Journal of Visualized Experiments

The Automated Crystallography Pipelines at the EMBL HTX facility in Grenoble

--Manuscript Draft--

Article Type:	Invited Methods Collection - JoVE Produced Video
Manuscript Number:	JoVE62491R1
Full Title:	The Automated Crystallography Pipelines at the EMBL HTX facility in Grenoble
Corresponding Author:	Jose Marquez EMBL Grenoble, Isere FRANCE
Corresponding Author's Institution:	EMBL
Corresponding Author E-Mail:	marquez@embl.fr
Order of Authors:	Irina Cornaciu
	Raphael Bourgeas
	Guillaume Hoffmann
	Florine Dupeux
	Anne-Sophie Humm
	Andrea Pica
	Damien Clavel
	Gael Seroul
	Peter Murphy
	Jose Marquez
Additional Information:	
Question	**Response**
Please indicate whether this article will be Standard Access or Open Access.	Open Access (US$4,200)
Please specify the section of the submitted manuscript.	Biochemistry
Please indicate the **city, state/province, and country** where this article will be **filmed**. Please do not use abbreviations.	71 Avenue de Martyres, 38000 Grenoble, France
Please confirm that you have read and agree to the terms and conditions of the author license agreement that applies below:	I agree to the Author License Agreement
Please provide any comments to the journal here.	
Please indicate whether this article will be Standard Access or Open Access.	Standard Access ($1400)

TITLE:

The Automated Crystallography Pipelines at the EMBL HTX facility in Grenoble.

AUTHORS AND AFFILIATIONS:

Irina Cornaciu[1,2], Raphael Bourgeas[1], Guillaume Hoffmann[1], Florine Dupeux[1,3], Anne-Sophie Humm[1], Andrea Pica[1,2], Damien Clavel[1,2], Gael Seroul[1,2], Peter Murphy[1], José Antonio Márquez[1,2]

[1]European Molecular Biology Laboratory, Grenoble, France
[2]ALPX S.A.S. 71, Avenue des Martyrs, Grenoble, France
[3]Institut de Biologie Structurale, Grenoble, France

cornaciu@embl.fr
rbourgeas@embl.fr
ghoffman@embl.fr
fdupeux@embl.fr
ahumm@embl.fr
apica@embl.fr
dclavel@embl.fr
seroul@embl.fr
pmurphy@embl.fr
marquez@embl.fr

Corresponding author:
José Antonio Márquez

SUMMARY:

Here, we describe how to use the automated macromolecular crystallography pipelines for protein-to-structure, rapid ligand-protein complex analysis and large-scale fragment screening based on the CrystalDirect technology at the HTX Laboratory in EMBL Grenoble.

ABSTRACT:

EMBL Grenoble operates the High Throughput Crystallization Laboratory (HTX Lab), a large-scale user facility offering high throughput crystallography services to users worldwide. The HTX lab has a strong focus in the development of new methods in macromolecular crystallography. Through the combination of a high throughput crystallization platform, the CrystalDirect technology for fully automated crystal mounting and cryocooling and the CRIMS software we have developed fully automated pipelines for macromolecular crystallography that can be remotely operated over the internet. These include a protein-to-structure pipeline for the determination of new structures, a pipeline for the rapid characterization of protein-ligand complexes in support of medicinal chemistry, and a large-scale, automated fragment screening pipeline enabling evaluation of libraries of over 1000 fragments. Here we describe how to access and use these resources.

INTRODUCTION:

Automation has been introduced at all steps of the macromolecular crystallography experimental process, from crystallization to diffraction data collection and processing[1–6] and including a number of technologies for sample mounting[7–11]. This has not only accelerated the pace at which crystallographic structures are obtained but has contributed to streamline applications like structure guided drug design[12–15]. In this manuscript we describe some of the aspects of the automated crystallography pipelines available at the HTX lab in Grenoble as well as the underlying technologies.

The HTX lab at EMBL Grenoble is one of the largest academic facilities for crystallization screening in Europe. It is co-located at the European Photon and Neutron (EPN) campus together with the European Synchrotron Radiation Facility (ESRF), which produces some of the world's most brilliant X-ray beams and the Institut Laue Langevin (ILL), which provides high flux neutron beams. Since the start of operations in 2003 the HTX lab has provided services to over 800 scientists and processes more than 1000 samples per year. The HTX lab has a strong focus in the development of new methods in macromolecular crystallography, including methods for sample evaluation and quality control[16–18] and the CrystalDirect technology, enabling fully automated crystal mounting and processing[19–21]. The HTX lab has also developed the Crystallographic Information Management System (CRIMS), a web-based laboratory information system that provides automated communication between crystallization and synchrotron data collection facilities, enabling uninterrupted information flow over the whole sample cycle from pure protein to diffraction data. Through the combination of the capacities of the HTX facility, the CrystalDirect technology and the CRIMS software, we have developed fully automated protein-to-structure pipelines integrating crystallization screening, crystal optimization, automated crystal harvesting processing and cryocooling and X-ray data collection at multiple synchrotrons into a single and continuous workflow that can be remotely operated through a web browser. These pipelines can be applied to support rapid determination of new structures, the characterization of protein-ligand complexes and large-scale compound and fragment screening through X-ray crystallography.

