Tutor/s

Dr. Daniel Sainz García
Departament de Química Inorgànica

Dr. Roeland Boer
Sincrotró ALBA

Grau de Química

Treball Final de Grau

Crystallization and diffraction of single-crystals at ALBA Synchrotron to structure determination of a Relaxase protein.

Cristal·lització i difracció de monocristalls al Sincrotró ALBA per a la determinació estructural d'una proteïna Relaxassa.

Tibisay Guevara Puig

11 de Gener del 2016

Universitat de Barcelona

Campus d'Excel·lència Internacional
BKC Barcelona Knowledge Campus

"La perseverança és la virtut
per la qual totes les altres virtuts donen fruit."
Arturo Graf (poeta italià 1848-1913)

Abans de res vull agrair al Dr. Roeland Boer la oportunitat que m'ha donat de treballar amb ell a un centre com és el Sincrotró ALBA; també a l'Albert Castellví i l'Isidre Castro que m'han acollit al BioLab amb grans atencions i enorme generositat. Estic especialment agraïda a la Dra. Núria Valls, una gran amiga que sempre sap transmetre pau, seny y ordre en el moment precís.

Voldria remarcar la gran tasca del Dr. Daniel Sainz com a tutor durant tots aquests anys i em fa molta il·lusió tancar aquesta estapa de la seva mà: *Daniel*, s*in tus sinceras opiniones y consejos el camino no hubiera sido el mismo. Sin duda, has sido el mejor tutor que podía tener.*

L'espai en blanc que em queda es fa diminut. Hi ha tantes persones que han estat al meu costat regalant-me consells, experiències, temps o aguantant-me la plorera...un UN MILIÓ DE GRÀCIES A TOTA LA MEVA GENT. Dir-vos que m'encarregaré personalment de que tots tingueu el que us mereixeu, no és una amenaça...és una promesa.

Finalment agrair a la meva família...els meus incondicionals...la seva comprensió i el seu respecte. A partir d'ara només espero saber trobar la manera de cuidar-vos com vosaltres ho heu sabut fer amb mi. Y a ti Cristian: *las palabras se quedan cortas para agradecerte el "apoyo logístico", los abrazos y los ánimos que me han dado la energía suficiente para llegar al final.* Tots sou el meu motor.

No sóc una estudiant exemplar, no tinc una intel·ligència brillant ni tampoc cap capacitat extraordinària però tinc una companyia immillorable, sóc tossuda com ningú...i...

... ho he aconseguit!

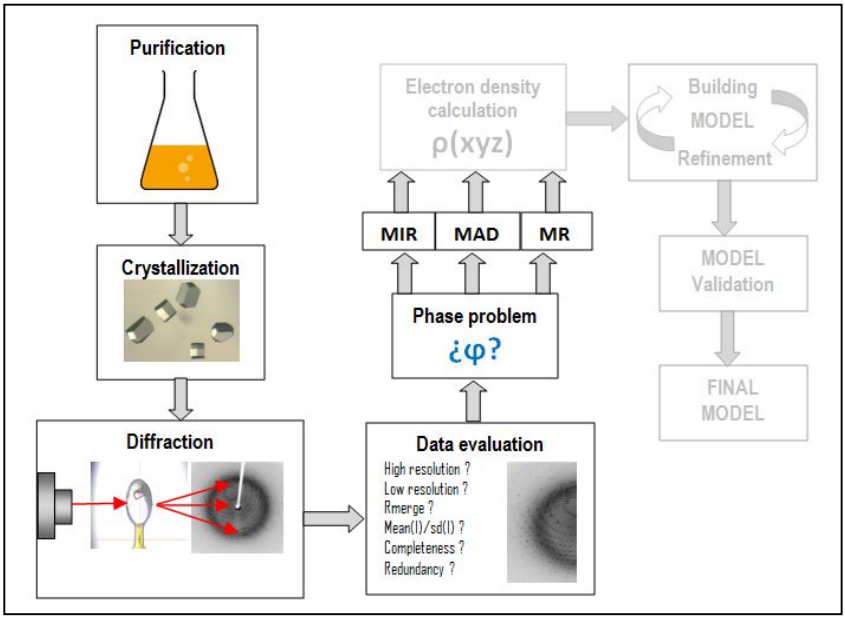
REPORT

CONTENTS

1. SUMMARY

The aim of the present work is the structure description of two variant (full length and N-terminal) of a protein with Relaxase activity. Both pure variants were crystallized using sitting drop vapor diffusion technique. Crystals obtained were optimized increasing its quality and some of them were treated to derivatize it with metals.

Diffraction of crystals and pre-treatment of protein were integrally performed in ALBA Synchrotron (Barcelona) installations. In BL-13 XALOC beamline were collected the data which were processed and evaluated as good quality data. The phase problem solution was not possible by applying Molecular Replacement nor detecting the anomalous signal of derivatives crystals. The lack of phases did not permit to solve the electron density equation, and therefore, it was not possible to complete the model building.



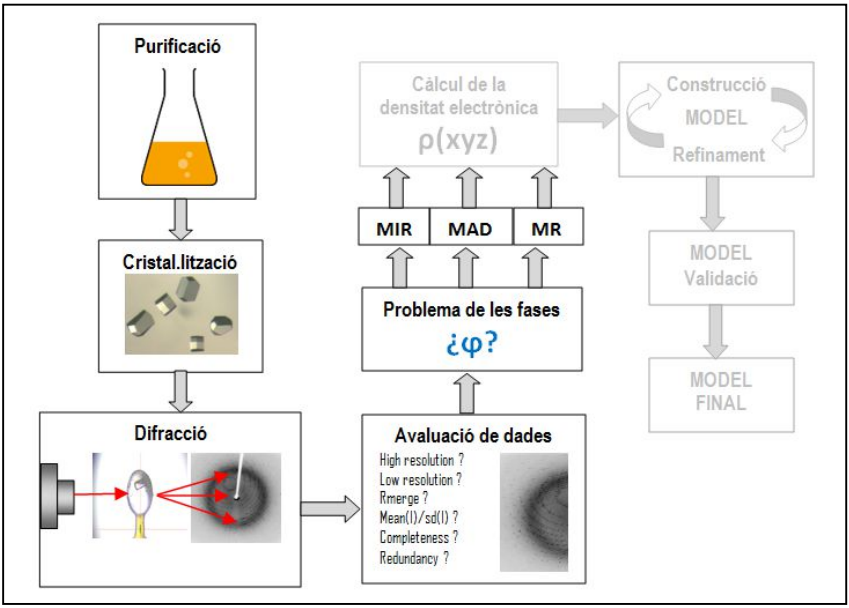
The figure describes in a simple manner the steps followed. Unsuccessful stages are indicated in light grey on which it is necessary to continue working.

Keywords: ALBA Synchrotron, BL-13 XALOC, structural model, crystallization, protein, diffraction.

2. RESUM

El treball que es presenta té com a principal objectiu descriure l'estructura de dues variants (seqüència completa i N-terminal) d'una proteïna amb activitat Relaxassa.

Es van cristal·litzar ambdues variants pures aplicant la tècnica de la difusió de vapor. La difracció dels cristalls així com el pre-tractament de la proteïna es va realitzar íntegrament a les instal·lacions del Sincrotró ALBA (Barcelona). A la línea BL-13 XALOC van ser col·lectades dades que es van processar i avaluar com a dades de qualitat. La resolució del problema de les fases no va ser possible aplicant Reemplaçament Molecular ni detectant la senyal anòmala dels cristalls derivatitzats. La manca de fases no va permetre resoldre l'equació de densitat electrònica i per tant la confecció del model no es va poder completar.



La figura descriu de forma simplificada les etapes realitzades. En color gris clar, s'indiquen les etapes sense èxit i sobre les que cal seguir treballant.

Paraules clau: Sincrotró ALBA, BL-13 XALOC, model estructural, critsal·lització, proteïna, difracció.

3. INTRODUCTION

3.1. PROTEINS: FUNCTION AND STRUCTURE

Proteins are the cellular molecules which are responsible for the cell activity and are codified by genes. This activity requires the coordination of a lot of different protein types.

Although proteins are made by linear chains of just 20 different amino acids, they carry out an incredible number of diverse tasks.

A protein chain folds to acquire a unique shape, stabilized by noncovalent interactions between regions of the linear amino acid sequence. This three dimensional organization of the protein, the conformation, is the basis to understanding its function. Only when a protein is in its correct three dimensional structure, it can function efficiently.

The proteins function derives from the tridimensional structure and this is established by the amino acids sequence, which is the key to comprehend how proteins work.

3.1.1. How to study protein structures

Nowadays, several methods are used to study protein structures. Every methodology has advantages and disadvantages, and sometimes it is necessary to combine some of them to obtain a more complete information.

The creation of an atomic model needs a combination of many pieces of information. These pieces can be obtained principally using three different techniques: NMR Spectroscopy, Electron Microscopy and X-ray Crystallography.

X-ray Crystallography requires the crystallization of the protein. Then the crystal is submitted to an intense X-ray beam (2.5-0.55 Å wavelength). The molecules that form the crystal diffract the X-ray beam in a characteristic spots pattern. The intensities of the diffracted spots are required to calculate the electron distribution in the protein. The obtained electron density map is then interpreted to determine the position of each amino acid.

X-ray crystallography can provide very detailed atomic information but the crystallization process is difficult and it can limit the sort of proteins that can be studied by this method. It is a good method to study rigid proteins but has inconvenience for proteins with mobile parts in their structure.

