Journal of Visualized Experiments

Workflow and Tools forCrystallographic Fragment Screening at the Helmholtz-Zentrum Berlin

--Manuscript Draft--

Article Type:	Invited Methods Collection - JoVE Produced Video
Manuscript Number:	JoVE62208R1
Full Title:	Workflow and Tools forCrystallographic Fragment Screening at the Helmholtz-Zentrum Berlin
Corresponding Author:	Manfred S Weiss, Ph.D. Helmholtz-Zentrum Berlin fur Materialien und Energie GmbH Berlin, GERMANY
Corresponding Author's Institution:	Helmholtz-Zentrum Berlin fur Materialien und Energie GmbH
Corresponding Author E-Mail:	msweiss@helmholtz-berlin.de
Order of Authors:	Jan Wollenhaupt
	Tatjana Barthel
	Gustavo Lima
	Alexander Metz
	Dirk Wallacher
	Elmir Jagudin
	Franziska Huschmann
	Thomas Haußß
	Christian Feiler
	Martin Gerlach
	Michael Hellmig
	Ronald Förster
	Michael Steffien
	Andreas Heine
	Gerhard Klebe
	Uwe Mueller
	Manfred S Weiss, Ph.D.
Additional Information:	
Question	**Response**
Please specify the section of the submitted manuscript.	Biochemistry
Please indicate whether this article will be Standard Access or Open Access.	Open Access (US$4,200)
Please indicate the **city, state/province, and country** where this article will be **filmed**. Please do not use abbreviations.	Berlin, Germany
Please confirm that you have read and agree to the terms and conditions of the author license agreement that applies	I agree to the Author License Agreement

below:	
Please provide any comments to the journal here.	iNEXT Discovery Issue

TITLE:

Crystallographic Fragment Screening at the Helmholtz-Zentrum Berlin

AUTHORS AND AFFILIATIONS:

Jan Wollenhaupt[a], Tatjana Barthel[a,b], Gustavo M. A. Lima[c], Alexander Metz[d], Dirk Wallacher[e], Elmir Jagudin[c], Franziska U. Huschmann[a,d], Thomas Hauß[a], Christian G. Feiler[a], Martin Gerlach[a], Michael Hellmig[a], Ronald Förster[a], Michael Steffien[a], Andreas Heine[d], Gerhard Klebe[d], Uwe Mueller[a], Manfred S. Weiss[a*]

[a]Helmholtz-Zentrum Berlin, Macromolecular Crystallography, Albert-Einstein-Str, Berlin, Germany
[b]Freie Universität Berlin, Institute for Chemistry and Biochemistry, Structural Biochemistry Group, Takustr, Berlin, Germany
[c]MAX IV Laboratory, BioMAX, Fotongatan,Lund, Sweden
[d]Philipps-Universität Marburg, Institute of Pharmaceutical Chemistry, Drug Design Group, Marbacher Weg, Marburg, Germany
[e]Helmholtz-Zentrum Berlin, Department Sample Environment, Hahn-Meitner-Platz, Berlin, Germany

Email addresses of co-authors:

Jan Wollenhaupt	(jan.wollenhaupt@helmholtz-berlin.de)
Tatjana Barthel	(tatjana.barthel@helmholtz-berlin.de)
Gustavo M. A. Lima	(gustavo.alvares.lima@gmail.com)
Alexander Metz	(alexander.metz@idorsia.com)
Dirk Wallacher	(dirk.wallacher@helmholtz-berlin.de)
Elmir Jagudin	(elmir.jagudin@maxiv.lu.se)
Franziska U. Huschmann	(franziska.huschmann@gmx.de)
Thomas Hauß	(hauss@helmholtz-berlin.de)
Christian G. Feiler	(christian.feiler@helmholtz-berlin.de)
Martin Gerlach	(gerlach-reise@gmx.de)
Michael Hellmig	(michael.hellmig@helmholtz-berlin.de)
Ronald Förster	(foerster@helmholtz-berlin.de)
Michael Steffien	(michael.steffien@helmholtz-berlin.de)
Andreas Heine	(heinea@mailer.uni-marburg.de)
Gerhard Klebe	(klebe@staff.uni-marburg.de)
Uwe Mueller	(uwe.mueller@helmholtz-berlin.de)
Manfred S. Weiss	(msweiss@helmholtz-berlin.de)

*Corresponding author:

Manfred S. Weiss	(msweiss@helmholtz-berlin.de)

KEYWORDS:

crystallographic fragment screening, compound library, crystal handling, ligand soaking, FBDD, macromolecular crystallography, data processing, hit identification

SUMMARY:

Crystallographic fragment screening at the Helmholtz-Zentrum Berlin is performed using a workflow with dedicated compound libraries, crystal handling tools, fast data collection facilities, and largely automated data analysis. The presented protocol intends to maximize the output of such experiments to provide promising starting points for downstream structure-based ligand design.

ABSTRACT:

Fragment screening is a technique that helps to identify promising starting points for ligand design. Given that crystals of the target protein are available and display reproducibly high-resolution X-ray diffraction properties, crystallography is among the most preferred methods for fragment screening because of its sensitivity. Additionally, it is the only method providing detailed 3D information of the binding mode of the fragment, which is vital for subsequent rational compound evolution. The routine use of the method depends on the availability of suitable fragment libraries, dedicated means to handle large numbers of samples, state-of-the-art synchrotron beamlines for fast diffraction measurements, and largely automated solutions for the analysis of the results. Here, the complete practical workflow and the included tools on how to conduct crystallographic fragment screening (CFS) at the Helmholtz-Zentrum Berlin (HZB) are presented. Preceding this workflow, crystal soaking conditions as well as data collection strategies are optimized for reproducible crystallographic experiments. Then, typically in a one-to-two-days procedure, a 96-membered CFS-focused library provided as dried ready-to-use plates is employed to soak 192 crystals, which are then flash-cooled individually. The final diffraction experiments can be performed within a day at the robot-mounting supported beamlines BL14.1 and BL14.2 at the BESSY II electron storage ring operated by the HZB in Berlin-Adlershof (Germany). Processing of the crystallographic data, refinement of the protein structures, and hit identification is fast and largely automated using specialized software pipelines on dedicated servers, requiring little user input. Using the CFS workflow at the HZB enables routine screening experiments. It increases the chances for successful identification of fragment hits as starting points to develop more potent binders, useful for pharmacological or biochemical applications.

INTRODUCTION:

The first step in drug development is the screening of compounds against a target of interest. Traditionally, large compound libraries in the order of 10^5–10^6 entries are used in high-throughput biochemical assays in the pharmaceutical industry. This strategy was complemented by fragment-based drug design (FBDD), a newer method that took a steep rise during the last 20 years and became a mainstream strategy to generate high-quality lead candidates due to several inherent advantages[1]. The term fragment refers to a small organic molecule containing typically less than 20 non-hydrogen or heavy atoms (HAs). Thus, a fragment is significantly smaller than the drug- or lead-like molecules (usually less than 30 HAs) explored in conventional high-throughput screening. Fragments are weak-affinity binders. However, compared to larger molecules, fragments are more versatile, since even a small collection of them can represent the respective chemical space of molecules of the same size much better[2]. Also, evolving fragment

screening hits into lead molecules is considerably more effective than optimizing already larger molecules[2–5]. That means, pending sufficient sensitivity of the detection, screening of fragments can be employed efficiently and yields high-quality starting points for further compound evolution. Several biophysical methods may be applied for fragment screening, the most popular being nuclear magnetic resonance, X-ray crystallography, surface plasmon resonance and thermal shift assays. These methods are used either in a parallel or a sequential way, with the aim to increase the confidence in the hits and reduce the number of false positives or false negatives, respectively. However, a recently conducted comparative study[6] suggested that sequential screening cascades are to be avoided due to the low overlap between the different methods.

X-ray crystallography is well established as a method for structure determination at atomic detail but has recently also been developed as a tool for screening purposes[7,8]. As protein crystals tolerate high fragment concentrations (e.g., 100 mM), crystallographic fragment screening (CFS) can compete with other biophysical methods for screening fragments or even outperform them as a first-step screening method[6,9]. However, a vital pre-requisite for CFS is a validated crystallization system of the target protein reproducibly delivering crystals with diffraction properties to considerably high resolution, typically better than 2 Å.

An exclusive benefit of CFS compared to all other fragment screening methodologies is the provision of detailed 3D information about the binding mode of the identified fragments. This structural information is absolutely crucial for the rational optimization of the fragment hits to higher-affinity binders. Established elaboration strategies are growing, merging, and linking fragment hits[5]. Thereby, relatively high ligand efficiency is provided from the start, and the introduction of unnecessary or spatially not suitable groups can be avoided, thus reducing chemical synthesis costs. All in all, CFS has unrivaled advantages as a starting strategy for drug design.

Given that a particular biological target meets the high requirements of CFS regarding crystal quality, what are the main factors maximizing the chances for a successful outcome of such screening campaigns? It depends on the quality of the fragment library used, on an efficient workflow to carry out the experiments before the diffraction experiment, on synchrotron beamlines with sufficient automation and data collection speed, as well as ways and means for largely automated data processing and analysis. Here, the complete workflow from the crystal soaking experiments to the hit identification is presented (**Figure 1**), in the way it is successfully established at the macromolecular crystallography beamlines at BESSY II. The facility is open to academic and industrial users for collaboration. Additionally, academic users of the countries under EU outside Germany can straightforwardly apply for funding via the iNEXT Discovery project.

There are indispensable prerequisites to be able to start a CFS campaign and conduct the protocol outlined in this work: Well-diffracting crystals of the target protein are available that can be reproducibly grown in large numbers, that are stable at ambient temperature, and that the crystallization cocktail employed does not contain highly volatile ingredients. Another

prerequisite is the suitability of the crystal lattice for the experiment. In an appropriate lattice, the interesting sites of the target protein must be exposed toward the solvent channels and thus be accessible. Another preceding step that is optional but nevertheless highly recommended to ensure success in the workflow of the CFS campaign is the optimization of the soaking condition for the experiment. Vital benchmark statistics here are the diffraction power of the crystal and the relevant data quality indicators, which are determined during the data scaling statistics procedure. Typical factors to optimize are Dimethyl sulfoxide (DMSO)-tolerance, buffer concentration, and cryo-protectant. Although not a strict prerequisite as further detailed below, DMSO as a co-solvent can help to increase fragment solubilization. Typical tests should include soaking of 0%, 3%, 6%, or 10% (v/v) DMSO overnight. An increase of the buffer concentration to 200 or 300 mM helps to prevent loss in diffraction quality due to occasional pH-shifting effects arising from the high fragment concentrations to be used. Finally, it is decisive to find out whether and which additional cryoprotectant is required and whether it can be already included in the soaking condition. In many cases, however, an additional cryo-protectant is not needed, because DMSO itself can act as a cryo-protectant. If so, this will save one handling step in the final experiment. Most crystals need less cryo-protectant if flash-cooled on appropriately sized loops, minimizing or avoiding surrounding mother liquor as much as possible. However, in rare cases, a layer of the mother liquor is indeed necessary to prevent damage to the crystal upon flash cooling.

The number of hits obtained in a CFS campaign is not only dependent on the druggability of the target protein and the suitability of the crystal lattice (see above) but is also dependent on the quality of the library. Library quality comprises two aspects: the selection of the compounds for the library and the confectioning of the compounds, i.e., in which physical form they are presented for the experiment. For compound selection, different strategies can be employed. Most library designs include the maximization of the chemical diversity of the fragments. A strategic focus could be to include the chemical tractability of the fragments for follow-up design, which has been applied, for instance, in the DSI poised library[10]. Yet another strategic focus for library design could be to maximize the representation of commercially available chemical space of fragments by shape- and pharmacophore-based clustering, as has been exemplified by the F2X libraries developed at HZB[11]. More specifically, the 1103-membered F2X-Universal Library and representative 96-compound subset for initial CFS campaigns, which is called F2X-Entry Screen, have been developed and the F2X-Entry Screen has been validated successfully[11]. The F2X-Entry Screen is the primary choice for CFS campaigns at HZB. Subsequent, larger campaigns can then be carried out using the F2X-Universal Library or the 1056-membered EU-OPENSCREEN fragment library[12] that is also being offered at HZB. At present, these libraries are available for users of the macromolecular crystallography beamlines of the BESSY II synchrotron in Berlin free-of-charge on the basis of a collaboration contract. That also applies to users via iNEXT Discovery proposals. Moreover, the F2X-Entry Screen is available to all interested scientists on the basis of a material transfer agreement.