The HTX lab is equipped with a nonvolume crystallization robot (including an LCP module that enables crystallization of both soluble and membrane proteins), crystal farms (at 5 °C and 20 °C), two robotic liquid handling stations to prepare crystallization screens, and two automated CrystalDirect crystal harvesters with capacity to produce and store up to 400 frozen sample pins per operation cycle. Scientists send their samples to the facility by express courier, which are then processed by dedicated technicians at the HTX lab. Scientists can remotely design crystallization screening and optimization experiments through a web interface provided by the CRIMS system. Through this interface, they can choose from a wide range of parameters and experimental protocols available at the facility to fit their specific sample requirements. Results together with all experimental parameters are made available to users in real time through CRIMS. All samples received are assayed through a specifically developed method that enables to estimate the crystallization likelihood of the sample[16, 17]. Based on the results of this assay specific recommendations are made to users concerning the optimal incubation temperature and possible sample optimization experiments. Once crystallization experiments are set up, scientist can evaluate the results by looking at crystallization images collected at different time points

through the web. When crystals suitable for X-ray diffraction experiments are identified, scientists can use a dedicated interface to establish a crystal mounting plan that will then be executed by the CrystalDirect robot.

The CrystalDirect technology is based on the use of a modified vapor diffusion crystallization microplate and a laser beam to mount and cryo-cool crystal samples into diffraction compatible supports closing the automation gap existing between crystallization and data collection[19–21]. Briefly, crystals are grown in a modified vapour diffusion plate, the CrystalDirect microplate. Once crystals appear the CrystalDirect harvesting robot automatically applies a laser beam to excise a film piece containing the crystal, attach it to a standard diffraction data collection pin, and cryo-cool it in a nitrogen gas stream (see Zander *et al.* 2016 and https://www.youtube.com/watch?v=Nk2jQ5s7Xx8). This technology has a number of additional advantages over manual or semi-automated crystal mounting protocols. For example, the size and shape of the crystals is not an issue, making it equally easy to harvest large crystals or microcrystals, it is often possible to avoid the use of cryo-protectants, due to the special way in which the technology operates (see reference 6, Zander et al.), making X -ray diffraction analysis much more straightforward. The laser-beam can also be used as a surgical tool to select the best parts of a sample when crystals grow on clusters or show epitaxial growth for example. The CrytalDirect technology can also be used to automated soaking experiments[21] , the delivery of solutions with small molecules or other the chemicals to crystals. Thereby it enables to support fully automated, large-scale compound and fragment screening. Once Crystals are harvested and cryocooled by the CrystalDirect robot, they are transferred to either SPINE or Unipuck pucks which are compatible with most synchrony macromolecular crystallography beamlines around the world. The system can harvest up to 400 pins (the capacity of the cryogenic storage Dewar) in a fully autonomous manner. CRIMS communicates with the harvester robot during the process and provides automated tracking of crystal samples (pucks and pins). Pucks are marked with both barcodes and RFID tags to facilitate sample management[22].

CRIMS provides an application program interface (API) supporting automated communication with the ISPyB system supporting X-ray data collection management and processing at many synchrotrons in Europe and the world[23]. After automated crystal harvesting is completed, scientists can select crystal samples (pucks) and create sample shipments for the macromolecular crystallography beamlines at either the ESRF (Grenoble, France)[5, 6, 24] or Petra III synchrotrons (Hamburg, Germany)[25, 26]. CRIMS transfer the data corresponding to the selected beamline samples to the synchrotron information system along with pre-selected data collection parameters. Once the samples arrive at the selected synchrotron beamline, X-ray data collection is carried out either manually, through remote beamline operation or in a fully automated manner (i.e., at the MASSIF-1 beamline of the ESRF[6] operated by the joint EMBL ESRF Joint Structural Biology Group (JSBG)). After data collection CRIMS retrieves automatically information about the results of data collection along with initial data processing results carried out by the synchrotron data processing systems and presents it to the scientist through a convenient user interface.

The HTX lab applies these automated pipelines to support three different applications, rapid

determinations of new structures, rapid characterization of protein-ligand complexes and large-scale compound and fragment screening. Below we describe the how to use and operate them.

PROTOCOL:

NOTE: Funded access to these pipelines for scientists worldwide is supported through a series of funding programs. At the moment of writing this manuscript applications for access are accepted through either the iNEXT Discovery program (https://inext-discovery.eu), an European facility network to stimulate translational structural biology[27] funded by the Horizon 2020 programme of the European Commission or INSTRUCT-Eric (https://instruct-eric.eu/). Contact the corresponding author for the current modalities and routes for funded access at a particular time. This protocol describes operation of the protein-to-structure pipeline and includes steps common to all our pipelines while specificities for the other two pipelines are discussed in the following section. The instructions here refer to CRIMS V4.0.

1. High Throughput Crystallization Laboratory

1.1. Before starting, ask for registration at the HTX lab through the CRIMS system https://htxlab.embl.fr/#/. The user credentials provide remote access to all experimental design and evaluation interfaces.

1.1.1. Log in to CRIMS through a web navigator (Firefox, Chrome and Safari are supported). The CRIMs web server is encrypted to prevent third parties from accessing data while it travels through the web. Once in CRIMS, a series of menus to the left of the screening help manage and create samples, request crystallization experiments, manage and visualize plates, etc. A series of video tutorials are available at https://medias01-web.embl.de/Mediasite/Showcase/embl/Channel/a2168bcaa36b4564851663e5b69594014d.

NOTE: Users sending samples by courier need to register the samples and the requested crystallization experiments through CRIMS before sending the sample to ensure they can be processed without delay upon arrival. Please, send the shipment details to htx@embl.fr.

1.2. Log in into CRIMS (https://htxlab.embl.fr) with a web browser and click on the **Samples Menu**. This will open an interface with project and sample management tools.

1.3. Click on the **New Sample** button and provide the information requested. CRIMS allows to organize samples under different projects, targets and constructs. Assign the sample to existing ones or create new ones at this point.

1.4. Once the requested information has been entered click on **Save & Make Request**. Select the crystallization protocol, the crystallization screens to be used, the incubation temperature and the desired date for the experiments.

