Diamond Light Source

Outline Proposal for a Phase III beamline

i-plate: in-situ diffraction screening beamline

Document to be no longer than 4 pages
(not including cover page)

Times New Roman 12 pt

i-plate: in-situ diffraction screening beamline

Case prepared by: Jonathan Grimes and Frank von Delft

1. Summary

The purpose of the *i-plate* beamline is to exploit the unique properties of synchrotron radiation to perform *in situ* analysis of macromolecular crystallization experiments, and thereby deliver diagnostics on success or failure that are far more immediate and sensitive than is possible if crystals must first be harvested to assess quality, as is currently the case. The beamline would greatly expand Diamond's user base, and thus its science, to include large numbers of biologists with sufficiently pure sample but little crystallography background, to attempt crystallization successfully. We envisage three modes of operation:

- *Scanning*: Identification of crystallizability of a sample (protein/DNA/virus/etc)
- *Probing*: Identifying false positives and quantifying crystal variation
- *Collecting*: Generating complete datasets for difficult-to-harvest crystals

An intense, highly focussed X-ray beam (variable 0.5-30um) will yield sufficient diffraction from crystalline material to allow reliable characterization even in liquid. Additionally, for larger crystals refractory to harvesting (e.g. membrane proteins, viruses or large fragile complexes), the beamline will enable collecting full diffraction datasets *in situ*.

A fully unattended beamline is envisioned: crystallization plate storage- and imaging robotics, integrated in the beamline, will allow remote users to specify regions of interest (including crystals) that will be queued and sampled by X-ray at high speed. A mail-in user program will provide rapid turn-around (<2 days), allowing users to optimize their crystallization experiments very efficiently. Added value will be gained by engaging with ongoing activities at the OPPF-UK and Membrane Protein Laboratory, both of which are co-located. The beamline will be flexible toward crystallization format, in order to be compatible with continuing developments in crystallization methodology.

2. Scientific Case

Crystallography remains the analytical technique in biology that is individually most richly informative, as well as the cheapest route to structural information. Accordingly, its use and scope in biology has grown exponentially, and as it has become increasingly accessible thanks to improved methods, it is now being routinely seen as a tool, and therefore attempted by researchers with little or no experience in the technique.

Synchrotron diffraction is currently employed only at a late experimental stage of a project, namely for collecting optimal datasets when crystals could already be generated and cryogenically harvested. Indeed, beamtime usage is most efficient only if the preceding stages, crystal harvesting and testing, were already accomplished before arrival at the synchrotron, even on beamlines with robotic sample changers.

Yet it is rather earlier in the experimental sequence (Figure 1) that projects tend to fail: by far the most time and money is spent obtaining a preparation of the sample (*i.e.* macromolecule, complex or assembly) that can crystallize; and even if this is achieved, success is not guaranteed. In the worst case, a sample's ability to crystallize may remain unrecognized, since we rely on the appearance of visible crystals in crystallization trials. Crystals that do appear may be false positives (salt or detergent), or else may not diffract sufficiently well after harvesting – while it is impossible to tell whether harvesting itself may have degraded the crystals: harvesting is a laborious manual technique requiring considerable dexterity and practice, and when crystals are small (<20um) is difficult even for "experts". And if crystals simply are intrinsically variable, that too can only be assessed through yet more laborious harvesting.

These multiple potential failures are currently very difficult to diagnose, especially for less experienced experimenters, leading in turn to extensive speculative follow-up experiments that consume precious sample for poorly defined outcomes – the likely reason structures are expensive. There is thus a great need for better diagnostics surrounding the crystallization experiment.