With respect to the physical presentation of a library, two approaches are commonly adopted: the fragments are either used as DMSO stock solutions or the fragments are dried and immobilized on ready-to-use plates. At HZB, both the F2X-Entry Screen and the non-volatile

compounds of the F2X-Universal Library are presented as dried-in compounds in a 3-lens 96-well MRC low profile crystallization plate. The presentation of the fragments immobilized in crystallization plates has two vital advantages: Firstly, it allows transport of the screening plates to the user's home lab. Therefore, the soaking and crystal handling steps of the workflow presented here (steps 1–3) can be carried out anywhere. Secondly, DMSO-free solution can be employed. DMSO-sensitive targets can thus be screened easily, largely retaining expected hit rates[11]. However, DMSO does increase fragment solubility, hence, it is worthwhile to check the DMSO tolerance of a crystal system of choice beforehand as outlined above.

The protocol outlined below will describe a typical experiment with a 96-compound screen such as the F2X-Entry Screen. For that, approximately 250 crystals need to be prepared in time to be used freshly. It is highly advisable to prepare the soaks for all 96 compounds in duplicate. It is recommended but optional to prepare additional mock-soaks that will later help with data analysis using the pan-data density analysis (PanDDA) approach for hit identification[13]. Mock-soaks are defined as soaking experiments on protein crystals using the same soaking solution as the fragment soaks for the same incubation time, with no fragments present. If the soaking solution is equal to the crystallization condition, the crystals may be directly harvested from the crystallization plate.

Dependent on the capabilities of the robotic sample changer, different puck formats may have to be used. At the moment, samples for the HZB-operated beamline BL14.1 need to be prepared in Unipuck format, while samples for the HZB-operated beamline BL14.2 need to be prepared in SPINE puck format. In this protocol, preparation in Unipuck format is assumed.

PROTOCOL:

1. Soaking crystals

1.1. Take the screening plate (**Figure 2**) from the -20 °C freezer and place it onto the bench/table for about 30 min to pre-warm it to room temperature to avoid condensation moisture.

1.2. Set up the working place with two closely arranged microscopes and all tools needed (**Figure 3A**).

1.3. Choose 3–4 loops of the appropriate size for transfer of the crystals to be soaked and place them close to the microscopes.

1.4. Fill the glass spot plate cavities with de-ionized or distilled water.

1.5. Prepare a 5 mL soaking solution.

1.6. Cut open the bag of the screening plate pre-warmed to room temperature.

1.7. Remove the lid and the foil from the screening plate and place the plate on the surface of the bench/table.

1.8. Decant the 5 mL soaking solution in the reagent reservoir.

1.9. Fill each of the 96 reservoirs with 40 µL soaking solution using a 12-channel pipette.

1.10. Place the EasyAccess Frame on top of the screening plate and secure it with the included clamps by sliding them onto the left and the right side of the device.

NOTE: The EasyAccess Frame is a special device for handling multiple crystals, which was developed at the HZB[14]. It enables easy access to each well by shifting the movable tiles while protecting the other wells from evaporation.

1.11. Place the screening plate (including the EasyAccess Frame) under one microscope and the crystallization plate (including the crystals to be soaked) under the other microscope.

1.12. Slide open well A1 of the screening plate by moving the respective acrylic glass tile of the Easy Access Frame either with a finger or the supplied pen tool.

1.13. Add 0.4 µL soaking solution from the reservoir to the fragment containing well (upper-left lens) using a fresh pipette tip. Ensure through the microscope that the drop covers the dried-on fragment to allow it to dissolve.

NOTE: Alternatively, this step can be carried out using a pipetting robot before the assembly of the EasyAccess Frame. This way the soaking drops of all wells could be placed in one automatic procedure. However, it is recommended to add the soaking solution directly before the soaking step as described to ensure that the fragment solubilizes slowly and in the presence of the crystal. This avoids the crystal from experiencing a sudden shock upon transfer of a drop with a high fragment concentration.

1.14. Under the second microscope, cut open the sealing foil of the crystallization plate at one of the wells that contains the target crystals.

1.15. Transfer two crystals using an appropriately sized loop mounted on the crystal wand to the well A1 of the screening plate under the first microscope.

1.16. Wash the loop in the prepared glass spot plate and dry it by gently touching it with the tissue. Do this after every transfer to avoid cross contamination with the fragment containing soaking solutions.

1.17. Use the microscope to ensure that the crystals have been properly placed.

1.18. Move on to next well (e.g., B1).

1.19. Repeat steps 1.13–1.18 with all 96 wells of the screening plate until each soaking drop contains two crystals.

1.20. Remove the screening plate along with the EasyAccess Frame from under the microscope and place it onto the bench/table.

1.21. Remove the EasyAccess Frame from the screening plate.

1.22. Seal the screening plate with sealing foil and place it in the crystallization incubator or cupboard, respectively, where the crystals were grown.

1.23. Incubate for the optimized soaking time. Overnight is usually convenient.

1.24. Preparation of approximately 40 apo crystals (i.e., mock soaking) (optional step)

1.24.1. Take a 3-lens 96-well low-profile crystallization plate and fill two columns with 40 µL soaking solution per well using the 12-channel pipette.

1.24.2. Place the EasyAccess Frame on top of the crystallization plate and secure it with the included clamps by sliding them onto the left and right side of the device.

1.24.3. Slide open the acrylic glass tile of well A1.

1.24.4. Place a drop of 0.4 µL soaking solution in each of the two left lenses of the well.

1.24.5. Transfer 2–3 crystals to each drop. After each transfer, wash the loop in the prepared glass spot plate and dry by gently touching it with the tissue.

1.24.6. Move to the next well (e.g., B1).

1.24.7. Repeat steps 1.24.4–1.24.6 until about 40 crystals are ready for incubation.

1.24.8. Remove the crystallization plate (including the EasyAccess Frame) from under the microscope onto the bench/table and remove the EasyAccess Frame.

1.24.9. Seal the crystallization plate with sealing foil and place it the aforementioned crystallization incubator or cupboard.

1.24.10. Incubate for the same time as the screening plate.

2. Harvesting crystals

2.1. Take out the incubating plate(s) from the incubator or cupboard.

2.2. Set up the working place with one microscope and all the tools needed (**Figure 3B**).

2.3. Prepare the Unipuck foam dewar with 3 unipuck lids (i.e., sample enclosures and fill it with liquid nitrogen (LN2).

NOTE: Beware of the appropriate safety precautions for working with LN2, i.e., wear safety goggles and use suitable protective equipment. It is best to get fresh LN2 several times during the session to avoid water condensation in the LN2 storage can. Through the entire procedure, make sure the LN2 level in the foam dewar is always reaching the upper edge of the dewar. Also ensure that the LN2 is ice-free, i.e., frequently replace the LN2 (e.g., once per 45 min or latest if ice starts to accumulate). Then, fill the second foam dewar and transfer the Unipucks there. Empty the icy foam dewar and remove residual ice and moisture with the blow dryer.

2.4. Remove foil from the screening plate and place the EasyAccess Frame on top.

2.5. Slide open well A1.

2.6. Harvest two crystals from the drop and flash-cool them in LN2 (one by one) by plunging with a fast vertical movement into the LN2 and then insert the sample in the proper puck position. Take relevant notes on the sample tracking sheet.

2.7. Cryoprotection step (if necessary, for the target crystals). In such a case, perform this step instead of 2.6.

2.7.1. Place 0.4 µL soaking solution, including cryo-protectant on the lower left lens of the well.

2.7.2. Pull the loop with a crystal mounted from the drop in the upper-left lens slowly through the solution in the lower-left lens. Ensure that the crystal is in the loop, then flash-cool in LN2. Harvest two crystals in this manner.

NOTE: In steps 2.6 and 2.7, make sure to keep the time short so that the crystal is exposed to air while being in the loop. The plunging, i.e., the vertical drop of the sample in the LN2-filled dewar should be performed as fast as possible. This ensures high sample quality and prevention of ice rings in the data. Track the samples, i.e., note whether the crystals have damages etc., to prioritize either of the duplicates for the following X-ray measurements using the template. Even if the crystals have cracks, “hairs” or other defects due to the soaking, they can still be used and should always be harvested. In case crystals have broken into several pieces, two of the biggest/best looking pieces should be harvested. **Figure 5** shows some examples of what such crystals can look like. All the shown crystals gave still useful datasets in the respective campaign[11], underlining that it is worth to harvest crystals after soaking treatment, even if substantial morphologic changes occurred.

2.8. Go to the next well and repeat steps 2.5–2.6/2.7 until all three pucks are full.

2.9. Add the Unipuck bases on top of the lids after pre-cooling them in LN2.

2.10. Store the Unipucks in storage racks in a transport dewar or storage dewar.

2.11. To process all the wells of the screening plate, repeat the preceding steps.

2.12. In case mock-soaked crystals were prepared, harvest them in a similar fashion as described beforehand (optional step).

NOTE: If two crystals for each of the 96 conditions of the screening plate could be flash-cooled, there will be space for 32 mock-soaked apo crystals, to fill up the 14 Unipucks.

2.13. Store the Unipucks in LN2 until measurement.

3. Data collection

3.1. Transfer the universal pucks to the beamline BL14.1.

NOTE: In case SPINE pucks have been used in step 2, transfer them to beamline BL14.2.

3.2. Carry out standard measurements on the beamline, specific recommendations are given below.

NOTE: Details about the facility and the experiment control program MXCuBE2 have been presented previously[15,16]. **Figure 4** shows the interior of the experimental hutches of beamlines BL14.1 and BL14.2 as well as an example screenshot of the MXCuBE2 control software at beamline BL14.1.

3.3. Specific data collection recommendations deviating from standard experiment

3.3.1. To maximize time efficiency and throughput, do not collect test images.

3.3.2. Fix the sample-to-detector distance to a value that is suitable for the upper resolution limit of the crystal system determined in earlier experiments.

3.3.3. If the data collection strategy was not optimized beforehand, collect 1,800 images of 0.2 degrees each with an exposure time of 0.1 s per image.

3.3.4. Test the data collection strategy in the prior experiments using mock-soaked apo crystals.

NOTE: For higher symmetry space groups, 1,200 images or even 900 images, i.e., 240° or 180°, respectively will already give complete datasets with good statistics, independent of the starting

angle of data collection. Although higher redundancy and finer slicing can yield superior quality data[17], using this "enough but not more" strategy is an excellent trade-off between quality, data collection time, as well as computational requirements for analysis later on. Following the above steps, 200 data collections in 24 h is possible at beamlines BL14.1 and BL14.2. Nevertheless, samples should be prioritized.

3.3.5. To prioritize the sample

3.3.5.1. First collect diffraction datasets for one sample per fragment condition, based on the prioritization in step 2.6/2.7, i.e., collect the data for the higher prioritized duplicate.

3.3.5.2. For those experiments in 3.3.5.1 that suffered from data collection failure, loss of diffraction or severe icer rings, collect data for the second duplicate sample for the respective fragment condition.

