


PERSPECTIVE OPEN ACCESS

The HZB F2X-Facility—An Efficient Crystallographic Fragment Screening Platform

Tatjana Barthel¹  | Laila Benz^{1,2} | Yara Basler^{1,2} | Thomas Crosskey¹ | Alexander Dillmann¹ | Ronald Förster^{1,2} | Paula Fröling^{1,2} | Camilla G. Dieguez¹ | Christine Gless¹ | Thomas Hauß¹ | Michael Hellmig¹ | Lea Jänisch^{1,2} | David James¹ | Frank Lennartz^{1,3} | Jelena Mijatovic^{1,2} | Melanie Oelker^{1,3} | James W. Scanlan^{1,3} | Gert Weber¹ | Jan Wollenhaupt¹ | Uwe Mueller¹ | Holger Dobbek³ | Markus C. Wahl^{1,2} | Manfred S. Weiss¹

¹Macromolecular Crystallography, Helmholtz-Zentrum Berlin für Materialien und Energie, Berlin, Germany | ²Structural Biochemistry Group, Institute for Chemistry and Biochemistry, Freie Universität Berlin, Berlin, Germany | ³Strukturbiologie/Biochemie, Institut für Biologie, Humboldt-Universität zu Berlin, Berlin, Germany

Correspondence: Manfred S. Weiss (manfred.weiss@helmholtz-berlin.de)

Received: 14 June 2024 | **Revised:** 6 August 2024 | **Accepted:** 19 August 2024

Funding: These developments were supported by the German Ministry of Science and Education (BMBF) via the projects Frag2Xtal (No. 05K13RM1) and Frag4Lead (No. 05K16RM1), as well as by iNEXT-Discovery, project No. 871037, funded by the Horizon 2020 program of the European Commission, the German Research Foundation (DFG) via the project NECESSITY (FE2166/1-1) and the German Academic Exchange Service (DAAD) via the project CFS-Pipeline (57654470).

Keywords: crystallographic fragment screening | EasyAccess Frame | F2X compound libraries | F2X facility | fragment-based lead discovery | structure-based drug design | synchrotron beamlines

ABSTRACT

Crystallographic fragment screening (CFS) has recently matured into an important method for the early stages of drug discovery projects. It is based on high-throughput structure determination and thus requires a high degree of automation as well as specialized workflows and robust analysis tools. Consequently, large-scale research facilities such as synchrotrons have embraced the method, and developed platforms to perform CFS campaigns with the help of crystallography experts and specific tools. The BESSY II synchrotron, operated by the Helmholtz-Zentrum Berlin (HZB), is one of these synchrotron facilities that offer a CFS platform, named the F2X-facility. Here, the specialized F2X workflow is described along with the relevant differences to other existing CFS platforms, and the ongoing developments aimed at supporting users of the facility. The different stages of a CFS campaign including requirements, beamline capabilities, and the software environment are detailed and explained. A unique F2X-GO kit is featured, which allows users the possibility of performing all sample preparation in their home laboratories. Furthermore, at the HZB a computational workflow has been built to support users beyond the hit identification stage. The advantages of the F2X-facility at HZB are described and references are provided to successfully conduct CFS.

1 | Introduction

The development of a new drug often comprises a decade-long journey with several distinct phases. After target selection, the initial phase in the process involves the discovery of chemical structures binding to the target, the so-called starting points. Currently, there are two prominent ways to identify starting

points for drug development: high-throughput screening (HTS) and fragment screening (FS). HTS entails the screening for larger, drug-like molecules (~500 Da) [1], while FS involves screening of smaller, fragment-sized (~300 Da) molecules [2]. Fragments are weak binders with affinities typically in the higher micromolar to millimolar range [2]. Their weak binding affinity is outweighed by the manifold advantages they offer as

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Author(s). *Applied Research* published by Wiley-VCH GmbH.

screening probes. Due to their small size, fewer molecules are needed to sample chemical space compared to using drug-like molecules [3]. Therefore, while screening hundreds to thousands of fragments is adequate for a single campaign, a HTS campaign may well entail the screening of millions of compounds [4]. Additionally, fragments engage in higher quality interactions with the protein surface, as their lower complexity allows them to avoid unfavorable structural motifs (Figure 1). It has been shown over the last years that FS is indeed a very promising method, with seven fragment-based FDA-approved drugs being already in the market and several more in clinical phases I–III ([5, 6] edition). These drugs address diverse protein targets, for example, protein–protein interactions, enzymes, and signaling proteins that have been thought to be undruggable before [7–10].

Weakly binding fragments may be identified using several biophysical methods. Prominent methods here are nuclear magnetic resonance (NMR), X-ray crystallography, and surface plasmon resonance (SPR). Of these, only X-ray crystallography enables the immediate recognition of the fragment’s binding site and pose. Furthermore, crystallographic fragment screening (CFS) has been demonstrated to be particularly efficient in identifying fragments bound to a protein target [11]. In addition, CFS is not only interesting for drug identification, but also as a basic research tool. It has been shown that fragments are able to recognize interesting functional surface areas on proteins that might hint towards putative PPI sites, allosteric sites, or other biologically relevant binding sites [12–15]. Thus, even when the protein target characteristics are not fully understood, conducting a CFS campaign can be beneficial in obtaining

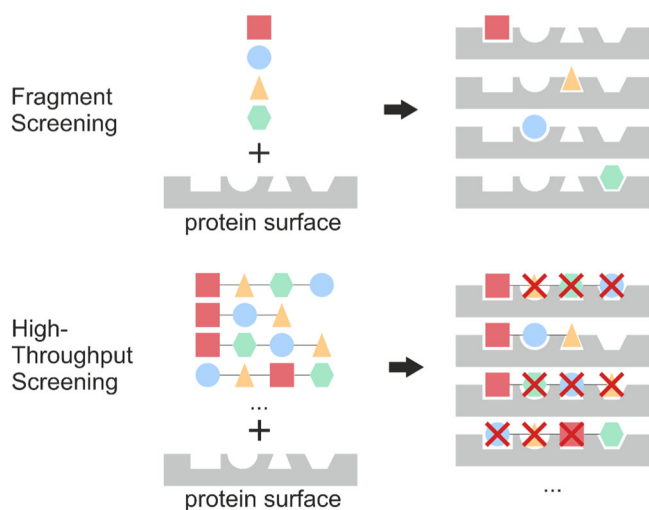


FIGURE 1 | Comparison of fragment screening (FS) and high-throughput screening (HTS). Fragment-sized structural features are shown as different shapes and are colored accordingly. The protein surface is shown in gray with indentations for possible interactions with molecules. In the case of FS, given that the correct structural features are present, fragments fit well to surface pockets of the protein without introducing unfavorable interactions. In case of HTS, the screening compounds are larger and contain several structural features. Consequently, it is necessary to screen significantly more compounds to find a candidate that fits all present indentations on a protein surface without introducing unfavorable interactions.

initial insights into potential target sites for future tool compounds or drug candidates.

