Stephen R. Wasserman
Senior Research Fellow
Eli Lilly and Company
June 21, 2012

Summary

Eli Lilly and Company is a long-standing user of the Advanced Photon Source, one of four X-ray synchrotron light sources operated by the US Department of Energy. We currently operate our own X-ray beamline for protein structure at the APS, the Lilly Research Laboratories Collaborative Access Team, LRL-CAT. The partnership between our company and the APS is an important part of our effort to deliver innovative, new medicines to the patients who need them.

Lilly has more than 10 experimental compounds in Phase I and Phase II clinical trials that were developed with the aid of the Advanced Photon Source. Experiments at the APS support research on one-third of the protein targets in Lilly's early stage drug discovery portfolio. The therapeutic research areas that utilize the APS include cancer (oncology), diabetes, autoimmune, psychiatric disorders, and neurological conditions such as neurodegeneration (Alzheimer's) and pain.

National User Facilities such as the Advanced Photon Source are essential for the nation's technological development. They are, however, too large for any one organization, corporate or academic, to consider building on its own. In creating the User Facilities the government has provided a great service to the nation. Continued high-level funding to keep them operating and at the state of the art is important for the economic and technological advancement of the United States.

The relationship between the National User Facilities and their users is strong. This relationship can be enhanced by further development of the technical and organizational environment the facilities provide. Potential enhancements include modifications to the agreements between user and facility, especially for proprietary users and operators of individual beamlines from outside the DOE. The addition of automation to speed the execution of experiments and reduce future costs would maximize scientific value from the facilities. Implementation of upgrades for the core machines and, where present, ancillary experimental stations will ensure that users continue to have the opportunity to employ the unique and powerful capabilities of the User Facilities in their scientific investigations.

Stephen R. Wasserman
Eli Lilly and Company
Advanced Photon Source
Argonne National Laboratory
Argonne, Illinois 60439

Stephen R. Wasserman is a Senior Research Fellow in the Translational Science and Technologies Department of Lilly Research Laboratories, the research arm of Eli Lilly and Company. He is the Director of the Lilly Research Laboratories Collaborative Access Team at the Advanced Photon Source of Argonne National Laboratory. Wasserman received the bachelor of science degree in Chemistry from Yale University in 1979. He received the master's and Ph.D. degrees in Chemistry from Harvard University in 1981 and 1988 respectively. Prior to joining Lilly, Wasserman was a Senior Director of SGX Pharmaceuticals, Inc., a company acquired by Lilly in 2008. Earlier he was a Senior Director for deCode Genetics, Inc. and Managing Director of their Advanced X-ray Analytical Services subsidiary. Dr. Wasserman has been a staff scientist for the original Polaroid Corporation and Lord Corporation. From 1992 to 2001 he was a researcher in the Chemistry Division and at the Advanced Photon Source of Argonne National Laboratory. Wasserman is currently a member of the steering committee for the National User Facility Organization and chairs their working group on Industrial Access and Interactions.

Statement of
Dr. Stephen R. Wasserman
Director, Lilly Research Laboratories Collaborative Access Team
Senior Research Fellow, Translational Science and Technologies
Eli Lilly and Company

Before

The United States House of Representatives
Committee on Science, Space, and Technology
Subcommittee on Energy and Environment

June 21, 2012

Chairman Harris, Ranking Member Miller and Members of the Subcommittee,

It is a pleasure to be here this morning to describe Eli Lilly's work at the Advanced Photon Source (APS) of Argonne National Laboratory, one of the four X-ray synchrotron user facilities operated by the United States Department of Energy. The partnership between our company and the APS is an important part of our effort to deliver innovative, new medicines to the patients who need them.

We urge Congress to continue to support our country's National User Facilities and the National Laboratories in which many are located. We strongly agree with the sentiment recently expressed by the Director of Argonne National Laboratory, Eric Isaacs: "The work we do in the national laboratories promises to dramatically accelerate the discovery and development of new materials, technologies, and processes–and ultimately, those efforts will power the expansion of the American economy." As we will illustrate today, these new materials include pharmaceuticals.

National User Facilities such as the Advanced Photon Source are too large for any one organization, corporate or academic, to consider building on its own. The United States government had the foresight to recognize that it alone could construct this scientific infrastructure. By creating such facilities, it provides an essential service for the nation's technological development. Continued high-level funding to keep these facilities operating and at the state of the art is important for the economic and technological advancement of the United States.

Lilly and the Advanced Photon Source

Lilly has been a continual user of the Advanced Photon Source since the first days of the facility. Today, we operate our own x-ray beamline for protein crystallography, the Lilly Research Laboratories Collaborative Access Team (LRL-CAT). Each year we analyze more than 10,000 crystalline samples. Most of these crystals contain both proteins that are targets for the treatment of disease and small chemical compounds of interest in the development of potential new medicines. The

experiments at the APS permit us to examine the interaction between the protein and small molecule atom-by-atom and to develop innovative new ways to optimize that interaction. Through this detailed, microscopic view, we seek to maximize the efficacy of new pharmaceuticals and minimize side effects.

Today our company has more than 10 experimental compounds in Phase I and Phase II clinical trials that were developed with the aid of the Advanced Photon Source. In addition, experiments at the APS support research on one-third of the protein targets in Lilly's early stage drug discovery portfolio. The therapeutic research areas that utilize structure-based drug design are diverse, including cancer (oncology), diabetes, autoimmune, psychiatric disorders, and neurological conditions such as neurodegeneration (Alzheimer's) and pain.

Our work on the protein known as β-secretase, a potential target for the treatment of Alzheimer's Disease, is illustrative of the interface between experiments at the APS and Lilly's drug discovery research. The crystallographic effort that included the APS has, to date, resulted in the determination of the three-dimensional structures of more than 400 different compounds bound to the protein. But the total effort in developing a molecule that can be tested in clinical trials extends far beyond our experiments in crystallography. Considerable effort was required to design the properties of the candidate molecules, in order that the final compound could be administered orally but still enter the brain. Our biological colleagues tested the compound and its precursors for efficacy, while computational chemists developed models for the physical properties of early stage molecules and their interactions with the protein. We also tested molecules against other proteins that are fairly similar to β-secretase, in order to predict and diminish side effects. Not surprisingly, this diverse research extended over many years. It has resulted in an investigational new drug, whose phase II clinical testing will soon commence (www.clinicaltrials.gov).