3.3.5.3. Collect the diffraction data sets of apo crytsals (if prepared according to step 1.24).

3.3.5.4. Collect the diffraction datasets of the remaining duplicates of each fragment condition.

3.3.6. In the MXCuBE program, match the dataset identifiers of a CFS campaign to the following pattern: <protein>-<library>-[ABCDEFGH][01][0123456789][abcdef] (e.g., MyProtein-F2XEntry-B05a, where "B05" stands for the well, i.e., the fragment condition in the screen and the following "a" for the first duplicate).

4. Data treatment

4.1. For data analysis of the CFS campaign, use FragMAXapp (**Figure 6**) (Lima, G. M. A. et al. FragMAXapp, unpublished data), a web-based solution to control a multiplex analysis for processing auto-refinement and PanDDA hit evaluation of CFS data[18].

NOTE: In the FragMAXapp version deployed at HZB the following programs/pipelines are available: XDSAPP[19], Xia2-DIALS and Xia2-XDS[20], fspipeline[7], DIMPLE[21], Phenix LigFit[22], and PanDDA[13,23].

4.2. Use a well refined input model of the target as input for automatic refinement, otherwise perform meticulous refinement of one high resolution mock-soaked crystal that was collected during the campaign.

4.3. PanDDA hit identification

4.3.1. Use PanDDA to automatically calculate electron density maps of a set of data sets in a CFS campaign[13,23].

4.3.2. PanDDA assumes the electron density maps as non-binding fragment conditions and generates the average to form ground state model.

4.3.3. Furthermore, it uses the ground state model to derive local discrepancies between each electron density map and the ground state map, using voxel-associated Z-scores.

4.3.4. It performs a fine-tuned subtraction of ground state density from the respective map to create PanDDA-map for areas of high Z-scores to largely enhance the visibility of fragment binding events.

NOTE: To maximize the outcome of PanDDA, the authors recommend a two-step approach. Firstly, performing a PanDDA run (pandda.analyze) with standard settings. Even if mock-soaked crystals have been collected, their identity will not be included as a parameter (which is possible nonetheless) in order to enable an unbiased generation of the ground state model by PanDDA from all the available data. After that, the output data is evaluated by the user via a so-called PanDDA inspection in Coot[24]. Here, hits with relatively high confidence should be noted, concluding the first step. Secondly, the authors recommend re-running the pandda.analyze step excluding the preliminary hits (determined in the first step) from the ground state model via the --exclude_from_charactersation = **<list-of-bound-dataset-ids>** command line option. Further details are described on the PanDDA help pages (https://pandda.bitbucket.io/). This way, datasets that are clear hits and thus would obscure the ground state model if included there, are disregarded, leading to an improved ground state model and thus to improved results overall.

4.3.5. Finally, perform a thorough PanDDA inspection to complete the hit identification.

NOTE: FragMAXapp also includes an output option to save the modeled bound states or prepare data for PDB submission; for further details see FragMAX webpages (https://fragmax.github.io/).

REPRESENTATIVE RESULTS:

As part of the previously reported validation campaigns of the F2X-Entry Screen[11], three campaigns were conducted at the BioMAX beamline at MAX IV and one of the campaigns was conducted at beamline BL14.1 at HZB. In the latter campaign, a particular set of F2X-Entry Screen conditions using a soaking condition that did not contain DMSO was screened against the protein-protein complex of yeast Aar2 and the RNaseH-like domain of yeast Prp8 (AR). The selected set of conditions comprises the hits that were found in an earlier campaign of the F2X-Entry Screen against AR in a soaking condition containing DMSO[11], i.e., in the campaign performed at HZB those hits were re-screened in the absence of DMSO. **Figure 7** shows an overview of the hits obtained after analyzing the data with the FragMAXapp combination of XDSAPP for processing, fspipeline for auto-refinement, and subsequent hit finding using PanDDA.

FIGURE AND TABLE LEGENDS:

Figure 1: **Schematic representation of the workflow of a crystallographic fragment-screening (CFS) experiment with a focus on the special environment at the Helmholtz-Zentrum Berlin.**

Figure 2: **Formulation and packaging of the F2X-Entry Screen.** (**A**) The 96-compound screen is available on an MRC low-profile plate, sealed with foil and vacuum-packed. (**B**) The 96 compounds of the screen are dried from DMSO solutions in 2 of the 3 lenses of each well.

Figure 3: **Photography of the CFS workbench in the HZB preparation lab.** Assemblies of necessary tools for (**A**) soaking and for (**B**) crystal fishing are displayed.

Figure 4: **Data collection end stations and control software.** (**A**) Photograph of the experimental hutch of HZB-MX beamlines BL14.1 (left) and BL14.2 (right). (**B**) Screenshot of the MXCuBE experiment control interface used at BL14.1 for diffraction data collection. At BL14.2 a very similar interface is used.

Figure 5: **Photographic snapshots of some crystalline samples in cryogenic environment before data collection.** This illustrates the variability of morphologies of the crystals after performing the fragment soaking and crystal harvesting. The photographs were taken on the BioMAX beamline (MAX IV synchrotron, Lund, Sweden) for AR samples collected there as part of the F2X-Entry validation[11].

Figure 6: **Screenshot of the FragMaxApp installed at the HZB for convenient data analysis.**

Figure 7: **Overview of the results of the CFS campaign F2X-Entry vs. AR (without DMSO).** The AR protein complex is shown in cartoon view, with Aar2 colored in gray and the RNaseH-like domain of Prp8 colored in blue. The fragment hits of the campaign are colored in element colors (C: yellow, O: red, N: blue, S: orange, Cl: light cyan).

DISCUSSION:

For a successful CFS campaign, it is vital to adhere to the described prerequisites (see **Introduction**). A reliable crystallization system is needed for the reproducible growth of many well-diffracting crystals, and a well-refined structure is needed as the input apo model for automated refinement. It is also important to check that the target site on the protein (active site, or interface area) is accessible for fragments in the crystal lattice. It is crucial to optimize the soaking conditions beforehand to ensure that the soaking does not significantly deteriorate the crystal quality. Neglecting these conditions will very likely lead to a suboptimal experiment, which will be of limited use and will, in the worst case, require a repetition of the entire experiment.

The protocol described above outlines the procedures that are followed during a standard CFS campaign. If all prerequisites are met, at least 90% of all soaked crystals should display diffraction to high resolution in a diffraction experiment. If this is not the case, the soaking times may be shortened to a few hours or even minutes. Due to the good solubility of most of the fragments, this should suffice to obtain decent occupancy values. Also, a typical CFS campaign will result in a hit rate of roughly 10% or above. For the F2X-Entry Screen validation campaigns[11] and ongoing user campaigns with the same library hit rates, even higher have been observed (20% and above,

data not shown).

A general caveat of crystallographic fragment screening is the presence of crystallographic contact sites. These could either occlude priori known active sites (to be checked before the screening, see above), or these contact sites also often provide pockets and hot spots where fragments can bind. Such fragment hits will be artifacts of the crystallization lattice and will likely not bind to the protein in solution. However, this is for most projects usually the minor portion of the hits obtained. For example, in the F2X-Entry Screen validation campaign using endothiapepsin (EP) and the spliceosomal protein-protein complex of $Prp8^{RNaseH}$ and Aar2 (AR), most of the hits occurred in promising sites[11]. For EP, 27 out of the 37 observed binding events were located in the active site, i.e., the peptide cleft of this protease. The 10 remote binding events comprise two solvent exposed binding events and eight crystal contact binding events (corresponding to five unique hits). Excluding those crystal contact hits would still reflect an overall rate of 24% unique hits for the EP campaign. It is also important to notice that binding events remote of a known active site (except crystal contact binders) could also potentially be interesting, e.g., revealing new hot spots or allosteric sites of the protein. For the AR campaign (in the same publication), of the 23 observed binding events, seven were located at crystal contacts, one was located at the direct interface of the two proteins, seven were located at known protein-protein interaction sites with other binding partners of the larger biological context (hence different assembly stages of the spliceosome), eight binding events revealed two hot spots on AR of yet unknown function and one being at a solvent exposed surface of $Prp8^{RNaseH}$. Therefore, excluding the events at crystal contacts and the $Prp8^{RNaseH}$ singleton, the number of potentially useful binding events is 15 (corresponding to 14 unique hits), thus a hit rate of 15.6%. These hits can be starting points for design of protein-protein interaction modulators or for tool compounds aimed to explore the two discovered Aar2 hot spots. Taken together, also in line with conducted user campaigns, often only a minor portion of hits in crystallographic fragment screening must be disregarded as artefacts. However, this will also be largely target dependent.

If the hit rate is significantly lower, this may indicate one of the following problems related to the target protein. For instance, in a CFS campaign against a viral cysteine protease a hit rate of only 3% was observed (data not shown). It turned out that the protein used was likely chemically modified in its active site. In such a case, a different protein preparation may solve the problem. If crystals are very DMSO intolerant, the F2X-Entry Screen may also be used without DMSO, although the results may differ slightly. Most of the hits obtained in the presence of DMSO will also show up in its absence. There will also be some hits that cannot be observed in the absence of DMSO, even though they can be observed in its presence. And finally, there will be some that only show up in the absence of DMSO.

The most severe difficulty occurs if the protein undergoes an induced-fit motion upon substance binding. Most likely, the crystal lattice will not tolerate the protein motion and the crystals will disintegrate. In such a case, the only choice is to resort to co-crystallization of the protein and the fragments. This may, however, lead to new crystal forms. Therefore, much of the automation of the entire process will not work efficiently anymore. Luckily, in most CFS campaigns conducted

at the HZB so far, this kind of problem has not been encountered. It may be, that the weak binding of a fragment, does not provide enough energy to induce a protein motion, in particular if the crystallized conformation is stabilized by crystal packing forces. Another serious limitation of the method that the authors have encountered so far is when the crystallization cocktail (and thus the soaking solution) contains volatile compounds. Then, it becomes close to impossible to perform all the crystal handling in a meaningful way.

Different proteins may contain druggable sites to a greater or lesser extent. For example, protein-protein interactions are usually mediated by extended flat surfaces that are more difficult to target. The fragment binding hit rate will therefore likely depend on the structure of the protein's molecular surface. In an extreme case, a protein might not contain any suitable surface hot spots that serve as target sites for fragment binding. Thus, despite a meticulously performed experiment, no fragment hits will result from the screening. However, the authors have so far not encountered such a situation.

In principle, using the protocol outlined above, the crystal soaking and harvesting part of a CFS campaign can be performed in any laboratory which is equipped for crystal handling. This distinguishes the methodology at HZB from other CFS facilities and can be an advantage in some cases. For example, if the crystals cannot be easily re-produced at another site or if the travelling of the experimenters is limited, e.g., in a world-wide pandemic situation. Users at HZB are therefore provided with the entire equipment (pucks, tools, EasyAccess Frame, sample holders etc.) as a portable set.

However, the requirements for large numbers of sample holders and cryogenic storage capacities are still more conveniently met at dedicated CFS facilities. Moreover, the need for collection of many diffraction data sets strongly advocate for localizing these facilities close to beamlines, which are geared toward a high sample throughput. Examples for this are the beamlines I04 at the Diamond light source and the associated XChem facility in the UK[8,25], the MASSIF beamlines at the ESRF in France[26], or the FragMAX facility at the BioMAX beamline at MAX IV in Sweden[18].

In the future, it could be envisioned to design CFS experiments without the need for crystal handling altogether. First advances in this direction have been reported. For instance, by acoustic liquid transfer allowing the mixing of both the crystal-containing solutions and the fragment solutions directly on mesh-type sample holders[27]. Another approach was used for XFEL-based ligand-screening. In a proof-of-principle experiment, crystal slurry was prepared in batch, and soaking and diffraction data collection were performed on a silicon fixed target chip[28]. However, these approaches are still under development and far from being applicable to a wide range of protein targets or feasible for CFS facilities as a routine.

With the protocol outlined in this work, detailed instructions to successfully perform CFS campaigns straight-forwardly at HZB (and elsewhere) have been outlined and general guidance and useful hands-on tips in preparing and conducting such experiments with higher chances for success have been given. Ultimately, better odds and success rates in CFS screening largely

contribute to efficiently providing starting points for downstream development of tool compounds or drug candidates.

