

# Synchrotron-based macromolecular crystallography module for an undergraduate biochemistry laboratory course

Kyle M. Stiers,<sup>a‡</sup> Christopher B. Lee,<sup>a‡</sup> Jay C. Nix,<sup>b</sup> John J. Tanner<sup>a\*</sup> and Lesa J. Beamer<sup>a\*</sup>

Received 19 August 2016

Accepted 19 October 2016

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

‡ These authors contributed equally to this work.

**Keywords:** undergraduate education; macromolecular crystallography; synchrotron sources; remote data collection.

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

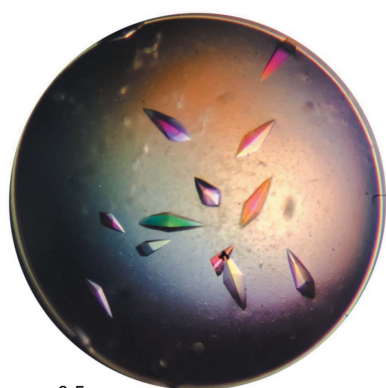
<sup>a</sup>Biochemistry Department, University of Missouri, 117 Schweitzer Hall, Columbia, MO 65211, USA, and <sup>b</sup>Molecular Biology Consortium, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. \*Correspondence e-mail: tannerjj@missouri.edu, beamerl@missouri.edu

This paper describes the introduction of synchrotron-based macromolecular crystallography (MX) into an undergraduate laboratory class. An introductory 2 week experimental module on MX, consisting of four laboratory sessions and two classroom lectures, was incorporated into a senior-level biochemistry class focused on a survey of biochemical techniques, including the experimental characterization of proteins. Students purified recombinant protein samples, set up crystallization plates and flash-cooled crystals for shipping to a synchrotron. Students then collected X-ray diffraction data sets from their crystals *via* the remote interface of the Molecular Biology Consortium beamline (4.2.2) at the Advanced Light Source in Berkeley, CA, USA. Processed diffraction data sets were transferred back to the laboratory and used in conjunction with partial protein models provided to the students for refinement and model building. The laboratory component was supplemented by up to 2 h of lectures by faculty with expertise in MX. This module can be easily adapted for implementation into other similar undergraduate classes, assuming the availability of local crystallographic expertise and access to remote data collection at a synchrotron source.

## 1. Introduction

Technological developments over the past several decades in molecular biology and X-ray diffraction instrumentation have enabled an explosion of structural information on macromolecules, especially proteins (Berman *et al.*, 2013; Helliwell & Mitchell, 2015). However, while the principles of protein structure are now commonly taught in undergraduate classrooms, student exposure to the methodologies used to obtain three-dimensional structural information [*e.g.* macromolecular crystallography (MX) and multi-dimensional NMR] on macromolecules is rare. It is well understood that insight into such techniques is critical for an appreciation and understanding of structural data, their applications and their limitations (Jaskólski, 2001; Faust *et al.*, 2008, 2010). It is also important to prepare undergraduate students in the life sciences for the many potential career directions that will utilize such information. In addition to graduate school, those planning careers in medicine, pharmaceutical sciences, biotechnology, patent law and many other sectors in technology/engineering are potential future users or beneficiaries of three-dimensional structural data.

As a discipline, crystallography offers a variety of excellent educational resources that benefit the community, including downloadable teaching pamphlets and links to various online tutorials. These are easily accessible through the educational sections of the web sites of professional organizations,



**Table 1**

Beamlines listing remote data collection capability on their web sites.