During the last years, CFS has transformed into a high-throughput method owing to improvements in modern synchrotrons facilities, robot-assisted sample mounting, and fast read-out detectors [16, 17]. Based on these advances, CFS has gained quite high popularity among various scientific communities. However, as CFS experiments require working with hundreds to thousands of crystals, a specialized workflow and knowledge are needed for successful CFS campaigns. Several platforms at different synchrotron sites are currently available, for example, XChem at the Diamond Light Source [18], FragMAX at MAX IV [19], FFCS at Swiss Light Source [20], the fully automated FS platform at ESRF [21–23], the HiPhaX beamline at DESY for compound screening (P09 HiPhaX) [24] and the F2X-facility at BESSY II [25]. At BESSY II (Helmholtz-Zentrum Berlin, HZB), the Macromolecular Crystallography group operates three beamlines, with two beamlines in use for standard experiments and regular CFS campaigns, and one beamline for room-temperature experiments [26]. The F2X-platform there has been developed over the last 12 years with the goal to make CFS at BESSY II a straightforward and rapid process where external users are supported from final crystal optimization steps until the first optimization rounds of hit-to-lead development.

2 | The CFS Workflow

In general, CFS campaigns can be divided into six steps (Figure 2):

1. Preparation of the CFS campaign

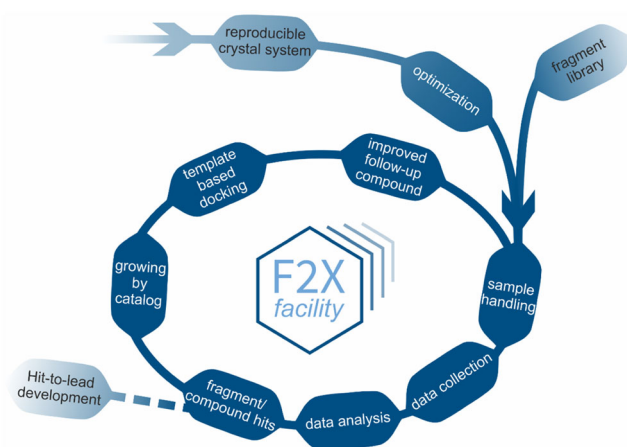


FIGURE 2 | Overview of the CFS campaign steps at the HZB BESSY II F2X-facility. First, users will optimize their sample and choose a fragment library to screen. Afterward, the sample is prepared, and data are collected and analyzed. Next, the hits are identified, and the user can either take this information and data to their home lab and continue with hit-to-lead development or they can use the Frag4Lead workflow to apply the “growing by catalog” approach to find follow-up compounds and validate them via X-ray crystallography. This optimization of fragment-sized hits to higher affinity binders can be repeated until the user can continue with their own optimization approaches.

2. Optimization of the CFS parameters
3. Crystal soaking
4. Crystal harvesting
5. Data collection
6. Data analysis.

In the following chapters, each of these steps as they are handled at the F2X-facility will be described in more detail and tips are given as to how novice CFS users are best supported. Additionally, the fragment optimization as provided will be discussed.

3 | Preparation of a CFS Campaign

A requirement for embarking on a CFS campaign at HZB is the availability of an optimized crystal system that allows the reliable production of several hundred crystals, which diffract to better than 2.5 Å resolution. These crystals should be of sufficient size of 50 μm or larger. Additionally, it is advantageous if several crystals appear per crystallization drop. Since only 1–2 crystals will be exposed to one fragment, it is imperative that each crystal possesses a high level of diffraction quality. A system which produces only one high-diffracting crystal out of 10 crystals, or yields crystals with similar morphology but varying space groups, does not qualify for CFS. These requirements are typical across the various CFS platforms at synchrotron facilities. Additionally, the crystals should take no longer than 2 weeks to grow, to keep the logistical effort manageable. Additionally, suitable crystal packing is crucial for a successful CFS campaign [27–29]. The users should investigate their crystal system for accessibility of their target sites and the size of their solvent channels, for example using the recently published LifeSoaks tool [30].

Concerning the cryoprotection of crystals, it is best practice to test different cryoprotectants at different concentrations to assure high-quality diffraction. This is not a strict requirement but should be considered by the user to achieve the highest quality diffraction data. Additionally, it should be tested how the crystals can be harvested optimally, for example, whether crystals can be scooped up by a loop to minimize surrounding solution, or whether they need to be harvested with more surrounding solution for optimal diffraction quality.

When these requirements are met, users can apply for a CFS campaign via a beamtime proposal through the user platform GATE (<https://www.helmholtz-berlin.de/pubbin/hzbgate>). After approval, the users will meet with one of the CFS user support at the F2X-facility and have three options to choose from. They can send protein to HZB and on-site user support will reproduce the crystallization, or they bring the crystals in the crystallization plate themselves to HZB, or they choose the F2X-GO option and prepare the samples in their home lab. After the crystallization option of choice has been established, beamtimes will be organized for further soaking optimization if needed and for the data collection of the CFS samples itself.

4 | Optimization of CFS Parameters

During the optimization process, several parameters will be tested to design a robust soaking experiment and data collection. Ideally, the soaking experiment should not impair crystal quality, that is, the average resolution and quality indicator values. In case of the soaking optimization, soaking times will be tested, different percentages of DMSO will be tried, and an increase in the buffer concentration and soaking in a cryo-protected solution will be tested. Longer soaking times will lead to a more comfortable and practical handling of crystals. For example, if overnight soaking is established, 1 day can be used exclusively for soaking and the next day for harvesting. The DMSO content will increase solubility of fragments and thus the final concentration of the fragments during soaking [31]. However, the presence of DMSO is not a strict requirement if our F2X-libraries are chosen for the CFS campaign [32] since all compounds have a high predicted solubility in water. In this case, the buffer concentration is increased to counteract possible pH shifts due to acidic or basic fragments dissolving in the soaking solution. Soaking in a cryo-protected solution will decrease the number of crystal transfer steps. It is also necessary to check the crystallographic data after performing the soaking experiment for any twinning, translational noncrystallographic symmetry (tNCS), space group ambiguity, or other special cases. Such data might lead to cumbersome and prolonged data analysis. Ideally, overnight soaking in about 5%–15% (v/v) DMSO with 100–300 mM of buffer component and cryo-protectant should be possible for the CFS campaign. The increased buffer concentration is important for maintaining the pH of the soaking solution in presence of high fragment concentration.

Depending on the space group, an optimized data collection routine can be established for the CFS campaign to result in complete data with minimal measurement time and exposure. Based on the user's experience of their crystal system a fixed detector distance is chosen. Test or characterization images are typically not collected.

5 | Crystal Handling and Data Management

The soaking and harvesting procedure for CFS campaigns at HZB has been extensively described in our publication in the *Journal of Visualized Experiments* [25]. In summary, a fragment plate will be provided with dried-on fragments on an MRC 3-lens 96-well low-profile plate. Forty microliters of soaking solution is added to the reservoir and 0.4 μL of the same soaking solution is added to the dried-on fragment. Per fragment, two crystals are then transferred from the crystal plate to the fragment plate. After the successful transfer of all crystals, the plate is sealed and incubated overnight at 20°C. The next day the crystals are harvested into unipucks and directly measured at the beamline or stored for the upcoming beamtime.