1.4.1. Use the comment fields to provide indications about the samples that are important for

the HTX lab operators to know. Custom screens can also be selected (see below). After submitting the crystallization request it will be validated by the HTX Lab Team and confirmation about the scheduling of the experiments will be sent via e-mail. Make sure to select experiment dates that are compatible with times required for sample shipment.

1.5. Once the samples arrive in the facility operators at the HTX lab will carry out the experiments as requested. Once crystallization experiments are set up, confirmation will be sent via e-mail and crystallization trays will be transferred to the automated imagers. CRIMS provides access to all experimental parameters and will automatically track new imaging sessions. E-mail notifications will be automatically sent when new images are available. An thermofluor-based[28] sample quality assessment experiment based on a protocol developed at the facility[16, 17] is carried out with every sample at this point and will be available through CRIMS.

1.6. Images of the crystallization experiments along with the results of the sample quality assessment will be available in CRIMS shortly after the crystallization trays are set up. Click on the **Thermofluor** menu and navigate to the sample to see the results of the sample quality assessment experiment[16].

1.6.1. Click on the **Plates** menu to see the images from the Crystallization plates. Navigate to the sample and either click on **View** to see the last imaging session or on the **+** (expand) symbol to select a different imaging session. A series of tools help to find and navigate easily through the samples. For example, clicking in a project box at the top of the screens filters samples for that project and search functions are available for most of the table columns.

1.7. Use the plate **View** interface to help evaluate and score the results of the crystallization experiments. It allows navigation through the different wells of the crystallization plates, select image types (i.e., Vis, UV), select image quality or record scores, for example. This interface also provides all experimental parameters used for the crystallization experiments including the composition of the crystallization solutions.

1.8. Click on the **Refinement** menu to design crystal optimization screens based on primary hit conditions identified through the initial screening. The **Chemicals** and **Stock Solutions** submenus allow one to register and manage the crystallization stock solutions. The **Screens** submenu provides access to an interface to design your own optimization or custom screens.

1.8.1. Select the plate type, stock solutions or gradient configurations that best fit the experimental design. It is possible to ask CRIMS to output a file directly compatible with the Formulator Robot (Formulatrix) to automatically pipette the screens into the plate or to output a printable document with the volumes for manual operation.

1.9. Iterate through steps 1.2-1.8 to carry out crystal optimization experiments.

1.10. Once crystals suitable for X-ray diffraction experiments are identified, navigate to the plate **View** interface and select the image corresponding to the right crystallization drop. Pre-

stored scores will help you do this easily.

1.10.1. Click either on **Crystal Harvesting** to record an automated crystal harvesting plan for the CrystalDirect harvester robot or on **Manual Harvesting** for traditional manual crystal mounting, if using CRIMS at a facility that is not equipped with CrystalDirect. Both interfaces will guide the user through the crystal harvesting process. CRIMS will automatically record and store the location of harvested crystals into either SPINE or Unipucks[29].

1.11. Select the **Crystal Manager** menu in CRIMS. Click on the **Harvested Crystals** submenu to inspect the frozen samples. When using the CrystalDirect harvester, images of the harvesting process are presented, including images of the pins with the harvested crystals.

1.12. Select the **Shipments** menu to connect to either ESRF or Petra III synchrotrons and create sample shipments for X-ray diffraction analysis. Click on the **Create Shipment** button and select the synchrotron you want to use and the bag number (the bag password at the synchrotron is necessary here). The next series of interfaces are used to select the pukcs to be included in the shipment. The system makes it possible to provide comments to support data collection and determine data collection parameters for automated beamlines like MASSIF-1.

1.13. If data collection is being carried out at ESRF or Petra III HTX lab, operators will transfer the samples to the beamline, data collection at other synchrotrons will be done at the user's own expense. It is possible to collect data by traveling to the synchrotron, through remote beamline operation or at MASSIF-1. In the latter case, the data collection process is fully automated. At the synchrotron, specific interfaces in ISPyB[23] allow users to recover the information sent by CRIMS and associate sample pucks to it so that results of data collection are automatically tracked. For the experiments described here, data collection at synchrotrons was typically carried out with the MXcube[30] software, while data processing and structure refinement was carried out with atuoPROC[31], Staraninso[32], BUSTER[33], Pipedream[31–33] and Coot[34].

1.14. Once data collection experiments have been carried out CRIMS retrieves summary information along with results of initial data processing at the synchrotron from the ISPyB[23] system. Go to CRIMS **Crystal Manager** menu and click on the **Crystal Diffraction Data** sub menu. All the information and metadata regarding diffraction data collection is available. It is also possible to download processed data from the synchrotron as well as raw diffraction images. View multiple data collections or select specific datasets. Sample management tools make it possible to navigate and select samples for specific projects constructs.

NOTE: This pipeline provides fully automated operation over the internet from pure protein to X-ray diffraction results and can be operated with one or multiple samples at the same time. It can be applied to different context and project types in structural biology.

REPRESENTATIVE RESULTS:

The automated crystallography pipeline described above has been applied to support a big number of internal and external projects with remarkable success. A few highlights include the

project from Djinović-Carugo and co-workers from the Max Perutz Laboratories (Vienna) focusing on the structural and functional analysis of a dipeptidyl peptidase essential for the growth a bacterial pathogen. The rapid succession of crystallization screening, diffraction evaluation, crystal optimization and X-ray data collection cycles (up to 8 iterations for this project) enabled to obtain structural models for three different conformational states of the protein in just a few weeks, which provided key mechanistic understanding on the function of this class of proteins[35] (see **Figure 1**).