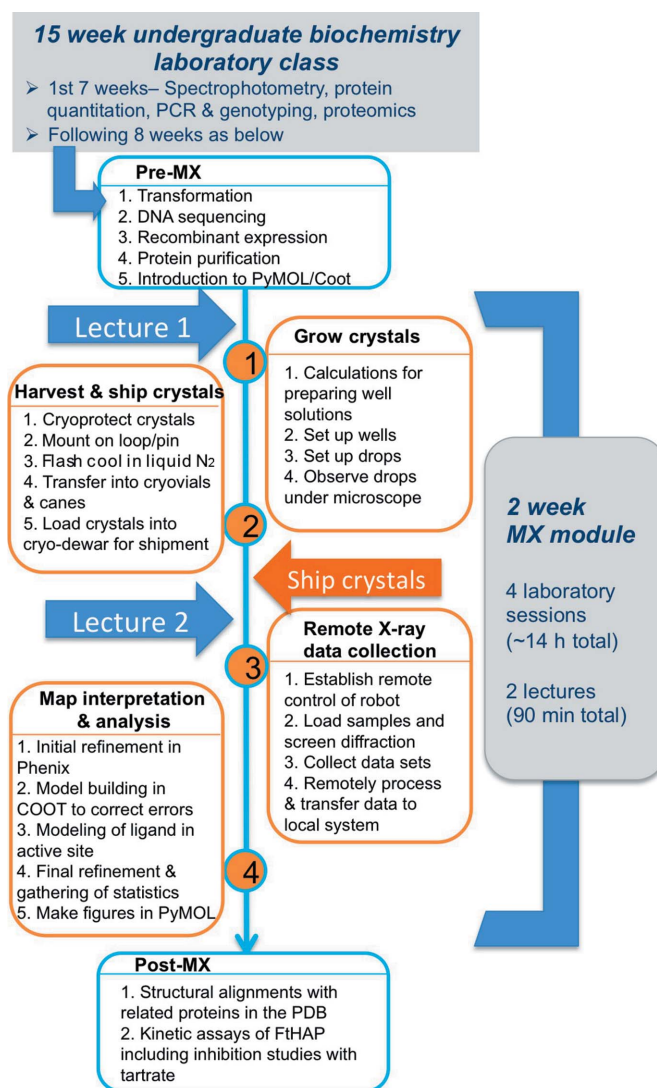
Synchrotron	Beamlines with remote access	URL
ALS – Berkeley, CA, USA	4.2.2, 5.0.1, 5.0.2, 5.0.3, 8.2.1, 8.2.2, 12.3.1	<a href="http://als.lbl.gov/">http://als.lbl.gov/</a>
APS – Argonne, IL, USA	17-ID-B, 19-ID-D, 21-ID-D, 21-ID-G, 22-BM-D, 22-ID-D, 23-ID-B, 23-ID-D, 24-ID-C, 24-ID-D	<a href="https://www1.aps.anl.gov/">https://www1.aps.anl.gov/</a>
CHESSE – Ithaca, NY, USA	A1, F1	<a href="http://www.chess.cornell.edu/">http://www.chess.cornell.edu/</a>
NLSL II – Upton, NY, USA	17-ID-1	<a href="https://www.bnl.gov/ps/">https://www.bnl.gov/ps/</a>
SSRL – Stanford, CA, USA	7-1, 9-1, 9-2, 11-1, 12-2, 14-1	<a href="http://smb.slac.stanford.edu/">http://smb.slac.stanford.edu/</a>
Australian Synchrotron – Melbourne, Australia	MX1	<a href="http://synchrotron.org.au/">http://synchrotron.org.au/</a>
ESRF – Grenoble, France	ID30A-1 / MASSIF-1	<a href="http://www.esrf.eu/MASSIF1">http://www.esrf.eu/MASSIF1</a>
CLS – Saskatchewan, Canada	08B1-1, 08ID-1	<a href="http://www.lightsource.ca/">http://www.lightsource.ca/</a>
Diamond – Didcot, UK	I02, I03, I04, I04-1	<a href="http://www.diamond.ac.uk/Home.html">http://www.diamond.ac.uk/Home.html</a>

including the American Crystallographic Association and the International Union of Crystallography. A number of intensive on-site educational courses are also regularly offered to help train young crystallographers (*e.g.* CCP4 workshops, the RapiData course in Stanford, M2M in Hamburg, X-ray methods courses at Cold Spring Harbor *etc.*). However, these resources are largely geared towards those conducting (or intending to conduct) crystallographic research, such as graduate or postdoctoral students. Typical undergraduate students, especially those in the life sciences, do not have sufficient time, background knowledge or interest to engage in such rigorous studies.

Practical considerations also constrain crystallographic experiments at the undergraduate level, which are typically limited to crystallization screens and possibly a diffraction experiment, if a home X-ray system is available (Gražulis *et al.*, 2015). If X-ray diffraction data are collected, this will most likely be done using crystals of small molecules, as part of a chemistry class (Pett, 2010; Aldeborgh *et al.*, 2014; Campbell *et al.*, 2015), rather than using crystals of biological macromolecules, which are less robust. Other features inherent to MX serve as additional barriers to undergraduate participation. These may include the lack of appropriate faculty MX expertise, issues related to the cost, fragility and safety of the instrumentation, and limited availability of data collection time on home X-ray systems. Consequently, it is often impractical (and potentially expensive) to introduce young inexperienced students and their samples into a structural biology X-ray facility, even though these exist at most research institutions.

The advent of remote synchrotron data collection and fast detectors has the potential to transform the teaching of MX at the undergraduate level. Freely available at multiple synchrotron facilities around the world (Smith *et al.*, 2010; Grochulski *et al.*, 2012; Beteva *et al.*, 2006; Warren *et al.*, 2008) (Table 1), remote data collection overcomes many of the potential barriers to student access. After crystals have been grown in the teaching laboratory, they can be shipped to a synchrotron, where remotely controlled robotic systems handle the samples without the need for intensive support by beamline personnel. Complete data sets collected with pixel array and complementary metal–oxide semiconductor (CMOS)-based detectors can often be obtained in a few minutes, allowing potentially dozens of data sets to be

collected and processed within the time frame of a typical 3 h laboratory class. Use of automated scripts for data processing provides rapid access to finalized data sets, which can be easily transferred back to the teaching laboratory. Thus, in addition



**Figure 1**  
Overview of the MX module and how it fits into the one-semester laboratory course. Orange boxes show the activities performed during the four laboratory meetings of the MX module. Blue boxes show the activities performed before and after the MX module.

to its benefits to researchers, remote synchrotron data collection can be a powerful tool in the teaching of MX to students either at an introductory level, as shown in this case, or involved in more advanced crystallographic studies.

Here we describe the integration of a 2 week introductory MX module (Fig. 1) into an existing senior biochemistry laboratory course at the University of Missouri, made possible through the use of remote synchrotron data collection. The module was conceived as an overview of the key experimental steps in MX, as a comprehensive treatment was beyond the scope and time frame of the class. Two faculty members (Beamer and Tanner) with active research programs in protein crystallography were responsible for module design and experimental oversight. Implementation of the methodology and day-to-day operations of the class were organized by the course instructor (Lee) and a graduate student teaching assistant (Stiers). Synchrotron data collection was arranged *via* the University of Missouri's membership in the Molecular Biology Consortium (beamline 4.2.2) at the Advanced Light Source (ALS) with the enthusiastic cooperation and assistance of the beamline 4.2.2 director (Nix). This introductory module serves as one example of many possible ways that new technology can facilitate the teaching of MX at the undergraduate level, within the constraints of the existing curricula.

