool and fix the problem before the final report is due. I typically need to correct some data from 5-6 groups every year, so the fact that GENI allows me to do this provides an important checkpoint in the process of authentic research.

One additional strength of GENI's centralized, online nature is that a timestamp is associated with each submission, so the instructor can track how the students proceeded through the project. The students are given two weeks to work independently on the project before the data are due. As the instructor, I can see if a group has been working steadily on the project or if it has been completed in a final rush of activity immediately before the due date, and this can be used to analyze the organization of the students' own research process or to calculate late penalties.

Biochemistry II: Protein Purification and Crystallization

GENI organizes an eight-week preparative protein chemistry exercise in the Biochemistry II (BIO/CHM 4362) laboratory scheduled in the winter quarter. About thirty students take this course in two or three lab sections. The details of how protein purification protocols from my post-doctoral research experience were adapted to the context of the teaching laboratory have been previously described (*18*). The primary challenge in this class is managing time constraints and bottlenecks so that each group of students can incubate liters of media for eight hours on limited shaker space, and so that the groups can share instruments like glass-bead lysis homogenizers efficiently. Some steps are conducted by grouping pairs of students into groups of four to accomplish tasks like column chromatography. With this arrangement, each lab section can make four novel recombinant proteins. GENI's collection of data in a database that is immediately accessible and transparent to all online allows ready data sharing within groups, within sections, and among the entire class.

Several small-scale protein design projects have been completed with this format, building on previous protein design projects. In 2011, a student used an online protein design program called HyPare (*19*) during summer research to design electrostatic stabilization into the MICA-NKG2D protein-protein interface. Plasmids were synthesized and students made mutant proteins in the winter 2012 teaching lab, which were tested for binding in later independent research. Some mutants resulted in significant interfacial stabilization (*20*). Then a research student used an online linker design program to link the two domains of the NKG2D homodimer, and in 2013 six designs were produced in the winter Biochemistry II teaching lab, and then tested for binding in the spring Survey of Physical Chemistry teaching lab. The NKG2D interdomain

interface was optimized with Rosetta in various ways, and these were produced in Biochemistry II in 2014 and 2016. These projects had varying degrees of success in completing the design objectives: the HyPare mutants stabilized the interface, but the NKG2D interfacial mutants did not have a significant effect on binding. In the pedagogical sense, all were a success because they brought students into an authentic research project in the teaching lab, giving them patterns of thought and laboratory skills that provided the basis for success in graduate school, medical school, and industry. The primary goal of this process is educational, and generation of new knowledge is a welcome but secondary effect.

As with the bioinformatics project in Biochemistry I, the location of these projects on GENI allows annual modification of a basic framework. The framework of Biochemistry II protein purification experiments implemented on GENI has been used for two different novel projects, showing how the teaching lab can become a place for genuine scientific exploration. In one project, a pre-veterinary biochemistry major found gene sequences of previously unexpressed MIC and NKG2D proteins from mammalian genomes. These proteins were produced in Biochemistry II labs and tested for binding in Survey of Physical Chemistry labs, showing significant cross-species binding (data to be published). In another project, previously unexpressed lipocalins that were modeled in Biochemistry I labs were produced in Biochemistry II labs. These proteins required significant modification of the previous protocols because lipocalins are produced by periplasmic expression, not in inclusion bodies. Students working with the modified protocols produced good amounts of lipocalin proteins, showing that GENI can be used in large classes with different kinds of protein production protocols.

Survey of Physical Chemistry: Protein Interaction Thermodynamics

GENI organizes a three-week protein-protein interaction thermodynamics exercise in the Survey of Physical Chemistry (CHM 3410) laboratory scheduled in the spring quarter. This course is small, ranging from four to twelve students, so the exercise is less structured, and the instructor is able to interact with students in the personal ways that scientists lead their research groups. In this context, GENI provides protocols and collects data like before, but the instructor can be present in the lab to help the students discern the quality of data, indicating which proteins show binding strengths appropriate to the surface plasmon resonance (SPR) technique. For several years, this course has been able to use the proteins made in the previous quarter in Biochemistry II. GENI's use as a single, universal online archive means that students are able to access Biochemistry II data online to see the results for their assigned protein in Physical Chemistry.

