Paulo Alexandre Correia Mamede

Bachelor degree in Biochemistry

Crystallographic studies of serine β-lactamase and serine proteases with novel inhibitors

Dissertation to obtain the Master degree in Biochemistry for Health

Supervisor: Dra. Margarida Archer Frazão, Principal Investigator, ITQB NOVA

Co-supervisor: Dr. José Artur Brito, Post-Doctoral Fellow, ITQB NOVA

January 2019

Paulo Alexandre Correia Mamede

Bachelor degree in Biochemistry

Crystallographic studies of serine β-lactamase and serine proteases with novel inhibitors

Dissertation to obtain the Master degree in Biochemistry for Health

Supervisor: Dra. Margarida Archer Frazão, Principal Investigator, ITQB NOVA

Co-supervisor: Dr. José Artur Brito, Post-Doctoral Fellow, ITQB NOVA

Jury:

Dr. Pedro Matias

Dr. João B. Vicente

Instituto de Tecnologia Química e Biológica António Xavier

January 2019

Crystallographic studies of serine β-lactamase and serine proteases with novel inhibitors

Acknowledgments

I would like to sincerely thank to Professor Margarida Archer for letting me develop my master thesis in the laboratory of Membrane Protein Crystallography (MPX), for the opportunity on working with this new target and ligands, for the supervision that provided me, and also all the encouragement. Besides, to the group of Professor Rui Moreira and their students Rita Félix and Luís Carvalho from Instituto de Investigação do Medicamento at Faculty of Pharmacy of University of Lisbon, for providing us the compounds and allowing me to give my contribution on this project.

Also, a special thanks to José Brito for all the supervision and availability, as well as the dedication and friendship on helping me to reach this goal. Thank you to my labmates, for their help and patience, for sharing their knowledge, and of course for all the time full of funny moments and jokes.

I also would like to acknowledge all members of Macromolecular Crystallography Unit at ITQB, for all tips and help with my experiments and results along this challenging year.

To Isabel Pacheco and João Carita from ITQB Research Facilities, I also would like to express my deepest thanks for the kind and precious help during some difficult steps of my work and for sharing their expertise with me.

Likewise, thanks a lot to my friends that I made along these exceptional years in faculty and also to my hometown friends, for the fellowship, support, trust and the fun. They made me became better at several levels and to think in an even more demanding way.

[PT]

Por último, mas igualmente fundamental, queria agradecer à Viviana pela força incondicional, por acreditar em mim todos os dias e pela paciência nos momentos mais atribulados ao longo desta etapa, por me ter ajudado a manter o foco e a saber do que sou capaz. À minha Família, pelo incomensurável apoio, por sempre acreditarem e confiarem nas minhas decisões e nos objectivos que tracei, especialmente aos meus pais e avós pela motivação e inspiração que sempre serão.

Abstract

Serine-based enzymes are present in many biological processes and are often chosen as promising drug targets, since their deregulation lead to diseases, and study the 3D structure is essential to understand how they interact with molecules capable of inhibiting their activity. This work aims for structure determination by X-ray crystallography of protein complexes with novel inhibitors and results from a collaboration established with Prof. Rui Moreira from *iMed.ULisboa*, who provided the ligands.

This thesis is organized in two chapters: the first includes studies with TEM-1, a serine β-lactamase from *E. coli* that is involved in β-lactam antibiotics resistance; the second describes studies with Porcine Pancreatic Elastase (PPE), a serine protease that damages the connective tissue and shares the same catalytic triad of Human Neutrophil Elastase (HNE). When misregulated, HNE triggers inflammatory and respiratory diseases. Both PPE and TEM-1 have similar reaction mechanisms.

The gene coding for TEM-1 was kindly provided by Prof. Christopher Schofield (*University of Oxford*), expressed in *E. coli* and the purification protocol was optimized to obtain recombinant protein in good conditions. Crystallization trials produced crystals, which did not diffract. Further conditions were optimized but no suitable crystals were obtained, so experiments with complexes were not performed. Thus, a different serine enzyme was crystallized with other ligands belonging to the same family of the β-lactam-based inhibitors.

PPE is commercially available and was structurally characterized with different ligands to investigate the inhibition mechanism, which could occur by a nucleophilic attack on the sulfonyl group of the ligand, rather than the common attack on the carbonyl group. PPE crystals in complex with ligand LMC249 diffracted in synchrotron at around 1.4 Å. Analysis of the electron density maps around the active site did not show clear density for the ligand.

Keywords: Serine-based enzymes; β-lactamase; Elastase; X-ray crystallography; β-lactam-based inhibitors.

Resumo

Enzimas de serina estão presentes em diversos processos biológicos e são frequentemente escolhidas como alvos terapêuticos, dado que a sua desregulação leva a doenças, e estudar a sua estrutura 3D é essencial para perceber como estas interagem com moléculas capazes de as inibir. O objectivo deste trabalho experimental foi a determinação estrutural, por cristalografia de raios-X, de complexos proteicos com novos inibidores e resultou de uma colaboração estabelecida com o grupo do Prof. Rui Moreira no *iMed.ULisboa*, que forneceu os ligandos.

Esta Tese é composta por dois capítulos: o primeiro aborda os estudos com a TEM-1, uma β-lactamase de serina que está envolvida na resistência a antibióticos β-lactâmicos; o segundo descreve os estudos com a Elastase Pancreática Suína (PPE), uma protease de serina que danifica o tecido conjuntivo e partilha a mesma tríade catalítica da Elastase Neutrofílica Humana (HNE). Quando mal regulada, a HNE desencadeia doenças inflamatórias e respiratórias. Tanto a PPE como a TEM-1 apresentam mecanismos de reação semelhantes.

O gene que codifica para a TEM-1, gentilmente fornecido pelo Prof. Christopher Schofield (*University of Oxford*), foi expresso em *E. coli* e o protocolo de purificação foi optimizado para se obter proteína recombinante em boas condições. Os ensaios de cristalização produziram cristais, os quais não difractaram. Outras condições foram optimizadas mas não se obteve qualquer cristal, pelo que os ensaios com complexos não foram realizados. Assim, uma enzima de serina diferente foi cristalizada com outros ligandos pertencentes à mesma família de inibidores baseados no anel β-lactâmico.

A PPE está disponível comercialmente e foi caracterizada estruturalmente com diferentes ligandos para se investigar o seu mecanismo de inibição, que pode ocorrer através de um ataque nucleofílico no grupo sulfonilo do ligando, ao invés do ataque normal no grupo carbonilo. Cristais de complexos da PPE com o ligando LMC249 difractaram em sincrotrão a cerca de 1.4 Å. A análise dos mapas de densidade electrónica junto ao centro activo não mostrou de forma clara que há densidade para o ligando.

Palavras-chave: Enzimas de Serina; β-lactamase; Elastase; Cristalografia de raios-X; Inibidores baseados no anel β-lactâmico.

Index of contents

Index of figures

Figure 4.6 – Chromatograms for the best TEM-1 histrap purification. Flow rate was 3 mL/min. Injected sample (F.T. 1 – Flowthrough 1) was reinjected again in the column (F.T. 2 - Flowthrough 2) before starting the gradients. Main peaks are highlighted with corresponding imidazole concentration and fraction number.

Figure 4.7 – 15% SDS-PAGE after the best purification attempt by doing a histrap purification, performed with a 5 mL histrap column. M - Precision Plus Protein™ Dual Color Standards.

Figure 4.8 – Chromatogram for the best TEM-1 ionic exchange purification. Flow rate was 2 mL/min. Main peak is highlighted with the concentration value and fraction number.

Figure 4.9 – 15% SDS-PAGE of the previous best ionic exchange chromatography. M - Precision Plus Protein™ Dual Color Standards. Injected (previously named pool +) had 9 µg protein, in 5 mL. Arrows indicate where is expected to be TEM-1 band.

Figure 4.10 – Chromatogram for the best TEM-1 size-exclusion chromatography. Flow rate was 0.5 mL/min. The peak is highlighted with the identification of its fraction.

Figure 4.11 – 15% SDS-PAGE of previous size-exclusion chromatography, as well as one sample from the purer pool 1 after its concentration up to 200 µL. M - Precision Plus Protein™ Dual Color Standards. F.T. – Flowthrough after concentration. Arrows indicate where is expected to be TEM-1 band.

Figure 4.12 – BSA calibration curve to quantify TEM-1 from pool 1. Blue points represent the experimental mean values obtained to trace the calibration curve. Red cross represents point that was not considered to do the calculations (outlier). Blue line is the trendline of the linear regression, with the corresponding equation represented at the top right.

Figure 4.13 – Image of the drop with a native TEM-1 crystal (hexagonal shape), obtained after 10 days in the condition 1-32 of BCS Screen. Protein concentration was 16.5 mg/mL and had a length of about 50 µm. The picture of the whole drop was taken with a ~~zoom of 8x.~~

Figure 4.14 – Crystal of native TEM-1 obtained after 7 days in the well E6 of the optimization matrix of BCS Screen. Protein concentration was 16.5 mg/mL and had about 60 µm. The picture of the whole drop on the left was taken with ~~a zoom of 6x.~~

Figure 4.15 – Crystals of native TEM-1 obtained after optimization of the initial hit in a BCS screen condition. It appeared after 7 days, with a protein concentration of 8 mg/mL and had about 120 µm of length. The picture was ~~zoomed 5x~~. White arrows point to the two crystals that was tested, one *in-house* and the other in beamline I24 of Diamond synchrotron.

Figure 4.16 – Data collection for the native crystal of TEM-1 β-Lactamase. On the left is the loop mounted in the beamline with the crystal centered and ready to diffraction screening, and on the right is one of the diffraction patterns obtained with no spots visible. The diffraction was tested in beamline I24 at Diamond Light Source.

Figure 4.17 – Crystal of native TEM-1 obtained in condition 1-30 of PACT premier™ screen, after 11 days, with protein from pool 2.0. The crystal had about 40 µm. White arrow points to the crystal at the edge of the drop, and the picture was taken with a ~~zoom of 11x.~~

Figure 5.1 – The six main classes of enzymes, organized by the type and its general reaction, the classes with the respective EC number, and an example of one enzyme for each of them. It is ordered by the hierarchical level of enzyme commission number from top to bottom.

Figure 5.2 – Overview of protease families identified and with PDB entries, until 2017. Arrow highlights the highest number that belong to serine proteases. It comprises not only the types from MEROPS classification, but also others.

Figure 5.3 – General catalytic triad of serine proteases, formed by serine, histidine and aspartic acid.

Figure 5.4 – General catalytic mechanism of serine proteases, with the arrows indicating the reversibility of each reaction.

Figure 6.1 – Structures of the ligands originally designed for human neutrophil elastase (HNE), the LMC223 and LMC249, synthesized at *iMed.ULisboa*, and built with MarvinSketch® (from *ChemAxon*) software.

Figure 6.2 – Structures of the three ligands designed and already tested for porcine pancreatic elastase (PPE), named LMC188, LMC240 and LMC269, synthesized at *iMed.ULisboa*, and built with MarvinSketch® (from *ChemAxon*) open-source software.

Figure 7.1 – Image of two different drops with native PPE crystals (orthorhombic shape) formed after 5 days in 70% MPD with 10 mM sodium phosphate buffer pH 5.9, with the bigger crystals present in the drop on the left. Protein concentration was 20 mg/mL and by using the sitting-drop method at 20°C. The picture on the left side was taken with a ~~zoom of 2.5x~~, and the other one on the right side with a ~~zoom of 5x.~~

Figure 7.2 – Drops from the first plate with PPE co-crystallization experiments, by sitting-drop method. In general, all crystals appeared after 12 days in the same reservoir condition of the native crystals and the incubation with ligands (tenfold of molar excess) was 45 minutes at room temperature. **(1)** Drop A2 – with LMC223, ratio of 1:2 µL, [P]=15 mg/mL and the picture with a zoom of 5x; **(2)** Drop B4 – with LMC249, ratio of 1:2 µL, [P]=12.5 mg/mL and the zoom of the picture is 3x; **(3.1)** and **(3.2)** Drop B3 – with LMC249, ratio of 1:1 µL, [P]=12.5 mg/mL and the zoom of the picture is 5x.