## 2. The undergraduate laboratory class

### 2.1. Course structure and scope

The introductory MX module was implemented in an existing senior laboratory course required for all biochemistry majors at our institution. It is a one-semester (15 week) writing-intensive capstone laboratory experience. Each laboratory section has a maximum enrollment of 16 students, to keep the student-to-instructor ratio low for the required applied instruction. Two 1 h lecture periods and four 3.5 h laboratory sessions (2 weeks of the 15 week semester) were used, in addition to outside preparation time by the teaching assistant. All of the experiments were performed in a standard biochemistry undergraduate laboratory facility.

The MX module was incorporated into a section of the laboratory class that is popular with students planning on pursuing graduate or professional schools. It surveys a number of biochemical techniques and also introduces the students to a variety of instrumentation (Fig. 1). During the semester, the basic principles of optical spectroscopy, chromatography, mass spectrometry proteomics and confocal microscopy are surveyed in the context of biochemical pathways and the relationships between macromolecules. The course is designed to be a fully integrated research experience, utilizing student-prepared samples whenever possible for hands-on experiments, data analysis and interpretation.

The protein selected for the MX unit was histidine acid phosphatase from *Francisella tularensis* (FtHAP), which had been previously crystallized and structurally characterized in the laboratory of one of the authors (Felts *et al.*, 2006; Reilly *et al.*, 2006). In addition to its robust crystals (see §3.1 below), FtHAP was convenient for integration into the existing

curriculum, which already included purification and enzymatic studies of another acid phosphatase, that from wheat germ. An expression vector for FtHAP was initially introduced into the class as part of a study requiring the students to perform DNA sequence analysis to identify the encoded gene. Subsequently, recombinant protein was expressed and purified as described in §3.1. Student-purified samples of FtHAP were used for SDS-PAGE to evaluate protein purity and also for protease digestion followed by peptide identification *via* MALDI-TOF (matrix-assisted laser desorption/ionization-time-of-flight) mass spectrometry. Specific activity was measured and used to compare and contrast native *versus* recombinant protein purification techniques (wheat germ acid phosphatase *versus* FtHAP). More detailed enzyme kinetic experiments were also conducted on FtHAP, examining a variety of parameters, including pH, temperature and the effects of inhibitory molecules. This work was done either prior to the 2 week MX module or following it (enzyme kinetics).

To prepare for the MX module, students were also introduced to the molecular graphics and modeling package *PyMOL* (DeLano, 2002). Simple tutorials were designed to give them experience of viewing and displaying protein structures.

### 2.2. Equipment and supplies

In addition to typical laboratory reagents and pipettes, students had access to stereo zoom microscopes, pins/loops for mounting crystals and plates for crystal growth (both hanging- and sitting-drop formats). Cold-resistant gloves, cryo-canons and tongs were also provided for handling crystals during the flash-cooling process. Several Dewars of various sizes were used for holding liquid nitrogen. Frozen crystals were shipped in a Taylor Wharton CX100 dry shipping Dewar (pre-cooled with liquid nitrogen) equipped with a set of Universal V1-Pucks (Uni-pucks) from MiTeGen, LLC (Ithaca, NY, USA). Each Uni-puck holds 16 crystals. For a comprehensive list of reagents and their sources, see Table S1 in the supporting information.

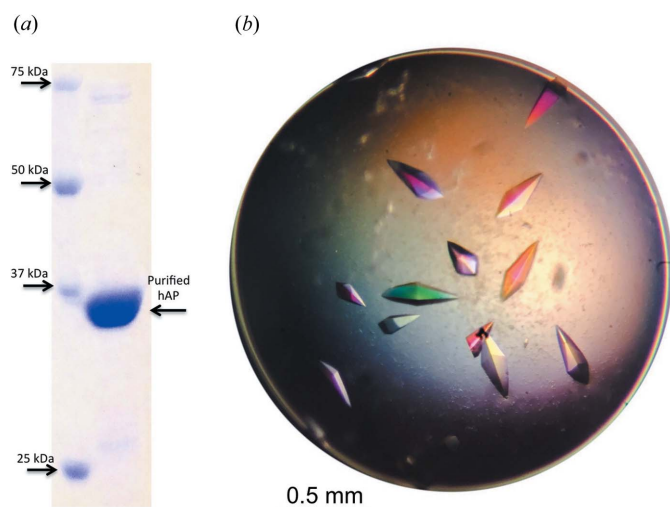
*Coot* (Emsley & Cowtan, 2004), *Phenix* (Adams *et al.*, 2010) and *PyMOL* were installed on eight laboratory laptops for structural analyses. Software for the remote data collection interface was run on a single laptop and projected onto a large screen in the laboratory, so all students could view the robot position, diffraction patterns *etc.*

We estimate the overall cost of consumables, including crystallization plates, reagents, cryogenics, pins and shipping expense, to be ~25 USD per student. Costs of the re-useable MX components (shipping Dewar, Uni-pucks *etc.*) would be substantially higher. However, some synchrotron beamlines have loaner puck programs.