3.2. CRYSTALLIZATION

For protein crystallization, it is essential to have a stable, proper concentrated and pure enough protein sample. This is often a difficult task, because proteins are easily affected by their environment and often are especially delicate when handling them.

The difficulty in controlling crystallization is based on the weak interactions that occur between the irregularly shaped and flexible protein molecules in order to self-assemble into a regular periodic crystal lattice.

In literature, crystallization is considered more an art than a science because crystallization tips are often based on anecdotal evidence and single events. At the same time, there are not too many rules that can be generally applied in practice.

Given the impossibility to know in advance the crystallization condition and the extremely wide range of potential crystallization experiments, the most common strategy is to tackle the problem in a systematic trial-and-error manner, as will be discussed later.

3.2.1. Crystals

Crystals are formed by elementary units (unit cells) placed regularly, following a determined scheme that is reproduced, in shape and orientation, in the whole crystal, thus forming a tridimensional lattice, the crystal structure.

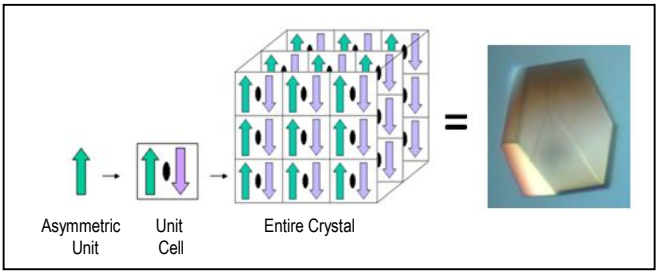


FIG.1. Relation between asymmetric unit, unit cell and the entire organization in a crystal

(Modified from free PDB material, Ref.19; Lysozyme Crystal picture (T.Guevara))

The asymmetric unit is the smallest portion of a crystal structure to which symmetry operations can be applied in order to generate the complete unit cell, the crystal repeating unit.

A crystal asymmetric unit may contain one biological assembly, a portion of a biological assembly or multiple biological assemblies. This content depends on the oligomerization state of the molecule that has crystallized and its conformations within the unit cell.

Protein single crystals often are very beautiful but, unfortunately, external appearance is not necessarily an indicator of absence of internal structural defects. Many protein crystals exhibit low quality diffraction or even no diffraction.

The following section explain a general strategy to obtain crystals and optimize it to improve its quality in order to obtain good diffraction data.

3.2.2. Crystallization strategy

To crystallize a protein, the protein solution needs to be brought to supersaturation, at which point the protein will either precipitate or crystallize, depending on the components of the condition crystallization solution. As was said, a good strategy to start crystallization experiments is to systematize the search for a suitable crystallization conditions. In general, every condition is composed by a precipitant, a salt and a buffer but it can present variations.

Some commercial crystallization kits (screenings) are available, which seem to be organized randomly, but most of these are in fact based on the experience of the crystallizing community over many years. Others are systematic combinations of precipitants, salts and buffers.

Precipitants	Salts	Buffers
Peg (MW 100-2000)	$(NH_4)_2SO_4$	TRIS
$(NH_4)_2SO_4$	HCOONa	Sodium citrate
2-propanol	CH_3COOK	HEPES
Jeffamine ED-600	NaF	MES

TAB.1. Some examples of compounds type which form crystallization conditions.

Crystallization experiments are set using nanoliter volumes of the protein solution, requiring high precision nano-dispensation robot. This equipment permits to screen lots of conditions with

a minimum quantity of protein and, at the same time, ensures high reproducibility, which is a basic concept to consider in crystallization.

Crystallization plates (96 well), screenings (96 conditions) and robot (96 dispensation tips), are all specially designed for high throughput work. Using 0.5 ml of protein it is possible to setup more than 3.500 crystallization experiments.

Plates set need to be incubated at a stable temperature, which is another important parameter in crystallization experiments.

Periodically, this plates are evaluated under a mycroscope to find crystals or some promising result susceptible to be optimized to obtain a good monocrystal. Inspection of crystallization experiments is one of the most tricky tasks, because it is important to know how to discriminate and to interprete correctly what it is being seen but there are no general rules to identify a Crystal. As a result, experience is the best tool.

When a hit is identified, optimization and scale up are necessary to obtain a good number of diffraction-quality crystals. This step implies the systematic screening of variations on the initial parameters using higher volumes of the protein solution. Generally, scaling up crystallization experiments are 10 times bigger than the ones set using the robot.

Optimization can be a difficult task because the possible variations to be introduced to the crystallization system with this aim can be infinite (protein concentration, protein buffer, incubation temperature, pH, crystallization solution composition and concentration, crystallization technique, plate type,...)

3.2.3. Vapor diffusion technique

There are different crystallization techniques to obtain quality crystals, some based on the use of different stands and others based on diverse principles. Vapor diffusion technique is the most commonly used.

The principle of vapor diffusion is simple. A drop composed of a mixture of protein and crystallization condition is placed in vapor equilibration with a liquid reservoir of condition. Typically the drop contains a lower reagent concentration than the reservoir. To achieve equilibrium, water vapor leaves the drop and eventually ends up in the reservoir.

As water leaves the drop, the sample undergoes an increase in relative supersaturation. Both the sample and condition increase in concentration as water leaves the drop for the reservoir.

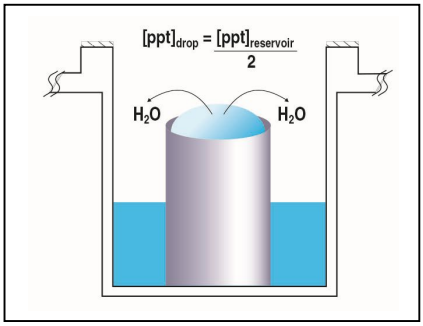


FIG 2. Vapor diffusion technique scheme. When the system is sealed, concentration in drop starts to increase due to concentrations difference in between drop and reservoir.

(Kindly loaned by Hampton Research company; Ref.16)

Equilibration is reached when the reagent concentration in the drop is approximately the same as that in the reservoir.

Once the solubility limit of protein is exceeded, the solution is supersaturated. This creates a metastable state where the system is not in equilibrium and in front a nucleation event, the excess protein molecules separate from solution by forming a crystal.

Solubility phase diagram plots the protein solubility changes with precipitant concentration and show, in a simplified way, the different situations that can occur during a crystallization experiment.

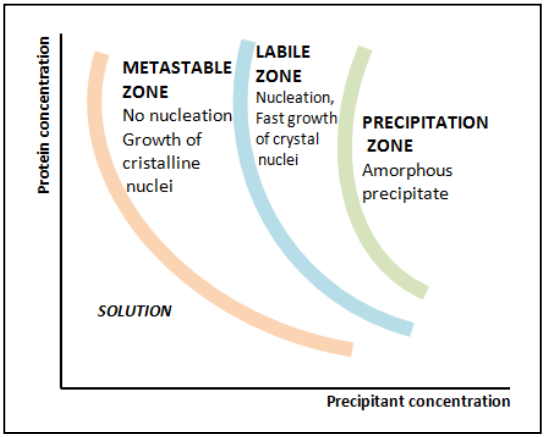


FIG.3. Phase Diagram *(Inspired in Ref.18)*

The purpose of crystallization screening and subsequent optimization is to find a condition at which the system is in the right zone of the phase diagram.

3.2.4. Crystals preparation to diffraction

Protein crystals are very small, ranging in size between few 100 μm to about 10 μm, and due to the high solvent content and weak, noncovalent intermolecular interactions they are very fragile and susceptible to radiation damage under the intense X-ray beams.

Cooling the crystal to cryogenic temperatures minimizes the radiation damage during data collection.

During the crystals flash-cooling under liquid nitrogen (-196 ºC), crystalline ice can be formed and destroy the delicate protein crystal, so the addition of cryoprotectants is needed.

Ethylene Glycol, Glycerol, Peg, high salt concentrations and Sucrose are some of the typical cryoprotectants. These compounds, added to the crystal, minimize the water content and vitrify the remainder (forming non crystalline structure) avoiding damaging ice crystals.

Handling, cryoprotectant solutions and protocol must be optimized to obtain a well prepared crystal to be diffracted. This step is determinant to good data collection.

Protein crystal characteristics
Composed ~50 % solvent on average, it can range from 25 to 90 %
Large pores in its internal structure. Small solvent and molecules can diffuse freely into these pores
Molecules form a relatively small number of bonds with their neighbors
Mechanically fragile, soft and easy to crush
Sensitive to dehydration
Sensitive to temperature changes
Sensitive to radiation damage

TAB.2. Protein crystals handling and freezing must to be done with special care and considering the special characteristics of it. It is usual to break and dehydrate accidentally some crystals in the first attempts.

3.3. RADIATION SOURCE

3.3.1. ALBA Synchrotron

ALBA is the Spanish synchrotron light source. It is a complex of electron accelerators to produce synchrotron light, which allows visualization and analysis of matter and its properties at atomic and molecular levels. It is in operation since May 2012 and has seven beamlines which are able to perform experiments in different scientific fields.

FIG.4. ALBA Synchrotron is a peculiar building with 270 m of circular perimeter located in Campus of Universitat Autònoma de Barcelona. *(ALBA intranet, material available for students)*

A synchrotron is a particle accelerator where magnetic and electric fields are combined to obtain radiations of high energies from electrons acceleration in a closed and circular trajectory.