Figure 7.3 – Drops from the second plate with PPE co-crystallization experiments, where the incubation conditions and reservoir like the previous plate. Several crystals were visible after 10 days. **(1)** Drop C1 – with LMC249, ratio of 1:1 µL, [P]=12.5 mg/mL, sitting-drop and the zoom of the picture is 3.5x; **(1.1)** and **(1.2)** different close-up's of drop C1 to highlight the best crystals and the zoom of the pictures are 6x and 9x, respectively; **(2)** Drop C4 – with LMC249, ratio of 1:1 µL, [P]=12.5 mg/mL, hanging-drop and the picture with a zoom of 5x.

Figure 7.4 – Drops also from the second plate with PPE co-crystallization experiments, but now showing trials where after 10 days appeared smaller crystals in **(1)** drop B1 and **(2)** drop B2. Both contain the protein with LMC223, ratio of 1:2 µL, protein at 15 mg/mL, sitting drop and the pictures were taken with a zoom of 8x.

Figure 7.5 – Data collection for PPE, crystal **2)**, in complex with the ligand LMC249. On the left is an image of the loop mounted in the beamline with the crystal centered, and on the right is shown an image of the diffraction pattern screening with several spots visible. The diffraction was tested in beamline I24 at Diamond Light Source.

Figure 7.6 – Electronic density map of PPE complex with LMC249 focused on the active site region, obtained after the final refinement. The $2F_{obs}$-F_{calc} difference map is in blue at a contour level of 1.2 rmsd, the F_{obs}-F_{calc} map is in green for positive density, red for the negative one and contoured at 3.6 rmsd. The residues His57, Ser195 and Asp102 are in ball & stick representation and they belong to the catalytic triad of the enzyme, and the data resolution of the crystal is 1.35 Å. The image was produced with *COOT* software.

Figure 7.7 – Representation of the electrostatic surface for the final atomic model of PPE+LMC249, with active site Ser195 properly highlighted. Image produced with COOT software, where surface is colored by the electrostatic potential, red for the regions with the negative charges and blue for the ones with the positive charges.

Figure B.1 – Histrap purification of 3rd attempt of TEM-1 purification. Elution was done with linear gradients of [I] imidazole concentration: 1st with no [I] for 30 mL; 2nd up to 25 mM [I] for 30 mL; 3rd up to 50 mM [I] for 30 mL; 4th from 50 mM to 250 mM [I] for 75 mL; 5th a final elution with 500 mM [I] for 30 mL. Flow rate was 3 mL/min. Total volume of injected sample was 38 mL and it was reinjected in column before starting the gradients. Main peaks are highlighted with [I] corresponding values.

Figure B.2 – 15% SDS-PAGE after 3rd attempt of affinity chromatography, performed by two methods: a histrap purification and with a batch/gravity-flow purification, with different [I] imidazole concentrations. M - Precision Plus Protein™ Dual Color Standards. Arrows indicate where is expected to be TEM-1 band. Pooled fractions: **Pool ++**: F 90→F 104; **Pool (from batch purification)**: 50 mM [I] Wash2 → 300 mM [I] Wash2.

Figure B.3 – Chromatograms of the 3rd attempt of TEM-1 purification with a HiLoad 16/10 Q-Sepharose High Performance column and the buffers used were: buffer A with 20 mM HEPES pH 7.5; buffer B with 20 mM HEPES pH 7.5 + 1 M NaCl.

Final volumes of injected samples were: 5 mL Sample + 5 mL buffer A in **C1**, and 10 mL Sample in **C2**. Purification was done with linear gradients of sodium chloride (NaCl) concentration (0 → 1 molar). First, the column was washed with washed with buffer A for 20 minutes and then was done a 1st gradient up to 300 mM [NaCl] for 1 hour. After that, a 2nd gradient between 300 mM and 1 M [NaCl] was performed for 25 minutes. Flow rate was 2 mL/min. Main peak is highlighted with the NaCl concentration and fraction number.

Figure B.4 – 15% SDS-PAGE with fraction from ionic exchange purification after the 3rd attempt of protein purification.

M - Precision Plus Protein™ Dual Color Standards. The arrows indicate the molecular weight where is expected to be TEM-1 band. Pooled fractions: **Pool +** (from initial pool ++ of histrap purification) - F14 → F21; **Pool -** (from pool of the initial batch purification) - F32 → F41 + F4 + F5.

Figure B.5 – Chromatograms of 3rd attempt of TEM-1 purification, after two SEC's with a HiLoad 16/10 Sephadex 200 pg column and the buffer used was 20 mM HEPES pH 7.5 plus 150 mM NaCl. Injected sample volume was 1.5 mL for both runs. Flow rate was 0.5 mL/min. Main peaks are highlighted with the identification of the fraction.

Figure B.6 – 15% SDS-PAGE to analyze the two SEC's of 3rd purification attempt. Gel on the left refers to samples obtained after purification of previous **pool +** and gel on the right has the samples after purification of **pool –**. M - Precision Plus Protein™ Dual Color Standards.

Figure B.7 – Chromatogram of 5th attempt of TEM-1 purification with a 5 mL histrap fastflow column. Elution was performed with linear gradients of imidazole ([I]): 1st with no [I] for 10 minutes; 2nd with 50 mM [I] for 10 minutes; 3rd from 50 mM to 250 mM [I] for 25 minutes; 4th a final elution with 500 mM [I] for 5 minutes. Flow rate was 3 mL/min. Total volume of injected sample (F.T. 1 - Flowthrough 1) was 85 mL and reinjected in the column (F.T. 2 - Flowthrough 2) before starting the gradients. Main chromatogram peaks are highlighted with imidazole concentration and corresponding fraction number.

Figure B.8 – 15% SDS-PAGE after 5th attempt of the histrap purification, performed with a 5 mL histrap fastflow crude column. M - Precision Plus Protein™ Dual Color Standards. Arrows indicate the molecular weight where is expected to be the TEM-1 band. Pooled fractions: F 60 → F 86.

Figure B.9 – Chromatogram of 5th attempt of TEM-1 purification with a HiLoad 16/10 Q-Sepharose High Performance column. Volume of injected sample (pool made after the previous histrap purification) was: 5 mL sample + 5 mL buffer A. Purification was done with linear gradients of sodium chloride (NaCl) concentration (0 → 1 molar). First, the column was washed with

buffer A for 20 minutes and then was done a 1st gradient up to 300 mM [NaCl] for 1 hour. After that, a 2nd gradient between 300 mM and 1 M [NaCl] was performed for 25 minutes. Flow rate was 2 mL/min. Main peak is highlighted with the NaCl concentration and fraction number.

Figure B.10 – 15% SDS-PAGE of the results from ionic exchange purification after 5th attempt of protein purification. M - Precision Plus Protein™ Dual Color Standards. Arrow indicates the molecular weight where is expected to be TEM-1 band. Pooled fractions: **Pool +** (F 18→F 30); **Pool -** (F 31→F 34).

Figure B.11 – Chromatograms of the 5th attempt of TEM-1 purification, after two SEC's with a Superose 12 10/300 GL column and the buffer used in both was 20 mM HEPES pH 7.5, 150 mM NaCl, 5% glycerol and 0.2 mM PMSF. For **Pool +** the injected volume was 500 µL with a final concentration of 9 mg/mL, and the **Pool –** was injected with 6.8 mg/mL and a volume of 1 mL. Flow rate was 0.5 mL/min. Main peaks are highlighted with the identification of the fraction.

Figure B.12 – 12% SDS-PAGE gels with the results from the two SEC's performed in the 5th purification attempt. M - Precision Plus Protein™ Dual Color Standards. Arrows indicate the molecular weight where is expected to be TEM-1 band.

Final pooled fractions: **Pool 1.0** (C1: F21 + F22); **Pool 2.0** (C2: F21 + F22); **Pool 2.1** (C2: F23 + F24).

Figure C.1 – List of conditions of the BCS (Basic Chemical Space) Screen, from *Molecular Dimensions*. Available on: https://www.moleculardimensions.com/applications/upload/MD1-104%20The%20BCS%20Screen.pdf.

Figure C.2 – Matrix of optimization for TEM-1: from one hit in BCS Screen, condition #1-32.

Figure C.3 – List of conditions of PACT premier™ screen (a pH, Anion, Cation crystallization screen), from *Molecular Dimensions*. Available on: https://www.moleculardimensions.com/applications/upload/MD1-29%20PACT%20premier%20v2.pdf

Index of tables

Abbreviations

Å	Ångström ($1Å = 10^{-10}\ m$)
Abs	Absorbance (unitless)
AMR	Antimicrobial resistance
AU	Arbitrary units, applied in absorbance values
BLAST	Basic Local Alignment Search Tool
BSA	Bovine serum albumin
Da	Dalton
DLS	Diamond Light Source
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
EC	Enzyme commission
E. coli	*Escherichia coli*
HEPES	N-(2-hydroxyethyl)-piperazine-N'-2-ethanesulfonic acid
HNE	Human Neutrophil Elastase
IEX	Ionic exchange chromatography
IMAC	Immobilized metal-affinity chromatography
IPTG	Isopropyl-β-D-thiogalactopyranoside
MPD	2-Methyl-2,4-pentanediol
MW	Molecular weight
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PCTP	Buffer composed by sodium propionate, sodium cacodylate trihydrate and Bis-Tris propane, with pH varying between 4.0 and 9.5
PDB	Protein data bank
PEG	Polyethylene glycol
PMSF	Phenylmethylsulfonyl fluoride
PPE	Porcine Pancreatic Elastase
PVDF	Polyvinylidene fluoride
Q-Sepharose	Quaternary ammonium (Q) strong anion exchange groups matrix, crosslinked to 6% agarose beads
rpm	Rotations per minute
SDS	Sodium dodecyl sulfate
SEC	Size exclusion chromatography
TEM	Temoniera, the name of a greek patient
USD	United States Dollar
V/V	Volume per volume
w/V	Weight per volume
3D	Three-dimensional

1. Introduction: biological background

1.1 The problem of antimicrobial resistance (AMR)

Dated from 1928, in a British hospital, it is known that Alexander Fleming discovered a novel substance on a piece of mould that fortuitously contaminated a petri dish under his bench, while the bacteria he was examining were killed by. [1] This substance, nicked by Fleming as Penicillin, was intensively studied by Fleming itself, but not alone, and in the course of a little more than 10 years became a sensational drug capable of cure efficiently patients with bacterial infections. That type of healing we can designate as the first approach to an antimicrobial treatment, or, more accurately, the first antibiotic if we consider just infections caused by bacteria. [1, 2] There are various types of resistance that can be comprised in antimicrobial resistance (AMR), namely if the infection is bacterial, viral, fungal or caused by parasites (e.g. malaria). In the context of this Thesis, and from here on, the term AMR will only refers to bacterial infections. [1, 2]

Besides penicillin, other antibiotics were discovered along time and made a revolution on healthcare, establishing the greatest medical improvements of the 20th century. Simple infections as small wounds or cuts no longer had to be fatal, as well as childbirth deaths were practically extinguished. [1] In recent years, resistance has become a huge public health problem because of the low pace at which we are developing novel antibiotics and the fact that antibiotic usage is rising exponentially. All of these phenomena have led to the rise of '*superbugs*', a term often used to describe strains of bacteria that are resistant to most antibiotics commonly used today. [1, 2]

The problem today reaches the entire world, and the damaging effects of infections caused by AMR are coming to at least 50 000 deaths each year, across Europe and USA. Despite the contrast, AMR should be a concern to every country regardless of its level of income. Even in developed healthcare systems, with have access to second and third-line antibiotic treatments, the mortality for patients infected by resistant bacteria is very high, with the effectiveness of the treatments being progressively lower. [1, 2] In figure 1.1, we can easily see the real impact of the number of deaths due to AMR every year in the future, when compared with other leading causes of death of this century. In 2050 it is estimated this number will be 14 times bigger than the actual, even more than those dying of cancer.