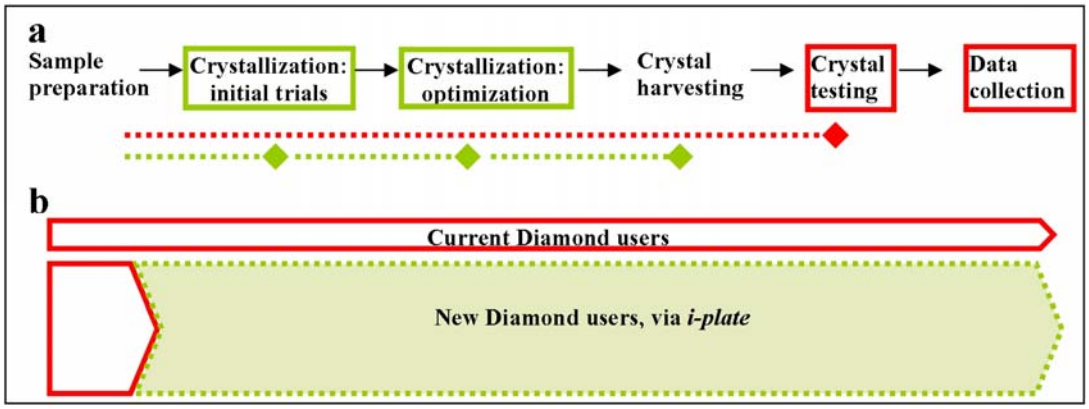

Figure 1. **a.** The experimental sequence in crystallographic structure solution. Steps already using synchrotron radiation are in red, those covered by *i-plate* in green. Dotted lines indicate unresolved uncertainty over success or failure in preceding steps, under current (red) and *i-plate* (blue) procedures. **b.** Schematic representation of the volume of potential users that would be served by *i-plate*.

The *i-plate* beamline aims to harness synchrotron radiation earlier in the experimental sequence, by implementing X-ray exposure of crystallization experiments *in situ* and at room temperature, thereby addressing the need for quick, reliable analyses prior to the harvesting step. Three types of operations are envisaged (described in detail below):

- *Scanning*: Identification of crystallizability of a sample (protein/DNA/virus/etc)
- *Probing*: Identifying false positives and quantifying crystal variation
- *Collecting*: Generating complete datasets for difficult-to-harvest crystals

These analytical tools will be set up in a remote-use, mail-in beamline so as to be easily accessible to a wide range of scientists and projects, so that they can help change fundamentally how crystallization experiments are designed and executed in the UK. The tools would enable replacing the traditional inefficient approach of broad random scans, by a rationally directed process that relies on clear analytical outcomes for the intermediate stages, and thus allow pursuing a far more effective approach to the true problem, namely identifying which variations of a sample will crystallize and yield data.

Scanning mode

Rationale: The scanning mode is aimed at assessing the sample in crystallization experiments where no obvious crystals have grown. Growth of *single large crystals* is a stochastic and statistically rare event. Thus, the common approach to crystallizing a sample (macromolecule or complex) is to set it up in hundreds or thousands of individual crystallisation trials, and observe them over time for the appearance of crystals big enough to be visible. It is still not known how efficient this approach is.

In contrast, a statistically common occurrence in these trials is *crystalline precipitation* of the sample. Knowledge of its presence is exceptionally useful: it confirms the ability of the sample to crystallize; and it indicates which chemical conditions must be optimized to grow single crystals (Figure 2a). It is however visually almost indistinguishable from non-crystalline precipitate, which is even more common. However, nanocrystals do diffract (Figure 2b), and it is this diffraction that will be probed in scanning mode.

Operation: A sub-micron beam will be scanned across areas of interest in crystallization trials, with potential diffraction captured by a fast detector at several Herz. The presence of diffraction will be unambiguously identified by comparing all images from across an area, likely requiring a non-trivial clustering algorithm. The regions to scan will initially be specified by users, but image recognition should eventually be able to select them automatically. Requested scans will be placed on a queue and executed unattended.

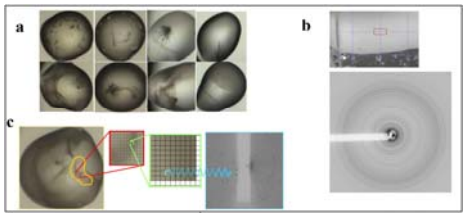

Figure 2: Features of crystalline precipitation, and the *scanning mode*. **a.** Pairs of drops (above and below) illustrate how easy it is to miss a crystal condition: drops differ only in the relative quantity of protein. **b**. Nanocrystals invisible to the eye (4-800nm lysozyme) yield measurable powder diffraction even when using a large, non-optimized beam ($60x40um^2$) (SLS beamline PXIII; R. Bingel-Erlenmeyer and V. Olieric). **c**. Schematic of the scanning process and its spatial resolution. The yellow line indicates the user-defined region-of-interest to be scanned.