The number of scientists needed to execute these experiments and analyses across the entire Lilly portfolio is large. The subset that interacts with the APS and the data from the synchrotron is more than 150. These researchers are involved directly in preparing the samples that are sent to LRL-CAT, analyzing the data that we return to them, and using the conclusions from these experiments in their pursuit of innovative pharmaceuticals.

The experimental medicines undergoing clinical trials represent only the tip of the iceberg in the use of structural biology within Lilly's drug discovery efforts. Even negative results that do not detect an interaction between compound and protein often influence future scientific directions. In other cases, the association that is found is different from the hypothesis that directed the original experiment. In a recent example, such a result led to a reassessment of the approach to be pursued with a protein target.

We are able to rapidly disseminate the results of our work at the APS throughout the company. On average, evaluated experimental results are available to our Lilly

colleagues in San Diego, Indianapolis, the United Kingdom, Spain and China within 14 minutes of completion of the analysis at the APS. During normal operations, the median time between when a sample is created and when the experiment at LRL-CAT is finished is less than 1.6 days, including the time required to ship the sample overnight to the synchrotron. This speed allows us to execute crystallographic analyses as quickly as other assays used in discovery pharmaceutical research. Virtually all of the data acquisition process is automated, permitting us to execute up to several hundred experiments each day, day in and day out. In 2011, using this system, Lilly solved more than 940 structures of proteins and protein-ligand complexes, including 29 novel discovery targets.

At the APS, we obtain data of a quality that cannot be duplicated elsewhere in the United States, including our own laboratories. We recognize the great value of this quality for the pharmaceutical discovery process. We are not alone in this recognition. Virtually every large pharmaceutical and biotechnology company operating in the United States uses the APS or one of the other DOE-funded synchrotrons. Indeed x-ray light sources are the *de facto* standard for protein crystallography. Of the approximately 8300 x-ray structures of biological macromolecules publicly disclosed worldwide in 2011, more than 85% utilized data acquired at synchrotron sources (source: http://biosync.sbkb.org). 35% of these structures came from the four DOE x-ray synchrotrons, making the United States the world-leader in this scientific area.

The power and capabilities offered by the Advanced Photon Source are even more critical for the class known as membrane proteins, which includes the G-protein coupled receptors that are the targets for a significant fraction of the pharmaceuticals available today. These proteins present significant difficulties in crystallization and the crystals obtained are extremely small. Because of their small dimensions, crystallographic analysis of these materials is only possible using high-intensity light sources such as the APS.

Lilly has committed its own resources for its research at the APS. SGX Pharmaceuticals, a company Lilly acquired in 2008, built the original beamline. We have a dedicated staff based at Argonne National Laboratory that maintain and operate LRL-CAT. In 2011, we completed an upgrade of the facility. This investment increased our sample capacity to 540 crystals at a time and doubled the speed at which we can execute X-ray measurements.

Lilly pays the DOE mandated fees for all its proprietary experiments at the Advanced Photon Source. These fees fully reimburse the Department of Energy for the cost of generating the X-rays we use. In addition, following DOE regulations, we provide, at no cost, up to 25% of the available time at the beamline to non-proprietary users from universities and other organizations. In this way, we effectively pay back the Department of Energy for its original investment in building the synchrotron.

Challenges and Opportunities

The most significant challenge we face in the use of the Advanced Photon Source is the uncertainty in federal funding for the APS. While this uncertainty is understandable given the current federal budgetary climate, the user facilities need a reliable funding stream so that they can continue to operate at the current level. If the APS and the other US synchrotron sources were not available or their operating schedules substantially reduced because of funding cuts, we would be forced to consider moving our X-ray measurements to light sources in other countries. We have performed recently, or are scheduled to perform in the near future, experiments in Canada, the United Kingdom, France and China. Reliance on facilities outside the United States, however, would slow the pace of our research and impact how soon new treatments become available to patients. It would also affect competitiveness and possibly employment here in the United States.

An opportunity for improvement can be found in the user agreements for the National User Facilities, particularly the intellectual property provisions contained therein. The DOE has recently modified these agreements. The current terms do offer some enhancement in interactions between facility staff and users. However, the new agreements are not appropriate for beamlines operated by organizations outside the DOE. We have been working with Argonne to rectify this inadvertent oversight. The provisions on intellectual property and ownership of inventions continue to have significant ambiguities for proprietary users, even though they have paid the proprietary fee. Our agreements with light sources in Canada, England, and France exhibit much greater clarity in this area: “if you pay, you own”, even when facility staff directly participate in the experiment.

Another possibility for enhancement is in the efficiency of experimental execution. In developing LRL-CAT, we have emphasized automation and efficiency of beamline operations and data collection. As discussed above, this capability permits us to rapidly return data to our scientific colleagues. We recognize that facilities such as the Advanced Photon Source require significant fiscal resources for each hour of operation. By minimizing the time for each experiment, however, we can reduce the cost for the measurement, even within a fixed hourly cost. A benefit of such an approach is that the scientists’ can focus their efforts on the most value-added activities.

Finally, the APS is currently engaged in the early stages of an upgrade to the facility. We ourselves have seen how, with time, operations can be held hostage to deprecated and aging equipment. Components purchased more than 10 years ago for LRL-CAT are no longer manufactured or are approaching their end of life. Indeed, that was one of the motivations for our recent upgrade at the beamline. The APS and other user facilities have similar issues, though on a much larger scale. Investments in upgrades, and ongoing continuous improvement afterward, will ensure the operations of the National User Facilities into the future.

Conclusion

The National User Facilities, including the Advanced Photon Source, are a scientific resource of which the nation should be justly proud. No other country has an equivalent variety of capabilities for investigation and analysis.

Science usually has long time horizons. 10 to 15 years can pass before an initial result yields a useful application. It can be difficult to discern the effectiveness of an investment made today. For this reason, scientific research undertaken now may not appear important. It is, something we at Lilly know well as we pursue new pharmaceuticals.

When we ask the government to provide capabilities that facilitate innovation, we in turn take on a responsibility to use these capabilities prudently, both scientifically and fiscally. In this way we can continue an environment of public trust that will guarantee our future technological health.

June 21, 2012

Argonne, IL

Appendix to the
Statement of
Dr. Stephen R. Wasserman
Director, Lilly Research Laboratories Collaborative Access Team
Senior Research Fellow, Translational Science and Technologies
Eli Lilly and Company

Before

The United States House of Representatives
Committee on Science, Space, and Technology
Subcommittee on Energy and Environment

June 21, 2012

"Rapid-access, high-throughput synchrotron crystallography for drug discovery", *Trends in Pharmacological Sciences*, May, 2012, Vol. 33, No. 5, pp. 261-267.