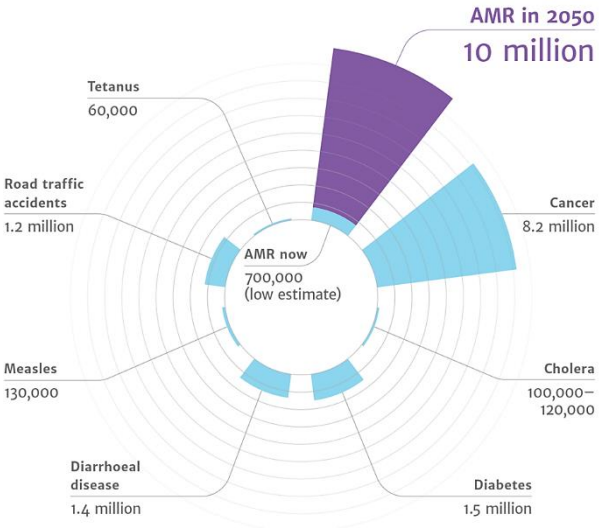


Figure 1.1 – Diagram with deaths attributed to AMR every year, when compared to other causes of death around the world. Blue represents the actual data and purple if for the data predicted for 2050. Estimates for 2011 and updated in 2016. [2]

Globally, the consumption of antibiotics in human medicine rose by nearly 40% between 2000 and 2010, but this behavior masks patterns of declining usage in some countries and rapid growth in others. One of the greatest worries about AMR is that modern health systems and treatments that rely heavily on antibiotics could be severely damaged, for instance when in most surgeries the patients need prophylactic treatment with antibiotics to minimize the risk of bacterial infections. If the antibiotics do not work, this measure would become largely useless and surgery would become far more dangerous. Many procedures, such as joint replacements, which currently allow people to live physically active for longer and may enable them to stay in the workforce, might become too risky to perform. Other current example is the modern cancer treatments that often take down the patients' immune systems, making them more susceptible to infections. Hence, without effective antibiotics to prevent or treat infections, chemotherapy would become riskier. [1, 2]

1.1.1 The current dissemination and rising of AMR on a global scale

The potential impacts outlined above show that it is critical that healthcare systems all over the world do not let themselves be undermined by resistance to antimicrobial drugs. Likewise, resistance is not just a main health worry, which will lead to millions more people dying each year, but it is also an economic issue. Because of that, it is needed to find new strategies and drugs capable of stopping and inhibiting the mechanisms of resistance, allowing us to preserve one of the most precious medical resources the world has ever had. [1, 2]

The development of resistance is inevitable, evolutionary speaking, even with the antimicrobials always used properly and sparingly. However, studies show that quick actions are crucial, and we must do everything to slow the spread of resistance and to moderate its impact with effective new treatments. If it could be possible to delay the progress of widespread resistance by just 10 years, 65 trillion USD of the world's output could be saved between now and 2050, showing the importance of not just treating infections but also reducing and controlling them. [1, 2]

Figure 1.2 shows the estimated number of deaths due to AMR across the world by 2050. There are variations between regions, as expected, since the level of healthcare and medical expertise are not equal in the different parts of the world. Africa and Asia are the continents that will suffer the most. [1]

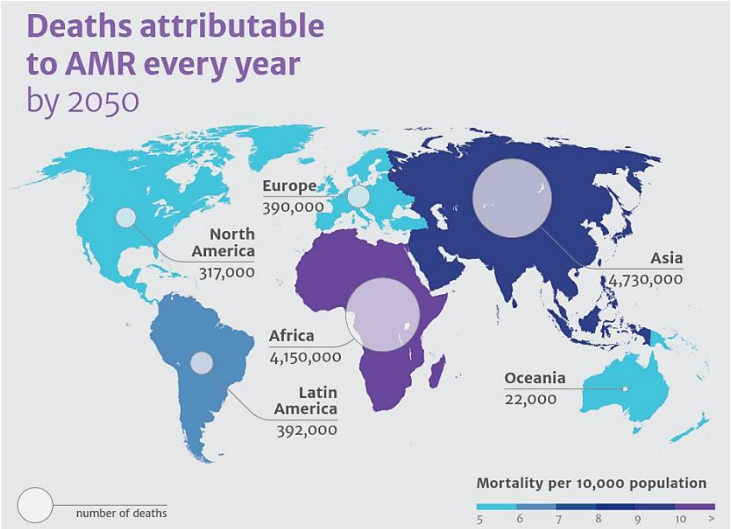


Figure 1.2 – Diagram with the number of deaths worldwide due to AMR every year, predicted for 2050. Different colours for the several areas represent the variation of mortality rate, per each 10 thousand people. Values estimated in 2011. [1]

1.1.2 The future of antibiotics towards an AMR decrease

Drug-resistant infections already cost a huge amount of lives. After antibiotic fails in specific cases, there are no more available treatments. The AMR problem is a serious and deep clinical problem, with more than 70% of the bacteria that cause hospital-acquired infections being resistant to at least one of the drugs that are currently used for treatment. Indeed, the resistance to only one specific class of antibiotic is rare and the multiple drug resistance (MDR) has grown dramatically. [8]

Moreover, for some types of infections when the doctors are running out of options, they need to choose antibiotics that were once avoided due to their extreme side effects (e.g. colistin) which can cause kidney disfunction and so was avoided for many years. Nevertheless, over the past decade it has re-entered on clinical use as a last option of treatment for particularly hard-to-treat gram-negative bacterial infections, and the colistin resistance it was already detected. [1, 2]

Finding new options of antibiotics to prevent the deaths caused by AMR it is not a simple task. In a few decades we change from 30 big pharmaceutical companies to just 6 of them, through processes of merging or acquisitions from other companies, and all of that do not encourage a good market competition. To have a real idea, between 1983 and 1987 were approved 16 new types of antibiotics to combat antibacterial infections, however between 2008 and 2011 there were only 2 new ones. [3]

According to data of *The Pew Charitable Trusts*, a non-governmental organization, on May of the last year were developed 41 new possible antibiotics, yet the most probable scenario is that just up to 20% will reach the markets. Besides, just 10 of them will be considered innovative drugs meaning they are not derivatized or from other compounds already available. [3]

1.2 β-lactamases: triggers of resistance

β-lactamases (EC 3.5.2.6), also known as penicillinases, are enzymes that belong to the class of hydrolases and perform the hydrolysis of a β-lactam ring, generally present in β-lactam antibiotics, by breaking up the (N-C=O) nitrogen-carbonyl bond in the ring (Figure 1.3). This is the most common mode of action for β-lactam resistance in bacteria rendering the drugs inactive. β-lactamases exhibit wide range of substrate preferences, which often include penicillins, cephalosporins and carbapenems, among others. [9, 11]

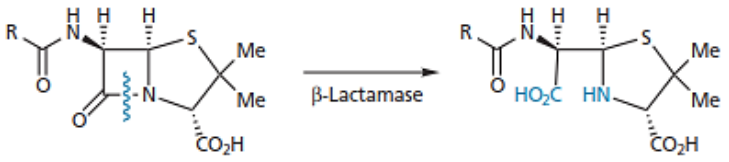


Figure 1.3 - Example of deactivation of penicillin by β-lactamase. [9]

The natural production of these enzymes by bacteria is the most common mechanism that confer AMR, even with several different types of therapeutics. A handful of β-lactamases were known in the early 1970's, but the number grew rapidly and today are known over 470 β-lactamases. [4] It supports indirectly the survival of others microorganisms that are sensitive to the antibiotics during infection processes, and therefore that contributes to the high level of antibiotic prescriptions to combat those microorganisms. [8, 9]

Related to its cell location, they can be identified in gram-positive bacteria, in the surrounding environment between the cell membrane and the peptidoglycan multi-layered cell wall. They are also present in gram-negative bacteria (Figure 1.4), as *E.coli*, freely in the periplasmic space and attached to the thin peptidoglycan. [9, 12]

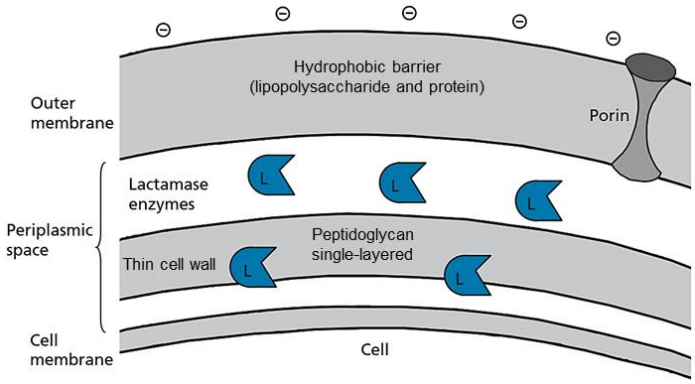


Figure 1.4 - Architecture of a gram-negative bacteria membrane, with the β-lactamase enzymes location. Adapted from 9

Generally, gram-negative bacteria are more resistant to antibiotics, because the released enzymes are trapped between the cell membrane and the outer membrane and they cannot pass through the latter, forming a barrier of enzymes around all the cell. As a result, any penicillin molecule that enters will encounter a higher concentration of β-lactamase and they can effectively hydrolyze 1000 penicillin molecules per second. [9]

After penicillin was introduced, bacterial resistance to that was detected and eventually shown to be due to the production of these enzymes. To inactivate them, mechanism-based inhibitors such as clavulanic acid (Figure 1.5) were introduced in the 1980's to clinics. [9]

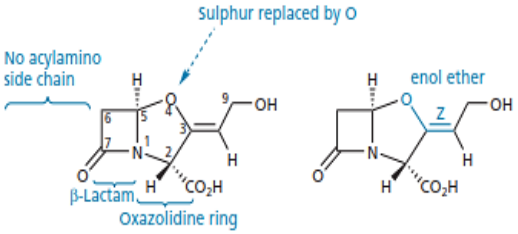


Figure 1.5 - Structure of Clavulanic acid/Clavulanate. [9]

Clavulanic acid has a slight antibiotic effect when used as inhibitor *per se*, which means that it is often administrated in combined therapies with several penicillins, such as amoxicillin. Nowadays, this medication is used as a first approach to treat mild cases of bacterial infections. [8, 9]

1.2.1 Types of classification

β-lactamases are increasing in number due to its high rate of mutations, and thus the diversification of this group of enzymes that inactivate antibiotics is an actual key point in the AMR area. [4]

Two major classification schemes exist for categorizing β-lactamases considering the characteristics of the enzymes and their substrate profiles: [5 - 7, 12, 13]

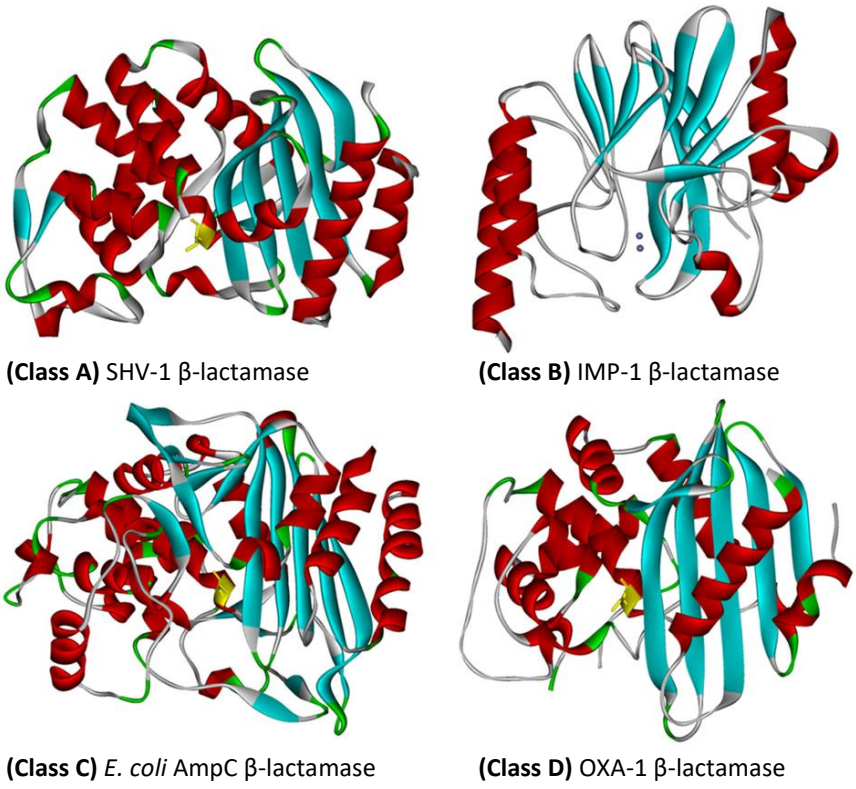
1) Bush-Jacoby-Medeiros classification is based on the biochemical and functional characterization (substrate and inhibitor profile) with 4 different groups from 1 to 4, and also sub-groups (from a to f).

