








## ARTICLE

# A practical guide to teaching with Proteopedia

Claudia Castro<sup>1</sup>  | R. Jeremy Johnson<sup>2</sup>  | Bruno Kieffer<sup>3</sup>  | John A. Means<sup>4</sup> |  
Ann Taylor<sup>5</sup>  | Jason Telford<sup>6</sup> | Lynmarie K. Thompson<sup>7</sup> |  
Joel L. Sussman<sup>8</sup>  | Jaime Prilusky<sup>9</sup>  | Karsten Theis<sup>10</sup> 

<sup>1</sup>National University of Cuyo, School of Medical Science, Biochemical and Biotechnology Institute, Mendoza, Argentina

<sup>2</sup>Department of Chemistry, Butler University, Indianapolis, Indiana, USA

<sup>3</sup>IGBMC, Illkirch-Graffenstaden, France

<sup>4</sup>School of Natural & Social Sciences, University of Rio Grande, Rio Grande, Ohio, USA

<sup>5</sup>Department of Chemistry, Wabash College, Crawfordsville, Indiana, USA

<sup>6</sup>College of Arts and Sciences, Maryville University of Saint Louis, St. Louis, Missouri, USA

<sup>7</sup>Department of Chemistry, University of Massachusetts Amherst, Amherst, Massachusetts, USA

<sup>8</sup>The Weizmann Institute of Science, Department of Chemical and Structural Biology, Rehovot, Israel

<sup>9</sup>The Weizmann Institute of Science, Life Sciences Core Facilities, Rehovot, Israel

<sup>10</sup>Department of Chemical and Physical Sciences, Westfield State University, Westfield, Massachusetts, USA

## Correspondence

Karsten Theis, Department of Chemical and Physical Sciences Westfield State University, Westfield, MA.  
Email: ktheis@westfield.ma.edu

## Abstract

Proteopedia (proteopedia.org) is an open resource to explore the structure–function relationship of proteins and other biomolecules. This guide provides practical advice on how to incorporate Proteopedia into teaching the structure and function of proteins and other biomolecules. For 11 activities, we discuss desired outcomes, setting expectations, preparing students for the tasks, using resources within Proteopedia, and evaluating student work. We point out features of Proteopedia that make it especially suitable for teaching and give examples of how to avoid common pitfalls.

## KEYWORDS

molecular visualization, protein structure function and action mechanism, visual literacy, web-based learning

## 1 | INTRODUCTION

To understand biochemistry and molecular biology, students have to learn about structure–function relationships.<sup>1</sup> One difficulty is that biomolecules have complex three-dimensional structures, and students need to acquire spatial skills to understand them.<sup>2</sup>

Having access to three-dimensional (3D) models (either physical or rendered on a computer display) helps with understanding, especially when students can manipulate these models.<sup>3,4</sup> With practice, students learn how to visualize three-dimensional biomolecules internally while looking at two-dimensional depictions, a skill that is not automatic.<sup>5</sup> Once students

become familiar with the conventions of representing the 3D structures of biomolecules on paper, they can transition to deciphering figures they encounter in the scientific literature.

Proteopedia is an interactive, collaborative 3D-encyclopedia of proteins and other biological molecules.<sup>6–8</sup> Proteopedia implements 3D interactivity in a way that is accessible because it only requires internet access and a browser, and its use is free.\* Individual entries are connected by links, providing a web of information (as such, Proteopedia is part of the web that leverages the natural ways that humans learn<sup>9</sup>). Beyond traditional blue hypertext links, green links will open interactive 3D scenes right on the page. Viewers can rotate and zoom the structures and click on different parts for identification. Writing, revising, sharing, and reusing entries is easy because Proteopedia, like other wikis, is a collaborative environment by design.<sup>10</sup> A scene-authoring tool serves as a user-friendly way to create the 3D scenes integrated into the text. Proteopedia contains 11,051 entries written by 4725 registered users (as of February 2021). Because contributors are registered users, they take responsibility for the content they create. Most entries are editable by all users, but there are also reserved spaces for individuals (entries starting with “User:”), spaces used for work in progress (“Sandboxes”), spaces for groups, and spaces for interactive 3D complements to journal articles with limits to editing and, in some cases, viewing.<sup>11</sup> In addition, every structure in the protein database is available for exploration through computer-generated pages. While the mission of Proteopedia is to communicate structural biology to the public, it is also an excellent tool for teaching.<sup>12,13</sup>

Leveraging this platform and its content, the authors of this guide—from four different countries working at a variety of higher education institutions—have taught using Proteopedia in a wide range of contexts, using diverse lesson plans (Table 1), and using diverse biomolecules as examples (Figure 1 and Appendix S2). While these example lessons were all designed for undergraduate students of biochemistry, Proteopedia has also been used in secondary education and graduate schools. The content available on Proteopedia is sufficiently comprehensive to act as a 3D supplement to an introductory biochemistry course, as the page [Proteopedia: Biochemistry\_in\_3D]<sup>†</sup> illustrates. Reflecting on the combined teaching experience gained in the example lessons, we provide a practical guide describing how students can learn about biomolecular structure and function by studying existing Proteopedia content and how they can create Proteopedia content themselves as a way to delve deeply into structural biology.

## 2 | LESSONS THAT ASK STUDENTS TO READ AND VIEW PROTEOPEDIA PAGES

### 2.1 | Overview of activities

There are numerous reports that students who use 3D interactive models, on either tablet or computer, have increased learning outcomes in a number of disciplines.<sup>14–16</sup> The integration of 3D scenes with explanatory text (via “green links”) in Proteopedia provides interactivity paired with content. Learning goals for biomolecular visualization may be met by either consuming or creating content.<sup>17</sup> Here, we first discuss lessons where students *read* existing Proteopedia pages (example lessons prefixed with “R-”, Table 1).

*Lesson R-CC* was developed for a 15-week integrated Biochemistry/Cell Biology course (“From the Molecules to a Cell”) at the School of Medical Science, National University of Cuyo, Mendoza, Argentina. The goals were to facilitate the understanding of the relationship between 3D structure and protein function and to discover how conformational changes may lead to a misfolding protein that triggers disease. Students were asked to access Proteopedia, along with a study guide consisting of a series of motivating questions (for the full assignment, see Appendix S2). Students had the freedom to move into the discovery and understanding of different protein structures, ensuring they acquired the essential learning objectives for the topic. This activity allowed cooperative learning, and the development of this lesson allowed students to study proteins with more enthusiasm. In addition, the lesson helped to develop skills of information search and selection; students acquired the ability to search, discriminate, and select, which will help them in their learning process.

