Journal of Visualized Experiments

Fully autonomous characterization and data collection from crystals of biological macromolecules

--Manuscript Draft--

Article Type:	Invited Methods Article - JoVE Produced Video
Manuscript Number:	JoVE59032R2
Full Title:	Fully autonomous characterization and data collection from crystals of biological macromolecules
Keywords:	Fully automatic beamline, Massively Automated Sample Selection Integrated Facility (MASSIF-1), macromolecular crystallization, diffraction data collection and analysis, undulator beamline, robotic sample changer
Corresponding Author:	Stephanie Hutin, PhD ESRF Grenoble Cedex 9, France, Isere FRANCE
Corresponding Author's Institution:	ESRF
Corresponding Author E-Mail:	stephanie.hutin@esrf.fr
Order of Authors:	Stephanie Hutin, PhD
	Bart Van Laer
	Christoph Mueller-Dieckmann
	Gordon Leonard
	Didier Nurizzo
	Matthew Bowler
Additional Information:	
Question	**Response**
Please indicate whether this article will be Standard Access or Open Access.	Open Access (US$4,200)
Please indicate the **city, state/province, and country** where this article will be **filmed**. Please do not use abbreviations.	Grenoble, Isere, France

TITLE:

Fully Autonomous Characterization and Data Collection from Crystals of Biological Macromolecules

AUTHORS AND AFFILIATIONS:

Stephanie Hutin[1] *, Bart Van Laer[1] *, Christoph Mueller-Dieckmann[1], Gordon Leonard[1], Didier Nurizzo[1], Matthew W. Bowler[2]

[1]European Synchrotron Radiation Facility, Structural Biology Group, Grenoble, France
[2]European Molecular Biology Laboratory, Grenoble Outstation, Grenoble, France

*These authors contributed equally to this work.

Corresponding Author:

Matthew W. Bowler (mbowler@embl.fr)
European Molecular Biology Laboratory
Grenoble Outstation
France

Email Addresses of Co-authors:

Stephanie Hutin (stephanie.hutin@esrf.fr)
Bart Van Laer (bart.van-laer@esrf.fr)
Christoph Mueller-Dieckmann (muellerd@esrf.fr)
Gordon Leonard (leonard@esrf.fr)
Didier Nurizzo (didier.nurizzo@esrf.fr)

KEYWORDS:

Fully automatic beamline, Massively Automated Sample Selection Integrated Facility (MASSIF-1), macromolecular crystallization, diffraction data collection and analysis, undulator beamline, robotic sample changer

SUMMARY:

Here, we describe how to use the automated screening and data collection options available at some synchrotron beamlines. Scientists send cryocooled samples to the synchrotron, and the diffraction properties are screened, the data sets are collected and processed and, where possible, a structure solution is carried out—all without human intervention.

ABSTRACT:

High-brilliance X-ray beams coupled with automation has led to the use of synchrotron-based macromolecular X-ray crystallography (MX) beamlines for even the most challenging projects in structural biology. However, most facilities still require the presence of a scientist on site to perform the experiments. A new generation of automated beamlines dedicated to the fully automatic characterization of, and data collection from, crystals of biological macromolecules has recently been developed. These beamlines represent a new tool for structural biologists to

screen the results of initial crystallization trials and/or the collection of large numbers of diffraction data sets, without users having to control the beamline themselves. Here we show how to set up an experiment for automatic screening and data collection, how an experiment is performed at the beamline, how the resulting data sets are processed, and how, when possible, the crystal structure of the biological macromolecule is solved.

INTRODUCTION:

Determining the three-dimensional structure of specific proteins is crucial in biology. The information that is derived from doing so sheds light on the biological function and on the shape and specificity of active and/or binding sites contained in the molecule under study. In many cases, this allows mechanisms of action to be determined or, where appropriate, potential therapeutic molecules to be developed. MX is the technique most commonly used to obtain structural information, but a bottleneck is the determination of the optimal conditions to obtain well-diffracting crystals. Therefore, crystallization trials are carried out in numerous different conditions and are then screened, to find the best crystals to be used for diffraction data collection. The automation of the setup of crystallization trials[1] has clearly helped in this regard. However, the subsequent steps (i.e., crystal mounting, diffraction screening, and diffraction data collection) are usually carried out manually, taking up a lot of time, effort, and resources. The automation of diffraction screening and data collection would, therefore, mean an enormous gain in time and efficiency.

Diffraction screening and data collection in MX is most often carried out at synchrotron MX beamlines at which automation has largely facilitated this process. However, in most cases, it is necessary for the scientist to be present at the beamline during an experiment or to operate it remotely. Recently, a new generation of completely automated MX beamlines has been developed[2]. Here, users do not need to be present, either physically or remotely, during an experimental session. This allows scientists to spend more time on less routine tasks, rather than spending entire days, and often nights, screening crystals and collecting diffraction data. The world's first fully automated beamline is the Massively Automated Sample Selection Integrated Facility (MASSIF-1, ID30A-1)[2,3] at the European Synchrotron Radiation Facility (ESRF). It has a unique sample environment in which a high-capacity sample-containing dewar operates in tandem with a robotic sample changer that also acts as the beamline's goniometer[4,5]. MASSIF-1 is an undulator beamline equipped with a single-photon-counting hybrid pixel detector[6], that operates at a fixed wavelength of 0.969 Å (12.84 keV) with an intense X-ray beam (2 x 10^{12} photons/s). The beam size at the sample position can be adjusted between a minimum of 10 µm (round beam) to a maximum of 100 µm x 65 µm (horizontal by vertical beam size). On average, the beamline can process, in a completely automatic fashion (see below), 120 crystals in 24 h. The operation of the beamline is based on a series of workflows[7], each of which takes intelligent decisions based on the outcome of previous steps in the workflow, to ensure the measurement of the best possible data from the sample under study. In particular, the evaluation of the diffraction characteristics of an individual sample takes into account crystal volume and flux and ensures, where the crystal is larger than the X-ray beam, that only the best region of the crystal is used for subsequent data collection. Diffraction data sets are, thus, optimized for maximum resolution with minimized radiation damage[2,3]. Demanding data collection protocols, such as

pseudo-helical (multi-position) data collection strategies for both native and single-wavelength anomalous diffraction (SAD) data collection, are also available[8].

Completely automatic experiments at MASSIF-1 involve cryocooling and mounting the crystals on a magnetic sample mount suitable for the desired beamline equipment standard pins SPINE[9], entering the desired experimental parameters in the 'diffraction plan' table in the Integrated System for Protein Crystallography beamlines (ISPyB)[10], a web-based information management system for MX experiments, and sending the samples to the beamline. At the ESRF, all costs of the transport of the samples to/from the beamline are supported by the ESRF User Office (see the website of the ESRF[11] for details). At MASSIF-1, no restrictions are placed on the loop size or crystal quality. When choosing a diffraction plan for a given crystal, the user can either use default settings or choose from specific workflows, which can be customized for each sample. Several preprogrammed workflows are available. In the **MXPressE**[3] workflow, the sample-containing loop is first aligned to the sample position using optical centering. Then, X-ray-based centering ensures that the best region of the crystal is centered to the X-ray beam. Data collection strategies are then calculated using eEDNA, a framework for developing plugin-based applications especially for online data analysis in the X-ray experiments field, taking into account crystal volume and the real-time flux at the beamline. Following the collection of a full diffraction data set, this is then processed using a series of automatic data processing pipelines[12] and the results are made available for inspection and download in ISPyB. The **MXPressE SAD**[3] workflow is aimed at selenomethionine-containing crystals of the target protein and exploits the fact that the operating energy of MASSIF-1 is just above the Se K edge. Here, the MXPressE eEDNA data collection strategy is optimized for SAD data collection (i.e., high redundancy, and with the resolution set to where the R_{merge} between Bijvoet pairs is below 5%). To screen the diffraction properties of a series of crystals without subsequent data collection, the **MXScore**[3] workflow can be used to produce a full quality assessment of the crystals analyzed. In the **MXPressI**[3] workflow, 180° of rotation data are collected using 0.2° oscillations and using the starting phi angle and the resolution determined by an eEDNA strategy. **MXPressO**[3] includes a preobserved resolution into the workflow (default: d_{min} = 2 Å). To make an initial assessment of the crystals resulting from a crystallization trial, the **MXPressM**[3] workflow is offered. This performs a high-dose mesh scan over the widest orientation of sample support with no data collection or centering. Recently, two new experiment workflows, **MXPressP** and **MXPressP_SAD**, which perform pseudohelical data collections, have been implemented[8]. The execution of all steps in all workflows can be followed online and in real-time by the user, via ISPyB.