2) Ambler classification is based on their amino acid sequence homology with four classes: A is the serine β-lactamases, B the metallo-β-lactamases, C the cephalosporinases, and D is for the oxacillinases.

Presently, the second is the most used and facilitates the addition of new discovered enzymes. In the revised Ambler nomenclature, classes A, C and D contain serine enzymes that have a high level of sequence conservation and whose reactions involve an acyl-enzyme intermediate, while class B do not form such intermediates. [5 - 7, 4, 13]

Class A β-lactamases, where TEM-1 is included, generally prefer penicillins as substrates and are susceptible to β-lactamase inhibitors (as clavulanic acid). [5 - 7, 13]

Figure 1.6 allows to have a clearer idea of the differences of folding and what type of secondary structure predominates between all Ambler classes, with an example of structure for each class. [13]



(Class A) SHV-1 β-lactamase **(Class B)** IMP-1 β-lactamase

(Class C) *E. coli* AmpC β-lactamase **(Class D)** OXA-1 β-lactamase

Figure 1.6 – Structural representations of the four different Ambler classes. In classes A, C and D the active-site serine is highlighted in yellow. Secondary structure is differentiated with colors: red for α-Helix; blue for β-Sheet; grey for loops/unstructured regions. Based on PDB entries 1SHV, 1DDK, 2BLS, and 1M6K, respectively. Adapted from [13]

These four representations show a higher structural similarity between class A, C, and D. Particularly, various TEM enzymes family have conservation of their folding, regardless of whether they are from gram-negative or gram-positive bacteria. [10] Besides, classes A and D present the largest number of unique enzymes of serine β-lactamases, so it supports the need to look forward and find new inhibitors for these two β-lactamase classes. [5 – 7]

1.2.2 The genetics role

The spread of β-lactamases is mostly plasmid mediated on the clinical environment, and the genes for some enzymes – TEM-1 included – are primarily found on mobile elements named transposons.

Beyond the resistance acquired by mutation, a second way in which bacterial cells can be drug resistant is by gaining resistance from another bacterial cell. This occurs because the genetic information required to synthesize β-lactamases can be passed, rendering bacteria resistant to penicillins. This problem is particularly prevalent in hospitals where several types of infections caused by gram-positive bacteria are resistant to antibiotics such as penicillin and tetracycline. [9]

1.3 Serine β-lactamases: structure and function

1.3.1 TEM-1 β-lactamase: origin and nomenclature

TEM-1 β-lactamases are capable of hydrolyzing penicillins and first generation cephalosporins. *Escherichia coli* is one of the organisms that hyperproduce TEM-1, more exactly around 50% of *E. coli* produce these enzymes. They are also found in *Klebsiella pneumoniae* and other members of the *enterobacteriaceae* family. [5 - 7]

This enzyme was first discovered in the mid-1960's in Greece, in a single strain of *E. coli* isolated from a blood culture from a patient named *Temoniera*, hence the designation **TEM**. The number present after the acronym is attributed by sequential order of its discovery (isolation and identification from patients), with this rules valid for the other classes. [16] TEM β-lactamases were the first to be identified in penicillin-resistant clinical isolates, and they continue to be the most common today. Serine β-lactamases have around 1300 entries on β-lactamase Database, Structure and Function (BLDB), of which 140 are TEM-1 enzymes and from them 76 structures are deposited in Protein Data Bank (PDB). [4] So, the knowledge of compounds and mechanism of inhibition of TEM's can have high importance since they are the most widespread cause of resistance to β-lactam antibiotics. [11]

1.3.2 Active-site architecture and catalytic mechanism

Over 140 variants of TEM β-lactamases showed only a significant homology for the [**Ser-X-X-Lys**] motif, that is observed in the majority of them. [9, 15] So, this serine-based hydrolytic mechanism is composed by an acylation of substrate, and then a deacylation step can proceeds with the presence of a hydrolytic water. That highly conserved lysine is obviaously important for hydrolysis. [10, 13 – 15]

The irreversible inhibition mechanism of a β-lactam substrate (e.g. penicillin) by a serine β-lactamase is shown in figure 1.7, in a simplified version.

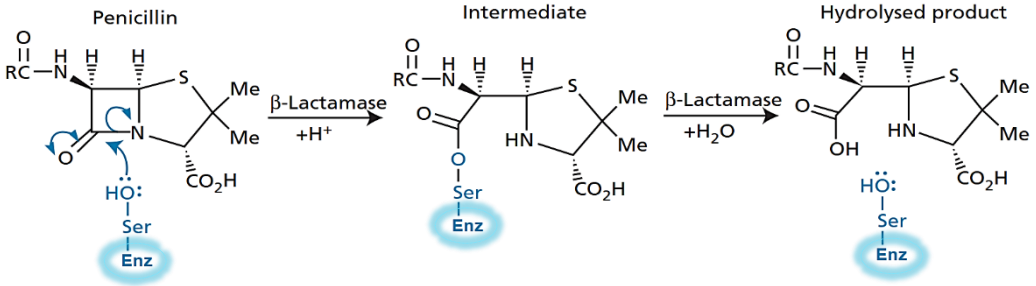


Figure 1.7 - Scheme of the inhibition mechanism by bacterial β-lactamase enzymes. The formation of the acyl-enzyme intermediate is a long-lived step. Adapted from 9 and 15

The detailed inhibition mechanism has the following steps: **1)** formation of an acyl-enzyme complex aided by a water molecule, followed by a nucleophilic attack by the (Ser)-oxygen of active-site on the carbonyl of substrate which results in a tetrahedral acylation intermediate; **2)** protonation of a nitrogen from the substrate and subsequent cleavage of the C-N bond of the β-lactam ring, opening it; **3)** a water molecule attacks the covalent complex and create a new tetrahedral deacylation intermediate; **4)** final hydrolysis between the carbonyl of the substrate and the nucleophilic oxygen of serine, regenerating the active-site and releasing the final inactive β-lactam. [9, 10, 13 - 15]

This mechanism is supported by quantum mechanics and an ultra-high resolution (0.85 Å) structure of a complex with an acylation transition state analog (PDB entry 1M40). This β-lactamase mechanism is effective against most of those enzymes, but not against all. [13]

1.4 β-lactam antibiotics: fighting β-lactamase production

Almost as soon as a new β-lactam antibiotic is introduced into the medical practices, some previously unrecognized β-lactamases with the capability of destroying that antibiotic activity are identified, thus making them a serious threat to public health. Moreover, β-lactams are the largest and most common group of antibacterial agents used worldwide. This type of antibiotics are mainly semi-synthetic compounds, produced from fungi and bacteria. They can block the transpeptidation of the cell wall's peptidoglycan, through inhibition of penicillin-binding proteins (PBPs), however the exact mechanism how it leads to cell death is still unknown. Some of them have a very narrow antimicrobial spectrum, while others have a very broad spectrum to fight ESBLs. [8]

Resistance against β-lactams is primarily mediated by a structural change that brings lower drug affinity, or by production of β-lactamases to cleave the β-lactam ring. Yet, this process is not quite simple and there are cases of bigger susceptibility depending on the quantity of porins present in the outer membrane, and others where the decreasing of porin production and β-lactamases overproduction can result in enhanced resistance. [9] To overcome β-lactamase-mediated resistance it is generally used inhibitors of β-lactamase, and these inhibitors usually act as suicide substrates. There are three β-lactamase inhibitors in current usage: clavulanic acid (Figure 1.5), tazobactam and sulbactam. Mainly, the inhibitors are like penicillin substrates by retaining the amide bond of the β-lactam group, but with a certain modification in the side chain, and these features allow the β-lactamases to be inactivated irreversibly. [8, 9] As mentioned before, they show a weak antibacterial activity when are administered alone and consequently, they are co-prescribed with β-lactam antibiotics. Several combinations are commercially available, although some are only available for severe infections due to their oral bioavailability limitations. [8, 9]

1.5 Inhibitory mechanism: reaction in two possible ways

It is reported and highly studied that most inhibitors of enzymes with a catalytically serine residue, such as proteolytic enzymes or transpeptidases, are based on peptidyl fluoroketones or ketones strongly linked to an electron leaving group. As alternative strategy, acylating agents can be used that will generate an acyl enzyme which is not able to turnover, making the desired inhibition. Generally, the nucleophilic hydroxyl group of the serine is the one that is acylated and this type of inhibition is well described for the monocyclic β-lactams.

Furthermore, sulfonation of serine enzymes is mentioned as an interesting alternative to common

acylation mechanism but it is still not well explored because sulfonyl derivatives are much less reactive than their acyl counterparts. However, more recently, some studies and assays with β-Sultams, cyclic compounds similar to the β-lactam ring but with a sulfonyl group as substituent. They are great candidates to explore the mechanism of sulfonation and its possible inhibitory role, with the probable ring opening of β-Sultams giving the sulfonate ester. The formation of sulfonates enzymes could itself promote the inactivation due to its resistance to hydrolysis or the sulfonate ester could undergo a further reaction giving rise to loss of activity. [17]

These two inhibitory mechanisms are represented in figure 1.8, where is present the general reaction for each one: **(1)** Acylation - attack on the carbonyl (reversible); **(2)** Sulfonation - attack on the sulfonyl (irreversible). [17]

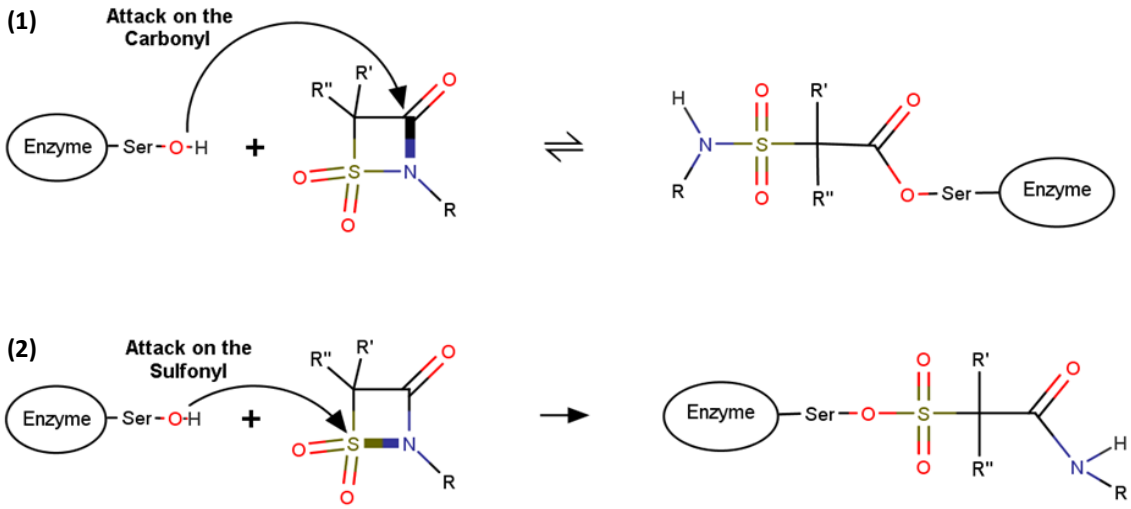


Figure 1.8 – Mechanism of inhibition through enzyme **(1)** acylation and through enzyme **(2)** sulfonation. The substrate presented is a β-Sultam, structurally similar to the β-lactam ring. The structures and arrows were produced with the MarvinSketch© (from *ChemAxon*) open-source software. Adapted from 17

One of the possible explanations for that, chemically speaking, is because the reaction of nucleophilic substitution in the acylation proceeds through the formation of an unstable tetrahedral intermediate, while the mechanism for sulfonyl group transfer involves an intermediate with a stable trigonal bipyramidal geometry. The sulfonation mechanism is well supported by structural data from X-ray crystallography and biochemical data from enzyme activity assays, performed for porcine pancreatic elastase (PPE). [17]

2. Introduction to methodology

Next, it will be introduced the main methodology used along this Thesis: the crystallization process itself and the fundamentals of X-ray crystallography.