Review

Rapid-access, high-throughput synchrotron crystallography for drug discovery

Stephen R. Wasserman[1], John W. Koss[1], Sonal T. Sojitra[1], Laura L. Morisco[1] and Stephen K. Burley[2]

[1] LRL-CAT, Advanced Photon Source, Argonne National Laboratory, 9700 South Cass Avenue, Building 438, Argonne, IL 60439, USA
[2] Lilly Biotechnology Center, 10300 Campus Point Drive, Suite 200, San Diego, CA 92121, USA

Synchrotron X-ray sources provide the highest quality crystallographic data for structure-guided drug design. In general, industrial utilization of such sources has been intermittent and occasionally limited. The Lilly Research Laboratories Collaborative Access Team (LRL-CAT) beamline provides a unique alternative to traditional synchrotron use by pharmaceutical and biotechnology companies. Crystallographic experiments at LRL-CAT and the results therefrom are integrated directly into the drug discovery process, permitting structural data, including screening of fragment libraries, to be routinely and rapidly used on a daily basis as part of pharmaceutical lead discovery and optimization. Here we describe how LRL-CAT acquires and disseminates the results from protein crystallography to maximize their impact on the development of new potential medicines.

The challenge

Pharmaceutical and biotechnology companies are currently facing enormous pressure to improve research and development productivity. This pressure reflects rapidly declining revenues due to loss of patent exclusivity and other pricing constraints, and historic lows in the number of annual approvals of new chemical and biological entities [1]. Recent estimates suggest that ~30% of the attrition in drug discovery and development can be attributed to toxicity detected during preclinical animal testing or safety concerns that arise in subsequent human trials [2]. Most failures are thought to result from binding of drug candidates to one or more undesirable off-targets. A further ~30% of the attrition of new clinical candidates results from efficacy failures, when engagement of the target protein is inadequate or fails to produce the desired clinical outcome [2].

Efforts have been under way for more than a decade to make structural biology central to the drug discovery process [3–6]. The goal has been to use structures of proteins (drug targets and off-targets) and protein–ligand complexes to directly and rapidly influence the discovery and optimization of lead compounds and the selection of drug candidates. As the premier method for visualizing the interaction between compound and protein, crystallography can help to minimize off-target effects by guiding medicinal chemistry efforts towards specific and selective interaction with the target. Such an approach to innovation seeks to combine what is now technically feasible in structural biology with what we must accomplish if the industry is to continue to prosper. The challenge, however, is twofold. Traditional crystallography pipelines in pharmaceutical and large biotechnology companies rarely do justice to the speed at which structures of protein–ligand complexes can now be determined. Because of current economic realities within the industry, this situation is unlikely to change. In addition, routine daily access to synchrotron X-ray sources, the most efficient route to high quality data, is uncommon.

The infrastructure

The past decade has seen dramatic advances in the infrastructure available for structural guidance of drug discovery. Rapid crystallographic data collection from small samples (~10–100 μm for the longest dimension) is now routinely available at an ever-growing number of third-generation synchrotron sources (BioSync: A structural biologist's guide to high energy data collection facilities; http://biosync.sbkb.org/). These sources exploit insertion devices to provide very small, intense and highly directional X-ray beams [7,8]. Unlike in-house laboratory sources, which are limited to X-ray wavelengths corresponding to the K_α emission lines of various metals, synchrotron facilities offer access to a continuous range of X-ray energies. With this flexibility and the relative ease with which we can now prepare samples that substitute Se-methionine for methionine, determination of a new protein structure via measurement of X-ray phases can often be accomplished with just one crystal [9,10]. In 2011, publicly disclosed experimental structures of biological macromolecules exceeded 9200 worldwide (Protein Data Bank, http://www.pdb.org). Approximately 93% of the structures came from X-ray experiments, the overwhelming majority of which (~90%) were performed at synchrotron sources (http://biosync.sbkb.org/). Although deposition of structures to the PDB by industry represents a small fraction of all public disclosures (<10%), most industrial structures are not published. Extrapolating

Corresponding author: Wasserman, S.R. (swasserman@lilly.com).

0165-6147/$ – see front matter doi:10.1016/j.tips.2012.03.009 Trends in Pharmacological Sciences, May 2012, Vol. 33, No. 5

from internal efforts, we estimate that industry determines in excess of 10 000 macromolecular structures annually.

X-ray data collection for co-crystal structure determination of protein–ligand complexes has become incredibly efficient. For most complexes, minimal upstream sample preparation time is required to produce modest-sized crystals (~50 μm for the longest dimension). At third-generation synchrotron sources, the time for acquisition of diffraction data is typically no more than 15 min from start to finish. With state-of-the-art detectors, the process is complete within 5 min. The quality and speed advantages of synchrotron sources for this mainstay experiment have long been recognized [11].

Towards fully-integrated structure-guided drug discovery

At Lilly Research Laboratories (LRL), we are focused on using structure to improve the prospects of discovering molecules that engage the target with minimal binding to other, off-target proteins. Making this happen has entailed improving the odds of success for challenging *de novo* structure determinations and increasing the speed with which we can characterize target–ligand interactions in three dimensions. Using our proprietary LRL-Collaborative Access Team X-ray beamline (LRL-CAT), located at the Advanced Photon Source of Argonne National Laboratory, we have integrated structure determination into the Lilly lead discovery and optimization pipeline by providing co-crystal structure data on the same time scale as routine biochemical assays or biophysical measurements (such as surface plasmon resonance [12]) of compound binding. Our approach seeks to maximize the impact of structural information on the discovery of new drug candidates. Following this model, LRL determined more than 900 structures of proteins and protein–ligand complexes during 2011, including 29 novel discovery targets, two of which were integral membrane proteins.

The what

Accomplishing this end involved:

1. Minimizing upstream efforts in sample preparation by enabling data collection from the smallest possible crystals that exhibit acceptable diffracting power;
2. Providing near-immediate access to the synchrotron;
3. Sharing information regarding sample provenance between the laboratory creating the sample and the beamline;
4. Streamlining crystal handling and mounting at the beamline;
5. Minimizing the need for redundant data collection from replicate samples;
6. Maximizing the accuracy and diffraction resolution limits of data collected from a given sample; and
7. Automating data reduction and interpretation to deliver protein–ligand co-crystal structure information with minimal, if any, human intervention immediately following data collection.

