Report for MX-2415
Proposal Title: The Institute of Structural and Molecular Biology (ISMB) application for beam time, comprising groups from Birkbeck College, UCL and the Queen Mary University of London

Overall summary
MX-2415 was the first proposal after the covid disruption. Using data from ESRF we have a new publication in Science Advances, while few more publications from the groups Kozielski, Irving, and Larsen are already submitted, or they are in revision. Most importantly all groups participated in the current proposal have made significant progress in their projects and we are expecting more publications in the near future.

Below we are providing 2 highlights of our research as also a more detailed summary of the BAG activities.

Research Highlights

1. **Crystal structure of phospholipase Cγ with IP3 reveals key features of the active enzyme**

Katan M[a], Bunney T[a], Pinotsis N[b]
a. University College London, b. Birkbeck College

Phospholipase γ (PLCγ) enzymes are key components of intracellular signalling that are also linked to disease development. Additionally, to elucidation of an autoinhibited form, insights into physiological and aberrant activation of PLCγ requires understanding of an active, membrane bound form, capable of hydrolysing inositol-lipid substrates.
PLCγ enzymes are characterized by an array of domains, referred to as "γ-specific array (γ-SA)" or as "regulatory region", encompassing the sPH, nSH2, cSH2 and SH3 domains. The autoinhibitory interfaces lock the γ-SA on top of the PLC-core. The structures also suggest mechanisms of activation via physiologically relevant phosphorylation of PLCγ by tyrosine kinases and by gain-of-function mutations discovered across diverse pathologies. Despite initial studies suggesting the role of the PLCγ1 nPH domain in selective binding of PI(3,4,5)P3 and the involvement of PLCγ2 C2 domain in Ca2+-dependent translocation to the membrane, the overall, direct information about membrane interactions and their possible roles are lacking for PLCγ enzymes.
To gain further structural insights, we have obtained a new structure of PLCγ1, resolved to 2.0 Å, in its autoinhibited form and in the complex with inositol trisphosphate (IP3) that has not been previously described for this PLC (Fig. 1). Our structure reveals that the active site remains accessible to inositol phosphate, despite autoinhibition. The geometry of the active site is largely conserved compared to the structure of the related PLCδ1 - coordination of phosphate 1 of IP3 involves H335, H380 and Ca^{2+}, the position 4 and 5 phosphates engage in electrostatic interactions, including interaction of the position 4 phosphate with K462 and R1010, and the 2-hydroxyl group is ligated to Ca^{2+}. Despite the presence of excess IP3 in the crystallization buffer, IP3 binding to the nPH domain was not observed. This suggests that unlike PLCδ1, the nPH of PLCγ1 lacks a canonical PIP2 binding site. Similarly, Ca^{2+} binding to the C2 domain was not observed.
The structure was further used for molecular dynamics simulations demonstrating that PLCγ1 cannot bind membranes unless the autoinhibition is disrupted.