ACKNOWLEDGMENTS:

We thank the numerous user groups that have performed CFS campaigns at the HZB. Their feedback led to the incremental improvement of our workflow. We want to thank the drug design group at the University of Marburg and the FragMAX group at MAX IV, as the close collaborations were the foundation for several developmental leaps for improved CFS. We are thankful for the support by the German Federal Ministry of Education and Science (BMBF), via the projects Frag2Xtal and Frag4Lead (numbers 05K13M1 and 05K16M1). We are additionally grateful for support via iNEXT-Discovery, project number 871037, funded by the Horizon 2020 program of the European Commission.

DISCLOSURES:

A patent application regarding the EasyAccess Frame has been filed by Helmholtz-Zentrum Berlin with the German Patent and Trademark Office with the registration number DE 10 2018 111 478.8. Additionally, an international patent application via the PCT route, using the priority of the German patent, has been filed.

REFERENCES:

1. Erlanson, D. A., Fesik, S. W., Hubbard, R. E., Jahnke, W., Jhoti, H. Twenty years on: the impact of fragments on drug discovery. *Nature Reviews Drug Discovery*. **15** (9), 605–619 (2016).
2. Hall, R. J., Mortenson, P. N., Murray, C. W. Efficient exploration of chemical space by fragment-based screening. *Progress in Biophysics and Molecular Biology*. **116** (2–3), 82–91 (2014).
3. Erlanson, D. A. Introduction to fragment-based drug discovery. *Topics in Current Chemistry*. **317**, 1–32 (2012).
4. Scott, D. E., Coyne, A. G., Hudson, S. A., Abell, C. Fragment-based approaches in drug discovery and chemical biology. *Biochemistry*. **51** (25), 4990–5003 (2012).
5. Lamoree, B., Hubbard, R. E. Current perspectives in fragment-based lead discovery (FBLD). *Essays in Biochemistry*. **61** (5), 453–464 (2017).
6. Schiebel, J. et al. Six biophysical screening methods miss a large proportion of crystallographically discovered fragment hits: a case study. *ACS Chemical Biology*. **11** (6), 1693–1701 (2016).
7. Schiebel, J. et al. High-throughput crystallography: reliable and efficient identification of fragment hits. *Structure*. **24** (8), 1398–1409 (2016).
8. Krojer, T. et al. The XChemExplorer graphical workflow tool for routine or large-scale protein-ligand structure determination. *Acta Crystallographica Section D: Structural Biology*. **73** (3), 267–278 (2017).
9. Radeva, N. et al. Active site mapping of an sspartic protease by multiple fragment crystal structures: versatile warheads to address a catalytic dyad. *Journal of Medicinal Chemistry*. **59** (21), 9743–9759 (2016).
10. Cox, O. B. et al. A poised fragment library enables rapid synthetic expansion yielding the first reported inhibitors of PHIP(2), an atypical bromodomain. *Chemical Science*. **7** (3), 2322–2330

(2016).
11. Wollenhaupt, J. et al. F2X-Universal and F2X-Entry: Structurally Diverse Compound Libraries for Crystallographic Fragment Screening. *Structure*. **28** (6), 694–706 (2020).
12. EU OPENSCREEN fragment library. at <https://www.eu-openscreen.eu/services/compound-collection/fragment-library.html>.
13. Pearce, N. M. et al. A multi-crystal method for extracting obscured crystallographic states from conventionally uninterpretable electron density. *Nature Communications*. **8**, 15123 (2017).
14. Barthel, T. et al. Facilitated crystal handling using a simple device for evaporation reduction in microtiter plates. *Journal of Applied Crystallography*. **54**, Accepted (2020).
15. Mueller, U. et al. The macromolecular crystallography beamlines at BESSY II of the Helmholtz-Zentrum Berlin: Current status and perspectives. *The European Physical Journal Plus*. **130** (7), 141 (2015).
16. Oscarsson, M. et al. MXCuBE2: The dawn of MXCuBE collaboration. *Journal of Synchrotron Radiation*. **26** (2), 393–405 (2019).
17. Mueller, M., Wang, M., Schulze-Briese, C. Optimal fine ϕ-slicing for single-photon-counting pixel detectors. *Acta Crystallographica Section D: Biological Crystallography*. **68** (1), 42–56 (2012).
18. Lima, G. M. A. et al. FragMAX: The fragment-screening platform at the MAX IV Laboratory. *Acta Crystallographica Section D Structural Biology*. **76** (8), 771–777 (2020).
19. Sparta, K. M., Krug, M., Heinemann, U., Mueller, U., Weiss, M.S. XDSAPP2.0. *Journal of Applied Crystallography*. **49** (3), 1085–1092 (2016).
20. Winter, G. xia2: an expert system for macromolecular crystallography data reduction. *Journal of Applied Crystallography*. **43** (1), 186–190 (2010).
21. Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D: Biological Crystallography*. **67** (4), 235–242 (2011).
22. Terwilliger, T. C., Klei, H., Adams, P. D., Moriarty, N. W., Cohn, J. D. Automated ligand fitting by core-fragment fitting and extension into density. *Acta Crystallographica Section D: Biological Crystallography*. **62** (8), 915–922 (2006).
23. Pearce, N. M., Krojer, T., Von Delft, F. Proper modelling of ligand binding requires an ensemble of bound and unbound states. *Acta Crystallographica Section D Structural Biology*. **73** (3), 256–266 (2017).
24. Emsley, P., Lohkamp, B., Scott, W. G., Cowtan, K. Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography*. **66** (4), 486–501 (2010).
25. Collins, P. M. et al. Achieving a good crystal system for crystallographic X-Ray fragment screening. *Methods in Enzymology*. **610**, 251–264 (2018).
26. Bowler, M. W. et al. MASSIF-1: a beamline dedicated to the fully automatic characterization and data collection from crystals of biological macromolecules. *Journal of Synchrotron Radiation*. **22** (6), 1540–1547 (2015).
27. Cuttitta, C. M. et al. Acoustic transfer of protein crystals from agarose pedestals to micromeshes for high-throughput screening. *Acta Crystallographica Section D Biological Crystallography*. **71** (1), 94–103 (2015).
28. Moreno-Chicano, T. et al. High-throughput structures of protein-ligand complexes at room temperature using serial femtosecond crystallography. *IUCrJ*. **6** (6), 1074–1085 (2019).

Figure 1

Fragment screening workflow at HZB

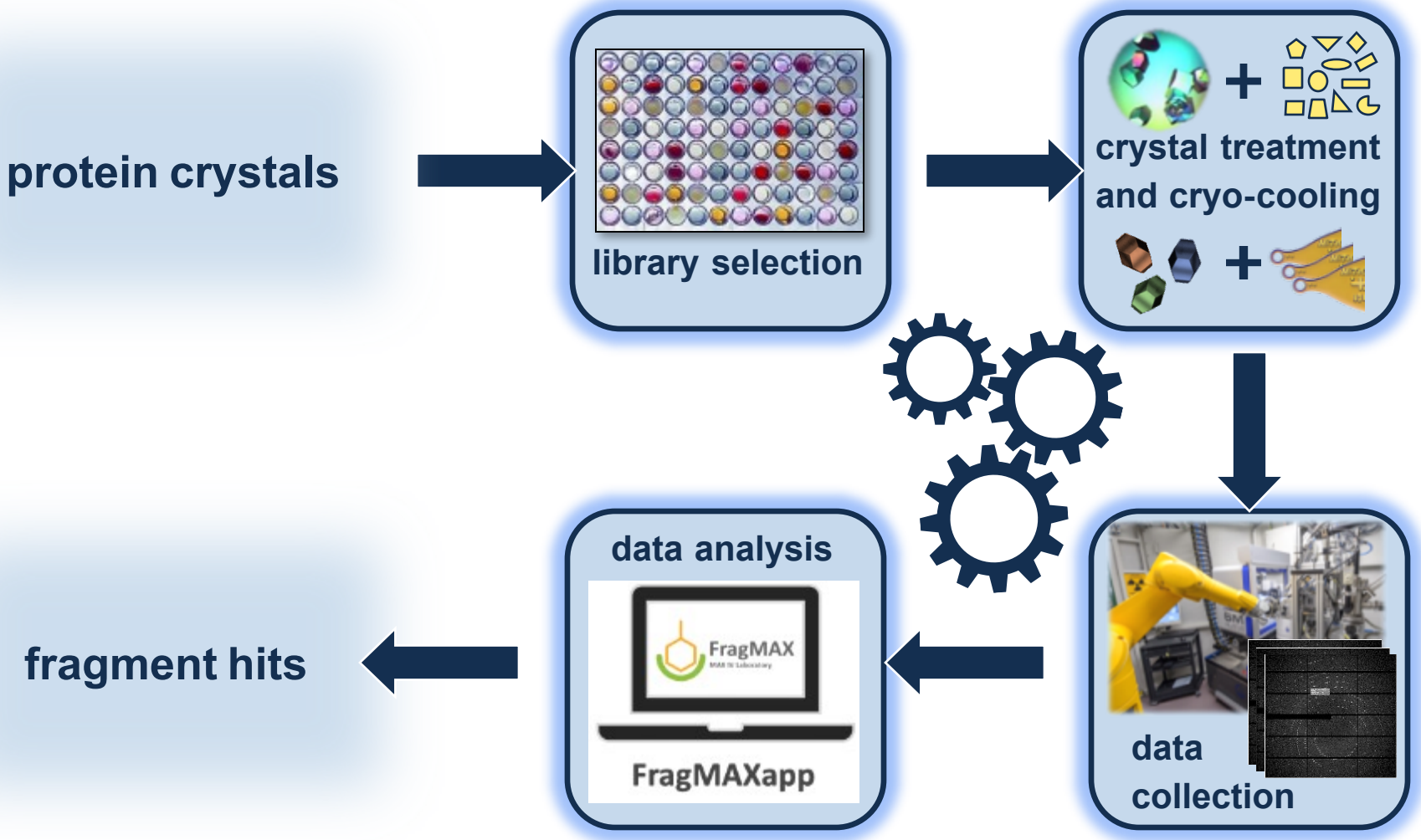
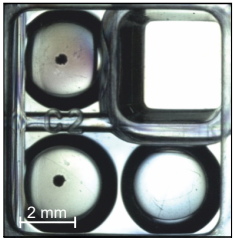
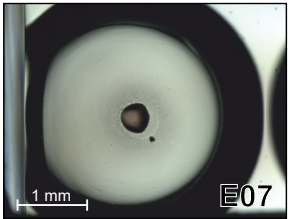


Figure 2

B

Click here to access/download;Figure;Figure2_v2_revised.





