



HAL
open science

Structural Analyses of Bacterial Effectors by X-Ray Crystallography

Chloé Dugelay, Virginie Gueguen-Chaignon, Laurent Terradot

► **To cite this version:**

Chloé Dugelay, Virginie Gueguen-Chaignon, Laurent Terradot. Structural Analyses of Bacterial Effectors by X-Ray Crystallography. Laure Journet, Eric Cascales. Bacterial Secretion Systems, 2715, Springer US, pp.485-502, 2023, Methods in Molecular Biology, <10.1007/978-1-0716-3445-5_29>. <hal-04793247>

HAL Id: hal-04793247

<https://hal.science/hal-04793247v1>

Submitted on 20 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Structural analyses of bacterial effectors by X-ray crystallography.

Chloé Dugelay ¹, Virginie Gueguen-Chaignon ² and Laurent Terradot ^{1,*}.

Affiliations:

1. Microbiologie Moléculaire et Biochimie Structurale (MMSB), Université Lyon 1, CNRS, UMR5086, 69007, Lyon, France
2. Protein Science Facility, SFR Biosciences, Centre National de la Recherche Scientifique UAR3444, Université de Lyon, Lyon, France

*** To whom correspondence should be addressed**

Email: laurent.terrادت@ibcp.fr

Tel: +33(0)472722652

Abstract

X-ray crystallography is a method of choice to determine and analyse protein structures. Although large complexes are challenging to crystallise and cryo-electron microscopy is thus better suited for these, crystallography can still be efficient in solving structures of single components of secretion systems or effectors. Many of the different steps leading to structure determination by X-ray crystallography have been automatised. Here we describe a generic approach to obtain crystals, solve the structure of a given protein and perform a preliminary analysis, highlighting novel and efficient possibilities offered by automatization and contribution of Alpha Fold 2 structure prediction.

Running head title: X-ray crystallography of bacterial effector proteins

Key words

X-ray crystallography, bacterial effector, secretion system, X-ray diffraction, protein data bank, Alpha Fold.

1. Introduction

Bacterial secretion systems are of considerable interest from a number of perspectives. They are involved in many biological processes such as conjugation, transformation, biofilm formation or infection. Some of them deliver protein effectors that modulate host defences and metabolism, thereby promoting infection. Structural biology and in particular X-ray crystallography has played a crucial role in deciphering the function of these molecules. The development of cryo-electron microscopy has enabled the study of large, multi-components flexible assembly that were particularly difficult (or impossible) to crystallise (1). However, for single protein components (or low complexity complexes) such as bacterial effectors, X-ray crystallography is still a very efficient method to determine the structure of a protein.

The past decade has seen spectacular advances in automation for several steps of the X-ray crystallography process, including crystallisation conditions screening, data collection and processing, structure determination, model building and refinement (2-7). As a consequence, determining a structure by X-ray crystallography can be very rapid, even as fast as a couple of weeks and the recent COVID pandemic has also demonstrated that Drug discovery using automated X-ray crystallography structure determination can be remarkably efficient (8). The process of solving a crystal structure is entirely dependent on the availability of diffracting crystals for a given protein and this remains the major bottleneck of the method. But once the diffraction data is available and at resolutions higher than $\sim 3.5 \text{ \AA}$, structure determination can be, in most cases, rather straightforward. To solve the phase problem of X-ray diffraction, protein crystallographers rely on two main methods: i) molecular replacement (MR) or ii) experimental phases determination using derivatives or intrinsic sulfuric atoms. On the one hand, producing derivative crystals can be achieved by incorporating heavy metals in the crystals or producing Seleno-methionine protein to generate new crystals. In that case, solving the structure requires careful data collection and wavelength-tunable beamlines to optimise the

diffraction data collection. On the other hand, MR uses an available 3D model of a homologous protein to place it in the asymmetric unit so that to obtain preliminary phases. With Alpha fold *de novo* prediction from a single amino-acid sequence, structural biology enters a new era, where these prediction models can be used in MR to solve the phase problem of virtually any diffracting crystal, even if their diffraction is at lower resolution.

In previous chapter(s), protein purification and characterization protocols are described and are thus not discussed here, although these are important steps for protein crystallisation. Assuming that you have pure, monodispersed protein, we describe the different steps (**Fig.1**) to solve the structure of a given protein or domain by X-ray crystallography at a high brilliance synchrotron. This chapter assumes that the reader has a good knowledge of the principles of X-ray diffraction. We incorporate the use of Alpha fold 2 (9) to generate models for structure determination by MR and a list a number of procedures for preliminary structural analysis of the model obtained that can identify key properties and guide future functional experiments.

2. Materials

2.1. Drop set-up using mosquito robot

- Mosquito Robot (STP Labtech <https://www.sptlabtech.com/products/mosquito>)
- RockImager 182 (Formulatrix <https://formulatrix.com/protein-crystallization-systems/rock-imager-crystallization-imagers/>)
- MRC plates 2-Lens 96-wells Crystallisation plate (Swissci #MRC96T-PS)
- Nunc 96-wells plates, conical (Thermofisher Scientific #249662)
- Pipet 1-10 µl, 20-200 µl and 100-1000 µL and corresponding pipet tips
- 8-channels pipet and corresponding pipet tips
- Ultrafree MC GV Centrifugal filter (Millipore UFC30GV00)

- Sealing films: visualisation film Greiner #67670 and covering film before experiment

Greiner #67601

- 96 well crystallisation screens (Table 1)

2.2. Manual optimization of crystallisation conditions

- 24-wells pre-greased plates Calibre Scientific #MDL-MD3-21
- MRC 48-wells plates Swissci #MMX02-PS
- Siliconized coverslides 18 mm for 24-wells plate Hampton Research #HR3-239
- Pipet 1-10 μ l and 100-1000 μ L and corresponding pipet tips
- Stock solutions of precipitant, salts and buffers. These can be home-made or purchased

directly at Hampton Research or Calibre Scientific: <https://hamptonresearch.com/>,
<https://calibrescientific.com/>)

- Binocular Olympus SZX10

2.3. Crystal harvesting freezing

- Freezing loop: 18 mm Mounted CryoLoop Hampton Research #HR4-918 - 20
micron (30 pack: 5 of each diameter size 0.1 - 1.0 mm)

- Freezing bucket: Cryogenic Foam Dewar Hampton Research #HR4-673

- Storage dewar: CX100 Taylor-Wharton cryogenic dry shipper #MDL-MD7-21

Calibre scientific

- Pucks and accessories Calibre Scientific: Uni-puck #MDL-MD7-601, magnetic
wand #MDL-MD7-606, Puck-Shelved Shipping Cane #MD7-603, Angled Cryo-Tongs
#MD7-604, Cane Lifting Tool (Hook) #MD7-609, Puck Dewar Loading Tools set (#MD7-
607)

- Liquid nitrogen

2.5. Data collection, processing, model building and analysis

- Access to synchrotron facility X-ray crystallography beamline
- Computer with installed ccp4 (10) and Phenix (11).
- Protein analysis and visualisation softwares: Coot (12), Pymol (Schrodinger) and Chimera (13).
- Storage place 500 Mo minimum.