Here we show how to prepare a fully automated MX experiment at MASSIF-1 and how to retrieve and analyze the data resulting from the experiment. As an example, we use human mitochondrial glycine cleavage system protein H (GCSH). This lipoic acid-containing protein is part of the glycine cleavage system responsible for the degradation of glycine. This system further includes the P protein, a pyridoxal phosphate-dependent glycine decarboxylase, the T protein, a tetrahydrofolate-requiring enzyme, and the L protein, a lipoamide dehydrogenase. GCSH transfers the methylamine group of glycine from the P protein to the T protein. Defects in the H protein are the cause of nonketotic hyperglycinemia (NKH) in humans[13].

PROTOCOL:

NOTE: The production, purification, and crystallization of GCSH are described in **Supplementary File 1**.

1. Brief description of the offline preparation and crystal mounting

1.1. Position a nylon loop or another crystal mounting support already fixed to a SPINE pin under one or more crystals and lift them out of the precipitation solution (20 µL of 0.5 M sodium formate pH 4.0 + 25 µL of protein solution).

1.1.1. Remove the bulk liquid around the crystal(s) by touching the mount with a paper wick to suck off any excess liquid.

1.2. Soak the crystal(s) in the cryoprotective solution containing the precipitation solution plus 30% glycerol; then, remove both the crystal support and the crystal(s).

1.2.1. Remove the bulk liquid around the crystal(s) by touching the mount with a paper wick to suck off any excess liquid.

1.3. Plunge the mount into a SPINE vial filled with liquid nitrogen and store it, along with any other crystals similarly prepared, in a European Molecular Biology Laboratory (EMBL)/ESRF sample changer puck[9] at liquid nitrogen temperature.

NOTE: The crystal(s) are stable in this condition until the beamtime is available.

2. Requesting beamtime on MASSIF-1

2.1. Request beamtime as early as possible on the ESRF homepage (at http://www.esrf.eu/UsersAndScience/UserGuide/Applying).

NOTE: There are a number of possible modes of access to the ESRF MX beamlines. Laboratories can apply collectively as part of a Block Allocation Group (BAG), to have beamtime allocated for 2 years. If groups wish to apply individually, they can apply for rolling access, which allows them rapid access to the beamlines after peer review. The group's proposal will be reviewed and cleared by the ESRF Safety Group who may request additional details. If the proposal is accepted, an experiment number and password will be communicated. Proprietary research can be performed by purchasing beamtime.

2.2. Complete the required safety training online (at http://www.esrf.eu/UsersAndScience/UserGuide/Preparing/SafetyTraining).

2.3. Book beamtime on the MASSIF-1 Calendar.

NOTE: It is possible to book up to a maximum of 50 sample holders to be analyzed per shift.

2.4. Fill in the **A-form** to declare a mail-in experiment (http://www.esrf.eu/UsersAndScience/UserGuide/Preparing/new-a-form), along with the required safety information, for the samples that are to be measured.

3. Creation of a diffraction plan in ISPyB

NOTE: The diffraction plan holds all the information needed for a sample in ISPyB and can contain additional information to tailor the experiment performed for each sample.

3.1. Open ISPyB (at https://exi.esrf.fr/).

3.2. Choose **MX experiments**.

3.3. Log in with the experiment number and password.

3.4. Click on **Shipment** | **Add new** and provide the necessary information. Click **Save**.

3.5. Click **Add parcel** and fill in the information requested. Click **Save**.

3.6. Then, click **Add container**, give the puck barcode as the name, and choose **SPINE puck**. Click **Save**.

3.7. Click on the container symbol and **Edit**, and fill in the necessary information, like protein name, preferred workflow, crystal position in the puck, etc., concerning the samples.

3.8. Choose the protein (for example, GCSH or lysozyme) that has been approved by the ESRF safety group.

3.9. Enter a unique sample name to identify each individual sample. It is possible to optionally scan the pin barcode. The rest of the information below is optional.

3.10. Enter the optional information.

3.10.1. For each individual sample, enter the experiment type (i.e., MXPressE_SAD, SCORE, or MXPressO, etc., default MXPressE) under **Exp. Type**. This defines which automatic workflow will be used to process each crystal. Given that the GCSH crystals are needles, choose **MXPressP**.

3.10.2. Enter a space group (for example, P1, C2, or $P2_12_12_1$), if known. If present, this will be used for data collection strategy calculations and by the automatic data processing pipelines available.

3.10.3. Enter the desired resolution (default: d_{min} = 2.0 Å). This defines the crystal-to-detector distance for the initial mesh scans, characterization, and default data collection.

3.10.4. Set the desired threshold resolution (for example, 1.5 Å or 2.3Å), to prevent the collection of full datasets from crystals which do not diffract to this limit. This can save data storage space and analysis time.

3.10.5. Set the required completeness (default: 0.99). Set the required multiplicity (default: 4). If more than one crystal is contained on the sample support, set the maximum number of crystals to be analyzed. The default value is 1 or 5 for MXPressP.

3.10.6. Select the appropriate beam size (default: 50 µm). If a specific value is not selected, the X-ray-centering and data collection strategy calculations will be performed with a beam size of 50 µm.

NOTE: During any subsequent collection of complete data sets, the beam size will be adapted automatically.

3.10.7. Put in the space group, if known, in the forced space group column. Set the radiation-sensitivity of the crystals (0.5–2.0 for low to high sensitivity, with a default value of 1).

3.11. If desired, set the total rotation angle to be collected for the full dataset collection (default: the total rotation angle determined by eEDNA).

3.12. Save the values. Click on **Return to Shipments**. Press **Send Shipment to ESRF**.

3.13. Print the shipping label and send the samples. Users should arrange a pickup with a courier, using the ESRF account details.

NOTE: It is very important to select **Include return label** to allow the seamless return of samples (see https://www.esrf.eu/MXDewarReimbursement).

4. Data collection, viewing, and retrieval

NOTE: On the day of the experiment, samples are transferred to the MASSIF-1 High Capacity Dewar (HCD). Beamline scientists then launch the data collection, which can be followed by users remotely. For each different sample type users receive an e-mail informing them that the data collection has started. As previously noted, the execution of all steps in all workflows can be followed online and in real-time by the user *via* ISPyB, from which the results can be viewed and downloaded.

4.1. For each sample analyzed, examine the results of the automatic experiment in ISPyB (https://exi.esrf.fr/).

4.1.1. Log in, using the experiment number and password, and click on the desired experimental session at **ID30A-1**.

4.1.2. Select the preferred (top-scoring) autoprocessing pipeline (for example, Grenades or XDS_APP) and download the data written out in the correct space group with the highest completeness and highest resolution by clicking on **Last Collect Results** and, then, **Download**.

NOTE: All mesh, line, and characterization images are in a subdirectory for each sample, called /MXPressE_01. The ESRF automatically runs five separate processing packages, namely EDNA_proc[12], Grenades[12], XDS_APP[14], autoPROC[15], and XIA2[16]. Data integration is based on XDS, with the exception of XIA2, which is based on DIALS. All packages are also run in anomalous and nonanomalous modes, allowing the automatic detection of an anomalous signal, if present in the data, to be used in SAD phasing protocols. Each package uses different parameters and decision trees, meaning that some packages run better with certain samples. However, this can make for a large number of results when the number of packages and possible space groups is accounted for. The results are, therefore, ranked based roughly on resolution and other quality metrics, such as R_{merge} in the lowest resolution shell, CC(1/2), and completeness. This is aimed at guiding the user to the best data sets, but all possible space groups and results should be inspected carefully.

4.2. Unzip the downloaded folder, which will include all log files and unmerged XDS_ASCII.HKL and merged and scaled .mtz files.