The how

Traditional modes of synchrotron utilization are not compatible with a requirement that structural data be

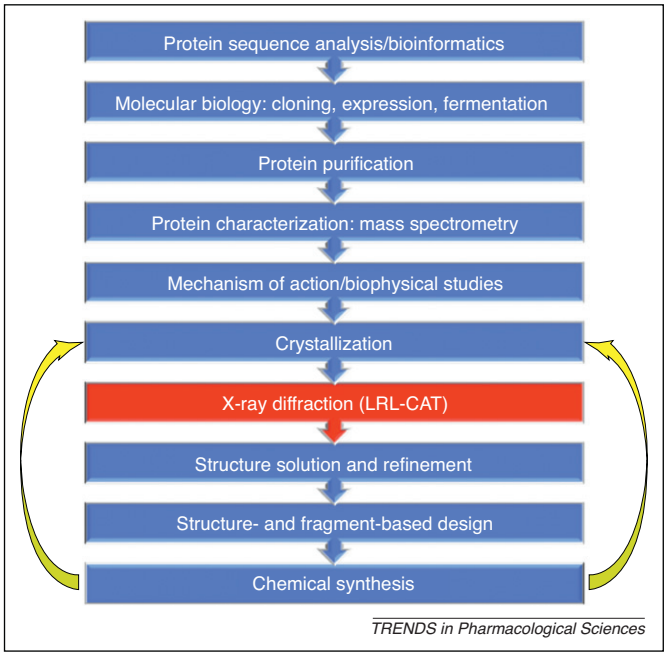

Figure 1. The structural biology process within drug discovery. The compound design cycle is applied iteratively to optimize the interaction between ligand and target.

available within days of compound synthesis or biochemical assay. Even so-called rapid access mechanisms at synchrotron sources take far too long for the lead discovery and optimization process, which ideally has a cycle time (compound design, chemical synthesis, characterization and molecular redesign) of no more than a few weeks. Figure 1 shows the structural biology process for drug discovery and the location of the compound design cycle within the overall paradigm.

Lilly has addressed this medicinal chemistry imperative by creating a just-in-time system for synchrotron protein crystallography. The LRL-CAT beamline operates without a pre-determined user schedule. Crystals are examined as they come through the door by an experienced full-time staff who operate and maintain the beamline and perform all crystallographic experiments.

In 2011, LRL-CAT evaluated 12 270 crystalline samples for diffraction quality and collected 4282 X-ray datasets. On average, a crystal completes its beamline odyssey in less than 2 days following its creation at a Lilly research site in San Diego or Indianapolis. With the aid of robust information and crystal tracking systems, LRL-CAT routinely manages several hundred samples at any given time. The Lilly Structural Biology Laboratory Information Management System (LIMS) stores all information pertaining to each protein crystal, from the original DNA construct design through protein expression, purification and crystallization to the completed structure. The LIMS system uses Oracle® for the database component, ensuring scalability to meet future needs. Each sample sent to the beamline is identified through a barcode system that includes failsafe redundancy. The barcode provides the link between the physical sample and the LIMS database information.

In addition to the LIMS data management system, the design of the hardware for the facility minimizes the need for

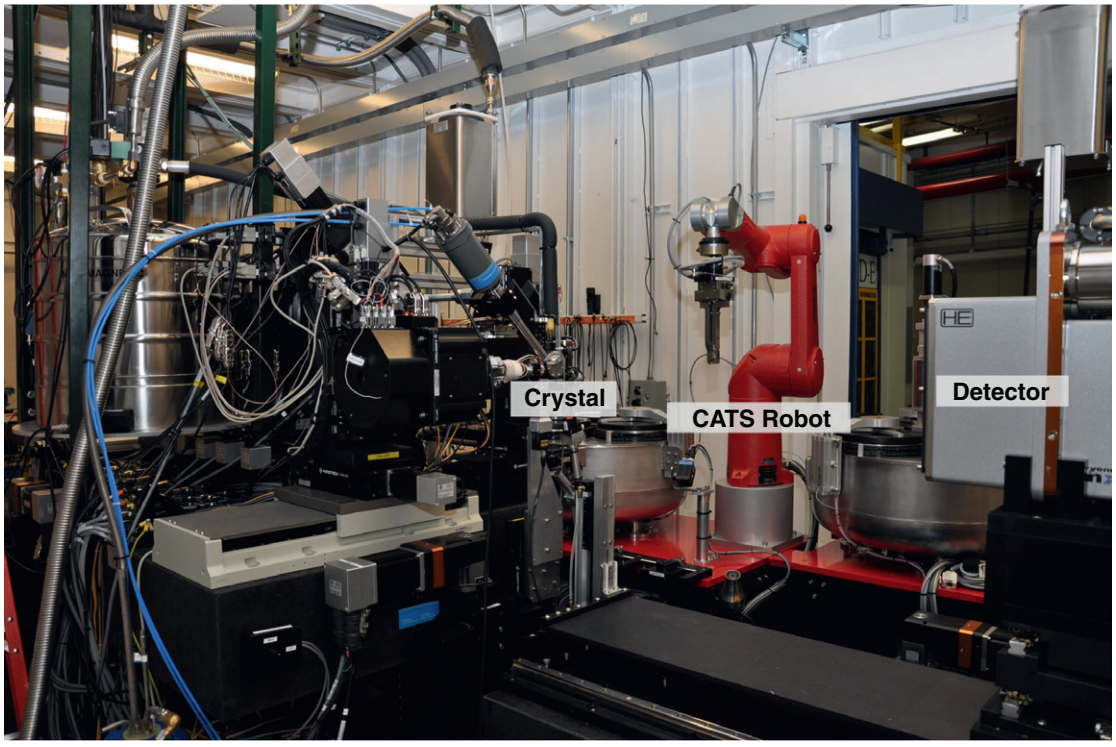

Figure 2. The crystallography end station at LRL-CAT, showing the crystal position, CATS robot for crystal mounting and CCD detector.

human intervention. LRL-CAT was constructed with commercial and custom robotic hardware and software optimized to fully automate X-ray crystallography (Figure 2). Sample queuing, mounting, centering, crystal quality evaluation, data collection, data reduction and transmission to remote Lilly research facilities are all managed by a single control system. With a high-capacity crystal-handling robot, the system can operate unattended for days at a time.

Sample queuing utilizes LIMS data to prioritize samples for evaluation of crystal quality and data collection. Lilly structural biologists assign a priority to each crystal based on the current status of the drug discovery portfolio and whether the sample is a co-crystal or an attempt at a new protein structure. This priority and the age of the sample are combined with a requirement to minimize the time expended on robotic manipulations to create the experimental queue for crystallographic analysis. The latter requirement recognizes that sequential analysis of crystals that are located near each other in the robot is more efficient. Manual overrides are available for handling special cases when necessary.