*Lesson R-JJ* asked students in a first-semester biochemistry course at Butler University, Indiana, USA, to use the integrated 3D representations of Proteopedia to answer structural questions about HIV protease. In this module, students were introduced to Proteopedia as a resource for exploring protein structure and function. As part of a semester-long, interwoven story on HIV protease,<sup>13</sup> students individually explored the Proteopedia page on HIV protease [HIV-1 protease] and answered questions about its structure and dynamics. What the Proteopedia page provides that students have not seen previously is the ability to see the dynamic movement of the flaps on HIV protease that facilitate polypeptide and inhibitor binding. Multiple structural representations were presented for the loop dynamics, with and without inhibitors bound, and students were asked to interpret the biological role of these dynamics. Protein structural

**TABLE 1** Summary of goals, extent and tasks for the lessons discussed here

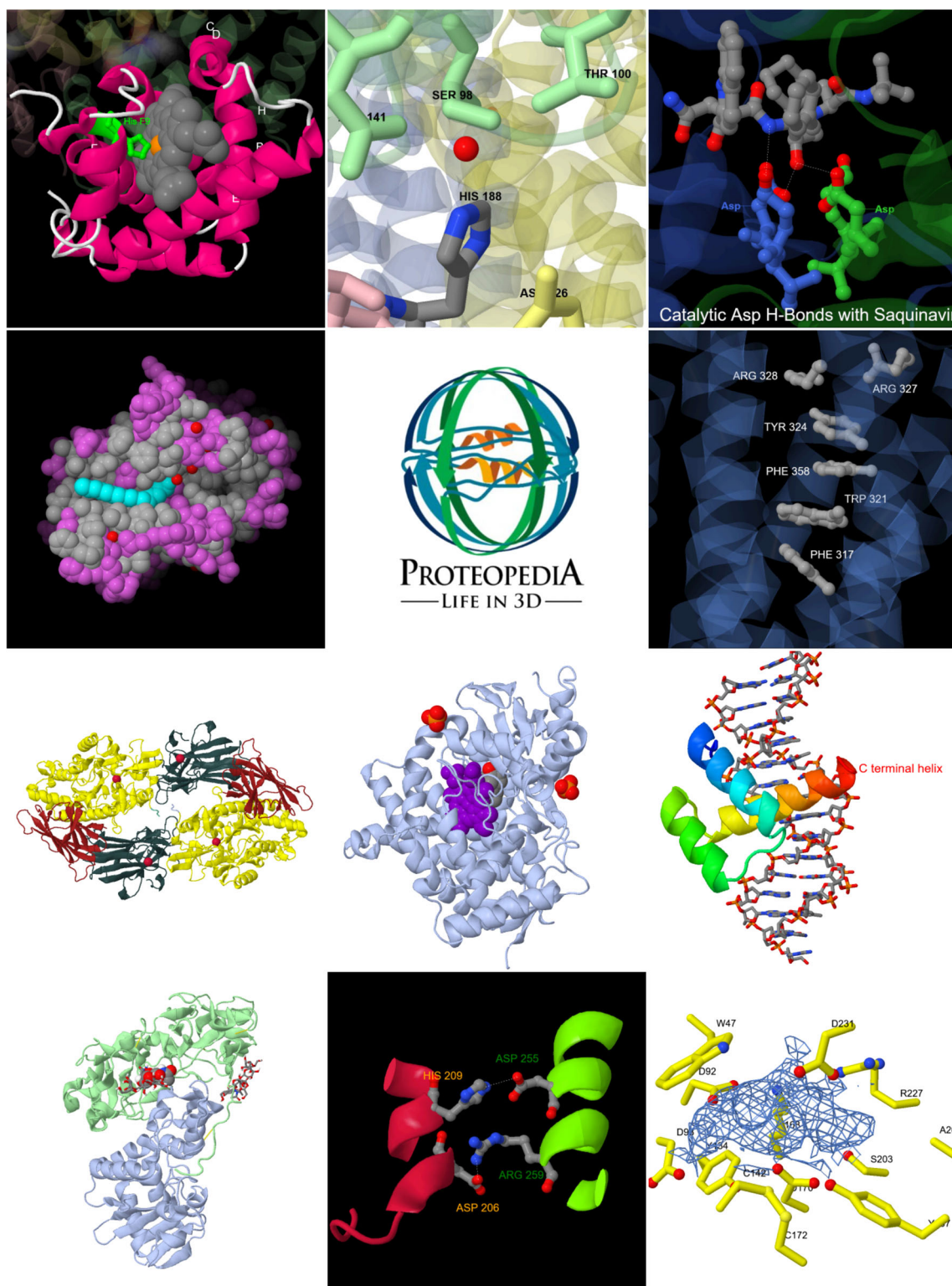
Lesson	Course type	Goals (among others)	Extent	Task
R-CC	Biochem for Medicine students	Focus on process instead of knowledge; 3D structure of proteins and their relationship to their function	One week	Associate protein's structures with their physiopathological functions in health and diseases
R-JJ	Biochem I	To investigate the structure, function, and dynamics of HIV protease.	Homework and in-class session	Individually read Proteopedia page and answer questions
R-AT	Biochem lab	Independent learning	Pre-lab	Read before lab, use in lab to design mutations
R-KT	Biochem I	Evaluate 3D figures	Three class sessions	Constructive review of existing Proteopedia figure
W-JJ	Biochem II	To have students develop their own stories relating protein structure and function.	Semester	In groups, create a Proteopedia page
W-BK	Structural Biology for Engineers	Support flipped classroom, produce rather than consume information, appreciate open questions	Semester	Create a Proteopedia page on chosen protein
W-JM	Biochem I	Make clear connections between the structure and function of biological macromolecules	Semester	Research biomolecule without Proteopedia presence in lieu of traditional research paper; biomolecules chosen from instructor-created list
W-AT	Biochem	Appreciation for the structural nuances of proteins and structure determination method	Semester	Create a Proteopedia page on chosen protein
W-JT	Biochem	Information literacy	semester	Write research paper, code web site
W-KT	Biochem I	Deeper understanding of 2D figures in publications	Two 3-h labs	Recreate published 2D figure in 3D
W-LT	Biochem I	Become familiar with 3D visualization, increase depth of engagement with subject matter	Semester	Create Proteopedia page

*Note:* There are four lessons based on existing Proteopedia pages (lessons starting with R-) and seven lessons where students create new pages (lessons starting with W-).

dynamics were discussed in the textbook, but students found them more tangible when they could manipulate these dynamics in three-dimensions. As an output from this exploration, students received a small point incentive for completing a short assignment (see Appendix S3).