In the Physical Chemistry lab, a three-week structure is used that may be transferred to other types of biochemical projects. The class meets in the lab during the first week to learn how to use the instrument and to measure preliminary "preconcentration" data on the SPR instrument. Then, over the next two weeks, the instrument is made available for students to sign up for about six hours of

instrument time. Students are assigned multiple protein-protein interactions to screen and choose one interaction to test in triplicate for publication-quality data. This allows for the possibility that some protein-protein pairs may not bind well enough to be detected by SPR; the one-out-of-three chance for binding has been sufficient for all students to have a stable protein-protein interaction to test. If one student is unlucky enough to test three protein-protein pairs that fail, that student can be given the data posted on GENI from another student who tested two or three protein-protein pairs that worked. This is another form of collaboration facilitated by GENI's accessibility for all students with an internet connection and simplified by the fact that all data is located in one project-centered database. We have collected binding data for multiple projects, including redesigned MICA proteins binding NKG2D, single-chain and redesigned NKG2D proteins binding MICA, multiple mammalian species of recombinant MICA binding NKG2D, and serum antibodies binding recombinant microbial adhesion proteins (data to be published).

SPR is an expensive technique that is not commonly accomplished in the undergraduate curriculum. We purchased a used SPR systems from the mid-1990s, and found that it is sufficient to measure nanomolar to micromolar binding constants such as are found with the MICA-NKG2D system and typical antibody-antigen systems. This older instrument is not automated, but the students can compensate for that by completing repeated sample injections by hand, while refining their pipetting techniques. Other protein-protein interaction measurement instruments can be substituted on GENI with minimal modifications within this adaptable framework and used at any institution with an internet connection.

Conclusion

The online nature of GENI facilitates electronic links to assessment tools that can survey the entire populace of students using GENI. The GENI Consortium has collaborated with faculty and graduate students from schools of education to develop online surveys that take advantage of this fact, which is another sense in which the online nature of GENI fosters collaboration. Three surveys have been collected in each class, and the results are now being validated and interpreted to help us find best practices for using GENI and implementing authentic research in the teaching lab.

Ultimately, as GENI connects students to protocols and collects data online, it helps promote a host of deeper connections: among institutions, between courses, and to the community at large through publication. The flexibility of the GENI website has allowed me as a scholar to keep learning about protein design and protein-protein interaction chemistry as I have fulfilled my teaching responsibilities. The most important connections that GENI facilitates are internal to each student, as the individual makes new connections between what is done in the lab and what is learned in the classroom, motivated by the prospect of accomplishing something truly novel and unique. Others have noted that this kind of open collaboration and reasoning is a main goal of CUREs in general (*21*). Because GENI is at heart a collaborative technology, it builds new scientific

thinkers by giving students the chance to collaborate as scientists within the bounds of their required curriculum.

