JOURNAL OF APPLIED CRYSTALLOGRAPHY

ISSN 1600-5767

Remote laboratory training for high school students: grocery store based hands-on project in protein crystallography

Ali Lillian Fox, Alyssa R. Teteris and Irimpan I. Mathews*

Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA. *Correspondence e-mail: iimathew@slac.stanford.edu

Received 16 May 2023
Accepted 3 August 2023

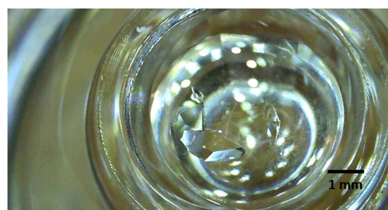
Edited by J. M. García-Ruiz, Instituto Andaluz de Ciencias de la Tierra, Granada, Spain

Keywords: remote learning; high school projects; protein crystallography; protein function; drug design.

PDB reference: hen egg-white lysozyme grown with kitchen recipe, 7knk

Supporting information: this article has supporting information at journals.iucr.org/j

The COVID-19 pandemic measures forced students to stay home and confined them to remote learning. This had a large impact on laboratory experiments, which are often impossible to complete from home. This article is a resource for instructors/educators to introduce the topic of structural biology and crystallographic methods. The main focus is to describe a hands-on crystallization laboratory exercise that can be carried out remotely at home with safe household products. X-ray crystallography is a vital technique for determining protein structure and function. This information can be used to understand fundamental biological processes and to help in the design of life-saving medications. Here, a method was developed to teach crystallography using reagents and equipment that can be found in grocery stores. The steps involved in a crystallography experiment are detailed with links and references to additional resources.



1. Introduction

For high school science students, visualizing proteins and molecular-level interactions can be a major barrier to understanding biochemical concepts. Because they cannot be seen, molecular-level interactions often seem abstract. Even when depicted as two-dimensional pictures in a textbook, it is easy for these concepts to feel disconnected from the real world and from other parts of science. Having a project that works to visualize macromolecules can help make these concepts feel more concrete (Bethel & Lieberman, 2014; Garcia-Ruiz *et al.*, 2015).

One promising project in this vein is a hands-on laboratory experiment focused on X-ray crystallography. X-ray crystallography is one of the most prevalent ways of determining a protein structure. There are a number of educational sites that explain crystallographic concepts in simple terms (*e.g. Crystallography 101*: https://www.ruppweb.org/Xray/101index.html). This structural and chemical information can be used to better understand fundamental biological processes on the molecular level, and can inform drug design (Anderson, 2003; Hammond, 2015).

Owing to its importance in research, use in biochemistry and other wide-ranging applications, X-ray crystallography presents a compelling way to study proteins at high school level. Learning about the usefulness of protein structure as well as the importance of structure in drug design creates a tangible way for high school science students to connect their learning to the real world. It also meets one of the goals of the US Next Generation Science Standards (NGSS) for K-12 education [*i.e.* kindergarten to twelfth grade (usually 6 or 7 to

18 years old)], which is to teach scientific practices that reflect those used by professional scientists (Achieve Inc, 2013).

The idea of teaching crystallography basics in high school is not new; there are several papers already published that outline lesson plans and laboratory experiments for this very purpose (Blattmann & Sticher, 2009; Luft *et al.*, 2010; Bethel & Lieberman, 2014). However, these plans often carry a significant cost barrier, given the high prices of lab kits and reagents. In this project, we sought to develop a method to teach crystallography using minimal resources, such as reagents that can be found in grocery stores and software that is free to the public. We show that lysozyme can be crystallized with items available in a grocery store. The methods can be carried out safely from home, which has become vital in the wake of the COVID-19 pandemic and the necessity of remote learning. This project would be very advantageous to home-schooled students. Measurements are also given in kitchen measuring spoons to help students perform this experiment at home.

Lysozyme is a small enzyme that protects us from bacterial infection by destroying the protective cell walls of bacteria. Lysozyme is abundant in our tears, saliva, mucus and human milk. Lysozyme is easy to work with and is considered to have no chemical hazards, making it an attractive option for these projects. After successfully forming lysozyme crystals with common reagents in crystallization plates, we expanded the project to include co-crystallization and soaking with food dye. Co-crystallization and soaking, which are used to determine the structures of proteins bound to ligands, are methods commonly used in drug design (Anderson, 2003). This can provide an additional connection of the project to current scientific research. In the hopes of making this experiment even more accessible, we also crystallized lysozyme without crystallization plates, in common containers such as contact lens cases and paint pots. The relative success of this method means that lysozyme powder is the only necessary ingredient for this project that could not easily be found at a grocery or drug store.

Finally, we solved the structure of lysozyme from crystals formed using the 'kitchen crystallization recipe', which allowed us to confirm the success of the crystallization method, as well as to directly connect protein visualization to this project. Most of the synchrotron facilities in the world have remote data collection capabilities that could allow data collection by the students from schools or homes. They also provide training for new users and offer support during data collection. This fosters student interaction and engagement with professional scientists. Information about various synchrotrons is provided in Table 1. The Stanford Synchrotron Radiation Lightsource (SSRL) has recently developed facilities for remote room-temperature data collection (https://www-ssrl.slac.stanford.edu/smb-mc/content/users/manuals/remote-access-at-elevated-temperatures-and-controlled-humidity) and designed shipping containers to ship crystals at room temperature or 5°C (Crystal Positioning Systems: https://www.crystalpositioningsystems.com/product-category/crystallography-at-room-temperature-controlled-humidity/; MiTeGen: https://www.mitegen.com/product/thermal-shipper/). Since lysozyme crystals are generally stable, students can ship their crystals at room temperature and collect the diffraction data from a computer terminal at their school or home. This circumvents the safety practices needed to cryocool crystals using liquid nitrogen for the synchrotron data collection.

Table 1
Synchrotron facilities around the world.

Name	Website
Advanced Light Source	https://als.lbl.gov/
Astrid2	https://www.isa.au.dk/facilities/astrid2/astrid2.asp
Advanced Photon Source	https://www.aps.anl.gov/
Canadian Light Source	https://www.lightsource.ca/
Cornell High Energy Synchrotron Source	https://www.chess.cornell.edu/
Deutsches Elektronen-Synchrotron DESY	https://www.desy.de/index_eng.html
ELETTRA	https://www.elettra.eu/
European Synchrotron Radiation Facility	https://www.esrf.fr/
LNLS	https://lnls.cnpem.br/
MAX IV	https://www.maxiv.lu.se/
National Light Source	https://www.bnl.gov/nsls2/
National Synchrotron Radiation Laboratory	https://en.nsrl.ustc.edu.cn/main.htm
PETRA III	https://www.desy.de/research/facilities__projects/petra_iii/index_eng.html
SESAME	https://www.sesame.org.jo/
SOLEIL	https://www.synchrotron-soleil.fr/en
SSLS	https://ssls.nus.edu.sg/
Stanford Synchrotron Radiation Lightsource	https://www-ssrl.slac.stanford.edu/
Swiss Light Source	https://www.psi.ch/en/sls
Taiwan Light Source	https://tls.nsrrc.org.tw/index.aspx?lang=en
Organization for world light sources	https://lightsources.org/

The availability of the solved structures in the Protein Data Bank (PDB) will allow students to download structures, analyze them and learn from them using the tools provided in this paper. New high school projects can be designed that focus on various aspects of enzyme function, understanding how antibiotics work, visualizing the mechanics of cell machinery, analyzing the details of virus structures *etc.* using the coordinates available in the PDB. This project enables students to ask questions, define problems, develop a research mentality, evaluate research methods and engage in evidence-based discussions. These are in line with the practices suggested by NGSS for science and engineering.