*Lesson R-AT* was developed for a majors' level biochemistry course at Wabash College, a small, liberal arts college in Indiana, USA. Students used Proteopedia to identify the locations of previously studied sites in fumarate to form hypotheses and design their own site-directed mutants to study in the laboratory (see Appendix S4). Because the outcome of their multi-week lab experiment hinged on a reasonable analysis of the information provided in the instructor-written Proteopedia page [Fumarate 2], they were motivated to take a close look and understand the structural data in depth.

*Lesson R-KT* was developed for a biochemistry course at Westfield State University, Massachusetts, USA. The goals are to improve critical reading skills and to learn about biomolecular visualization. Students were asked to choose a topic from Proteopedia entries, evaluate one 3D scene in the entry according to a set of criteria, and present the topic to the class (for the full assignment, see Appendix S5). With permission from the students, constructive critiques were posted on Proteopedia on the discussion pages associated with the entries. For example, one student evaluated the entry on Ricin and suggested visualizing the hydrogen bonds between protein and ligand [Talk:Ricin]. Aided by the instructions and supported with a sample critique, students gave meaningful feedback on the entries and 3D figures, demonstrating their critical reading skills.



**FIGURE 1** Selected Proteopedia scenes from the activities described in this guide (scenes are either created by students or used in assignments). A wide range of topics and a wealth of structural insight are represented in these materials. Lessons associated with these scenes are (by rows) R-CC, R-AT, R-JJ; R-KT, n/a, W-JJ; W-BK, W-JM, W-AT; W-JT, W-LT, W-KT, with “R” indicating reading and viewing assignments and “W” indicating assignments where students write Proteopedia content. An interactive 3D version is available as Figure S1

To summarize, while all four example lessons made use of existing Proteopedia entries, lesson R-CC focused on structural principles, lessons R-JJ and R-AT focused on

structure–function relationships, and lesson R-KT focused on understanding what information an author can communicate through a figure. Clearly, these distinct





approaches dictate their own unique lesson plans and target distinct learning outcomes.

## 2.2 | Learning outcomes

When planning a new lesson, it is good practice to start with the desired learning outcomes, to consider how the assignment supports the overall goals of the course, and to share goals and outcomes with the students.<sup>18,19</sup> This way, the students know what the instructor wants them to take home from the information presented in the assigned Proteopedia pages and how they will be assessed. The example lessons presented here allow students to learn about specific biomolecules and to explore principles of molecular structure and functions. They also help students acquire skills, such as familiarity with conventions of biomolecular visualization, utilizing technical terms that are employed in biochemistry, correlating structural features with function, and developing strategies to tackle the complexity of biomacromolecules using interactive 3D models. Proteopedia can be included in problem-based learning (PBL), and be very useful for stimulating reasoning skills over the ability to memorize a concept. Table 2 (top) summarizes the combined learning outcomes for the example lessons that focus on reading and viewing existing Proteopedia pages.

## 2.3 | Choosing which pages to assign

Proteopedia contains many types of entries. Some entries showcase a single structure, while others give an overview (with many links) of a class of biomolecules, explain general concepts, or discuss the methods of obtaining structural data. While some pages are similar to original research papers (but with 3D figures) and require a higher level of expertise in the reader (e.g., [Hemoglobin]), others were written with a general audience in mind (e.g., [Tutorial: How do we get the oxygen we breathe]).

Depending on the context of the course, some pages will be more effective in supporting specific student learning outcomes than others. For many lessons, an existing page will suffice. For some lessons, however, it makes sense for the instructor to create a new page or adapt an existing one. Lesson R-AT is one such example where students work with fumarase in the lab (to test a specific hypothesis on the location of an active site histidine). The instructor-created page shows two views of fumarase, highlighting the two histidines in the context of the overall fold [Fumarase 2]. In lesson R-JJ, on the other hand, the instructor decided to use an existing page [HIV-1 protease] with rich information on the HIV-1

TABLE 2 Learning outcomes

Activities	Learning outcomes associated with at least one of the example lessons. After the lesson, students will be able to...
Read text and view 3D figures (lessons starting with R-)	<ul style="list-style-type: none"> <li>Identify visual elements used in images of molecular structures</li> <li>Interpret 2D depictions of 3D molecular structures</li> <li>Find non-covalent interactions and predict their relative strength</li> <li>Measure and interpret bond angles and distances</li> <li>Identify structural features and correlate them with function</li> <li>Interpret presented material when answering quiz questions</li> <li>Design lab experiments based on structural data</li> <li>Appraise clarity of figures and judge quality of Proteopedia entries</li> </ul>
Create text and 3D figures (lessons starting with W-)	<ul style="list-style-type: none"> <li>Find and interpret structural data from the protein data bank</li> <li>Identify and appraise published 3D structural evidence</li> <li>Apply copyright and plagiarism rules to their own writing</li> <li>Organize biochemical information into major topics</li> <li>Design, create and annotate figures</li> <li>Assemble figures and text to support an argument</li> <li>Critique drafts of their research papers and revise them</li> </ul>

protease structure, function, and inhibition. Before an assignment is given to students, we recommend checking whether the relevant Proteopedia entries have been edited. Each Proteopedia entry has a history page that makes earlier versions of the entry available; if the current page no longer matches the activity, it is possible to assign an earlier version.

Some of our sample lessons allow students to choose a page for the assignment. In that case, the instructor provides an introduction to searching pages in Proteopedia [Help:Searching]. For lesson R-KT, students

were asked to choose from pages featured in the journal *Biochemistry and Molecular Biology Education*<sup>20</sup>[Category: Featured in BAMBED]; as these are hand-picked (and peer-reviewed), they are outstanding examples of high-quality Proteopedia pages.

The choice of a Proteopedia entry to assign to students depends on the learning goals for the specific situation. If the instructor is looking for something specific that does not yet exist, it is straightforward to create an entry specifically for the intended lesson. The examples above illustrate how Proteopedia entries cover a broad range of curriculum content and diverse formats to cater to different course objectives.