2.1 X-ray crystallography

X-ray crystallography continues to be one of the most useful and successful techniques for the determination of three-dimensional structures of biological macromolecules, such as proteins, protein complexes, virus and viral particles, and nucleic acids. First, the process of structure determination requires single and well-organized crystals from purified protein or other macromolecule under study.

2.1.1 Background of crystal growth

A crystallization experiment typically begins with the sample in a stabilizing buffer solution and possibly other additives such as, salt, reducing agents, and other additives (inhibitors, cofactors, substrate/products or their analogues), known to stabilize the protein.

In macromolecular crystallography, crystals production is the first and most unpredictable stage, since it includes a series of thermodynamic and kinetic variables to produce a unique crystalline three-dimensional packing. Precipitants are added to the sample solution to reduce its solubility until supersaturation is achieved, leading to nucleation where small crystals will be formed or then to an amorphous precipitate. So, it is important to understand the best protein/precipitant ratios to have good transitions between the nucleation and growth stages. [18]

Often many hundreds of crystallization conditions are tested, mainly using commercially available crystallization screens to test a wide variety of initial conditions to search for the best condition to crystallize the sample. A systematic search is done by varying in a conjugated manner parameters like the protein concentration, precipitating agents, buffers, pH, temperature, ionic strength or add some additives. Most laboratories already have crystallization robots that can make nanoliter volume drops to perform the screens and optimization trials and thus they can test a huge number of conditions without spend a lot of protein. [18]

Prior to mixing the sample with precipitant, sample solution is undersaturated with the biomolecule(s) in study, and in an undersaturated sample solution no nuclei can be formed. Upon addition of precipitant, the supersaturation of the sample is increased. Thus, the water diffusion will happen from the drop to the reservoir until the concentration of the precipitant in the drop equals the one in the well. So, the concentrations of both protein and precipitant will slowly increase in the drop.

The solution should become supersaturated in the drop with protein, then nucleation can occur, and the crystals grow. All these steps can be associated to a phase diagram (Figure 2.1) with four divided zones: 1) stable or undersaturated – where crystal nucleation and growth cannot happen, having clear drops; 2) metastable or supersaturated – where the growth of crystals is possible, but nuclei are not viable; 3) labile or supersaturated – where nucleation are highly promoted and crystals can grow too; 4) precipitation zone – where drop solution forms precipitates, and neither nucleation nor growth are possible. [18, 19]

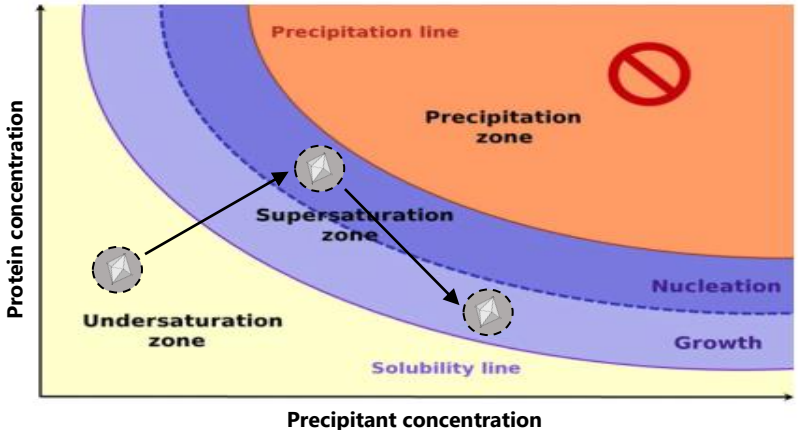


Figure 2.1 – Protein crystal phase diagram. Yellow represents the region where the solution is not saturated with protein, light purple is the growth phase, purple matches the nucleation events, and orange region is the precipitation zone. An ideal crystallization experiment is when the protein starts in a supersaturated stage (metastable zone) where the nucleation can occur and next should go to the growth region. Adapted from 18, 20

2.1.2 Crystallization techniques

The methods for crystallization include vapour diffusion – hanging or sitting drop – that is the most common one, but also microbatch, microdialysis and free-interface diffusion. Since the crystal appearance is a stochastic event, different methods to promote the equilibrium can be necessary to have higher chances of success in this process. Figure 2.2 illustrates in more details the vapour diffusion method, where the protein/precipitant solution will equilibrate in a sealed system with a larger aqueous reservoir whose precipitant concentration is optimal for producing crystals. [18, 20]

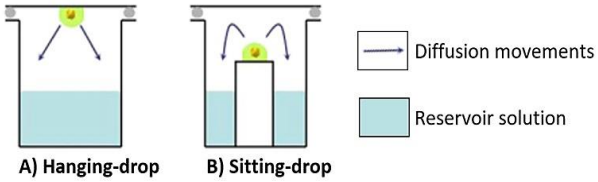


Figure 2.2 – Crystallization methods of vapor diffusion techniques. The drop has the protein depicted in orange, in green is the mixture of reservoir and protein, in blue is the reservoir, and the arrows represent the direction of diffusion. [18]

Choose the crystallization method is of high importance because the surface area, tension forces of the drop and the shape itself affect the rate of equilibrium and the number of nuclei formed. The search for an ideal equilibrium time is imperative because longer duration of equilibrium tends to from micro-crystals and lower time produce amorphous precipitates. [21]

After finding promising initial conditions in the positive hits of the screens tested, it is necessary to do a scale-up of the crystallization volumes, in similar or equal conditions, to obtain bigger crystals. Protein crystals have typically about 0.1 – 0.2 mm, but smaller ones can be measured at synchrotron facilities using proper microfocus beamlines. However, sometimes it is not straightforward to obtain crystals with enough dimension to be exposed to X-rays and so it is necessary to do crystal optimization trials, sometimes achieved by using seeding techniques like streakseeding and microseeding. Once nucleation is essential, these methods will promote it via the addition of seeds/nuclei from crystals of previous assays under close conditions and placed in the metastable region, where the growth in a slow and controlled way will happen. Both techniques will be described in more detail below.

- **Streakseeding** (Figure 2.3):

It is applied a thin seeding tool - generally a cat whisker or a synthetic fiber - on the crystalline material, with a single streaked touch, to remove and transfer tiny crystals (seeds) and then swiped across the new drop(s) that will support the growth of larger crystals. Sometimes more than a few passages are necessary through several drops until a small number of crystals be obtained. [18, 22]

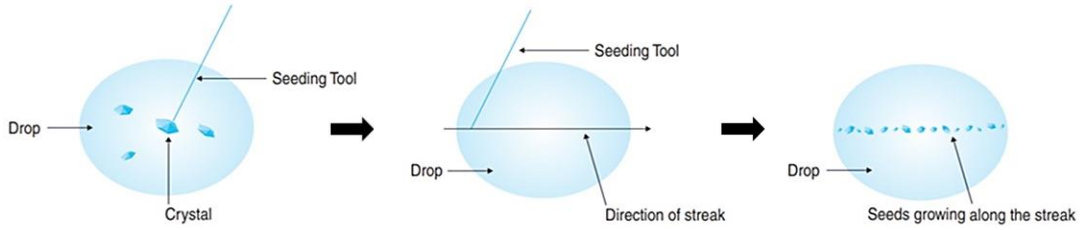


Figure 2.3 – Illustration of the procedure for the streakseeding technique. Scheme of the technique in which the seeding tool used is adapted from *Hampton Research* (HR8-133). [22]

♦ **<u>Microseeding</u>** (Figure 2.4):

For this, the first step is to prepare a seed stock solution by crushing previous grown crystals in the drop well with an appropriate tool, e.g. a crystal crusher (*Hampton Research*). If you have plenty of crystals, next is used a seed bead (can be made of steel, plastic, ceramic or glass) to create a stock of micro seeds for performing subsequent seeding experiments. That solution will be used to prepare serial dilutions and so is added to a clear drop to promote nucleation and subsequent growth. [18, 19]

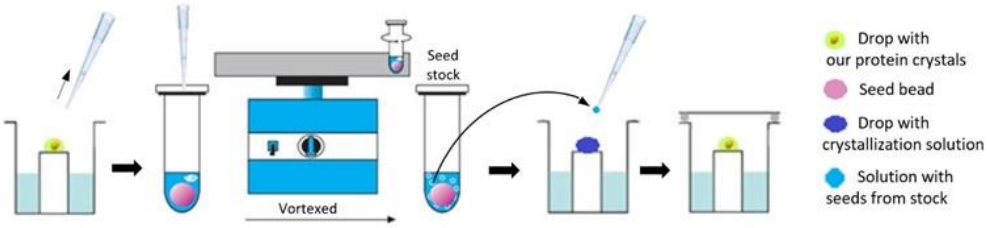


Figure 2.4 – Illustration of the procedures for the microseeding technique. Scheme of the technique in which is used a seed bead (*Hampton Research*) to produce the seed stock. Adapted from 18, 19

When crystallization conditions are optimized for native crystals, it is possible to give a step forward to studies with complexes, having the protein interacting with: ligands (usually organic compounds, e.g. a drug candidate), cofactors, substrate analogues, peptides, other proteins or even nucleic acids. Structures of a protein–ligand complexes can be achieved by soaking or co-crystallization (Figure 2.5):

a) <u>Soaking</u>

It is based on the diffusion of small molecules that can pass through the solvent channels, that typically have about 20 – 100 Å wide, of a native crystal already formed; the solution with the ligand is placed directly in the crystal drop and then sealed again. [18]

b) <u>Co-crystallization</u>

This second strategy requires a pre-incubation of the protein with the ligand/molecule desired, mostly at room temperature, in order to promote their binding before making the drops. It can be more useful than soaking when the ligands are too big to diffuse through the solvent channels. [18]

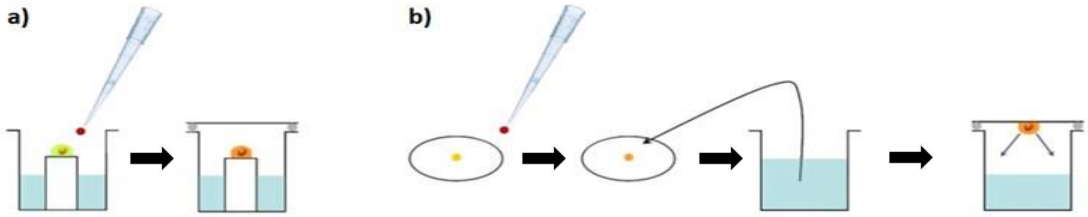


Figure 2.5 – Techniques of **a)** Soaking and **b)** Co-crystallization to produce crystal complexes. Schemes are colored with blue for the reservoir, red for ligand solution, yellow for protein solution, orange for protein-ligand complex. Adapted from 18

It should also be highlighted that the crystallization conditions to obtain protein–ligand complexes may vary from those obtained for the protein alone. [18]

2.2 X-ray diffraction and data collection

Next, the chosen crystals are harvested from the drop and mounted on a microscopic loop, usually made of a nylon fiber. When crystals are first placed on the diffractometer the initial tests of crystal quality are performed and its preliminary diffraction pattern, crystal lattice and resolution limit are checked. The crystal will then proceed to X-ray exposure during a rotational angle enough to have a complete data set. This process will produce free radicals that could damage rapidly the crystal and to avoid that the measurements can be done under a cryogenic nitrogen gas stream (≈ -173,15ºC), keeping the low temperature during data collection. Consequently, complete data sets can be recorded from a single crystal. Usually, crystals have a solvent content of 50%, but can vary from 20% to 80%, thus the water molecules should be replaced by cryoprotectant solutions – which can contain glycerol, PEGs, glucose or other sugars – otherwise upon flash freezing ice will be formed inside the crystal and the diffraction quality of the crystal will be affected. [18]