In ALBA, electric fields are used to give energy to the electrons. Trajectory of electrons is bended by magnetic fields produced by 32 dipole magnets placed along the round and 112 quadrupole and hexapole magnets focalize the beam. When electrons are deflected to follow a circular trajectory they emit synchrotron radiation. This radiation is emitted in a linear trajectory to the experimental hutch while electrons remain inside the storage ring. Before the beam reach the sample there is some equipment which select and condition the ray to specific characteristics of the experiment.

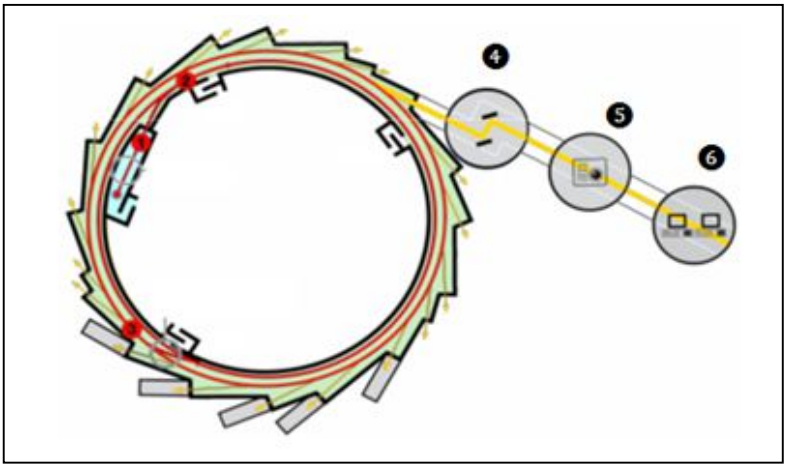


FIG.5. Scheme of ALBA electron accelerator. The 3 GeV electron beam energy is achieved by combining a Linear Accelerator (LINAC) **(1)** and a low-emittance, full-energy BOOSTER ring **(2)** placed in the same tunnel as the STORAGE RING **(3)**. There are seven beamlines completely operative. Each line is composed by conditioning beam station **(4)**, experimental hutch **(5)** and control room **(6)**.

(ALBA intranet, material available for students)

3.3.2. XALOC beamline

All protein crystals diffracted in the present work were diffracted at BL13-XALOC beamline. This beamline is devoted to the structural determination of macromolecules and macromolecular complexes using X-ray crystallography. The beam is fully tunable in 5-22 keV (2.5-0.55 Å) and the experimental hutch is equipped with diffractometer with high-accuracy oscillator axis, Pixel detector Pilatus 6M, automatic sample changer (90 crystals capacity), fluorescence detector, accurate humidity and temperature control and more other elements which characterize the beamline as a high technology installation.

The wet-lab part of the reported work was performed in the Biological Lab. This space is equipped with the necessary hardware to express, purify and crystallize proteins for in-house research and at the same time offer facilities to beamlines users in their samples preparation.

FIG.6. Experimental Hutch of XALOC beamline. The detector is where the diffraction pattern is collected, the diffractometer has all the elements to move and keep the crystal in good conditions under the beam during collection and the sampler changer permit to handle the crystals from control room.
(ALBA intranet, material available for students)

3.4. CRYSTAL STRUCTURE SOLUTION

3.4.1. Symmetry and diffraction pattern

When electrons are irradiated with X-rays part of the photons are scattered elastically. When crystals are illuminated with X-rays, the constructive and destructive interferences between elastically scattered photons will give rise to secondary (diffracted) beams. These diffracted beams can be recorded on a suitable detector, which is called a diffraction pattern. It consists on many spots which correspond to the diffracted rays (reflections).

From the diffraction images, the intensities of the diffracted beams must be determined. As a first step, the unit cell has to be identified from one of the possible 14 Bravais lattices (see Appendix 1).

Bragg's law is a mathematical model that explains the geometry of the incoming and diffracted X-rays by representing a crystal as a series of parallel mirrors which cut the unit cell in equal sections (see Appendix 2). These mirrors reflect X-ray and the reflections which satisfy Braggs law in a certain crystal orientation can be measured on the detector.

To unequivocally identify of each spot, Miller indexes are assigned. These are set of three numbers represented by (h k l) that are associated with a set of parallel (imaginary) mirrors.

It is important to keep in mind that a complete data set can be formed by thousands of spots (in the order of 10^4), so the quantity of information that is necessary to manage to solve a structure is huge.

In addition to the unit cell, a space group, compatible with the Bravais lattices (see Appendix 1), has to be assigned. There are 230 space groups but due to enantiomorphic characteristics of L-amino acids which form proteins, just 65 of them are possible for proteins.

The information on the intensity of the diffracted X-ray waves allows the calculation of the electron density within the crystal. The calculation is performed using the Fourier Transform (FT), where each diffracted wave represents one of the terms of it.

3.4.2. Electronic density equation

To calculation of the electron density at each point (x,y,z) in the unit cell is done using the following mathematic function.

$$\rho(xyz) = \frac{1}{V}\sum_{\substack{hkl \\ -\infty}}^{+\infty} |F(hkl)| \cdot e^{-2\pi[hx + ky + lz - \varphi(hkl)]}$$

EQ.1. Electron density equation

ρ(xyz): electron density

F(hkl): Structure factor, it is the intensity of the reflections

h,k,l: Miller indexes

φ(hkl): phase, these are the relative phases of the diffracted waves

This equation represents the Fourier Transform of the real space (where the atoms are), represented by the electronic density (ρ), from the structure factors (intensity and phases) which represent the reciprocal space (diffraction pattern).

The equation explains the 'holistic' character of diffraction, meaning that the electron density in one position (xyz) has a contribution from all the structure factors.

The intensity of the structure factors F(hkl) are experimentally obtained from the diffraction images, but there is no technique to experimentally determine the phases directly from the diffracted waves. This is called the Phase Problem.

3.4.3. The solution to the Phase Problem

When an analogous structure of the target protein is known, Molecular Replacement (MR) permits to use it as an initial model that later must be adjusted (refinement step). This procedure is based on the observation that proteins with homologous amino acid sequences show very similar folding.

The difficulty in this case is to find an orientation and position of the model protein that is compatible with the measured data. In this case, is not necessary to specifically determine phases because the position of the known model provides an initial estimate of the phases, which can then be used to adjust the model to fit the experimental data better.

When there's no known homologous protein and structure must be solved from scratch, it is necessary to prepare crystals derivatives that incorporate high atomic number atoms (heavy atoms: Au, Ni, Zn, Pt,...) which disperse X-rays more efficiently and therefore have a bigger effect on the intensity of the structure factors of the derivative structure compared to the native diffraction pattern.

These differences permit the determination of the heavy atoms positions and, finally, the calculation the structure factor (F(hkl)) and phases (φ(hkl)) of the protein structure. This is called Multiple Isomorphous Replacement (MIR) method.

Multi-wavelength Anomalous Diffraction (MAD) is other method based on intensity changes of reflections caused by absorption of the incident X-rays by the heavy atoms. Using different wavelengths and evaluating the differences of the data is possible to calculate amplitudes (structure factor) and phases.

Finally, each method gives a preliminary complete set of information to apply an iterative procedure to build the model.

It is important to know that quality of the model determined is completely dependent of the quantity and quality of the data collected. Bad data give a bad model, but also with good data, it is possible to build an incorrect model.

To control the quality of both, data and model, it is important to evaluate the quality indicator parameters obtained in each step, for instance Rmerge, Mean(I)/sd(I), completeness for data and Z-score for the model.

3.5. INTEREST OF THE PROJECT

The studied protein, P58, is a Relaxase, this type of proteins have an important role in the DNA processing during conjugative cell-to-cell DNA transfer. This kind of mechanisms give the bacteria cells the possibility to acquire antibiotic multiresistance and the description of its structure can suggest novel strategies for therapeutic intervention.

The structure of P58 protein will permit to know and describe how it takes part in the complex process of bacterial conjugation. In literature, it has been reported that Relaxase activity is localized at the N-terminal domain. For all of this, the present work includes both variants of P58 protein, N-terminal (N-term) and Full Length (FL).

The experimental part of the present work starts with the E.coli bacteria cells properly grown to express and contain the protein to study. This bacteria cultures were performed by an experimented Biochemist in Biological Lab of ALBA Synchrotron.

4. OBJECTIVES

The aims of the present work are:

-The obtaining of P58-FL and N-term variants quality crystals using different crystallization techniques and the optimization strategies necessary.

-Get the best possible data diffracting the crystals with synchrotron radiation at XALOC beamline in ALBA Synchrotron.

-The study of how to manage the data and which are the strategies to follow to solve a protein structure.

- Make hypotheses about P58's function based on the solved structure.

5. EXPERIMENTAL SECTION

5.1. PROTEIN EXTRACTION, PURIFICATION AND CONCENTRATION

The following steps are the same for obtaining the FL protein and the N-term variant.

The recovery of the protein from the inside off the cells requires the lysate of them. The cells were resuspended in lysis buffer (see Table 3) in a ratio of 5 ml/g of pellet cells and they were softly mixed during 30 min at 4 ºC. The mixture was sonicated at 4 ºC setting the power at 200 W with a VibraCell 75041 (Bioblock Scientific) during 12 cicles of 10 s ON followed by 59 s OFF.

The lysate was poured in Oak Ridge centrifuge 50 ml tubes (Nalgene) and centrifugated in a Megafuge 40R centrifuge (Heraeus) at 4 ºC during 20 min at 13.000 xg speed using a F15 (Heraeus) carbon fiber rotor.

The protein of interest was contained in the supernatant and the pellet were discarded. Before purification the supernatant was filtered across 0.22 µm PES syringe filter (Fisher).