2. Genes to structure

Genetic material is passed down through DNA, but the way these genes are put into action is through proteins expressed by the information encoded in DNA (Berg *et al.*, 2002). Proteins are found in all organisms and are vital to many biological processes. They provide mechanical and structural support in cells, transport and store molecules, and work as catalysts, among other functions (Berg *et al.*, 2002). Proteins help us see, smell, hear and feel. Pathogens use proteins to attack us. Proteins fold into three-dimensional structures determined by hydrogen bonds between amino acids, van der Waals and hydrophobic interactions between their side chains,

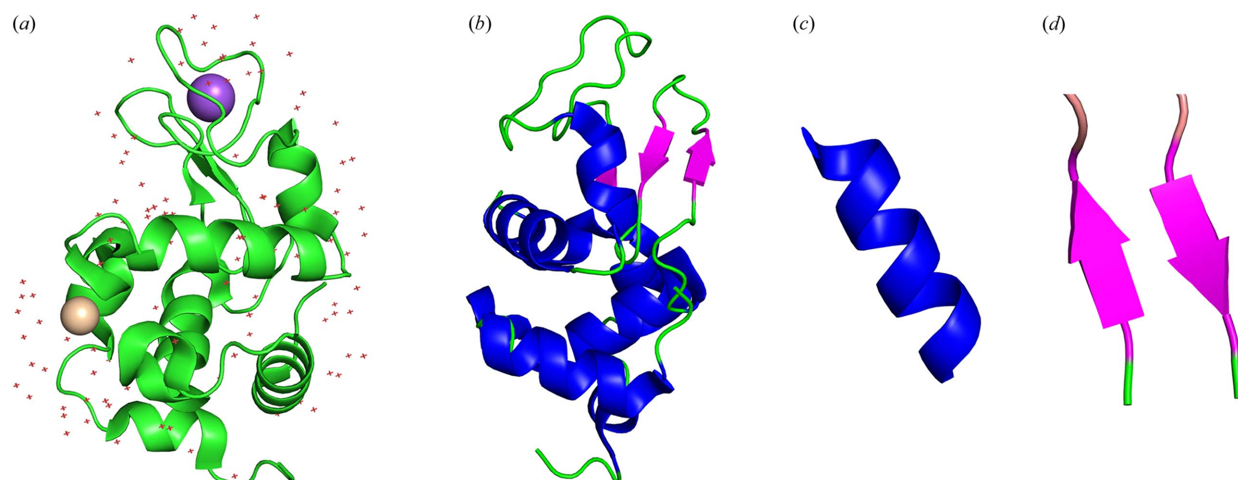


Figure 1
The structure of lysozyme. (*a*) Full structure including water molecules (red 'plus' shaped spots), a chlorine atom (wheat sphere) and a sodium atom (purple sphere). (*b*) Different view of the protein structure, with the secondary structure color coded: α-helices (blue), β-sheets (magenta) and loops (green). (*c*) Sample α-helix. (*d*) Sample β-sheet. The images were generated using the *PyMOL* software.

and interactions that the protein has in complex with other proteins, DNA or RNA. This final three-dimensional structure determines the protein function, which makes understanding protein structure vitally important to understanding proteins themselves (Fig. 1) (Rupp, 2009; Müller, 2017). The general shape of the protein is represented by a line or ribbon corresponding to the backbone and secondary structure features are represented by spirals (α-helices) and arrows (β-sheets) (Fig. 1).

Proteins are too small to be seen with visible light and we need to use a light that can see the atoms in the proteins. X-rays with wavelengths in the 1 Å range are ideal. One of the most prevalent ways of determining a protein structure is X-ray crystallography (Figs. 2 and 3). X-rays are diffracted by electrons, and the location of the electron cloud can be used to define the positions of atoms in a molecule. Crystallography makes use of crystals – three-dimensional periodic arrangements of molecules – for X-ray diffraction. Essentially, if an X-ray beam is directed at a crystal, the X-ray will scatter in a specific pattern depending on the molecules in that crystal. This diffraction pattern can be used, with some additional phase information, to reconstruct the electron density of a molecule. From there, the molecule itself can be reconstructed and analyzed (Sands, 1994; Hammond, 2015; Rupp, 2009).

The structural information gained from X-ray crystallography allows us to better understand biological processes on a molecular level, and can also inform drug design. Through structural analysis, researchers look for well defined binding pockets or active sites, and can design chemicals to alter the function of the enzyme. In some cases, this means temporarily attenuating the effect of human proteins. There are also drugs designed against pathogenic organisms, such as bacteria. In this case, medications are meant to inhibit necessary proteins in these microbes to the point where they cannot survive, all while leaving human proteins unaffected. In both cases, specificity is necessary to ensure the health of those taking medications, which underscores the importance of understanding structure when designing unique drug leads. Further experiments and structural analysis on this lead molecule can guide the design of life-saving medications (Anderson, 2003).

Crystallography starts with selecting a protein of interest and producing that protein using bacteria or other organisms. Given the restraints of an at-home laboratory, however, we will use a protein called lysozyme, which is readily available in a powder form (hen egg-white lysozyme, HEWL). We used HEWL purchased from Sigma–Aldrich (https://www.sigmaaldrich.com/US/en/sds/mm/4403-m). Lysozyme has a conserved fold and HEWL shares the same fold as human lysozyme. The purification of lysozyme from egg white using bentonite and crystallization has also been reported (Alderton *et al.*, 1945; Olieric *et al.*, 2007). All the crystallization experiments described in this paper were performed by a student in her apartment in Vermont (USA) using items from a local grocery store and the lysozyme powder and Cryschem plates shipped from SSRL, California (Hampton: https://hamptonresearch.com/product-Cryschem-Plate-84.html). In this project, we will go through all the steps of crystallography research that a scientist follows after obtaining an isolated protein: (1) protein crystallization (includes optional co-crystallization and soaking methods), (2) X-ray diffraction data collection, (3) phasing (to obtain the electron density),

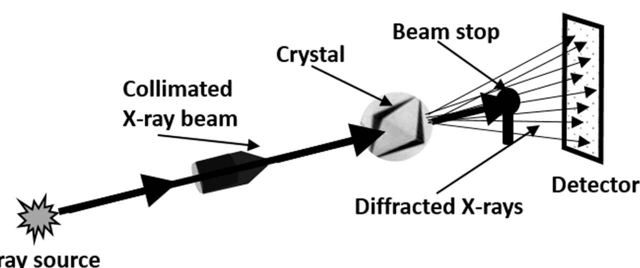


Figure 2
Schematic of the main components of an X-ray diffraction experiment.