3. Methods

3.1. Crystal screening

Six crystallisation screens are routinely used in the first screening approach: Crystal screen or JCSG+, PEGIon, MbClassII, Natrx, Morpheus, BCS or Pact1er (Table 1). Crystal screening is performed on MRC 2 lens 96-wells plates using a Mosquito robot.

1. First, manually fill the reservoirs of the MRC plates with 70 μ l of the screens using the 8-channels pipet.
2. Temporarily cover the plates with sealing film while preparing the protein.
3. Concentrate the pure and homogeneous protein (Note 1) to approximately 10 mg/ml and filter on a 0.22 μ m centrifugal filter.
4. As MRC plate contains two drops per well, one can test two concentrations 5 and 10 mg/ml for example or, if not enough protein is available, test only one concentration, 10 mg/ml, and the purification buffer in the second drop as a control. This allows to check whether the crystals obtained are protein crystals or salt crystals from the purification buffer.
5. Deposit the protein in a Nunc plate on an 8-wells column, which will be used by the robot for dispensing in crystallisation plates. To make sure that enough protein to transfer to the 6 crystallisation plates, 18 μ l is deposited in each well of this column.
6. Program the Mosquito robot to set up drops of 200 nl of protein + 200 nl of crystallisation solution. It first transfers 200 nl of protein from a column of a Nunc plate to the first well,

then 200 nl of crystallisation solution from the reservoir are added to the protein drop. In the software program, specify to run half of the plate, then the other half to avoid evaporation of the drops. Fig.2 illustrates the flowchart of this screening step.

7. Run the Mosquito robot. Once the two drops are made, immediately cover the plate with the visualisation sealing film.
8. On the RockImager imaging robot, design a barcode corresponding to the plate to identify the owner, the project, the date, the screen and the schedule of imaging.
9. Incubate the plate in the robot at 19°C. Images are taken regularly. Using a Fibonacci sequence to set the imaging schedule allows for good tracking of crystal growth.

3.2. Optimization

Once conditions of crystallisation have been identified, crystals often need to be upgraded in size and quality. Several approaches can be considered according to the quantity of pure protein available, the shape of crystals, and the parameters that have to be optimized. You may need to increase crystal size, slow crystal growth, improve monodispersity and shape. For all approaches, optimization can be performed manually or by using automated systems. An example of possible optimization steps is given in Fig.3.

The first optimization step can be to use a screen dedicated to the main precipitant which led to the first crystals. For example, PEGsI and PEGsII can be used if PEG is the main precipitant or AmSO₄ or MPD screens can be used to optimize AmSO₄ or MPD conditions respectively. Then the crystallisation conditions are optimized by varying several parameters: the percentage of precipitating agent, the pH of the buffer, the concentration of salt additive. The protein concentration can also be optimized by reducing it if too many precipitates are visible or by increasing it if the crystals are small and /or appear late. The crystallisation

temperature is also an optimization factor if the crystals appeared very quickly, then redoing the plate at 4°C can slow down the growth and increase the crystal size.

1. In 24-wells plates, vary the concentration and pH around the initial crystallisation condition.
Fill the reservoir with 500 µl of these crystallisation solutions.
2. Deposit 1 µl of protein solution on silicone glass cover slide placed on a plastic support (or a paper).
3. Deposit gently 1 µl of the corresponding reservoir solution on the protein drop.
4. Take the cover slide, turn them upside down and place them on the corresponding well in the crystallisation tray. Press slightly on the cover slide to generate vacuum with the grease.
5. Process like this for the remaining wells, close the tray and leave it in the temperature control chamber or room.
6. The days after, inspect the plates at regular period using the binocular microscope and look for single crystals.

Other methods can be applied to increase the quality, regularity and size of crystals (Note 2).

3.3. Crystal harvesting and freezing

1. Prepare the material, cool the puck into liquid nitrogen (Fig.4)
2. Look for the crystals that meet quality criteria, single crystals, with a size above 30 µm.
3. Prepare your cryo-protectant solution according to the reservoir solution (Note 3).
4. Open the crystallisation tray and deposit your cover slide with the drop up on the transparent lid of the tray. Deposit 1 µl of reservoir solution near your drop as to make a second drop on the cover slide. Fish and transfer a single crystal using the mounted loop and magnetic wand. Make sure you transfer a single crystal at this stage. Once the crystal is in the drop, add 0.5 µl of cryo-protectant solution using the pipet. Let it equilibrate for one minute.

5. Repeat this step 3 times, increasing the percentage of cryoprotectant solution. Transfer the crystal equilibrated in the nearby drop of cryo-protectant with the loop (Note 4).
6. Fish a single crystal and plunge it rapidly into the liquid nitrogen up to covering the loop and base (Fig.4). Once frozen, place the loop in the puck and write down the position for future data collection.

3.4. X-ray diffraction screening and data collection

We perform X-ray diffraction screening at the high brilliance European Synchrotron Radiation Facility (ESRF) or at the synchrotron SOLEIL.

1. Pucks with harvested crystals are placed in the robot storage (for instance Automatic sample changer FlexHCD) by the local contacts of the synchrotron. Here we use ESRF set-up as an example for data collection.
2. Mount the frozen crystal using beamline robot from the control cabin or remote access using Mxcube (14). Test the crystal for diffraction individually by collecting 1° rotation images at position 0, 90° , 180° and 270° . Crystal characterization is performed automatically using the program BEST (15), which determine the space group of the crystal, the best orientation for the goniometer and number of images and oscillation angle. BEST also optimizes data collection strategy to consider radiation damage. If diffraction is good (individual spots, lattice identified) move to the next step. If not, try another crystal.
3. If you suspect a metal bound to your protein, it is a good practice to perform an Extended X-Ray Absorption Fine Structure (EXAFS) scan of your crystal to identify or confirm the metal present. If present, adjust the wavelength of data collection so that you can measure the anomalous signal (Note 5).