A 5 ml Hi Trap Chelating HP column (GE Healthcare), charged with Nickel (see Appendix 3) and connected to an Econogradient Pump (BioRad) was cleaned with 5 column volumes (CV) with water to eliminate the ethanol of the storage solution. The equilibration of the column with Binding buffer (see Table 3) was done with 10 CV. These operations were performed at 5 ml/min of flow rate.

The lysate supernatant is loaded into at 3 ml/min onto the column and the flow through was collected in a tube. The column was washed with 20 CV of Binding buffer at a flow rate of 3 ml/min. The wash was collected in a different tube.

To elute the protein, the pump was programmed with a gradient from 0% to 50% Elution buffer (see Table 3). The buffers mixture was done by a Gradient Mixer (BioRad) connected in line before the column. The run was performed at 3 ml/min and 2 ml fractions were collected along the gradient with a Fraction Collector F-200 (Pharmacia).

EXTRACTION AND PURIFICATION BUFFERS COMPOSITION	
Lysis buffer	50 mM Tris pH8, 5 % Glycerol, 0.5 M NaCl, 1 mM EDTA, 20 mM Imidazole, 1 mg/ml Lysozime, Spatula tip of DNAse
Binding buffer	20 mM Tris pH8, 0.5 M NaCl, 5 mM Imidazole
Elution buffer	20 mM Tris pH8, 0.5 M NaCl, 1 M Imidazole

TAB.3. Composition of buffers used to extract and purify the proteins.

The absorbance of each fraction was measured with NanoDrop spectrophotometer (Thermo). Spectrophotometric measures were done at 280 nm wavelength. The chromatogram was made plotting the Absorbance vs. Fraction number.

To identify the fractions containing the pure protein, two SDS-PAGE of 15 % acrylamide (see Appendix 4) were prepared (see Table 4). 5 µl of PageRuler™ Prestained Protein Marker (Invitrogen) was loaded to identify molecular weight of the proteins in sample.

16 µl of fractions were mixed with 4 µl of loading buffer (see Table 4) and loaded into the gel which was run in Verti-Gel Mini electrophoresis system (Fisher) with Anode and Cathode buffers (see Table 4).

The gel run was performed, during 1 h and 30 min, connecting the electrophoresis cuvette to a power supply Consort EV243 (Broser) and setting the voltatge to 400 v. The resulting gel was stained with a Coomassie dye solution (see Table 4) during 15 min with soft agitation using a See-Saw rocker shaker (Stuart) and then distained with Distain solution (see Table 4) shaking it again during 15 min.

Based on the gel, fractions containing the protein were identified and joined to quantify the pure protein.

The quantity of protein was obtained measuring the absorbance at 280 nm and following the Lamber-Beer Law using the protein extinction coefficient. Protein extinction coefficient was obtained using an Protparam (online tool of ExPAsy site (see Appendix 5)) using the protein sequence.

The fractions quantified were concentrated with a concentrator device Amicon Ultra-4 30 kD MWCO (Millipore) using a Megafuge 40R centrifuge (Heraeus) with 750067 swinging rotor with

15 ml tubes adapter (Termo). The buffer was exchanged to 20 mM Tris pH 7.5 and 100 mM NaCl using a PD-10 Desalting column (GE Healthcare).

A PCT Pre-Crystallization Test (Hampton Research) was used to evaluate if the concentration is optimal to crystallize (see Appendix 6).

ELECTROPHORESIS ELEMENTS COMPOSITION	
SDS-PAGE gel	Stacking: 0.67 ml Proteogel, 1 ml Buff.C, 2.3 ml H_2O, 30 µl PSA (10 %), 5 µl TEMED Resolving: 6 ml Proteogel, 3 ml Buff.B, 3 ml H_2O, 72 µl PSA (10 %), 6 µl TEMED
Buffer B	1.5 M TRIS pH 8.8, 0.4 % SDS
Buffer C	0.5 M TRIS pH 6.8, 0.4 % SDS
Loading buffer	0.125 M TRIS, 7.5 % SDS, 25 % β-Mercaptoethanol, 50 % Glycerol
Anode buffer	0.02 M TRIS
Cathode buffer	0.01 M TRIS, 0.01 M Tricine, 0.01 % SDS
Stain sol.	45 % Methanol, 10 % Acetic Acid, 0.1 % Coomassie R-250
Distain sol.	45 % Methanol, 10 % Acetic Acid

TAB.4. Composition of buffers used to prepare SDS-PAGE gel and samples for electrophoresis analysis.

5.2. CRYSTALLIZATION

5.2.1. Initial sparse/matrix Crystallization Screens and optimization

A similar procedure was followed for the crystallization experiments on the P58FL protein and the P58N-term domain. A high precision nano-dispensation robot located in Applied Proteomics and Protein Engineering Unit from UAB (Art Robbins) was used. The drop size was set to 200 nl of protein plus 200 nl of the crystallization buffer.

Crystallization experiments were set on MRC 96 well Crystallization Plate™ (Molecular Dimensions) and the conditions screened were commercial kits of 96 conditions: Index and

Peg/Ion from Hampton Research; Structure, Morpheus, JCSG and PACT from Molecular Dimensions. The plates were sealed with OptiClear Seals for MRC Plates from Molecular Dimensions.

Plates were placed in incubators (Comersa) at 18 ºC and the appearance of crystals was evaluated during the following weeks by visual inspection using a NZ12 binocular microscope (Leica).

Depending on the evaluation and the results obtained in initial crystallization screen plates, optimization screens around successful conditions must be designed and set.

5.2.2. Crystal handling and freezing

All the crystal handling was done using as a spoon a 18 mm Mounted CryoLoop™- 20 µm (Hampton Research) mounted in a CryoCap (Molecular Dimensions). The size of the chosen CryoLoop must be adequate to the crystal size.

The last step is performed placing a CryoVial under liquid nitrogen in a Dewar vessel, then, the CryoLoop with the crystal were vertically inserted to the Cryo Vial taking care not to hit the Crystal to the vial wall.

At this moment the crystals were inserted in ESRF/EMBL Sampler Changer Basket (Molecular Dimensions) and stored under liquid nitrogen in a dry shipper CX100 (Molecular Dimensions) until the moment of being diffracted.

5.3. DATA COLLECTION AND PROCESSING

All data were collected from frozen crystals at 100K with PILATUS M6 detector at XALOC beamline (ALBA Synchrotron, Barcelona)

The CryoVials with crystals were placed in a sampler changer robot (under liquid nitrogen), manufactured by IRELEC. From the control room, the crystals were mounted with the sample changer arm on the goniometer using in-house software .

The crystal were centered in the beam trajectory and diffracted. Initially were checked two images of the diffraction pattern, positioning the crystal at 0º and 90º. These were evaluated by the diffraction characterization program EDNA.

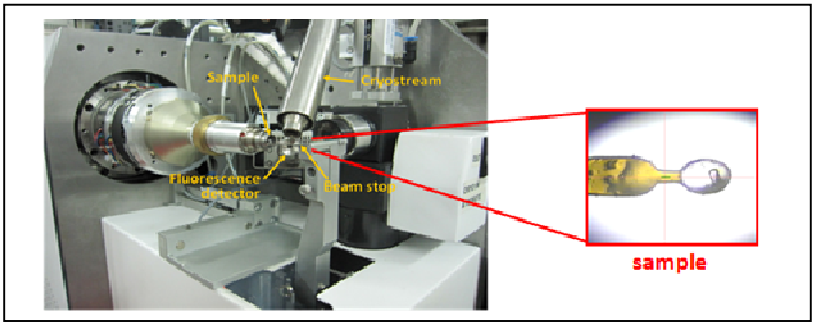


FIG.7. Sample placed in the diffractometer

(adapted from ALBA intranet, material available for students)

Crystals which gave indexable images (i.e. assignment of the Miller indexes) were collected following the strategy suggested by EDNA (degrees, images, detector distance,...).

Indexing of data was done with imosphilm program. The evaluation of statistical parameters of imosphilm (for example: Rmerge, I/sigma(I), completeness) was done to check quality of the data. The indexes list is converted with an CCP4 interface program in a ".mtz" file to apply Molecular Replacement method.

5.4. SOLUTION OF PHASE PROBLEM

5.4.1. Molecular Replacement (MR)

Imposing the sequence of the studied protein in the alignment tool from PDB were found some protein sequences with high identity that are possible models to apply MR.

The homologous protein model (.pdb file) and the reflections file (.mtz file) were used to apply MR method using Phasing program. Z-score, Rcryst, Rfree and other geometrical parameters of the model were checked to assess the quality of the model determined (see Table 5).

5.4.2. Scanning of anomalous signal

Crystals soaked with metal compounds were placed, as natives ones with the robotic arm on the goniometer. An energy scan was performed with the in-house software on the crystal and

the emitted fluorescence was plotted to determine a maximum which correspond to a maximal absorbance. This peak must be characteristic of the metal which is supposed in the crystal.

TF Z-score	Structure solution
<5	no
5-6	unlikely
6-7	possibly
7-8	probably
>8	definitely

TAB.5. Values of Z-score determine if MR permit to solve the structure with the model and data used. Z-score is an statistical parameter which represent the number of standard deviations over the mean. *(Adapted from Ref.22)*

6.RESULTS AND DISCUSSION

6.1. EXTRACTION, PURIFICATION AND CONCENTRATION

6.1.1. Full length variant

6 g of cells from 1 liter culture were lysate and purified obtaining 34 mg of full length protein. The chromatogram profile show the fractions which content more total protein (see Figure 8). In fractions from 15 to 25 there is a small peak overlapping with the largest. SDS-PAGE gel show the fractions which accumulate more FL protein and less contaminants.