Briefly, a crystal is a periodic and tiny structure composed by unit cells that are the smallest repeating units that can produce one crystal just by doing translation operations. Related to the unit cells, each one contains the asymmetric unit – smallest fraction that can be subjected to rotations and translations (using only symmetry operators allowed in the crystallographic symmetry) to generate a unit cell. [18] Inside a crystal, each constituent atom is defined by a set of three-dimensional Cartesian coordinates, *x*, *y*, and *z*. It is mandatory defined one of the vertices to use as the origin of the unit cell's coordinate system (x, y, z = 0, 0, 0), and there are also three angles $(\alpha, \beta \text{ and } \gamma)$ defined by each two planes of the unit cell (Figure 2.6 – 1 and 2). When it comes to protein crystals, there are only seven crystal systems (Figure 2.6 - 3), composed by the different possibilities of the angles and the edges dimensions. [18, 20, 23]

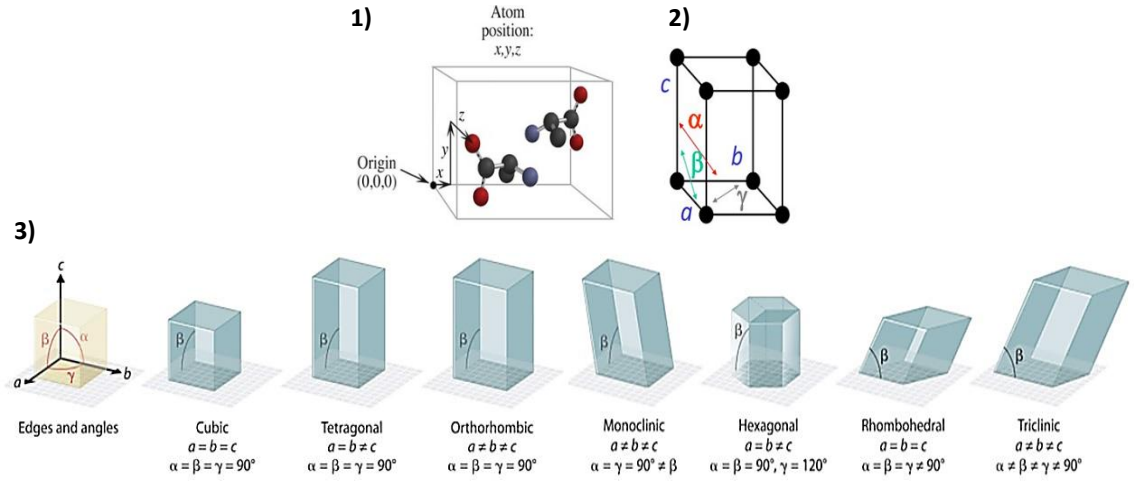


Figure 2.6 – **1)** and **2)** are general representations the unit cell and **3)** has schemes for the seven possible systems for protein crystals, depending on the edges and angles of the unit cell. Adapted from 20, 23

Many X-ray sources are available but two of them are the most used: rotating anode X-ray tube, micro-source X-ray tube, and the synchrotron x-ray sources. The latter one is what provides X-ray beams with higher intensity and provides the possibility of tunable energy (select a specific wavelength is important in some cases for anomalous dispersion experiments). Regarding the detectors, there are several types but the most used are the CCD (charge-coupled device) detectors or the most recent hybrid pixel array detectors (like PILATUS or EIGER detectors) often present at synchrotron facilities.

Since the crystal is rotated relative to the X-ray beam, a complete diffraction pattern can be recorded on the detector generating the diffraction spots (Figure 2.7). A single diffraction data set can contain tens to hundreds of thousands of reflections, and sometimes more than one data set is needed per crystal. [18, 20]

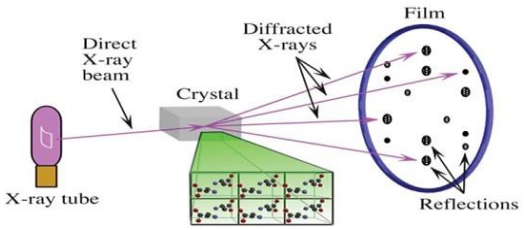


Figure 2.7 – Process of data collection in X-ray crystallography. The crystal diffracts the source of X-ray beams into a set of reflections (diffraction spots) that are recorded on the detector. [20]

X-ray crystal diffraction is only possible because the X-rays wavelength is from the same magnitude order of the bond distance between atoms (≈ 1 – 1.5 Å). When electromagnetic radiation from X-rays interacts with the electrons from atoms present in the crystal, wave's interference of two types can be produced: constructive and destructive. The first is the one that produces spots on the diffraction patterns, with different intensities (I_{hkl}) according to Bragg's law: $\lambda = 2d.\sin\theta$. It has associated a reflection angle (θ), the resolution of the diffraction in Ångström (d), θ is the diffraction angle and λ the wavelength of the radiation used. There is mandatory for this type of interference that the waves be in phase. [18]

After obtaining a few test images at 0° and 90° about the perpendicular axis regarding the beam source, the indexing step consists in determining the reciprocal unit cell. This will allow elucidation of the Bravais lattice and space group identification. It is possible to list the following options for centered unit cells and its matching symbol: **P** – Primitive (no centering); **A** – A-face centered; **B** – B-face centered; **C** - C-face centered; **F** - All-face centered; **I** - Body centered; **R** - Rhombohedrally centered. The data obtained after diffraction is considered the experimental raw data, from which the structure is derived. Following data collection, all images are computationally integrated, and intensities are scaled. Combining the diffraction patterns in many different crystal orientations allows to produce a three-dimensional dataset from the original bi-dimensional results. [20, 24]

2.3 Phase determination

The atomic positions of the crystal structure are defined in the three-dimensional space by the coordinates (x, y, z) and their matching diffracted beams are defined by Miller indices (h, k, l). Each diffracted beam is defined by a vector sum of the individual contributions from all atoms that are characterized by a phase (α_{hkl}) and an amplitude $|F_{hkl}|$. The amplitude is related to the number of electrons of each atom, and the phase is resultant of the atom position within the unit cell. Yet, the information of the phase for all dataset is lost during the process of data collection. [18, 20]

An important concept to be introduced at this moment is the structure factor (F_{hkl}). It is a mathematical description of the contribution of all the atoms in the unit cell to the total of scattered radiation, at each reciprocal lattice point (h, k, l), after interaction with the structure and is represented by: $F_{hkl} = \sum h \sum k \sum l\, |F_{hkl}| \exp(i\boldsymbol{\alpha_{hkl}})$, where the amplitudes of the structure factors $|F_{hkl}|$ come from experimental measurements, and with i being the imaginary unit. The phase α_{hkl} is

lost during the experiments and therefore there are methods to estimate that information. After this problem is solved (the last challenge towards a 3D-structure of macromolecule), it is possible to determine the electron density map (p) for any coordinate within the unit cell. The equation is represented by: $p(x,y,z) = \frac{1}{V}\sum h \sum k \sum l \ |F_{hkl}| \ \exp(i\alpha_{hkl}) \exp[-2\pi i(hx+ky+lz)]$. So, it is critical to determine the phases to proceed with the calculations since the electron density is a function of the structure factors. Also, the mathematical correlation between electron density and structure factors is the Fourier transform. [18, 20]

There are three main strategies for obtaining the phase information via indirect methods: molecular replacement, anomalous scattering and isomorphic replacement. Also directs methods, named *Ab initio* method, may be an option too if the dataset has atomic resolution. About the indirect methods, the molecular replacement (MR) method – one of the most used – is based on an identification of a search model through sequence alignment with known structures that will be the base to estimate the phases. In this case, one main rule for its total success is that the sequence identity should be above of 25% to guarantee the two proteins are structurally homologues. Another way to do the phase determination is using the anomalous scattering method which is possible if the energy of the X-ray beam is close to the natural absorption energy of an atom within the crystal structure, and the presence of heavy atoms in the protein confers that ability (here the selenomethionine is very useful). These modifications might be used to estimate the phases, and more than one wavelength can be used. It requires an accurate collection of information related to the diffracted intensities. If it uses just one wavelength that corresponds to the heavy atom peak is called single anomalous dispersion (SAD), but using two or more wavelengths is named multiple anomalous dispersion (MAD) that typically includes the peak of the heavy atom but also data of the absorption curve and a remote wavelength of high or low energy. [18]

The third method is the isomorphous replacement with interference effects on the intensities of the diffracted beams caused by the presence of heavy atoms (e.g. iron, zinc, molybdenum or silver) and it will allow to estimate the phase angles. Phase determination is done by at least one derivatized crystal where the heavy atoms are located on different positions inside the crystal or using changes relatively to native crystals. Often, multiple crystal derivatizations are needed to know the phase, named as multiple isomorphous replacement (MIR), but sometimes may be enough only one good derivative to obtain the structure, single isomorphous replacement (SIR). [18]

2.4 Model building and refinement process

Electron density maps are calculated after the phase determination, and an initial model is built. This model will need to be refined to improve the agreement between the experimental and calculated values of the structure factor amplitudes ($|F_{hkl}|$). The quality of the electron density maps is proportional to the quality of the diffraction data obtained and their phase estimation. So, the examination of these maps is done throughout refinement processes. The resolution of a model is defined by the level of detail associated to the electron density map and there are these four resolution levels: atomic ≤ 1.2 Å; 1.2 Å $\leq$ high ≤ 1.8 Å; 1.8 Å $\leq$ medium ≤ 2.5 Å; 2.5 Å $\leq$ low ≤ 3.0 Å.

Moreover, the refinement also optimizes factors/parameters related with the preliminary structural model that contains errors related with stereochemical restraint and geometry. Several computational strategies are available, and they are applied as an iterative process of cycles of refinement against the improved electron density map until the convergence be achieved.

Generally, parameters such as the temperature factor/B-factor that describe the atomic displacements, atomic positions and occupancies are usually refined. In the end, a suitable three-dimensional model of the studied protein is obtained. [18, 20]

2.5 Model validation

It is important to proceed with a quality assessment and validation of the model. Several statistical parameters are often reported and used to evaluate the diffraction quality of the dataset, including:

- **Resolution (Å):** defines the level of detail associated to the electron density map.
- **Multiplicity**: ratio between the total number of measured reflections and the number of unique reflections existing within the resolution range. This will give the average number of independent measurements for each reflection.
- **Completeness** (%): relation between measured reflections and the total number of theoretical unique reflections. The goal is near to 100%.
- **Ratio of signal-to-noise ($I/\sigma(I)$)**: explains how many times above noise level was done the measurements of intensities of the diffracted beams.
- R_{merge}: measures the agreement between the multiple independent observations of the same reflection. This value should be low in the case of good datasets, so high values possibly indicate problems with the data processing.
- R_{work}: measures the discrepancy between the observed/experimental (F_{obs}) structure factor amplitudes and the calculated (F_{calc}) ones from the model.
- R_{free}: R-factor that also uses structure factor amplitudes values from a 'test-set', a set not used during the refinement, that is used for cross-validation and it shows how the current atomic model is comparing to that subset of measured structure factor amplitudes that were not present in the refinement calculations. The higher the resolution, the lower is expected to be the R-factor. When is a structure with atomic resolution, it is desirable that R_{free} might be at least $1-2\%$ higher than R_{work}. [18]

Model geometry validation can be achieved with tools that will do a detailed analysis on stereochemistry, local chemistry environments, among others. Besides, other properties of the residues are usually checked, such as: bond angles and lengths, Φ (phi) and Ψ (psi) torsion angles, side-chain torsion angles, peptide flips, C_β deviations, or Ramachandran and rotamer outliers. [18]

Figure 2.8 illustrating the steps involved in a pipeline of macromolecular structure determination.