Sample mounting is performed by the Cryogenic Automated Transfer System (CATS) robot [13], a commercial system with a customized capacity of 540 crystal samples. The robot contains two storage dewars, each of which can store 27 EMBL/ESRF-type baskets [14]. Unlike the original CATS robot, which used a static configuration, the plate holding the baskets rotates into position for access by the multi-axis robot that transfers the sample onto the goniostat. The robot has a very low failure rate (<0.1%). Most failures are due not to the robot, but to defects in the materials used to mount the crystals, particularly the base on which the crystal is mounted and the plastic cryovial in which it is stored. In virtually all cases, problematic samples can be rescued through operator intervention. Operational errors with the robot are minimized by requiring the use of just one type of base and cryovial, both from a single manufacturer. The robot includes an autofill system for liquid nitrogen. Software prevents the robot from running in the event of a failure of the liquid nitrogen supply. The storage dewars maintain the samples at cryogenic temperatures for more than 12 h after loss of liquid nitrogen. Automatic text messages to the staff ensure that the cryogenics will be restored before loss of samples can occur.

Once a crystal has been placed on the sample stage, a vision recognition system identifies the center of the nylon loop containing the crystal (Figure 3) and places the center of this minute sample stage within the incident X-ray beam. In its current incarnation, the centering process requires 24 s using a single camera. Despite variations in loop size and orientation, the system is highly robust. It correctly places more than 97% of the samples in the X-ray beam without manual intervention. The success of the vision system software relies in part on a strong commitment from the upstream crystallization laboratories to use loops of a size commensurate with that of the crystal.

For each diffraction experiment, whether for crystal quality evaluation or data collection, 14 parameters are needed. These parameters include a crystal identification number, location of the sample within the CATS robot (four parameters), X-ray energy, setting of the undulator insertion device for beam attenuation, specimen-to-detector distance, initial phi angle for the crystal goniostat, number of oscillation images, oscillation range for each frame,

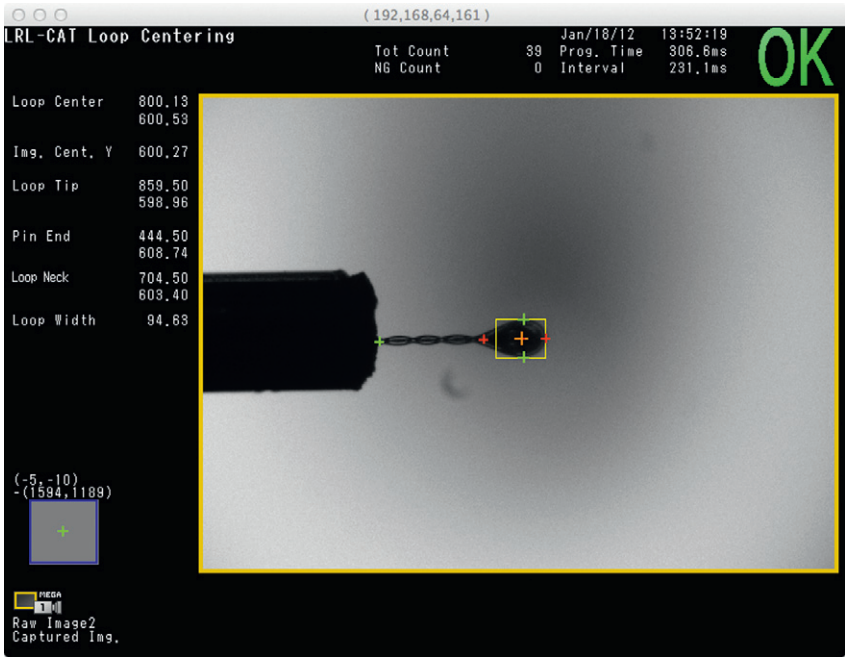

Figure 3. Automated sample alignment in the X-ray beam. The vision software system identifies several reference points for the sample mount (red and green crosses). The centroid of the sample mount (orange cross) is brought to the position of the X-ray beam.

spacing in phi for successive images (45 minus the oscillation range for screens, 0 for data sets), next frame to be collected and exposure time. The total number of parameters required to control automatic collection throughout the course of a typical day at LRL-CAT is in the thousands. Manual entry of such a volume of data by beamline operators is simply not feasible. Within the LRL-CAT paradigm, the required information is either stored in the LIMS database or can be calculated from entries resident therein. Thus, the mechanics of data collection for several hundred samples at a time can be defined in a matter of seconds by the LRL-CAT staff.

The imperative of rapid delivery of protein–ligand structures to our chemistry design teams (consisting of crystallographers, medicinal and synthetic organic chemists and computational experts) dictates that diffraction experiments focus only on samples likely to yield useful information. After an initial series of diffraction images has been acquired from a crystal, another software system, based on interpretation of output from standard software (d*trek [15] and mosflm [16]), provides a quality score and estimated diffraction resolution limit for each crystalline sample. From October 2005, when the scoring system was first deployed, to the end of 2011, more than 65 000 crystals have been evaluated at LRL-CAT. Scoring results from each of these crystals are permanently resident in the beamline database.

Within the LIMS database, replicate samples are linked as a group. Such linkage permits selection of the crystal within the group that has the highest quality score for subsequent data collection. Only crystals that meet the required minimum quality, reach the requested diffraction resolution and represent the best crystals within a group of duplicate samples, progress to data collection. During measurements, care is taken to optimize data quality through consideration of a crystal's diffraction limit when selecting the specimen-to-detector distance. The X-ray dose is matched to the diffracting power of the crystal to control the number of overloaded reflections. In addition, previous experience on susceptibility to radiation damage is used to adjust the incoming X-ray beam, particularly for anomalous experiments. Empirical data, derived from examination of thousands of crystals, have been used to create algorithms that automatically calculate the exposure time used for each image and the intensity of the X-ray beam. This approach, which contrasts with *ab initio* calculations on acceptable X-ray doses [17], has proven effective.

Once a diffraction measurement is complete, an automatic data reduction system transforms the recorded oscillation images into experimental structure factor amplitudes. Experience has demonstrated that none of the four commonly used programs (xds [18], mosflm [16], d*trek [15] and HKL2000 [19]) successfully indexes and integrates every dataset. LRL-CAT uses the first three of these programs in combination. Following integration, the data reduction pipeline sorts, scales and truncates the data. For well-defined discovery projects, known crystal symmetry and unit cell dimensions are furnished automatically from LIMS to the data reduction system. The scaling results are evaluated by internally developed quality control software, which examines R factors, data multiplicity, completeness and intensity. The resolution of the scaled data is compared to that possible given the X-ray wavelength and sample-to-detector distance to ensure that all relevant data have been collected. Overall, more than 80% of the data sets collected at LRL-CAT reach the desired resolution limit and meet other standards of quality. Samples that do not pass quality control after automatic processing are flagged for individual evaluation by LRL-CAT staff.

