Diamond-II Proposal for flagship project Beamline K04 for Ultra High Throughput XChem Fragment Screening

Science Group: **Macromolecular Crystallography**

Case prepared by: **Frank von Delft**

José Brandao-Neto

Dave Hall

Beamline K04 for Ultra High Throughput XChem Fragment Screening - Achieving Routine Preclinical Impact

1. Summary/Impact statement

I04-1 has evolved hand-in-hand with the development and growth of the XChem facility to be a world unique offering for structure-based drug design (SBDD), heavily oversubscribed by academic and industrial scientists alike. The reconfiguration of the machine for Diamond II presents the opportunity to make a step change to the impact of this facility. The resulting beamline, K04, will underpin the next revolution in rational drug discovery over the coming decade, by resolving the long-standing challenge in structure-based drug design: allowing all key 3D information to be measured up front to ensure a clinic-ready drug candidate can be designed from scratch. Crucially, Diamond II delivers the flux and beam properties required for the order-of-magnitude increase in throughput to accommodate all types of samples that are highlighted as key drug targets by modern genomics approaches. This premise now has preliminary demonstration from the recent COVID-19 XChem experiments[1] and the resulting global "Moonshot" project[2]; and supporting methodologies are in active development through various strategic collaborations (Rosalind Franklin Institute[3], EUbOpen[4], Fragalysis[5]).

2. Scientific Case

The Diamond II machine upgrade makes it unfeasible to retain beamline I04-1 in its current state, as the I04 and I04-1 undulators will no longer be able to occupy the same straight section in their current canted configuration. In addition, the layout of beamline I04-1 is such that rebuilding *in situ* at its current K03 position would be neither cost- nor time effective. To enable I04-1/XChem to meet demand right up to the dark period, and re-emerge with the required step-change in capabilities immediately after the opening of Diamond II, it is proposed to build a replacement beamline at position K04 in the run up to the closure for the machine upgrade. This minimises the downtime of a key UK infrastructure and maintains its colocation within the life science beamline cluster and the XChem laboratory space.

The driver of Diamond's 7-year investment in I04-1 and XChem is the possibility, by accelerating tool generation in chemical biology, to transform the impact of structural biology on a wide spectrum of biological questions. These range from early drug discovery in non-communicative and infectious diseases, through the understanding of chronic conditions and ageing (one of the four pillars of UK governments industrial strategy[6]), agricultural biology and indeed biological processes in general. The impacts go beyond the direct and indirect disease burden on society and encompass food security and wealth creation through increased productivity especially in the UK's biopharmaceutical and agrichemical sector.

The opportunities for both increased throughput and new experiment modalities afforded by Diamond-II are directly aligned with the requirements of a UK academic and industrial sector that has been foundational in, and remains at the forefront of, structure-based drug design. It was stressed in the 2018 user consultation

[1] https://www.diamond.ac.uk/covid-19/for-scientists/Main-protease-structure-and-XChem.html
[2] https://covid.postera.ai/covid
[3] https://www.rfi.ac.uk
[4] https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/topic-details/imi2-2019-17-02. Project launch May 12th 2020.
[5] https://fragalysis.diamond.ac.uk and https://diamondlightsource.atlassian.net/wiki/spaces/FRAG/overview
[6] https://www.gov.uk/government/publications/industrial-strategy-the-grand-challenges/industrial-strategy-the-grand-challenges

that Diamond's existing offering of high-quality, high-throughput MX analysis must if anything be expanded, and certainly not be compromised by the development of new experimental modalities; the current impact on biological discovery is enormous and should continue that way.

K04 will directly address multiple aspects of the Diamond-II vision:

- Retain a world leading research facility status
- Enable a step change in generating ideas to solve 21st century challenges
- Increase the capacity to serve the user community
- Improve the efficiency and productivity of industrial partners

Specifically, K04 will increase XChem throughput 20-100-fold, allowing screening experiments 10-30x larger than currently possible (3x5000 datasets per proposal, 3-5 per week): we now know this is what's required for an experiment to produce the quantity and quality of data that permits the direct design of molecules that can immediately be used in pre-clinical and other *in vivo* studies.