## 3. Specifics of the MX unit

### 3.1. Protein expression, purification and crystallization

As noted above, FtHAP was selected for the MX module because of the ease of crystallization, the robustness of the



**Figure 2**  
 (a) SDS-PAGE of FtHAP purified according to the protocol described in the text. (b) Photograph of FtHAP crystals grown with the protocol used by the students. Crystals were grown as both hanging/sitting drops, from 2  $\mu\text{l}$  of protein (10 mg ml<sup>-1</sup>) and 2  $\mu\text{l}$  of well buffer solution, and typically appear overnight (see text). Crystals in this photograph (<0.5 mm) are representative of those used to collect the diffraction data in Table 2. Photograph by Kyle Stiers.

crystals and their high-resolution diffraction (Felts *et al.*, 2006; Singh *et al.*, 2009). Since its original crystallization, FtHAP has been utilized locally as a model system for new graduate students to learn the techniques of recombinant protein expression, protein purification, crystallization and X-ray diffraction. It was also convenient to adapt to the laboratory

class because its enzymatic activity is easily assayed using *p*-nitrophenylphosphate as the substrate (Reilly *et al.*, 2006), which was helpful during purification and for the kinetic studies noted above.

Purification and crystallization of FtHAP by students followed published protocols with several minor variations to simplify the procedure and make it compatible with the time constraints of a laboratory class. Briefly, frozen pellets of bacterial cultures were lysed with a French press and centrifuged and the supernatant collected. The supernatant was batch bound to Ni<sup>2+</sup> affinity resin, equilibrated for 30 min and eluted with 400 mM imidazole. The purified protein was dialyzed into final buffer of 20 mM sodium phosphate pH 7.4, 0.5 M NaCl, then flash-frozen and stored at -80°C (193 K). To avoid the extra laboratory time associated with concentration of the individual student samples for crystallization, a large batch of protein was prepared outside the class by the teaching assistant using a slightly modified protocol (see the supporting information) and concentrated to 10 mg ml<sup>-1</sup>. Aliquots of this protein, which had also been pre-tested for successful crystallization, were given to the students for crystallization screens; student-prepared protein (see Fig. 2a for SDS-PAGE) was used for all other experiments in the class, including kinetic analyses.

Crystallization conditions from Felts *et al.* (2006) were used to develop an optimized standard condition of 10% Tacsimate pH 7.0 (Hampton Research), 0.1 M HEPES-Na pH 7.0, 19% (w/v) PEG 3350 that reliably produced crystals (Fig. 2b). Variations around this standard condition were designed to demonstrate differences in crystallization related to Tacsimate

and PEG concentrations, pH, protein concentration and various additives. Students were provided with plates for hanging/sitting-drop vapor-diffusion experiments (Fig. 3), a written crystallization protocol for a 24-well screen (different for each team; see the appendix in the supporting information) and stock solutions (Table S2) for the well buffer components. Students were responsible for calculating the required concentrations of the various components needed to set up the screen.

A detailed protocol for protein expression and purification, as well as the four parts of the MX module, is available in the supporting information.

### 3.2. Flash-cooling of crystals and shipment to the beamline

FtHAP crystals appear within 1 d and typically reach maximal size in less than a week. Crystals were first observed by the students 2 d after the drops were set up. In this case, owing to an intervening spring break of 9 d, harvesting of



**Figure 3**  
 Photographs of students participating in the MX laboratory class (clockwise from upper left): setting up crystallization plates, viewing crystals, examining electron-density maps and performing remote data collection at beamline 4.2.2 (ALS). Photographs by Lesa Beamer.

**Table 2**  
Processing statistics for student-collected data sets at ALS beamline 4.2.2.

Data set	1	2	3	4	5	6	7†	8
X-ray source	ALS 4.2.2	ALS 4.2.2	ALS 4.2.2	ALS 4.2.2	ALS 4.2.2	ALS 4.2.2	ALS 4.2.2	ALS 4.2.2
Wavelength (Å)	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Total oscillation angle (°)	180	180	180	180	180	180	70	180
Resolution range (Å)	61.9–1.87 (1.91–1.87)	61.9–1.90 (1.94–1.90)	61.9–1.83 (1.86–1.83)	62.0–2.10 (2.16–2.10)	61.8–1.88 (1.92–1.88)	62.0–2.03 (2.08–2.03)	61.2–2.80 (2.95–2.80)	61.5–2.20 (2.27–2.20)
Space group	$P4_12_12$	$P4_12_12$	$P4_12_12$	$P4_12_12$	$P4_12_12$	$P4_12_12$	$P4_12_12$	$P4_12_12$
Unit-cell parameters (Å)	$a = b = 61.9,$ $c = 209.6$	$a = b = 61.9,$ $c = 209.6$	$a = b = 61.9,$ $c = 210.7$	$a = b = 62.0,$ $c = 210.1$	$a = b = 61.8,$ $c = 209.3$	$a = b = 62.0,$ $c = 210.7$	$a = b = 61.2,$ $c = 205.0$	$a = b = 61.5,$ $c = 207.6$
No. of observations	469452	453827	491485	258343	457744	342662	52010	283606
Unique reflections	34857	33295	37680	24995	34159	27699	10091	21261
Multiplicity	13.5	13.6	13	10.3	13.4	12.4	5.2	13.3
$I/\sigma(I)$	26 (2.0)	19 (1.9)	25 (1.9)	14 (1.8)	21 (2.0)	20 (2.2)	10.3 (1.9)	17 (2.0)
$R_{\text{merge}}$	0.076 (1.254)	0.112 (1.449)	0.079 (0.972)	0.103 (0.677)	0.085 (1.195)	0.100 (0.823)	0.156 (0.838)	0.109 (1.215)
$R_{\text{meas}}$	0.079 (1.314)	0.117 (1.505)	0.082 (1.032)	0.108 (0.739)	0.089 (1.249)	0.105 (0.890)	0.174 (0.930)	0.113 (1.266)
$R_{\text{pim}}$	0.021 (0.391)	0.031 (0.406)	0.023 (0.335)	0.031 (0.287)	0.024 (0.361)	0.029 (0.330)	0.076 (0.397)	0.031 (0.352)
CC1/2	1.0 (0.78)	1.0 (0.70)	1.0 (0.70)	1.0 (0.71)	1.0 (0.80)	1.0 (0.76)	0.99 (0.70)	1.0 (0.77)
Completeness (%)	100 (100)	100 (100)	100 (100)	99.9 (99.3)	100 (100)	100 (99.8)	98.7 (99.5)	100 (100)