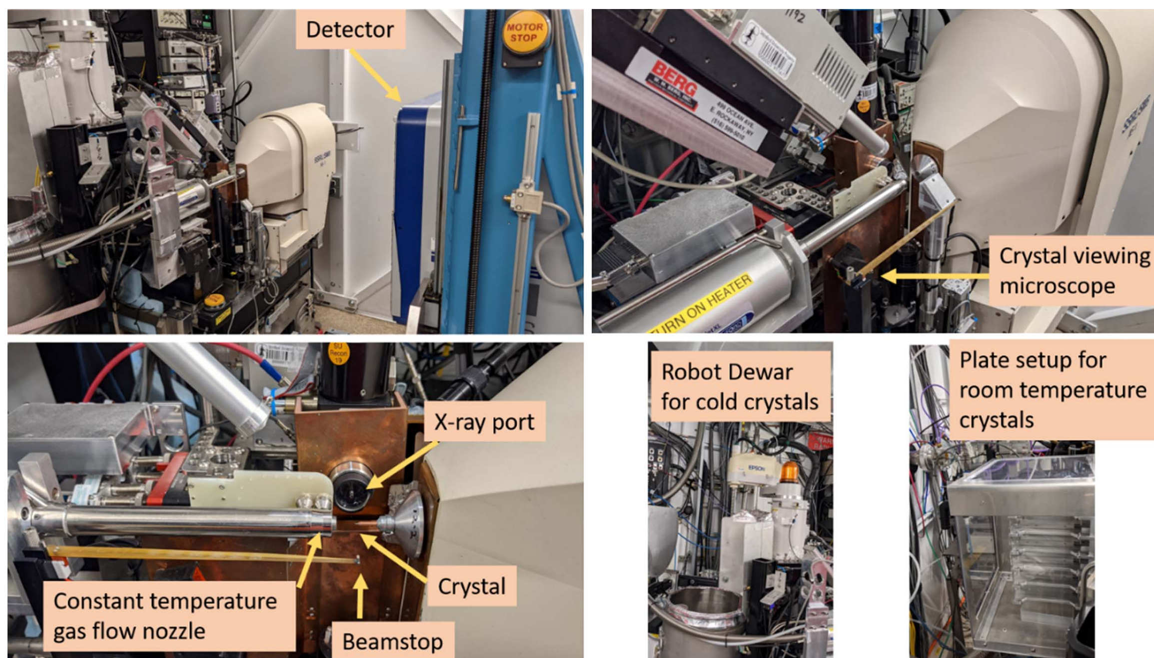


Figure 3
Synchrotron beamline setup (SSRL) showing the various modules used for the diffraction experiment.

(4) model building and refinement, and (5) functional interpretation.

Though X-ray diffraction requires a high-powered X-ray beam, many components of crystallography can be achieved without high-tech equipment.

3. Protein crystallization

The X-ray signal from a single protein molecule is weak and we need to pack the protein molecules in three dimensions to increase the signal. The three-dimensional packed array of protein molecules is called the crystal. When crystallizing a protein researchers look for the specific chemical conditions that will cause periodic arrangement to happen spontaneously. These conditions will differ from protein to protein, but generally each crystallization solution contains a salt, a precipitant and a buffer to set the pH. The lysozyme crystallization uses common table salt (NaCl) as the precipitant and sodium acetate as the buffer. Here, we react rice vinegar and baking soda to generate sodium acetate. Note that this is just one of the crystallization solutions that can be used to form lysozyme crystals. The reagent options for specific crystallization solutions are endless, and crystallographers usually try many crystallization solutions and run experiments to find which ones work best through an initial screening of new protein solutions (Hammond, 2015; Dessau & Modis, 2011).

Using these crystallization solutions, there are two main options to grow protein crystals, known as the hanging-drop method and the sitting-drop method. Both utilize the fact that the protein/precipitant mixture in the drop is less concentrated than that of the well solution, so water will evaporate out of the drop and into the well. As the water diffuses out of the drop, the concentration of the protein and the precipitant in the drop increases. This eventually brings the protein to a state called supersaturation, *i.e.* above its saturation limit. This means that its concentration is too high to maintain it all in solution, and the excess of protein will leave the solution to form a solid. If the chemical conditions are right, the protein molecules in the solid will be arranged in a three-dimensional structure, forming a crystal (Figs. 4 and 5). More detailed information on this process is available elsewhere (Hammond, 2015). Using a Cryschem well plate allows the sitting-drop method to be employed, and most alternatives will involve the hanging-drop method.

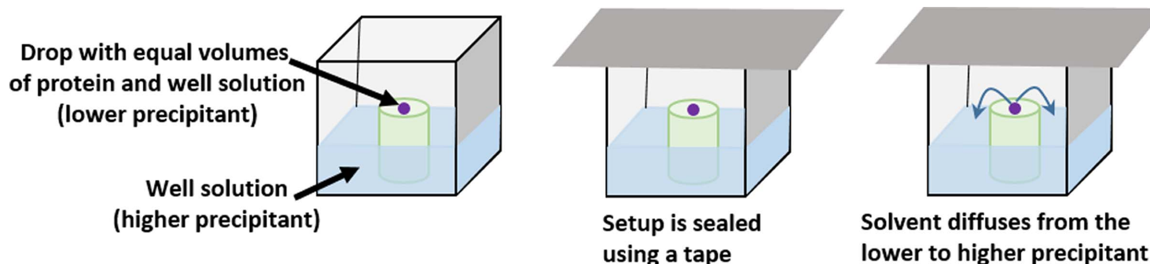


Figure 4
Schematic view of the crystallization process for the sitting-drop method.

Figure 5
Crystals formed in a sitting-drop crystallization plate over 72 h from 100 mg ml^{-1} lysozyme solution and a well containing a solution of 3.0 *M* sodium chloride, 0.5 *M* baking soda and 0.5 *M* rice vinegar. Scale bar is shown with a black line. Crystal photographs were taken with a STPCTOU wireless digital microscope (50X-1000X 1080P) purchased from Amazon.

3.1. Materials

Materials used included: table salt (NaCl), rice vinegar, baking soda, distilled water, lysozyme, food dye (co-crystallization and soaking methods only), paper towels or coffee filters, tablespoons (1/2 and 1, where 1 tbsp = 15.0 ml), teaspoons (1/8, 1/4, 1/2 and 1, where 1 tsp = 5.0 ml), Cryschem M-Microplate crystallization plate, 0.5 ml Pasteur pipettes (online retailers like Amazon sell inexpensive disposable plastic transfer pipettes), clear packing tape, weighing balance (milligram scale with USB supply, NEWACALOX reloading scale 100 × 0.001 g, high-precision portable multifunction lab powder scales with calibration tare weights, tweezers, weighing pans).

3.2. Crystallization

(1) Prepare 3.0 *M* sodium chloride by adding 1/2 tbsp of table salt (10 g m) to 3 tbsp of distilled water (45 g m). Mix well and filter the solution using a clean paper towel or coffee filter. All filtration steps were carried out by placing the paper towel/coffee filter in a conical funnel.

(2) Prepare 0.5 *M* baking soda by dissolving 1/4 tbsp of baking soda (3.3 g) in 2 tbsp of distilled water (30 g). Mix and leave for several minutes to completely dissolve the baking soda. Mix well and filter the solution using a paper towel (pH 9.6). There will be some undissolved impurity.

(3) Prepare 0.5 *M* rice vinegar by adding 2 tbsp of rice vinegar (30 g) to 1 and 1/8 tbsp of water (16.9 g). Mix well (pH 2.7). Vinegars successfully tested: Trader Joe's Rice Vinegar, Nakamo Rice Vinegar (Safeway) and Organic White Vinegar (Whole Foods).

(4) Prepare sodium acetate by mixing an equal volume of 0.5 *M* baking soda and 0.5 *M* rice vinegar in a beaker. The solution will be bubbling. Do not close the container while the solution is bubbling. Keep it for several minutes and mix well by rotating the container/test tube (pH 7.0).