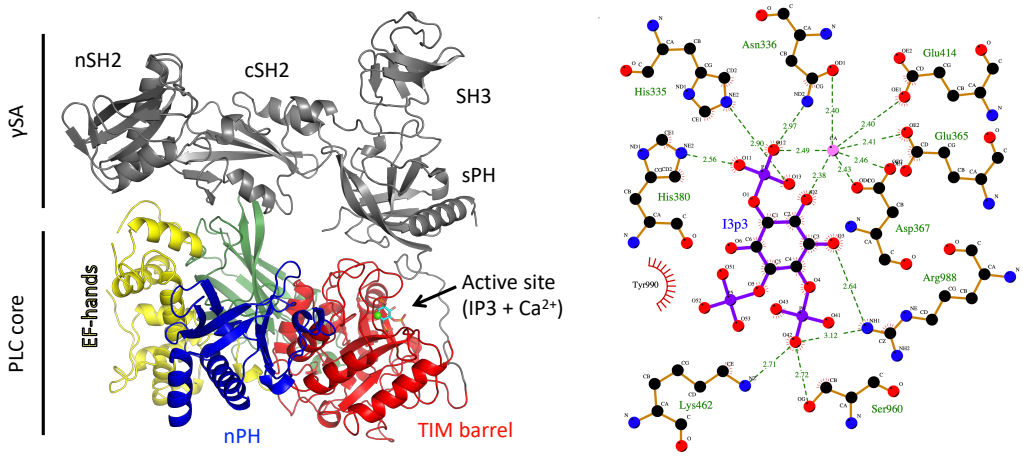

Figure 1: Domain layout of the rat PLCγ1 structure in complex with IP3 and Ca^{2+} at 2.0 Å resolution. Interactions of IP3 at the active site are shown on the right using a ligplot representation.
Reference: Le Huray K et al, Science Advances 2022, PDB ID: 7Z3J

2. Fragment-based screening of SARS-CoV-2 nsp1 reveals ligand-binding sites and cross-inhibition with other medically relevant coronaviruses

School of Pharmacy: Shumeng Ma, Shymaa Damfo, Jiaqi Lou, Haider and Frank Kozielski;
Birkbeck College: Nikos Pinotsis; ESRF: Matthew Bowler

Non-structural protein 1 (nsp1), a unique viral protein, has been shown to be a crucial virulence factor causing host mRNA degradation and suppressing interferon (IFN) expression as well as host antiviral signaling pathways. In view of the essential role of nsp1 in the coronavirus (CoV) life cycle, it is regarded as a potential target for antiviral agent discovery. No inhibitors of nsp1 have been reported do far.

Our project aims at identifying fragment hits as a starting point for the development of more potent analogues by fragment-based screening of SARS-CoV-2 nsp1 via x-ray crystallography. Subsequently, ligands binding sites were identified and characterised, followed by chemical optimisation of hits using structure-based drug design.

We expressed, purified and crystallised SARS-CoV-2 nsp1 and determined its structure improving the recently reported resolution from 1.6 Å to 0.95 Å. The space group is $P4_32_12$ with one molecule in the asymmetric unit. The crystals diffract routinely to below 1.6 Å resolution and are stable in the presence of DMSO for several hours, excellent conditions for fragment-based screening using x-ray crystallography at ESRF beamline Massif-1. After optimising soaking and cryo-conditions we collected datasets for ca. 600 nsp1-fragments soaks and ca. 40 native crystals for the ground state in batches of 200 crystals in automated mode. Automated data collection parameters were further optimised between batches to increase data quality.

We identified, verified, and confirmed 9 nsp1-fragment complexes so far through PanDDA analysis. This number is at the lower end of success rates reported for fragment-based screening, however, it should be noted that nsp1 is a relatively small protein of 17 kDa (Figure 1).

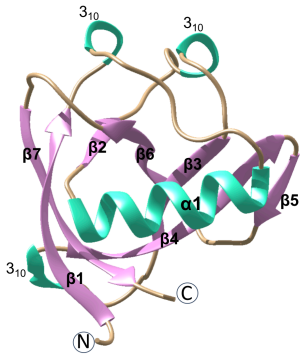

Figure 1: *The secondary structure elements at 0.95 Å resolution are depicted in different colours with α-helixes coloured in green, β-strands shaded in purple, and loops are shown in tan.*

We characterised the nsp1-fragment interactions using orthogonal biophysical methods such as thermal shift assays (TSA) and microscale thermophoresis (MST) assays. In particular, MST allowed us to determine K_d values and rank the fragment hits. We have also expressed and purified nsp1 from SARS and MERS and set up MST assays, allowing us to detect potential cross-inhibition with other CoVs of medical relevance. In parallel we initiated "SAR-by-catalogue" and started testing commercially available fragment analogues as a first step to improve potency of the fragments. We hope to be able to contribute to the development of potent inhibitors targeting SARS-CoV-2 nsp1 and to the better understanding of this important protein [1-2].