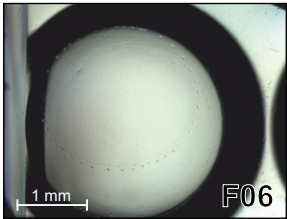


Figure 3

Click here to access/download;Figure;Figure3_v1.pdf

A

B

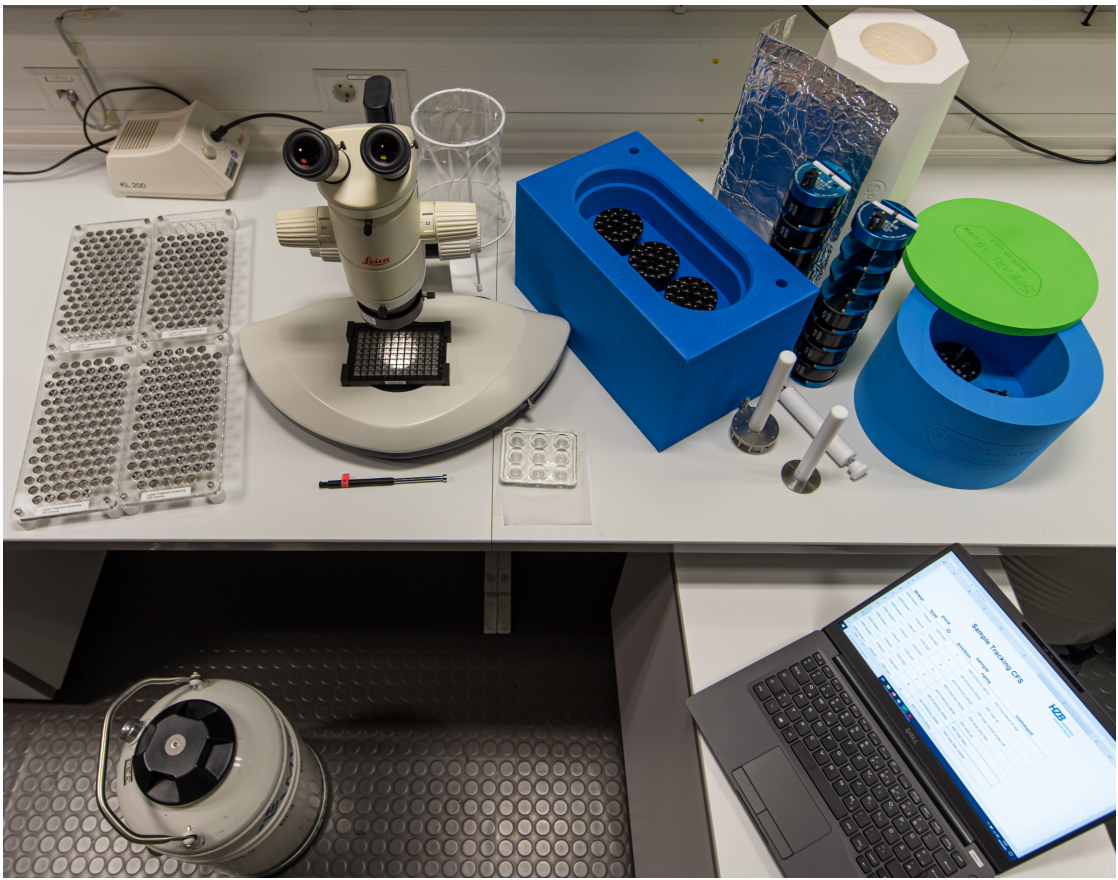
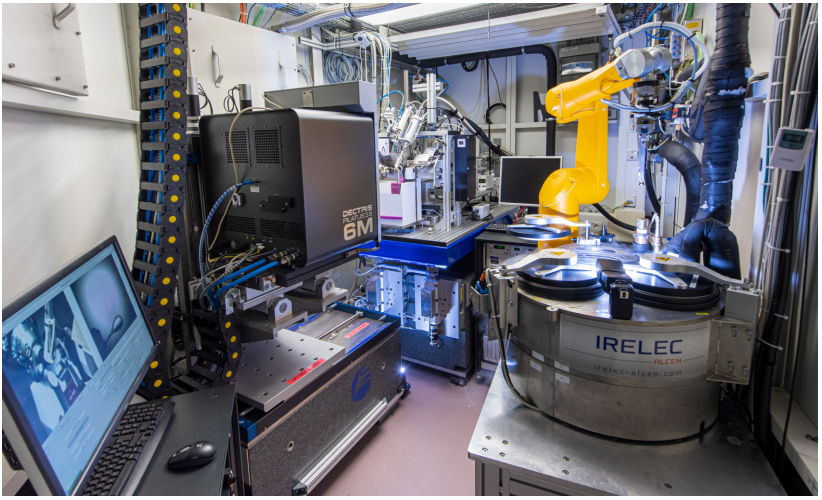


Figure 4

Click here to access/download;Figure;Figure4_v1.pdf

A



B

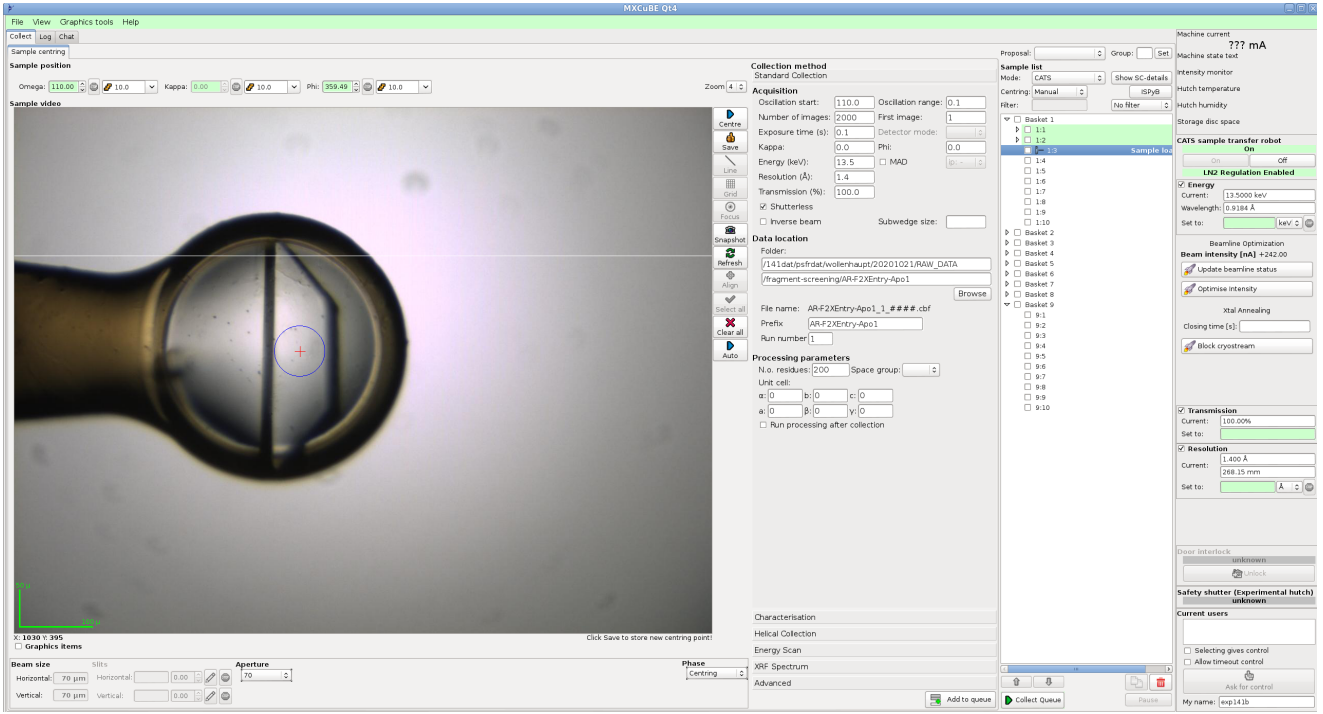
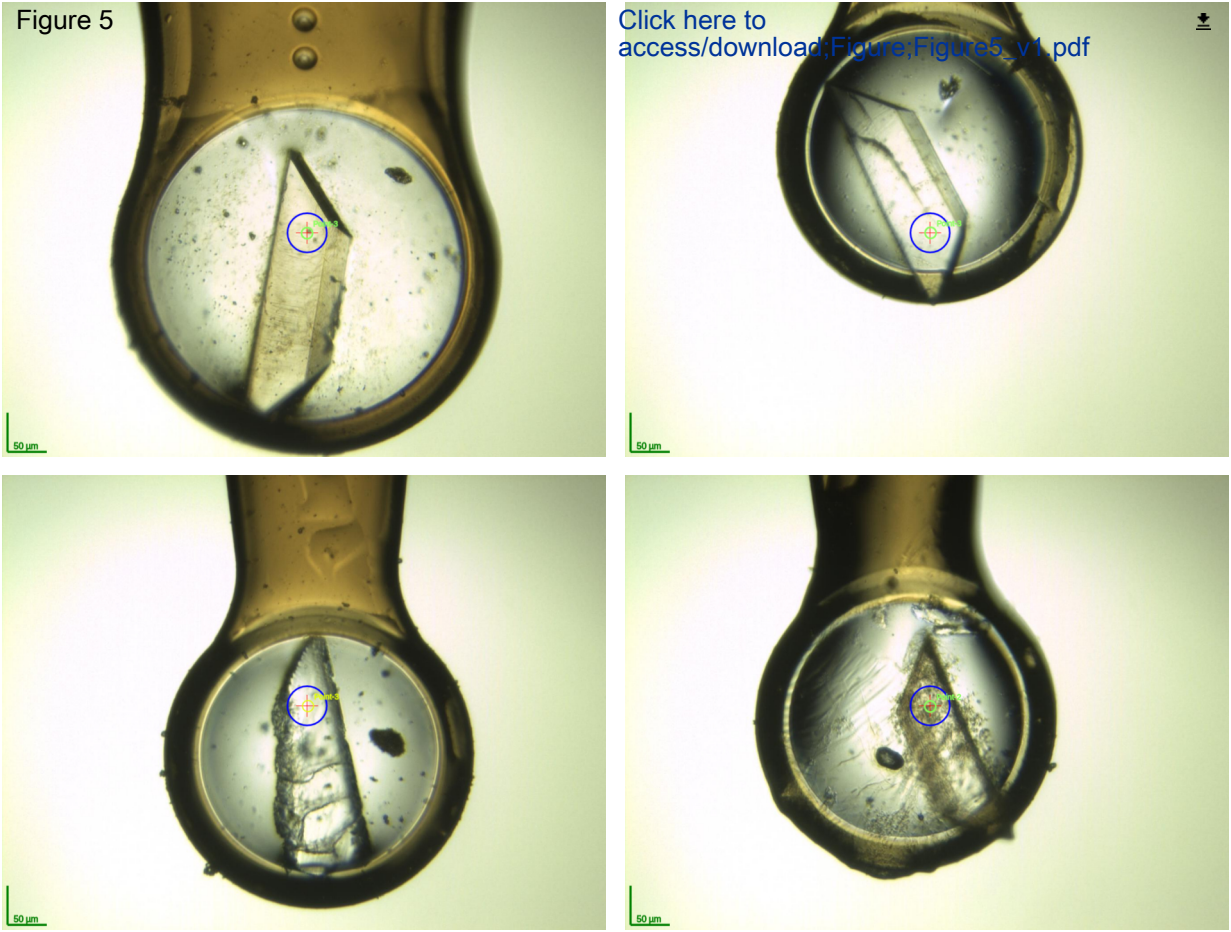
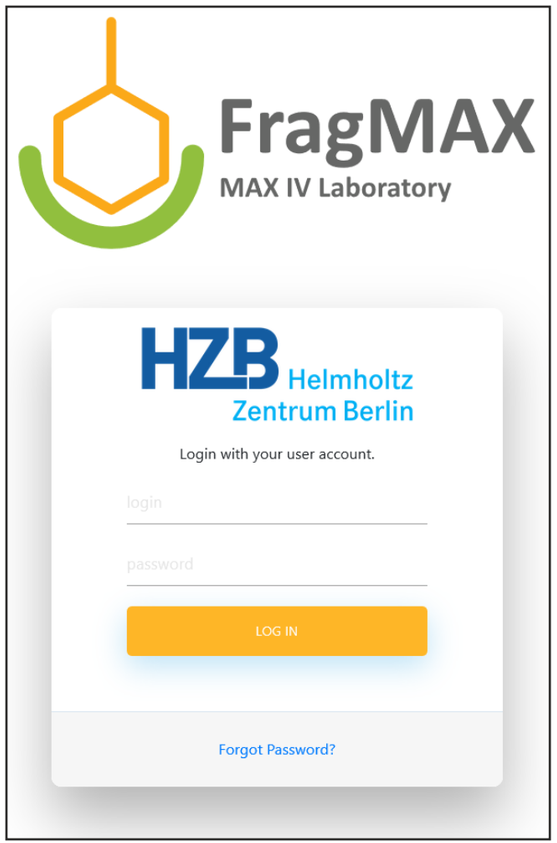


