Using Crystallographic Data to Facilitate Students' Discovery of How Protein Models Are Produced- an activity illustrating the effect of resolution on model quality.

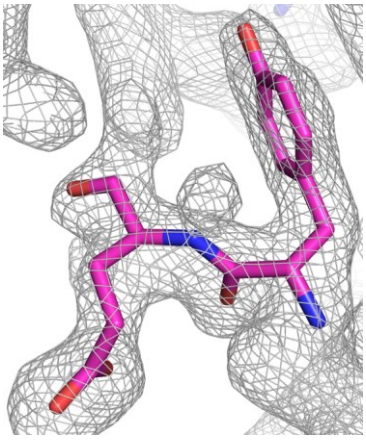
Hazel. R. Corradi*

Department of Biology and Biochemistry, University of Bath, Bath, BA2 7AY, United Kingdom

ABSTRACT

X-ray crystallography is a core technique underpinning many important results in the field of biochemistry. Although most biochemists will not become experts in this technique, many will use the structural models deposited in the protein data bank for designing and interpreting other experiments. As there are a number of limitations to these models, it is important that undergraduate biochemists, as potential end-users of this technique, have some understanding of how these models are produced. A computer activity is described in which the program *WinCoot* is used to build part of a protein model to interpret an electron density map. This activity allows students to experience the nature of crystallographic data first-hand and discover how protein structural models are produced.

TABLE OF CONTENTS/GRAPHIC ABSTRACT



KEYWORDS

Second-Year Undergraduate, Biochemistry, Computer-Based Learning, Hands-On Learning/Manipulatives, Biophysical Chemistry, X-ray Crystallography

INTRODUCTION

Protein x-ray crystallography is an imaging technique that visualizes diffraction from electrons within crystals, after processing, as an electron density map. Viewed through graphical software, electron density maps indicate the positions of the atoms

within the protein, enabling a chemical model to be built. These models are a key resource to many biochemical researchers who use them to plan and interpret their own experiments. The level of detail observed in electron density maps varies with the resolution of the diffraction data, and this is, in turn, limited by the nature of protein crystals, which have high solvent content and often diffract poorly.[1] For high resolution data, atoms can be placed accurately into the electron density map. For low resolution data useful protein models can still be built, using prior knowledge of the amino acid sequence and expected bond lengths and angles for peptides, but contain less detail.[2] This resulting difference in the quality of the models that are deposited in the Protein Data Bank, generates the need to educate the end-users in their limitations. It is, therefore, vital that graduate biochemists have some understanding of how models of protein structure are produced.[3]

The illustration of techniques through authentic activities is important in deepening student learning and can give students a stronger sense of fulfillment and satisfaction.[4] Other authentic activities that illustrate protein crystallography include crystallisation[5-6] and collecting diffraction data.[7] The activity described here complements these, giving students an opportunity to work with real experimental data to explore how molecular models are built. The activity has been completed by three cohorts of 65-80 biochemistry students at a UK university. It is one of several activities that support a second-year compulsory module in protein structure. After a basic introduction to amino acid chemistry and levels of protein structure in the first year, this unit includes key ideas about protein folds, accompanied with an introduction to both crystallography and nuclear magnetic resonance spectrosocopy, two key techniques in determining protein folds.

SOFTWARE REQUIRED

The freely available graphics program *WinCoot*[8] is a popular molecular modeling tool through which electron density maps may be displayed as a three dimensional mesh, allowing a ball and stick chemical model to be built. *WinCoot* is controlled intuitively through pull down menus and GUI elements, of a type familiar to the students, and the program contains many features that make it easier for the inexperienced to interpret the maps. For example, standard conformations for amino acids are offered, and atoms can be moved into the electron density map by clicking and dragging with the program maintaining standard bond angles and lengths.

DESCRIPTION OF THE ACTIVITY

In this activity, students build part of a three dimensional protein model into an electron density map. Students are provided with two sets of maps and models for the same protein structure, one at 2 Å resolution and one at 3 Å. The resolution indicates the distance between objects that can be clearly resolved in the map, so for a 2 Å map it is often possible to see individual bonds such as those to the carbonyl oxygen atoms that protrude from the peptide backbone. However at 3 Å, the peptide backbone is more featureless and the side chains appear as ill-defined blobs of different sizes. The maps and models used for this activity were published data, but the models were edited to remove six residues for rebuilding. The protein chosen was the N domain of Angiotensin Converting Enzyme (ACE) for which there are two models available at different resolutions (PDB codes are 2C6F,[9] 3NXQ[10]). ACE is the target of many drugs for hypertension,[11] providing an obvious interest and application for the technique. An activity schedule guides the students through the activity.