4. Collect data in a separate folder with the crystal name. Adjust the diffraction resolution (position of the detector) and the wavelength. If you intend to use anomalous signal for phasing, optimise your data collection strategy accordingly (highly redundant, shorter exposition).

3.5. Data processing and analysis

1. At the ESRF, using the data server IspyB (16), collected data will be automatically analysed, reflection integrated using XDS (17), scaled and merged using AIMLESS (18). Diffraction anisotropy will be also assessed and included in the Star aniso file. You will obtain a .mtz file with your intensities in the most probable space group. This file contains your experimental data truncated at the appropriate resolution. Verify in the .log file that the dataset parameters are satisfactory (intensities, Rpim, CC1/2, completeness, etc.). Note 6.
2. Alternatively, you can use the XDS output XDS_ASCII.hkl. Use this file in ccp4i with the program AIMLESS to determine the space group and also generate the free set of reflection (5% of total) that will be kept aside with the label FreeR_flag. A good practice is to save all log files from program for future pdb deposition.

3.6. Structure determination and model building

1. Determine the number of molecules to search for in your asymmetric unit (AU) using your .mtz file in the program MATTHEWS in ccp4. This will also give you the solvent content of your crystal (in percentage).
2. If you have found a suitable homologous protein structure download its .pdb file coordinates as search model (for instance, name it MRsearchmodel.pdb)
3. Open a web browser Alpha Fold window and input your sequence

<https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb>

-). Execute all. Wait until modeling is done and download your model (.pdb file). In Pymol, open the .pdb, display as cartoon and color it by spectrum/b-factors. Here the model will be colored in rainbow according to the prediction quality (pLLD) from very good (red) to poor (blue).
4. Examine the model and cut parts that are protruding out of the core and/or are badly predicted (with low pLLD). Save your model as the MRsearchmodel.pdb.
 5. Open ccp4>Molecular replacement>Phaser. Load your structure factor (mtz) file. The space group and resolution should load automatically. Use the MR search parameter to upload your MRsearchmodel.pdb from the homologous protein structure or Alpha fold modified model. Indicate the solvent content of your crystals (obtained in step 1) and the number of molecules to search for. You can do the MR search in the current space group or in enantiomers (recommended). Run the program.
 6. If successful, you will have as output a .pdb file containing all the models (if you asked for several) placed in the AU and an .mtz file with the refined structure factors (read the manual here [https://www.phaser.cimr.cam.ac.uk/index.php/Phaser-2.7.17: Manual](https://www.phaser.cimr.cam.ac.uk/index.php/Phaser-2.7.17:Manual) for details). If unsuccessful, check that you have indicated the right number of molecules to search for or try different space groups (Note 7).
 7. Open the map (the function autopen.mtz will display both the 2Fo-Fc in blue and difference map Fo-Fc in green (positive difference) and red (negative difference)). Open the .pdb file in COOT and examine the quality of the density map (contoured at around 1.3 σ) and the fitting of your model in the density. Modify your model according to the electron density by manually delete and rebuild the model. Alternatively, use automatic model building. For this you can use ccp4 software Buccaneer (19), ARP/wARP or phenix Autobuild with the original .mtz file, from phaser and the .pdb model output from phaser.

3.7. Refinement/validation/deposition

1. Once you have a complete model, refine your model against your X-ray data using Phenix or Refmac. This step is iterative and should be done carefully with proper refinement restraints so that $R_{\text{work}}/R_{\text{free}}$ diminish together. Many refinement parameters can be adjusted manually (restraints, non-crystallography symmetry, R_s difference values). Using Phenix for refinement will provide you with many validation parameters that you can monitor/check to make sure that your model is being ameliorated to fit the X-ray experimental data. Geometric parameters should all be valid and monitored carefully.
2. Check your model/data refinement with PDB-redo automated refinement (20).
3. Once your model is complete, all waters and ligands are built and refined, validate and deposit your structure using the Protein Data Bank One Dep server (<https://deposit-pdbe.wwpdb.org/deposition>).

3.8. Structure analysis

3.8.1. PBDSUM: general properties of your protein

A good start for protein structure analysis is to use the server PDBSum from EMBL-EBI (21). This will provide you with a number of structural properties.

1. Log into the PDBsum and go to the “generate” on the left menu or go directly here (<http://www.ebi.ac.uk/thornton-srv/databases/pdbsum/Generate.html>)
2. Upload your coordinate file and start. Indicate the name of your structure and your e-mail. You will receive a link after one hour or less.
3. Go to your results page and you will find all the results
4. You can analyse the different elements: secondary structure elements, topology diagram and putative clefts/binding sites.

3.8.2 PISA analysis, interface and multimeric assemblies

1. Log into the PISA server (<https://www.ebi.ac.uk/pdbe/pisa/>) and launch PISA (22).
2. Use “coordinates” and click on “upload” to input your coordinate file. PISA will find the number and names of chains in your structure.
3. Click on “interfaces”. Wait for the job to be finished. The output is a list of interfaces analysed by PISA and provide you with characteristics (surface buried, solvation free energy gain upon formation of the interface, etc...).
4. Select an individual interface and click on “details”. This will provide you with all the information about the given interface, notably the interacting residues and the characteristics of the interactions between them. PISA will also list each hydrogen bond, salt bridge, etc..
5. PISA will also give a “significance” score for the interface (between 0 and 1). Interfaces with score of 1 are generally genuine.

3.8.3. Comparing structures using DALI

1. Log into the DALI server (<http://ekhidna2.biocenter.helsinki.fi/dali>) (23) and select the PDB search tab to compare the tridimensional structure of interest with all structures available in the Protein Data Bank. Upload the PDB file into DALI server and submit your request.
2. Once the job is done, the DALI server will send you an email with a link to access at the results page. Click on the link and click on “Matches against full PDB”.
3. The results are presented as a list of structures ranked by the Z-score defined as the sum of equivalent residue-wise C α -C α distances among two proteins (24). The higher the Z-score, the more the structure of interest shares similarities with the uploaded structure. To analyse all results, DALI server also provides you with additional information: the Root-mean-

Square-Deviation, the number of residues of the structure of interest aligned with the proposed structure “lali” and the number of total residues of the proposed structure “nres” (Note 8). Finally, the last column of the summary results is a “description” corresponding to the function of the structure.