Another example is the from Macias and co-workers from the Institute of Biomedical Research (IRB, Barcelona) that combined bioinformatics tools and structural approaches to identify new DNA binding motifs for the SMAD3 and SMAD4 transcription factors involved in cell fate regulation. This work has produced 6 high resolution structures of SMAD3 & 4 in complex with different DNA binding motifs[36, 37] revealing a so far unsuspected capacity of these transcription factors to recognize and bind to a diverse array of DNA sequences, which is key for the interpretation of their function in different biological contexts. These technologies have also been applied to support proprietary research in the context of drug design projects from research groups in pharma and biotech companies. For example, thanks to the rapidity contributed by these pipelines, the structural analysis of multiple ligand-target complexes can be achieved within days, which is of great value to support successive rounds medicinal chemistry optimization in the context of drug development. Finally we have also applied this infrastructure for large-scale X-ray based fragment screening[38].

FIGURE AND TABLE LEGENDS:

Figure 1: Automated Crystallography Pipelines. Integrated operation of the EMBL HTX lab including the CrystalDirect technology and the CRIMS software with the MASSIF-1 beamline at ESRF and automated communication between the CRIMS and ISPyB software enable to support fully automated, remote controlled protein-to-structure pipeline integrating crystallization screening and optimization, automated crystal harvesting and cryo-cooling and automated data collection and processing. The structural models correspond to three different conformational states of a protease from a pathogenic bacterium identified in a record time by applying these pipeline[39].

DISCUSSION:

The automated crystallography pipelines described here are available to researchers worldwide through different funding programs. Currently, funded access for crystallization experiments and the CrystalDirect technology can be obtained by applying to the iNEXT Discovery program and INSTRUCT-ERIC, while access to macromolecular crystallography beamlines at the ESRF is supported through the ESRF user access program. This approach minimizes the delay between crystal growth and measurement, accelerating the progression of very challenging projects that require diffraction-based optimization of protein production and crystallization conditions and frees scientists from complex operations associated with crystallization, crystal handling and beamline operation, rendering crystallography more accessible to non-expert groups. It can also be used for rapid exploration of crystallization additives, phasing agents or for compound screening through co-crystallization experiments. While most crystallography projects could

potentially benefit from this approach, some samples may require special protocols not amenable to automation or to the pipelines presented here, for example those requiring microfluidic systems or highly specialized crystallization devices or samples that are extremely labile and would not tolerate shipment.

The CrystalDirect technology also enables automated crystal soaking[21] for the characterization of of small molecule-target complexes. For this, a small aperture is created with the laser prior to the harvesting process and a drop of a solution containing the desired chemicals (i.e., phasing agents or potential ligands) is added on top, so that it enters in contact with, and diffuses into the crystallization solution eventually reaching the crystal. Chemical solutions can be formulated in water, DMSO or other organic solvents. After a certain incubation time the crystals can be harvested and analyzed by diffraction as described above. This approach has been applied to the rapid characterization of ligand-protein complexes in the context of structure-based drug design as well as to large-scale compound and fragment screening. In the latter case fragment libraries with hundreds to over a thousand fragments can be rapidly analyzed. Specific CRIMS interfaces not presented here facilitate the design and automated tracking of crystal soaking experiments, while integration between the CRIMS software and the Pipedream software suite, developed by Global Phasing Ltd (U.K) enable automated data processing, phasing, ligand identification and structure refinement over hundreds of datasets in parallel, streamlining data analysis and interpretation[31, 33]. For example, this pipeline was recently applied to the identification of fragments binding both to the active site and several allosteric sites of *Trypanosoma brucei* farnesyl pyrophosphate synthase, a key enzyme of the parasite causing human African trypanosomiasis.

The pipelines presented here can contribute to accelerate the pace of discovery in structural biology and make macromolecular crystallography more accessible to a larger number of research groups. Moreover, by facilitating large-scale compound and fragment screening they can contribute to foster translational research and speed up the process of drug discovery, contributing to facilitate the development of better and safer drugs against a larger number of targets.

ACKNOWLEDGMENTS:

We want to thank the joint EMBL-ESRF Structural Biology Group (JSBG) for support in the use and operation of the ESRF macromolecular beamlines. We are thankful to Matthew Bowler for support with data collection at the MASSIF-1 beamline of ESRF and Thomas Schneider and the EMBL Hamburg Team for excellent support with data collection at P14 of the PetraIII synchrotron (DESY, Hamburg, Germany). The CrystalDirect harvester is developed in collaboration with the Instrumentation Team at EMBL Grenoble. This project was supported by funding from the European CommunityH2020 Programme under the projects iNEXT (Grant No 653706) and iNEXT Discovery (Grant No 871037) as well as the Région Auvergne-Rhône-Alpes through the Booster programme.

DISCLOSURES:

JAM is co-author of a patent on the CrystalDirect system

REFERENCES:

Mariaule, V., Dupeux, F., Márquez, J.A. Estimation of Crystallization Likelihood Through a Fluorimetric Thermal Stability Assay. *Structural Genomics*. **1091**, 189–195, doi: 10.1007/978-1-62703-691-7_13 (2014).

2. Dupeux, F., Röwer, M., Seroul, G., Blot, D., Márquez, J.A. A thermal stability assay can help to estimate the crystallization likelihood of biological samples. *Acta Crystallographica Section D Biological Crystallography*. **67** (11), 915–919, doi: 10.1107/S0907444911036225 (2011).

3. Dimasi, N., Flot, D., Dupeux, F., Márquez, J.A. Expression, crystallization and X-ray data collection from microcrystals of the extracellular domain of the human inhibitory receptor expressed on myeloid cells IREM-1. *Acta Crystallographica Section F Structural Biology and Crystallization Communications*. **63** (3), 204–208, doi: 10.1107/S1744309107004903 (2007).

4. Cipriani, F., Röwer, M., Landret, C., Zander, U., Felisaz, F., Márquez, J.A. CrystalDirect: a new method for automated crystal harvesting based on laser-induced photoablation of thin films. *Acta Crystallographica Section D: Biological Crystallography*. **68** (10), 1393–1399, doi: 10.1107/S0907444912031459 (2012).