(5) Prepare the crystallization solution by mixing 1/8 tsp of 3 *M* NaCl (0.7 g), 1/8 tsp of sodium acetate (0.6 g) and 1/4 tsp of water (1.25 g). This solution can be used to fill two wells.

(6) Prepare 100 mg ml^{-1} lysozyme by dissolving 10 mg of lysozyme in 0.1 ml of water. Low-cost milligram scale weighing balances are available to buy from online retailers like Amazon. Mix the lysozyme powder carefully to ensure the solid lysozyme is fully dissolved. Once the lysozyme is in solution, keep it on ice. We also successfully tested the crystallization of lysozyme kept at room temperature. However, it is better to keep lysozyme solution on ice if it needs to be stored for more than 8 h. If the lysozyme solution is more than 2 days old, it seems to generate crystals with rounded corners.

(7) Each well in the Cryschem crystallization plate contains an outer well in addition to a pillar with a smaller well at the center. To set a drop for crystallization, first fill the outer well with one pipette-full of crystallization solution. Next, add one drop of protein solution and one drop of crystallization solution to the small center well so that the volume of protein and crystallization solution in the well are equal.

(*a*) The Pasteur pipettes can be used to make these drops. First cut the bulb of the Pasteur in the middle. Seal the cut section with a finger and gently squeeze the bulb. Insert the tip into the well solution and release the pressure on the bulb. This will suck some fluid into the tip. Remove the finger to open the bulb and touch the tip in the well solution again. This keeps a small amount of solution in the tip. Move the tip into the center well of the Cryschem plate and release the solution by closing the pipette bulb with a finger and gently squeezing it. The same pipette can be used to put well solutions into other wells that have crystallization solution in the outer well. Trials with various people showed that the transfer of solution with a Pasteur pipette requires some practice. We suggest practicing the liquid transfer a few times with water. The 0.5 ml Pasteur pipette holds ~8 µl of solution in the tip. (Note: if the well and protein solutions are enough deep to dip the cut transfer pipette into the solution by at least 1.5 cm, dipping the cut transfer pipette into the solution will automatically get the solution into the pipette. Raising the tip will keep a small amount of solution in the tip.)

(*b*) Similarly, use a new cut pipette to drop protein into the center well. Try to avoid generating bubbles while transferring the solution to the drop. Use new pipettes for placing protein solution into other drops. This prevents contamination of the protein stock solution.

(*c*) If using an alternative well plate, first fill the well of the container (*e.g.* a contact lens case or paint pot) with 1/4 tsp (1.25 g) or ~2 full pipettes of crystallization solution. Next, add one drop of protein solution and one drop of crystallization solution to the inside of the cap of the container, so that the volumes of protein and crystallization solutions in the well are equal. Close the container carefully, ensuring that the drop does not fall into the well solution. Seal the wells with clear packing tape. Colorless lysozyme crystals should appear in 2–3 days.

Crystal mounting and shipping the crystals for data collection are not covered in this paper (Yamano, 2012; Pflugrath, 2015; video links are provided in the supporting information).

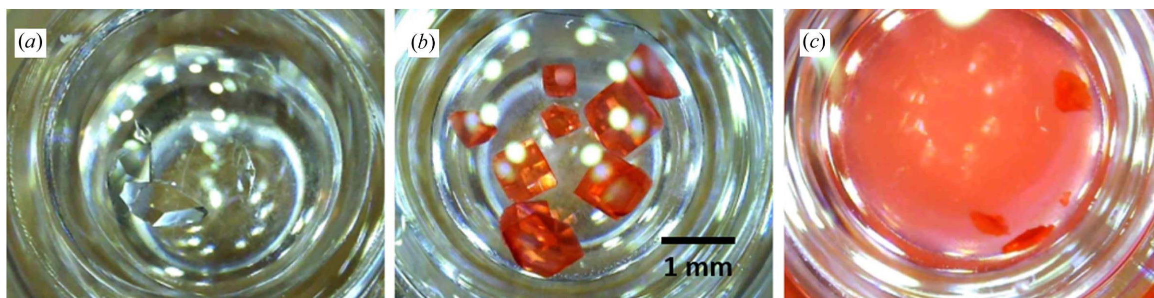


Figure 6
Crystals formed through (*a*) the basic crystallization method, (*b*) the soaking method and (*c*) the co-crystallization method. Images taken with a low-cost portable STPCTOU Wi-Fi USB digital microscope.

3.3. Co-crystallization and soaking

When researching the chemical interactions of proteins, crystallographers need to obtain diffraction data for proteins that are bound to ligands. This is especially necessary when it comes to drug design. When drug leads are created, they work in theory, but researchers must obtain crystallographic and other biochemical information for the drug molecule bound to the target protein in order to ensure that new molecules also work in practice (Anderson, 2003).

There are two main methods to form a crystal with a protein–ligand complex: co-crystallization and soaking. Co-crystallization refers to crystallizing the protein and the ligand together. Depending on the protein and ligand, this may mean that crystals are less likely to form. Soaking is the method of adding the ligand to the protein after the crystal has formed. With this method, there is some risk that the crystal will break due to the osmotic shock caused by the addition of the ligand solution. It is also possible that the protein molecules in the original crystal are packed in such a way that there are no open active sites available for the soaked ligands to bind to. Each method has advantages and drawbacks, so understanding both becomes an important aspect of X-ray crystallography (Müller, 2017).

Both co-crystallization and soaking can also be simulated using food dye as the ligand. Experiments with colored dye enable the students to gain a clear understanding of ligand-soaking experiments performed by scientists at pharmaceutical companies.

To add co-crystallization and soaking to this experiment:

(1) Follow basic crystallization steps (1)–(5) to make at least 7.5 tsp of crystallization solution.

(2) Split the crystallization solution into five 1.5 tsp portions. Put one drop of food coloring of each color into each of four of the five crystallization solutions. Leave one solution clear.

(*a*) Note: sizes of the drops of food dye may vary. Additional crystallization solution may be added to dyed crystallization solutions if the color appears too dark. The dyed solutions should generally not appear opaque, as too much food dye can dilute the crystallization solution and make it less effective.

(3) Follow steps (6)–(7) of the basic crystallization protocol.

(*a*) Use the clear crystallization solution for at least four wells.

(*b*) Use each of the different colored crystallization solutions in at least one well. (Note: the drop of crystallization solution in the center well should be the same as the crystallization solution in the outer well, *i.e.* if the outer well contains red crystallization solution there should be one drop of red crystallization solution and one drop of protein solution in the center well.)

(*c*) Seal the wells with clear packing tape.

(4) Dyed crystals should form within 3–4 days. Clear crystals should form within 2–3 days. Dyed crystals have a lower concentration of the precipitant and take longer to attain supersaturation.

(5) Once clear crystals have formed, add one drop of dyed crystallization solution to each crystal drop and re-cover with packing tape. Within 24 h, the crystals should have absorbed some of the dye and will no longer be clear (Fig. 6). The solution with red dye showed precipitation during soaking. However, crystals were not affected by the precipitation.