Figure 5





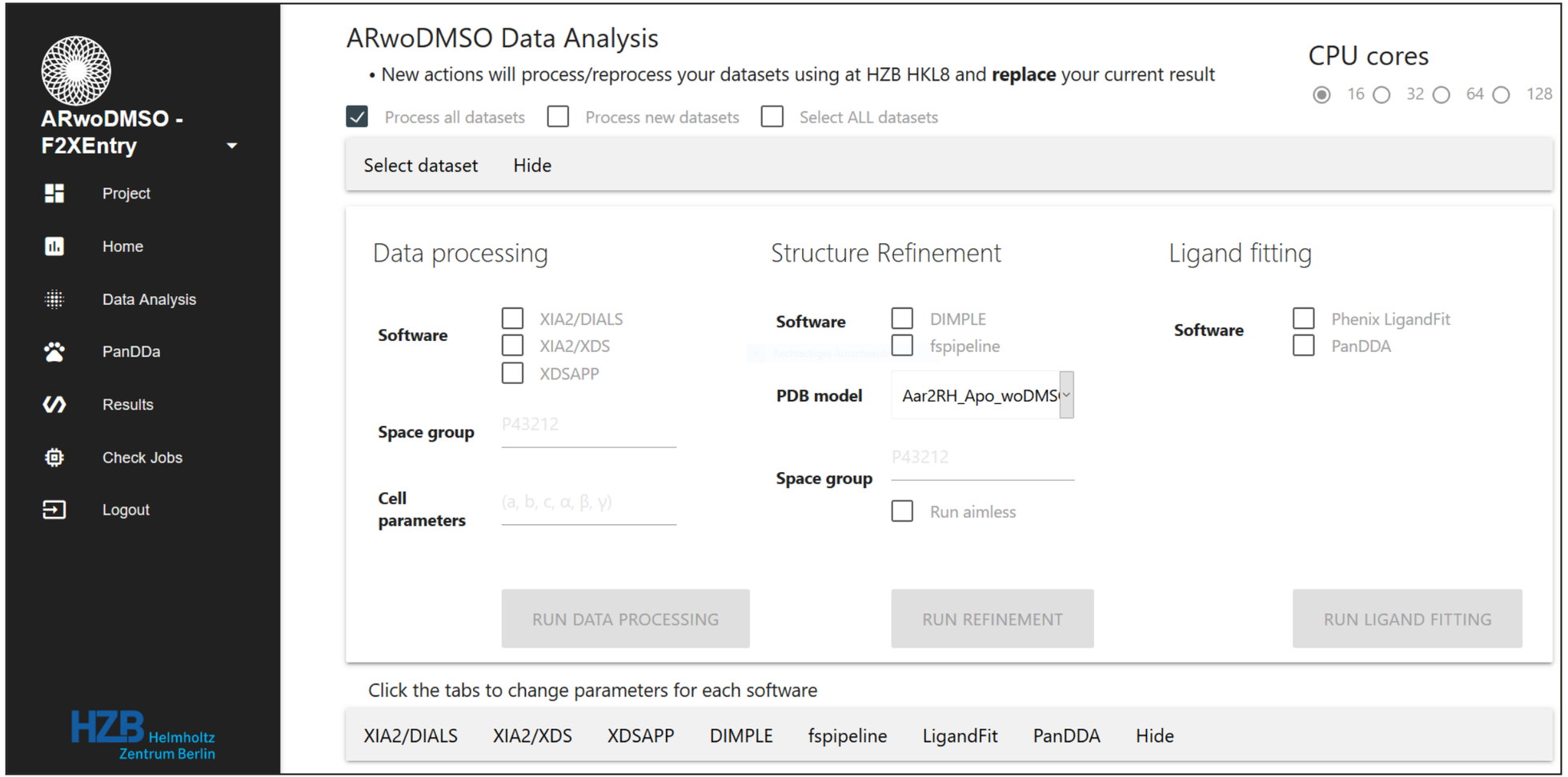
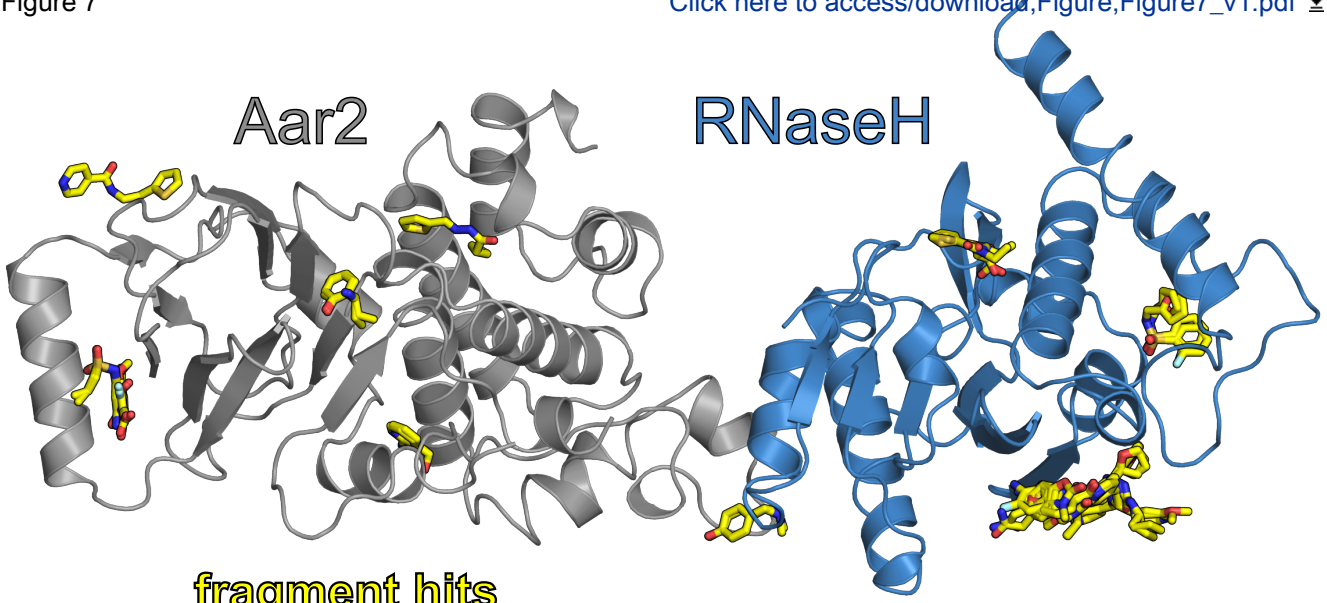


Figure 7

Click here to access/download;Figure;Figure7_v1.pdf



Name of Material/ Equipment	Company	Catalog Number
1µL pipet	Eppendorf	EP3123000012
12 channel pipet, 100µL	Eppendorf	EP4861000791
Blow dryer	TH-Geyer	9.106 788
Crystal containing crystallization plates		
Crystallization incubator		
Dual Thickness MicroLoops (LD) of different aperture sizes	MiTeGen	various, e.g. M5-L18SP-75LD
EasyAccess Frame	HZB	
F2X-Entry Screen plate	HZB	
Glas spot plate	VWR	MARI1406506
Liquid nitrogen		
Liquid nitrogen storage can	n.a.	n.a.
Magentic crystal wand	MiTeGen	M-R-1013198
Microscopes	Leica	n.a.
MRC 3-lens 96well low profile crystallization plate	SwissCI	3W96TLP-UVP
Reagent reservoir	Carl Roth	EKT6.1
Sample tracking template		
Scalpel	B. Braun	BA825SU
Sealing foil for microtiter plates	GreinerBioOne	676070
Shelved puck shipping canes (for Unipucks)	MiTeGen	M-CP-111-065
Soaking solution		
Soaking solution including cryo-protectant, 150µL		
Tissues	Roth (Kimberly Clark Professional)	AA64.1
Transport dewar (Whartington dry shipper)	MiTeGen	TW-CX100

Unipuck foam dewars with lid	MiTeGen	M-CP-111-022
Unipuck starter set	MiTeGen	M-CP-UPSK001
Unipucks	MiTeGen	M-CP-111-021

Comments/Description

Contains crystals to be soaked
Providing constant temperature for crystallization experiment, at HZB: 20°C

250 loops in the appropiate size needed for the protocol, can be provided by HZB
The EasyAccess Frame is a special device for handling multiple crystals, which was developed at the HZB (Barthel et al., 2021).
Developed F2X-Entry Screen (Wollenhaupt et al., 2020)

At least a filled up 5 L can

For mock-soaked crystals (optional)
25 ml volume
Provided from HZB, see supplementary infomation

2 canes made of aluminum; can be provided by the HZB
At least 5 ml are needed
Only needed if soaking solution is not cryo-protectant already

2 Travel dewars for storage of the 2 unipuck canes, alternatively a storage dewar of type VHC35 or similar could be used.

two foam dewars especially suited for unipuck handling described in the protocol
if SPINE pucks are used, different foam dewars might have to be applied.

Can be provided by the HZB
14 unipucks; can be provided by the HZB

HZB Helmholtz Zentrum Berlin

Helmholtz-Zentrum Berlin für Materialien und Energie GmbH | Hahn-Meitner-Platz 1 | 14109 Berlin

To the
Editors of
JoVE

Dr. Manfred S. Weiss
Macromolecular Crystallography (HZB-MX)
Helmholtz-Zentrum Berlin
Albert-Einstein-Straße 15
D-12489 Berlin, Germany
Tel +49-30-8062-13149
Fax +49-30-8062-14975

msweiss@helmholtz-berlin.de

Berlin, Jan 22, 2021

Ref.: Revised Manuscript Submission

Dear Editors,

Thank you very much for your comments on our manuscript on the workflow for crystallographic fragment-screening at the HZB. In the following we would like to answer all points raised by the referees and by you. We hope that you find our comments and the concomitant changes we applied to our text and the figures satisfactory and that our manuscript is now suitable for publication in JoVE.

Editorial comments:

Changes to be made by the Author(s):

1. Please take this opportunity to thoroughly proofread the manuscript to ensure that there are no spelling or grammar issues. Please define all abbreviations at first use.

We acknowledge the editor's comment and have thoroughly revised the entire manuscript in this respect.

2. Please revise your title to "Workflow and Tools for Crystallographic Fragment Screening".

We are a bit surprised by your request to change the title of our manuscript, but are willing to comply with it to some extent. we would propose a compromised title version such as: "**Workflow and Tools for Crystallographic Fragment Screening at the Helmholtz-Zentrum Berlin**". The addition of "at the Helmholtz-Zentrum Berlin" is absolutely crucial as this protocol (as any suchlike protocol) is in parts very specific to the experimental site in Berlin. This is because these kinds of experiments require the use of large-scale facilities, with adjusted protocols for each of them (hence see the contribution by the DIAMOND light source and the ESRF to this methods collection). We definitely want to avoid giving the impression that our manual "as is" is applicable to every other laboratory/facility, also out of respect for our colleagues at other CFS facilities. The protocol outlined here concerns the crystallographic fragment screening at the HZB only. Several steps of the protocol (by the nature of large-scale facilities) are bound to be carried out at the HZB. Nonetheless, the reader not intending to use our facility may still extract the transferrable parts of the protocol and use the know-how in his/her own laboratory. This is even an advantage of our protocol, e.g. steps 1 and 2 are indeed transferrable to any laboratory. However, the other steps are site-specific, which needs to be reflected in the title of the manuscript.