Fractions from 26 to 48 were joined to concentrate, in these, FL protein is majority as was checked by gel. The gel bands (See Figure 9) shows in the small peak there is some higher molecular weight proteins mixed with less quantity of FL protein.

In final concentration step, absorbance (280nm) measure value was 19.6 and using the theoretical molar extinction coefficient (70250$M^{-1}cm^{-1}$) the concentration was determined in 14 mg/ml.

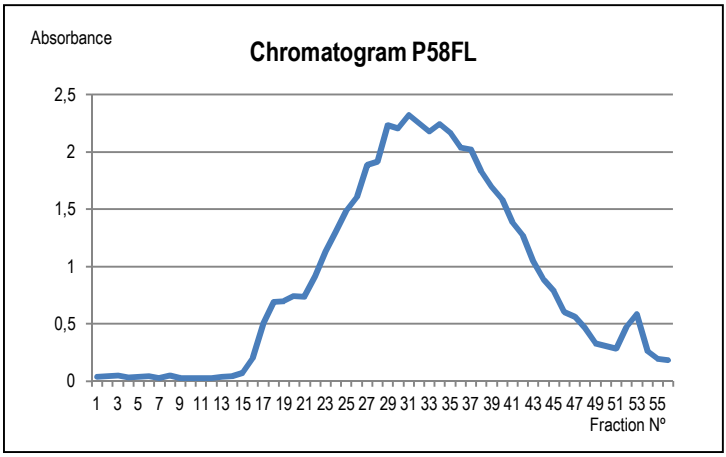


FIG.8. P58FL purification chromatogram

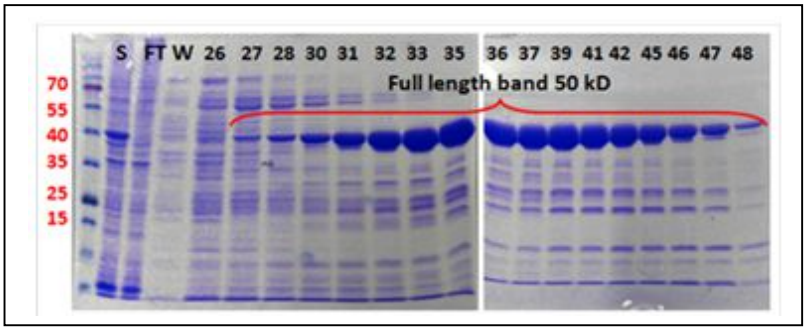


FIG.9. P58FL SDS-PAGE purification gel

In red on the left: MW markers (in kD) S: sample before purification, FT: flow through; W: wash, 26-48: collected fractions. FL protein has a 50 kD MW but it migrates as 40 kD marker maybe due to its shape.

Approximately 250 µl of P58FL at 14 mg/ml in 20 mM Tris pH 7.5 and 100 mM NaCl was finally obtained.

The Pre-Crystallization Test showed that 14 mg/ml was a good concentration to start crystallization experiments.

6.1.2. N-Terminal variant

6g of cells from 1 liter culture were lysate and purified obtaining 40 mg of N-term protein. As with the previous purification, a small overlapping peak is again observed which the SDS-PAGE

gel reveals as small quantities of N-term protein and other contaminant with bigger molecular weight.

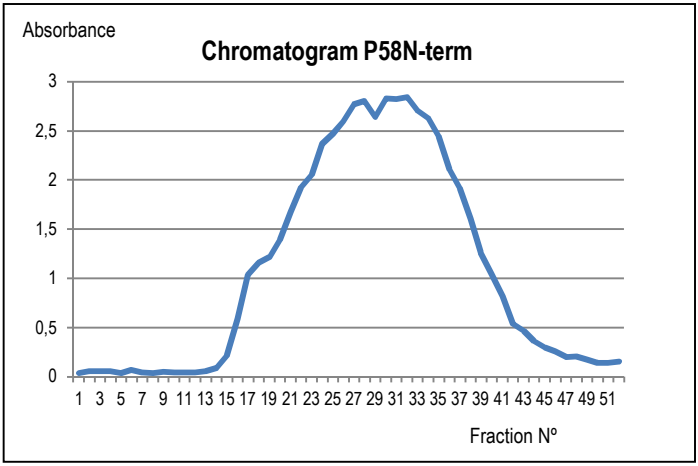


FIG.10. P58N-term purification chromatogram

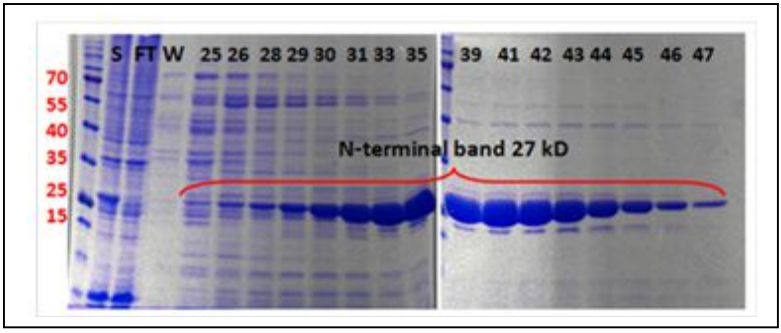


FIG.10. P58N-term SDS-PAGE gel

In red on the left: MW markers (in kD) S: sample before purification, FT: flow through; W: wash, 25-47: collected fractions. N-terminal protein has a 27kD MW as migration show.

Fractions from 25 to 47 were joined and absorbance measured in last concentration step was 7.8 which calculated with molar extinction coefficient correspond (35560 $M^{-1}cm^{-1}$) to 6 g/ml concentration.

400 µl of P58N-term at 6 mg/ml in 20 mM Tris pH 7.5 and 100 mM NaCl was finally obtained.

The Pre-Crystallization Test showed that 6 mg/ml was a good concentration to start crystallization experiments.

6.2. CRYSTALLIZATION

6.2.1. P58 Full length

After 10 days in plates set at 14 mg/ml and incubated at 18 ºC crystals grew in two crystallization conditions of JCSG screen: C10: 10 % Peg 20 000, 2 % Dioxane, 0.1 M Bicine pH 9 and E10: 10 % Peg 6000, 0.1 M Bicine pH 9.

Crystals from E10 condition were better shaped and bigger than C10 condition, so initial optimization and reproduction steps were based on it (See figure 12).

All optimization plates were set at 14 mg/ml concentration of protein and using the same protein buffer as the initial plates.

Optimization plates around E10 condition were designed by varying the Peg concentrations around the initial ones, by varying the molecular weight of the PEG compounds and by modifying the pH.

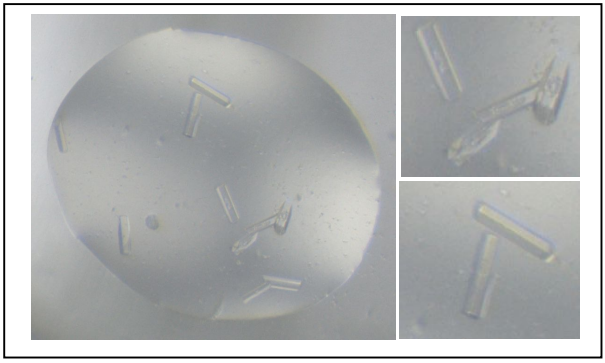


FIG.12. P58FL crystals JCSG E10 condition

	1	2	3	4	5	6	
A	**4**	**7**	**10**	**13**	**16**	**19**	**% Peg 6000** 0.1 M Glycine pH9,5
B	**4**	**7**	**10**	**13**	**16**	**19**	**%Peg 6000** 0.1 M Bicine pH9
C	**4**	**7**	**10**	**13**	**16**	**19**	**%Peg 6000** 0.1 M Bicine pH8.5
D	1:1	2:1	1:2	1:1	2:1	1:2	Crystallization condition directly from commercial JCSC screen tube using different proportions protein : condition
	JCSG E10			JCSG C10			

TAB.6. Plate-1, optimization plate designed. The table describe the composition of each well. The plate was incubated at 18 ºC and the protein used was at 14 mg/ml. Successful conditions in violet color.

Two more plates were set using the same A to C design but changing Peg 6000 by Peg 8000 (Plate-2) and Peg 4000 (Plate-3). No crystals appeared in rows A to C neither with Peg 8000 or 4000, just some crystalline precipitate, some aggregates and one crystal in well D2 of Plate-1 were found.

In Plate-4 a different gradient of precipitant and a mixture of both initial conditions were tried

	1	2	3	4	5	6	
A	**7**	**8**	**9**	**10**	**11**	**12**	**% Peg 6000** 0.1 M Bicine pH9
B	**13**	**14**					
C	**7**	**8**	**9**	**10**	**11**	**12**	**% Peg 6000** 0.1 M Bicine pH9 2% Dioxane
D	**13**	**14**					

TAB.7. Plate-4. The table describe the composition of each well. Successful conditions in violet color.

In Plate-4 some successful conditions gave some crystals. We observed that the appearance of crystals was not systematic along the concentration gradient.

As will be explained later, to solve the structure of this protein, a derivative heavy atom crystal is probably needed to find phases and solve electric density equation. With this aim Plate-5 was set, in which some metal salts were added to the condition composition in order to incorporate some metal atoms in crystals lattice (see Table 8). No protein crystal appeared in this plate, just some salt crystal.