The special features of the crystal handling step in the CFS workflow at HZB are the high fragment concentrations used for soaking and the possibility to soak without the addition of organic solvents such as DMSO. High concentrations are

achieved by dissolving the dried-on fragments in small volumes of the soaking solution and transferring the crystal into the soaking drop. Through such reconstitution of dried-on fragments it is possible to achieve high fragment concentrations (up to 100 mM if 0.4 μ L drops are used), since the fragments can dissolve to concentrations up to their solubility limit in the specific soaking solution, independent of the presence of organic solvent. At other synchrotron sites fragments are transferred as small droplets from their organic solvent stock onto the crystal drops for soaking [18, 20]. This is a fast way of preparing fragment-soaked crystals, though it makes the fragment concentration dependent on the organic solvent content. If the protein crystals cannot tolerate organic solvents well, the achievable fragment concentration is reduced, or the campaign is not feasible at all.

The crystal handling is performed with the EasyAccess Frame [33] and is recorded via a tracking sheet provided by the CFS facility [25]. The EasyAccess Frame acts as an evaporation protection lid on a typical crystallization plate. This means that the crystallization foil can be removed completely from the plate. The EasyAccess Frame is then placed on the plate and cumbersome cutting through foil to open wells and resealing them is avoided. The big advantage of moving crystals from the crystallization plate to the soaking plate is that the use of crystals is more efficient. This comes at the expense of being slightly more manual than when the Crystal Shifter [34] or the CrystalDirect system [22, 23] is used. In comparison to the Crystal Shifter, the EasyAccess Frame allows more flexibility in handling protein crystals. As the device is very small and transportable it can be moved between microscopes, and the plate with the EasyAccess Frame on top can be rotated under the microscope to change the angle of the sample loop entering the soaking drop. Based on our previous experience, it is possible to handle up to 70 crystals per hour with the EasyAccess Frame. A comparison of the features of each device are given in Table 1.

The soaking drops will have varying appearances, depending on the solubility of the fragment in the soaking solution and the influence of the fragment on the crystal integrity. However, based on several CFS campaigns performed at HZB, even if this has changed after soaking, a decision on the diffraction quality based on the crystal morphology is not recommended. Each crystal that can still be harvested should be harvested.

Data collection is then performed either at BL14.1 or BL14.2 [26]. To achieve maximum data completeness in the minimum amount of time, the same data collection strategy is applied to

all crystals. This standard strategy is based on data collections of crystals soaked in soaking buffer but without fragment, which were performed before the actual CFS campaign. The detector distance is determined based on the highest achievable resolution of the crystal system and the rotation range is determined based on the space group of the crystal. These parameters will then remain fixed for all data collections of the campaign. This way, up to 240 samples can be measured in 24 h. All information following the data collection is then tracked in *FragMAX-app* [19], which allows for easy automatic processing, automatic refinement, and semi-automated hit identification via *PanDDA* [35]. The big progress provided by the development of *PanDDA* is that lower occupancy ligands can be identified in the ligand complex structures. While traditional difference electron density maps may show the ligands weakly or not at all, the *PanDDA* event map may still provide statistically significant evidence for the presence of the ligand [36, 37]. The inspection of the event map by an experienced crystallographer constitutes typically the end of a CFS campaign at HZB. It is of course clear that following the CFS campaign, refinement may be attempted to improve the overall structure, binding of the fragments may be further validated, and larger fragment analogs may be tested for binding. These steps are typically left to the users. The complete workflow of sample preparation, data collection, and data analysis as described here typically takes about 1–2 weeks.

6 | F2X-GO

At all synchrotrons offering a CFS platform, the usual approach is to send a protein sample to the facility, and crystallization is repeated on site along with soaking, crystal harvesting, and data collection. For this to work, it is vital that crystallization can be established on site at the facility. This is not a trivial endeavor and can often lead to lower-quality crystals or no crystal growth at all. Additionally, establishing crystallization on site of the facility can substantially increase the timeline of the campaign. It can also be the case that the protein sample needs special requirements like an anaerobic environment that cannot be guaranteed at the synchrotron site. Further, there might be travel restrictions as during the SARS-CoV-2 pandemic which prevents scientists traveling to the synchrotron site for soaking and harvesting experiments. To tackle such challenges, at HZB we developed the F2X-GO kit which includes all necessary tools to perform all CFS steps before the actual data collection in the user's home laboratory [25] (Figure 3).

The fragment libraries offered for CFS campaigns at HZB are the F2X-Entry Screen (96 fragments) as an easy start for new

TABLE 1 | Comparison of crystal handling devices.

	Crystal Shifter	CrystalDirect	EasyAccess Frame
Manual harvesting	Yes	No	Yes
Specific plate type	No	Yes	Yes
Cryo-cooling with solution	Yes	No	Yes
Transportable	No	No	Yes
Size	Medium	Large	Small
Price	High	Very high	Low

F2X-GO package:

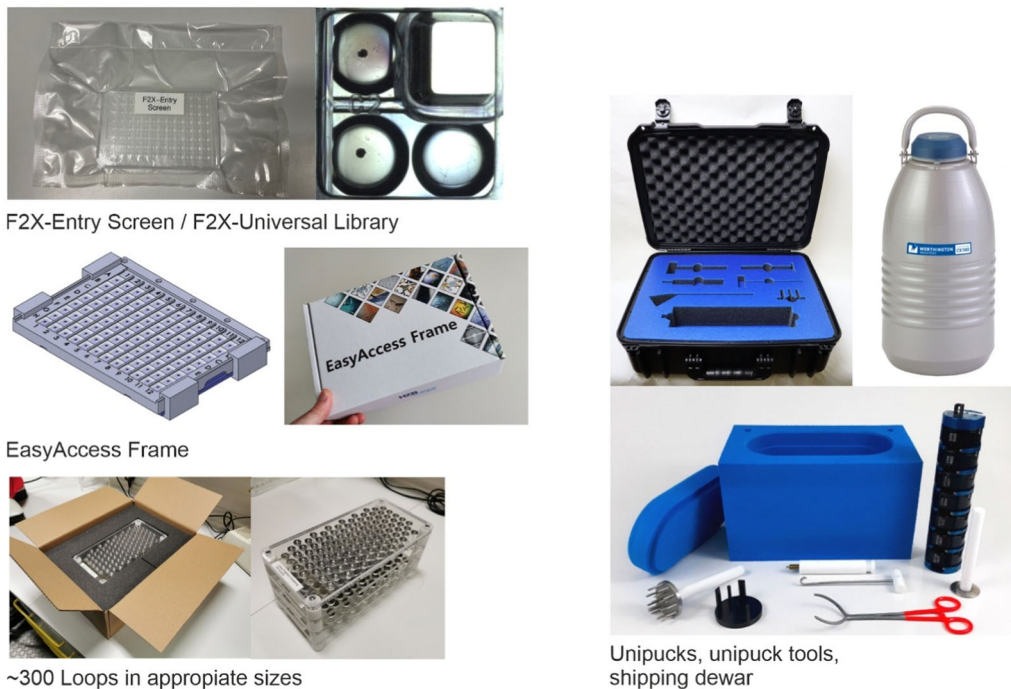


FIGURE 3 | The equipment when choosing the F2X-GO option is shown. All required equipment will be shipped to the user. It includes the ready-to-use soaking plate(s) either of the F2X-Entry Screen or of the F2X-Universal Library. The fragments are dried on the 96-well plate(s) without any disturbance even during transportation at room temperature. The EasyAccess Frame is also included if the user does not have one. Additionally, about 300 mounting loops in the appropriate sizes are included in the package. Inhouse developed loop racks secure the loops during travel. The necessary number of unipucks, unipuck tools, and dry shippers are sent to the user in case they do not have this equipment themselves.