## 2.4 | Teaching students how to read a Proteopedia article and view its figures

Proteopedia has been designed to have a low barrier of entry—visitors can look at the material just like they would view a web page, and clicking on the green links will provide rich structural information through the 3D figures. However, students can benefit from explicit training, giving them a more interactive experience. Instructors may wish to introduce Proteopedia by walking students through a page in class such as the Introduction to Protein Structure entry, which integrates tips for manipulating structures and using various Proteopedia tools. Students can also reference a viewing guide [Help: Viewing Pages] when they explore Proteopedia after class, reviewing how to use the mouse to reposition the 3D structures, identify atoms and residues, and zoom in and out, as well as how to use the right-click menu to alter the representation of the structure to explore it further. For a more comprehensive introduction, Proteopedia features a series of written and video tutorials to introduce its features [Help:Contents]. Thus, Proteopedia provides not only content on biomolecules but also the training material to get the most out of that content. Through a combination of directed instruction and independent learning using the resources on Proteopedia, students will become proficient in exploring the assigned entries and will also gain the skills to study material on Proteopedia independently.

## 2.5 | Ensuring deep interaction with the material

Assigned reading is a classic example of passive learning. However, there are various strategies to encourage students to engage actively with the material, fostering learning. One strategy is to assign questions about the material.

Questions can be anything from simple recall of material presented, to analyzing a structure, to evaluating the quality of the text and figures. For lesson R-CC, the instructor gave students a list of questions to ponder while studying the page on hemoglobin [Hemoglobin]; the goal was to help students understand how alterations in the structures of some molecules may lead to a misfolding protein that is responsible for a certain disease. For example, one question asked students to identify an amino acid change leading to an anomalous structure (see Appendix S2). Another strategy is to integrate questions into the text itself, as illustrated by the entries written by participants in example lesson W-LT [Student\_Projects\_for\_UMass\_Chemistry\_423\_Spring\_2016]. This ability to incorporate interactive exercises, quizzes, and other interactive features specifically for students separates Proteopedia from a traditional publication or textbook. Finally, students working on lesson R-AT directly applied what they learned from a Proteopedia page on fumarase [Fumarase 2] to design and make mutations in this enzyme. This models how a trained biochemist would first research the literature before planning and doing experiments.

Figures (in research papers and Proteopedia articles alike) carry crucial information about biomolecules, but students sometimes do not appreciate this, focusing instead on the text. To encourage readers to engage with the figures, some Proteopedia entries reference figures through probing questions in the text. For example, the entry [Tutorial:The opioid receptor: a molecular switch] asks readers to identify non-bonding interactions or to measure distances, angles, and torsion angles. One way to ensure that students engage with the figures is to have students evaluate those figures. For lesson R-KT, students were asked to evaluate the quality of a 3D scene and to suggest improvements (example prompts: “From which figure did you learn the most? What is good about the figure? What could be improved?”). In a second part of the assignment, students then had to explain a major point of the page to their peers, referring to the 3D figure projected in the classroom (see Appendix S4).

In summary, the four example lessons above suggest several strategies that turn a passive reading assignment into an active learning activity.

## 3 | LESSONS THAT ASK STUDENTS TO CREATE PROTEOPEDIA CONTENT

### 3.1 | Overview of activities

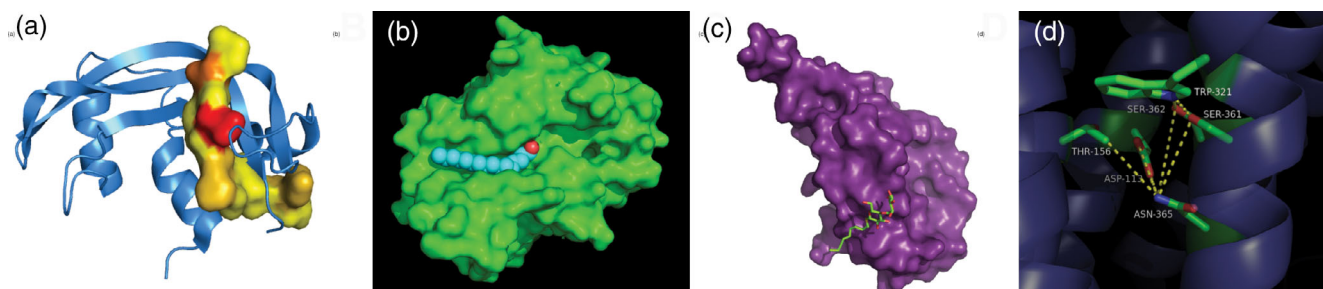
The seven lessons prefixed with “W-” (Table 1) all tasked students with *writing* new entries, i.e. creating new

content. This is more challenging, takes more time, and is more difficult to get started. On the other hand, students can be creative, delve even deeper into the subject matter, and end up with a product they can proudly share. They also learn about visualization techniques to create 3D figures, either through the built-in scene authoring tool or by writing scripts for the underlying Jmol 3D engine, a free, open source viewer for three-dimensional structures.<sup>21,22</sup> Finally, they interact with repositories of structural information (Protein Data Bank, PDB<sup>23,24</sup>). To make the expectations more tangible, it helps to give specific guidelines to students (see Appendices S6, S8, S9, and S10 for example grading rubrics).

*Lesson W-JJ* was developed for students in the second semester of undergraduate biochemistry at Butler University, Indiana, USA, after they had some exposure to Proteopedia already (see lesson R-JJ described above). For this course, Proteopedia is a central instructional component, as students applied their protein structure knowledge to complex biochemical systems. In a weekly 2-h computational lab, students developed Proteopedia pages relating the structure and function of a protein family that reinforced the lecture material for Biochemistry II.<sup>25–27</sup> Over 6 weeks, students worked in groups to learn about the design of a Proteopedia page (week 1), to prepare rough drafts of their Proteopedia pages (weeks 2–4), and to finalize their Proteopedia pages (weeks 5–6) (see Appendix S6). Proteopedia provides a scalable resource where students can transition from passive explorers of protein structure to active creators of stories about the interconnection between protein structure and function. Diverse topics including protein dynamics, ligand

binding, bioinformatics, structural representations, and literature analysis are all interwoven into the Proteopedia assignments. The finished Proteopedia entries provide students with a high-quality resource that can be shared with friends, family members, and professional school applications. The entries also provide a clear example of their understanding of how a protein's structure determines its biological function (Figure 2).