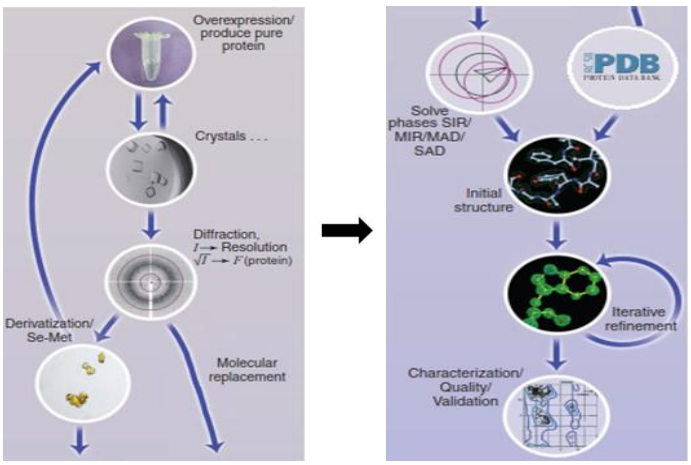


Figure 2.8 – Diagram with the pipeline, from top of image on the left to bottom of image on the right, for macromolecular structure determination. Blue arrows represent the possible ways to follow a strategy in this pipeline. Adapted from 25

3. Materials and methods

Solutions and buffers preparation that needed to have a defined pH value were prepared and their pH adjusted using a pH meter (pH 1100L, *VWR*). The compounds were weighted on a precision balance or an analytic balance (model 440 and ADB, both from *KERN*). All steps that needed sterile conditions were performed in a laminar flow chamber (*Telstar*) or near a Bunsen burner. All the measurements of optic density at 600 nm (OD_{600}) were done in a spectrophotometer (Ultraspect 10, *Amersham Biosciences*), with LB media used as blank.

3.1 Techniques of DNA manipulation: cell transformation

Professor Christopher Schofield, Department of Chemistry, University of Oxford (UK), kindly provided a plasmid pQE-30 (*Qiagen*) with the coding sequence of TEM-1 β-lactamase (UniProt P62593) without the signal peptide sequence (residues 24–286), with a *His-tag* at the N-terminus, one disulfide bridge, and theoretical molecular weight around 32 kDa. [26] Competent cells of *Escherichia coli* (BL21 (DE3)) were transformed and then placed in solid growth medium (LB-Agar) supplemented with 100 mg/mL Ampicillin (*Carl Roth*), the resistance marker provided by the plasmid, at 37°C. After ≈16 hours of incubation, isolated colonies were observed in the Petri dishes.

In parallel, a sample of DNA was prepared and sent to *GATC Biotech* to confirm the nucleotide sequence of the pQE30-TEM1 plasmid by Sanger sequencing. The sequence obtained with plasmid plus the gene of interest is presented in the Appendix A.

3.2 Protein expression

In the context of this work, one isolated colony from freshly transformed cells (pQE30-TEM1 in *E. coli* BL21 DE3) were grown overnight, at 37°C and 220 rpm, in LB medium supplemented with 100 mg/mL Ampicillin. Cell growth was initiated with a 1:10 dilution of overnight saturated culture in LB medium supplemented with proper antibiotic, at 37°C and 220 rpm. Gene overexpression was induced by the addition of 0.2 mM IPTG (*Carl Roth*) to the media when cells reached an $OD_{600} \approx 0.6 - 0.7$. Temperature was then lowered to 28°C and the cells grown for another $4 - 4.5$ hours at 220 rpm. Cells were harvested by centrifugation at 6 000 rpm at 4°C, for 20 minutes, and stored at -80°C. The yield of cell pellet was about 30 grams of cells from 4 liters of culture.

3.3 TEM-1 purification

Thawed cells were resuspended in 50 mM HEPES buffer pH 7.5, 500 mM NaCl, 5 mM Imidazole pH 7.5, 5 mM $MgCl_2$, 25 U/mL of DNAse (BitNuclease, from *Biotool*) and one tablet of Pierce™ protease inhibitor (*Thermo Scientific*) and homogenized using a rotor-stator homogenizer. Cell lysis was performed using a cell disruptor (APV® 2000 model, from *SPX*) at 800 bar ($\approx 11\,603\ psi$) and doing three passages. Disrupted cells were separated from cell debris by centrifugation at 30 000 g at 4°C, for 20 minutes. Pellet was discarded and supernatant containing the soluble proteins was kept.

3.3.1 1st batch of protein

The soluble cell lysate was injected in a 5 mL Histrap™ fastflow crude (*GE Healthcare*) affinity chromatography column. The protein was eluted with an imidazole concentration linear gradient from 5 to 500 mM, in 50 mM HEPES pH 7.5, 500 mM NaCl, 0.2 mM PMSF (Phenylmethylsulfonyl fluoride) and 5% (V/V) glycerol. Protein purity was assessed with a 15% (V/V) SDS-PAGE and the most enriched fractions were pooled together. The fractions were then concentrated and further purified using a 20 mL HiLoad 16/10 Q-Sepharose high performance (*GE Healthcare*) ionic exchange chromatography column. This second purification step was performed with a NaCl concentration linear gradient from zero to one molar, in 20 mM HEPES pH 7.5, 0.2 mM PMSF and 5% glycerol. Both chromatography columns were connected to an ÄKTAprime plus system (*GE Healthcare*). All steps were performed at 4°C. The fractions containing the target protein in high quantity and purity were pooled, concentrated and its concentration was determined mainly by Nanodrop™ measurements or by Bradford method. A second pool of protein was made, which had TEM-1 protein in good quantity but with slight contaminants (less pure fractions from Q-Sepharose purification). This pool was submitted to a third purification step, a size exclusion chromatography, using a Superose™ 12 10/300 GL column (*GE Healthcare*). Final buffer concentrations were 20 mM HEPES pH 7.5, 150 mM NaCl, 0.2 mM PMSF and 5% glycerol. The protein concentrations were 33 mg/mL (in ≈200 µL) and 13 mg/mL (in ≈40 µL), for pool 1 (the purer) and pool 2 (less pure, after the size exclusion chromatography), respectively.

The described purification strategy was based on previous work [26, 27] but with some modifications to optimize the protocol.

3.3.2 2nd batch of protein

Furthermore, a second batch of TEM-1 (starting from 26.5 grams of cells in 4 liters of culture) was produced using the same three columns but at this time both protein-enriched pools, after ionic exchange chromatography, were injected on a size-exclusion chromatography, originating then three different pools, according to their purity in the end of the third purification step, namely: pool 1.0, pool 2.0 (the purer) and pool 2.1. Results concerning this purification are present in Appendix B, named as 5th attempt of purification.

3.3.3 Protein purity (SDS-PAGE & Western blot)

After each purification step, the protein purity was assessed by SDS-PAGE and Western blot.

For SDS-PAGE analysis, it was used a Mini-PROTEAN® Tetra cell Electrophoresis System (*Bio-Rad*). The samples were previously prepared in 4x Laemmli Sample Buffer (*Bio-Rad*) and then denatured at 95°C for 10 minutes. The gel was run at 100 Volts (V) for 15 minutes and then at 160 V for 1h20, for the stacking and resolving parts of the gel, respectively. No amperage limit was imposed in the system during the run. The molecular weight marker used in these experiments was the Precision Plus Protein™ Dual Color Standards (*Bio-Rad*). Protein bands were stained with 1x Bio-Safe™ Coomassie Stain (*Bio-Rad*). The recipes of the gels and buffers are available in table A.1 of Appendix A.

Western blot was carried out on some gels. First, each gel is incubated for 3 minutes in 1x Transfer buffer while a PVDF membrane is activated with methanol for 30 seconds. Both are then put together in the TransBlot® Turbo™ (*Bio-Rad*), in a transfer cassette according to the figure 3.1.

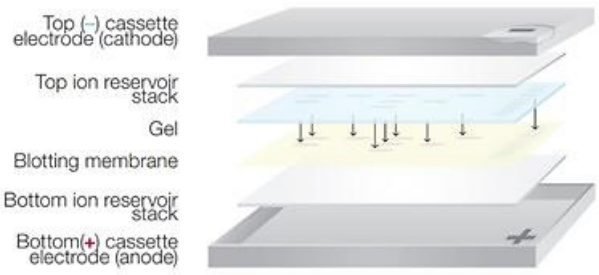


Figure 3.1 – Scheme of the correct assembling of the Western blot sandwich. Adapted from the quick start guide of Trans-Blot® Turbo™ Transfer system, from *Bio-Rad*.

Blotting was achieved running the preprogrammed protocol "Mixed MW" for 7 minutes and up to 25 V. The membrane was removed from the sandwich, washed one time in 25 mL 1x TBS-T for 10 minutes and blocked with 3% (w/V) of low fat powder milk that was prepared in 25 mL 1x TBS-T, at room temperature for 1 hour in a roller shaker. Next, is washed twice in 25 mL 1x TBS-T, for 10 minutes and incubated with 10 mL of a (1:5000) dilution HisProbe-HRP (*Thermo Scientific*) in 1x TBS-T for 1 hour. After, the washing steps are repeated two times in 10 mL 1x TBS-T for 10 minutes and incubated for 5 minutes with 10 mL (1:1 proportion) SuperSignal west pico chemiluminescent plus substrate (*Thermo Scientific*). Afterwards go to a ChemiDoc™ XRS+ machine (*Bio-Rad*) to observe the results. The information about the solutions used is available in table A.1 of Appendix A.

3.3.4 Protein quantification methods

Two methods were applied to quantify the protein samples:

1) Nanodrop measurements, in a NanoDrop™ 2000 and NanoDrop™ One spectrophotometers (*Thermo Scientific*), using 2 µL sample drops and the protein buffers as blanks.

2) Bradford assay, a colorimetric method followed at 595 nm, is performed with the Pierce™ coomassie (Bradford) protein assay kit (*Thermo Scientific*). First thing to do is to determine a calibration curve with a series of dilutions of a standard protein (in this case the BSA - Bovine Serum Albumin), with the same buffer of the protein to be analyzed. In total seven different dilutions of BSA were prepared, in triplicates, with one being the blank sample (no protein added). A 96-well immune flat-bottom plate was used with standards and protein samples. Each well had 150 µL of coomassie bradford reagent and 5 µL of respective standards dilution. A similar procedure was done to prepare the dilutions of the protein solution to be quantified, with the original protein concentration and the dilutions of 1:10 and 1:100. After 30 minutes of incubation in the dark, the absorbances of all samples were measured at 595 nm using a FLUOstar® Optima plate reader (*BMG Labtech*). Data were processed using the MARS Data analysis software (version 2.1) to obtain a calibration curve based on the mean value of the triplicates for each dilution. Therefore, it was possible to calculate the linear regression of that curve. The same rational was applied to obtain the mean value of the protein samples, so it was was possible to estimate, using the equation of the calculated linear regression, the concentration of protein in the samples.

3.4 X-ray crystallography: crystallization methodologies

3.4.1 Crystallization of recombinant TEM-1 [1st batch of protein]

The first approach, after obtaining the first batch of pure protein (pool 1), is to find the best crystallization conditions for TEM-1. Therefore, the first experiments were done based on literature of the most common TEM-1 crystallization conditions, after the results of a Protein BLAST between the sequence of our plasmid and the sequence of UniProt code P62593 (TEM-1 β-lactamase). The results of the alignment are presented in table D.1 of Appendix D.

In a 48-well screen plate (MRC Maxi type) from *TTP Labtech* and using the nano-drop robot (*Mosquito®, TTP Labtech*), it was performed the Quik Screen™ (*Hampton Research*) testing 24 different crystallization conditions of sodium-potassium phosphate (Na/KP_i) buffer in a range of different pH values (varying between 5.0, 5.6, 6.3, 6.9, 7.5 and 8.2) and different buffer molarities. It was done for protein concentration at 10 mg/mL and 20 mg/mL. All information about this plate is present on the next table 3.1.

Table 3.1 – Summary of crystallization conditions of the 48-well plate testing Quik Screen™ with the first protein batch.

Protein buffer	**20 mM HEPES pH 7.5, 150 mM NaCl**
Protein concentration (mg/mL)	10 20
Temperature (℃)	20
Technique	Sitting − drop
V_{Drop} (nL) ***(Protein + reservoir)***	(100 + 100)
$V_{Reser.}$ (μL)	200

No crystals appeared in this plate, just light precipitates at 10 mg/mL of protein and heavy precipitates at 20 mg/mL.