The underlying technology of the K04 and XChem facility remains Fragment-Based Lead Design (FBLD) embedded at the synchrotron(*1, 2*). This approach to drug design was already proposed in the 1980s, following the impact of molecular biology advances and the acceleration of structure determination by X-ray crystallography; FBLD has since led to over 30 clinical drug candidates with 3 FDA-approved drugs in oncology (*3*). Nevertheless, over the last half-decade, Diamond's XChem facility redefined the paradigm for FBLD, and even relying only on the significantly flux-limited beamline I04-1, has achieved an order-of-magnitude increase in the efficiency of early stage drug discovery. This has transformed the take-up of fragment screening by X-ray crystallography and has led to the democratisation of FBLD in academia and small companies: the accessibility and ease of use of XChem has enabled over 150 experiments since 2016.

Nevertheless, the full potential of FBLD, to quickly (in weeks), cheaply (<£10k) and routinely generate potent and bioactive compounds against any target of interest, remains unrealized. Such a capability would potentially unlock a flood of drug discoveries and new chemical probes (*4*), which are powerful tools for exploring fundamental biological mechanisms and validating the druggability of targets. Nevertheless, they remain rare due to the prohibitive cost of generating them. Bespoke probes, essentially on demand, would transform biological research in a similar way to gene editing. Thus, for instance, mechanisms of antimicrobial resistance (AMR) could be quickly dissected in chemical detail and then by-passed through nimble compound design, allowing rapid restocking of the antibiotic arsenal to side-step the looming AMR crisis.

The XChem developments were complemented with a number of strategic collaborations to establish the complementary downstream platforms and technologies. These include the Fragalysis Cloud[5] for rapid compound design, heavily supported through the IMI-funded ULTRA-DD project[4]; streamlined and robotic synthesis methods through the Rosalind Franklin Institute and the new IMI-funded EUbOpen project, which Diamond joined as partner organization; and most recently the "COVID Moonshot"[2], which is demonstrating how very large screening experiments can accelerate the progression of data toward clinical impact.

3. Benefit to the Diamond research community

The redevelopment of I04-1 at position K04 enables

- *Continuity of service for the "in demand" XChem facility before and after the dark period with the shortest feasible downtime*
- *Close collaboration with the end-station development of MXBridge, which will bring lessons from running the high-throughput facility at I04-1 to a next level end-station that can carry the UK community, including XChem, through the dark period and beyond*

- *Supercharging XChem by delivery of an order of magnitude more flux - this will increase the opportunity for exploring much greater druggable space and open up access to a much broader community*
- *Moving up a gear all other MX activity through the village, including addressing the vision of directly observing virus-host interactions and drug effects with vastly improved X-ray imaging methods (working alongside the rapidly improving methods of cryo electron microscopy and tomography)*

4. Outline Specification

For the beamline, the throughput and data quality can be achieved initially with refinement of existing technology, although ultimately aiming at a standard energy of ~25keV will necessitate the arrival of cost-effective high Z, large area integrating detectors.

Flux in excess of ~10^{13} ph/sec in a 1-50 microns beam with a narrow bandpass being acceptable, and a high-speed, integrating detector along with a fast, high-density sample changer will enable the aims of K04-XChem. Development of fully automated sample preparation technology, to ensure the necessary large numbers of samples can be generated for large numbers of experiments annually, should run in parallel to development of the beamline.

- Coupling a 1.5m HPMU (transferred from I04-1) in straight section K04 with a DMM/DCM and the new Diamond-II lattice will provide over a broad energy range (10 – 27 keV) at least two orders of magnitude higher flux (>10^{13} ph s^{-1}) than currently available at I04-1. This provides a flexible platform for experiment adaptability and provisions for future exploitation of high energy data collection.
- Variable focussing options (~1-50 micron) should be available to adapt to crystal properties of the target campaign.
- Diagnostic feedback systems and beam position control loops should be built in from the ground up for continuous beam delivery and alignment.
- High capacity, ultra-high-throughput end-station capable of 5,000 data collections per day (developed for MXBridge and refined for K04)
    - High speed goniometer equipped with cold nitrogen stream.
    - High speed robot exchange.
    - High capacity cryogenic sample storage system able to exchange large number of samples rapidly to minimise hutch open time.
- A silicon integrating detector will be appropriate initially but longer term a high Z sensor detector will be needed if routine use above 20 keV is shown to be effective for the smallest crystals.
- Data analysis pipelines will be in place to cope with the ultra-high data rates. Reduction of diffraction data on the fly will be essential for both x-ray centring and integration. Consideration of data model for storage will need to be assessed and implemented. Cloud based capabilities will need to be baked in at the outset for downstream XChem analysis pipelines.