5. Márquez, J.A., Cipriani, F. CrystalDirect^{TM}: A Novel Approach for Automated Crystal Harvesting Based on Photoablation of Thin Films. *Structural Genomics*. **1091**, 197–203, doi: 10.1007/978-1-62703-691-7_14 (2014).

6. Zander, U. *et al.* Automated harvesting and processing of protein crystals through laser photoablation. *Acta Crystallographica Section D Structural Biology*. **72** (4), 454–466, doi: 10.1107/S2059798316000954 (2016).

7. Hiraki, M., Matsugaki, N., Yamada, Y., Hikita, M., Yamanaka, M., Senda, T. RFID tag system for sample tracking at structural biology beamlines. 060074, doi: 10.1063/1.5084705 (2019).

8. Delageniere, S. *et al.* ISPyB: an information management system for synchrotron macromolecular crystallography. *Bioinformatics*. **27** (22), 3186–3192, doi: 10.1093/bioinformatics/btr535 (2011).

9. McCarthy, A.A. *et al.* ID30B – a versatile beamline for macromolecular crystallography experiments at the ESRF. *Journal of Synchrotron Radiation*. **25** (4), 1249–1260, doi: 10.1107/S1600577518007166 (2018).

10. 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, doi: 10.1107/S1600577515016604 (2015).

11. von Stetten, D. *et al.* ID30A-3 (MASSIF-3) – a beamline for macromolecular crystallography at the ESRF with a small intense beam. *Journal of Synchrotron Radiation*. **27** (3), 844–851, doi: 10.1107/S1600577520004002 (2020).

12. Cianci, M. *et al.* P13, the EMBL macromolecular crystallography beamline at the low-emittance PETRA III ring for high- and low-energy phasing with variable beam focusing. *Journal of Synchrotron Radiation*. **24** (1), 323–332, doi: 10.1107/S1600577516016465 (2017).

13. Gati, C. *et al.* Serial crystallography on *in vivo* grown microcrystals using synchrotron radiation. *IUCrJ*. **1** (2), 87–94, doi: 10.1107/S2052252513033939 (2014).

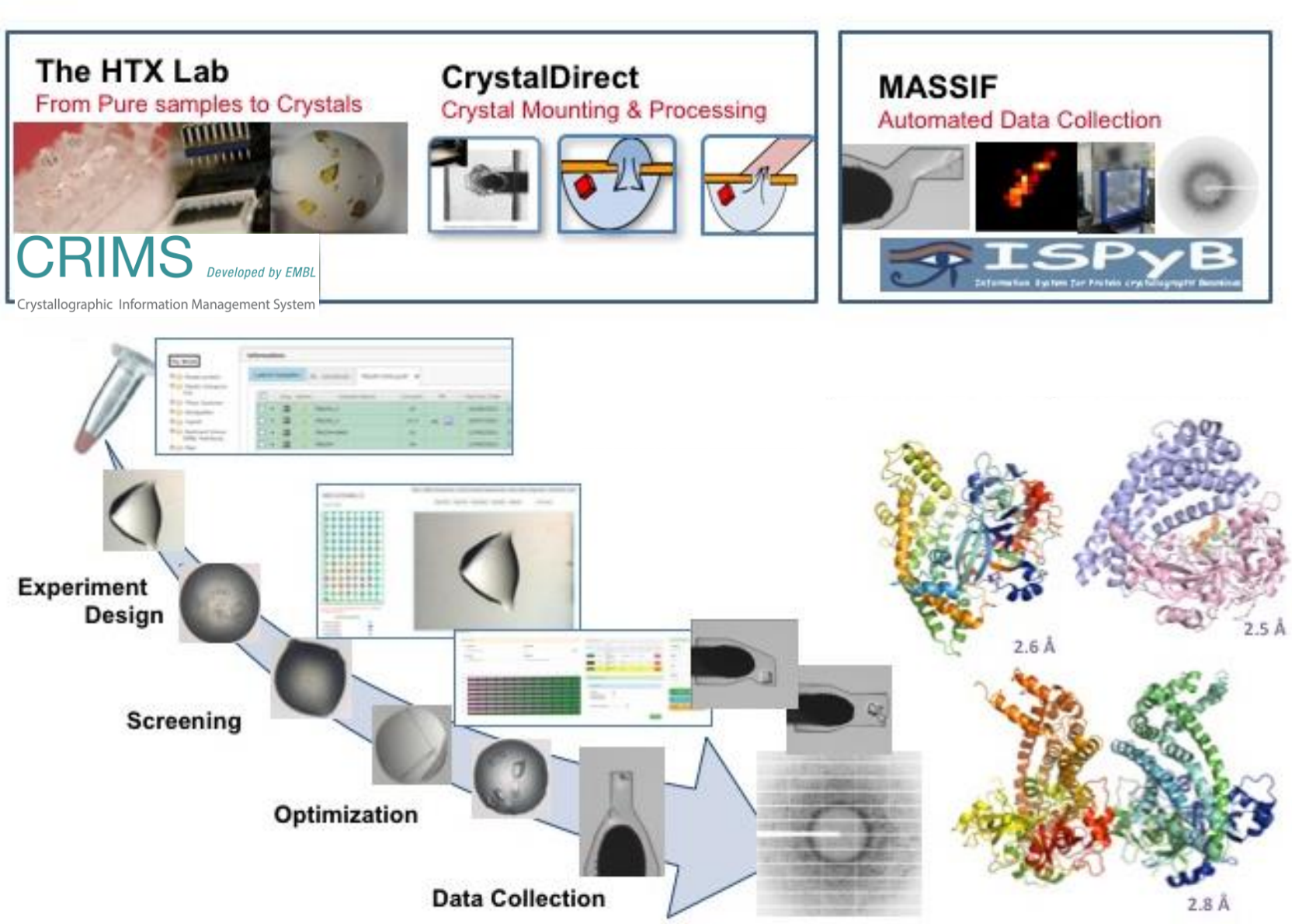
14. iNEXT Consortium iNEXT: a European facility network to stimulate translational structural