Probing mode

Rationale: The probing mode is aimed at quickly assessing any crystals produced by a crystallization experiments without going to the extensive trouble of harvesting them first. This is the common usage of existing *in situ* systems, including the PX Scanner, PXIII at SLS, and ESRF. Although this mode degrades the crystal and leaves it unsuitable for subsequent harvesting or data collection, in a significant subset of time-consuming cases, the direct read-out is experimentally more valuable.

The first case is the identification of false positives, namely crystals of salt, small-molecule ligands, or detergent. The latter are very common for membrane proteins solubilised with detergent, and since crystal optimization spans many months and involves many detergents, rapid and reliable identification of the many incorrect leads is vital.

Many crystals are observed to have large variation in quality when harvested, or are very sensitive to growth conditions; this is typical for membrane proteins, large multi-protein complexes, and viruses. Finding the best diffracting crystal requires harvesting and testing large numbers of crystals; this is much simplified if the goal is known, and probing allows the best diffraction to be determined quickly by exposing *in situ* a large number of crystal still undisturbed by harvesting or else those too small to harvest.

Operation: Objects (crystals) for testing are marked by the user, and are exposed to an X-ray beam adjusted in size to match the crystal (Figure 3). Requests for probing, which can include simple rotations, will also be placed on a queue and executed unattended.

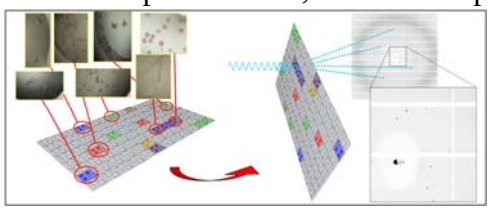

Figure 3: Schematic of the *probing mode*: users flag selected objects in the crystallization plate, which are then exposed with X-rays. It is quick to test many crystals on one plate or in one set of capillaries.

Collecting mode

Rationale: The collecting mode is aimed at collecting complete datasets from crystals that consistently lose all diffraction when conventionally harvested, or else are hazardous and may not be harvested for safety reasons. Each dataset will typically require a large number of crystals, because of the rate that crystals decay in X-rays at room temperature; on the other hand, at room temperature isomorphism between crystals is typically good.

In proof-of-principle experiments on the microfocus beamline I24 with both protein and virus crystals, we showed that complete datasets could be collected from a number of crystals at *in situ* room temperature (Figure 4). Crystals of Vaccinia virus capsid protein D13 were grown from 100+100nl protein+precipitant drops in standard SBS format Greiner plates. Over 20 crystals were examined and data were collected from multiple positions, resulting in a dataset complete to 3.2Å with good redundancy and signal-to-noise, and yielding a final model with R/Rfree of 19/25%. Likewise, 30um crystals of picornavirus, with a 400Å cell, yielded diffraction to resolutions similar to that measured for the original structure 20 years ago; crystal lifetime in this case was 1sec.

Operation: This mode is similar to the probing mode, but because data collection requires oscillation images, every crystal will need careful centring in the beam; thorough calibration of parallax will be vital. By adjusting flux, dose-rate effects at room temperature can be exploited to extend crystal lifetime. This type of experiment will be non-trivial and will need to be actively driven by the user.

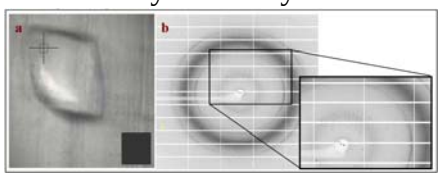

Figure 4: Proof-of-principle of *in situ* data collection on I24. **a**. Image of a D13 crystal in the I24 beam ($10x10um^2$, square upper left). **b**. A diffraction image of the dataset.

