Journal of
Applied
Crystallography
ISSN 0021-8898

Received 19 March 2010
Accepted 26 August 2010

# RapiData: a practical course in macromolecular X-ray diffraction data measurement and structure solving at the NSLS

**R. M. Sweet* and A. Soares**

Biology Department, Brookhaven National Laboratory, Upton, NY 11786, USA. Correspondence e-mail: sweet@bnl.gov

RapiData provides two days of high-level lectures, then two more of experimental work on several beamlines of the National Synchrotron Light Source, for about 50 students. Students are invited to bring their own research projects for measurement, and about half of them do. The students frequently solve half a dozen structures during the course. Tutorials by the lecturers run throughout the data-collection period. The crystal-preparation laboratory is popular for tutorials and practice, and often there is a beamline available for practice. This article provides details about the organization of the course and tells some of the reasons for its success.



## 1. Objectives of the course and some history

Synchrotron radiation is an indispensable tool for many macromolecular crystallography groups. Some synchrotron-specific skills are difficult to learn without intensive hands-on training, but not all research groups can provide comprehensive training for students and postdoctoral staff. The RapiData course is designed to provide this training, thereby increasing the capabilities of research groups who visit the X-ray beamlines at the National Synchrotron Light Source to perform diffraction measurements. Supported largely by a grant from the National Institutes of Health (NIH) National Center for Research Resources, the course is organized by the PXRR – Macromolecular Crystallography Resource at the NSLS, funded also by the Department of Energy (DOE), Office of Biological and Environmental Research. In this course we develop the ideas behind conventional methods for diffraction-data collection, we describe several different structure-solving methods, and we give students practical experience working on their own specimens.

RapiData was intended to complement the long-running annual Cold Spring Harbor Laboratory (CSHL) course in macromolecular crystallography (MX) created by Jim Pflugrath and others, which is given each fall. The CSHL course treats all of MX, from growing crystals to preparing drawings of the refined structures. We hold RapiData in the spring, and we focus on the synchrotron-based component of modern MX: measuring the best possible diffraction data and then obtaining phase information to create an interpretable electron-density map.

We first presented the course in April 1999, and have given it annually since then. By now there are about 600 alumni. At first, our admission policies aimed to promote a high initial level of readiness, and we required that our students demonstrate a substantial background in the theory and practice of crystallography. However, this was at a time when the paradigm was shifting towards users with an increasingly modest crystallography background because many new crystallography laboratories were forming where the thrust of the research, and the background of the investigator, were very biological. They had relatively meager MX background.

To help fill the gap in background we added a day-long lecture course on the fundamentals of crystallography to sharpen the skills of the students, preparing them for a rigorous two days of lectures on data collection and structure solving, then two more days of tutorials and data collection. We also softened our admission policies for theoretical background, and we began to favor applicants from laboratories that demonstrated an emerging project-driven need for increased experience in the theory and practice of MX.

In 2002, the US National Committee for Crystallography helped us launch a Latin American initiative, complementing that of the American Crystallographic Association. They gave us seed money for a few travel scholarships to students in Latin American crystallography laboratories. By 2005, we were able to win a solid grant from the International Union of Crystallography for the same purpose, and we have been able to renew that each year since. To win the scholarship the applicants must be working in MX, must be under 40-years old, and must both demonstrate the need for the course and promise that they have sufficient background to profit from it. We have funded about half a dozen students each year; most are graduate students, some hold postdoctoral positions and a few are junior faculty.

## 2. The structure of the course

The main RapiData course begins on a Monday morning, in April, following an optional five-hour Sunday lecture course on the fundamentals of crystallography. There are lecture

courses (and fine evening meals) on Monday and Tuesday. Then starting Tuesday evening data collection begins, running until breakfast time on Friday. At the synchrotron we are on 'beamline rations': bagels and fruit for breakfast, sandwiches for lunch, pizza or heroes for supper, and lots of coffee and snacks all other times. Tutorials, run by the lecturers and often a helper, run essentially continuously (day and night) through the two full days of data collection.