5. State of the art benchmark

Since the beginning of the XChem operations in 2016, I04-1 has continually evolved to ensure fast, reliable and fully unattended data collection. The feasibility had previously been demonstrated at the Lilly-CAT[7] beamline (APS), which ran an unattended programme since the early 2000s. The fully automated beamline MASSIF-1[8] (ESRF) evolved over time to encompass broader capabilities, initially screening and collecting from

[7] http://lrlcat.lilly.com/
[8] https://www.esrf.eu/MASSIF1

user generated crystals, and now offering a crystallisation-to-structure service linked with the HTX platform[9] of EMBL-Grenoble.

I04-1 supports the regular MX user programme, alongside the high-throughput XChem user programme. The high throughput MX beamlines at Diamond will collectively offer a rapid response service, mixing fully automated collection with interactive remote experiments in response to COVID-19 restrictions. The K04-XChem facility will be positioned to take the developments driven from this to another level for FBLD.

The emergence of protein structure determination by cryoEM allows a repositioning of crystallography in the landscape of structural biology (*5, 6*). The transformation in throughput delivered by MX beamlines such as K04, linked to the unique capabilities of the XChem facility, opens up possibilities not currently available, increasing the relevance of expanded capacity and rapid response to existing and emerging health needs. Coupled even with low-flux I04-1, XChem is already capable of delivering thousands of structures per year (Table 1); an order-of-magnitude expansion of output will be truly transformative and cannot be matched by any other structure biology technique. This is an opportunity also embraced by the new FragMAX[10] facility (MAX IV); what will be unique however through Diamond-II is that capacity for other MX experiments can be sustained at full capacity, thanks to upgrades throughout the rest of the MX village.

Year	Depositions in PDB from Diamond Beamlines	Est. structures from XChem	**Est. Total output from Diamond Beamlines**
2015	1002	0	**1002**
2016	916	2500	**3416**
2017	2153	3000	**5153**
2018	1250	4000	**5250**
2019	631	5000	**5631**

Table 1. Numbers recorded in PDB for depositions imply lower output since 2017 however after depositing some XChem results as separate depositions, counting changed and underestimates output. A conservative estimate of structures from XChem is outlined above and therefore we can draw an estimate of total structures from Diamond over the preceding five years. A good proportion of activity at MX beamlines and XChem is proprietary research and these outputs are unlikely to be reflected in these numbers.

I04-1/XChem is now unique in its capabilities for joined up FBLD activities for proprietary and non-proprietary research and recognised as world leading (*6*). The delivery of beamline K04 will ensure continued pre-eminence in this field for Diamond and the UK.

6. Community engagement

The 2018 Diamond-II workshops confirmed overwhelming demand in the user community for the increased capacity for the currently available experiments that will also be supported by K04. As for the XChem experiment, growth in demand continues unabated, even for the current smaller flux-limited screens: waiting times for industrial access are almost 6 months, beyond the time horizon of typical early discovery projects. Given the clear relevance of this proposal to both academia and industry proposal champions will be sought who can strongly represent both camps.

A user working group for XChem is already in existence and will form part of the consultation process. In parallel a joint user working group for beamline development for ultra-high-throughput data collection for K04, MXBridge, I03 and I04 could comprise academic and industrial users, beamline scientists (in particular with experience of automation) and computational scientists.

[9] https://www.embl.fr/services/ht_crystallisation/
[10] https://www.maxiv.lu.se/fragmax/

7. References

1. S. E. Thomas *et al.*, Structure-guided fragment-based drug discovery at the synchrotron: screening binding sites and correlations with hotspot mapping. *Philos Trans A Math Phys Eng Sci* **377**, 20180422 (2019).
2. P. J. McIntyre *et al.*, Characterization of Three Druggable Hot-Spots in the Aurora-A/TPX2 Interaction Using Biochemical, Biophysical, and Fragment-Based Approaches. *ACS Chem Biol* **12**, 2906-2914 (2017).
3. T. L. Blundell, Protein crystallography and drug discovery: recollections of knowledge exchange between academia and industry. *IUCrJ* **4**, 308-321 (2017).
4. C. H. Arrowsmith *et al.*, The promise and peril of chemical probes. *Nat Chem Biol* **11**, 536-541 (2015).
5. J. M. Grimes *et al.*, Where is crystallography going? *Acta Crystallogr D Struct Biol* **74**, 152-166 (2018).
6. A. Förster, C. Schulze-Briese, A shared vision for macromolecular crystallography over the next five years. *Structural Dynamics* **6**, (2019).