† Only 70° of data could be processed owing to an apparent pathology, most likely a split crystal.

crystals for data collection was done after two calendar weeks. Nearly all student teams successfully produced crystals, in both hanging- and sitting-drop plates. In a few cases, the lack of crystals was traced back to errors in following the protocol, such as not including protein in the crystallization drops.

Students were given advice on how to select optimal crystals for data collection (large size, good morphology *etc.*) and provided with pins/loops for harvesting. Cryoprotection was done through successive additions of the known cryoprotection buffer (Singh *et al.*, 2009) to the crystallization drops: 10% Tacsimate pH 7.0, 0.1 M HEPES-Na pH 7.0, 24% (w/v) PEG 3350, 25% (v/v) PEG 200. Two stations for flash-cooling crystals were set up, each with a microscope, a small low-form Dewar of liquid nitrogen for flash-cooling crystals, a tall Dewar for storing vitrified crystals on cryo-canes, a pair of cold-resistant gloves, pins for harvesting crystals and a magnetic wand for manipulating the pins.

Student-mounted crystals were loaded onto cryo-canes and transferred outside the laboratory class to Uni-pucks, which were shipped overnight by Fedex in a dry shipping Dewar to beamline 4.2.2 at the ALS. A shipment of crystals grown outside the laboratory class had been previously sent to the beamline, to serve as a backup in case of shipping delays or damage to the crystals during shipment.

### 3.3. Remote data collection, processing and transfer

Data collection time that coincided with a laboratory class period was pre-arranged with beamline staff. Prior to the start of data collection, Dr Nix gave the students a brief introduction to the ALS and the beamline *via* Skype. Uni-pucks containing student crystals were loaded into the robot Dewar by beamline staff. Login to the beamline was done using the remote desktop NoMachine (NX Technology) and data collected using the *Blu-Ice* interface (McPhillips *et al.*, 2002), which was displayed on a large screen in the laboratory. After an initial demonstration, each student team (two persons)

used the software interface to select their crystals for mounting by the robot and screen them for diffraction. Instructions were given to students on simple evaluations of diffraction, including resolution, looking for a single lattice and checking for ice rings. Students collected complete data sets on suitable crystals using a CMOS-based Taurus-1 detector (Thompson *et al.*, 2013) in shutterless mode (Hasegawa *et al.*, 2009). Each set consisted of 180° of fine-sliced data collected over 3 min of X-ray exposure, with frames written every 0.1 s.

Diffraction data were processed in space group  $P4_12_12$  with  $a = b = 62.0$ ,  $c = 211$  Å; the asymmetric unit contains one FtHAP molecule. Data processing was done using an automated script developed by Dr Nix that integrates and scales with *XDS* (Kabsch, 2010) and converts intensities to amplitudes with *Aimless* (Evans & Murshudov, 2013). This is the same script utilized for processing by all research users of beamline 4.2.2, and requires only a single click of the mouse, from start to finish, with approximately 10 min of waiting (processing time) per data set. (Students watched the setup for processing, although an instructor clicked the start button.) The high-resolution cutoff was selected so that  $I/\sigma(I)$  in the high-resolution bin was approximately 2.0 (Table 2). [In classroom lecture 2, the students learned that  $I/\sigma(I)$  is a measure of signal to noise, which is an intuitive metric that applies to most experiments.] Processed reflection files were transferred by secure file transfer protocol immediately after data collection to laboratory computers for refinement and model building. See Table 2 for processing statistics for the data sets obtained during the MX module, showing the good-quality diffraction data obtained from the student-grown crystals and student-collected data sets.

### 3.4. Refinement and model building

Processed data sets in MTZ format, a starting PDB (Protein Data Bank) coordinate file and an amino acid sequence file

were provided to each student team. Various three-dimensional models were prepared ahead of time, which were modifications of the known FtHAP structure (PDB entry 3it1; Singh *et al.*, 2009); three different models were assigned among the student teams. Modifications included deletions of side chains or multiple residues (*e.g.* one turn of an  $\alpha$ -helix). Students used the provided input files, ran one round of refinement in *Phenix* and then examined electron-density maps in *Coot* (Fig. 3). Students interpreted the maps for missing atoms and used *Coot* to make appropriate corrections to the model. Several rounds of iterative refinement and model building were performed and  $R/R_{\text{free}}$  monitored. Before finishing refinement, students were also asked to interpret density for an unknown ligand bound in the active site of the protein. [FtHAP crystals grown from Tacsimate (a mixture of carboxylic acids) have the competitive inhibitor L(+)-tartrate bound in the active site (Panjikar *et al.*, 2005); this compound was also utilized in kinetic inhibition studies of FtHAP.] At the end of refinement, students were asked to make figures of their structure using *PyMOL*.