The course should accommodate 48 students total. Entry to the course is by application, open from mid-December to early February. Each student submits appropriate personal data, and then a statement of what his/her preparation for the course is in terms of courses and laboratory experience, and what he/she hopes to gain from it. In addition, each student's supervisor/mentor must submit a supporting letter, giving his/her view of the student's preparation and capabilities to be gained, and how this student's attending the course would help the project leader's research program. All students are encouraged to bring their own specimens for data collection, and to bring stubborn data that can be tamed during the data-reduction and structure-solving tutorials. We divide students into 10–12 sections, depending how many beamlines are available. Every section comprises some students with crystals and some without. Students who want to bring a project must describe it, warts and all. We receive between 70 and 90 applications each year.

The course students are diverse. Typically one-third will be each of these: US citizens; those who come from other countries, essentially the rest of the world; and those of non-US origin who are living and working in the US. The largest fractions are graduate students and those in postdoctoral positions. We also have a few junior faculty who are just starting out and want to sharpen their skills before they run a group, some are senior scientists wanting to find out what it is all about, and the rare, treasured candidate is the tenured NMR investigator who is ready to turn to the dark side (really the side of en*light*enment).

We have been able each year to allocate about six NSLS dipole beamlines to the 12 sections. Each section will get a total of about 30 h on a beamline. In recent years we have made our two high-brightness undulator beamlines available for especially demanding projects. These beamlines, which run at nearly the pace of some of the hottest beamlines at third-generation synchrotrons, give students a taste of the feeling of 'drowning in data' that has become typical these days. On the other hand, every 'ordinary' dipole beamline has at least two tutors to help students puzzle out whether their crystals are good enough for data collection, how best to get the data, and then how to reduce and scale the data.

Those without crystals will have a period of several hours to practice on sample specimens they will have mounted themselves. In addition to conventional cryo-mounting, students may consider employing automounting and use of an optical spectroscopy system for colored crystals.

Anywhere between five and 12 structures are typically solved during the course, from the crystals that roughly half of the students bring. Each of these could be a publishable result, and typically one or two publications each year attribute RapiData as having been the source of their data. On Wednesday and Thursday afternoons during the data-collection portion of the course, and then on the final Friday morning, the entire course, students and instructors, gather for about a hour to discuss their results and to talk about things learned. Stimulated by H. Klei, one of our regular helpers, and of course made possible by the remarkable advances in crystallographic structure-solving and display software, the reports of results often are mini-lectures, with an introduction to the subject and pictures to show the result and the quality of the data.

## 3. The lectures

The goal for the lectures on Monday and Tuesday has been to teach a few fundamental concepts that students might not encounter in other places, then to teach the nuts and bolts of data reduction and structure solving. Here are typical topics, with the names of those who commonly present the lecture in square brackets.

(*a*) Some issues in data collection [R. Sweet]: How does the 'rotation' method work? What is collimation? How do we improve the signal-to-noise ratio?

(*b*) Modern CCD-based and pixel-array X-ray detectors (Phillips *et al.*, 2002) [C. Nielsen, W. Phillips]: What is inside the box and how does it work? What are the limitations? What can we expect for the future?

(*c*) Tactics in data collection, and X-ray damage to crystals (Dauter, 1999) [Z. Dauter, E. Garman]: How far do I have to rotate? When should I throw my crystal away?

(*d*) Special properties of synchrotron radiation [L. Berman]: Where do X-rays come from? What is an undulator? What does the future hold?

(*e*) Specimen preparation (Garman & Schneider, 1997) [E. Garman, D. Thiel, S. Parkin]: What are all the tricks in cryo-preservation?

(*f*) Data reduction with *d*TREK* (Molecular Structure Corporation, 1997) [J. Pflugrath].

(*g*) Data reduction with the *Mosflm* (Leslie, 1992) package [F. von Delft, M. Szebenyi].

(*h*) Data reduction with the *HKL* (Otwinowski & Minor, 1997) suite [W. Minor, Z. Otwinowski].

(*i*) Issues in phasing – multiple- and single-wavelength anomalous diffraction [H. Robinson]: How does one tune the data collection to get the best phases?

(*j*) Structure solving with *SHELX* (Sheldrick, 2008) [B. Schierbeek, M. Benning].

(*k*) Structure solving with *SOLVE* (Terwilliger, 2004) and *Phenix* (Adams *et al.*, 2010) [T. Terwilliger].

(*l*) Structure solving with *autoSHARP* (Vonrhein *et al.*, 2007) [C. Vonrhein, C. Flensburg].