3.4. Alternatives to crystallization plates

Two containers that have been used successfully with the hanging-drop method are contact lens cases and paint pot strips (Fig. 7). It is very important to keep the containers air-tight to avoid evaporation of the drops. If the drops become dry, use grease on the contact areas with the lid (vacuum grease preferred). It was found that this method can also be used with co-crystallization and soaking procedures, but has shown lower levels of success with co-crystallization than the

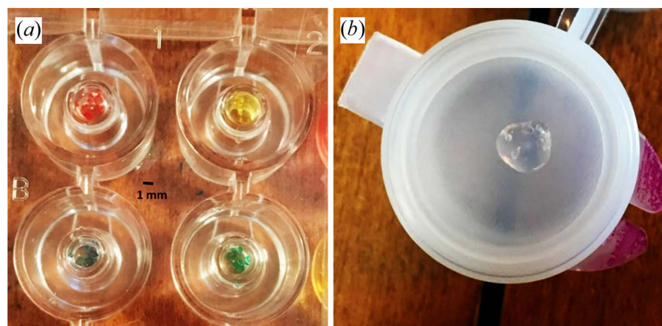


Figure 7
Representative images of (*a*) crystals formed in a sitting well crystallization plate and dyed using the soaking method, and (*b*) crystals formed in a paint pot using the hanging-drop method. Photographs taken with an iPhone SE (first generation).

crystallization plate. Additionally, the containers tested with this protocol do not constitute an exhaustive list. Other containers may be usable, provided they contain a well of some sort and a sealable cap. This provides students with another avenue to experiment and determine the best containers using the scientific method.

Crystals for X-ray data collection need to be cryo-cooled (see supporting information).

4. X-ray data collection

The electrons present in atoms interact with X-rays, so when a crystal is placed in a beam of X-rays, it will diffract them (Fig. 2). When a plane containing atoms diffracts X-rays, it produces spots in the data collecting detector. This can be conceptualized by visualizing the reflection spot from a mirror placed in sunlight. Diffraction from a protein crystal placed in the X-ray beam is shown in Fig. 8. The majority of the spots in this image come from a plane of atoms in the crystal. Data collection involves collecting hundreds of diffraction images of the crystal by rotating it and shining X-rays on different orientations of the crystal (Fig. 9). Most synchrotrons have remote data collection capabilities and training sessions to help users collect their data. The simulated beamline data collection setup at SSRL allows the user to learn all aspects of data collection using the data collection software in a simulated environment without a crystal (Smith *et al.*, 2010). Most synchrotron facilities have remote video viewing options for real time monitoring of the instruments during the diffraction experiment.

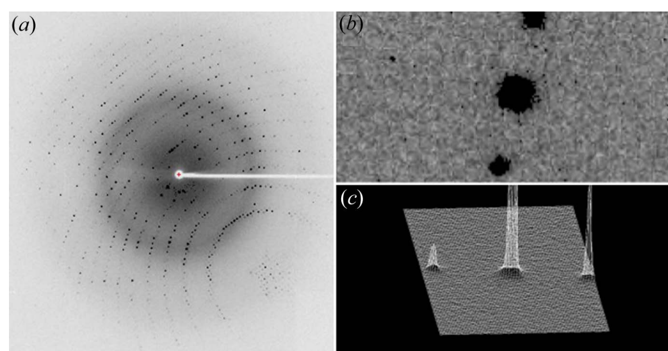


Figure 8
(*a*) Diffraction image from a crystal. (*b*) Zoomed-in view of spots. (*c*) Picture illustrating the darkness/intensity of the spots shown with a three-dimensional peak profile.

5. Phasing

All types of electromagnetic radiation are waves and have an amplitude and a phase. Both are required for the reconstruction of the enlarged images of the protein from diffracted rays. However, when we record diffracted X-rays using a detector, we lose the important phase information contained in the diffracted X-rays. In other words, the detector can only detect amplitudes, but not phases. There are several methods for generating the phase information (Sands, 1994; Hammond, 2015; Rupp, 2009; Crystallography 101). We used the

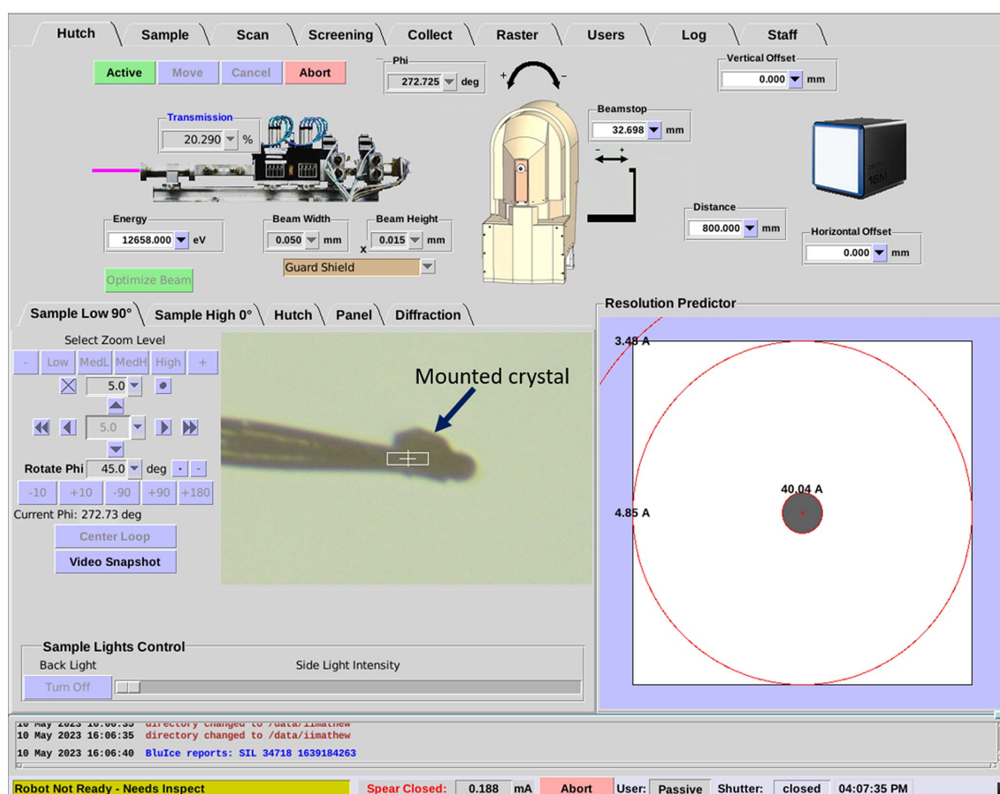


Figure 9
View of one section of the data collection software at SSRL. The crystal mounted for data collection is labeled.

Table 2
Freely available crystallography programs that are useful for this project.

Name	Website	Description
CCP4	https://www.ccp4.ac.uk/	Collection of various software for crystallography
Phenix	https://phenix-online.org/	Performs various steps in crystallography
Coot	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/	Graphics program for model building
PyMOL	https://pymol.org/2/	Graphics program

molecular replacement phasing method to construct a lysozyme structure from the diffraction images. Essentially, this means we used the structure of an identical molecule (PDB entry 193l; Vaney *et al.*, 1996) as a way to anchor our structure. This method works well when there is an already-known structure similar to the one being investigated. The successful phasing of diffracted data generates a three-dimensional view of the electron density in the crystal. Since atoms are located at the center of the electron cloud, the electron density map allows modeling/fitting of all atoms of the molecule present in the crystal. The availability of free crystallographic software and tutorials enables teachers to take these projects to a deeper level with hands-on refinement and model-fitting exercises (Table 2).