CFS users and new projects, and the F2X-Universal Library (1103 fragments) for more advanced users and projects [32]. The F2X fragments are presented as dried-on fragments on MRC-3 96-well 3-lens low-profile plates. In each well only one fragment is presented. The fragments are then dissolved in a soaking solution, and crystals are transferred from the crystallization drop into the soaking drop. Co-crystallization can also be performed but might result in deformed crystal growth or no growth, making soaking the preferred option. This soaking plate can be shipped to other sites and has the advantage that no expensive and specialized machinery is necessary for crystal transfer onto this plate. To speed up crystal transfer and harvesting, another tool, the EasyAccess Frame, is shipped to the user as well [33]. This tool acts as a lid to protect crystal drops from evaporation during crystal manipulation. The EasyAccess Frame is lightweight and small, allowing easy shipment, and prevents evaporation for up to 6 h for most solutions. Additionally, as most academic groups will not have hundreds of loops in the same size available, an appropriate number of loops in inhouse-developed storage racks will be sent to them. Also provided are the necessary unipucks, transport dewar, and unipuck tools. The respective user support for CFS campaigns will be available remotely to answer any questions. Additionally, a detailed step-by-step explanation in the Journal of Visualized Experiments is available, which will help users to perform the experiments on their own [25].

The F2X-GO kit is available for all CFS users. It can be requested initially or be shipped in case crystallization on site did not result in useful crystal growth.

6.1 | Beamlines and IT Infrastructure at HZB

The MX-group at BESSY II operates two high-throughput beamlines (BL 14.1 and BL 14.2) suitable for CFS campaigns and one beamline (BL 14.3) that offers the option for more specialized experiments [26] (Figure 4). The specific characteristics of each beamline are shown in Table 2.

BL 14.1 is a variable-wavelength beamline, and it is equipped with a mini-kappa goniometer, which is particularly helpful for low-symmetry space groups or large unit cells. It can hold up to 144 samples in the unipuck system, so that during the measurement of a CFS campaign with redundant data collection for the F2X-Entry Screen, the samples would need to be exchanged only once. It is also possible to perform an energy-scan and select smaller beam sizes if needed.

BL 14.2 is also a variable-wavelength beamline and allows performing energy-scans. The recently installed ISARA2 sample dewar holds up to 464 samples in the unipuck system, allowing the collection of a full CFS campaign with redundant data collection for the F2X-Entry Screen. It is also feasible to collect data of light-sensitive samples at this beamline.

BL 14.3 is typically not utilized during CFS campaigns but offers the unique possibility to perform dehydration experiments and room-temperature experiments. Therefore, it can be used in the early stages of establishing a crystal system for CFS campaigns and for further analysis of bound fragments

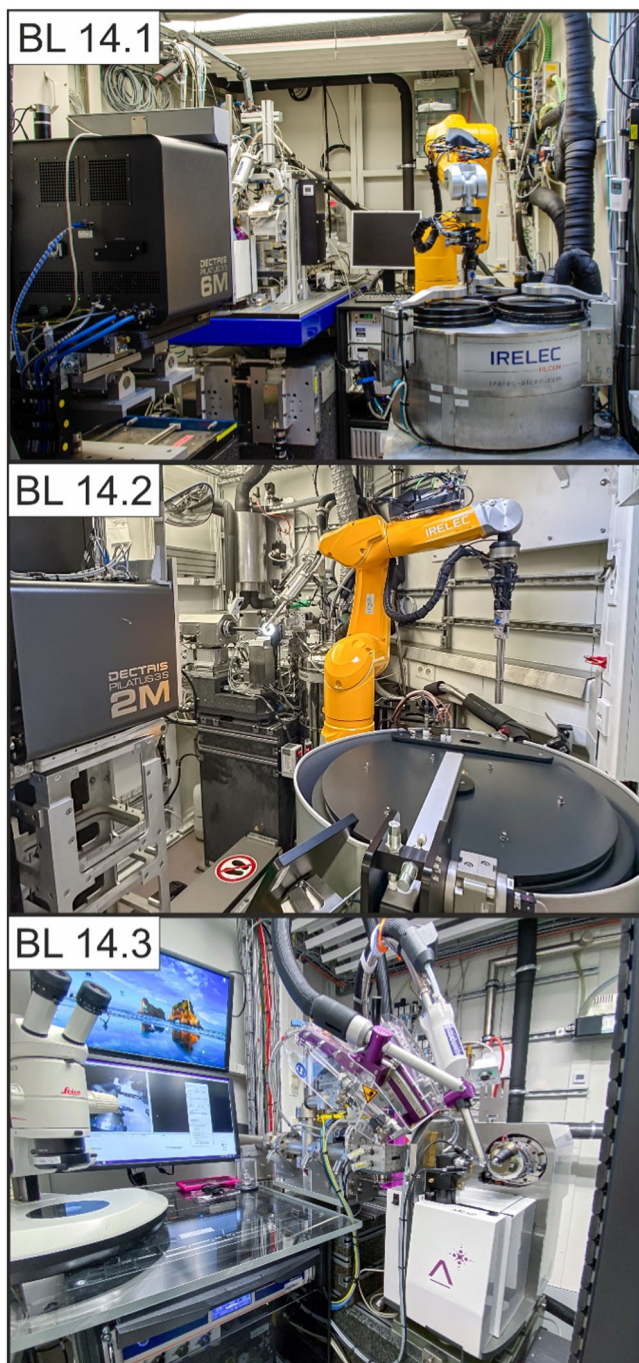


FIGURE 4 | Photos of the three beamlines operated by the macromolecular crystallography group at HZB.

like room-temperature data collection. The installed dehydration device allows cryoprotectant-less cryo-cooling of crystals.

All diffraction data arriving from the beamlines are transferred via private 10 GB-LX ethernet uplink to a dedicated HP-DL380GEN10 file server. From there, all data are transferred to the HZB unified experimental data storage SAN-system, which uses a HUAWEI Dorado 8000 v6. The MX group has currently a total of 100 TB storage capacity share dedicated for FS only, which can be expanded on demand.

For interactive data processing, each beamline has a dedicated Apple MAC studio ULTRA M1/M2 machine, which allows for extremely fast manual XDS-based diffraction data processing using XDSAPP [38]. In parallel, all incoming diffraction datasets are processed automatically via XDSAPP, and the results are presented to beamline users 10 to 15 min after data collection finished.

Our *FragMAXapp* [39] installation runs on a HP-DL580GEN10 compute server, which is equipped with 96 Intel XEON Platinum CPUs and has 790 GB RAM as well as several 10 GB-LX and SAN-interfaces connecting to all required experimental and metadata storage nodes.