(*m*) Structure solving with molecular replacement methods [P. Fitzgerald, P. Adams].

(*n*) Automatic control of data collection with the *EDNA* (Incardona *et al.*, 2009) system [A. Soares, G. Winter].

(*o*) Concurrent optical spectroscopy with X-ray diffraction [A. Orville].

(*p*) Operation of beamline software: *CBass, PXDB* (Skinner *et al.*, 2006) [J. Skinner].

There has been some shifting of lecture topics over the years. For many years *Shake-and-Bake* (Miller *et al.*, 1994) was taught by C. Weeks. We added *SHELX* and molecular replacement rather recently. We have emphasized crystal damage increasingly in several lectures.

## 4. The tutorials

Each of the major software packages is presented also at a tutorial. We provide over 20 computer seats around the floor of the NSLS, most in quiet but sometimes crowded rooms. All computers are networked to the central disk farm, which also carries all of the data measured during the course. For each tutorial room there is some provision for a large display that a group of up to ten might be able to follow. Most of the tutors will provide a sign-up sheet so that students can reserve a space in advance. Typically the tutorials are packed Tuesday night, then from morning to well into the night on Wednesday and Thursday.

A critical tutorial is held in our crystal-mounting room. The cryopreparation lecturer will hold court with microscopes, dewars, safety glasses, and assorted clamps, loops and other gizmos. Students may practice mounting either their own crystals or crystals that our staff have prepared. In recent years we have dedicated a nearby dipole beamline to this tutorial, so that each student can evaluate their specimen preparation and cryo-search efforts using before-and-after diffraction patterns. There is also an automounter at that station for those who want to give it a try.

There have been *ad-hoc* tutorials in the use of *Coot* (Emsley *et al.*, 2004) [A. Héroux] and writing Fortran jiffy programs [M. Rould].

**Figure 1**
Although each RapiData student attends for unique reasons, the stories frequently interweave into a rich tapestry of science, education and humor. During RapiData 2008, Yi Jun's hemoglobin work (bottom) gravitated to beamline X26-C, where Allen Orville mentored her in correlated X-ray and spectroscopy studies. The resulting science (top left) was unique and so interesting that Orville chose to make it a center piece for his RapiData 2009 lecture (center). Each year's commemorative T-shirt is based on an interesting or humorous structure solved during the previous year, and Allen cannot contain his smile as he holds up the proof of his student's hard work.

## 5. The relationship between the data collection and the tutorials

It was a surprise to us during the 1999 course that a structure was actually solved during the course itself. Although we now try not to emphasize any race to get the first structure, there is an informal competition between the several software packages each for data reduction and structure solving to have a role in each structure solved. Students, to their credit, will sometimes take their data from tutor to tutor to learn how best to use each tool to do their work.

## 6. The beamlines, their operators and other helpers

The PXRR beamlines X8-C, X12-B, X12-C, X25, X26-C and X29 have been operated by A. Soares, A. Héroux, D. Schneider, A. Saxena, H. Robinson, A. Orville, M. Becker, H. Lewis and D. Stoner-Ma. The New York Structural Biology Center beamlines at X4-A and X4-C have been operated by J. Schwanof, R. Abramowitz and C. Ogata. The Case Western Reserve beamlines at X9/X3 have been operated by B. Ramagopal, W. Shi and K. Rajashankar. One year, the NSLS beamline X6-A was operated by V. Stojanoff. The course would simply not exist without the administrators who have put each course together: S. Sclafani, D. Kranz Robertson, A. Baittinger and A. Emrick. Others who have helped in core lecturing or operations but are not named above have been H. Klei, J. Birktoft, L. Hung, R. Grosse-Kunstleve, L. Flaks, B. McGrath, K. Hölzer, K. Schroer, W. Nolan, P. Dunten, J. Berendzen, C. Lukacs, T. Radhakannan, J. Jiang, M. Allaire, R. Jayaraman, Q. Liu, Y. Chen, M. Coincon, M. Echols, R. Edayathumangalam, M. Garcia-Diaz, P. Jeffrey, E. Enemark, G. Leonard, M. McMillan, R. Page, N. Silvaggi, N. Nassar, R. Jayaraman and Q. Liu. Local workers, whose enthusiastic help with setup, operation, emergency problem solving and final clean-up have been critical, are M. Cowan, C. Dropp, R. Jackimowicz, T. Langdon, J. Lara, S. Myers, G. Shea-McCarthy, N. Whalen, R. Buono, M. Carlucci-Dayton, S. Vadai and R. McNally.