This work just submitted to the International Journal of Molecular Sciences

Additional reports from usage of beamlines:

1. Irving group (UCL)

Use of the beamtime over the last 6 months

(1) Crystals were screened of a conformation ('latent') of AAT that mimicks the subunits of the pathogenic aggregates that underlie alpha-1-antitrypsin deficiency. These ordered aggregates are not suitable for crystallography because they are flexible and have heterogenous size. Diffraction was successfully obtained to ~4/3 Å when processed isotropically/anisotropically, respectively. Molecular replacement provided a solution in the P1 space group with 8 molecules in the asymmetric unit.

(2) We have developed a toolkit of monoclonal antibodies that are able to recognise this protein selectively in different conformational states. As well as informing on the structural changes that result in aggregation and loss of activity, these antibodies have significant potential as clinical diagnostic tools. The 9C5 monoclonal antibody is used routinely in our lab in immunoassays, cryo-EM and has potential as a clinical diagnostic tool. Unsuccessful attempts were made to obtain a dataset of a complex between Fab9C5 and AAT with a resolution better than the current ~3.9 Å.

(3) Based on an earlier dataset collected at the ESRF that provided a high-resolution structure of a complex between a monoclonal antibody and AAT, small molecules were shortlisted in silico and screened by NMR with the aim of developing a non-invasive clinical imaging reagent. Small molecules with binding activity were soaked into crystals of native and cleaved AAT. Datasets were collected between 2-4 Å, with the identification of an improved cryo-condition at the same time, however no unequivocal small molecule density has yet been identified.

(4) Neuroserpin is a member of the same serpin family as AAT and is expressed in developing neurons. Mutants of this protein are associated with severe familial dementia. In support of a cryo-EM project in which aggregates of neuroserpin are labelled with an antibody ('1A10') Fab domain, we have performed crystallisation trials of the Fab alone and in complex with wild-type neuroserpin. A dataset was obtained of 1A10 Fab alone to 3.6 Å(iso)/3.0Å(aniso), and we identified very low resolution diffraction (>10 Å) in crystals of the complex which now need to be improved.

2. Frank Kozielski (School of Pharmacy, UCL)

Group's additional report details

Human MCAK structures in the absence (nucleotide-free) and presence of various nucleotides (Mg^{2+}ADP, ATP-like states represented by AMP-PNP and ATPγS) have been determined and the structures are being prepared for submission. To finish the manuscript, we are attempting to increase the resolution of one of the structures, the MCAK-AMP-PNP complex, which is currently at 3.3 Å to below 3.0 Å, which delays the manuscript a bit. We are also attempting to obtain crystal structures of MCAK-AMP-PCP and MCAK-'ready-to-split' complexes. The manuscript is in preparation [3]. For MCAK fragment screening, we have previously identified 60 fragment hits by biophysical assays and plan to obtain the MCAK-fragment complexes in the near future.

We have collected data on a large variety of crystals of ***M. tuberculosis*** **FtsZ** with fragments. These complex structures reveal novel inhibitor binding sites and will allow conducting structure-based drug design (in progress). These complex structures will serve as a basis to develop inhibitors with higher potency and efficacy.

We conducted fragment-based screening of **dengue virus** and **zika virus methyltransferase (MTase)** at DLS collecting diffraction data of around 700 MTAse-fragment soaks. The data allowed us to identify a range of novel ligand binding sites on MTase. As the crystals showed twinning and contained eight molecules in the asymmetric unit, which led to very long refinement for each individual complex, we searched for novel crystallisation conditions. Shymaa Damfo, the PhD student working on the project was able to grow a new crystal form, space group $C222_1$, diffracting routinely to better than 2.0 Å resolution. We therefore switched to this untwinned crystal form collecting native diffraction data and three MTase-fragment analogue complexes at the ESRF. These structures will allow structure-based drug design (in progress) and will serve as a basis to develop MTase-targeting inhibitors with higher potency and efficacy. We also remeasured untwinned crystals of dengue MTase in the presence of all 39 fragment hits at the ESRF and processing of all datasets is currently in progress. The data collected at the ESRF will be included in the second manuscript on dengue MTase, once we published the initial fragment screen.