NOTE: In case the structure (in PDB format) of the protein of interest or a close homolog was uploaded to ISPyB at the start of the experiment, the autoprocessing pipeline at ESRF will automatically perform a molecular replacement (MR) run using this structure as the search model on the best scoring solution. The results of the MR pipeline are displayed in ISPyB and can be found in the processed data folder (for example, /data/visitor/mx2112/id30a1/20180711/PROCESSED_DATA/GCSH/GCSH-x5/autoprocessing_GCSH-x5_run1_1/grenades_fastproc/user_nohet.pdb_mrpipe_dir/). Here, the final model will be named coot1.pdb and the reflection data sidechains.mtz. Note that the pipeline might reduce the symmetry of the cell (primitive cell reduction) in order to increase the likelihood of finding a solution. In the case here, the MR pipeline wrote out the solution in a monoclinic cell (C2) rather than in an orthorhombic cell ($C222_1$). Details on how to perform a molecular replacement run manually (exemplified for the second best-scoring autoprocessing solution) are included in the **Supplementary Files**.

REPRESENTATIVE RESULTS:

The MXPressP workflow was used at the ESRF beamline MASSIF-1 to, fully automatically, mount, center in the X-ray beam, characterize, and collect full diffraction data sets from a series of crystals of human GCSH. The samples were mounted and the loop analyzed for an area to scan (**Figure 1**, left). After the diffraction analysis, four points were selected within the crystal for data collection (**Figure 1**, right). Subsequent processing by automated data analysis pipelines, including the MR pipeline, yielded high-quality datasets (**Table 1**) for which an MR solution was found. The latter allows users to rapidly evaluate whether the obtained dataset and the used

search model are suitable for phasing by molecular replacement. In addition, the presence of ligands can be judged, thus permitting the user to focus only on the most promising datasets for further analysis. Manual structure determination by MR yielded a high-quality electron density map after a single automated refinement cycle (**Figure 2a**). For this dataset, the automated pipeline cut the data at a 1.32 Å resolution; however, users can still decide to cut the data at a lower resolution to arrive at different quality statistics ($CC_{1/2}$, $\langle I/\sigma(I)\rangle$, R_{meas}) in the highest resolution shell. The crystal structure of human GCSH structure is similar to that of the bovine protein (3KlR)[16].

Continuous electron density is visible for the entire amino acid chain, apart from the N-terminal histidine tag. Of the four substitutions that distinguish human and bovine GCSH, three are readily identifiable in the electron density (Ile/Val66, Asp/Glu98, and Leu/Phe149; **Figure 2b-d**). This is less clear for the Asp/Lys125 substitution for which the electron density of the side chain is only partially resolved due to flexibility (**Figure 1e**). The currently obtained model has R_{work} and R_{free} values of 20.4% and 23.8%, respectively, and can be further optimized by further cycles of automated and manual model building and refinement.

FIGURE AND TABLE LEGENDS:

Table 1: X-ray diffraction data collection, refinement, and validation statistics. Values for the highest resolution shell are given in brackets.

Figure 1: Sample analysis before data collection. (**A**) The region selected for scanning is shown by a red box. (**B**) The analysis of diffraction images is shown as a heat map. Four positions within the located crystal were selected for data collection.

Figure 2: Visual validation of electron density maps obtained after refinement. Electron density maps contoured at 2x r.m.s. level around (**a**) Trp143, (**b**) Val66 (Ile in human GCSH), and (**c**) Glu98 (Asp in human GCSH) and maps contoured at 1x r.m.s level around (**d**) Phe149 (Leu in human GCSH) and (**e**) Lys125 (Asp in human GCSH).

DISCUSSION:

Fully automatic beamlines provide automated characterization and data collection from large numbers of macromolecular crystals without the presence of a scientist, either at the beamline or remotely, being required. Using completely automated beamlines has many advantages compared to manual operation. For example, the automated sample centering, based on X-ray mesh and line scans, is more precise than that performed with the human eye as it is not affected by thermal or optical effects. Indeed, these mesh and line scans provide additional data (i.e., detailed dimensions of the crystal and the best diffracting region of the crystal) which are important in determining the correct beam size to use for data collection—especially for small crystals[18]—and often result in an improved quality of the obtained diffraction data. Moreover, by taking advantage of the user-defined parameters in the setup of automatic experiments, the steps in specific workflows can be tailored to best suit the system under study, thus further optimizing the experiment success rate.

Taking together, the reliability of the workflows available, the straightforward access to the beamline (users self-schedule, using a calendar [see above]), and the fully automated approach of MASSIF-1 provides a rigorous, high-throughput, and time-saving alternative to classical hands-on MX experiments and the potential to implement more advanced procedures and applications into automatic workflows. In the near future, crystal cartography in 3D[19] will be implemented to improve the accuracy of X-ray centering, while more complex protocols, such as crystal dehydration experiments[20], will be automated. It is hoped that fully autonomous data collection will become a standard method in MX, providing high-quality data for small-molecule fragment screens, optimizing the screening of large numbers of poorly diffracting crystals and automatically providing phase information to solve crystal structures de novo. In combination with developments in the automated harvesting of crystals[21], the possibility of protein crystal structure solution as an automated service could well become a reality.

ACKNOWLEDGMENTS:
The authors thank the ESRF for beamtime.

DISCLOSURES:
The authors have nothing to disclose.

REFERENCES:

1. Hui, R., Edwards, A. High-throughput protein crystallization. *Journal of Structural Biology*. **142** (1), 154–161, doi: 10.1016/S1047-8477(03)00046-7 (2003).
2. 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).
3. Svensson, O., Malbet-Monaco, S., Popov, A., Nurizzo, D., Bowler, M.W. Fully automatic characterization and data collection from crystals of biological macromolecules. *Acta Crystallographica Section D Biological Crystallography*. **71** (8), 1757–1767, doi: 10.1107/S1399004715011918 (2015).
4. Cipriani, F. et al. CrystalDirect: a new method for automated crystal harvesting based on laser-induced photoablation of thin films. *Acta Crystallographica Section D Biological Crystallography*. **68**, 1393–1399, doi: 10.1107/S0907444912031459 (2012).
5. Nurizzo, D. et al. RoboDiff: combining a sample changer and goniometer for highly automated macromolecular crystallography experiments. *Acta Crystallographica Section D Structural Biology*. **72** (8), 966–975, doi: 10.1107/S205979831601158X (2016).
6. Bowler, M.W., Svensson, O., Nurizzo, D. Fully automatic macromolecular crystallography: the impact of MASSIF-1 on the optimum acquisition and quality of data. *Crystallography Reviews*. **22** (4), 233–249, doi: 10.1080/0889311X.2016.1155050 (2016).
7. Brockhauser, S. et al. The use of workflows in the design and implementation of complex experiments in macromolecular crystallography. *Acta Crystallographica Section D Biological Crystallography*. **68** (8), 975–984, doi: 10.1107/S090744491201863X (2012).
8. Svensson, O., Gilski, M., Nurizzo, D., Bowler, M.W. Multi-position data collection and dynamic