## 7. Our view of the impact of RapiData

It is always exciting to watch the students' skills develop throughout the week. Many students arrive with little or no experience in crystallography, but they leave knowing how to obtain and process real data, as well as how to locate and fix problems as they arise. That is the advantage of having such a hands-on program.

Many components contribute to the success of the course. It is crucial to have the lecture section of the course operate like clockwork: every lecture starts on time, so everyone knows what to expect and no-one is ever late for a meal. On the other hand, the 60 h of data collection and tutorials are flexible, with each student having a chance to try many things. Being able to dedicate so many beamlines to the effort gives every student a chance to try whatever comes to mind. Having so many world-class scientists as lecturers and tutors provides a critical mass of high-level scientific thinking that really defines the course, and the software designers/lecturers/tutors take RapiData as the opportunity to try out new things and to test them on new data. We do provide many teachers, tutors and helpers – there are over 35 on the roster for each course. A feature of RapiData that we adopted from the Cold Spring Harbor course, which we believe contributes a lot to the success of both courses, is that the students and teachers all have meals together. The value of this small feature, in stimulating interactions, cannot be overestimated.

We now have hundreds of alumni. We see them at meetings all over the world, and we see their names on papers and MX bulletin boards. Many RapiData alumni have now blossomed into crystallography experts, and they continue to refer their less-experienced colleagues to BNL to polish their skills. We regularly get applicants who are graduate students of former RapiData students. (We think of them as grandchildren.) Former students frequently volunteer to support the course, usually helping the beamline scientists.

Finally, one of the great sources of satisfaction for us is that at the end of the course the teachers come to ask us if they may come again next year. It is all great fun (Fig. 1).

## 8. Funding

The course is funded in part by students' fees, which until now have paid only for their room, board and supplies, and by grants from NIH's National Center for Research Resources and the DOE's Office of Biological and Environmental research. In addition to the fees and the DOE and NIH funding, a special grant was provided by the International Union of Crystallography to assist Latin American students in attending the course. Additional support was provided by Brookhaven Science Associates and of course the NSLS. We have enjoyed significant, critical financial support from Hoffmann–La Roche and Area Detector Systems Corporation, and important and enduring 'in-kind' support from Hampton Research Corporation, Rigaku Americas Corporation, Bruker AXS, Bristol-Myers Squibb, Global Phasing and MiTeGen. Finally, the employers of all the visiting teachers are donating a week of these people's time, and we warmly acknowledge that gift.

For more information about RapiData, go to http://www.bnl.gov/RapiData. For more information on the PXRR, go to http://www.px.nsls.bnl.gov/.

## References

Adams, P. D *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Dauter, Z. (1999). *Acta Cryst.* D**55**, 1703–1717.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Garman, E. F. & Schneider, T. R. (1997). *J. Appl. Cryst.* **30**, 211–237.
Incardona, M.-F., Bourenkov, G. P., Levik, K., Pieritz, R. A., Popov, A. N. & Svensson, O. (2009). *J. Synchrotron Rad.* **16**, 872–879.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF-EACBM Newsl. Protein Crystallogr.* **26**.
Miller, R., Gallo, S. M., Khalak, H. G. & Weeks, C. M. (1994). *J. Appl. Cryst.* **27**, 613–621.
Molecular Structure Corporation (1997). *d*TREK*. MSC, The Woodlands, Texas, USA.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Phillips, W. C., Stewart, A., Stanton, M., Naday, I. & Ingersoll, C. (2002). *J. Synchrotron Rad.* **9**, 36–43.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Skinner, J. M., Cowan, M., Buono, R., Nolan, W., Bosshard, H., Robinson, H. H., Héroux, A., Soares, A. S., Schneider, D. K. & Sweet, R. M. (2006). *Acta Cryst.* D**62**, 1340–1347.
Terwilliger, T. (2004). *J. Synchrotron Rad.* **11**, 49–52.
Vonrhein, C., Blanc, E., Roversi, P. & Bricogne, G. (2007). *Methods Mol. Biol.* **364**, 215–230.