*Lesson W-BK* was developed as a support for a course in structural biology at the École Supérieure de Biotechnologie de Strasbourg, France, for students preparing for their bachelor's degree. The goals were to help students develop their ability to synthesize information from various sources (institutional databases, publications, etc.), to describe structure–function relationships in depth, and to recognize protein structure as an essential component of innovation in the field of biotechnology. The lesson was the conclusion of a formal course where fundamental aspects of protein structure and folding were presented. Organized in small groups, the students were asked to choose a topic related to human health and prepare a written project (see Appendix S7). The purpose of this exercise was to help students define their interests, to provide guidelines to avoid common pitfalls, and to provide the students with elementary project management methods. After 6 weeks, the Proteopedia pages, written in sandboxes, were submitted to and evaluated by the tutors, and feedback was provided to the individual groups. At this stage, students who contributed high quality examples were encouraged to publish their work as full Proteopedia entries. A meeting between tutors and students concluded the lesson, and various aspects of this authoring exercise, such as



**FIGURE 2** Example images from undergraduate student prepared Proteopedia pages. Undergraduate students work in pairs to prepare Proteopedia pages about diverse protein families and topics. In addition to interactive three-dimensional structures, students prepare static two-dimensional images. Examples of these images from student Proteopedia pages are provided ([https://proteopedia.org/w/CH462\\_Butler](https://proteopedia.org/w/CH462_Butler)). (a) Semi-synthetic RNase a, noncovalent complex of RNase a residues 1–118 (blue) with a synthetic peptide comprising residues 111–124 (yellow; colored by atom type) (PDB ID: 1SSC). (b) Palmitoyl protein thioesterase 1 (PPT1). Complex of PPT1 (green) with its natural ligand palmitic acid (blue; colored by atom type) (PDB ID: 1EI9). (c) Glucagon binding. Extracellular domain of the glucagon G-protein coupled receptor (purple) bound to the glucagon (sticks; colored by atom type) (PDB ID: 3IOL). (d) Neurotensin receptor. Closed form of the sodium binding pocket where the residues shown in green sticks cap the entrance of the sodium into the top of the binding pocket on the neurotensin receptor (blue) (PDB ID: 4XEE). Detailed descriptions of the Proteopedia assignments and grading rubrics are provided in the Appendix S6

copyright issues, knowledge sharing, and scientific responsibility were openly discussed.

*Lesson W-JM* was developed for a 15-week, first-semester biochemistry course at the University of Rio Grande, Ohio, USA. The goal was for students to explore the correlation between structure and function by developing a Proteopedia page for a biochemical molecule that did not already have a dedicated page. This assignment was provided to students as a more modern version of a classic research paper (see Appendix S8). Two class meetings were dedicated to familiarizing students with the RCSB PDB website and getting students accustomed to Proteopedia and its authoring tools, with an emphasis on the scene authoring tools. Students were then given most of the semester to complete their pages prior to the opening of a window of time for anonymous peer-grading. Students were asked to select a molecule from an instructor-curated list of proteins and nucleic acids (with PDB IDs) that did not already have a Proteopedia page. The final product for each student was a researched and referenced article that included background, significance, specific function, and other interesting information, such as connections to disease states. In addition, the articles included at least two 3D structural scenes that highlighted specific structural features that were tied to the function of the molecule. Utilizing an instructor-provided rubric, student Proteopedia pages were independently peer-graded by two classmates who only knew the sandbox number of the page that they were reviewing. (The instructor also graded every Proteopedia page.) Finally, students received the feedback from the peer-graders and the instructor so that they could address it and publish their pages to the permanent Proteopedia collection if they wished.

*Lesson W-AT* was developed for an advanced Biochemistry course at Wabash College, Indiana, USA. During the course, students were asked to choose a protein related to the topic of the course and give a short presentation on it in class. Examples include DNA-protein interactions [DNA-protein interactions] and proteins that are often involved in cancer [Proteins involved in cancer]. As their final project in the course, students each individually developed a page for their protein, and the individual pages were compiled into a summary page. Individual pages were peer-evaluated after a final presentation (for details of the lesson and rubrics, see Appendix S9).

*Lesson W-JT* was developed for a 15-week, second semester biochemistry course at Maryville University of St. Louis, St. Louis, Missouri, USA. The goals are for students to create and author a new content, as a page in Proteopedia, to examine the structure, function and significance of a protein or another biological molecule, and to gain experience at a basic level of web

development. To accomplish this, three graded assignments build up throughout the course of the semester (see S10). Students first decide on a biomolecule, second write a research paper on that biomolecule and finally convert the paper into a Proteopedia page, which is done as a workshop during class (see Appendix S8 for a more detailed description of the lesson). Examples of student-authored pages include pages on rhodopsin [Rhodopsin Structure and Function], the opioid receptor [Mu\_Opioid\_Receptor], lactate dehydrogenase [Human Lactate Dehydrogenase] and cockroach milk protein [Lili-Mip:Milk Protein].

*Lesson W-KT* was developed for a biochemistry lab at Westfield State University, Massachusetts, USA. The goal was for students to engage deeply with figures in a primary publication and to hone their biomolecular visualization skills. Students worked in groups of three during two 3-h lab sessions. The task was to recreate published 2D figures in a 3D format, with the goal of collaboratively creating a new Proteopedia page (see S11). In the first session, students accessed the relevant PDB files, loaded them into Jmol and changed the representations and colors to match that of the paper. In the second session, students worked with the instructor to tackle more sophisticated steps (superpositions, showing surfaces and electron density) and wrote figure captions. Finally, the authors of the original 2D figures were invited to share their impressions of the 3D recreations. The assembled 3D scenes are available on Proteopedia [Schubert lab: bacterial InIC disrupts human Tuba complexes], [Garman lab: Interconversion of lysosomal enzyme specificities], [Kisker lab: 5B5Q].

In *lesson W-LT*, student teams in a biochemistry course at the University of Massachusetts Amherst created Proteopedia pages illustrating a protein-drug complex. As they created a page that explored and illustrated a protein complex they found important, they discovered the concepts discussed in class at work in a particular protein. The communication objective of their page required them to see and overcome the challenge of finding a clear and attractive way to illustrate key points about complex molecules. Students worked in teams of 4–6 students, ideally containing a mix of chemistry and chemical engineering majors. Project deadlines were spread throughout the semester, beginning with choosing a topic of interest (see Appendix S12 for the complete assignment). The next deadline was for each team to create a 3D scene during an in-class Proteopedia workshop. The rest of the project was done outside of class throughout the semester, with each team member responsible for one of the sections of the Proteopedia page. The team as a whole was responsible for coordinating sections to achieve continuity and avoid overlap before the page was complete.



These seven lessons are examples of how asking students to create material on Proteopedia supports teaching and learning in a range of different courses and modalities.