Crystals grew in Plate-4 were treated by soaking them for few seconds in the crystallization condition in presence of metals compounds (1mM $NiSO_4$, 1mM $CaCl_2$, 1mM $MnCl_2$, 1mM $Zn(NO_3)_2$) to try to incorporate this atoms in crystal lattice before freezing it.

To analyze the protein content of the crystals, some crystals were disolved in water as preparation for gel electrophoresis. The SDS-PAGE showed that the composition of P58FL protein crystals contained mostly the P58N-term variant. So, during crystallization the degradation of protein has occurred. As a result, crystals are just containing the N-term fragment.

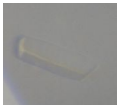
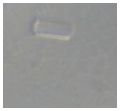
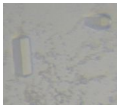
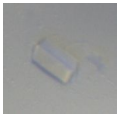
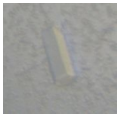
	1	2	3	4	5	6	
A	**4**	**7**	**10**	**13**	**16**	**19**	**%Peg 6000** 0.1M Bicine pH9 **1mM $Zn(NO_3)_2$**
B	**4**	**7**	**10**	**13**	**16**	**19**	**%Peg 6000** 0.1M Bicine pH9 **1mM $CaCl_2$**
C	**4**	**7**	**10**	**13**	**16**	**19**	**%Peg 6000** 0.1M Bicine pH9 **1mM $MnCl_2$**
D	**4**	**7**	**10**	**13**	**16**	**19**	**%Peg 6000** 0.1M Bicine pH9 **1mM $NiSO_4$**

TAB.8. Plate-5 . Optimization plate designed including metal salts in the conditions.

The table describe the composition of each well.

6.2.2. P58 N-terminal

After 4 days in plates set at 6 mg/ml and incubated at 18 ºC crystals grew in several conditions of different initial crystallization screens (see Table 9).

Index F1 0.2M L-Proline 0.1M Hepes pH7.5 10% Peg 3350		**JCSG B4** 0.1M Hepes pH7.5 10% Peg 8000 8% Ethylene glycol	
Morpheus D10 0.12M Alcohols 0.1M Buffer Syst. 3 50% Precipitant Mix2		**Morpheus E10** 0.12M Ethylene glycols 0.1M Buffer Syst. 3 50% Precipitant Mix2	
Morpheus F10 0.12M Monosaccharides 0.1M Buffer Syst. 3 50% Precipitant Mix2		**PACT C5** 0.1M PCTP Buff. pH8 25% Peg 1500	
PACT C6 0.1M PCTP Buff. pH9 25% Peg 1500		**PACT D8** 0.2M $(NH_4)Cl$ 0.1M Tris pH8 20% Peg 6000	
PACT E1 0.2M NaF 20% Peg 3350		**Peg/Ion B9** 0.2M HCOONa 20% Peg 3350 pH7.2	
Peg/Ion B10 0.2M HCOONa 20% Peg 3350 pH7.3		**Peg/Ion C5** 0.2M CH_3COOK 20% Peg 3350 pH8.1	
Peg/Ion C6 0.2M $CH_3COO(NH_4)$ 20% Peg 3350 pH7.1		**Peg/Ion G3** 0.1M HCOONa 12% Peg 3350 pH7	
Peg/Ion G4 0.2M HCOONa 20% Peg 3350 pH7		.	

TAB.9. P58 N-terminal crystals and conditions composition

(details of Morpheus and PACT conditions in Appendix 7)

Due to the good appearance, size and quantity of crystals, no optimization plates were set for this variant. Some of the crystals were prepared as P58FL to incorporate metal atoms.

6.2.3. Crystal handling and freezing

P58FL, N-terminal variant and derivatives metal crystals of both were treated following the same steps.

Handling was performed using 0.05-0.025 mm or 0.05-0.1 mm loops as a spoon, depending on the crystal size.

Before freezing a crystal it was stabilized dipping it in the same solution of the well where it was grew. Each crystal was freezed it in a cryoprotectant final solution composed by 20 % Glycerol anhydrous and 80 % crystallization condition from the well.

The cryoprotectant procedure was done in steps of 5 % Glycerol in order not to shock the crystal but in some cases the number of steps was reduced to minimize the handling and avoid damaging the fragile crystals. All crystals were stored under liquid nitrogen until the moment of diffraction.

6.3. DATA COLLECTION AND PROCESSING

After diffraction of several crystals from FL and N-term variant the best data was obtained from one of P58N-term protein. The collection was performed rotating the crystal from 0º to 180º and 1image /1º was took.

The parameters obtained from processing were evaluated satisfactorily and data were determined as a good data collection (see Figure 12).

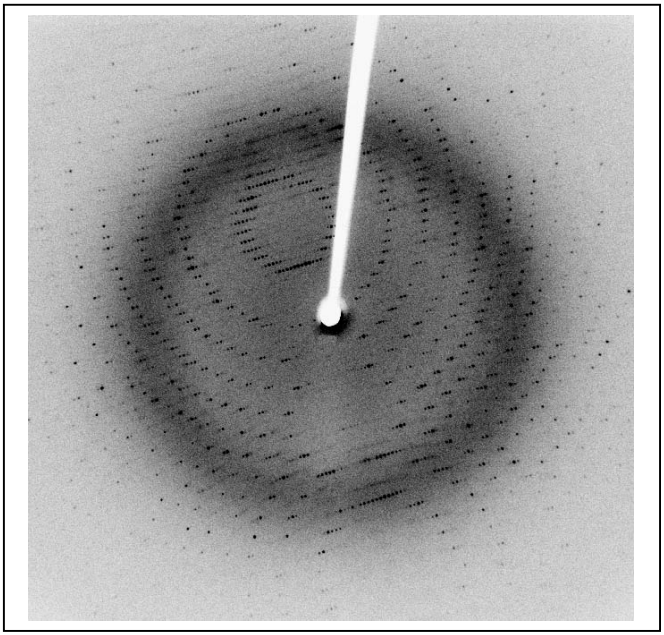


FIG.12. Image of P58N-term data collection.

Data collection Report			
Low resolution limit	30.181	Rpim (all I+ & I-)	0.018
High resolution limit	1.961	Total number of observations	111792
Rmerge	0.032	Total number unique	26537
Ranom	0.028	Mean(I)/sd(I)	22.6
Rmeas (within I+/I-)	0.036	Completeness	99.7
Rmeas (all I+ & I-)	0.037	Redundancy	4.2
Rpim (within I+/I-)	0.022	CC(1/2)	0.902

TAB.10. Data parameters report, in red some of the most significant.

High resolution limit give an idea of the detail that we can reach with the data collected. The value near to 1 indicates a good data. Rmerge is a measure of agreement among multiple measurements of the same reflections. The low value is indicative of good data. Mean(I)/sd(I) Give an idea of the accurancy of the data related to the measurement error. An average of 1.5 I/sd(I) in the highest resolution shells is still acceptable.

Completeness value is the percentage of crystallographic reflections measured in a data set. The value near 100% give us the idea that we have all the information from the crystal. Redundancy is the number of times each reflection has been measured.

The parameters obtained from processing were evaluated satisfactorily and data were determined as a good data collection.

After processing, the space group proposed is P622 and the cell unit dimensions are: a: 104.5505, b: 104.5505, c :59.2093, α: 90.000º, β: 90.000º and γ: 120.000º, so the crystal lattice correspond to Hexagonal P Bravais lattice (see Appendix1).

6.4. SOLUTION OF PHASE PROBLEM

6.4.1. Molecular replacement

In PDB Blast were found 13 proteins with 28% sequence identity to P58N-term, but only for very short stretches of sequences (see Appendix 8). This is a low homology percentage to apply MR method. Knowing this situation, it was tried with the 13 models with low identity, for which it is known that they exhibit relaxase activity as P58, independently of the sequence identity.

The program phaser was used to adapt the model selected to the data collected and to obtain an approximately value of phases (see Figures 13).

Phaser program tried to fit the data in the chosen model by moving it and rotating it and supposing different space groups of the data collected based on the one determined in processing. In Figure 13, detailed in blue, are some of the different space group used as a variable in the fitting process. In the report shows the best 130 solutions of many more that were used to fit the model. The fitting values obtained (red in Figure 13) were less than 5, (Z-score, return to Table 5) establishing that is not possible to solve the structure by molecular replacement using that model.

The process described was done with the PDB model 4LVT chosen.

```
Solution #1 annotation (history):
  SOLU SET  RFZ=3.3 TFZ=4.9 PAK=0 LLG=99 RFZ=3.0 TFZ=6.0 PAK=6 LLG=125 LLG=175
  SOLU SPAC P 62
  SOLU 6DIM ENSE ensemble1 EULER 169.7 83.6 134.1 FRAC 0.02 -0.35 -0.31 BFAC -8.06
  SOLU 6DIM ENSE ensemble1 EULER 168.6 146.9 7.1 FRAC 0.78 0.54 0.25 BFAC 16.80
  SOLU ENSEMBLE ensemble1 VRMS DELTA -0.2534 RMSD 1.43 #VRMS 1.34
   (.........)
Solution #50 annotation (history):
  SOLU SET  RFZ=3.1 TFZ=4.6 PAK=0 LLG=99 RFZ=3.0 TFZ=5.5 PAK=0 LLG=115 LLG=164
  SOLU SPAC P 61
   (.........)
Solution #130 annotation (history):
  SOLU SET  RFZ=3.3 TFZ=4.6 PAK=2 LLG=99 RFZ=3.0 TFZ=5.1 PAK=4 LLG=114 LLG=162
  SOLU SPAC P 64
```

FIG.13. Report of the different solutions proposed by Phaser.