The model based on the 2 Å map is built first, as this is the easiest to complete using the standard tools. The students insert the residues one by one, ensuring each is located optimally in the map, and adjusting the bond angles to ideal ones using local refinement tools. Once all six residues have been built for the 2 Å map, this model is saved. The procedure is then repeated using the 3 Å map and partial model, although here it is substantially harder to arrange the new amino acids correctly, due to the poorer quality of the map. Both models are then superposed with the 2 Å map, allowing the student to see the degree to which the two models differ. Figure 1 shows a student's attempt for both models, highlighting the difference in the features between the 2 Å and 3 Å maps.

LEARNING ASPECTS

This activity can be summarized as three basic learning outcomes, namely to use *WinCoot* to build six residues of known sequence into a 2 Å map, to repeat this task for the same six residues but using a 3 Å map, and to compare the the two models by producing an overlay. As this is not the principal coursework of the Protein Structure unit, the outcomes are simply assessed on a pass/fail basis by requiring the students to submit a screenshot of their overlay (see Supporting Information for examples). In addition, to encourage students to think about the concepts behind the activity whilst they are doing it, the schedule contains a number of comprehension questions that students are required to submit. The responses to these questions show that completing the activity helps students to engage with the concepts more fully. Firstly, it helps them destinguish between the collected data (i.e., the map) and the chemical model that is built to interpret it. Secondly, it helps them remember that the 2 Å map is the highest resolution of the two,

as the map contains more features to place the atoms accurately.[13] Finally, some students are also able to develop a more nuanced understanding of the nature of the data in relation to the crystal. For example, that only one model is built to interpret the repeating unit of the crystal (the asymmetric unit), and that this is an average of the many copies of this unit within the crystal.

STUDENT ENGAGEMENT

This activity has taken place in three hour sessions with a cohorts between 65 and 80 students. All students were able to achieve the learning outcomes by submitting an overlay within three hours, although the quality of the 3 Å models was variable. Many students seemed to enjoy the activity, and the ability to 'drag the model into the electron density' using the mouse was often greeted with expressions of wonder. End of module feedback indicated that the students think the quality of the instructions and supervision provided were very good. Additional comments suggest that they found schedule clear and helpful. (see Supporting Information). Although the schedule contains full instructions, the novelty of the activity means that it is useful to have a ratio of one teacher/demonstrator to 10-15 students.

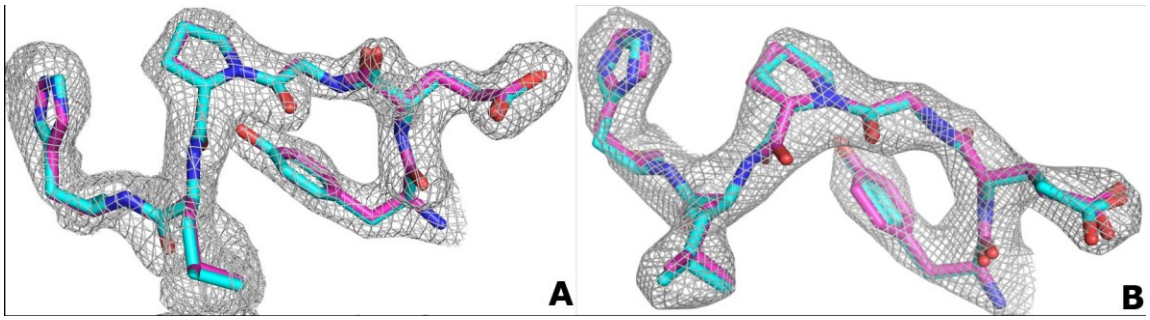


Figure 1. Overlay of the published structure (cyan) and a model from a student (pink) for both the 2 Å map (A) and the 3 Å map (B) (the rest of the structure is not shown for clarity). Bonds are shown as sticks with carbon represented by the native color (cyan/pink), nitrogen in navy blue and oxygen in red. The 2Fo-Fc map used to build the models is shown contoured at 1 sigma as a grey mesh. The picture is generated in PyMol[12] using files used and created during the practical.