6. Model building and refinement

The model building and refinement consists of analyzing the electron density map and tracing the protein model in the map manually and by using computer programs. During this process, electron density maps are inspected for any deviations of the structure. Electron density maps are calculated using the experimental diffraction data. The electron density maps are colored depending on whether the the electron density in the fitted region is correctly or incorrectly modeled. Protein parts that are missing in the model are often displayed with green density and other regions with blue density. If the model contains atoms that are missing in the protein in the crystal, the electron density for this region will be colored red. Fig. 10(*a*) shows properly fitted arginine and tryptophan side chains in the blue electron density. The globular green density in this figure shows solvent water molecules that are not included at this stage of the refinement. Figs. 10(*b*) and 10(*c*) shows the electron density features for a misplaced Tyrosine side chain. A researcher would manually fix these issues and perform refinements before publishing the structure. Software programs like *Coot* and *PyMOL* could be used for visualizing the electron density, model building and making illustrations (Table 2). There are various software tools available from the PDB to evaluate the quality of the refined structure (Table 3). The structure of the lysozyme grown with the kitchen recipe presented in this paper has been deposited under PDB entry 7knk.

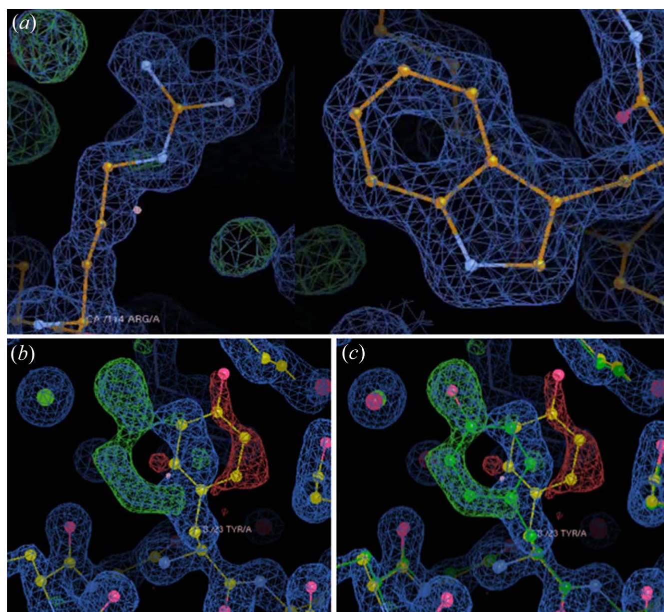


Figure 10
Electron density maps. (*a*) Refined structure in electron density. (*b*) A structure requiring manual fitting for the Tyrosine side chain. (*c*) Correct model of Tyrosine shown with a model colored in green. The images are generated using *Coot*.

7. Functional interpretation and drug design

The next phase of drug development is dependent on modeling/docking and databases searches, so the process scientists use and the process students use will closely resemble each other. Scientists will use the crystallographic data to try to determine the function and mechanism of the protein (Anderson, 2003; Rupp, 2009). In addition to crystallographic data from the unbound protein, data from co-crystallization with ligands, mutational studies and sequence comparisons are performed to help determine the function. The mutational studies, co-crystallization and ligand analog study data are combined and visualized using the same programs discussed above to provide a picture of the protein structure at different points in the reaction so that the scientist can piece together the enzyme mechanism and function. After the function is known, simulations are used to determine which ligands might be used as drugs to inhibit the function of the protein. The promising compounds are then crystalized with the protein and the process is repeated with kinetic and physiological data. Modifying these procedures for the classroom can be easily done so that this activity encompasses the entire rational drug design process and allows students to understand where drugs come from through use of scientific practices (screening of different compounds, studying effects of amino acid mutations *etc.*).

Mutational studies have many applications and access many ideas important in biochemistry that may go beyond the scope of this lesson. Mutational studies work by introducing point mutations (changing one amino acid to another amino acid) in the active site, so side chains that participate in or greatly stabilize the reaction are identified (Peng *et al.*, 2019). This works by systematically substituting side chains in the active site with a similar side chain and comparing the rate of product

Table 3
Free resources and software for docking and drug design.

Some have the same function and the best fit for your classroom will have to be determined.

Resource	Description	Notes and links
ZINC	Database of ligands	Missing many drugs currently in use; useful for turning this into a longer project (https://zinc.docking.org/)
Asinex	Database of ligands	Has many of the drug molecules; you need to download the entire file to access ligands (https://www.asinex.com/)
Protein Data Bank (PDB)	Database of proteins and resources for educators	Molecule of the month and other features provide many resources for understanding the database (https://www.rcsb.org/)
Proteins Plus/DoGSiteScorer	Visualize proteins and identify binding sites	There are many tools with good tutorials from the Center for Bioinformatics (https://proteins.plus/)
Autodock Vina	Docking simulation	Requires download on Windows, Linux or Mac (https://vina.scripps.edu/)
iGEMDOCK	Docking simulation	Uses PDB files to perform docking; requires download on Windows or Linux machine (http://gemdock.life.nctu.edu.tw/dock/igemdock.php)
DockThor	Docking simulation	Not as powerful as the other two docking simulators but does not require download (https://dockthor.lncc.br/v2/)
Openbabel	File format converter	Useful for converting files from those in databases to those accepted by docking simulators (https://openbabel.org/wiki/Main_Page)

formation with that of the original enzyme. If the rate is lower or the enzyme stops catalyzing the reaction completely, then the original side chain is integral to the mechanism and/or stability of the protein. If the activity is maintained, the protein complex with the compound might be crystallized to see if the mutated side chain is still interacting or if another side chain in the active site is substituting for the normal side chain function. There are resources for teaching about expressing desired genes with mutations in bacteria such as CU Engineering's *Introduction to genetic engineering and its applications* lesson (Zelisko & Anderson 2013). Mutational studies are also important when considering how the target organism might mutate to gain resistance to the drug. If common mutations are anticipated and designed against, the drug can be used effectively within the population for a longer time (Hao *et al.*, 2012).

Co-crystallization studies, on the other hand, can be assessed by students. To get an idea of how these functional determinations are carried out, refer to structures in the PDB from co-crystallization. For example, a signaling protein, cAMP-dependent protein kinase is activated by cAMP binding (Fig. 11) (Goodsell, 2012). The structures of bound and unbound protein kinase A can be compared to understand the changes caused by cAMP binding by looking at PDB entries 1j3h, 2cpk, 1atp, 1jlu and 1bx6 (Akamine *et al.*, 2003; Knighton *et al.*, 1991; Zheng *et al.*, 1993; Madhusudan *et al.*, 1994; Narayana *et al.*, 1999). Students might go through and systematically note differences, or a quick comparison can be used as a jumping off point for learning to look at ligand binding before moving on to ligand design (Fig. 11). Protein visualization tools that are built into the PDB should suffice for this exercise, and there is other visualization software available from the PDB website. To focus the inquiry, first any large-scale conformational changes should be noted and then individual amino acids in the binding site considered. They may form bonds or have varying strength intermolecular forces with the ligand which are important for the function. Though some nuances are likely to be lost, changes in bonding are readily apparent in these comparisons and are often the most important for the function.