Acknowledgments

References

1. Clark, T. M.; Ricciardo, R.; Weaver, T. Transitioning from Expository Laboratory Experiments to Course-Based Undergraduate Research in General Chemistry. *J. Chem. Educ.* **2016**, *93*, 56–63.
2. Kerr, M. A.; Yan, F. Incorporating Course-Based Undergraduate Research Experiences into Analytical Chemistry Laboratory Curricula. *J. Chem. Educ.* **2016**, *93*, 658–662.
3. Danowitz, A. M.; Brown, R. C.; Jones, C. D.; Diegelman-Parente, A.; Taylor, C. E. A Combination Course and Lab-Based Approach to Teaching Research Skills to Undergraduates. *J. Chem. Educ.* **2016**, *93*, 434–438.
4. Brown, A. M.; Lewis, S. N.; Bevan, D. R. Development of a Structured Undergraduate Research Experience: Framework and Implications. *Biochem. Mol. Biol. Educ.* **2016**, *44*, 463–474.
5. Lopatto, D.; Hauser, C.; Jones, C. J.; Paetkau, D.; Chandrasekaran, V.; Dunbar, D.; MacKinnon, C.; Stamm, J.; Alvarez, C.; Barnard, D.; Bedard, J. E. J.; Bednarski, A. E.; Bhalla, S.; Braverman, J. M.; Burg, M.; Chung, H.-M.; DeJong, R. J.; DiAngelo, J. R.; Du, C.; Eckdahl, T. T.; Emerson, J.; Frary, A.; Frohlich, D.; Goodman, A. L.; Gosser, Y.; Govind, S.; Haberman, A.; Hark, A. T.; Hoogewerf, A.; Johnson, D.; Kadlec, L.; Kaehler, M.; Key, S. C. S.; Kokan, N. P.; Kopp, O. R.; Kuleck, G. A.; Lopilato, J.; Martinez-Cruzado, J. C.; McNeil, G.; Mel, S.; Nagengast, A.; Overvoorde, P. J.; Parrish, S.; Preuss, M. L.; Reed, L. D.; Regisford, E. G.; Revie, D.; Robic, S.; Roecklien-Canfield, J. A.; Rosenwald, A. G.; Rubin, M.

R.; Saville, K.; Schroeder, S.; Sharif, K. A.; Shaw, M.; Skuse, G.; Smith, C. D.; Smith, M.; Smith, S. T.; Spana, E. P.; Spratt, M.; Sreenivasan, A.; Thompson, J. S.; Wawersik, M.; Wolyniak, M. J.; Youngblom, J.; Zhou, L.; Buhler, J.; Mardis, E.; Leung, W.; Shaffer, C. D.; Threlfall, J.; Elgin, S. C. R. A Central Support System Can Facilitate Implementation and Sustainability of a Classroom-Based Undergraduate Research Experience (CURE) in Genomics. *CBE Life Sci. Educ.* **2014**, *13*, 711–723.

6. Kappler, U.; Rowland, S. L.; Pedwell, R. K. A Unique Large-Scale Undergraduate Research Experience in Molecular Systems Biology for Non-Mathematics Majors. *Biochem. Mol. Biol. Educ.* **2017**, *45*, 235–248.

7. Sarmah, S.; Chism, G. W.; Vaughan, M. A.; Muralidharan, P.; Marrs, J. A.; Marrs, K. A. Using Zebrafish to Implement a Course-Based Undergraduate Research Experience to Study Teratogenesis in Two Biology Laboratory Courses. *Zebrafish* **2016**, *13*, 293–304.

8. Streu, C. N.; Reif, R. D.; Neiles, K. Y.; Schech, A. J.; Mertz, P. S. Drug Synthesis and Analysis on a Dime: A Capstone Medicinal Chemistry Experience for the Undergraduate Biochemistry Laboratory. *J. Chem. Educ.* **2016**, *93*, 2084–2088.

9. Chinn, C. A.; Malhotra, B. A. Epistemologically Authentic Inquiry in Schools: A Theoretical Framework for Evaluating Inquiry Tasks. *Sci. Educ.* **2002**, *86*, 175–218.

10. Bell, J. K.; Eckdahl, T. T.; Hecht, D. A.; Killion, P. J.; Latzer, J.; Mans, T. L.; Provost, J. J.; Rakus, J. F.; Siebrasse, E. A.; Bell, J. E. CUREs in Biochemistry—Where We Are and Where We Should Go. *Biochem. Mol. Biol. Educ.* **2017**, *45*, 7–12.

11. Ditty, J. L.; Kvaal, C. A.; Goodner, B.; Freyermuth, S. K.; Bailey, C.; Britton, R. A.; Gordon, S. G.; Heinhorst, S.; Reed, K.; Xu, Z.; Sanders-Lorenz, E. R.; Axen, S.; Kim, E.; Johns, M.; Scott, K.; Kerfeld, C. A. Incorporating Genomics and Bioinformatics across the Life Sciences Curriculum. *PLOS Biol.* **2010**, *8*, e1000448.