### 3.5. Supplementary lectures

In addition to the four laboratory class sessions utilized, two lectures (30–50 min each) on MX were given to the students. The first consisted of an overview of MX, including its applications to the life sciences, a discussion of protein crystallization methods and theory, and an introduction to synchrotrons as an X-ray source. The second lecture provided an introduction to X-ray diffraction and electron-density maps. The specific topics covered included Bragg's law, the definition of crystallographic resolution, how resolution impacts the appearance of electron-density maps, what to look for in data processing statistics tables, the phase problem, the mathematical relationships that relate electron density and structure factors, how to interpret  $2F_o - F_c$  and  $F_o - F_c$  electron-density maps, and the *R* factor as a measure of the agreement between experimental data and the atomic model.

### 3.6. Efforts to foolproof the module

Owing to the novelty of the MX module, significant efforts were made to avoid potential problems and pitfalls that would lead to failure of the experiments and a poor learning experience for the students. These included preparation of extra protein stocks by the teaching assistant, as well as sending a backup shipment of crystals to the beamline prior to the scheduled data collection time. Each of the different crystallization protocols was pre-tested by undergraduate researchers outside the laboratory class, to ensure that at least some of the conditions would produce crystals.

Other potential problems, including the possibility of a poor remote connection between the teaching laboratory and the beamline, were also considered. An alternative data collection time was pre-arranged in case of problems with the Dewar shipment, beamline instrumentation or network connection. Although network problems are unpredictable and difficult to

prevent entirely, the remote connection was pre-tested and worked successfully during the scheduled time.

## 4. Considerations for future implementations

### 4.1. Student feedback

After finishing the MX module, students were given a simple survey to complete, which addressed their overall learning experience and asked for opinions on different aspects of the module. For convenience, the MX module was divided into four sub-modules: growing crystals, harvesting crystals, collecting X-ray data, and electron-density map interpretation and structure analysis. (Each sub-module corresponded to approximately one laboratory session.) Students were asked which of these four were (*a*) most fun/interesting and (*b*) most helpful for understanding biochemistry. While each received a range of scores (both high and low), overall the students ranked harvesting crystals as the most fun/interesting and analyzing the structure as the most helpful with understanding biochemistry.

When asked to list something new learned from the unit, the top three responses were (i) learned about X-ray crystallography for the first time; (ii) learned about the difficulty of crystallizing proteins; and (iii) learned about protein structural analysis. When asked about difficulties or confusing parts of the MX module, the top responses were (i) analyzing the structure and/or using software; (ii) understanding the basis of diffraction; and (iii) integrating the various components of the module into a 'whole'. All of the respondents (15/16 total students) recommended keeping the MX module in future years of the class. Anecdotal student remarks confirmed very high enthusiasm for the module.

### 4.2. Improving laboratory-time utilization

Based on student evaluations and post-class discussions of the authors, the MX module is planned to recur on an annual basis in the same laboratory class at our institution. As the time constraints of this class are considerable, a major consideration for our future implementations is improving the utilization of laboratory time. The crystallization and electron-density map interpretation modules fit particularly well into the undergraduate laboratory format, because all the teams can work simultaneously. On the other hand, crystal harvesting was limited by the number of microscope stations (two in our case). The synchrotron data collection was even more limited since only one team can collect data at a time.

It is therefore important to plan activities for the students to perform during the harvesting and collection days to minimize idle time. Such efforts were moderately successful in this first implementation. For example, during the crystal harvesting day, the idle students visited the proteomics center in preparation for an upcoming laboratory class or completed a written assignment from a previous laboratory class. On the data collection day, the students started a new activity related to an upcoming laboratory class. In the next implementation of the MX module, we plan to better utilize existing downtime

during the laboratory sections. In particular, as the final part of the module (model building/refinement) seemed to be both quite challenging and also relatively more time consuming, we will introduce a model building tutorial with *Coot* prior to its use in the MX module. Various model building options will be demonstrated on a shared computer that is displayed on a large screen in the laboratory. This could be done, for example, at the start of the crystal harvesting day, and the students could then perform an exercise on model building while they are waiting for the harvesting station to become available.

We also plan to formalize a set of learning objectives for the MX module. Many of these were addressed informally with the student during the interactive laboratory sessions, but as this was not done in a systematic fashion, it is likely that some students missed certain points. In the future, we will give the students a list of questions for each sub-module that specifies key issues for the students to be aware of during their experiments. Samples of potential learning objectives are provided in the supporting information.

#### 4.3. Simplifications to reduce required class time, need for local expertise

Many variations of the MX module can be envisioned, which could help reduce its overall complexity and/or the required classroom time. Our MX module was designed to complement the curriculum of the existing senior biochemistry laboratory component, which emphasizes the use of student-prepared samples. In cases where this is not a priority, an obvious simplification would be to use commercially available proteins for the crystallization experiments. Several of these are available (*e.g.* Hampton Research) and are relatively inexpensive. In the case of hen egg-white lysozyme, a 10 min crystallization reagent is available (Hampton Research), which would allow crystal growth to be observed in real time by the students. Another potential time-saving option, where available, would be use of a crystallization robot for setting up the plates. This would allow the use of 96-well plates, which could cover a wider range of conditions and facilitate the use of commercial crystallization screens, rather than pre-prepared stock solutions. Such modifications would reduce the 'hands-on' effort by students, which could be advantageous in some settings.