4. To compare the structures, you can download the text file generated by DALI. Alternatively, download the PDB code of the best match(es) from the PDB database. Open Chimera 1.16. Click on the File tab and select open. Select your structure of interest and the structure you downloaded and that you want to compare. Use the Matchmaker tab to structurally align the models (Fig.5a). This will align your model and generate the corresponding sequence alignment.
5. To compare the conservation of residue, visualize the sequence alignment based on the structure alignment by Chimera. Click on tools structure comparison and Match align (Fig.5c, d).
6. Save the Chimera sequence alignment as ALN file. Log the ALN file into the Endscript 3.0 (<https://espript.ibcp.fr/ESPrIPT/cgi-bin/ESPrIPT.cgi>) (25). Add the PDB file of the structure of interest into the “secondary structure depiction” section and click on Submit.
7. The colour coding of ESPrIPT server 3.0 highlights the level of residue conservation (Fig.5e). The combined use of Chimera and ESPrIPT allows to visualize the conservation of residue and the conservation of folding.
8. To visualize details of the structures (metal coordination, side chain positions and conservation), you can also use Pymol as shown in the example (Fig.5f)

3.8.4. Analysis of surface properties

Charge and hydrophobicity

1. Open the PDB file in UCSF Chimera X software (26), select the molecule display tab and click on “hydrophobicity”. The surface of your model will be displayed and coloured according to hydrophobicity ranging from dark cyan (most hydrophilic) to white to dark goldenrod (most lipophilic).
2. To visualize the electrostatic potential of the surface of a protein, log the PDB file into UCSF Chimera X software. Select the molecule display tab and click on electrostatic. Negatively charged surface is coloured in red while positively charged surface is coloured in blue.

Metal ligand properties

1. Upload the PDB file into the “CheckMyMetal” server (<https://cmm.minorlab.org/metal-sites/showsite>) (27).
2. The server characterizes the metal ligand properties as the geometry of metal coordination, the metal ligand distances and the atomic contact (Fig.6). All properties are defined by a colour depending on the physicochemical characteristics of the metal-ligand interactions.

3.8.5. Conserved surface residues

1. Upload the PDB file of the structure into the Consurf server (https://consurf.tau.ac.il/consurf_index.php) (28) and submit your request.
2. Once the job is done, the Consurf server will send you an email with a link to access at the results page. Click on the link and go to the Results page.
3. Consurf server coloured the residues according to their conservation. The conservation is split into 9 levels of colours from Variable to Conserved.
4. Download the high-resolution Figures and open these with Chimera or Pymol software. Visualize conserved residues at the surface via the colour code of Consurf.

5. Consurf generates Multiple Sequence Alignment in FASTA format or on line with the colour code. It allows to access at the list of sequence homologues of the protein of interest.

4. Notes

1. Before any crystallisation experiment, the quality of protein should be checked with several biophysical approaches: absence of aggregation and monodispersity with dynamic light scattering (DLS) experiment, oligomerization state and monodispersity using a multiangle light scattering (MALS) coupled to a size exclusion chromatography, folding of the protein with circular dichroism. Stability of the protein of interest can also be checked in several buffers using differential scanning fluorimetry (NanoDSF) or thermal shift assay (TSA).

2. If the crystals are small and numerous, they can be used to seed new drops. Two techniques can be implemented: you can slide a cat's or horse's whisker through the drop containing the microcrystals which will cling to the whisker. Then the whisker is passed through a new drop and the microcrystals will serve as seeds for larger crystals. This approach is used in 24-wells plates. For the second technique, microcrystals are crushed with a small glass pestle and serial dilutions in crystallisation solution are prepared up to $1/10000^{\text{th}}$. Then, a new drop is set up containing both the crystallisation solution, the protein and the dilution with crushed microcrystals. This microseeding technique can also be performed in 96-wells plates, using the Mosquito, to multiply conditions.

3. Crystallisation conditions should be carefully considered and cryoprotection solutions prepared using 15-25% glycerol or ethylene glycol (final concentrations). Other cryoprotectants might be used such as sugars, high concentration salts or solvents. To have a good idea of glycerol concentrations to be used as efficient cryoprotectant, these have been assessed for 50 different crystallisation solutions (29).

4. Manipulate crystals very gently and avoid touching them with the loop. Try to place them in the middle of the loop without breaking them. If they are fragile, add the cryoprotectant solution directly in the crystallisation drop as to limit the manipulation steps.
5. By collecting data at the correct wavelength, you optimise the anomalous signal in the diffraction data and thus directly use this data to obtain experimental phases and solve the crystal structure. This was used recently in our group to solve the crystal structure of the effector domain Tre^{Tu} of the *Salmonella enterica* Typhimurium rearrangement hot spot (Rhs) toxin using Zn as metal present in the crystallisation conditions (30).
6. Expertise in X-ray crystallography is likely to be required at some point if MR does not work automatically. First data processing is generally better performed by automation using the beamline on-the-fly server. If you are to do it manually, then you should have a strong knowledge of how to interpret the results of your integration/scaling/space group attribution. Moreover, even if you find the correct space group and dataset at good resolution, MR is not always straightforward. It depends on the protein structure, the crystal system, the solvent content, the protein structure (multidomain, secondary structure content) and many other parameters.
7. At this point it is not always clear which space group the crystal belongs to. The final/definitive space group will be obtained with structure determination and sometimes also during refinement.
8. With the number of residues of the structure of interest aligned with the proposed structure “lali” and the number of total residues of the proposed structure “nres” it is possible to calculate the percentage of total residues aligned between the protein of interest and the DALI match.

Acknowledgments

This work was supported by the French National Research Agency (grant number ANR-19-CE11-0012), The authors declared no competing interests.