beam sizing: recent improvements to the automatic data-collection algorithms on MASSIF-1. *Acta Crystallographica Section D Structural Biology*. **74**, 433–440, doi: 10.1101/236596 (2017).
9. Cipriani, F. et al. Automation of sample mounting for macromolecular crystallography. *Acta Crystallographica Section D Biological Crystallography*. **62** (10), 1251–1259, doi: 10.1107/S0907444906030587 (2006).
10. Delageniere, S. et al. ISPyB: an information management system for synchrotron macromolecular crystallography. *Bioinformatics*. **27** (22), 3186–3192, doi: 10.1093/bioinformatics/btr535 (2011).
11. ESRF. Structural Biology samples: how to organise the transport of reimbursed dewars. http://www.esrf.fr/MXDewarReimbursement.
12. Monaco, S. et al. Automatic processing of macromolecular crystallography X-ray diffraction data at the ESRF. *Journal of Applied Crystallography*. **46** (3), 804–810, doi: 10.1107/S0021889813006195 (2013).
13. Kikuchi, G., Motokawa, Y., Yoshida, T., Hiraga, K. Glycine cleavage system: reaction mechanism, physiological significance, and hyperglycinemia. *Proceedings of the Japan Academy. Series B, Physical and Biological Sciences*. **84** (7), 246–263 (2008).
14. Sparta, K.M., Krug, M., Heinemann, U., Mueller, U., Weiss, M.S. *XDSAPP2.0*. *Journal of Applied Crystallography*. **49** (3), 1085–1092, doi: 10.1107/S1600576716004416 (2016).
15. 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).
16. Winter, G. *xia2* : an expert system for macromolecular crystallography data reduction. *Journal of Applied Crystallography*. **43** (1), 186–190, doi: 10.1107/S0021889809045701 (2010).
17. Higashiura, A. et al. High-resolution X-ray crystal structure of bovine H-protein at 0.88 Å resolution. *Acta Crystallographica Section D Biological Crystallography*. **66** (6), 698–708, doi: 10.1107/S0907444910010668 (2010).
18. Sanishvili, R. et al. A 7 µm mini-beam improves diffraction data from small or imperfect crystals of macromolecules. *Acta Crystallographica Section D Biological Crystallography*. **64** (4), 425–435, doi: 10.1107/S0907444908001741 (2008).
19. Bowler, M.W. et al. Diffraction cartography: applying microbeams to macromolecular crystallography sample evaluation and data collection. *Acta Crystallographica Section D Biological Crystallography*. **66** (8), 855–864, doi: 10.1107/S0907444910019591 (2010).
20. Bowler, M.W. et al. Automation and Experience of Controlled Crystal Dehydration: Results from the European Synchrotron HC1 Collaboration. *Crystal Growth & Design*. **15** (3), 1043–1054, doi: 10.1021/cg500890r (2015).
21. 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).

Figure 1

Click here to access/download;Figure;Mesh_figure_1.tif

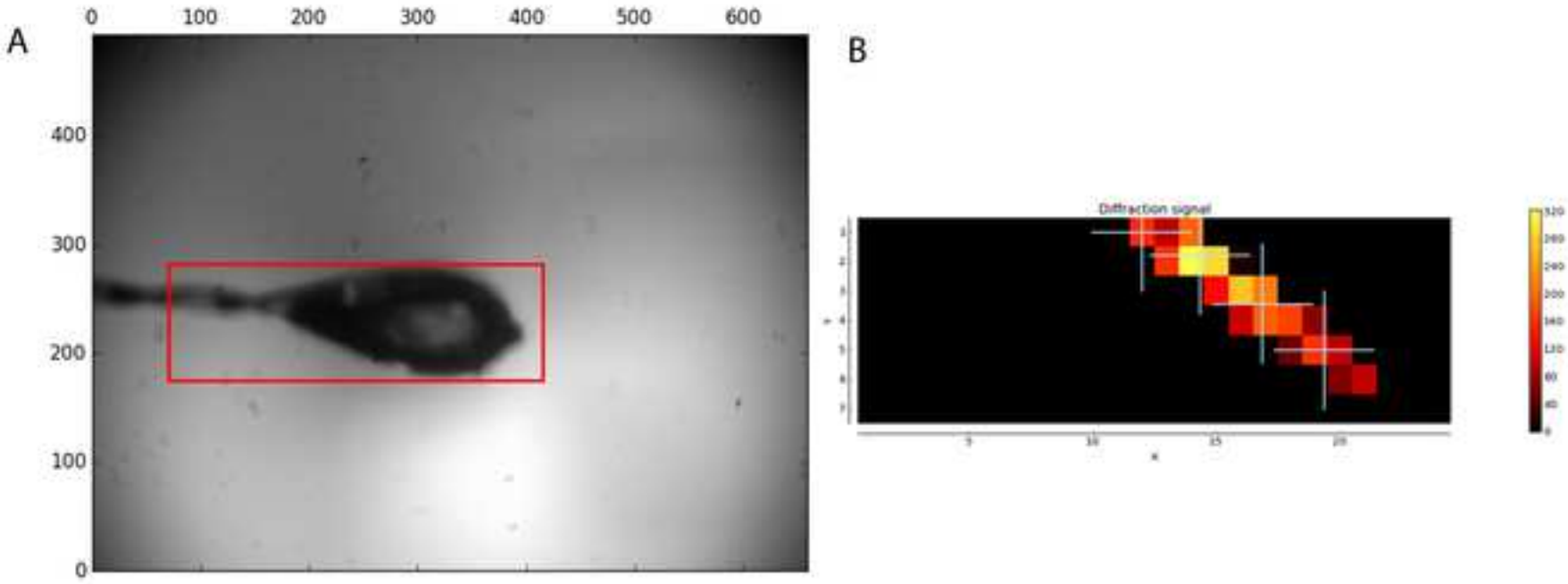
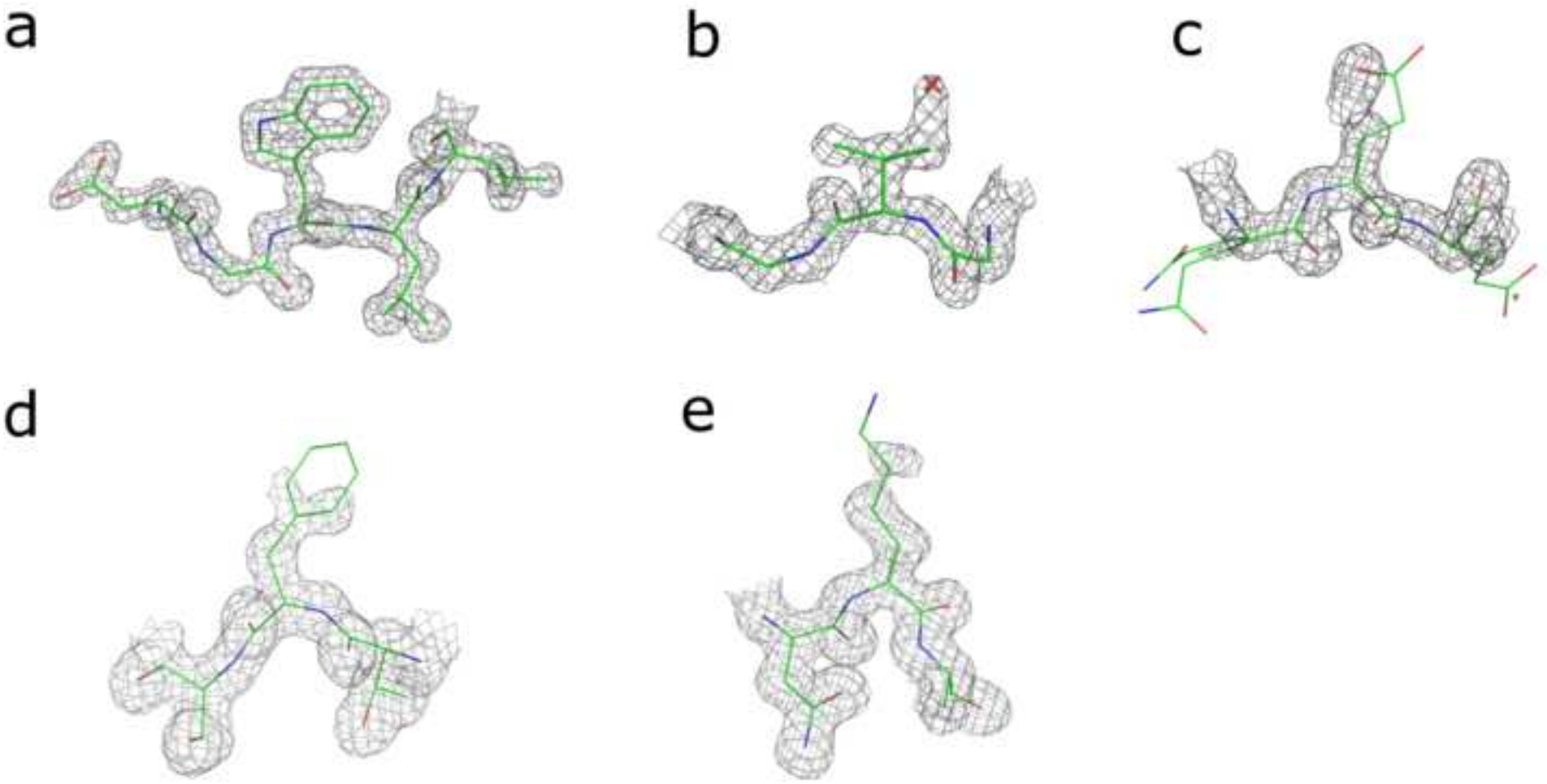