Other adaptations can be made to reduce data collection time or the need for local crystallographic expertise. For example, collection of a single good X-ray diffraction data set, rather than one per team or student, might suffice. If so, time for data processing and transfer would be correspondingly reduced. This might allow data processing to be done by beamline staff, rather than relying on local crystallographers.

#### 4.4. Modifications to increase challenges and learning

Many modifications can be imagined to increase the challenges of the MX module, if desired. The scope of our MX module was highly constrained by pre-existing parameters of the laboratory class (allotted time, knowledge base and

interest level of students, the need to minimize cost of supplies *etc.*). In the case of more advanced students or if additional time/resources are available, it will be easy to introduce modifications at different points in the module (*e.g.* prior to, during or after the MX experiments). As an example of the first, additional molecular biology or biochemical studies could be conducted on the protein that will be crystallized, such as making site-directed mutants. Students could then purify and crystallize the mutant protein(s), and interpret the effects of the mutation on structure and activity. This might also provide opportunities to assess changes in unit-cell parameters or space groups arising from the mutations. An example of additional work during the MX module would be to ask students to identify crystallization conditions from scratch, using a sparse matrix screen, and to then optimize those conditions.

If time and local expertise permit, additional instruction on data quality as well as hands-on data processing by the students could also be incorporated into the MX unit. Similarly, multi-wavelength/single-wavelength anomalous dispersion X-ray data sets could be collected, and used to solve the structure *via* various automated structure determination packages such as *Phenix* (Adams *et al.*, 2010) or *Auto-Rickshaw*, which runs on a web server easily accessible to students (Panjikar *et al.*, 2005). If additional time is available for model building, three-dimensional templates that require more extensive re-building (*e.g.* polyalanine chains or models missing one domain of a multi-domain protein) could be easily prepared. These are but a few of the many possibilities that could be implemented depending on the goals of the class and interests of the students.

As noted previously, our MX module was designed as an introductory experience for students with broad career interests in the life sciences (*i.e.* not just those interested in graduate school). At institutions where faculty and student interest are sufficient for a more extensive course on MX, other more comprehensive advanced tutorials are available (Faust *et al.*, 2008, 2010), which could be utilized as previously described, with the incorporation of real-time remote data collection as a part of the tutorial. This would be of particular benefit for institutions without local access to a beamline.

#### 4.5. Remaining hurdles to routine MX in undergraduate education

Currently, access to synchrotron beamtime is arranged through application to a general user program, or through collaborations with crystallographers whose institutions are part of a beamline consortium (as in our case). In places where such arrangements are not forthcoming, such as smaller colleges, easy access to beamtime remains a hurdle for implementation of an MX module. This barrier could be overcome, however, through development of 'teaching user' applications at beamlines, so that beamtime requests could be made directly by teaching faculty or undergraduate course directors. Support for implementing such a program might come from agencies such as the National Science Foundation

or the US Department of Energy, which have interests in science education as well as gaining public support for funding research infrastructure.

Other remaining barriers include the need for local crystallographic expertise at various stages of the process. Crystal growth, for example, is not difficult for novices given an amenable protein. However, mounting and flash-cooling crystals require more finesse, and probably at least initial oversight by an expert. Similarly, use of the data collection interface and knowledge of appropriate experimental settings would also benefit from expert input, at either the institution or the beamline. Some of these barriers could be reduced, however, through the use of instructional videos or other tutorials. Indeed, several YouTube (<https://www.youtube.com>) videos from well respected crystallographic laboratories are already online that provide instructions for freezing crystals, and others could be easily developed as needed. Alternatively, in situations where these barriers cannot be overcome, the module can be modified to simplify various steps, as noted earlier.

Other remaining practical issues include the expense and time commitment for implementing the module. As detailed above, the consumables are not prohibitive in price. However, some of the non-consumable items, such as shipping Dewars, are expensive (>\$1000) and would have to be purchased if not locally available. This particular cost could be avoided if Dewars were available to borrow from beamlines for those who are shipping crystals. Other costs, however, such as specialized tools for freezing crystals would still remain if they could not be borrowed. With regard to the labor and time involved in the MX module, this was quite significant for its initial implementation, particularly for the teaching assistant. This time commitment should be reduced in future years now that protocols are available.

Other constraints inherent to teaching MX in the undergraduate curriculum may be more difficult to overcome, particularly since students at this level are typically not interested in a full semester course on the topic. Thus, one significant challenge includes finding time within the curriculum of existing courses (e.g. the basis of our abbreviated 2 week module), many of which are already packed with required topics. In addition, tailoring the presented material on MX at a level that matches the background knowledge of potential students but also piques their interest is also critical, and will need to be done on a case-by-case basis.