5. References

1. Bai, X.C., McMullan, G. and Scheres, S.H. (2015) How cryo-EM is revolutionizing structural biology. *Trends Biochem Sci*, **40**, 49-57.
2. Nanao, M., Basu, S., Zander, U., Giraud, T., Surr, J., Guijarro, M., Lentini, M., Felisaz, F., Sinoir, J., Morawe, C. *et al.* (2022) ID23-2: an automated and high-performance microfocus beamline for macromolecular crystallography at the ESRF. *J Synchrotron Radiat*, **29**, 581-590.
3. Nurizzo, D., Bowler, M.W., Caserotto, H., Dobias, F., Giraud, T., Surr, J., Guichard, N., Papp, G., Guijarro, M., Mueller-Dieckmann, C. *et al.* (2016) RoboDiff: combining a sample changer and goniometer for highly automated macromolecular crystallography experiments. *Acta Crystallogr D Struct Biol*, **72**, 966-975.
4. de Sanctis, D., Beteva, A., Caserotto, H., Dobias, F., Gabadinho, J., Giraud, T., Gobbo, A., Guijarro, M., Lentini, M., Lavault, B. *et al.* (2012) ID29: a high-intensity highly automated ESRF beamline for macromolecular crystallography experiments exploiting anomalous scattering. *J Synchrotron Radiat*, **19**, 455-461.
5. Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H. and Adams, P.D. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr*, **68**, 352-367.
6. Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Echols, N., Headd, J.J., Hung, L.W., Jain, S., Kapral, G.J., Grosse Kunstleve, R.W. *et al.* (2011) The Phenix software for automated determination of macromolecular structures. *Methods*, **55**, 94-106.
7. Zwart, P.H., Afonine, P.V., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., McKee, E., Moriarty, N.W., Read, R.J., Sacchettini, J.C. *et al.* (2008) Automated structure solution with the PHENIX suite. *Methods Mol Biol*, **426**, 419-435.
8. Douangamath, A., Fearon, D., Gehrtz, P., Krojer, T., Lukacik, P., Owen, C.D., Resnick, E., Strain-Damerell, C., Aimon, A., Abranyi-Balogh, P. *et al.* (2020) Crystallographic and electrophilic fragment screening of the SARS-CoV-2 main protease. *Nat Commun*, **11**, 5047.
9. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zidek, A., Potapenko, A. *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature*, **596**, 583-589.
10. Hough, M.A. and Wilson, K.S. (2018) From crystal to structure with CCP4. *Acta Crystallogr D Struct Biol*, **74**, 67.
11. Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W. *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr*, **66**, 213-221.
12. Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr.*, **60**, 2126-2132. Epub 2004 Nov 2126.

13. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem*, **25**, 1605-1612.
14. Gabadinho, J., Beteva, A., Guijarro, M., Rey-Bakaikoa, V., Spruce, D., Bowler, M.W., Brockhauser, S., Flot, D., Gordon, E.J., Hall, D.R. *et al.* (2010) MxCuBE: a synchrotron beamline control environment customized for macromolecular crystallography experiments. *J Synchrotron Radiat*, **17**, 700-707.
15. Bourenkov, G.P. and Popov, A.N. (2006) A quantitative approach to data-collection strategies. *Acta Crystallogr D Biol Crystallogr*, **62**, 58-64.
16. Delageniere, S., Brenchereau, P., Launer, L., Ashton, A.W., Leal, R., Veyrier, S., Gabadinho, J., Gordon, E.J., Jones, S.D., Levik, K.E. *et al.* (2011) ISPyB: an information management system for synchrotron macromolecular crystallography. *Bioinformatics*, **27**, 3186-3192.
17. Kabsch, W. (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *Journal of Applied Crystallography*, **26**, 795-800.
18. Evans, P.R. and Murshudov, G.N. (2013) How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr*, **69**, 1204-1214.
19. Bond, P.S. and Cowtan, K.D. (2022) ModelCraft: an advanced automated model-building pipeline using Buccaneer. *Acta Crystallogr D Struct Biol*, **78**, 1090-1098.
20. Joosten, R.P., Long, F., Murshudov, G.N. and Perrakis, A. (2014) The PDB_REDO server for macromolecular structure model optimization. *IUCrJ*, **1**, 213-220.
21. Laskowski, R.A., Jablonska, J., Pravda, L., Varekova, R.S. and Thornton, J.M. (2018) PDBsum: Structural summaries of PDB entries. *Protein Sci*, **27**, 129-134.
22. Krissinel, E. and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol*, **372**, 774-797.
23. Holm, L., Laiho, A., Toronen, P. and Salgado, M. (2022) DALI shines a light on remote homologs: one hundred discoveries. *Protein Sci*, e4519.
24. Holm, L. and Rosenstrom, P. (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res*, **38**, W545-549.
25. Robert, X. and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res*, **42**, W320-324.
26. Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H. and Ferrin, T.E. (2021) UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci*, **30**, 70-82.
27. Zheng, H., Cooper, D.R., Porebski, P.J., Shabalin, I.G., Handing, K.B. and Minor, W. (2017) CheckMyMetal: a macromolecular metal-binding validation tool. *Acta Crystallogr D Struct Biol*, **73**, 223-233.
28. Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T. and Ben-Tal, N. (2016) ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res*, **44**, W344-350.
29. Garman, E.F. and Mitchell, E.P. (1996) Glycerol concentrations required for cryoprotection of 50 typical protein crystallization solutions. *Journal of Applied Crystallography*, **29**, 584-587.
30. Jurenas, D., Rey, M., Byrne, D., Chamot-Rooke, J., Terradot, L. and Cascales, E. (2022) Salmonella antibacterial Rhs polymorphic toxin inhibits translation through ADP-ribosylation of EF-Tu P-loop. *Nucleic Acids Res*.
31. Koropatkin, N., Randich, A.M., Bhattacharyya-Pakrasi, M., Pakrasi, H.B. and Smith, T.J. (2007) The structure of the iron-binding protein, FutA1, from *Synechocystis* 6803. *J Biol Chem*, **282**, 27468-27477.

32. Shouldice, S.R., Dougan, D.R., Williams, P.A., Skene, R.J., Snell, G., Scheibe, D., Kirby, S., Hosfield, D.J., McRee, D.E., Schryvers, A.B. *et al.* (2003) Crystal structure of *Pasteurella haemolytica* ferric ion-binding protein A reveals a novel class of bacterial iron-binding proteins. *J Biol Chem*, **278**, 41093-41098.