Figure 2



Table

	Human
	GRENADES pipeline
Data collection and processing	
X-ray source / Beam line	ESRF / I
Wavelength (Å)	0.
Resolution (Å)	41.88 – 1.48 (1.53 – 1.48)
Total/Unique reflections	127670 / 28644 (12178 / 2775)
Space group for indexing, scaling and merging	C2222
Cell dimensions a, b, c (Å)	 42.20, 83.75, 95.85
Mosaicity	0.05
R_{meas} (%)	10.0 (110.7)
$\langle I/\sigma(I) \rangle$	9.6 (1.3)
$CC_{1/2}$ (%)	99.7 (53.9)
Completeness (%)	99.6 (99.6)
Multiplicity	4.5 (4.4)
Molecular replacement and preliminary model refinement	
Space group for phasing	C2
Cell dimensions a, b, c (Å) α, β, γ (°)	 83.74, 42.18, 95.82 90, 90.03, 90
Search model for MR (PDB)	3KLR
Protein molecules / ASU	2
Protein residues	250
R_{work}/R_{free} (%) after 1ste refinement	24.3 / 26.5
RMSD bond length (Å) after 1ste refinement	0.01
RMSD bond angle (°) after 1ste refinement	1.2
Rotamer outlier (%) after 1ste refinement	1.07
Ramachandran favoured/allowed/disallowed (%) after 1ste refinement	95.93 / 4.07 / 0

n GCSH
XDSAPP pipeline
MASSIF-1
.966
41.86 - 1.32 (1.39 – 1.32)
177332 / 40134 (23772 / 5714)
$C222_1$
42.19, 83.72, 95,82
0.05
11.1 (198.2)
7.6 (0.7)
99.7 (19.1)
99.5 (98.6)
4.4 (4.2)
$C222_1$
42.19, 83.72, 95,82 90, 90, 90
3KLR
1
125
20.4 / 23.8
0.01
1.83
4.29
95.12 / 4.88 / 0

Name of Material/ Equipment	Company	Catalog Number	Comments/Description
Beamline MASSIF-1	ESRF		
BL21DE3	New England Biolabs	C2527I	
chloramphenicol	Roth	3886.1	
Concentrators: Amicon Ultra-4 Ultracel -30K	Merck Millipore	UFC803024	
Dialyzing membrane	Spectrumlabs	132655	
DMSO	Sigma-Aldrich	D8418	
Dnase	Roche	11284932001	
DTT	Euromedex	EU0006-B	
EDTA- free protease inhibitors	Roche	4,693,159,001	
glycerol	VWR Chemicals Prolabo	14388.29T	
His-trap HP	GE healthcare	17-5247-01	
imidazole	Sigma-Aldrich	56750-500G	
IPTG	Euromedex	EU0008-B	
LB medium	Sigma-Aldrich	L3022	
lipoic acid	Sigma-Aldrich	T5625	
loop	Hampton Research	HR8-124	
lysozyme	Roche	10 837 059 001	
MonoQ 5/50 GL	GE healthcare	17-5166-01	
NaCl	Fisher Chemical	S/3160/60	
Sonicator vibra cell 75/15	SONICS		
SPINE pucks	MiTeGen	SKU: M-CSM003-0001A	
Tris base	Euromedex	26-128-3094-B	
Sodium Formate	Sigma-Aldrich	1064430500	
GCSH purification buffer	20 mM TRIS pH 8, 200 mM NaCl		

GCSH cryo-protection buffer	0.25 M Sodium Formate pH 4, 30% glycerol		

Programs:

MxCube	Gabadinho, J. et al. MxCuBE : a synchrotron beamline control environment customized for macromolecular crystallography experiments. Journal of Synchrotron Radiation. 17 (5), 700–707, doi: 10.1107/S0909049510020005 (2010)	local development

ISPyB	ESRF	Solange Delagenière, Patrice Brenchereau, Ludovic Launer, Alun W. Ashton, Ricardo Leal, Stéphanie Veyrier, José Gabadinho, Elspeth J. Gordon, Samuel D. Jones, Karl Erik Levik, Seán M. McSweeney, Stéphanie Monaco, Max Nanao, Darren Spruce, Olof Svensson, Martin A. Walsh, Gordon A. Leonard; ISPyB: an information management system for synchrotron macromolecular crystallography, Bioinformatics, Volume 27, Issue 22, 15 November 2011, Pages 3186–3192, https://doi.org/10.1093/bioinformatics/btr535	local development

MXCube2	ESRF	Gabadinho, J. *et al. MxCuBE* : a synchrotron beamline control environment customized for macromolecular crystallography experiments. *Journal of Synchrotron Radiation* . **17** (5), 700–707, doi: 10.1107/S0909049510020005 (2010). De Santis, D., Leonard, G. Notiziario Neutroni e Luce di Sincrotrone,Consiglio Nazionale delle Ricerche. (19), 24–226 (2014).	local development
BES workflow server		Brockhauser, S. et al. The use of workflows in the design and implementation of complex experiments in macromolecular crystallography. Acta Crystallographica Section D Biological Crystallography. 68 (8), 975–984, doi: 10.1107/S090744491201863X (2012).	
DOZOR	ESRF	Bourenkov and Popov, unpublished	local development

BLISS beamline control	Guijarro, M. et al. BLISS - Experiments Control for ESRF EBS Beamlines. Proceedings of the 16th Int. Conf. on Accelerator and Large Experimental Control Systems, ICALEPCS2017, Barcelona, Spain. doi: 10.18429/jacow-icalepcs2017-webpl05 (2018).	local development
AUTO processing of images	Monaco, S. et al. Automatic processing of macromolecular crystallography X-ray diffraction data at the ESRF. Journal of Applied Crystallography. 46 (3), 804–810, doi: 10.1107/S0021889813006195 (2013)	local development
BEST and EDNA	Incardona, M.-F., Bourenkov, G.P., Levik, K., Pieritz, R.A., Popov, A.N., Svensson, O. EDNA : a framework for plugin-based applications applied to X-ray experiment online data analysis. Journal of Synchrotron Radiation. 16 (6), 872–879, doi: 10.1107/S0909049509036681 (2009).	local development

CCP4	Winn, M.D. *et al.* Overview of the *CCP* 4 suite and current developments. *Acta Crystallographica Section D Biological Crystallography*. **67** (4), 235–242, doi: 10.1107/S0907444910045749 (2011).
Phaser MR	McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., Read, R.J. *Phaser* crystallographic software. *Journal of Applied Crystallography*. **40** (4), 658–674, doi: 10.1107/S0021889807021206 (2007).
Coot	Emsley, P., Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*. **60**, 2126–32 (2004).
refmac5	Murshudov, G.N., Vagin, A.A., Dodson, E.J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallographica Section D*. **53**, 240--255 (1997).

Matthews

Matthews, B.W. Solvent content of protein crystals. Journal of Molecular Biology. 33 (2), 491–497 (1968).

1 Alewife Center #200
Cambridge, MA 02140
tel. 617.945.9051
www.jove.com

ARTICLE AND VIDEO LICENSE AGREEMENT

Title of Article: Fully autonomous characterization + data collection from crystals of biological macromolecules

Author(s): Rubin, VanLoer, Müller, Conrad, Nurizzo, BOWLER

Item 1 (check one box): The Author elects to have the Materials be made available (as described at http://www.jove.com/author) via: ☐ Standard Access ☒ Open Access

Item 2 (check one box):

☒ The Author is NOT a United States government employee.

☐ The Author is a United States government employee and the Materials were prepared in the course of his or her duties as a United States government employee.

☐ The Author is a United States government employee but the Materials were NOT prepared in the course of his or her duties as a United States government employee.