6.2 | Software for CFS

To enable a fast data analysis of hundreds to thousands of datasets and track the data quality during data collection, all data is automatically processed via XDSAPP 3.1 running in command line mode [38]. It is based on XDS [40] and determines the most probable space group through three cycles of integration based on a POINTLESS analysis. Thus, it has an integrated decision-making capability and cuts the resolution of a data set automatically based on the $CC_{1/2}$ value at a correlation significance level of 0.1% (*t*-test). The current version of XDSAPP 3.1 runs on Python 3.7 and newer, PyQt5 and matplotlib. The command line version runs fully independently on PyQt5 and matplotlib. XDS plugins are available to process EIGER data, namely dectris-neggia (Dectris) and durin-plugin (Diamond Light Source). An additional results file is created in cif format with entries according to PDBx/mmCIF standards. In the case of troublesome datasets, it is possible to process data manually via a user-friendly GUI, which allows tracking of the program through each integration cycle by observing plots like no. of rejected spots in each frame. All logfiles can be loaded in the GUI and the settings can be changed if needed. However, in CFS campaigns, processing is often straightforward, therefore manual processing is usually not necessary. XDSAPP is accessible for users via <https://www.helmholtz-berlin.de/xdsapp>.

The next step of automatic refinement performed at the F2X-facility is mainly via *fspipeline* [41]. The program runs only in a command line mode with an input model provided by the users and the mtz files from data processing. The pipeline utilizes the PHENIX package [42] and Coot [43]. Before refinement starts, if a twin law is found then *phenix.xtriage* is run to look for twin laws for the subsequent steps. Only one twin law can be used by PHENIX, thus if more are found a warning is issued and no twin law is used. Next, the input model is prepared for refinement. To this end, *fspipeline* determines the protein sequence and the space group of the input model, then removes any waters, ligands, and hydrogen atoms and sets the thermal displacement parameters to isotropic values with limited spread around the average B factor to avoid bias from extreme values. Additionally, if not given by the user, TLS groups are determined via *phenix.find_tls_groups*. Afterward, all mtz files in the underlying directories are located, except for directories that include the name “fspipeline.”

TABLE 2 | Characteristics of the HZB-MX beamlines.

	BL14.1	BL14.2	BL14.3
Wavelength range (Å)	0.8–2.25	0.8–2.25	0.89
Photon flux at sample (Phot/ s × 100 mA × 0.05% BW)	1.6×10^{11} (at $\lambda = 0.92$ Å)	1.5×10^{11} (at $\lambda = 0.92$ Å)	2.3×10^{10}
Energy resolution (eV)	< 2	< 2	< 5
Beam size (μm diameter)	50–100	100	50–200
Goniometer	MD2-microdiffractometer with MK3	Nanodiffractometer	MD2S-microdiffractometer with MK3
X-ray detector	PILATUS3 S 6M ^a	PILATUS3 S 2M	PILATUS2 6M ^a
Sample mounting	CATS robot	ISARA2 robot	Manual
No. of samples in sample dewar	144 Unipuck	464 Unipuck	—
Exposure times (s/°)	0.4–10 ^b	0.4–10	0.9–20
Detector distance range (mm)	140–649	57–800	110–501
Achievable resolution (Å)	0.84	0.71	0.85
Maximum unit cell length (Å)	600 ($d_{\min} = 2.0$ Å)	400 ($d_{\min} = 2.0$ Å)	600 ($d_{\min} = 2.0$ Å)
Special equipment and operations	<ul style="list-style-type: none"> • Crystal annealing • UV-pulsed laser for UV-RIP 	<ul style="list-style-type: none"> • Long wavelength and atomic resolution data collection • adjustable ambient light environment for light sensitive samples 	<ul style="list-style-type: none"> • RT data collection • Controlled dehydration • REX nozzle exchanger

Note: For each beamline typical parameters are given with the information of special equipment available at the beamlines.

^aDetector at BL14.1 will be exchanged for a PILATUS3 × 6 M from BL14.1 will be moved to BL14.3 and replace PILATUS2 6 M detector.

^bAfter the detector exchange the possible exposure times will be 0.1–10.

This exception allows *fspipeline* to be run several times in the same directory with different options if needed. Only data with a resolution better than 3.5 Å are considered for refinement. In the first refinement step, either a rigid body refinement or molecular replacement are performed. For this decision, *fspipeline* checks whether the input mtz and pdb files contain the same space group and similar cell parameters. If they match, a rigid body refinement is performed. In case they do not match, molecular replacement is performed. Additionally, if after the rigid body refinement, the R_{work} value is above 40% molecular replacement is performed automatically. A full molecular replacement can be performed from the start by giving the option `mr=True`. The next step, so-called “initial refinement” includes a standard coordinate and isotropic atomic displacement parameter (ADP, B factor) refinement. Additionally, simulated annealing can be performed during this step but is skipped as default. Afterwards, a TLS refinement is performed. During the next step, *Coot* [43] in the command line mode is used to place water molecules while ignoring electron density features that represent larger molecules. The water molecule addition can also be done via *PHENIX* with the option `water=phenix`. The following refinement step depends on the resolution of the data. In case the resolution is better than 1.6 Å, anisotropic ADP refinement of the protein and anisotropic ADP refinement of the water molecules is tested. If the final R_{work} and R_{free}

are lowered by at least 0.5%, anisotropic refinement is kept for the next refinement steps. A further refinement is performed with the addition of hydrogen atoms to the protein. In the default case that *Coot* is used for water addition, a final refinement is done with a second water addition and check cycle. *Fspipeline* decides that the next step is the last refinement step if the improvement of the R_{work} is less than 0.5% from the previous refinement. The output of *fspipeline* includes the final pdb file, final mtz file, and a log file. Additionally, each refinement step is tracked in individual directories with the respective pdb file, mtz file, log file, and python script. An overview of refinement statistics for each data set during one *fspipeline* run are given in the `results.html`.

The output of *fspipeline* can then be transferred to *PanDDA* for hit identification [35]. *PanDDA* utilizes the complete CFS data set made up of hundreds to thousands of datasets and identifies weak binding fragments. As *PanDDA* needs homogenous data for successful employment, another program called *cluster4x* can be used to identify clusters based on $C\alpha$ positions or amplitudes [44]. In this way, *cluster4x* can help to remove outliers from the CFS data set or split it into more homogenous parts [14].

All programs (except for *cluster4x*) and several more like *dials* [45] and *dimple* [46] are incorporated in a specialized

web-based data management platform called *FragMAXapp* [39]. This user-friendly platform is used to streamline auto-processing, auto-refinement, and hit identification. It greatly simplifies tracking CFS data and makes it easy for users to start processing jobs and to investigate the results. This platform has been developed at MAX IV with input by HZB and is maintained at MAX IV. *FragMAXapp* does not need to be installed by the user—projects can simply be set up with minimal input using the web-based platform, which can also be done remotely, allowing the user to handle their data as they see fit. Other synchrotron sites have their own development to help with data management like *XChemExplorer* at Diamond Light Source [47] or *HEIDI* at Swiss Light Source [48].