The final phase of the crystallographic process at the beamline involves transmission of the experimental structure factors to Lilly scientists in Indianapolis and San Diego. The median time for transmission is <14 min following completion of data collection. For most samples, further automatic processing at these remote locations is then used to convert the structure factors into an experimental electron density map, followed by a refined three-dimensional structure. In the case of protein–ligand co-crystals, the database stores the location of the appropriate model for molecular replacement to be used for each protein target during solution of the structure. For most LRL-CAT co-crystal structures, human intervention first occurs on visual inspection of how the ligand engages the target.

The Lilly system for synchrotron-based crystallography requires the ability to routinely and rapidly execute diffraction experiments, combined with robust information management. Tracking of the pipeline at LRL-CAT involves sifting through large amounts of data, including 55 individual pieces of information per sample. For a full complement of 540 samples in the CATS robot, the total number of database cells queried is ~30 000. Despite the volume of data, LRL-CAT personnel are able to determine the current status of experiments using a single web page. Furthermore, LRL-CAT is able to rapidly disseminate results to scientists at the originating laboratories. The Lilly structural biologists in San Diego and Indianapolis are able to see the images from the initial crystal evaluation within 1 min of completion of the experiment at the Advanced Photon Source.

LRL-CAT supports prodigious throughput. In comparison, Astex determined 54 structures in 80 h using in-house laboratory X-ray sources, robotic hardware and an automatic data reduction system [20].The same set of experiments can currently be done in 9 h at LRL-CAT and on a similar time scale at other synchrotron facilities. The synchrotron offers the added advantage of superior data quality, particularly for small crystals whose diffraction may not even be observed with a home source. Astex also examined approximately 160 crystals and acquired 50 datasets in 20 h at the European Synchrotron Radiation Facility [21]. The advantage at LRL-CAT is access to equivalent productivity throughout the year rather than on an occasional basis (Box 1).

On average, one-third of the crystals examined at LRL-CAT proceed to full data collection. The diffraction data for automatic quality evaluation are acquired in <2.2 min, including sample placement. Full datasets require ~10 min. For both types of experiment, data reduction is performed in the background while the next sample is being analyzed. Today, given the 3:1 ratio between total crystals and datasets collected, LRL-CAT can process (evaluate and, when appropriate, collect) more than 200 crystals in 24 h. Such bandwidth allows Lilly to use crystallography to screen small chemical fragments for binding to target proteins [22]. The core Lilly fragment library, consisting of ~2000 compounds, can be crystallographically screened against a target protein in a matter of days.

The functionality of LRL-CAT has been made available to scientists external to Lilly (http://lrlcat.lilly.com) [23–26]. Samples from academic general users of the Advanced Photon Source and industrial partners are tracked and analyzed using the same systems employed for internal crystals. Data reduction to experimental structure factors is performed after collection for both acceleration of subsequent structure determination and quality control. Direct delivery of the data, including diffraction images, to the external laboratory is accomplished through secure file transfer (sftp). The traditional requirement that samples be available for a prescheduled run at the synchrotron facility is eliminated, thereby providing our external users with data on a just-in-time basis.

Box 1. BACE

The automated sample handling, collection and data reduction process at LRL-CAT is optimized for protein–ligand co-crystals. The following example is drawn from our experience with the human β-secretase enzyme (BACE), a potential target for treatment of Alzheimer's disease [40]. In total, we have determined more than 400 co-crystal structures of Lilly compounds bound to BACE. Thus far, two drug candidates have been advanced to clinical trials (www.clinicaltrials.gov). The speed with which we can determine co-crystal structures using LRL-CAT is exemplified by the following typical timeline.

8:03 a.m.: Crystal placement on goniostat begins (LRL-CAT).
8:05 a.m.: Data set collection begins.
8:14 a.m.: Data reduction begins.
8:39 a.m.: Transmission of reduced data to Lilly San Diego.
8:47 a.m.: Molecular replacement begins (San Diego).
9:03 a.m.: Initial molecular replacement complete.
After the initial structure solution has been found, a preliminary electron density map of the active site of the protein is generated (Figure 4a).
9:50 a.m.: Automatic ligand refinement complete.

Further work may be required for structure deposition to the Protein Data Bank (Figure 4b).

What's next?

Minibeams for integral membrane proteins

Structure-based drug discovery for integral membrane proteins, including G protein-coupled receptors, is fast becoming a reality [27]. Crystals of these challenging targets tend to be quite small (<10 μm for the longest dimension), generally smaller than those produced with soluble proteins. Data quality for these systems can be improved by matching the size of the incident X-ray beam to that of the sample. Decreased beam sizes reduce the background coming from X-rays scattered by parts of the sample mount that do not contain the crystal, thereby improving the signal-to-noise ratio. Smaller beams may also reduce the effect of radiation damage on the data acquired [28]. Several X-ray beamlines have pioneered use of minibeams defined by pinhole collimators (diameter ~1–20 μm) [29,30]. The state of the art is being further refined with the advent of true microfocused beams [31]. These next-generation microbeams concentrate all the available X-rays coming from the synchrotron into a ~1-μm beam, providing advantages similar to mini-beams for even smaller crystals [32].