3. Please provide an email address for each author.

Jan Wollenhaupt jan.wollenhaupt@helmholtz-berlin.de
Tatjana Barthel tatjana.barthel@helmholtz-berlin.de
Gustavo M.A. Lima gustavo.alvares.lima@gmail.com
Alexander Metz alexander.metz@idorsia.com
Dirk Wallacher dirk.wallacher@helmholtz-berlin.de
Elmir Jagudin elmir.jagudin@maxiv.lu.se
Franziska U. Huschmann franziska.huschmann@gmx.de
Thomas Hauß hauss@helmholtz-berlin.de
Christian G. Feiler christian.feiler@helmholtz-berlin.de
Martin Gerlach gerlach-reise@gmx.de
Michael Hellmig michael.hellmig@helmholtz-berlin.de
Ronald Förster foerster@helmholtz-berlin.de
Michael Steffien michael.steffien@helmholtz-berlin.de
Andreas Heine heinea@mailer.uni-marburg.de
Gerhard Klebe klebe@staff.uni-marburg.de
Uwe Mueller uwe.mueller@helmholtz-berlin.de
Manfred S. Weiss manfred.weiss@helmholtz-berlin.de

4. JoVE cannot publish manuscripts containing commercial language. This includes trademark symbols (™), registered symbols (®), and company names before an instrument or reagent. Please remove all commercial language from your manuscript and use generic terms instead. All commercial products should be sufficiently referenced in the Table of Materials and Reagents.

For example: Kapton loops (MiTeGen Dual Thickness 191 MicroLoops (LD)); Hampton Research; Falcon tube; Unipuck foam dewars; beamline 14.1 or SPINE pucks to beamline 14.2 dewar, DIAMOND light source etc

We understand this point very well and revised the manuscript accordingly. However, we need to state that some of these names are used in the scientific community as such. There is simply no other description for the device. The comprises Beamline 14.1, SPINE pucks, Unipucks, DIAMOND light source. There are no replacements for this, and we are sure that you do not want us to replace DIAMOND Light Source for instance by “the synchrotron located in the vicinity of Cambridge”. Also, not detailing equipment parts by using the terms commonly used in the field can create utter confusion. For example, if we would call the unipucks only “pucks” this would be too general as the different types of pucks are fundamentally different in their specifications.

5. Please revise the text, especially in the protocol, to avoid the use of any personal pronouns (e.g., "we", "you", "our" etc.).

We acknowledge the editor’s comment and have thoroughly revised the manuscript in this aspect.

6. Please ensure that all text in the protocol section is written in the imperative tense as if telling someone how to do the technique (e.g., “Do this,” “Ensure that,” etc.). The actions should be described in the imperative tense in complete sentences wherever possible. Avoid usage of phrases such as “could be,” “should be,” and “would be” throughout the Protocol. Any text that cannot be written in the imperative tense may be added as a “Note.” However, notes should be concise and used sparingly. Please include all safety procedures and use of hoods, etc.

We recognize the editor's comment and changed the manuscript accordingly.

7. The Protocol should contain only action items that direct the reader to do something. Please move section 1 in the protocol to the introduction and sections 2 and 3 to the Discussion.

We acknowledge the removal of explanatory part from the protocol but found that the mentioned sections were all better suited in the introduction as these points need to be taken care of prior to execution of the protocol. We have adjusted the introduction accordingly.

8. Please note that your protocol will be used to generate the script for the video and must contain everything that you would like shown in the video. Please add more details to your protocol steps.

We acknowledge the editor's comment and added more details where it seemed necessary for understanding.

Please ensure you answer the "how" question, i.e., how is the step performed? Alternatively, add references to published material specifying how to perform the protocol action. Please add more specific details (e.g., button clicks for software actions, numerical values for settings, etc) to your protocol steps. There should be enough detail in each step to supplement the actions seen in the video so that viewers can easily replicate the protocol.

9. Wherever possible, list materials, reagents, software, equipment in the Table of Materials and cite the Table of Materials in the text. Hence, remove all lists of materials to be used from the protocol.

We acknowledge the editor's comment and removed the list of materials from the protocol part. The Table of Materials has been updated to include all materials there and cited appropriately at the points of the protocol where workplaces are arranged (i.e. 1.2 and 2.2)

10. Please include a one line space between each protocol step and then highlight up to 3 pages of protocol text for inclusion in the protocol section of the video.

We thank the editor for the reminder. We included the one-line spaces now for all protocol steps where they were missing and have highlighted the protocol parts to be filmed in yellow.

11. Please include a scale bar for all images taken with a microscope to provide context to the magnification used. Define the scale in the appropriate Figure Legend.

We are thankful for the reminder to provide a scale bar for microscope images. We added the bar scale in figure 2.

Reviewers' comments:

Reviewer #1:

Manuscript Summary:

This manuscript describes how to perform a Fragment screening campaign at Bessy (Helmholtz-Zentrum Berlin). Fragment screening is a powerful method to facilitate structure-based drug design and drug discovery. Here the authors clearly describe how to streamline the process. This significantly lowers the threshold for new users of this method and also helps and guides experienced users. The method is very clearly described and everything is presented with the clarity and rigor that this group is known for.

We are very happy about the overall enormously positive feedback by Reviewer #1.

Minor Concerns:
The authors should mention how the streamlined method also can be used to analyze prescreened fragment libraries. They could also mention advantaged and disadvantages of prescreening libraries using different, orthogonal methods.

We are grateful for the suggestion to include the topic of prescreening fragment libraries and orthogonality of different biophysical methods. We are aware that in the field of fragment screening several different methods can be applied. It has shown by us and others that screening the same fragment library with different methods could result in very low overlap of hits between the different methods applied. Nevertheless, we have included two more sentences about this in the introduction. There it reads now:

"Several biophysical methods may be applied for fragment screening, the most popular being nuclear magnetic resonance, X-ray crystallography, surface plasmon resonance and thermal shift assays. These methods are used either in an parallel or in a sequential way, with the aim to increase the confidence in the hits and reduce the numbers of false positives or false negatives, respectively. However, a recently conducted comparative study[6] suggested that sequential screening cascades are to be avoided due to the low overlap between the different methods."

It is clear that the topic of pre-screening is so large, that an entire article could be written on it. However, this would be widely out of scope for this manuscript, which is not a review article, but a practical protocol about **crystallographic** fragment screening as a first-step screening technique as performed at the HZB, with the Protocol/Video as the central part. And at the HZB, we have – for reasons detailed above – refrained from including pre-screening in our workflow.

What does the sporadic yellow highlighting of text mean?

We thank the reviewer for the question about the yellow highlighting. It is mandatory for a JoVE produced video manuscript to highlight protocol text in yellow that will be filmed later on, as described in the author guidelines/ the template. The video is actually the main and central part of the publication. We apologize that this was not communicated along with the manuscript.

Reviewer #2:
Manuscript Summary:
Very well-described and thorough protocol for fragment screening by X-ray crystallography.
Major Concerns:
The recommended soaking drop size of 0.4 microliters is too small. For the inexperienced hand it will lead to evaporation of the liquid before the crystal can be harvested, which may lead to crystal damage. Please address why such a small drop size has been selected. A usual protocol for such a screening would involve soaking in a drop at least 1 microliter in volume.

We recognize the concerns of the reviewer about the small drop size of 0.4 µl. However, in respect to screening campaigns in macromolecular crystallography, such small drop sizes are very common and also well established. In our publication concerning the validation campaigns of our library (Wollenhaupt et al., 2020, doi.org/10.1016/j.str.2020.04.019) and all conducted user campaigns, we always use 0.4 µl and it works very well, in our experience, even for users that are fairly new in the field of macromolecular crystallography. Additionally, at other synchrotron sites that perform such screening campaigns on a weekly basis similar drop sizes are used, for example at XChem (DIAMOND) 0.2 – 0.6 µl drops. Also, a recent collaboration headed by the group at DESY screening several thousand crystals of the SARS-CoV-2 protease using drop sizes of 0.5 µl (preprint: Günther et al., 2020, doi.org/10.1101/2020.11.12.378422). Many more examples could be given. Thus, it is very convincing that these drop sizes are well established in the crystallographic community and no changes or further elaboration on that is necessary in the manuscript.

Reviewer #3:

Manuscript Summary:

Wollenhaupt et al. report on the procedures of crystallographic fragment screening at the BESSY synchrotron. The article is generally very well written, sound and concise. The article is of high relevance for the crystallographic community and the JoVE journal is a very well suited medium for this information. I recommend publication after minor corrections, as detailed below.

General comments:

Some text is marked in yellow for reasons not obvious to the reviewer. I ignored this.

The journal seems to provide videos for the procedures. These were not accessible through the review system.

We thank the reviewer for noticing the yellow highlighting and lack of a video. JoVE is a video-based journal. The video is actually the main and central part of the publication. It will be based on the protocol part and will only be produced after the manuscript is peer-reviewed. It is mandatory for a JoVE produced video manuscript to highlight protocol text in yellow that will be filmed later on, as described in the author guidelines/ the template. We apologize that this was not communicated along with the manuscript.

Minor concerns:

Line 144: " The number of hits obtained in a CFS campaign is not only dependent on the druggability of the target protein and the suitability of the crystal lattice (see above) but it also on the choice of the library, its quality and its presentation."

Please specifiy in more detail, which factors are meant when referring to "choice", "quality" and "presentation" of the library.

We thank the reviewer for the suggestion to elaborate more on this particular topic. Nevertheless, we have to recognize that this manuscript is not an exhaustive review, but a practical guide to performing CFS experiments with the protocol (and the resulting, to be filmed video) as a central part. We have tried to incorporate the suggestion of the reviewer mildly in the introductory part of the manuscript.

The changed paragraphs read now:

“The number of hits obtained in a CFS campaign is not only dependent on the druggability of the target protein and the suitability of the crystal lattice (see above) but it is also dependent on the quality of the library. Library quality comprises two aspects: the selection of the compounds for the library and the confectioning of the compounds, i.e., in which physical form they are presented for the experiment. For compound selection different strategies can be employed. Most library designs include the maximization of the chemical diversity of the fragments. A strategic focus could be to include the chemical tractability of the fragments for follow-up design, which has been applied for instance in the DSI poised library [10]. Yet another strategic focus for library design could be to maximize the representation of commercially available chemical space of fragments by shape- and pharmacophore-based clustering, as has been exemplified by the F2X libraries developed at HZB [11].

[… and about the “representation” further below:]

“With respect to the physical presentation of a library, two approaches are most commonly adopted: the fragments are either used as DMSO stock solutions or the fragments are dried and immobilized on ready-to-use plates.”

Line 172: "We recommend preparing additional mock-soaks that will later...". Please define "mock-socks" at this point (e.g. "data sets without a soaked ligand"). It becomes clear later, but may be unclear at this point for many readers.

We appreciate the suggestion of the reviewer to define “mock-soaks” in more detail at an earlier stage. We therefore added the following paragraph as second-to-last paragraph in the introduction:

“Mock-soaks are defined as soaking experiments on protein crystals using the same soaking solution as the fragment soaks for the same incubation time, but no fragments present. If the soaking solution is equal to the crystallization condition, the crystals may be directly harvested from the crystallization plate.”

Line 216: " Add milli-Q or distilled water..." -> Add deionized water ... (milli-Q is a device from Millipore)

We are grateful for the reviewer pointing out the use of the word milli-Q. We replaced the term as suggested by the referee at every position where “milli-Q” was used in the former version of the manuscript.

Line 351:
" Nevertheless, samples should be prioritized:
a) First collect one diffraction dataset of each fragment soak, prioritizing the better of the two duplicates
b) Collect duplicate diffraction datasets in case of failed data collection, loss of diffraction or severe ice rings
c) Collect diffraction datasets of apo crystals.
d) Collect the remaining duplicates."
Even after repeated reading of this section, the procedure is not entirely clear to the reviewer. Is this the suggested procedure?:
a) First, collect a data set for one crystal of each fragment, based on the prioritization in step 6.6.
b) Collect a second diffraction dataset for those fragment data sets, that suffer from failed data collection, loss of diffraction or severe ice rings in the data collection of the first crystal.

c) Collect diffraction datasets of apo crystals.
d) Collect data sets of the remaining duplicates of each fragment.