Next, using these protein visualization tools, we can explore other proteins in the PDB and the way in which ligands bind to

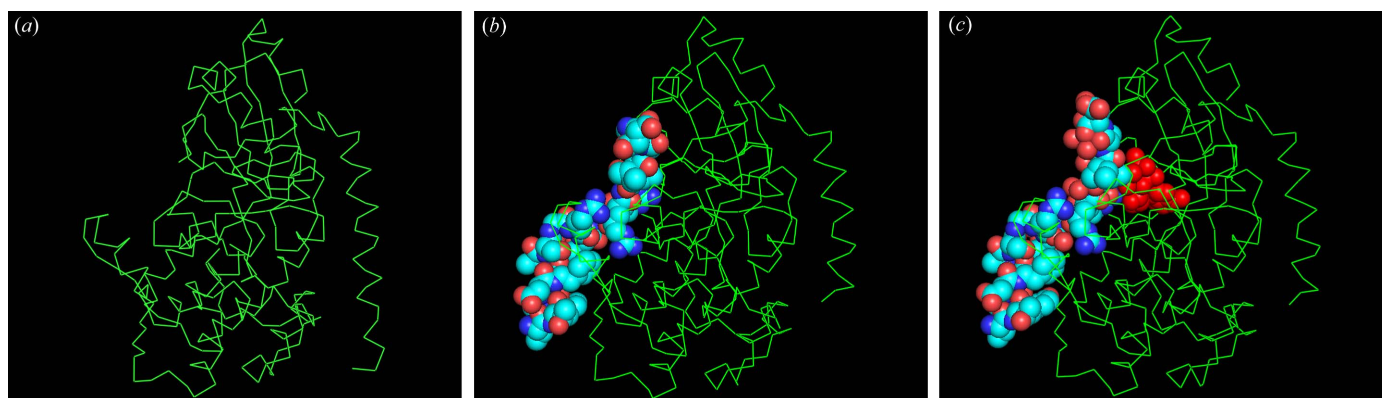


Figure 11
cAMP-dependent protein kinase A in (*a*) the unbound form (1j3h), (*b*) bound to cAMP (2cpk), and (*c*) bound to cAMP and ATP (1atp). Students might notice that the top part of the protein curls in to bind the ligands. They might zoom in to the ATP in panel 3 and notice it has interactions with ions in the protein, which is a stabilizing interaction that is important to the mechanism.

them. Students can perform these procedures on any protein in the PDB, though some good options with both unbound protein structures and ligand-bound structures are SARS-CoV-2 spike protein (PDB entry 7dwz; Yan *et al.*, 2021) or penicillin binding protein (PDB entry 3zg7; Jeong *et al.*, 2013), which is the target of the β-lactam antibiotics. Once students have chosen the protein they would like to investigate, they can use the *Proteins Plus* tools to add hydrogen atoms (most PDB structures do not have hydrogens) and find good binding pockets. The *DoGSiteScorer* (Volkamer *et al.*, 2012) finds binding pockets and assigns them a drug-ability score which students can use to try to design ligands and dock them to the protein (Table 3). The PDB educational portal, PDB-101, has many resources for framing a lesson about how drugs work.

There are many possibilities for ligand design and docking simulation. To assess whether a ligand is promising, the free energy in a docking simulation and the specificity of the ligand are checked. Students can use the Asinex or ZINC databases to download ligand files to load into one of the docking programs listed in Table 3. Then docking can be carried out and the free energy used to assess the binding to the target protein. A lower energy corresponds to tighter binding and in general a better drug, though students should think about why this might not always be the case. Additional resources can be found by a quick web search and in the work of Szarecka & Dobson (2019).

In a professional laboratory, the drug design process can take years and must also consider potential interference with human proteins (Anderson, 2003). In the classroom, multiple iterations of modeling could be carried out to get an idea of how to change ligands to better fit a target. Possibilities to engage students include having a competition for the best fit or choosing target proteins that are more personal to the students. After completing the simulations, students could then consider the quality of their results (inspect free-energy values and the interactions with the protein) and what their next steps would be if they had the research laboratory resources. There are more powerful paid versions of docking and visualization software available to researchers (Table 3), which also incorporate sequencing techniques and the ability to synthesize novel ligands. Engaging students through multiple channels to consider protein–ligand interactions, such as interacting with PDB files, conducting docking studies and representing proteins in different ways, allows students to engage with the scientific process, a goal of NGSS.

8. Results and discussion

The methods presented here showed successful crystal growth within 24 h, successful soaking of crystals in food dye within 24 h and moderately successful co-crystallization with food dye within 96 h of drops being set (Figs. 5–7). In most of our experiments, the crystals formed could be seen with the naked eye, negating the need for an expensive microscope. This was especially true when soaking with food dye was added to the experiment, as the colored crystals could be seen more easily (Figs. 6 and 7). However, some experiments, particularly the co-crystallization experiments, did produce crystals that were smaller and more difficult to detect without a microscope [Fig. 6(*c*)]. Although there is no perfect solution to this problem, it is likely that other magnification devices could be used if necessary. Even something as simple as a magnifying glass could allow students to detect smaller crystals and get a better view of the larger ones.

In this project, we go through the steps of X-ray crystallography, focusing on a hands-on crystallization laboratory and the visualization of finished protein structures. The intermediate steps of X-ray diffraction, phasing, and model building and refinement are described to add context to the experiment. In this way, educators can choose which parts of the project to emphasize. If the focus is more on crystallization, students can calculate their own dilutions and optimize the procedure. This is especially compelling because we did not optimize the hanging-drop methods or create an exhaustive list of common containers that could be used. This means that, rather than simply replicating our experiment, students can conduct their own research to find the conditions that work best in the common containers of their choosing. Examples of conditions to alter include the concentration of the protein solution, the concentration of salt in the crystallization solution and the pH of the buffer. Sample crystallization conditions for lysozyme can be found in the literature (Blattmann & Sticher, 2009; Dessau & Modis, 2011). Although the specific crystallization solutions differ from our method, these pH and salt concentrations can be used as a starting point for research. Increased access to synchrotron facilities around the world, remote access capabilities, and remote training and learning tools at these facilities enable students to gain valuable experience at these facilities with their projects and interact with scientists. These experiments expose students to these advanced facilities and help them engage in evaluating, communicating and analyzing data with scientists and learning from them.

The focus of this project can also be on three-dimensional models, of visualizing how structure determines function and on the possibilities of drug design. This method has the appeal of requiring nothing but a computer, making it free and easy to conduct in-person or remotely. The focus can also be on all of these aspects equally, as one cohesive project in which students crystallize a protein and then work with the structure of that same protein, crystallized using the same method they used. The lysozyme data collected from HEWL can be solved using the PDB coordinates for the turkey egg lysozyme (PDB entry 135l). The turkey egg lysozyme has seven mutations compared with HEWL, and students could use the electron density map to view and model build the mutated residues. There are also many creative ways to format deliverables. Artistic representations of results, academic writing or posters, gallery walks, and oral presentations can all be tailored to the delivery method and student needs (Marchak *et al.*, 2021). In these unprecedented times, keeping possible projects open-ended and flexible is more important than ever, so that at every juncture they can be tailored to fit the needs of each school and, further, each individual student. Regardless of

which parts of the project are emphasized, students should gain a more in-depth understanding of three-dimensional protein structures and the techniques that professional scientists use to determine them.

9. Conclusions

The primary emphasis of the project is to help high school students gain experience by conducting experiments carried out by scientists at major universities/facilities. We presented a method for growing crystals of lysozyme using materials available in a grocery store, which are safe and low cost. We also present protocols for performing the crystallization using tools available at home or school, increasing accessibility of these experiments and allowing for more remote learning options, which can bring science to underrepresented communities in STEM (science, technology, engineering and mathematics).