6.4.2. Scanning of anomalous signal

The energy scan done to crystals treated with metals gave no clear fluorescence peak, so that reveals that none of the metals used were incorporated to them.

6.5. ELECTRONIC DENSITY MAP

As was explained, during the period of the present work was no possible to solve the protein structure. But anyway it was decided to include this short section to show how an electronic density map looks like. This can support some of the topics treated and at the same time it will help to understand which information is needed to finally solve a protein structure.

An electron-density map is the end-product of crystallographic structure determination. The map is an image of the electron clouds surrounding the nuclei of the atoms. In a process called map interpretation, the crystallographer builds a model to fit the amino acid sequence.

The electronic density showed in Figure 14 is from white egg hen Lysozyme, one of the most used proteins as a standard in crystallization and crystallography due to its facility in crystallization and handling. The second inset shows part of the map where differences are found in the side chain conformation of an aspartate.

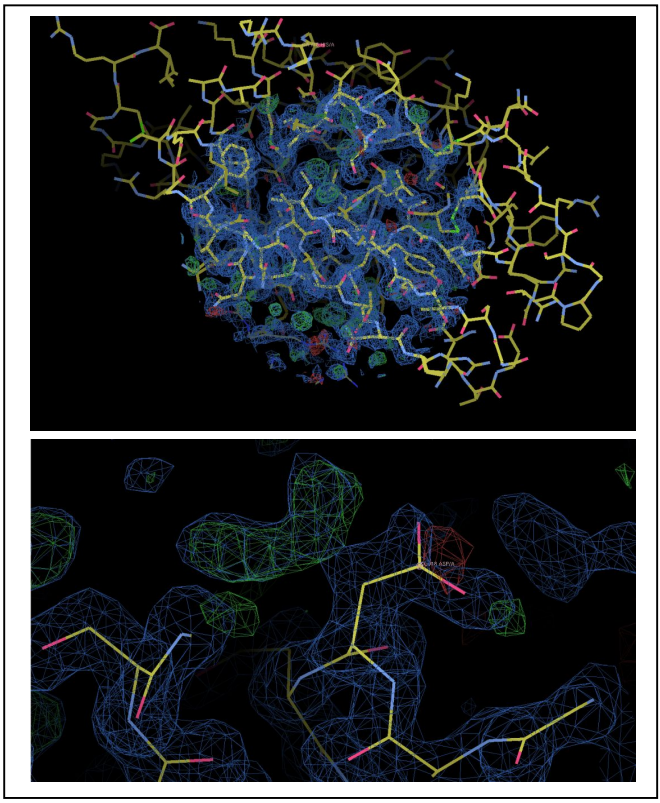


FIG.14. The three-dimensional structure of hen egg white Lysozyme was the first 2Å-resolution model via X-ray crystallography (David Chilton Phillips,1965)

(4N9R, PDB code Ref.7)

7. CONCLUSIONS

By the experimental work done for the present report is concluded the following items:

- The initial sparse/matrix crystallization screens performed with full length protein and N-terminal variant permitted obtaining of crystals.

-An electrophoresis analysis of full length protein crystals reveals that during crystallization a degradation process occurred that produced the N-term fragment. So finally, all crystals obtained were from N-term variant.

-Crystals were optimized and reproduced by the proper design of crystallization conditions to grow a good number of quality crystals. Best crystals were cryoprotected and stored to perform diffraction experiments.

- Crystals diffraction data were collected at XALOC beamline in ALBA Synchrotron. It was processed and evaluated by indication parameters values which qualified the collection as a quality data.

- To obtain phases, the Molecular Replacement method was applied by using different protein structures as a initial model. It was not successful due to low homology between the protein models and the target protein sequence.

-The obtaining of phases was not possible by diffracting metal derivative crystals. The presence of metals did not give any significant anomalous signal.

-The lack of phases did not permit to solve the electron density equation, so building the protein structure model was not possible. The phase problem solution has to be addressed by applying other techniques to derivatize the protein to grow and diffract true derivative crystals.

From the presented work It have learned which are the strategies to follow in order to crystallize a protein and how to handle and treat protein crystals to diffract them. Concerning the crystallography tasks, it was especially difficult the initiation in this subject, which mixes the crystals, real and palpable, with mathematics, statistics and modeling.

8. REFERENCES AND NOTES

1. Guasch A1, Lucas M, Moncalián G, Cabezas M, Pérez-Luque R, Gomis-Rüth FX, de la Cruz F, Coll M; *Recognition and processing of the origin of transfer DNA by conjugative relaxase TrwC.* Nat Struct Biol. 2003 Dec;10(12):1002-10. Epub 2003 Nov 16

2. Chris Larkin,Saumen Datta, Matthew J. Harley, Brian J. Anderson, Alexandra Ebie, Victoria Hargreaves, Joel F. Schildbach. *Inter- and Intramolecular Determinants of the Specificity of Single-Stranded DNA Binding and Cleavage by the F Factor Relaxase*. Structure Volume 13, Issue 10,

3. Gomis-Rüth FX1, Coll M.1544. *Cut and move: protein machinery for DNA processing in bacterial conjugation*. Curr Opin Struct Biol. 2006 Dec;16(6):744-52. Epub 2006 Oct 31.

4. Arthur F. Monzingo, Angela Ozburn, Shuangluo Xia, Richard J. Meyer, Jon D. Robertus. *The Structure of the Minimal Relaxase Domain of MobA at 2.1 Å Resolution*. J Mol Biol. 2007 February 9; 366(1): 165–178.

5. Jonathan S. Edwardsa, Laurie Bettsb, Monica L. Fraziera, Rebecca M. Polleta, Stephen M. Kwongc, William G. Waltonb, W. Keith Ballentine IIIb, Julianne J. Huangb, Sohrab Habibib, Mark Del Campod, Jordan L. Meiere, Peter. B. Dervane, Neville Firthc, and Matthew R. Redinboa. *Molecular basis of antibiotic multiresistance transfer in Staphylococcus aureus.* PNAS 2013 Feb 19;110(8):2804-9

6. The Protein Data Bank H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne (2000) Nucleic Acids Research, 28: 235-242. doi: 10.1093/nar/28.1.235

7. Valls Vidal, Núria; (2004) Tesi doctoral: *Estudi estructural d'oligonucleòtids en presència de farmacs intercalants i metalls de transició*. Departament d'Enginyeria Química de la UPC.

8. Mallorquí Fernández, Goretti; (2003) Tesina: *Caracterització estructural de Ribonucleases humanes.* Universitat de Girona.

9. Tilley, Richard; *Crystals and Crystal Structures.* England, Wiley, 2006.

10. Rupp, Bernhard; *Biomolecular Crystallography*. New York, Garland Science, 2010.

11. Grases Freixedas, F.,Costa bauzá, A., Söhnel, O. *Cristalización en disolución*. Barcelona, Editorial Reverté, 2000.

12. Lodish, Berk, Matsudaira, Kaiser, Kreiger, Scott, Zipursky, Darnell. *Biología Celular y Molecular*. Madrid, Editorial M´dica Panamericana, 2006.

13. A Petsko, Gregory, Ringe Dagmar. *Protein structure and function*. USA, NSP, 2004.

14. Giacovazzo,C., Monaco, H.L., Artioli, G., Viterbo, D., Ferraris, g., Gilli, G., Zanotti, G., Catti, M. *Fundamentals of Crystallography*. Oford, New York, 2002.

15. Bergfors, Terese M. *Protein Crystallization.* California, edited by T. Bergfords, 2009.

16. www.hamptonresearch.com

17. www.moleculardimensions.com

18. www.jenabioscience.com

19. www.rcsb.org/pdb

20. www.xtal.iqfr.csic.es/Cristalografia

21. www.expasy.org

22. http://spdbv.vital-it.ch/TheMolecularLevel/ModQual/#glossary *(Judging the Quality of Macromolecular Models. A Glossary of Terms from Crystallography, NMR, and Homology Modeling)*

23. www.gelifesciences.com. (Handbooks from GE Healthcare, *Affinity Cromatography principles and methods).*

9. ACRONYMS

Bicine: 2-(Bis(2-hydroxyethyl)amino)acetic acid

Buff.: buffer

Coomasie R250: blue dye triphenylmethane derivative

CV: colum volume

DNA : Deoxyribonucleic acid

DNAss: Deoxyribonucleic acid single strand

DNAse: Deoxyribonuclease

EDTA: Etildiaminotetraactic acid

FL: full length

HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

Jeffamine ED-600: O,O'-Bis(2-aminopropyl) polypropylene glycol-block-polyethylene glycol-block-polypropylene glycol

kD: kilo Dalton

MES: 2-(N-Morpholino)ethanesulfonic acid

MWCO: Molecular Weight Cut Off

N-term: Nitrogen terminal

PCT: Pre-Criystallization Test

Peg: polyethylene glycol

PES: polyethersulfone

PMSF: Phenylmethylsulfonyl fluoride

PROTEOGEL: 37.5:1 Acrylamide to Bisacrylamide Stabilized Solution

PSA: Ammonium persulfate

SDS: Sodium lauryl sulfate

SDS-PAGE: sodium dodecyl sulfate polyacrylamide gel electrophoresis

TEMED: tetramethylethylenediamine

Tricine: N-[Tris(hydroxymethyl)methyl]glycine

TRIS: 2-Amino-2-hydroxymethyl-propane-1,3-diol

UAB: Universitat Autònoma de Barcelona

APPENDICES

APPENDIX 1: BRAVAIS LATTICES

The repetition modes by translation in crystals must be compatible with the possible crystallographic point groups (the 32 crystal classes), and this is why we find only 14 types of translational lattices which are compatible with the crystal classes.