6.3 | Support Beyond the CFS Campaign

In a typical CFS campaign, the identification of the hits is the last step of the campaign. Following that, the users need to find ways to optimize their fragment hits. As this can be especially challenging for academic groups without ready access to medicinal chemistry expertise, a computational platform called *Frag4Lead* was established to support users in their first steps of fragment optimization [49] (Figure 5). The *Frag4Lead* workflow applies the “growing by catalog” approach for fragment hits, considering one fragment at a time. It has been built using *KNIME* [50] to streamline the computational steps. It is consequently rather easy to manipulate the input parameters without much programming knowledge. Since the publication, the *Frag4Lead* workflow has been adapted in several places to increase success rates. The new *Frag4Lead2* workflow incorporates the program *SeeSAR* by BioSolveIT [51]. The user provides a pdb structure of their hit binding to the protein. The workflow will then prepare the ligand and protein for template-based docking and perform a substructure search automatically in MolPort (Molport Website) [52] and Chemspace (Chemspace Website) [53]. To increase successful identification of promising follow-up compounds we use *SpaceMACS* [54] to search in larger make-on-demand spaces like the ENAMINE REAL space (REAL Space - Enamine) [55]. This substructure search performed via *SpaceMACS* can be inserted into the workflow and be filtered by Lipinski’s rule-of-5 [56] and known pan-assay interference compounds (PAINS) [57–59]. The user will then be asked to define the binding site for the docking program *FlexX* [60]. The prepared ligand, protein structure, and list of superstructures is then fed into *FlexX* for template-based docking. Afterward, the list of docking poses is filtered, as described previously [49], to a smaller list of promising docking poses. The final step is a rescoring via the program *HYDE* [61] and further filtering for *HYDE*-specific output like estimated affinity, torsion angles, and clashes. The output is a *SeeSAR* session which will then be investigated by the user in *SeeSAR*. This way we can conduct multiple rounds of docking and searching commercially available compound spaces to arrive at a higher affinity binder before engaging medicinal chemists for further optimization.

Other CFS platforms often do not support further steps after the CFS campaign. At Diamond Light Source it is possible to run the *Fragalysis* program (Fragalysis @ Diamond) [62]. It allows

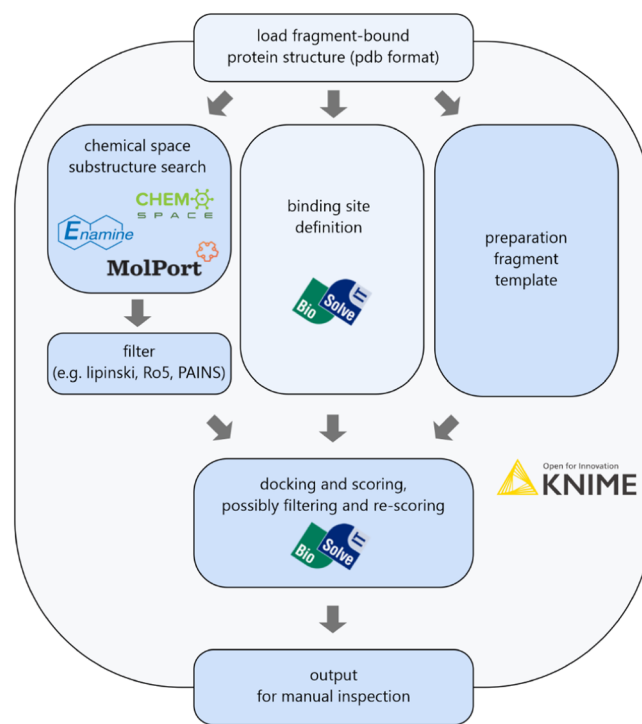


FIGURE 5 | Schematic overview of the computational *KNIME* workflow (*Frag4Lead2*). The applied software is indicated via the logo and inside the fields to indicate which step is covered by which software. User input is only needed to load a protein structure with the fragment bound and during the binding site definition in BioSolveIT *SeeSAR* (fields in lighter blue). The other steps of the workflow will be done automatically or with minimal input by the user support. The chemical space search can in most parts be done via *KNIME* APIs (ChemSPACE, MolPort), but it is also possible to load lists of compounds into the workflow (ENAMINE). The docking and scoring are performed with BioSolveIT Software. The output is handed back to the user.

the user to visualize their fragment hits and investigate their binding mode. It also includes the Astex Fragment Network [63], which helps users to perform SAR studies. Another program they developed is called *Fragmenstein* (GitHub—matteoferla/Fragmenstein) [64]. This program allows users to perform merging and linking of their fragment hits. However, so far it does not include synthetic tractability. The *Frag4Lead2* workflow also does not work with synthetic tractability but focuses on commercially available and made-on-demand compounds. In the case that linking and merging ideas come up after a CFS campaign the *Frag4Lead2* workflow can also be used to search for similar compounds that are commercially available based on the merging or linking idea. This way the users can find promising compounds to reach a lead structure.

7 | Conclusion

The CFS platform F2X-facility at HZB enables academic and industrial users to perform efficient CFS campaigns. Support to users is provided from the initial improvement of their crystal system through the first stages of optimizing their fragment hits. While the CFS workflow at HZB is slightly slower

compared to other synchrotrons with CFS platforms, due to the difference in sample preparation and the lower photon flux of the synchrotron, the benefits of the F2X-facility include: (1) enhanced soaking concentration by transferring a crystal to a soaking drop, potentially leading to higher hit rates; (2) organic solvent-independent soaking; (3) the flexibility to conduct sample preparation in the user's laboratory with the F2X-GO option; (4) additional assistance during the initial stages of hit-to-lead optimization. This comprehensive workflow allows us to assist our users throughout their entire CFS campaign, extending far beyond hit identification.

Acknowledgments

We would like to thank the BioMAX team at MAX IV, especially Tobias Krojer, Elmir Jagudin, and the former member Gustavo M. A. Lima for their immense help with employing the *FragMAXapp* at HZB. Additionally, we acknowledge the collaboration between the MX-group at HZB and the Drug Design Group at Marburg University for fruitful developments regarding the F2X-libraries, the EasyAccess Frame and the Frag4Lead workflow. Here, we would like to especially highlight the great work together with Gerhard Klebe and Alexander Metz. We would also like to thank all our fragment screening users for their great collaborations and feedback. Based on roughly 50 user fragment-screening campaigns plenty of experiences have been collected, which led to improvements at all steps of the process. These developments were supported by the German Ministry of Science and Education (BMBF) via the projects Frag2Xtal (No. 05K13RM1) and Frag4Lead (No. 05K16RM1), as well as by iNEXT-Discovery, project No. 871037, funded by the Horizon 2020 program of the European Commission, the German Research Foundation (DFG) via the project NECESSITY (FE2166/1-1) and the German Academic Exchange Service (DAAD) via the project CFS-Pipeline (57654470). Open Access funding enabled and organized by Projekt DEAL.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