Figure captions

Figure 1. Pipeline for protein structure determination by X-ray crystallography. Arrows in green and orange indicate user manipulation or automated procedure, respectively. Sizes of the arrows represent relative importance of automation versus human intervention.

Figure 2. Crystallisation screening. a) Manually fill MRC plates with the crystallisation screens using multi-channels pipette. b) Dispense nanodrops using Mosquito robot (STP Labtech). c) Place your plates in the RockImager system (Formulatrix). d) Follow crystal growth according to the defined schedule. Photos are from @CNRS/Virginie Gueguen-Chaignon and @CNRS/Vanessa Cusimano.

Figure 3. Example of screening and optimization of crystals. Pictures of crystal optimization obtained in four steps. a) Screening was performed in MRC 96-well plates with 300 μ l drops (150 μ l protein solution + 150 μ l reservoir). Needle crystals were obtained in PACT premierTM screen condition E9 (20% PEG3350, 0.2M Na K Tartrate) (maximum length 50 μ m). b) Crystals were improved by optimization in MRC 96-well plate with the screen PEGion F3 condition (12% PEG3350, 4% Tacsimate). c) Then crystals were optimized manually in 24-well plates in 3 μ l hanging drops containing 11-14 % PEG3350 and 3 -7% Tacsimate. d) Finally, larger crystals were obtained by using microseeding in 24-well plate. Photos are from @CNRS/Virginie Gueguen-Chaignon

Figure 4. Crystal cryo-protection and freezing

a) Prepare the material that you will need. From top to bottom: loops mounted on a magnetic base, and magnetic wand, the plate containing crystals, Uni-puck and cover, trays for liquid

nitrogen to accommodate pucks. Transport device for pucks. b) Incubate your crystals with cryo-protectant solution. Use the loop and magnetic wand to manipulate crystals. Transfer your crystals from the cryo-protectant solution to the liquid nitrogen. Be as fast as you can. Once all the crystals are frozen, you can place the puck filled and covered into the transport device and place it in the dry shipper or storage dewar. Photos are from @CNRS/Virginie Gueguen-Chaignon and @CNRS/Vanessa Cusimano.

Figure 5. Example of exploitation of DALI results

A search for homologue of the protein FutA1 (31) (PDB accession number 2PT1) using DALI gives the structure of the FbpA (PDB accession number 1Q35) as a significant hit. a) The two structures coordinates are uploaded and displayed as cartoon into UCSF Chimera. The FutA1 and the FbpA are coloured in cyan and magenta, respectively. Using Tools > Structure comparison > Matchmaker, the structures are superimposed b) The two structures are superimposed with a RMSD of 0.940 Å indicated in the dialog window. c) Using tools > Structure comparison > Match Align, you will generate a structure-based sequence alignment of the two models. d) Once the job is done, a window opens with the corresponding sequences aligned. e) Using this sequence alignment, you can generate publication quality pictures using ESPript 3.0 server. The picture will depict the alignment with secondary structures and conserved residues coloured in red. Here, the FbpA residues implicated in metal binding were identified by Shouldice *et al.* (32) and are Tyr 178, Tyr 198 and Tyr 199 (indicated by black dots). The alignment shows that the residues are conserved in FutA1. f) The two superimposed structures have be uploaded in Pymol and displayed as cartoon structures of FbpA and FutA1 are coloured in magenta and cyan, respectively. The side chains of Tyr 178, 198 and 199 (FbpA) and the side chains of Tyr 143, 199 and 200 (FutA1) are shown as ball-and-stick. The iron atom bound to FutA1 is shown as a grey sphere.

Figure 6. Example of the FutA1 CheckMyMetal server output

The holoprotein FutA1 (PDB accession number: 2PT2) is uploaded into the CheckMyMetal server. The left panel shows a ball-and-stick representation of FutA1 metal binding site (iron in orange, carbon in grey, nitrogen in blue and oxygen in red). The distances between the residues involved in the interaction and the metal are represented by dotted white line. Distances are specified in Angstrom. Residues are surrounded by the experimental $2f_o-f_c$ Map (blue wire). The right panel summarizes the characteristics of the protein - metal binding site evaluated by the CMM server. Each parameter is coloured according the accuracy of the value from « not applicable » to « acceptable ».

Table 1. High-throughput crystallisation screens routinely used for screening or optimisation.

Company	Reference	Kit name
<i>Screening</i>		
Hampton Research	HR2-116/HR2-117	Natrix 1 & 2
Calibre scientific	MDL-MD1-39	MemGold eco
Calibre scientific	MDL-MD1-63	MemGold2 eco
Hampton Research	HR2-082	PEGRx I
Hampton Research	HR2-084	PEGRx II
Calibre scientific	MDL-MD1-37 (eco)	JCSG+
Hampton Research	HR2-110/HR2-112	Crystal screen 1&2
Calibre scientific	MDL-MD1-106	MIDAS plus
Calibre scientific	MDL-MD1-104	BCS
Calibre scientific	MDL-MD1-29	PACT1er
Calibre scientific	NXT-130705	AmSO4
Calibre scientific	NXT-130706	MPD
Hampton Research	HR2-126/HR2-098	PEG/Ion
Calibre scientific	MDL-MD1-46	Morpheus
Calibre scientific	MDL-MD1-38	Proplex 1 & 2
<i>Optimisation</i>		
Hampton Research	HR2-138	Additives screen
Hampton Research	HR2-406	Detergent screen
Hampton Research	HR2-132	CryoPro
Calibre scientific	MDL-MD1-100	Angstrom additives
Calibre scientific	MDL-MD1-93	Morpheus additives
Calibre scientific	MDL-MD1-120	Durham additives