biology. *FEBS Letters*. **592** (12), 1909–1917, doi: 10.1002/1873-3468.13062 (2018).
15. Papp, G. *et al.* Towards a compact and precise sample holder for macromolecular crystallography. *Acta Crystallographica Section D Structural Biology*. **73** (10), 829–840, doi: 10.1107/S2059798317013742 (2017).
16. Bezerra, G.A. *et al.* Bacterial protease uses distinct thermodynamic signatures for substrate recognition. *Scientific Reports*. **7** (1), 2848, doi: 10.1038/s41598-017-03220-y (2017).
17. Martin-Malpartida, P. *et al.* Structural basis for genome wide recognition of 5-bp GC motifs by SMAD transcription factors. *Nature Communications*. **8** (1), 2070, doi: 10.1038/s41467-017-02054-6 (2017).
18. Aragón, E. *et al.* Structural basis for distinct roles of SMAD2 and SMAD3 in FOXH1 pioneer-directed TGF-β signaling. *Genes & Development*. **33** (21–22), 1506–1524, doi: 10.1101/gad.330837.119 (2019).
19. Münzker, L. *et al.* Fragment-Based Discovery of Non-bisphosphonate Binders of *Trypanosoma brucei* Farnesyl Pyrophosphate Synthase. *ChemBioChem*. **21** (21), 3096–3111, doi: 10.1002/cbic.202000246 (2020).
20. Bezerra, G.A. *et al.* Bacterial protease uses distinct thermodynamic signatures for substrate recognition. *Scientific Reports*. **7** (1), 2848, doi: 10.1038/s41598-017-03220-y (2017).
21. Smart, O.S. *et al.* Exploiting structure similarity in refinement: automated NCS and target-structure restraints in *BUSTER*. *Acta Crystallographica Section D Biological Crystallography*. **68** (4), 368–380, doi: 10.1107/S0907444911056058 (2012).
22. Vonrhein, C. *et al.* Data processing and analysis with the *autoPROC* toolbox. *Acta Crystallographica Section D Biological Crystallography*. **67** (4), 293–302, doi: 10.1107/S0907444911007773 (2011).

Figure 1

Click here to access/download;Figure;Marquez_Fig-1.pdf

Automated Crystallography Pipelines



Name of Material/ Equipment	Company	Catalog Number	Comments/Description
CrystalDirect harvester	Arinax		Automated crystal mounting and cryocooling
CrystalDirect Crystallization plate	Mitegen	SKU: M-XDIR-96-2	96-well crytsallization microplate
Formulator 16	Formulatrix		For the autoamted preparation of crystallization screens
Mosquito crystallization Robot	SPT Labtech		For the preparation of crystallization experiments
Tecan Evo Liquid handling station	Tecan		For the preparation of crystallization solutions
Spine Pucks	Mitegen	SKU: M-SP-SC3-1	SPINE-compatible cryogenic pucks for automated synchrotr
UniPucks	Mitegen	SKU: M-CP-111-021	Universal cryogenic pucks for automated synchrotron samp

·on sample exchangers
ıle exchangers

Dear Editor,

Please, find below a point by point answer to reviewers and editorial suggestions along with a description of changes to the MS. All changes described can be identified through track changes in the revised version of the MS.

We want to thank you and the referees for their suggestions, which we are convinced are a positive contribution the MS.

Best regards

Jose A. Marquez

Poin by point anwers:

Editorial comments:
Editorial Changes

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. E.g. line 155: "existing" instead of "exiting", etc.

The MS has been proofread, typo in line 155 and others have been corrected (see track changes

2. Please provide an email address for each author.

Corrected

3. Please check the symbols used in line 62.

Corrected

4. 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."

Corrections addressing these points have been introduced

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

Corrections addressing these points have been introduced

6. To provide a more balanced perspective, consider including the possible limitations of the approach, and compare it with other available approaches in the discussion.

Reference to potential limitations have been included in the discussion section. Comparison to traditional approaches was already included in the orignal text.

7. Please sort the Materials Table alphabetically by the name of the material.

Corrected

Reviewers' comments:

Reviewer #1:

In order to make the manuscript acceptable, it requires an introduction on automated structure determination which needs to make proper reference to systems developed elsewhere.

I should start by noting that the other reviewers disagree on this point with Reviewer 1. Nevertheless, in the revised version we have introduced a section at the start of the introduction with 15 citations to works on automated structure determination by other groups.

The second biggest concern I have is that the manuscript is written in the form of a manual. This is by itself not a problem, but there is very little if anything mentioned of the underlying technology. Which is the program that collects the data (reference?), which is the progam, which processes the data (reference?), which programs are used in the structure determination (reference?) and why are these used.

Description of the software with appropriate references have been introduced in line 453.

Along the same lines, it would be good to comment on the choice of programs, to compare the results to other pipelines or other efforts made elsewhere.

A comparison of different crystallographic data processing and refinement pipelines is clearly outside the scope of this work and would require a considerable amount of effort. This simply is not a reasonable request.

A third major issue is, what happens with the data at the end of the process? Will the user get access to all data including raw diffraction images and associated metadata? What is the timeline for such access?

This is already clearly stated on point 13 of the protocol. However, to clarify this further we have added as specific reference to metadata and raw diffraction images has been introduced. See line 460

Finally, the authors mention two projects, to which their pipeline was employed. What are the results for these two projects? Which fragments were soaked? What is the hit rate? How does this compare to other pipeline used elsewhere? Is is better Is it worse? Why? Etc. etc. etc.