ARTICLE AND VIDEO LICENSE AGREEMENT

1. Defined Terms. As used in this Article and Video License Agreement, the following terms shall have the following meanings: "**Agreement**" means this Article and Video License Agreement; "**Article**" means the article specified on the last page of this Agreement, including any associated materials such as texts, figures, tables, artwork, abstracts, or summaries contained therein; "**Author**" means the author who is a signatory to this Agreement; "**Collective Work**" means a work, such as a periodical issue, anthology or encyclopedia, in which the Materials in their entirety in unmodified form, along with a number of other contributions, constituting separate and independent works in themselves, are assembled into a collective whole; "**CRC License**" means the Creative Commons Attribution-Non Commercial-No Derivs 3.0 Unported Agreement, the terms and conditions of which can be found at: http://creativecommons.org/licenses/by-nc-nd/3.0/legalcode; "**Derivative Work**" means a work based upon the Materials or upon the Materials and other pre-existing works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which the Materials may be recast, transformed, or adapted; "**Institution**" means the institution, listed on the last page of this Agreement, by which the Author was employed at the time of the creation of the Materials; "**JoVE**" means MyJove Corporation, a Massachusetts corporation and the publisher of *The Journal of Visualized Experiments;* "**Materials**" means the Article and / or the Video; "**Parties**" means the Author and JoVE; "**Video**" means any video(s) made by the Author, alone or in conjunction with any other parties, or by JoVE or its affiliates or agents, individually or in collaboration with the Author or any other parties, incorporating all or any portion of the Article, and in which the Author may or may not appear.

2. Background. The Author, who is the author of the Article, in order to ensure the dissemination and protection of the Article, desires to have the JoVE publish the Article and create and transmit videos based on the Article. In furtherance of such goals, the Parties desire to memorialize in this Agreement the respective rights of each Party in and to the Article and the Video.

3. Grant of Rights in Article. In consideration of JoVE agreeing to publish the Article, the Author hereby grants to JoVE, subject to **Sections 4** and **7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Article in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Article into other languages, create adaptations, summaries or extracts of the Article or other Derivative Works (including, without limitation, the Video) or Collective Works based on all or any portion of the Article and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and (c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. If the "Open Access" box has been checked in **Item 1** above, JoVE and the Author hereby grant to the public all such rights in the Article as provided in, but subject to all limitations and requirements set forth in, the CRC License.

jove JOURNAL OF VISUALIZED EXPERIMENTS
1 Alewife Center #200
Cambridge, MA 02140
tel. 617.945.9051
www.jove.com

4. Retention of Rights in Article. Notwithstanding the exclusive license granted to JoVE in **Section 3** above, the Author shall, with respect to the Article, retain the non-exclusive right to use all or part of the Article for the non-commercial purpose of giving lectures, presentations or teaching classes, and to post a copy of the Article on the Institution's website or the Author's personal website, in each case provided that a link to the Article on the JoVE website is provided and notice of JoVE's copyright in the Article is included. All non-copyright intellectual property rights in and to the Article, such as patent rights, shall remain with the Author.

5. Grant of Rights in Video – Standard Access. This **Section 5** applies if the "Standard Access" box has been checked in **Item 1** above or if no box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby acknowledges and agrees that, Subject to **Section 7** below, JoVE is and shall be the sole and exclusive owner of all rights of any nature, including, without limitation, all copyrights, in and to the Video. To the extent that, by law, the Author is deemed, now or at any time in the future, to have any rights of any nature in or to the Video, the Author hereby disclaims all such rights and transfers all such rights to JoVE.

6. Grant of Rights in Video – Open Access. This **Section 6** applies only if the "Open Access" box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby grants to JoVE, subject to **Section 7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Video in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Video into other languages, create adaptations, summaries or extracts of the Video or other Derivative Works or Collective Works based on all or any portion of the Video and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and (c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. For any Video to which this Section 6 is applicable, JoVE and the Author hereby grant to the public all such rights in the Video as provided in, but subject to all limitations and requirements set forth in, the CRC License.

7. Government Employees. If the Author is a United States government employee and the Article was prepared in the course of his or her duties as a United States government employee, as indicated in **Item 2** above, and any of the licenses or grants granted by the Author hereunder exceed the scope of the 17 U.S.C. 403, then the rights granted hereunder shall be limited to the maximum rights permitted under such statute. In such case, all provisions contained herein that are not in conflict with such statute shall remain in full force and effect, and all provisions contained herein that do so conflict shall be deemed to be amended so as to provide to JoVE the maximum rights permissible within such statute.

8. Likeness, Privacy, Personality. The Author hereby grants JoVE the right to use the Author's name, voice, likeness, picture, photograph, image, biography and performance in any way, commercial or otherwise, in connection with the Materials and the sale, promotion and distribution thereof. The Author hereby waives any and all rights he or she may have, relating to his or her appearance in the Video or otherwise relating to the Materials, under all applicable privacy, likeness, personality or similar laws.

9. Author Warranties. The Author represents and warrants that the Article is original, that it has not been published, that the copyright interest is owned by the Author (or, if more than one author is listed at the beginning of this Agreement, by such authors collectively) and has not been assigned, licensed, or otherwise transferred to any other party. The Author represents and warrants that the author(s) listed at the top of this Agreement are the only authors of the Materials. If more than one author is listed at the top of this Agreement and if any such author has not entered into a separate Article and Video License Agreement with JoVE relating to the Materials, the Author represents and warrants that the Author has been authorized by each of the other such authors to execute this Agreement on his or her behalf and to bind him or her with respect to the terms of this Agreement as if each of them had been a party hereto as an Author. The Author warrants that the use, reproduction, distribution, public or private performance or display, and/or modification of all or any portion of the Materials does not and will not violate, infringe and/or misappropriate the patent, trademark, intellectual property or other rights of any third party. The Author represents and warrants that it has and will continue to comply with all government, institutional and other regulations, including, without limitation all institutional, laboratory, hospital, ethical, human and animal treatment, privacy, and all other rules, regulations, laws, procedures or guidelines, applicable to the Materials, and that all research involving human and animal subjects has been approved by the Author's relevant institutional review board.

10. JoVE Discretion. If the Author requests the assistance of JoVE in producing the Video in the Author's facility, the Author shall ensure that the presence of JoVE employees, agents or independent contractors is in accordance with the relevant regulations of the Author's institution. If more than one author is listed at the beginning of this Agreement, JoVE may, in its sole discretion, elect not take any action with respect to the Article until such time as it has received complete, executed Article and Video License Agreements from each such author. JoVE reserves the right, in its absolute and sole discretion and without giving any reason therefore, to accept or decline any work submitted to JoVE. JoVE and its employees, agents and independent contractors shall have

jove JOURNAL OF VISUALIZED EXPERIMENTS
1 Alewife Center #200
Cambridge, MA 02140
tel. 617.945.9051
www.jove.com

ARTICLE AND VIDEO LICENSE AGREEMENT

full, unfettered access to the facilities of the Author or of the Author's institution as necessary to make the Video, whether actually published or not. JoVE has sole discretion as to the method of making and publishing the Materials, including, without limitation, to all decisions regarding editing, lighting, filming, timing of publication, if any, length, quality, content and the like.

11. Indemnification. The Author agrees to indemnify JoVE and/or its successors and assigns from and against any and all claims, costs, and expenses, including attorney's fees, arising out of any breach of any warranty or other representations contained herein. The Author further agrees to indemnify and hold harmless JoVE from and against any and all claims, costs, and expenses, including attorney's fees, resulting from the breach by the Author of any representation or warranty contained herein or from allegations or instances of violation of intellectual property rights, damage to the Author's or the Author's institution's facilities, fraud, libel, defamation, research, equipment, experiments, property damage, personal injury, violations of institutional, laboratory, hospital, ethical, human and animal treatment, privacy or other rules, regulations, laws, procedures or guidelines, liabilities and other losses or damages related in any way to the submission of work to JoVE, making of videos by JoVE, or publication in *JoVE* or elsewhere by JoVE. The Author shall be responsible for, and shall hold JoVE harmless from, damages caused by lack of sterilization, lack of cleanliness or by contamination due to the making of a video by JoVE its employees, agents or independent contractors. All sterilization, cleanliness or decontamination procedures shall be solely the responsibility of the Author and shall be undertaken at the Author's expense. All indemnifications provided herein shall include JoVE's attorney's fees and costs related to said losses or damages. Such indemnification and holding harmless shall include such losses or damages incurred by, or in connection with, acts or omissions of JoVE, its employees, agents or independent contractors.