We also collected data at Massif beamtime for **dengue RdRp** in complex with fragments (about 200 data sets). Data have been processed using the PanDDA software from DLS and seven fragments binding to DENV RdRp have been identified [4]. In addition we also identified novel ligand binding sites on dengue virus MTase.

We collected several datasets of **SARS-CoV-2 main protease (nsp5)** to a resolution of 1.6 Å in space group $P2_12_12_1$. We soaked or co-crystallised nsp5 crystals with three potential new inhibitors, but did not observe any electron density for the inhibitors. We are in the process of establishing novel collaborations with medicinal chemists to develop potent nsp5 inhibitors using structure-based drug design.

We collected several datasets of **SARS-CoV-2 nsp15** to the best resolution of 1.9 Å in space group $P6_3$. However, datasets are not of high quality due to imperfect crystals. We are in the process of screening for better crystallisation conditions to obtian better diffraction data of the protein crystals.

We also conducted fragment-based screening via x-ray crystallography for SARS-CoV-2 **nsp10**. Overall, the nsp10 fragment screening project is slightly less successful than nsp1 as we profited from the mistakes done during the nsp10 project: although still routinely diffracting to better than 2.0 A, the resolution of the nsp10 crystals is not as good as for the nsp1 crystals. Nsp10 crystals are also more difficult to soak as drops sometimes form a skin preventing the fragments to soak into the drops. The reproducibility of nsp10 crystals is also an issue. Initially, we soaked cocktails of five fragments at 5 mM concentration (a total of 500 fragments), but did not identify any bound fragment. We therefore avoid measuring fragment cocktails, which is more work and requires more beamtime, but this represents the more successful approach. Nevertheless, we managed to collect ca. 600 individual fragments soaked with individual fragments and identified and verified ten fragment hits.

Deposited structures of SARS-CoV-2 nsp1 with fragments (HPUB):
8A55 (native, atomic resolution), 8ASQ, 8AYW, 8AZ8, 8AYS, 8AZ9, 8A4Y

List of potential publications attributable to ESRF time in the last 12 months

[1] Ma, S.; Damfo, S.; Lou, J.; Pinotsis, N.; Bowler, M. W.; Haider, S.; Kozielski, F., Two ligand-binding sites on SARS-CoV-2 non-structural protein 1 revealed by fragment-based x-ray screening. BioRxiv 2022, 2022.06.12.495816.

[2] Borsatto, A.; Akkad, O.; Galdadas, I.; Ma, S.; Damfo, S.; Haider, S.; Kozielski, F.; Estarellas, C.; Gervasio, F. L., Revealing druggable cryptic pockets in the Nsp-1 of SARS-CoV-2 and other β-coronaviruses by simulations and crystallography. BioRxiv 2022, 2022.05.20.492819. In revision in eLife

[3] Sandeep K. Talapatra, Jiaqi Lou and Frank Kozielski. Insights into the catalytic cycle of a kinesin depolymerase: MCAK – How structure changes affect microtubule binding and function. *Manuscript in preparation*.

[4] Sandeep K. Talapatra, Jiaqi Lou, Shymaa Damfo, Bethan Howells, Rick A. Davies and Frank Kozielski. Exploring the palm and thumb inhibitor-binding pocket of dengue virus RNA-dependent RNA polymerase using x-ray-based fragment screening. *Manuscript in preparation.*

[5] Jiaqi Lou, Shozeb Haider, Shymaa Damfo, Geoff Wells and Frank Kozielski. Fragment analogues against SARS-CoV-2 nsp10 reveal nsp14-nsp10 and nsp16-nsp10 inhibitors in the protein-protein binding interface. *Manuscript in preparation.*

3. Larsen Team (Towers/Selwood groups UCL)

The work of the Larsen team is focusing on new inhibitors for the HIV capsids. In a second project they are looking for interactions of the SARS-CoV-2 protein ORF6 (p6) with human proteins in order to develop efficient drugs for the treatment of the virus infection:
Ten data sets were collected on id30B and id30A-1 of the HIV-1(M) capsid protein (CA) soaked with peptides derived from known co-factors(Nup153, Sec24C) as well as a new drug. Unfortunately, despite using a proven soaking protocol, the ligands were not observed within the solved structures.
Furthermore, one dataset of Human Importin (KPNA2) soaked with a peptide derived from the HIV-1(M) protein Vpr were collected, but the Vpr peptide was not seen in the obtained crystal.