12. Wood, D. W.; Setubal, J. C.; Kaul, R.; Monks, D. E.; Kitajima, J. P.; Okura, V. K.; Zhou, Y.; Chen, L.; Wood, G. E.; Almeida, N. F.; Woo, L.; Chen, Y.; Paulsen, I. T.; Eisen, J. A.; Karp, P. D.; Bovee, D.; Chapman, P.; Clendenning, J.; Deatherage, G.; Gillet, W.; Grant, C.; Kutyavin, T.; Levy, R.; Li, M.-J.; McClelland, E.; Palmieri, A.; Raymond, C.; Rouse, G.; Saenphimmachak, C.; Wu, Z.; Romero, P.; Gordon, D.; Zhang, S.; Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, B.; Liao, L.; Kim, S.; Hendrick, C.; Zhao, Z.-Y.; Dolan, M.; Chumley, F.; Tingey, S. V.; Tomb, J.-F.; Gordon, M. P.; Olson, M. V.; Nester, E. W. The Genome of the Natural Genetic Engineer Agrobacterium Tumefaciens C58. *Science* **2001**, *294*, 2317–2323.

13. Gerlt, J. A.; Bouvier, J. T.; Davidson, D. B.; Imker, H. J.; Sadkhin, B.; Slater, D. R.; Whalen, K. L. Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A Web Tool for Generating Protein Sequence Similarity Networks. *Biochim. Biophys. Acta* **2015**, *1854*, 1019–1037.

14. Markowitz, V. M.; Chen, I. M. A.; Palaniappan, K.; Chu, K.; Szeto, E.; Grechkin, Y.; Ratner, A.; Jacob, B.; Huang, J.; Williams, P.; Huntemann, M.;

Anderson, I.; Mavromatis, K.; Ivanova, N. N.; Kyrpides, N. C. IMG: The Integrated Microbial Genomes Database and Comparative Analysis System. *Nucl. Acids Res.* **2012**, *40*, D115–D122.

15. Barendt, P. A.; Shah, N. A.; Barendt, G. A.; Kothari, P. A.; Sarkar, C. A. Evidence for Context-Dependent Complementarity of Non-Shine-Dalgarno Ribosome Binding Sites to Escherichia Coli Rrna. *ACS Chem. Biol.* **2013**, *8*, 958–966.

16. Yang, J.; Yan, R.; Roy, A.; Xu, D.; Poisson, J.; Zhang, Y. The I-TASSER Suite: Protein Structure and Function Prediction. *Nat. Methods* **2015**, *12*, 7–8.

17. Viet Hung, L.; Caprari, S.; Bizai, M.; Toti, D.; Polticelli, F. LIBRA: Ligand Binding Site Recognition Application. *Bioinformatics* **2015**, *31*, 4020–4022.

18. Peterson, M. J.; Snyder, W. K.; Westerman, S.; McFarland, B. J. Preparative Protein Production from Inclusion Bodies and Crystallization: A Seven-Week Biochemistry Sequence. *J. Chem. Educ.* **2011**, *88*, 986–989.

19. Schreiber, G.; Shaul, Y.; Gottschalk, K. E. Electrostatic Design of Protein–Protein Association Rates. In *Protein Design: Methods and Applications*; Guerois, R., de la Paz, M. L., Eds.; Humana Press: Totowa, NJ, 2006; pp 235−249.

20. Henager, S. H.; Hale, M. A.; Maurice, N. J.; Dunnington, E. C.; Swanson, C. J.; Peterson, M. J.; Ban, J. J.; Culpepper, D. J.; Davies, L. D.; Sanders, L. K.; McFarland, B. J. Combining Different Design Strategies for Rational Affinity Maturation of the MICA-NKG2D Interface. *Protein Sci.* **2012**, *21*, 1396–1402.

21. Lopatto, D.; Tobias, S. *Science in Solution: The Impact of Undergraduate Research on Student Learning*; Council on Undergraduate Research: Wasghinton, DC, 2010.