### 3.2 | Learning outcomes

Asking students to write a Proteopedia page combines multiple learning outcomes related to researching the literature and databases, writing about science, and creating 3D figures (Table 2, bottom). The background research may develop students' information literacy, for example, training them to go beyond a single Google search toward a discipline-specific search strategy.<sup>28</sup> As a tool to present and remix published information, Proteopedia provides an excellent way for students to analyze and synthesize information from diverse sources, as opposed to duplicating existing information.

Writing a Proteopedia page can serve as a replacement for the traditional research paper. Students have to carefully read the primary literature and critically evaluate the available structure data. If English is not the language of instruction, they will be exposed to technical terms in multiple languages. They are introduced to the genre of secondary literature within the manageable format of an encyclopedia entry (with a couple of sources rather than hundreds cited in a review article) where their writing has to be precise, concise, and tell a story within a given context. When working in groups, they collaborate in a low-stakes situation. Lessons may also provide an opportunity to discuss copyright laws and the ethical use of sources. For instance, before assigning activity W-BK, the instructor introduced legal issues raised by the reproduction of figures or reuse of text to prevent unauthorized use of published material. Guidelines for writing for Proteopedia are available on the site [Proteopedia: Guidelines for Ethical Writing].

Creating the 3D Jmol figures and integrating them into the text builds skills in coding (supported by Proteopedia's text editor and scene authoring tool) and allows students to internalize the structural concepts discussed in class, by engaging deeply with a structure to create 3D scenes that illustrate important points about the protein. For example, they have to correlate primary sequence with structural information when labeling residues in figures. For lesson W-KT, the assignment was to create a single figure using the scripting language of the Jmol 3D engine to access the advanced features not built into the user-friendly scene authoring tool. The scripts used to make the figures are available in the talk section, illustrating advanced techniques (i.e., superposition, depicting electron density, animations, interactivity) for

others to learn by example (e.g., [Schubert lab: bacterial InIC disrupts human Tuba complexes]). Students learn about the choices that experts (the authors of the original papers) make when creating 2D figures (i.e., color scheme, what to emphasize and what to leave out, choice of representation, placement of labels) and learn to what extent these decisions transfer to a 3D figure.

Overall, the ability to realize the full implications of their online-authored Proteopedia pages, from pure scientific aspects to ethical, organizational, and legal dimensions provides great educational value to the Proteopedia assignments.

### 3.3 | Helping with time management

As anyone who has written an article or made a figure knows, this requires much more time than reading an article or looking at a figure. The instructor has to make decisions regarding the extent of the assignment and the amount of class time to provide. For example, lesson W-KT set aside two lab sessions to make a single 3D scene (including training students how to use Jmol to do so). At the other end of the scale, lessons W-JJ, W-BK, W-JM, and W-JT are semester-long projects involving literature research and writing an entirely new Proteopedia page. For these longer projects, it helps students to have a timeline with milestones, and build in some form of a review for drafts. For example, the milestones for W-JT are:

*Day 1*, apply for a login and have students make their own sandboxes [Help: Sandboxes].

*Week 3*, choose a protein; write one or two sentences describing the function.

*Week 7*, start a paper: submit an outline of relevant points, references, and where the protein 'fits' in a system.

*Week 12*, storyboard the page, have most of the text and references done in a text editing program. Sketch, print, or describe the figures that will be embedded in the page.

*Weeks 14–15*, construct and publish a page on Proteopedia. Done as a hands-on, in-class exercise.

For lesson W-BK, in addition to a timeline similar to the one shown above, the progression of the different groups is followed by tutors on the learning management system. The request for regular progress reports is intended to help students in planning the different aspects of the Proteopedia assignment and provides an excellent introduction to project management questions. How much time students need for an assignment is related to whether they work individually or in groups, and this depends on how the activity is integrated into the course taught.

### 3.4 | Provide the appropriate training and infrastructure

Using technology with students always carries the risk of shifting the focus from learning about the content to troubleshooting technology problems. We find that providing upfront training, targeted to the lesson at hand, can minimize frustration. Training includes such things as information about the mechanics of writing text (wiki markup and editing interface, [Help: Editing]) and creating figures (scene authoring tool, [Scene Authoring Tools]). Existing Proteopedia pages are excellent starting points, and their underlying wiki markup is accessible by clicking on the “edit” tab. There are training videos showing how to make 3D scenes in Proteopedia using the scene authoring tool [Proteopedia:Video Guide], as well as written tutorials [Proteopedia:DIY:Scenes].

There are also resources within Proteopedia for learning the underlying Jmol scripting language to achieve advanced effects such as overlaying molecules [Superpositions with Jmol] or showing electron density [Electron density maps], as well as tutorials from other sources. For activity W-KT, Cohlberg's Jmol tutorial<sup>29</sup> is not only helpful to introduce Jmol scripting but also gives a very useful introduction to accessing structural data deposited in the protein database.

If students have already made 3D figures using PyMOL, an open source visualization application,<sup>30</sup> they can use a converter to reuse those figures in their Proteopedia page. Apart from knowing the basics of Proteopedia authoring, students should also have a solid understanding of how to find and use the structural data from the Protein Data Bank. Finally, technology frustration or anxiety can be avoided by deciding on the type of Proteopedia page students will create (e.g., sandbox, protected page, public page). These considerations are discussed in [Proteopedia: Practical\_Guide\_to\_Advanced\_Proteopedia\_Authoring\_and\_Its\_Use\_in\_Teaching].

### 3.5 | Choice of topic

A wise choice of the topic to write about is essential for student success. Our experience is that students need help choosing their topic. Different approaches have worked well in each of our lessons, ranging from assigning a specific topic, to letting students choose any topic and having it approved. If the instructor assigns the topic, there is the advantage that content information can be provided upfront in class. On the other hand, letting students make the choice will motivate them to share interests and knowledge they might have already.

Lessons W-JJ, W-AT, and W-KT focused on a single, pre-selected topic or paper for the entire class, while lesson W-LT employed topics that were selected by teams, lesson W-JM used individual topics that were selected from an instructor-developed list, and lesson W-JT used individual topics that were selected by the students with some clearly-delineated limitations. Lesson W-BK employed strategies that were a combination of those described above—self-selected topics per group.