Integral membrane proteins are often prepared in lipidic cubic phase (LCP) [33], which is optically opaque. Because the crystals cannot be visualized directly, alignment

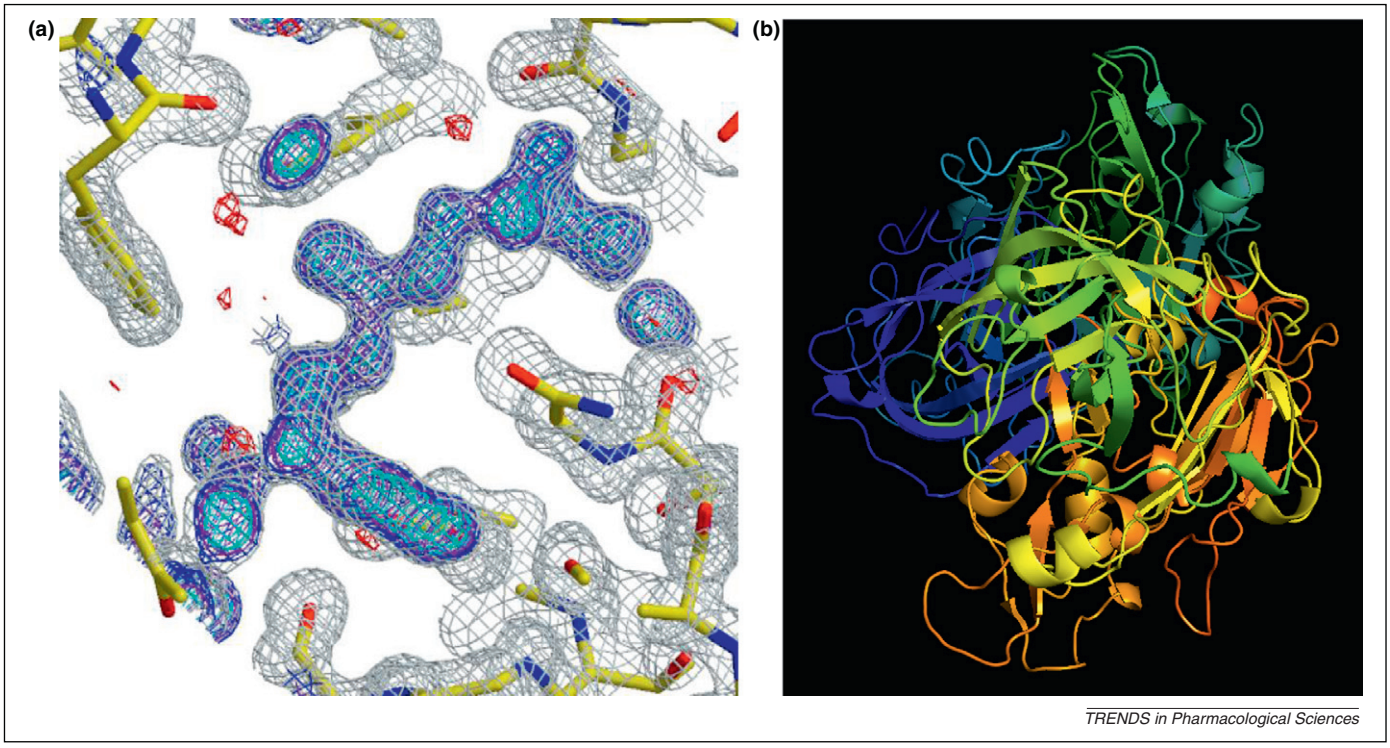

Figure 4. **(a)** Automatically generated electron density map of a ligand (blue contour) in the active site of human β-secretase. The atomic stick figure shows the enzyme structure (red, oxygen; blue, nitrogen; yellow, carbon). **(b)** Ribbon drawing of the β-secretase protein–ligand complex. The color in the protein chain follows the standard spectrum (red, N terminus; blue, C terminus).

of the crystals with small X-ray beams (≤10 μm) currently involves rastering across the sample mount to identify the precise location of the crystal [34]. Alternative detection methods such as second-harmonic generation from optically active crystals have been tested at synchrotron facilities, but are not yet routinely available [35]. LRL-CAT recently added a minibeam collimation system to its experimental arsenal. We are currently modifying the vision recognition software to provide guidance for the rastering system. The goal is fully automated positioning of crystals and mounts of all shapes and sizes.

Pixel array detectors (PADs)

PADs have recently been developed for protein crystallography [36,37]. These instruments offer two distinct advantages over the previous generation of detectors based on charge-coupled devices (CCD). First, the time required for detector readout is less than 5 msec compared with ~1 s readout and a total dead time of ~1.8 s for CCDs. Because typical X-ray exposures are of the order of 1 s, PADs can support continuous data collection without the need to open and close the X-ray beam shutter as the crystal begins and completes its rotation on the sample stage [38]. Data collection with a PAD can therefore be completed within 1–3 min instead of the 6–9 min required for CCDs. Second, PADs directly detect incident X-rays as opposed to measuring visible light generated from the X-rays by a phosphorescent film on the face of the CCD. PADs are more sensitive and when the rapid readout is used to fine-slice the measured diffraction, provide data with a superior signal-to-noise ratio. At LRL-CAT, installation of a PAD system would permit complete analysis of more than 345 crystals in 24 h.

PADs and other high-speed detectors will change how crystallography beamlines operate. Instead of initial evaluation of sample quality followed by prioritized data collection, PADs will be used to 'shoot first and ask questions later' (Oral history: Michael Rossman; http://virologyhistory.wustl.edu/rossmann.htm) [39]. Evaluation of crystal quality increasingly will be performed after data collection. The information management burden will increase commensurately, furthering reliance on sophisticated LIMS systems.

Future X-ray beamline access limitations

During the early to mid 1990s, synchrotron access for protein crystallography was the exception, not the rule as it is today. Worldwide, there are currently more than 130 synchrotron endstations for macromolecular crystallography (http://biosync.sbkb.org). These facilities offer more than sufficient capacity to meet the needs of both academia and industry. However, most such beamlines are funded and often owned and operated by governmental agencies that are now facing or soon will face significant financial pressures. Within the next 5 years, there is a very real possibility that time for macromolecular crystallography at synchrotron beamlines will again become a limited resource. How should our community respond? We can and should become better advocates for our science to governments and taxpayers, emphasizing the potential impact on human health and disease. We should also strive to improve the efficiency with which we use synchrotron X-rays. In principle, worldwide coordination among synchrotron facilities could result in on-demand access to beamlines. This goal remains elusive, even within a single country or geographical area. In the meantime, the LRL-CAT operational

model represents the most efficient way to access the advantages offered by synchrotron crystallography.

Concluding remarks

The benefits that accrue from intensive use of high-resolution structures of protein–ligand complexes in the drug discovery process are clear. At Lilly, structural biology is now used for approximately half of the discovery portfolio. We expect that the impact of synchrotron crystallography will become even more significant as discovery targets become more challenging and the innovation imperative becomes more pressing.

Acknowledgement

We thank the members of the Lilly structural biology and chemistry design teams for their constructive feedback on LRL-CAT operations. Use of the Advanced Photon Source is supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-06CH11357.