References

1. M. Entzeroth, H. Flotow, and P. Condrón, "Overview of High-Throughput Screening," *Current Protocols in Pharmacology* 44 (2009): 9.4.1–9.4.27.
2. D. A. Erlanson, "Introduction to Fragment-Based Drug Discovery," *Topics in Current Chemistry* 317 (2012): 1–32.
3. D. A. Erlanson, S. W. Fesik, R. E. Hubbard, W. Jahnke, and H. Jhoti, "Twenty Years On: The Impact of Fragments on Drug Discovery," *Nature Reviews Drug Discovery* 15 (2016): 605–619.
4. Y. Shi, and M. Itzstein, von, "How Size Matters: Diversity for Fragment Library Design," *Molecules* 24 (2019): 2838, <https://doi.org/10.3390/MOLECULES24152838>.
5. I. J. P. De Esch, D. A. Erlanson, W. Jahnke, C. N. Johnson, and L. Walsh, "Fragment-to-Lead Medicinal Chemistry Publications in 2020," *Journal of Medicinal Chemistry* 65 (2022): 84–99.
6. Practical Fragments: Fragments in the Clinic. 2024 Edition, <http://practicalfragments.blogspot.com/2024/02/fragments-in-clinic-2024-edition.html>.
7. J. M. Ostrem, U. Peters, M. L. Sos, J. A. Wells, and K. M. Shokat, "K-Ras(G12C) Inhibitors Allosterically Control GTP Affinity and Effector Interactions," *Nature* 503 (2013): 548–551.
8. J. Schoepfer, W. Jahnke, G. Berellini, et al., "Discovery of Asciminib (ABL001), An Allosteric Inhibitor of the Tyrosine Kinase Activity of BCR-ABL1," *Journal of Medicinal Chemistry* 61 (2018): 8120–8135.
9. W. D. Tap, Z. A. Wainberg, S. P. Anthony, et al., "Structure-Guided Blockade of CSF1R Kinase in Tenosynovial Giant-Cell Tumor," *New England Journal of Medicine* 373 (2015): 428–437.
10. G. Bollag, P. Hirth, J. Tsai, et al., "Clinical Efficacy of a RAF Inhibitor Needs Broad Target Blockade in BRAF-Mutant Melanoma," *Nature* 467 (2010): 596–599.
11. J. Schiebel, N. Radeva, S. G. Krimmer, et al., "Six Biophysical Screening Methods Miss a Large Proportion of Crystallographically Discovered Fragment Hits: A Case Study," *ACS Chemical Biology* 11 (2016): 1693–1701.
12. T. Krojer, J. S. Fraser, and F. von Delft, "Discovery of Allosteric Binding Sites by Crystallographic Fragment Screening," *Current Opinion in Structural Biology* 65 (2020): 209–216.
13. R. F. Ludlow, M. L. Verdonk, H. K. Saini, I. J. Tickle, and H. Jhoti, "Detection of Secondary Binding Sites in Proteins Using Fragment Screening," *Proceedings of the National Academy of Sciences United States of America* 112 (2015): 15910–15915.
14. T. Barthel, J. Wollenhaupt, G. M. A. Lima, M. C. Wahl, and M. S. Weiss, "Large-Scale Crystallographic Fragment Screening Expedites Compound Optimization and Identifies Putative Protein-Protein Interaction Sites," *Journal of Medicinal Chemistry* 65 (2022): 14630–14641.
15. I. A. Shumilin, M. Cymborowski, O. Chertihin, et al., "Identification of Unknown Protein Function Using Metabolite Cocktail Screening," *Structure* 20 (2012): 1715–1725.
16. R. L. Owen, J. Juanhuix, and M. Fuchs, "Current Advances in Synchrotron Radiation Instrumentation for MX Experiments," *Archives of Biochemistry and Biophysics* 602 (2016): 21–31.
17. M. C. Thompson, T. O. Yeates, and J. A. Rodriguez, "Advances in Methods for Atomic Resolution Macromolecular Structure Determination," *F1000Research* 9 (2020): 667, <https://doi.org/10.12688/F1000RESEARCH.25097.1>.
18. A. Douangamath, A. Powell, D. Fearon, et al., "Achieving Efficient Fragment Screening at XChem Facility at Diamond Light Source," *Journal of Visualized Experiments* (2021), <https://doi.org/10.3791/62414>.
19. G. M. A. Lima, V. O. Talibov, E. Jagudin, et al., "FragMAX: The Fragment-Screening Platform at the MAX IV Laboratory," *Acta Crystallographica Section D Structural Biology* 76 (2020): 771–777.
20. J. W. Kaminski, L. Vera, D. P. Stegmann, et al., "Fast Fragment- and Compound-Screening Pipeline at the Swiss Light Source," *Acta Crystallographica Section D Structural Biology* 78 (2022): 328–336.
21. I. Cornaci, R. Bourgeas, and G. Hoffmann, "The Automated Crystallography Pipelines at the EMBL HTX Facility in Grenoble," *Journal of Visualized Experiments* 172 (2021): e62491, <https://doi.org/10.3791/62491-V>.
22. J. A. Márquez and F. Cipriani, "CrystalDirect™: A Novel Approach for Automated Crystal Harvesting Based on Photoablation of Thin Films," *Methods in Molecular Biology* 1091, (2014): 197–203.
23. F. Cipriani, M. Röwer, and C. Landret, "CrystalDirect: A New Method for Automated Crystal Harvesting Based on Laser-Induced Photoablation of Thin Films," *Acta Crystallographica* 68 (2012): 1393–1399.
24. P09 HiPhaX, https://photon-science.desy.de/facilities/petra_iii/beamlines/p09_resonant_diffraction___hiphax/p09_hiphax/index_eng.html.