CONCLUSION

An activity has been designed to enable undergraduate biochemists to experience the building of protein models using authentic crystallographic data. The activity allows students to engage with the research technique and reinforces learning about the concepts surrounding crystallographic model building and resolution.

ASSOCIATED CONTENT

Supporting Information

The activity schedule as well as the files for the activity, the notes for instructors, a summary of student feedback and examples of student work are provided. This material is available via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*bsphre@bath.ac.uk

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

The author would like to thank Susan Crennell for her feedback on the practical design and Susan Crennell, Antje Kuhrs and Tadeo Corradi for their feedback on the manuscript.

REFERENCES

(1) Blundell, T. L.; Johnson, L. N. *Protein Crystallography*; Academic Press Inc. Ltd: London, 1976.

(2) Jones, T. A.; Zou, J.-Y.; Cowan, S. W.; Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **1991,** *A47* (2), 110-119.

(3) Wlodawer, A.; Minor, W.; Dauter, Z.; Jaskolski, M. Protein crystallography for aspiring crystallographers or how to avoid pitfalls and traps in macromolecular structure determination. *FEBS J.* **2013,** *280* (22), 5705–5736. DOI: 10.1111/febs.12495

(4) Meyers, N. M.; Nulty D. D. How to use (five) curriculum design principles to align authentic learning environments, assessment, students' approaches to thinking, and learning outcomes. *Assess. Eval. Higher Educ*. **2009,** *34* (5), 565–577. DOI:10.1080/02602930802226502

(5) Peterson, M. J.; Snyder, W. K.; Westerman, S.; McFarland, B. J. Preparative Protein Production from Inclusion Bodies and Crystallization: A Seven-Week Biochemistry Sequence. *J. Chem. Educ.* **2011,** *88* (7), 986-989. DOI: 10.1021/ed100594h

(6) Garrett, E.; Wehr, A.; Hedge, R.; Roberts, D. L.; Roberts, J. R. A Novel and Innovative Biochemistry Laboratory: Crystal Growth of Hen Egg White Lysozyme. *J. Chem. Educ.* **2002,** *79* (3), 366-368. DOI: 10.1021/ed079p366

(7) Jez, J. M.; Schachtman, D. P.; Berg, R. H.; Taylor, C. G.; Chen, S.; Hicks, L. M.; Jaworski, J. G.; Smith, T. J.; Nielsen, E.; Pikaard, C. S. Developing a new interdisciplinary lab course for undergraduate and graduate students: Plant cells and proteins. *Biochem. Mol. Biol. Educ.* **2007,** *35* (6), 410–415. DOI: 10.1002/bmb.99

(8) Emsley, P.; Lohkamp, B.; Scott, W.; Cowtan, K. Features and Development of Coot. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **2010,** *D66*, 486-501. DOI:10.1107/S0907444910007493

(9) Corradi, H. R.; Schwager S. L.; Nchinda, A. T.; Sturrock, E. D.; Acharya, K. R. Crystal structure of the N domain of human somatic angiotensin I-converting enzyme

provides a structural basis for domain-specific inhibitor design. *J. Mol. Biol.* **2006,** *357* (3), 964-974. DOI:10.1016/j.jmb.2006.01.048

(10) Anthony, C. S.; Corradi, H. R.; Schwager, S. L. U.; Redelinghuys, P.; Georgiadis, D.; Dive, V.; Acharya, K. R.; Sturrock, E. D. The N domain of human angiotensin-I-converting enzyme: the role of N-glycosylation and the crystal structure in complex with an N domain-specific phosphinic inhibitor, RXP407. *J .Biol. Chem.* **2010,** *285* (46), 35685-35693. DOI: 10.1074/jbc.M110.167866.

(11) Acharya, K. R.; Sturrock, E. D.; Riordan, J. F.; Ehlers, M. R. W. ACE revisited: A new target for structure-based drug design. *Nat. Rev. Drug Disc*. **2003,** *2* (11), 891-902. DOI:10.1038/nrd1227

(12) The PyMOL Molecular Graphics System, Version 1.3 Schrödinger, LLC. https://www.pymol.org/ (accessed Aug 2015)

(13) Wlodawer, A.; Minor, W.; Dauter, Z.; Jaskolski, M. Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures. *FEBS J.* **2008,** *275* (1), 1-21. DOI: 10.1111/j.1742-4658.2007.06178.x