References

1 Munos, B. (2009) Lessons from 60 years of pharmaceutical innovation. *Nat. Rev. Drug Discov.* 8, 959–968
2 Kola, I. and Landis, J. (2004) Can the pharmaceutical industry reduce attrition rates? *Nat. Rev. Drug Discov.* 3, 711–715
3 Blundell, T.L. *et al.* (2002) High-throughput crystallography for lead discovery in drug design. *Nat. Rev. Drug Discov.* 1, 45–54
4 Congreve, M. *et al.* (2005) Structural biology and drug discovery. *Drug Discov. Today* 10, 895–907
5 Scapin, G. (2006) Structural biology and drug discovery. *Curr. Pharm. Des.* 12, 2087–2097
6 Wasserman, S.R. *et al.* (2007) High-throughput crystallographic data collection at synchrotrons. In *Macromolecular Crystallography: Conventional and High-Throughput Methods* (Sanderson, M.R. and Skelly, J.V., eds), pp. 173–189, Oxford University Press
7 Winick, H. (1987) Synchrotron radiation. *Sci. Am.* 257, 88–99
8 Willmott, P. (2011) *An Introduction to Synchrotron Radiation*, John Wiley & Sons
9 Joachimiak, A. (2009) High-throughput crystallography for structural genomics. *Curr. Opin. Struct. Biol.* 19, 573–584
10 Hendrickson, W.A. (1991) Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. *Science* 254, 51–58
11 Hendrickson, W.A. (2000) Synchrotron crystallography. *Trends Biochem. Sci.* 25, 637–643
12 Huber, W. and Mueller, F. (2006) Biomolecular interaction analysis in drug discovery using surface plasmon resonance technology. *Curr. Pharm. Des.* 12, 3999–4021
13 Ohana, J. *et al.* (2004) CATS: a cryogenic automated transfer system installed on the beamline FIP at ESRF. *J. Appl. Crystallogr.* 37, 72–77
14 Cipriani, F. *et al.* (2006) Automation of sample mounting for macromolecular crystallograhy. *Acta Crystallogr.* D62, 1251–1259
15 Pflugrath, J.W. (1999) The finer things in X-ray diffraction data collection. *Acta Crystallogr.* D55, 1718–1725
16 Leslie, A.G.W. and Powell, H.R. (2007) Processing diffraction data with Mosflm. In *Evolving Methods for Macromolecular Crystallography* (Read, R.J. and Sussman, J.L., eds), pp. 41–51, Springer
17 Paithankar, K.S. and Garman, E.F. (2010) Know your dose: RADDOSE. *Acta Crystallogr.* D66, 381–388
18 Kabsch, W. (2010) *XDS*. *Acta Crystallogr.* D66, 125–132
19 Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. In *Methods in Enzymology*. vol. 276. Macromolecular Crystallography, Part A (Carter, C.W., Jr and Sweet, R.M., eds), pp. 307–326, Academic Press
20 Sharff, A.J. (2003) High throughput crystallography on an in-house source, using Actor. *Rigaku J.* 19–20, 5–10
21 Blundell, T.L. and Patel, S. (2004) High-throughput X-ray crystallography for drug discovery. *Curr. Opin. Pharmacol.* 4, 490–496
22 Blaney, J. *et al.* (2006) Fragment-based lead discovery and optimization using x-ray crystallography, computational chemistry and high-throughput organic synthesis. In *Fragment-based Approaches in Drug Discovery* (Jahnke, W. and Erlanson, D.A., eds), pp. 215–248, Wiley-VCH Verlag
23 Brownell, J.E. *et al.* (2010) Substrate-assisted inhibition of ubiquitin-like protein-activating enzymes: the NEDD8 E1 inhibitor MLN4924 forms a NEDD8–AMP mimetic *in situ*. *Mol. Cell* 37, 102–111
24 Kim, M-S. *et al.* (2011) Structure of the protein core of the glypican Dally-like and localization of a region important for hedgehog signaling. *Proc. Natl. Acad. Sci. U.S.A.* 108, 13112–13117
25 Jiang, F. *et al.* (2011) Structural basis of RNA recognition and activation by innate immune receptor RIG-I. *Nature* 479, 423–427
26 Marcotte, D.J. *et al.* (2010) Structures of human Bruton's tyrosine kinase in active and inactive conformations suggest a mechanism of activation for TEC family kinases. *Protein Sci.* 19, 429–439
27 Salon, J.A. *et al.* (2011) The significance of G protein-coupled receptor crystallography for drug discovery. *Pharmacol. Rev.* 63, 901–937
28 Nave, C. and Hill, M.A. (2005) Will reduced radiation damage occur with very small crystals? *J. Synchrotron Radiat.* 12, 299–303
29 Fischetti, R.F. *et al.* (2009) Mini-beam collimator enables microcrystallography experiments on standard beamlines. *J. Synchrotron Radiat.* 16, 217–225
30 Cusack, S. *et al.* (1998) Small is beautiful: protein micro-crystallography. *Nat. Struct. Biol.* 5, 634–637
31 Moukhametzianov, R. *et al.* (2008) Protein crystallography with a micrometre-sized synchrotron radiation beam. *Acta Crystallogr.* D64, 158–166
32 Hirata, K. *et al.* (2010) New microbeam beamline at SPring-8, targeting at protein micro-crystallography. *AIP Conf. Proc.* 1234, 901–904
33 Caffrey, M. (2009) Crystallizing membrane proteins for structure determination: use of lipidic mesophases. *Annu. Rev. Biophys.* 38, 29–51
34 Cherezov, V. *et al.* (2009) Rastering strategy for screening and centring of microcrystal samples of human membrane proteins with a sub-10 μm size X-ray synchrotron beam. *J. R. Soc. Interface* 6, S587–S597
35 Kissick, D.J. *et al.* (2011) Second-order nonlinear optical imaging of chiral crystals. *Annu. Rev. Anal. Chem.* 4, 419–437
36 Broennimann, C. *et al.* (2006) The PILATUS 1 M detector. *J. Synchrotron Radiat.* 13, 120–130
37 Hulsen, G. *et al.* (2006) Protein crystallography with a novel large-area pixel detector. *J. Appl. Crystallogr.* 39, 550–557
38 Marchal, J. *et al.* (2009) Synchrotron applications of pixel and strip detectors at Diamond Light Source. *Nucl. Instrum. Methods Phys. Res. A* 604, 123–126
39 Rossman, M.G. and Erickson, J.W. (1983) Oscillation photography of radiation-sensitive crystals using a synchrotron source. *J. Appl. Crystallogr.* 16, 629–636
40 May, P.C. *et al.* (2011) Robust central reduction of amyloid-β in humans with an orally available, non-peptidic β-secretase inhibitor. *J. Neurosci.* 16, 16507–16516