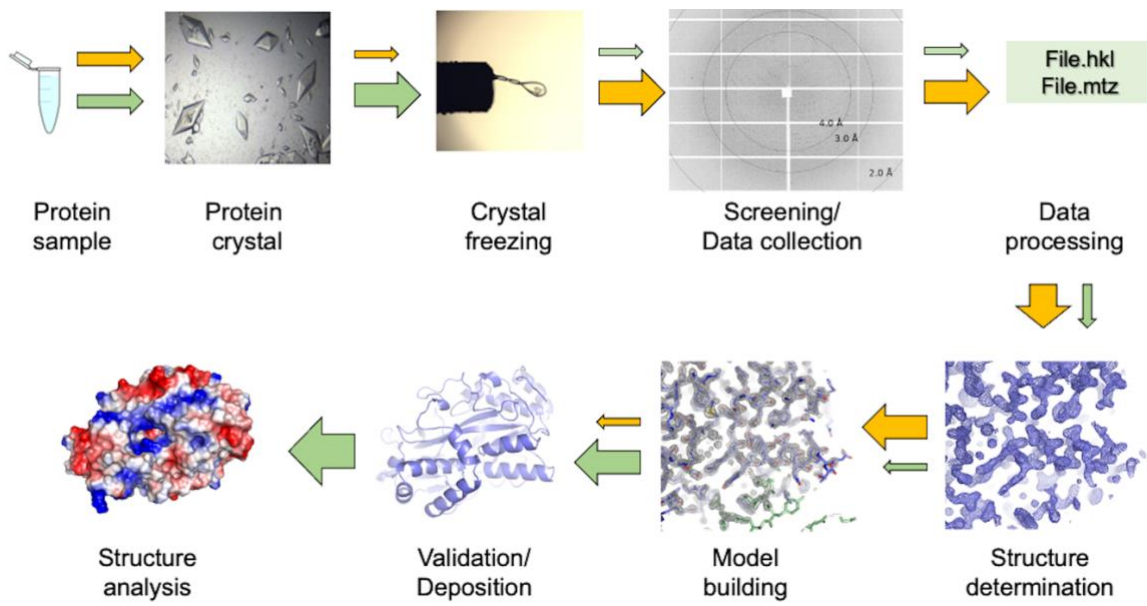


Figure 1

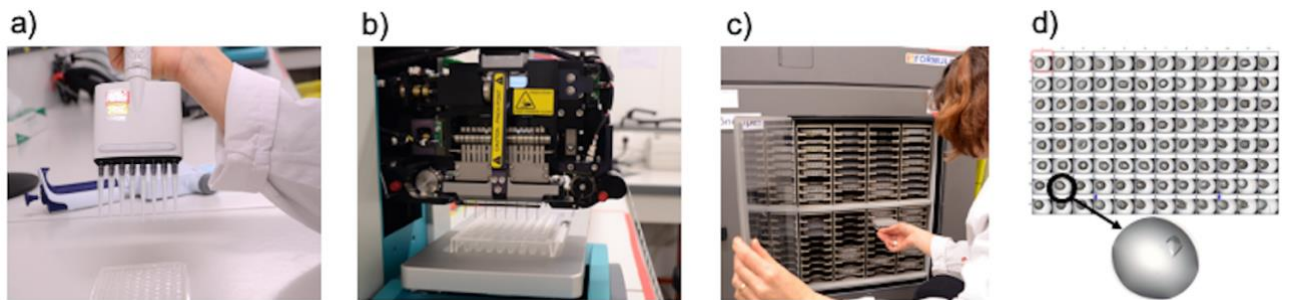


Figure 2

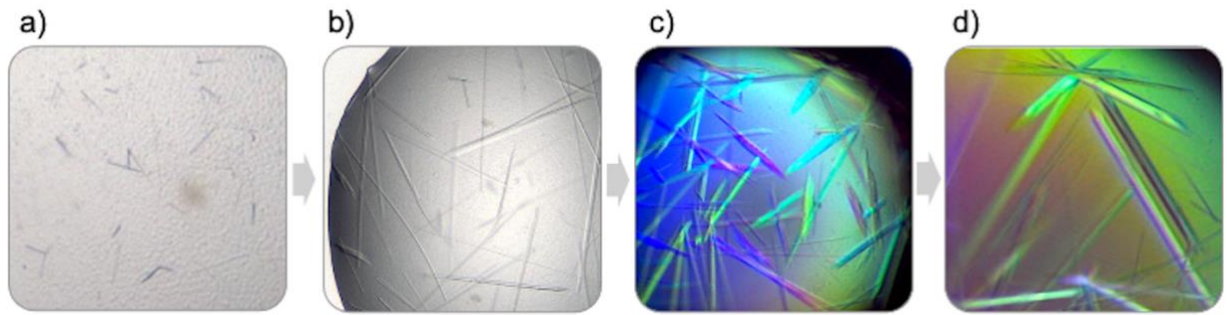


Figure 3

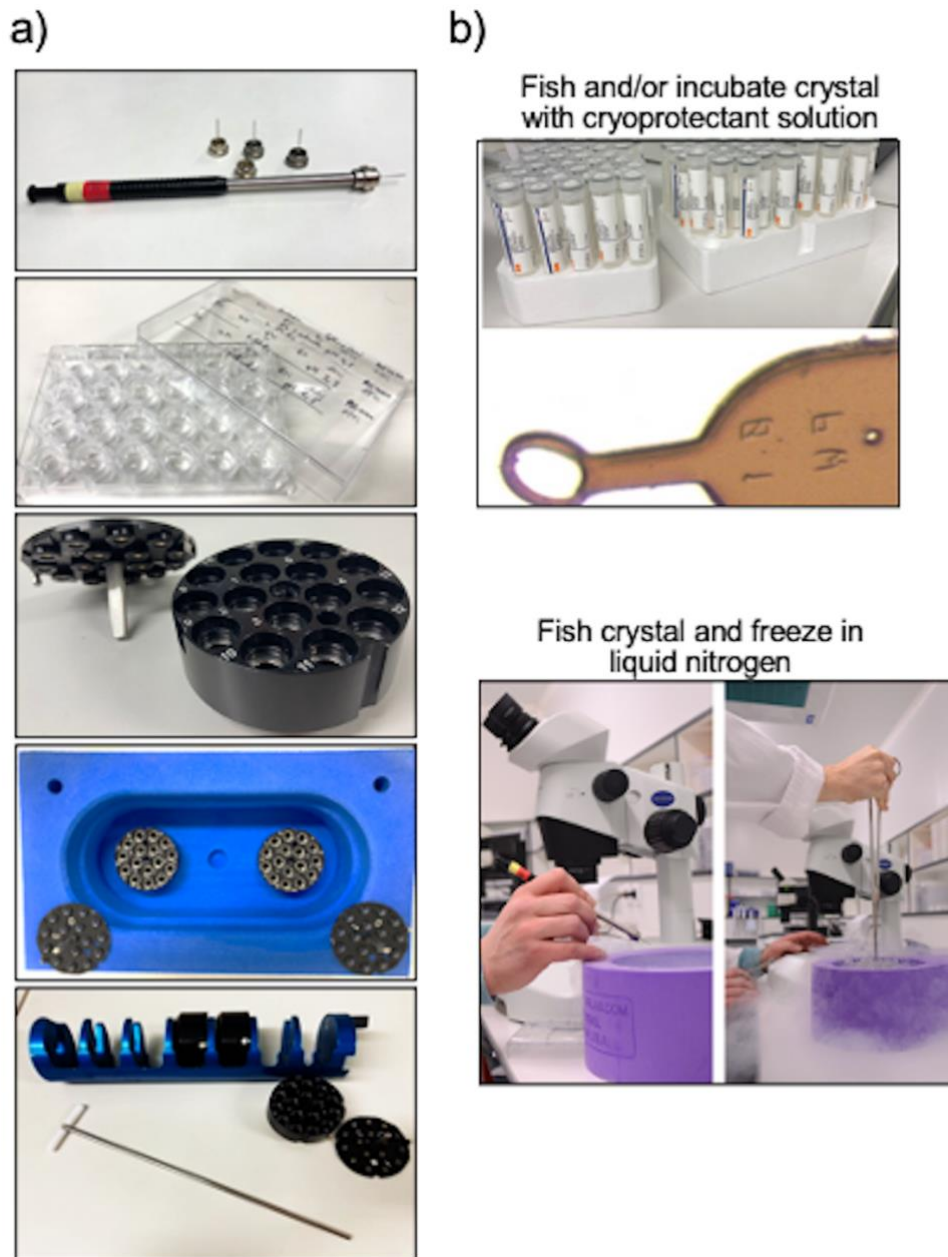


Figure 4

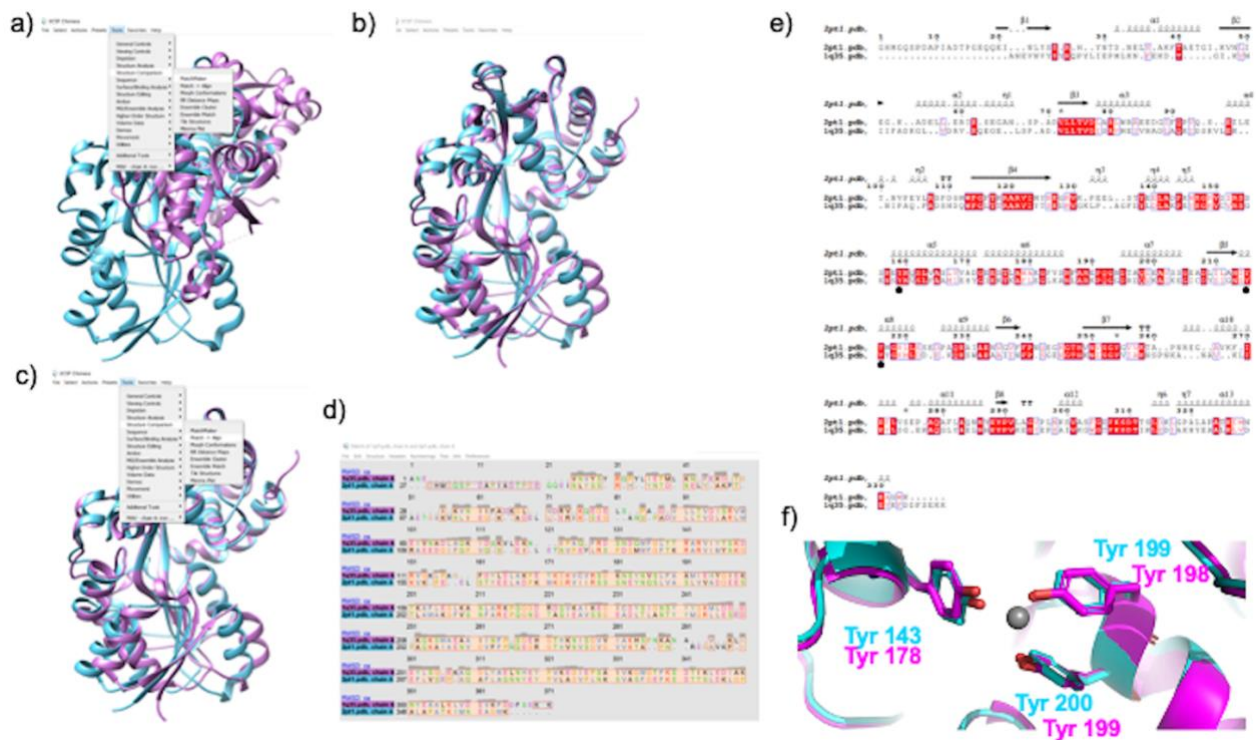
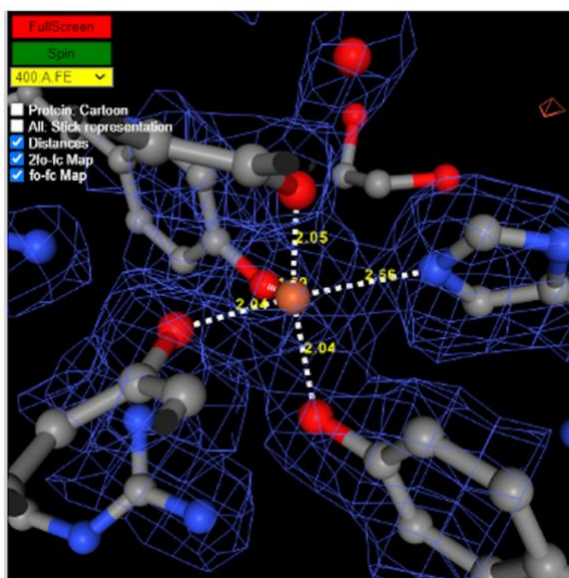


Figure 5



Metal	Fe
Occupancy	1
B factor (env.)	15,7 (15,2)
Atomic contacts	O ₄ N ₁
Valence	2,1
nVECSUM	0,2
Geometry	Trigonal Bipyramidal
gRMSD (°)	9
Vacancy	0
Bidentate	0

Legend (points in Proposed Metals):			
Not applicable (+0)	Dubious (-0.1)	Respective (1.1)	Acceptable (+2)

Figure 6