### 5. Conclusions

Familiarity with modern MX is a missing link in the preparation of undergraduates who wish to pursue careers related to the life sciences. Among other areas, that of three-dimensional structural data is playing a growing role in medicine, owing to the era of 'personalized biophysics' (Kroncke *et al.*, 2015), where three-dimensional structural information is sought to help understand the basis of disease in a single individual. Here we show that many of the hurdles previously associated with introducing MX into the under-

graduate curriculum can be overcome, or at least reduced, by utilizing the widely available option of remote X-ray data collection at synchrotrons. Use of the synchrotron demonstrates to students a major trend in cutting-edge scientific research whereby experiments are performed remotely at large shared experimental stations. Other examples include the Large Hadron Collider and the Very Large Telescope in Chile. Currently, local crystallographic expertise (within the institution or through close collaboration) is still important, but efforts can be made to reduce this requirement, either through utilizing expertise of the beamline staff or through the development of instructional media. The use of our federally funded research infrastructure in undergraduate education may help increase public support for the maintenance and development of these facilities.

The expression construct of FtHAP and detailed protocols for protein expression, purification and crystallization are available upon request from the authors.

### Acknowledgements

We thank the University of Missouri, including the Department of Biochemistry and MU Structural Biology Core, for ongoing support of institutional membership in the Molecular Biology Consortium of beamline 4.2.2 at the ALS. We also thank Abigail Graham for assistance with pre-testing crystallization conditions for the MX module. KMS was supported by NIH training grant T32 GM008396-26 from NIGMS. LJB is supported by grant No. MCB-0918389 and JJT is supported by grant CHE 1506206 from the National Science Foundation. Part of this work was performed at the Advanced Light Source. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the US Department of Energy under contract DE-AC02-05CH11231.

### References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
- Aldeborgh, H., George, K., Howe, M., Lowman, H., Moustakas, H., Strunsky, N. & Tanski, J. M. (2014). *J. Chem. Crystallogr.* **44**, 70–81.
- Berman, H. M., Coimbatore Narayanan, B., Costanzo, L. D., Dutta, S., Ghosh, S., Hudson, B. P., Lawson, C. L., Peisach, E., Prlić, A., Rose, P. W., Shao, C., Yang, H., Young, J. & Zardecki, C. (2013). *FEBS Lett.* **587**, 1036–1045.
- Beteva, A. *et al.* (2006). *Acta Cryst.* **D62**, 1162–1169.
- Campbell, M. G., Powers, T. M. & Zheng, S.-L. (2015). *J. Chem. Educ.* **93**, 270–274.
- DeLano, W. L. (2002). *The pyMOL Molecular Graphics System*. Version 1. Schrödinger LLC, <http://www.pymol.org/>.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst.* **D69**, 1204–1214.
- Faust, A., Panjikar, S., Mueller, U., Parthasarathy, V., Schmidt, A., Lamzin, V. S. & Weiss, M. S. (2008). *J. Appl. Cryst.* **41**, 1161–1172.
- Faust, A., Puehringer, S., Darowski, N., Panjikar, S., Diederichs, K., Mueller, U. & Weiss, M. S. (2010). *J. Appl. Cryst.* **43**, 1230–1237.
- Felts, R. L., Reilly, T. J., Calcutt, M. J. & Tanner, J. J. (2006). *Acta Cryst.* **F62**, 32–35.
- Grazulis, S. *et al.* (2015). *J. Appl. Cryst.* **48**, 1964–1975.



- Grochulski, P., Fodje, M., Labiuk, S., Gorin, J., Janzen, K. & Berg, R. (2012). *J. Struct. Funct. Genomics*, **13**, 49–55.
- Hasegawa, K., Hirata, K., Shimizu, T., Shimizu, N., Hikima, T., Baba, S., Kumasaka, T. & Yamamoto, M. (2009). *J. Appl. Cryst.* **42**, 1165–1175.
- Helliwell, J. R. & Mitchell, E. P. (2015). *IUCrJ*, **2**, 283–291.
- Jaskólski, M. (2001). *J. Appl. Cryst.* **34**, 371–374.
- Kabsch, W. (2010). *Acta Cryst. D* **66**, 125–132.
- Kroncke, B. M., Vanoye, C. G., Meiler, J., George, A. L. & Sanders, C. R. (2015). *Biochemistry*, **54**, 2551–2559.
- McPhillips, T. M., McPhillips, S. E., Chiu, H.-J., Cohen, A. E., Deacon, A. M., Ellis, P. J., Garman, E., Gonzalez, A., Sauter, N. K., Phizackerley, R. P., Soltis, S. M. & Kuhn, P. (2002). *J. Synchrotron Rad.* **9**, 401–406.
- Panjikar, S., Parthasarathy, V., Lamzin, V. S., Weiss, M. S. & Tucker, P. A. (2005). *Acta Cryst. D* **61**, 449–457.
- Pett, V. B. (2010). *J. Appl. Cryst.* **43**, 1139–1143.
- Reilly, T. J., Felts, R. L., Henzl, M. T., Calcutt, M. J. & Tanner, J. J. (2006). *Protein Expr. Purif.* **45**, 132–141.
- Singh, H., Felts, R. L., Schuermann, J. P., Reilly, T. J. & Tanner, J. J. (2009). *J. Mol. Biol.* **394**, 893–904.
- 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.
- Thompson, A., Nix, J., Achterkirchen, T. & Westbrook, E. M. (2013). *J. Phys. Conf. Ser.* **425**, 012018.
- Warren, J. E., Diakun, G., Bushnell-Wye, G., Fisher, S., Thalal, A., Helliwell, M. & Helliwell, J. R. (2008). *J. Synchrotron Rad.* **15**, 191–194.