Three examples from users of these pipelines are cited in the representative results section. Appropriate references describing the experiments performed, samples and ligands used and the results obtained are provided (see Representative results section and refs 36, 37, 38, and 39. These works speak for the quality of results and the pipelines presented

Again, while a comparison of results with other technologies would be an interesting subject it would require careful experimental design and is clearly outside the scope of this work. The goal of this MS is to present the capacities of the facility in Grenoble, along with those of other facilities in the series volume. We have not used other facilities and would prefer not to enter into subjective evaluations to compare one or the other. Readers and users would be able to draw their own conclusions.

Instead of the overview figure 1, I would prefer to see screenshots of the different views, the user sees at every step of the process. I understand that some of this will be provided by the video, but nonetheless I think the protocol would benefit from properly described screen shots.

Indeed, this will be presented in the video. Presenting screenshots at every step of the process would take considerable space and even make the text difficult to read. Fortunately, JoVE is unique in providing a video publication format that perfectly addresses this problem.

The materials table seems to be incomplete. Pucks? Sample holders? Crystallization plates? Screens? Etc. etc.

Materials have been added to the Materials table

Specific questions concerning the protocol:
1. How does one obtain a user name? Contact the authors? Please specify.

This is clearly *stated, see line 157*

2. What is a sample? A protein? A construct? A piece of DNA? Please specify.

This is defined by the user not by the facility. We believe this is quite obvious and does not need to be stated.

3. Which crystallization protocols and which screens are available? Please specify. How long dows it take until feedback from HTZ staff is provided to the user?

Available crystallization screens are displayed through the corresponding interface through which user can use the ones they prefer. The collection is very extensive and including them in the MS text would lead to a very long list, which doesn't seem necessary

As described in the text already experiments can be scheduled by the user so they can see when they will executed. Moreover, All results and experimental parameters are accessible in rea time, as clearly stated in line 88

4. What is the meaning of the quality assessment? Does it has any influence on the following steps? Will the sample be returned to the user, if the quality is not sufficient?

The Thermofluor based protocol for sample quality assessment is described in detail in reference 17. This also includes guidance for the interpretation of the results and possible actions to follow in the different cases are also described there.

5. Is there a way to check whether crystals are salt crystals or protein crystals?

As clearly stated in the text, the system enable UV imaging, but as it is well know, diffraction experiments are sometimes the definitive way to establish this point in a definitive way. These are well-know facts and do not represent novel aspects in relation to this MS

6. Is there a way to ask for room temperature diffraction experiments?

As stated in the line 450, "*It is possible to collect data by traveling to the synchrotron, through remote beamline operation or at MASSIF-1"* Therefore all diffraction experiments avaialble to these modalities are possible, including room temperature data collection. We believe this is rather obvious and does not need furthey clarificaiton

Minor Concerns:
Line 85: what is a standard diffraction data collection pine?
pin.

The typo has been corrected

Lines 85/86: a liquid nitrogen gas stream? hmmm. Liquid or gaseous? Please re-word.

Liquid has been deleted

Line 129: ISTRUCT --> INSTRUCT

Typo corrected

Line 178: Thermofluor should be spelled out as differential scanning fluorimetry and properly referenced.

While using the term DSF would be possible, the use of the term thermofluor is well established in the field, see for example doi: 10.1016/j.ab.2006.07.027. This reference has been added to the text. Line 193

Reviewer #2:

Manuscript Summary:

The manuscript provides an overview of how to use the CRIMS automated system/pipeline for remote high-throughput crystallization. The CRIMS system include HT screening, cryoprotection, ligand screening and crystal harvesting for structural characterization, either through shipment to beamlines at ESRF or Petra III. Data collection and experimental visualization can be fully automated, or scheduled for when a user has allocated beamtime. The system provides full documentation and user interaction to tailor crystallization experiments, collate data etc. The CRIMS automated pipelines described in this paper will be of use to both expert crystallography labs as well as non-expert users of structural biology resources.

The manuscript is well written, readily followed and is properly referenced etc. A very nice manuscript describing a very interesting and supportive crystallization pipeline for the structural biology community.

Reviewer #3:

Manuscript Summary:

This manuscript on the new high throughput method for macromolecular crystallography at the HTX facility presents a well-designed automated crystallography pipeline combining the CrystalDirect technology and the CRIMS system to provide a fully automated pipelines from the crystal growth, optimization, harvesting, towards the diffraction data collection. Every step in this pipeline can be requested and operated remotely by the end users. The authors provided a comprehensive rationale and background information for this automated crystallography pipeline, with the citation of the well-established CrystalDirect technology, and the protocol details of CRIMS system registration and operation step-by-step. The entire automated crystallography pipeline would accelerate the pace to the field of large-scale compound and fragment screening experiments, shortening the time gap between crystal growth in house and crystal diffraction measurement. However, the author doesn't discuss the limitations of this pipeline technique. For example, how to characterize if the sample properties change during the shipping? Though it is mentioned in the protocol section, step 4, that a thermofluor-based sample quality assessment experiment will be performed with every sample, some protein of interest requires the specific conformational states, which may be altered during the shipping. In this case, more assays are required to assess the protein properties. This pipeline development is just combining the existing instruments and experimental setups but does not show the high impact for the macromolecular crystallography field. It is more suitable for the application of an automation pipeline patent.

We, as the other two reviewers clearly disagree. As revier 2 states and we discuss in the MS, "*the CRIMS automated pipelines described in this paper will be of use to both expert crystallography labs as well as non-expert users of structural biology resources."* This in itself is of considerable impact, not to mention applications in drug design. Examples of the use technology and its impact are given, see refs 36, 37, 38 and 39.