Background scatter

The major experimental challenge is background scatter, because of the large volume of non-diffracting material the beam passes through: even the slimmest crystallization formats

(*e.g.* CrystalHarp) present more than 100um of material, whereas the crystalline material is typically between 50nm-20um thick. Three counteracting strategies are envisaged in *i-plate*:

1. The beam size will be adjustable to match at all times the size of the crystal, thereby minimizing the volume of non-crystal material that scatters non-productively.
2. As large as possible a sample-to-detector distance will be used, since scatter falls off as the square of the distance from its source. As large a detector as possible will thus be necessary, the best current candidate being the Pilatus 6M.
3. Bragg peaks must remain sharp even at large sample-detector distances, to maximise signal-to-noise; the beam therefore needs as low a divergence as feasible, even if this comes at the cost of flux.

3. Outline specification

i-plate will be a fully unattended beamline: crystallization plate storage- and imaging robotics, integrated in the beamline, will allow remote users to view crystallization images and specify regions of interest (including crystals) that will be queued and sampled by X-ray at high speed. A mail-in user program will provide rapid turn-around (<2 days), to allow users to optimize their crystallization experiments very efficiently.

a. *i-plate* will be a bending magnet beamline, with a beam size adjustable between 0.5-30um and a divergence <2mrad, with capillary optics used for the smallest beams.
b. The wavelength will be fixed and energy resolution is not important, which allows considerable flexibility in the optics: this way, fluxes of between 10^{10}-10^{11} ph/s into a 2x2um beam may be achievable, as indicated by initial discussions with Lucia Alianelli from the Diamond optics group.
c. The hutch will be fully automated with no user access: plates will be loaded into two imaging vaults (4 & 20°C) that are integrated into the hutch, and a robot will transfer plates to the X-ray beam location, which should also be temperature controlled (4-20°C).
d. The beamline will flexibly support multiple existing crystallization plates and formats, and be adaptable to developments of future formats.
e. The throughput in unattended operation, with remote crystal identification and drop queuing, must be more than 150 drops per hour; but it may be much higher if a higher flux can be achieved.
f. If necessary through demand, throughput can be doubled by taking a second beam (through using a different part of the bending magnet fan, or splitting the collimated beam) to a second X-ray station in the same hutch, which then support only one operating mode. The obstacles to this approach appear to be mainly ones of engineering, rather than of physics.
g. A robust remote-access user interface will allow users to view their crystallizations and queue further experiments.

4. Community

i-plate will at a stroke vastly expand Diamond's potential user base and thereby the science it supports: along with the existing set of structural biologists who are generally experienced enough to harvest and handle crystals, any biologist with a sample pure enough to attempt crystallization will now have reason to use Diamond (Figure 1b).

At the same time, the success of *i-plate* will depend most critically on the ease with which it can be accessed. For the large section of uses new to crystallography and without crystallization infrastructure, the co-location of both OPPF and the MPL within RAL is ideal; a close cooperation with both these labs is therefore intended, and the *i-plate* user program must also be tailored accordingly.

On the other hand, existing Diamond users that are able and prefer to set up crystallizations at home will need fully reliable mechanisms to transport their plates to Diamond rapidly. For instance, national courier companies might be approached to negotiate special transport arrangements.

Ultimately, the logistical challenge is to make it easier for an experimenter to send their inconclusive experiments to *i-plate* than for them to return to the lab to prepare more sample.

Appendix: Supporting letters

Letters of support were received from the following UK researchers:

Prof. David Barford – Institute of Cancer Research
Dr Richard Bayliss – Institute of Cancer Research
Prof Vilmos Fulop – University of Warwick
Dr Robert Gilbert – University of Oxford
Dr Stephen Graham – University of Oxford
Lesley Haire – MRC National Institute for Medical Research
Dr Karl Harlos – University of Oxford
Prof Bill Hunter – University of Dundee
Dr Marko Hyvonen – University of Cambridge
Prof Yvonne Jones – University of Oxford
Dr Adrian Lapthorn – University of Glasgow
Dr David Lawson – John Innes Centre
Dr Erika Mancini – University of Oxford
Prof Neil McDonald – Birkbeck College
Prof Simon Phillips – University of Leeds
Dr Katrin Rittinger – MRC National Institute for Medical Research
Mr Pierre Rizkallah – Cardiff University
Dr Christian Siebold – University of Oxford
Dr Roberto Steiner – King's College London
Prof Brian Sutton – King's College London
Dr Ian Taylor – MRC National Institute for Medical Research
Dr Jon Wilson – Insitute of Cancer Research