In lesson W-JJ, the entire class would study one larger topic (e.g., receptors) chosen by the instructor, with a group from two different lab sections being assigned overlapping proteins within this larger class. Each group then created sandboxes, which were merged through student and instructor collaboration into the final Proteopedia page (e.g., [Neurotensin receptor]). This resulted in a series of connected Proteopedia pages ([Glucagon receptor] and linked entries), providing a high-quality resource on the larger topic. Every time the course is taught, a new larger topic is chosen ([CH462 Butler]). A similar strategy is utilized in lesson W-AT, with the class writing a summary page that gave an overview of the topic ([DNA-protein interactions]) in addition to the individual pages. For lesson W-KT, the assignment focused on a research paper that was discussed by the entire class much like in a journal club. Then for their lab assignment, each lab group was tasked to recreate one of the figures from the paper in 3D using Jmol.

Lesson W-LT demonstrates a topic selection strategy that involves multiple, different topics in the class, as opposed to a common topic for the class. In this lesson, students work in teams to create a Proteopedia sandbox page illustrating a protein-drug complex for a disease of interest. They start from a template in Proteopedia that contains all of the sections and instructions for the assignment, and each team member is responsible for creating the text and 3D scenes for one of the sections.

Lessons W-JM and W-JT employed topic selections made by individual students, allowing for a greater degree of individualism in the classes but also introducing greater complexity for the instructor, since there is a greater variety of topics to keep track of throughout the students' work. Lesson W-JM utilizes a list of biological molecules for which no Proteopedia already exists just before the semester in which the lesson is utilized. The particular biomolecules that are made available to the students are selected by the instructor based on the quality of the structures in the PDB as well as the alignment of the residue numbering between the structure file and the peer-reviewed publication to which the structure file is tied. In previous iterations of the lesson, most students had significant difficulties working with structure files that did not follow the same residue numbering

scheme that was shown in the figures of the publication, so prior screening was determined to be necessary to avoid a high degree of student frustration.

In lesson W-JT, students are given the responsibility of selecting their own biomolecules with some limitations. Students are asked to choose proteins that have an entry in the Protein Data Bank so that one or more structural coordinate files are available. Very commonly used examples of proteins are excluded, specifically disqualifying most of the proteins that are discussed in the class. Students are encouraged to research areas that interest them. Sometimes it is a protein central to their career goals; for instance, a student who was admitted to optometry school worked on rhodopsin. For some, it is a protein involved in a disease that impacted their social network or family. Their motivations were as diverse as the students, and most came up with interesting subjects.

If students choose their own topic, the instructor should check that they are using the highest quality structures available for that topic and that students are aware of related structures with bound ligands, different conformation, etc., so that they have a more complete view of the known structural information. There will be students that have difficulty choosing a biomolecule at first, since there are ~174,000 structures in the Protein Data Bank, and the variety of structures can be overwhelming. Usually, through a few discussions, a subject can be pinned down based on the student's interest. This approach was implemented in lesson W-BK, where groups of students were asked to choose a structure according to their interests. The difficulty of this process is handled by assigning a tutor to each group to prevent common traps such as the use of low-quality PDB entries, or a specific member of a large family of proteins already described through other members in Proteopedia. This demanding step was formalized by requesting the students to write a small and early report about the rational considerations underlying their choices.

### 3.6 | Review and dissemination

The quality of student work will be higher if there is the opportunity for review while working on the project and for presentation after completion of the project. Review could come in the form of teacher's comments on drafts handed in or from (student) peer review. For lesson W-KT, researchers who published the structures commented on the 3D scenes, giving students a unique chance to interact with the researchers who solved the structures that the students were studying. For lesson W-BK, time was set aside at the end of the semester to provide

general feedback about student work, including social and ethical considerations. For lesson W-JM, finished pages were peer-reviewed, yielding constructive feedback from the instructor and other students, allowing students to learn how to both give and receive constructive critique in a professional manner. For lesson W-LT, finished presentations included a quiz-worthy question that could be used later for an online quiz for the rest of the class to engage with the structure and scenes to answer the question. Ways of disseminating student work could be a presentation, involve moving an excellent page from a sandbox to a public page on Proteopedia (some are featured in [Student Projects]), or even publishing in a peer-reviewed journal (for example lesson W-JJ,<sup>26</sup>). This allows students an opportunity to join the community of scholars through a course-based activity.

## 4 | CONCLUSION

Based on 11 example lessons, this guide is a sampling of how Proteopedia can be a versatile tool for teaching and learning structural biology. Among the 3D visualization tools available, Proteopedia is particularly well-suited for education. With its unique combination of explanatory text and 3D scenes of molecular structures, it directly connects structural data to its interpretation and functional consequences. Students can benefit from the rich content already available on the site, and they can learn by adding their own articles with integrated 3D scenes on a platform with a low barrier of entry. Students can access structural information on computers and mobile devices, on their own, or in the classroom. Because Proteopedia is an open educational resource,<sup>31</sup> the students' work is showcased and remains available to the public. Working with Proteopedia content introduces young scientists to several questions raised by current information technologies, such as open access to knowledge and authoring responsibility. Finally, Proteopedia's connection to journals and researchers gives students a glimpse of science as an ongoing process that textbooks typically do not provide. As illustrated by the variety of lessons presented here, teaching with Proteopedia caters to a range of courses and supports multiple learning goals. Teachers interested in incorporating Proteopedia in their courses can either adopt one of the example lessons discussed here or design their own based on the general recommendations of this guide.

## ACKNOWLEDGMENTS

We would like to acknowledge the pioneering work of Jane and David Richardson, bringing interactive 3D

figures to the classroom. Also, we would like to thank Eric Martz, Angel Herraéz, and Bob Hanson for their sustained effort to provide content and innovation to Proteopedia and the underlying Jmol engine and for all their work to maintain Proteopedia and support its user community. This work benefited from access to the Dana and Yossie Hollander Center for Structural Proteomics at the Weizmann Institute of Science, Rehovot, Israel, an Instruct-ERIC Centre.

## ORCID

Claudia Castro  <https://orcid.org/0000-0002-8428-2484>

R. Jeremy Johnson  <https://orcid.org/0000-0003-2694-3915>

Bruno Kieffer  <https://orcid.org/0000-0002-2033-4679>

Ann Taylor  <https://orcid.org/0000-0002-9594-8501>

Joel L. Sussman  <https://orcid.org/0000-0003-0306-3878>

Jaime Prilusky  <https://orcid.org/0000-0002-7019-0191>

Karsten Theis  <https://orcid.org/0000-0003-2697-6143>

## ENDNOTES

\* This ease of access was of great value for remote instruction during the COVID-19 pandemic.