Major Concerns:

The automated crystal harvesting equipped with a laser beam harvester is very interesting. However, nowadays, high-throughput crystallization screening and optimization is also in-house accessible in most crystallography laboratories worldwide. Optimization of protein crystallization is more convenient to be operated in the research laboratories, especially for the proteins of interest that are environmental sensitive. In-house crystallization screening and optimization is more time and efforts saving and can be carried out right after the primary hit conditions identified, without more sample shipping to the facility.

We find this statement is rather subjective and no hard evidence supporting it is provided here either. Optimization experiments done at the remote facility can be as fast as those done in house as communication with CRIMS is instantaneous. Moreover, evaluation of diffraction power is key for crystal optimization. We present how evaluation of diffraction properties can be done very rapidly given that the HTX facility and the Synchrotron are adjacent, while if crystals are grown in the home lab, weeks may be necessary before the next synchrotron experiments can be done. Not to count the other advantages of the system like automated harvesting.

I believe a few years back many researches might have felt the same as the reviewer concerning DNA sequencing, but now sequencing is seldom done in the local lab and samples are systematically shipped to an external facility. Similarly, many crystallographers thought that automated diffraction data collection would never produce the same data quality as manual beamline operation, but the MASSIF1 beamline has shown that in many cases automated data collection produces better results that manual data collection.

It is true that some samples may require special protocols that are not compatible with automation or be so labile that time of transfer to the facility is prohibitive, but those situations tend to be rare. Nevertheless, this point has been clarified in line 547.

The CRIMS system is almost same as other automated crystallization robot equipped with imaging instrument. Step 1-8 in the protocol is same as a standard protocol, which has been well established in most crystallography research labs.

We disagree imaging instruments do not include automated crystal harvesting, and automated communication with synchrotrons. These are unique and very substantial features.

The other concern is the short time consumption as mentioned in the representative results. Is the crystal data collection beamtime at ESRF or PETRA III guaranteed upon the registration of the automated crystallography pipeline at the HTX lab? If the answer is no, the time gap between the crystal growth and data collection is still a remaining question upon user's own application.

The answer is Yes. Through the iNEXT Discovery program Synchrotron beamtime at ESRF and PETRA III for diffraction measurements is also included as part of the same application so no additional delays are needed. As stated in line 152 of the MS we recommend potential users to contact the corresponding author for information on the current modalities and routes for funded access at a particular time.

Minor Concerns:
In the introduction section, line 61, it is mentioned that the HTX lab is equipped with an LCP module crystallization robot, but no further experimental protocol is stated about how to make it compatible with LCP crystallization set up and with the CrystalDirect technology to harvest crystals from LCP plates? If this has not been development yet, please delete the statement of LCP module in the manuscript without confusing audience that they can operate LCP crystallization and harvesting in the facility.

Yes, we can operate with LCP but covering this aspect here would to excessively lengthy MS. A separate MS describing this is currently in preparation.

Reviewer #4:
Manuscript Summary:
The manuscript describes the automatic pipelines for macromolecular crystallography that offers a remote procedure (over internet) to go from sample to protein structure. This can be done combining 1) the HTX (high throughput crystallization) lab in EMBL (Grenoble), to perform crystallography experiments, optimization and crystal harvesing, 2) the Crystal design technology, for crystal mounting and cryocooling and 3) the CRISM (crystallographic Information Management system) software that enables all communications from crystal platform, harvesting, freezing, shipping, data collection and processing. This procedure is set to determine new structures, rapid crystal-ligand complexes characterization and large scale automated compound and fragment screening using libraries with about 1000 fragments. This facility can be applied through inext-discovery (not clear if through instruct also). The equipment used in the process is well described as well as the different steps to perform the experiment. Similarly, the work is well referenced. Some examples of works performed with this protocol have been shown. The presentation is clear. This facility and description is very useful for crystallographers around world (and maybe non-crystallographers) since its use could have an impact on reducing effectively the time of an structural biology project. It constitutes a new approach for many crystallographers especially useful to find new compounds and with applications in drug design field. The HTX operates since 2003 and has given support to hundreds of scientist. The facility also offers quality control of the sample at arrival to determine the crystallization likelihood of the sample, what is very useful.

Minor Concerns:
1. At point 4 of the protocol, lane 172, it states that "A thermofluor-based sample quality experiment is carried out with every sample….."
It is not clear if this is the first thing to be done, before the crystallization experiments. If so, probably this should be moved at the beginning of point 4 or even before.

This is done at the same time as the crystallization experiments so its location in the protocol is adequate

2. Could the author discuss why the way of mounting crystals and cryo-cooling does not need the use of cryoprotectant in most of the cases?

These aspects are discussed in the original publication describing the CrystalDirect technology. The reference is included. See line 112

3. To make ligand soaking experiments, a drop with chemical is diffused into the crystal drop. Which solution is used for that? Is the compound dissolved in precipitant solution, water, DMSO? Probably this should be also indicated.

Water, DMSO or other organic solvents can be used. This has been clarified in line 554

4. The way of crystal mounting is not easy to visualize from the text explanations. Certainly, in the video this will be more clear.

Yes, the video should show this aspect in detail which is otherwise difficult to exemplify in a single figure

5. At the beginning, it is said that access to this pipeline is able through inext-discovery and instruct programs, but the discussion states that only can be accessed by inext-discovery and ESRF funded programs.

This has been corrected, see line 564

6. Could this facility help in co-crystallization of proteins with compounds? or, to search crystal conditions in presence of ligands, additives etc? As this is not the aim of this pipeline, I wonder if it could be also a good application.

Yes, certainly and users do this regularly at the facility. A phrase has been introduced in the new version of the MS clarifying this point See lines 569 to 571.