This study collects freely available tools that can be used by teachers and students to perform protein functional analysis. The students are exposed to advanced topics in enzyme structure and function by visualizing the active sites of various enzymes. We provide protocols and tools to engage the students in activities and help them understand the concept of structure-based drug design. The study also describes various steps for teachers and students to generate their own projects involving crystallization, enzyme functional studies and drug design. We describe how to connect and use the Protein Data Bank, which is an archive of experimentally determined three-dimensional structures of proteins, nucleic acids and other biological macromolecules.

We describe processes to bring a scientific experiment to a high school setup, and the methods presented offer ways to bring creative and active learning strategies to classrooms. The information presented in this work also provides tools for students to perform experiments in a remote format. High school teachers can use freely available software to teach difficult concepts of protein function by engaging students in the various projects outlined here. Teachers and students could have a project-based discussion of some disease targets, identify the enzymes involved, and download and visualize structures if they are available in the PDB. Depending on the level of the projects, they could dock compounds and try structure-based drug design approaches on the enzyme. These methods will allow students to learn about amino acids, understand their function, and expand their understanding to enzyme structure and function all with safety and accessibility in mind. Schools could collaborate with synchrotron facilities to expose students to advanced scientific methods.

The development of the hands-on experiments for students described in this paper could generate interest and motivation to make other advanced scientific concepts understandable and/or doable for general students.

Acknowledgements

ALF thanks the SULI program of the US Department of Energy (DOE) and ART thanks California Polytechnic State University's STAR program for support. We thank Michael Soltis, Enrique Cuellar, Aina Cohen, Doukov Tzanko, Silvia Russi, Clyde Smith, Nicole Terranova and Eddie Snell for the critical reading of manuscript. We thank Silvia Russi, Darya Marchany-Rivera, Ron Reyes and Renato Avelar for testing the crystallization. We thank Jim Pflugrath for providing the video links for crystal mounting. The authors want to thank the referees for their valuable suggestions that helped to improve the manuscript, and we specifically thank Dr Claude Sauter for his detailed and careful evaluation of the manuscript. We also thank members of the SMB group at SSRL for helpful discussions and support. The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of NIGMS or NIH.

Funding information

Use of the Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, is supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences (contract No. DE-AC02-76SF00515). The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research, and by the National Institutes of Health, National Institute of General Medical Sciences (contract No. P30GM133894). We also thank the STAR program from California Polytechnic State University (scholarship awarded to ALF and ART).

References

Achieve Inc. (2013). *Next Generation Science Standards.* Washington, DC: National Academies Press.

Akamine, P., Madhusudan,, Wu, J., Xuong, N.-H., Ten Eyck, L. F. & Taylor, S. S. (2003). *J. Mol. Biol.* **327**, 159–171.

Alderton, G., Ward, W. H. & Fevold, H. L. (1945). *J. Biol. Chem.* **157**, 43–58.

Anderson, A. (2003). *Chem. Biol.* **10**, 787–797.

Berg, J. M., Tymoczko, J. L. & Stryer, L. (2002). *Biochemistry*, 5th ed. New York: W. H. Freeman.

Bethel, C. & Lieberman, R. (2014). *J. Chem. Educ.* **91**, 52–55.

Blattmann, B. & Sticher, P. (2009). *Sci. School*, **11**, 30–36.

Dessau, M. A. & Modis, Y. (2011). *J. Vis. Exp.* **47**, 2285.

Garcia-Ruiz, J. M., Otalora, F., Garcia-Caballero, A., González-Ramírez, L. A. & Verdugo-Escamilla, C. (2015). *J. Appl. Cryst.* **48**, 1264–1275.

Goodsell D. (2012). *cAMP-Dependent Protein Kinase (PKA)*, https://pdb101.rcsb.org/motm/152.

Hammond, C. (2015). *The Basics of Crystallography and Diffraction*. Oxford University Press.

Hao, G., Yang, F. & Zhan, C. (2012). *Drug Discov. Today*, **17**, 1121–1126.

Jeong, J., Kim, Y., Rojviriya, C., Ha, S., Kang, B. S. & Kim, Y. (2013). *Antimicrob. Agents Chemother.* **57**, 3507–3512.

Knighton, D. R., Zheng, J. H., Ten Eyck, L. F., Ashford, V. A., Xuong, N. H., Taylor, S. S. & Sowadski, J. M. (1991). *Science*, **253**, 407–414.

Luft, J. R., Furlani, N. M., NeMoyer, R. E., Penna, E. J., Wolfley, J. R., Snell, M. E., Potter, S. A. & Snell, E. H. (2010). *J. Appl. Cryst.* **43**, 1189–1207.

Madhusudan, Trafny, E. A., Xuong, N. H., Adams, J. A., Ten Eyck, L. F., Taylor, S. S. & Sowadski, J. M. (1994). *Protein Sci.* **3**, 176–187.

Marchak, D., Shvarts-Serebro, I. & Blonder, R. (2021). *J. Chem. Educ.* **98**, 2809–2819.

Müller, I. (2017). *Acta Cryst.* D**73**, 79–92.

Narayana, N., Diller, T. C., Koide, K., Bunnage, M. E., Nicolaou, K. C., Brunton, L. L., Xuong, N. H., Ten Eyck, L. F. & Taylor, S. S. (1999). *Biochemistry*, **38**, 2367–2376.

Olieric, V., Schreiber, A., Lorber, B. & Putz, J. (2007). *Biochem. Mol. Biol.* **35**, 28–286.

Peng, Y., Alexov, E. & Basu, S. (2019). *Int. J. Mol. Sci.* **20**, 548.

Pflugrath, J. W. (2015). *Acta Cryst.* F**71**, 622–642.

Rupp, B. (2009). *Biomolecular Crystallography: Principles, Practice, Application to Structural Biology*, 1st ed. New York: Garland Science.

Sands, D. E. (1994). *Introduction to Crystallography*. Mineola: Dover Publications.

Smith, C. A., Card, G. L., Cohen, A. E., Doukov, T. I., Eriksson, T., Gonzalez, A. M., McPhillips, S. E., Dunten, P. W., Mathews, I. I., Song, J. & Soltis, S. M. (2010). *J. Appl. Cryst.* **43**, 1261–1270.

Szarecka, A. & Dobson, C. (2019). *Am. Biol. Teach.* **81**, 423–429.

Vaney, M. C., Maignan, S., Riès-Kautt, M. & Ducruix, A. (1996). *Acta Cryst.* D**52**, 505–517.

Volkamer, A., Kuhn, D., Grombacher, T., Rippmann, F. & Rarey, M. (2012). *J. Chem. Inf. Model.* **52**, 360–372.

Yamano, A. (2012). *Rigaku J.* **28**(2), 1–4.

Yan, R., Zhang, Y., Li, Y., Ye, F., Guo, Y., Xia, L., Zhong, X., Chi, X. & Zhou, Q. (2021). *Cell Res.* **31**, 717–719.

Zelisko, M. & Anderson, K. (2013). *Introduction to Genetic Engineering and Its Applications*, https://www.teachengineering.org/lessons/view/uoh_genetic_lesson01.

Zheng, J., Trafny, E. A., Knighton, D. R., Xuong, N., Taylor, S. S., Ten Eyck, L. F. & Sowadski, J. M. (1993). *Acta Cryst.* D**49**, 362–365.