We thank the reviewer for this very valuable comment, which helped us to improve the manuscript. We adopted the suggestions of the reviewer with slight modifications so that it reads now:

"Nevertheless, samples should be prioritized:

a) First, collect diffraction data sets for one sample per fragment condition, based on the prioritization in step 2.6. / 2.7, i.e. collect the data for the higher prioritized duplicate.

b) For those experiments in a) that suffered from failed data collection, loss of diffraction or severe ice rings, collect data for the second duplicate sample for the respective fragment condition.

c) Collect diffraction datasets of apo crystals (if prepared according to steps 1.24 and 2.12).

d) Collect diffraction datasets of the remaining duplicates of each fragment condition."

Line 375: Is this procedure of excluding fragment-bound data sets from the ground state models necessary if apo data sets have been collected for the purpose of determining the ground state for PanDDA analysis?

We are thankful for the question of the reviewer and would like to explain the procedure in more detail. The apo data sets **can** be used exclusively for the characterization of the ground state. In such a case, a second round of PanDDA excluding fragment-bound datasets would not lead to new information because the datasets used for the ground state remain the same. However, due to differences between crystals, even apo crystals, it is not recommended to solely depend on apo data sets for the ground state characterization. It is best to leave the decision of picking datasets used for the ground state characterization to the PanDDA algorithm, as it finds the datasets that are very similar. In this case rerunning PanDDA by excluding fragment-bound datasets results in an improved ground state model and can extract additional hits. As the reviewer pointed out with the question, it is not clearly stated in the manuscript how to use the apo data sets. Therefore, we reformulated the entire paragraph, explaining what a ground state model is and also make the suggested procedure more comprehensible. The paragraph reads now:

"A key element for hit identification is PanDDA. Details are explained in the respective publications [13, 23], in brief: PanDDA automatically calculates electron density maps of a set of data sets in a CFS campaign. These are then assumed as non-binding fragment conditions and averaged to generate the so-called ground state model. The ground state model is then used to derive local discrepancies between each electron density map and the ground state map, using voxel-associated Z-scores. Then, for areas of high Z-scores a so called PanDDA-map is created by fine-tuned subtraction of ground state density from the respective map. This largely enhances the visibility of fragment binding events.

To maximize the outcome of PanDDA, the authors recommend a two-step approach. Firstly, performing a PanDDA run (pandda.analyze) with standard settings. Even if mock-soaked crystals have been collected, their identity will not be included as a parameter (which is possible nonetheless) in order to enable an unbiased generation of the ground state model by PanDDA from all available data. After that the output data is evaluated by the user via a so-called PanDDA

inspection in Coot [24]. Here, hits with relatively high confidence should be noted, concluding the first step.

Secondly, the authors recommend re-running the pandda.analyse step excluding the preliminary hits (determined in the first step) from the ground state model via the

--exclude_from_charactersation="<list-of-bound-dataset-ids>"

command line option. Further details are described on the PanDDA help pages (https://pandda.bitbucket.io/). This way, datasets that are clear hits and thus would obscure the ground state model if included there, are disregarded, leading to an improved ground state model and thus to improved results overall. Finally, a thorough PanDDA inspection is performed to complete the hit identification."

Line 421: "The photographs were taken on the BioMAX beamline for AR samples collected there as part of the F2X Entry validation10." Include the information "BioMAX beamline (MAX IV synchrotron, Lund, Sweden)".

We thank the reviewer for pointing out the missing information and changed the manuscript as suggested.

Line 472: "It may be, that the weak binding of a fragment*,* does not provide enough energy to induce a protein motion." Suggestion, add: "...in particular if the crystallized conformation is stabilized by crystal packing forces." The protein conformation is also stabilized by crystal packing forces, which the fragment also needs to compensate. Omit the comma *,*.

We would like to thank the reviewer for the proposed addition to our statement. It is an important point to include the crystal packing forces and therefore we changed the text as proposed.

Line 478: "A rather theoretical limitation seems to be when a protein is not made to interact with anything and does not contain any surface..." I think, this is not so theoretical. It is very likely that the hit rate will depend on the type of interaction sites. Suggestion: "Different proteins may contain more or less druggable sites. Protein-protein interactions are usually mediated by extended flat surfaces, that are more difficult to target. The fragment binding hit rate will therefore likely depend on the stucture of the protein molecular surface.

We are grateful for the reviewer's suggestion to rephrase this part of the manuscript. We changed the respective paragraph in the discussion, largely adopting the proposed statement of the reviewer. It reads now:

"Different proteins may contain druggable sites to a greater or lesser extent. For example, protein-protein interactions are usually mediated by extended flat surfaces that are more difficult to target. The fragment binding hit rate will therefore likely depend on the structure of the protein's molecular surface. In an extreme case, a protein might not contain any suitable surface hot spots that serve as target sites for fragment binding. Thus, despite a meticulously performed experiment, no fragment hits will result from the screening. However, the authors have so far not encountered such a situation."

When specifying hit rates, the authors should also mention that a *large portion* of the hits can also occur at crystal contacts or sites that are unlikely to influence protein function (enzyme inhibition,

receptor modulation). Therefore the number of hits that are useful for drug design is significantly lower than the overall hit rate.

We appreciate the reviewer bringing up this topic. However, our experience with in-house and user campaigns does not suggest a general “large portion” of hits at crystal contacts or on site presumably “unlikely to influence the protein function”. However, we do acknowledge that there is a minor part of such hits. To this end we included the following paragraph in the discussion, after the paragraph specifying expected hit rates. The new paragraph reads:

“A general caveat of crystallographic fragment screening is the presence of crystallographic contact sites. These could either occlude a priori known active sites (to be checked before the screening, see above), or, these contact sites also often provide pockets and hot spots where fragments can bind. Such fragment hits will be artifacts of the crystallization lattice and will likely not bind to the protein in solution. However, this is for most projects usually the minor portion of the hits obtained. For example, in the F2X-Entry Screen validation campaign using endothiapepsin (EP) and the spliceosomal protein-protein complex of Prp8RnaseH and Aar2 (AR), most of the hits occurred in promising sites 11. For EP, 27 out of the 37 observed binding events were located in the active site, i.e. the peptide cleft of this protease. The 10 remote binding events comprise two solvent exposed binding events and eight crystal contact binding events (corresponding to five unique hits). Excluding those crystal contact hits would still reflect an overall rate of 24% unique hits for the EP campaign. It is also important to notice that binding events remote of a known active site (except crystal contact binders) could also potentially be interesting, e.g. revealing new hot spots or allosteric sites of the protein. For the AR campaign (in the same publication), of the 23 observed binding events, seven were located at crystal contacts, one was located at the direct interface of the two proteins, seven were located at known protein-protein interactions sites with other binding partners of the larger biological context (hence different assembly stages of the spliceosome), eight binding events revealed two hot spots on AR of yet unknown function and one being at a solvent exposed surface of Prp8RnaseH. Therefore, excluding the events at crystal contacts and the Prp8RnaseH singleton, the number of potentially useful binding events is 15 (corresponding to 14 unique hits) thus a hit rate of 15.6%. These hits can be starting points for design of protein-protein interaction modulators or for tool compounds aimed to explore the two discovered Aar2 hot spots. Taken together, also in line with conducted user campaigns, often only a minor portion of hits in crystallographic fragment screening has to be disregarded as artefacts. However, this will also be largely target dependent”

With best regards,

Manfred Weiss

HZB Helmholtz Zentrum Berlin

Helmholtz-Zentrum Berlin für Materialien und Energie GmbH | Hahn-Meitner-Platz 1 | 14109 Berlin

Supplemental Excel Table

Click here to access/download;Supplemental File (Figures, Permissions, etc.);TemplateCFSHZBSampleTracking.xlsx

Sample Tracking CFS

dewar	puck type	puck ID	sample position	sample name	comment
HZB-Dewar-1	Unipuck	HZB-0001	1	AR-F2XEntry-B01a	400 loop ok
HZB-Dewar-1	Unipuck	HZB-0001	2	AR-F2XEntry-B01b	200 loop piece of xtal
HZB-Dewar-1	Unipuck	HZB-0001	3	AR-F2XEntry-B02a	400 loop ok
HZB-Dewar-1	Unipuck	HZB-0001	4	AR-F2XEntry-B02b	200 loop corner broke off
HZB-Dewar-1	Unipuck	HZB-0001	5	AR-F2XEntry-B03a	500 loop jelly-like
HZB-Dewar-1	Unipuck	HZB-0001	6	AR-F2XEntry-B03b	500 loop jelly-like
HZB-Dewar-1	Unipuck	HZB-0001	7	AR-F2XEntry-B04a	400 loop hairy
HZB-Dewar-1	Unipuck	HZB-0001	8	AR-F2XEntry-B04b	500 loop fine cracks
HZB-Dewar-1	Unipuck	HZB-0001	9	AR-F2XEntry-B05a	400 loop one crack
HZB-Dewar-1	Unipuck	HZB-0001	10	AR-F2XEntry-B05b	500 loop with some fragment powder
HZB-Dewar-1	Unipuck	HZB-0001	11	AR-F2XEntry-B06a	400 loop one crack
HZB-Dewar-1	Unipuck	HZB-0001	12	AR-F2XEntry-B06b	200 loop piece of xtal
HZB-Dewar-1	Unipuck	HZB-0001	13	AR-F2XEntry-B07a	500 loop piece of xtal with cracks
HZB-Dewar-1	Unipuck	HZB-0001	14	AR-F2XEntry-B07b	400 loop piece of xtal
HZB-Dewar-1	Unipuck	HZB-0001	15	AR-F2XEntry-Apo01	500 loop ok
HZB-Dewar-1	Unipuck	HZB-0001	16	AR-F2XEntry-Apo02	400 loop ok
HZB-Dewar-2	SPINE	mx-004	1	AR-F2XEntry-B01a	400 loop ok
HZB-Dewar-2	SPINE	mx-004	2	AR-F2XEntry-B01b	200 loop piece of xtal
HZB-Dewar-2	SPINE	mx-004	3	AR-F2XEntry-B02a	400 loop ok
HZB-Dewar-2	SPINE	mx-004	4	AR-F2XEntry-B02b	200 loop corner broke off
HZB-Dewar-2	SPINE	mx-004	5	AR-F2XEntry-B03a	500 loop jelly-like
HZB-Dewar-2	SPINE	mx-004	6	AR-F2XEntry-B03b	500 loop jelly-like
HZB-Dewar-2	SPINE	mx-004	7	AR-F2XEntry-B04a	400 loop hairy
HZB-Dewar-2	SPINE	mx-004	8	AR-F2XEntry-B04b	500 loop fine cracks
HZB-Dewar-2	SPINE	mx-004	9	AR-F2XEntry-Apo01	500 loop ok
HZB-Dewar-2	SPINE	mx-004	10	AR-F2XEntry-Apo02	400 loop ok

HZB Helmholtz Zentrum Berlin

Sample Tracking CFS

	puck		sample		
dewar	type	ID	position	name	comment

HZB Helmholtz Zentrum Berlin

Sample Tracking CFS

	puck		sample		
dewar	type	ID	position	name	comment

HZB Helmholtz Zentrum Berlin

Sample Tracking CFS

	puck		sample		
dewar	type	ID	position	name	comment

HZB Helmholtz Zentrum Berlin

Sample Tracking CFS

	puck		sample		
dewar	type	ID	position	name	comment

HZB Helmholtz Zentrum Berlin

Sample Tracking CFS

	puck		sample		
dewar	type	ID	position	name	comment

Sample Tracking CFS

	puck		sample		
dewar	type	ID	position	name	comment

HZB Helmholtz Zentrum Berlin

Sample Tracking CFS

	puck		sample		
dewar	type	ID	position	name	comment

Sample Tracking CFS

dewar	puck		sample		comment
	type	ID	position	name	