12. Fees. To cover the cost incurred for publication, JoVE must receive payment before production and publication the Materials. Payment is due in 21 days of invoice. Should the Materials not be published due to an editorial or production decision, these funds will be returned to the Author. Withdrawal by the Author of any submitted Materials after final peer review approval will result in a US$1,200 fee to cover pre-production expenses incurred by JoVE. If payment is not received by the completion of filming, production and publication of the Materials will be suspended until payment is received.

13. Transfer, Governing Law. This Agreement may be assigned by JoVE and shall inure to the benefits of any of JoVE's successors and assignees. This Agreement shall be governed and construed by the internal laws of the Commonwealth of Massachusetts without giving effect to any conflict of law provision thereunder. This Agreement may be executed in counterparts, each of which shall be deemed an original, but all of which together shall be deemed to me one and the same agreement. A signed copy of this Agreement delivered by facsimile, e-mail or other means of electronic transmission shall be deemed to have the same legal effect as delivery of an original signed copy of this Agreement.

A signed copy of this document must be sent with all new submissions. Only one Agreement required per submission.

CORRESPONDING AUTHOR:

Name:	MATTHEW BOWLER
Department:	
Institution:	EMBL, Grenoble outstation
Article Title:	Fully automated data collection
Signature:	MBowler
Date:	31/8/2018

Please submit a signed and dated copy of this license by one of the following three methods:

1) Upload a scanned copy of the document as a pfd on the JoVE submission site;
2) Fax the document to +1.866.381.2236;
3) Mail the document to JoVE / Attn: JoVE Editorial / 1 Alewife Center #200 / Cambridge, MA 02139

For questions, please email submissions@jove.com or call +1.617.945.9051

Dear Editor,
Many thanks for giving us the opportunity to revise our manuscript. We have now edited it taking editorial and reviewer comments into account. A point-by-point rebuttal follows. Please contact us if you have any further questions.

Yours faithfully,

Matthew Bowler and Stephanie Hutin

Editorial Comments:

• Please take this opportunity to thoroughly proofread the manuscript to ensure that there are no spelling or grammatical errors.

Done

• **Protocol Language:** Please ensure that ALL text in the protocol section is written in the imperative voice/tense as if you are telling someone how to do the technique (i.e. "Do this", "Measure that" etc.) Any text that cannot be written in the imperative tense may be added as a "Note", however, notes should be used sparingly and actions should be described in the imperative tense wherever possible.
1) Examples NOT in imperative voice: parts of 1.1.1 , 1.1.7,

Section 1 has now been moved into supplementary data given reviewer comments and was changed accordingly.

• **Protocol Detail:** Please note that your protocol will be used to generate the script for the video, and must contain everything that you would like shown in the video. **Please add more specific details (e.g. button clicks for software actions, numerical values for settings, etc) to your protocol steps.** There should be enough detail in each step to supplement the actions seen in the video so that viewers can easily replicate the protocol. Some examples:
1) 1.1.1: It is unclear what exactly is done here because this appears to be a summary of several steps.
2) 1.1.4: Mention sonication frequency (Hz) and amplitude (W).
3) 1.1.5: Mention column specifications.
4) 1.1.6, 1.1.9: how is concentration performed? Mention filter cut off, centrifuge speeds (in g) and duration.
5) 1.1.7: Mention dialysis membrane specifications (dimensions, MWCO,material)
6) 1.1.8, 1.1.9: It is unclear what exactly is done here because this appears to be a summary of several steps.

Section 1 has now been moved into supplementary data given reviewer comments and was adjusted to the comments.

• **Protocol Numbering:** There must be a one-line space between each protocol step.

Done

• **Protocol Highlight:** After you have made all of the recommended changes to your protocol (listed above), please re-evaluate the length of your protocol section. There is a 10-page limit for the protocol text, and a 3- page limit for filmable content. If your protocol is longer than 3 pages, please highlight ~2.5 pages or less of text (which includes headings and spaces) in yellow, to identify which steps should be visualized to tell the most cohesive story of your protocol steps. Please see JoVE's instructions for authors for more clarification. Remember that the non-highlighted protocol steps will remain in the manuscript and therefore will still be available to the reader.

1) The highlighting must include all relevant details that are required to perform the step. For example, if step 2.5 is highlighted for filming and the details of how to perform the step are given in steps 2.5.1 and 2.5.2, then the sub-steps where the details are provided must be included in the highlighting.
2) The highlighted steps should form a cohesive narrative, that is, there must be a logical flow from one highlighted step to the next.
3) Please highlight complete sentences (not parts of sentences). Include sub-headings and spaces when calculating the final highlighted length.
4) Notes cannot be filmed and should be excluded from highlighting.
5) Please bear in mind that software steps without a graphical user interface/calculations/ command line scripting (e.g. section 6) cannot be filmed.

Please see manuscript

• **Discussion:** JoVE articles are focused on the methods and the protocol, thus the discussion should be similarly focused. Please ensure that the discussion covers the following in detail and in paragraph form (3-6 paragraphs): 1) modifications and troubleshooting, 2) limitations of the technique, 3) significance with respect to existing methods, 4) future applications and 5) critical steps within the protocol.

We think we cover this. Please see manuscript

• **Figures/Tables:**

1) Fig 1: Please add scale bars to all micrographs. Please make the panels uniform in dimensions. The right panel is much smaller than the left. Please increase the font size on the right panel to match the size of the font on the left panel.
2) Please upload each table as an individual Excel file.
3) Please expand the legends to adequately describe the figures/tables. Each figure or table must have an accompanying legend including a short title, followed by a short description of each panel and/or a general description.

Please see the figures

• **Commercial Language:**JoVE is unable to publish manuscripts containing commercial sounding language, including trademark or registered trademark symbols (TM/R) and the mention of company brand names before an instrument or reagent. Examples of commercial sounding language in your manuscript are DECTRIS PILATUS3 2M, SPINE, etc.

1) Please use MS Word's find function (Ctrl+F), to locate and replace all commercial sounding language in your manuscript with generic names that are not company-specific. All commercial products should be sufficiently referenced in the table of materials/reagents. You may use the generic term followed by "(see table of materials)" to draw the readers' attention to specific commercial names.

Dectris Pilatus has been removed – SPINE is not commercial but describes a European standard mount so has been left.

• **Table of Materials:**Please revise the table of the essential supplies, reagents, and equipment. The table should include the name, company, and catalog number of all relevant materials/software in separate columns in an xls/xlsx file. Please include items such as sonicator, columns, software,instruments, etc.

• Please define all abbreviations at first use.

Done

• Please use standard abbreviations and symbols for SI Units such as µL, mL, L, etc., and abbreviations for non-SI units such as h, min, s for time units. Please use a single space between the numerical value and unit.

• If your figures and tables are original and not published previously or you have already obtained figure permissions, please ignore this comment. If you are re-using figures from a previous publication, you must obtain explicit permission to re-use the figure from the previous publisher (this can be in the form of a letter from an editor or a link to the editorial policies that allows you to re-publish the figure). Please upload the text of the re-print permission (may be copied and pasted from an email/website) as a Word document to the Editorial Manager site in the "Supplemental files (as requested by JoVE)" section. Please also cite the figure appropriately in the figure legend, i.e. "This figure has been modified from [citation]."

Comments from Peer-Reviewers:

Reviewer #1:

Manuscript Summary:

The manuscript describes how to prepare for executing an experiment in the MASSIF beam line at the ESRf, a highly successful resource for x-ray crystallography.

Major Concerns:
I find steps 1 and 6 in the protocol out of scope. The first is a purification protocol for a single protein, and defocuses from the real issues, what to do once you do have crystals. The last step is a tutorial in using ccp4i, which again does not add much in the context of this paper.