4. Gouge Team (UCL/Birkbeck)

Cells are constantly subjected to oxidants that in turn can damage all their constituents. They therefore rely on complex systems to sense the redox balance and activate the detoxification machineries. Failing at maintaining the redox balance is a hallmark of human diseases, including cancer and neurodegenerative diseases. We are interested in the major regulator of the oxidative stress response in eukaryotes, the **Nrf2 pathway**. We are also interested in a newly identified post-translational modification occurring during oxidative stress. Under oxidative conditions, co-enzyme A (CoA) can react with cysteines in proteins to protect them against further oxidation, alter their function, their oligomeric state or cellular location. We are focussing our attention on two proteins shown to be derivatised by CoA, Prdx6 and NME-1.

Results

We obtained crystals of the repressed and active form of the promoters that encompass Nrf2, an accessory protein and DNA. Over the last year, we screened about 70 crystals with different DNA lengths. Unfortunately, we didn't observe diffraction below 7Å. Future

approaches will focus on screening more oligonucleotides, microseeding matrix screening, cryo-protection optimisations, and crystal dehydration.
Concerning Prdx6 and NME-1, we screened about 80 crystals and collected about 30 datasets between 2 and 2.5Å resolution for both proteins. Unfortunately, only residual density was found in the composite map for Prdx6. We did identify CoA binding site in NME-1, but couldn't observe covalent bond between CoA and the protein. We are currently exploring the relationship between oligomeric state and derivatisation by CoA.

5. Wallace group (Birkbeck College)

The Wallace Group's work is focused on elucidating the nature of the transmembrane drug interactions with voltage-gated sodium channels. These studies include comparisons of native and mutant sodium channels, where the mutations were designed to inhibit the drug binding. Many data sets have been collected at the ESRF to screen for drug complexes in the previous period, but none of these have led to data that was included in any publications. Most of our work in this period built on the previous period, where we developed methods to allow protein-drug interactions to occur but without devasting loss of diffraction. This has required at lot of screening. Using this approach, we have improved things to the point that we now have high quality datasets for some protein-drug interactions, and while others remain elusive data are gradually improving as we learn more using our time at the synchrotrons. Once we publish papers about these complexes, (even if the final datasets do not come from the ESRF), we will acknowledge use of all the facilities used in the evolution of the projects as they will have all contributed to the final outcomes.

6. Parkinson Group (School of Pharmacy, UCL)

Data collected on i-motifs in the current cycle (mx2415) relates to use of id30b on the 22nd of July 2022. Data on NOS i-motifs have been collected to 2.2 Å resolution and we are now attempting structure solutions. Screening was also undertaken on ILPR soaked with cobalt, but these crystals did not diffract. Additionally, HJ/protein complexes were screened but did not diffract to sufficient resolution of the collection of data.

7. Lamber group (Kingston University), Pinotsis (Birkbeck College)

Our research is focusing on the myomesin protein family which is an essential protein for striated muscle contraction. Myomesin proteins bind to the thick filaments and keep them align around M-band which is the centre of the sarcomeric unit.
Here we initially measured the C-terminal domain 12-13 of myomesin-2 a protein overexpressed in heart and fast twitch muscles. Although we did an extensive screening using several beamlines, we obtained a single data set to a maximum 3.41 Å resolution. Although the resolution is relatively low, we were able to solve the structure revealing a similar motif to myomesin-1. We also observe minor conformational differences of the overall architecture suggesting protein plasticity.
In a second project we measured the C-terminal domains 12-13 of myomesin-1 including the C-terminal linker which was not previously included in the already published structures. These data were collected in ID23-1 and they diffract to a maximum resolution of 2.35 Å. The structure was solved by molecular replacement and currently is getting refined.