25. J. Wollenhaupt, T. Barthel, G. M. A. Lima, et al., "Workflow and Tools for Crystallographic Fragment Screening at the Helmholtz-Zentrum Berlin," *Journal of Visualized Experiments* 2021, (2021): 1–19.
26. U. Mueller, R. Förster, M. Hellmig, et al., "The Macromolecular Crystallography Beamlines at BESSY II of the Helmholtz-Zentrum Berlin: Current Status and Perspectives," *The European Physical Journal-Plus* 130, (2015): 1–10.
27. E. Costanzi, M. Kuzikov, F. Esposito, et al., "Structural and Biochemical Analysis of the Dual Inhibition of MG-132 against SARS-CoV-2 Main Protease (Mpro/3CLpro) and Human Cathepsin-L," *International Journal of Molecular Sciences* 22 (2021): 11779, <https://doi.org/10.3390/IJMS222111779>.
28. C. Y. Huang, A. Metz, R. Lange, et al., "Fragment-Based Screening Targeting an Open Form of the SARS-CoV-2 Main Protease Binding Pocket," *Acta Crystallographica Section D Structural Biology* 80 (2024): 123–136.
29. G. D. Noske, A. M. Nakamura, V. O. Gawriljuk, et al., "A Crystallographic Snapshot of SARS-CoV-2 Main Protease Maturation Process," *Journal of Molecular Biology* 433 (2021): 167118.
30. J. Pletzer-Zelgert, C. Ehr, I. Fender, et al., "Lifesoaks: A Tool for Analyzing Solvent Channels in Protein Crystals and Obstacles for Soaking Experiments," *Acta Crystallographica Section D Structural Biology* 79 (2023): 837–856.
31. D. E. Danley, "Crystallization to Obtain Protein–Ligand Complexes for Structure-Aided Drug Design," *Acta Crystallographica Section D Biological Crystallography* 62 (2006): 569–575.
32. J. Wollenhaupt, A. Metz, T. Barthel, et al., "F2X-Universal and F2X-Entry: Structurally Diverse Compound Libraries for Crystallographic Fragment Screening," *Structure* 28 (2020): 694–706.e5.
33. T. Barthel, F. U. Huschmann, D. Wallacher, et al., "Facilitated Crystal Handling Using a Simple Device for Evaporation Reduction in Microtiter Plates," *Journal of Applied Crystallography* 54 (2021): 376–382.
34. N. D. Wright, P. Collins, L. Koekemoer, et al., "The Low-Cost Shifter Microscope Stage Transforms the Speed and Robustness of Protein Crystal Harvesting," *Acta Crystallographica Section D Structural Biology* 77 (2021): 62–74.
35. N. M. Pearce, T. Krojer, A. R. Bradley, et al., "A Multi-Crystal Method for Extracting Obscured Crystallographic States From Conventionally Uninterpretable Electron Density," *Nature Communications* 8 (2017): 15123.
36. M. Jaskolski, A. Wlodawer, Z. Dauter, W. Minor, and B. Rupp, "Group Depositions to the Protein Data Bank Need Adequate Presentation and Different Archiving Protocol," *Protein Science* 31 (2022): 784–786.
37. M. S. Weiss, J. Wollenhaupt, G. J. Correy, et al., "Of Problems and Opportunities—How to Treat and How to Not Treat Crystallographic Fragment Screening Data," *Protein Science* 31 (2022): e4391, <https://doi.org/10.1002/PRO.4391>.
38. K. M. Sparta, M. Krug, U. Heinemann, U. Mueller, and M. S. Weiss, "XDSAPP2.0," *Journal of Applied Crystallography* 49 (2016): 1085–1092.
39. G. M. A. Lima, E. Jagudin, V. O. Talibov, et al., "FragMAXapp: Crystallographic Fragment-Screening Data-Analysis and Project-Management System," *Acta Crystallographica Section D Structural Biology* 77 (2021): 799–808.
40. W. Kabsch, "XDS," *Acta Crystallographica Section D Biological Crystallography* 66 (2010): 125–132.
41. J. Schiebel, S. G. Krimmer, K. Röwer, et al., "High-Throughput Crystallography: Reliable and Efficient Identification of Fragment Hits," *Structure* 24 (2016): 1398–1409.
42. D. Lieschner, P. V. Afonine, M. L. Baker, et al., "Macromolecular Structure Determination Using X-Rays, Neutrons and Electrons: Recent Developments Inphenix," *Acta Crystallographica Section D Structural Biology* 75 (2019): 861–877.
43. P. Emsley, B. Lohkamp, W. G. Scott, and K. Cowtan, "Features and Development Ofcoot," *Acta Crystallographica Section D Biological Crystallography* 66 (2010): 486–501.
44. H. M. Ginn, "Pre-Clustering Data Sets Using Cluster4x Improves the Signal-to-Noise Ratio of High-Throughput Crystallography Drug-Screening Analysis," *Acta Crystallographica Section D Structural Biology* 76 (2020): 1134–1144.
45. G. Winter, D. G. Waterman, J. M. Parkhurst, et al., "DIALS: Implementation and Evaluation of a New Integration Package," *Acta Crystallographica Section D Structural Biology* 74 (2018): 85–97.
46. M. Wojdyr, R. Keegan, G. Winter, and A. Ashton, "DIMPLE - A Pipeline for the Rapid Generation of Difference Maps From Protein Crystals With Putatively Bound Ligands," *Acta Crystallographica Section A* 69 (2013): s299.
47. T. Krojer, R. Talon, N. Pearce, et al., "The XChem Explorer Graphical Workflow Tool for Routine or Large-Scale Protein–Ligand Structure Determination," *Acta Crystallographica Section D Structural Biology* 73 (2017): 267–278.
48. A. Metz, D. P. Stegmann, E. H. Panepucci, et al., "HEIDI: An Experiment-Management Platform Enabling High-Throughput Fragment and Compound Screening," *Acta Crystallographica Section D Structural Biology* 80 (2024): 328–335.
49. A. Metz, J. Wollenhaupt, S. Glöckner, et al., "Frag4Lead: Growing Crystallographic Fragment Hits by Catalog Using Fragment-Guided Template Docking," *Acta Crystallographica Section D Structural Biology* 77 (2021): 1168–1182.
50. M. R. Berthold, N. Cebron, F. Dill, et al., "KNIME—The Konstanz Information Miner: Version 2.0 and Beyond," *ACM SIGKDD Explorations Newsletter* 11 (2009): 26–31.
51. BioSolveIT. (2022), <https://www.biosolveit.de/SeeSAR/>.
52. Molport—Compound Sourcing, Selling and Purchasing Platform, <https://www.molport.com/shop/index>.
53. ChEMSPACE—The Provider of Integrated Discovery Services, <https://chem-space.com/>.
54. R. Schmidt, R. Klein, and M. Rarey, "Maximum Common Substructure Searching in Combinatorial Make-on-Demand Compound Spaces," *Journal of Chemical Information and Modeling* 62 (2022): 2133–2150.
55. REAL Space—Enamine, <https://enamine.net/compound-collections/real-compounds/real-space-navigator>.
56. C. A. Lipinski, F. Lombardo, B. W. Dominy, and P. J. Feeney, "Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings IPII of Original Article: S0169-409X(96)00423-1. The Article Was Originally Published in *Advanced Drug Delivery Reviews* 23 (1997) 3–25. 1," *Advanced Drug Delivery Reviews* 46 (2001): 3–26.
57. J. B. Baell, "Broad Coverage of Commercially Available Lead-Like Screening Space With Fewer Than 350,000 Compounds," *Journal of Chemical Information and Modeling* 53 (2013): 39–55.
58. J. B. Baell, L. Ferrins, H. Falk, and G. Nikolakopoulos, "PAINS: Relevance to Tool Compound Discovery and Fragment-Based Screening," *Australian Journal of Chemistry* 66 (2013): 1483–1494.
59. S. Saubern, R. Guha, and J. B. Baell, "KNIME Workflow to Assess PAINS Filters in SMARTS Format. Comparison of RDKit and Indigo Cheminformatics Libraries," *Molecular Informatics* 30 (2011): 847–850.
60. M. Rarey, B. Kramer, T. Lengauer, and G. Klebe, "A Fast Flexible Docking Method Using an Incremental Construction Algorithm," *Journal of Molecular Biology* 261 (1996): 470–489.

61. N. Schneider, G. Lange, S. Hindle, R. Klein, and M. Rarey, “A Consistent Description of Hydrogen Bond and Dehydration Energies in Protein–Ligand Complexes: Methods Behind the Hyde Scoring Function,” *Journal of Computer-Aided Molecular Design* 27 (2013): 15–29.
62. Fragalysis @ Diamond, <https://fragalysis.diamond.ac.uk/viewer/react/landing>.
63. R. J. Hall, C. W. Murray, and M. L. Verdonk, “The Fragment Network: A Chemistry Recommendation Engine Built Using a Graph Database,” *Journal of Medicinal Chemistry* 60 (2017): 6440–6450.
64. GitHub—Matteoferla/Fragmenstein, <https://github.com/matteoferla/Fragmenstein>.