We have now moved these sections into the supplementary data

In contrast, there is a missed opportunity on seriously expanding on 5.1.2: at that step several results are available, and novice users dont know what to download! The top-scoring is (sometimes) not the best choice. Depending on the processing program used the "mtz" file that the user mostly needs has a different name! I would actually expand on this section seriously and comment:
0. Why are there several choices for auto-processing?
1. What are the criteria for setting something as top score?
2. When should you think further than top-score (eg when a different space group is chosen by each processing program)
3. When do you want the anomalous-processed dataset and when not?
4. What are the content of each tarsal downloaded dependent on the processing program
5. Are there general recommendations for which auto-processing result to use/

We changed the entire section

The list for materials needs to be modified accordingly.

Minor Concerns:
The details of the paper are excellent, I have no comments.

Reviewer #2:
Minor Concerns:
A small point, but the special case of molecular replacement where the search model and the target crystal share the same unit cell and space group is often referred to a "Molecular Substitution" to prevent confusion with the more general method of Molecular Replacement that has no such limitations. I would not want to insist on this as a change, though, as referring to this case as Molecular Replacement is quite common. The description of how to carry out a molecular replacement manually are correct and helpful.

Reviewer #3:
Manuscript Summary:

This article describes automatic data collection at a MX beamline, ID30A-1 in ESRF. It covers all the protocols from obtaining a beam time to determining a crystal structure. This is quite informative to structural biologists and other researchers, and illustration by movies are convenient for synchrotron beginners.

Major Concerns:
Therefore, most parts of it are reasonable, however, we felt redundant in the following protocols:
1. Protocol 1. The purification and crystallization of GCSH. This part might be required as "materials and methods" in a general articles (not JoVE). However, it is a sample-dependent description. We felt it should be separately described as a supplemental.
2. Protocol 6. The protocol for molecular replacement. This part is conducted manually. We felt it should be separately described.
We felt both are not directly included in the pipeline. Instead of them, by adding a flowchart of the pipeline we can imagine its overview easily.

We have now moved these sections into the supplementary data

Minor Concerns:
Moreover, please confirm the following a minor point:
3. Line 277 search model must be in the same space group as the tested crystals.
Why it should have same space group?

This speeds up the calculations and was required for older versions but thanks to the primitive cell reduction this is no longer a strict requirement.

SUPPLEMENTARY INFORMATION FOR:

Fully autonomous characterization and data collection from crystals of biological macromolecules

AUTHORS AND AFFILIATIONS:

Stephanie Hutin[1*], Bart Van Laer[1*], Christoph Mueller-Dieckmann[1], Gordon Leonard[1], Didier Nurizzo[1] and Matthew W. Bowler[2].

[1]European Synchrotron Radiation Facility, Structural Biology Group, Grenoble, France
[2]European Molecular Biology Laboratory, Grenoble Outstation, Grenoble, France
[*] These authors contributed equally to this work.

Corresponding Author:
Bowler, Matthew
European Molecular Biology Laboratory
Grenoble Outstation,
France
mbowler@embl.fr

Protein expression, purification and crystallization adapted from Macherel *et al.* [1]

1.1 Express and purify GCSH

1.1.1 Inoculate three 1l LB flasks with BL21(DE3) *E. coli* cells containing a pET21d vector with the gene for the GCSH construct (His-tagged GCSH without the mitochondrial signal sequence (residue 49-173)). Add 100 µM lipoic acid (from a 1M stock in dimethylsulfoxide) and 100 µg/ml ampicillin to the media and grown cells at 37°C until an optical density at 600 nm of 0.6 is reached.

1.1.2 Induce the expression of the protein by addition of 0.4 mM IPTG.

1.1.3 Add 34 µg/mL chloramphenicol after 3 hours to stop the expression and enhance the lipoylation of GCSH. Maintain the cultures overnight at 20°C.

1.1.4 Harvest the bacterial cells (5000 x g, 10 min) and sonicate (10 sec on/20 sec off for a total time of 10 minutes at 40% amplitude, 20 kHz) in the presence of 50 mL 50 mM Tris pH 8.0, 150 mM NaCl, 1 mM DTT, 1 mg DNase, 1 mg lysozyme and EDTA-free protease inhibitors.

1.1.5 Purify the protein on a 5ml Ni-sepharose column (GE Histrap HP) and elute in 50 mM imidazole in the same buffer.

1.1.6 Pool the fractions containing protein and concentrate to 5 mL using a 15 kDa molecular weight-cutoff (MWCO) concentrator at 4000 g.

1.1.7 Dialyzed overnight at 4°C against 100 ml 20 mM Tris at pH 8.0, 100 mM NaCl using a 3 kDa molecular weight-cutoff (MWCO) cellulose dialysis membrane.

1.1.8 Perform an anion exchange chromatography (GE MonoQ 5/50 GL) using a gradient of 100 mM to 1M NaCl in 20 mM Tris at pH 8 over 30 minutes at a 1 ml/min flow rate.

1.1.9 Concentrate the protein peaks to 6.5 mg/mL using a 15 kDa molecular weight-cutoff (MWCO) concentrator at 4000 g and send to a mass spectrometry facility to find the protein fraction containing purified non-lipoylated GCSH (15562 Da).

1.2 Crystallization procedure

1.2.1 Mix 25 µl of GCSH solution with 20 µl 0.5 M Sodium formate pH 4.0 and store the sample at 4°C for a week.

1.2.1.1 Note: Crystal needles will appear.

Determination of phases by molecular replacement and refinement

2. Open the graphical interface of CCP4[2].

2.1 Determine the Matthew's coefficient, indicating the number of molecules in the asymmetric unit by opening the Matthews_coeff tab[3].

2.1.1 Press "Browse" and select the .mtz file of the collected data set.

2.1.2 Enter the molecular weight (15562Da) in the interface.

2.1.3 Press "Run now" and read the solvent content analysis to select most likely the composition of the asymmetric unit (P(tot) close to 1).

2.1.4 Press "close2.

2.2 Open the "Phaser MR" tab in the CCP4 interface[4].

2.2.1 Enter a job title.

2.2.2 Import again the .mtz file using "browse".

2.2.3 In the "Define ensembles (models) section enter an ensemble name and import the Protein Data Bank entry of the search model (bovine H-protein, PDB code: 3klr).

2.2.4 Give the sequence identity (which is 98%, but due to stringency we used just 90%).

2.2.5 Give the number of molecules in the asymmetric unit (1) in the "Define composition of the asymmetric unit" section and import a sequence file (.seq).

2.2.6 Select the search model in the "Search parameters" section and the number of copies to search for (1).

2.2.7 Press "Run now".

Note: The TFZ score should be at least 6 but above 8 indicates a good solution. The data yielded a single solution with a TFZ score of 64.1 and without any clashes).

2.3 Open a visualization program (Coot[5]) and inspect the density map.

2.4 Select refmac5[6] in the program list of ccp4. Run a restrained refinement (default

setting).

2.5 Import the mtz file and pdb file (output of Phaser) using `browse` in the corresponding field. Optional select “run Coot: find waters” to automatically add waters in the structure.

2.6 Press `run now`

References

1. Macherel, D., Bourguignon, J., Forest, E., Faure, M., Cohen-Addad, C., Douce, R. Expression, lipoylation and structure determination of recombinant pea H-protein in Escherichia coli. *European Journal of Biochemistry*. **236** (1), 27–33 (1996).
2. Winn, M.D. *et al.* Overview of the *CCP* 4 suite and current developments. *Acta Crystallographica Section D Biological Crystallography*. **67** (4), 235–242, doi: 10.1107/S0907444910045749 (2011).
3. Matthews, B.W. Solvent content of protein crystals. *Journal of Molecular Biology*. **33** (2), 491–497 (1968).
4. McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., Read, R.J. *Phaser* crystallographic software. *Journal of Applied Crystallography*. **40** (4), 658–674, doi: 10.1107/S0021889807021206 (2007).
5. Emsley, P., Cowtan, K. *Coot* : model-building tools for molecular graphics. *Acta Crystallographica Section D Biological Crystallography*. **60** (12), 2126–2132, doi: 10.1107/S0907444904019158 (2004).
6. Murshudov, G.N. *et al. REFMAC* 5 for the refinement of macromolecular crystal structures. *Acta Crystallographica Section D Biological Crystallography*. **67** (4), 355–367, doi: 10.1107/S0907444911001314 (2011).