† Text in bracket refers to Proteopedia entries, for example, [Hemoglobin], is at <https://proteopedia.org/w/Hemoglobin> or accessed by entering "Hemoglobin" into the search box found on the left-hand side of every Proteopedia page.

## REFERENCES

- Voet JG, Bell E, Boyer R, Boyle J, O'Leary M, Zimmerman JK. Recommended curriculum for a program in biochemistry and molecular biology. *Biochem Mol Biol Educ*. 2003;31:161–2.
- Oliver-Hoyo M, Babilonia-Rosa MA. Promotion of spatial skills in chemistry and biochemistry education at the college level. *J Chem Educ*. 2017;94:996–1006. <https://doi.org/10.1021/acs.jchemed.7b00094>.
- Howell ME, Booth CS, Sikich SM, Helikar T, Roston RL, Couch BA, et al. Student understanding of DNA structure–function relationships improves from using 3D learning modules with dynamic 3D printed models. *Biochem Mol Biol Educ*. 2019;47:303–17. <https://doi.org/10.1002/bmb.21234>.
- Richardson DC, Richardson JS. Teaching molecular 3-D literacy. *Biochem Mol Biol Educ*. 2002;30:21–6.
- Copolo CF, Hounshell PB. Using three-dimensional models to teach molecular structures in high school chemistry. *J Sci Educ Technol*. 1995;4:295–305.
- Prilusky J, Hodis E, Canner D, Decatur WA, Oberholser K, Martz E, et al. Proteopedia: a status report on the collaborative, 3D web-encyclopedia of proteins and other biomolecules. *J Struct Biol*. 2011;175:244–52.
- Hodis E, Prilusky J, Martz E, Silman I, Moulton J, Sussman JL. Proteopedia - a scientific "wiki" bridging the rift between 3D structure and function of biomacromolecules. *Genome Biol*. 2008;9:21.
- Voet JG, Voet D. Proteopedia and other exciting news. *Biochem Mol Biol Educ*. 2010;38:283–3.
- Brown JS. Growing up digital: how the web changes work, education, and the ways people learn. *Change: Mag High Learn*. 2000;32:11–20.
- Zheng B, Niiya M, Warschauer M. Wikis and collaborative learning in higher education. *Technol Pedagog Educ*. 2015;24:357–74.
- Prilusky J, Decatur W, Martz E. Publishing in Proteopedia: the guide, advancing methods for biomolecular crystallography, in NATO science for peace and security series a: chemistry and biology. Dordrecht, The Netherlands: Springer; 2013.p. 277–95.
- Jaswal SS, O'Hara PB, Williamson PL, Springer AL. Teaching structure: student use of software tools for understanding macromolecular structure in an undergraduate biochemistry course. *Biochem Mol Biol Educ*. 2013;41:351–9.
- Johnson RJ. Teaching foundational topics and scientific skills in biochemistry within the conceptual framework of HIV protease. *Biochem Mol Biol Educ*. 2014;42:299–304.
- McCullum BM, Regier L, Leong J, Simpson S, Sterner S. The effects of using touch-screen devices on students' molecular visualization and representational competence skills. *J Chem Educ*. 2014;91:1810–7.
- Belford R, Moore EB. ConfChem conference on interactive visualizations for chemistry teaching and learning: an introduction. *J Chem Educ*. 2016;93:1140–1.
- Wu H-K, Krajcik JS, Soloway E. Promoting understanding of chemical representations: students' use of a visualization tool in the classroom. *J Res Sci Teach*. 2001;38:821–42.
- Dries DR, Dean DM, Listenberger LL, Novak WRP, Franzen MA, Craig PA. An expanded framework for biomolecular visualization in the classroom: learning goals and competencies. *Biochem Mol Biol Educ*. 2016;45:69–75.
- Wiggins G, McTighe J. Understanding by design. 2nd ed. Alexandria, VA: Association for Supervision and Curriculum Development; 2005.
- Herraéz A, Costa MJ. Biochemical visual literacy with constructive alignment: outcomes, assessment, and activities. *Biochem Mol Biol Educ*. 2013;41:67–9.
- Hodis E, Prilusky J, Sussman JL. Proteopedia: a collaborative, virtual 3D web-resource for protein and biomolecule structure and function. *Biochem Mol Biol Educ*. 2010;38:341–2.
- Hanson RM. Jmol—a paradigm shift in crystallographic visualization. *J Appl Cryst*. 2010;43:1250–60.
- Herráez A. Biomolecules in the computer: Jmol to the rescue. *Biochem Mol Biol Educ*. 2006;34:255–61.
- Sussman JL, Lin D, Jiang J, Manning NO, Prilusky J, Ritter O, et al. Protein data bank (PDB): a database of 3D structural information of biological macromolecules. *Acta Crystallogr Sect D: Biol Crystallogr*. 1998;54:1078–84.
- Berman H, Henrick K, Nakamura H, Markley JL. The worldwide protein data Bank (wwPDB): ensuring a single, uniform archive of PDB data. *Nucleic Acids Res*. 2007;35:D301–3.
- Johnson RJ, Andorfer M, Chaffee A, Clark M, Clarke N, Douglass G, et al. Proteopedia entry: bovine pancreatic ribonuclease a. *Biochem Mol Biol Educ*. 2012;40:75.
- Jeremy Johnson R, Applegarth J, Bennett S, Caskey S, Coatney C, Corsbie K, et al. Proteopedia entry: G-protein coupled receptors. *Biochem Mol Biol Educ*. 2017;45:277–8.





27. Jeremy Johnson R, Bartels A, Erkilli R, Green N, Han S, Holt N, et al. Proteopedia entry: mammalian serine hydrolases. *Biochem Mol Biol Educ*. 2015;43:60–1.
28. Shultz GV, Zemke JM. "I wanna just Google it and find the answer": student information searching in a problem-based inorganic chemistry laboratory experiment. *J Chem Educ*. 2019;96:618–28.
29. Exploring proteins and nucleic acids with Jmol. <http://web.csulb.edu/~cohlberg/Jmolmanual.html>.
30. PyMOL. <https://pymol.org/2/>.
31. Hilton J III. Open educational resources and college textbook choices: a review of research on efficacy and perceptions. *Educ Technol Res Dev*. 2016;64:573–90.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Castro C, Johnson RJ, Kieffer B, et al. A practical guide to teaching with Proteopedia. *Biochem Mol Biol Educ*. 2021;49: 707–719. <https://doi.org/10.1002/bmb.21548>