These types of lattices (translational repetition modes) are known as the Bravais lattices.

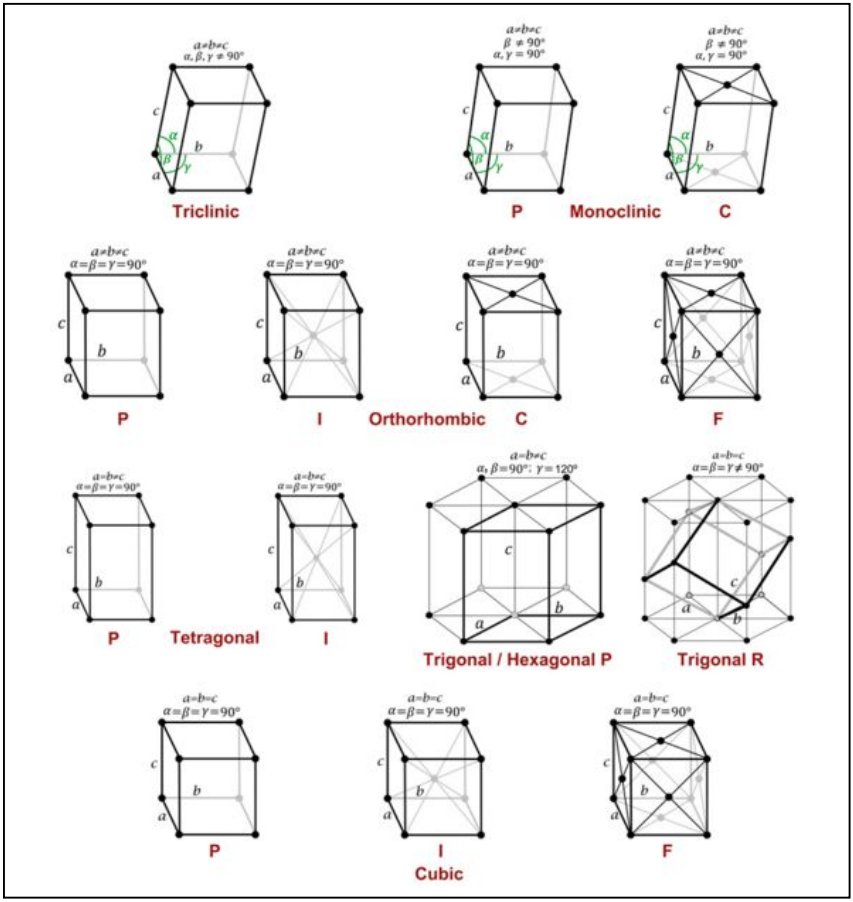


FIG. Bravais lattices with its characteristics cell parameters
(Figure adapted from Ref.19)

APPENDIX 2: BRAGG'S LAW

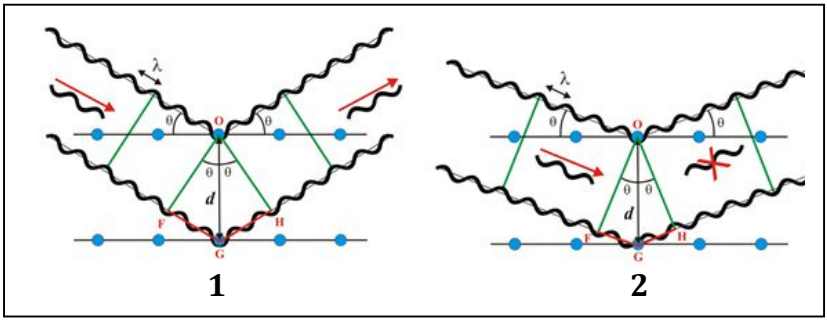


FIG.14. Bragg's Law scheme.**1**: beam which satisfy the law; **2**: beam which not satisfy the law
(Figure adapted from Ref.19)

Bragg's Law consists in imagining that the diffraction of an X-Ray reflection is caused by "imaginary" mirrors (horizontal black lines in Fig.14) formed by planes of atoms (blue circles in Fig.14) which acts as disperser centers of the crystalline lattice. Due to periodic nature of the crystal, these planes are separated by constant distances, d.

The two parallel X-Ray beams (λ, wavelength) impinge in phase on the imaginary planes at an θ angle causing a waves front (first green line from the left in Fig.14)

If the reflected beams are in phase, the interference is constructive (last green line on the right in Fig 14), and the reflected beam is observed as a spot in the diffraction pattern.

This situation occur when the trajectory of fronts (before and after reflection) is an integer number of times λ (n·λ). This condition is equivalent to the following expression (Braggs law):

$$FG + GH = n\cdot\lambda \rightarrow \text{aplying trigonometric} \rightarrow 2\,d\sin\theta = n.\,\lambda$$

(FG an GH segments in red in Fig.14)

However, when the angle of incidence of X-rays does not satisfy the Bragg law, the emergent beams are no longer in phase and cancel each other, so no reflected intensity will be observed.

APPENDIX 3: CHELATING COLUMN

HiTrap Chelating chromatography column has a matrix composed by highly crosslinked agarose beads to which iminodiacetic acid has been coupled by stable ether groups via a spacer arm. Several amino acids, for example histidine, form complexes with many metal ions. This column, charged with suitable metal ions, Ni in this case, will selectively retain proteins if complex forming amino acid residues are exposed on the surface of the protein.

Normally, proteins purified with this method must have a small segment typically formed by 6 histidines in one of the extremes of its amino acid sequence, this is called Histidine-tag. When sample is loaded into the column is retained by the matrix and separate from the impurities. Elution is performed adding to mobile phase a compound with more affinity by the metal than the Histidine-tag, often imidazole.

APPENDIX 4: SDS-PAGE

SDS page gel electrophoresis is a routine method to separate proteins by molecular weight due to the effect of an electric field. Proteins must be treated with SDS to make their charge negative so that they migrate to the Anode. The separation takes place in a matrix which is a molecular sieve (in the present work a 15 % mixture acrylamide-bis-acrylamide) where smaller proteins migrate faster than larger ones.

To visualize the result, the gel must be stained and with a dye which binds non specifically to proteins. These appear like blue bands on the gel. To check the molecular weight of the bands a patron of known molecular weight is loaded in each gel next to the protein target sample to compare its displacements.

APPENDIX 5: EXPASY AND PROTPARAM

ExPASy is the Swiss Institute of Bioinformatics (SIB) Bioinformatics Resource Portal which provides access to scientific databases and software tools in different areas of life sciences including proteomics, genomics, phylogeny, systems biology, population genetics, transcriptomics etc. On this portal it is possible to find resources from many different SIB groups as well as external institutions.

ProtParam is a protein identification and analysis tools on the ExPASy Server. ProtParam computes various physico-chemical properties that can be deduced from a protein sequence.

APPENDIX 6: PRECRYSTALLIZATION TEST (PCT)

The PCT kit is a commercial kit (Hampton Research) that contains 4 unique preformulated reagents used to evaluate protein concentration for crystallization screening. Initially, the sample protein is mixed with two of the reagents to determine if the protein concentration is appropriate for crystallization screening. If the protein is very sensitive to salt and polymer concentration, based on initial PCT results, the protein may be evaluated using a second set of PCT reagents. PCT results will then provide insight to either the appropriate sample concentration or indicate that other diagnostic could be necessary to evaluate sample homogeneity.

APPENDIX 7: COMPOSITION (MORPHEUS AND PACT)

MORPHEUS screen

ALCOHOLS: 0.2M 1,6-Hexanediol; 0.2M 1-Butanol; 0.2M 1,2-Propanediol; 0.2M 2- Propanol; 0.2M 1,4-Butanediol; 0.2M 1,3-Propanediol. BUFFER SYSTEM 3: 1M Tris, BICINE pH8.5 PRECIPITANT MIX 2: 40% v/v Ethylene glycol; 20 % w/v PEG 8000. ETHYLENE GLYCOL: 0.3M Diethylene glycol; 0.3M Triethylene-glycol; 0.3M Tetraethylene glycol; 0.3M Pentaethylene glycol. MONOSACCHARIDES: 0.2M D-Glucose; 0.2M D-Mannose; 0.2M D-Galactose; 0.2M L-Fucose; 0.2M D-Xylose; 0.2M N-Acetyl-D-Glucosamine.

PACT screeen

PCTP buffer; Sodium propionate, Sodium cacodylate trihydrate, Bis-Tris propane

APPENDIX 8: PROTEIN DATA BANK AND BLAST TOOL

The Protein Data Bank (PDB) has served as the single repository of information on the 3D structures of proteins, nucleic acids and complex assemblies. These structures are generally obtained by X-ray Crystallography and Nuclear Magnetic Resonance.

The PDB organization manages the PDB archive and ensures that the PDB is freely and publicly available to the global community.

On the PDB WebPages, there are a lot of useful tools available. BLAST is a suite of programs provided by National (U.S.) Center for Biotechnology Information (NCBI) for aligning query sequences against those present in a selected target database, the PDB in this case. The NCBI BLAST homepage (http://blast.ncbi.nlm.nih.gov) provides an access point for these tools to perform sequence alignment on the web.

In bioinformatics, a sequence alignment is a way of arranging the sequences of DNA, RNA, or protein to identify regions of similarity that may be a consequence of functional, structural, or evolutionary relationships between the sequences.