In parallel, other trials were done to test those conditions with (Na/KP_i) buffers. In a 24-well Linbro plate using the sitting-drop method and the solutions from Quik Optimize™ (*Hampton Research*), it was only done a pH screening from 7.6 to 8.2 and for buffer molarity from 1.0 M to 2.0 M. Two plates were done, one at 33 mg/mL of protein and another at 10 mg/mL. Protein buffer, technique and temperature were the same of previous table 3.1, and the volumes of the drop were 1+1 µL (Protein + Reservoir) and a reservoir with 500 µL. No crystals have appeared in those two plates.

Additionally, crystallization screens were performed with the *Mosquito®* robot and 96-well plates (SWISSCI 3-drop type) from *TTP Labtech*, such as: SaltRx (*Hampton Research*), JCSG-plus™ and BCS Screen (*Molecular Dimensions*). The protein and setup conditions are summarized in table 3.2.

Table 3.2 – Summary for setup conditions for crystallization experiments with SaltRx, JCSG-plus™ and BCS Screen.

Protein buffer	$\mathbf{20\ mM\ HEPES\ pH\ 7.5, 150\ mM\ NaCl}$
Protein concentration $(\boldsymbol{mg/mL})$	33 16.5 8.25
Temperature (°C)	20
Technique	Sitting − drop
$\boldsymbol{V_{Drop}}$ $(\boldsymbol{nL})$ ***(Protein + reservoir)***	$(100 + 100)$
$\boldsymbol{V_{Reser.}}$ $(\mu\boldsymbol{L})$	45

After 10 days, in condition 1-32 of BCS Screen, one crystal appeared at a protein concentration of 16.5 mg/mL in a reservoir solution with 28% V/V PEG smear broad, 5% V/V Glycerol, 0.05M L-Arginine and 0.05M L-Glutamic acid monosodium salt hydrate. A list with all conditions of this specific screen is presented in figure C.1 of Appendix C.

Next, a 48-well plate (MRC Maxi type) from *TTP Labtech* was done to scale-up and do a condition optimization from the initial positive hit from BCS Screen. This plate was made with protein at 16.5 mg/mL, the same technique, temperature and protein buffer of table 3.2, and the drop ratios were (500 + 500) nL and a 160 µL reservoir. The stock solutions used to prepare the reservoir were: 50% V/V PEG Smear Broad (*Molecular Dimensions*); 100% V/V Glycerol (*VWR Chemicals*); 1 M L-Arginine (*Carl Roth*); 0.2 M DL-Glutamic acid monohydrate (*Molecular Dimensions*). The optimization matrix of this plate is shown in figure C.2 of Appendix C. One crystal appeared after 7 days in the well E6.

Still to try to optimize the first positive result from condition 1-32 of BCS Screen, it was prepared a 24-well Linbro plate. Protein buffer and temperature were the same as before, but now testing both vapor diffusion techniques. A reservoir with 500 µL was prepared only with 28% PEG Smear Broad and 5% Glycerol, for all drops. The final volume of each drop was 2 µL and they were made as shown below in table 3.3, repeating this to three different protein concentrations (16.5, 12 and 8 mg/mL).

Table 3.3 – Summary of drops made in a 24-well plate to do the optimization from the initial hit in BCS Screen, repeated for three different protein concentrations at 16.5, 12 and 8 mg/mL.

	1	2	3
A **(Sitting-drop)**	$(1.0 + 1.0)\ \mu L$ [P] + **[x]**	$(1.1 + 0.9)\ \mu L$ [P] + **[x]**	$(1.2 + 0.8)\ \mu L$ [P] + **[x]**
B **(Hanging-drop)**	$(1.0 + 1.0)\ \mu L$ [P] + **[x]**	$(1.1 + 0.9)\ \mu L$ [P] + **[x]**	$(1.2 + 0.8)\ \mu L$ [P] + **[x]**

[x] = solution of BCS Screen, directly from the tube of condition 1-32, composed by: 28% V/V PEG Smear Broad, 5% V/V Glycerol, 0.05M L-Arginine, 0.05M L-Glutamic acid monosodium salt hydrate.

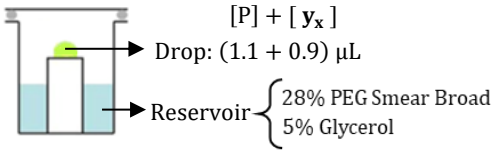
After 7 days, two crystals appeared in condition A2 for a protein concentration at 8 mg/mL.

More optimizations to replicate those last results of condition A2 were performed in order to obtain more crystals with good size and shape, by doing: a new 24-well Linbro plate with sitting-drop method

at 20°C, same protein buffer as before, a protein concentration at 8 mg/mL and a 500 µL reservoir with 28% PEG Smear Broad and 5% Glycerol. Drop ratios and volumes to add on each drop is explained in table 3.4.

Table 3.4 – Summary for a new 24-well plate to do an optimization after the previous results in table 3.3, with the volumes and solutions to add on each drop. The drop ratio was constant, 1.1 µL of protein and 0.9 µL of solution.

	1	2	3	4
A	[P] + **[y_1]**	[P] + **[y_1]**	[P] + **[y_1]**	
B	[P] + **[y_2]**	[P] + **[y_2]**	[P] + **[y_3]**	[P] + **[y_3]**



$\mathbf{y_1}$ = solution of BCS Screen, directly from the tube of condition 1-32, composed by: 28% PEG Smear Broad, 5% Glycerol, 0.05M L-Arginine, 0.05M L-Glutamic acid monosodium salt hydrate.
$\mathbf{y_2}$ = solution of reservoir.
$\mathbf{y_3}$ = "homemade" solution composed by: 28% PEG Smear Broad, 5% Glycerol, 0.05M L-Arginine and 0.05M DL-Glutamic acid monohydrate.

This plate did not produce positive results in the formation of any crystal. It was not simple to optimize the conditions to have native TEM-1 crystals, hence was also tried two seeding techniques to promote a quicker nucleation of the drops.

First a streakseeding, using as source of seeds the first native crystal obtained in condition 1-32 of BCS Screen (Table 3.2). By sitting-drop, 20°C and the same protein buffer as previous one, was prepared 8 drops to each protein concentration at 16.5 and 8 mg/mL. Drop ratio was again (1.1 + 0.9) µL for protein plus the solution directly from the condition of BCS Screen, and a 500 µL reservoir with 28% PEG Smear Broad and 5% Glycerol. Each line of the plate represents a different time where the drop was swiped across. Four drops were passed through, one at a time, two drops from each protein concentration. All details about this plate are summarized in table 3.5.

Table 3.5 – Streakseeding for optimization of TEM-1 native crystals. Protein concentrations tested were 16.5 and 8 mg/mL. Drop ratio was constant, 1.1 µL of protein and 0.9 µL of solution from BCS Screen, condition 1-32.

Time ↓	1	2	3	4
A **(After 3 hours)**	@16.5 mg/mL [P] + **[x]**	@16.5 mg/mL [P] + **[x]**	@8 mg/mL [P] + **[x]**	@8 mg/mL [P] + **[x]**
B **(After 7 hours)**	@16.5 mg/mL [P] + **[x]**	@16.5 mg/mL [P] + **[x]**	@8 mg/mL [P] + **[x]**	@8 mg/mL [P] + **[x]**
C **(After 22 hours)**	@16.5 mg/mL [P] + **[x]**	@16.5 mg/mL [P] + **[x]**	@8 mg/mL [P] + **[x]**	@8 mg/mL [P] + **[x]**
D **(After 26 hours)**	@16.5 mg/mL [P] + **[x]**	@16.5 mg/mL [P] + **[x]**	@8 mg/mL [P] + **[x]**	@8 mg/mL [P] + **[x]**

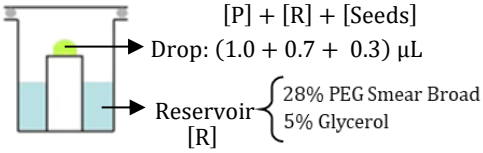
[x] = solution of BCS Screen, directly from the tube of condition 1-32, composed by: 28% V/V PEG Smear Broad, 5% V/V Glycerol, 0.05M L-Arginine, 0.05M L-Glutamic acid monosodium salt hydrate.

Once streakseeding did not produce any crystal, it was tried microseeding with the same crystal as source of seeds. Six drops were prepared, by sitting-drop, 20°C, same protein buffer as before, protein at 8 mg/mL, a 500 µL reservoir with 28% PEG Smear Broad and 5% Glycerol and drop ratio of

$(1.0 + 0.7 + 0.3)$ µL for the protein plus reservoir plus seeds solution. Each drop represents one of the dilutions prepared from the initial seeds stock, and one drop with the stock itself. This information is present in table 3.6.

Table 3.6 - Microseeding for optimization of TEM-1 native crystals.

	1	2	3
A	10^0 *dilution* $(seeds' \ stock)$ [P] + **[*x*]**	10^{-1}*dilution* [P] + **[*x*]**	10^{-3} *dilution* [P] + **[*x*]**
B	10^{-4} *dilution* [P] + **[*x*]**	10^{-5} *dilution* [P] + **[*x*]**	10^{-6} *dilution* [P] + **[*x*]**



These trials of the second seeding attempt also did not present new crystals.

3.4.2 Crystallization of recombinant TEM-1 [2nd batch of protein]

Before the crystallization screens with this new batch of protein, it was decided to firstly do a Pre-Crystallization Test (PCT™, from *Hampton Research*) to assess the best protein concentration to use. It was chosen to test only the solutions A1 and A2, with a drop ratio of $(0.5 + 0.5)$ µL for Protein + Reservoir, and a 200 µL reservoir. The best result pointed to use a protein concentration at 6 mg/mL. Afterwards, two crystallization screens were tested, BCS Screen – that already had produced one crystal – and PACT premier™ (*Molecular Dimensions*), using the best protein concentration estimated by the PCT. For each, in a 96-well plate (SWISSCI 3-drop type) from *TTP Labtech*, each drop had a different protein pool in a total of the three mentioned pools (sub-topic 3.3.2). The other conditions were the same of those in table 3.2. After 11 days, in condition 1-30 of PACT premier™ screen, just one tiny crystal appeared at a protein concentration of 6 mg/mL in a reservoir with 0.1 M PCTP buffer pH 9.0 and 25% w/V PEG 1500. Additionally, a complete list with the conditions of this screen is showed in figure C.3 of Appendix C. This second batch, specifically the pool 2.0 that originated one crystal, was also tested with similar crystallization conditions of those summarized in table 3.3, but only for the sitting-drop method. It was done 3 drops, with only a protein concentration at 8 mg/mL. In none of them appeared crystals.

3.5 X-ray crystallography: data collection

The two crystals that appeared in table 2 – condition A2 were plunged in a cryo-solution composed by 15% V/V glycerol and 28% V/V PEG smear broad and then one of them tested *in-house* with $Cu \ \mathrm{K}\alpha$ radiation ($\lambda = 1.5418$ Å), at room temperature, using a *Bruker AXS* Proteum Platinum[135] CCD detector coupled to an *Incoatec* IµS high brilliance microfocus source and a Kappa (K) goniometer. No diffraction pattern was detected.

Thus, the other crystal was freezed by plunging into liquid nitrogen and saved to be tested at Diamond Light Source synchrotron (Didcot, Oxfordshire, UK), beamline I24 (*microfocus MX*) using a Pilatus3 6M detector. Any diffraction patterns were obtained.

4. Results and discussion

4.1 TEM-1 purification

The first approach to study TEM-1 started with protein expression and purification in order to obtain a batch of pure protein to work with. Then, the main goal is to obtain protein crystals in complex with putative ligands to inhibit this β-lactamases. These were synthesized at *iMed.ULisboa* in Faculty of Pharmacy of University of Lisbon, in collaboration with Medicinal Chemistry Group, headed by Professor Rui Moreira.

Initially the entire protocol of protein production referred on the publication of *Inglis et al* [27] was followed. Yet, the protein after the purification with the Profinity™ IMAC Ni^{2+}-charged resin (*Bio-Rad*) was not so pure, when compared with what is shown by the gel in the supplementary material of that paper. Incubation time of 55 mL of lysate with resin was 1 hour at 4ºC, and buffer was 50 mM HEPES pH 8.0 plus 500 mM NaCl and imidazole pH 7.5. The purity was evaluated from results of 15% SDS-PAGE and a Western blot (Figure 4.1).

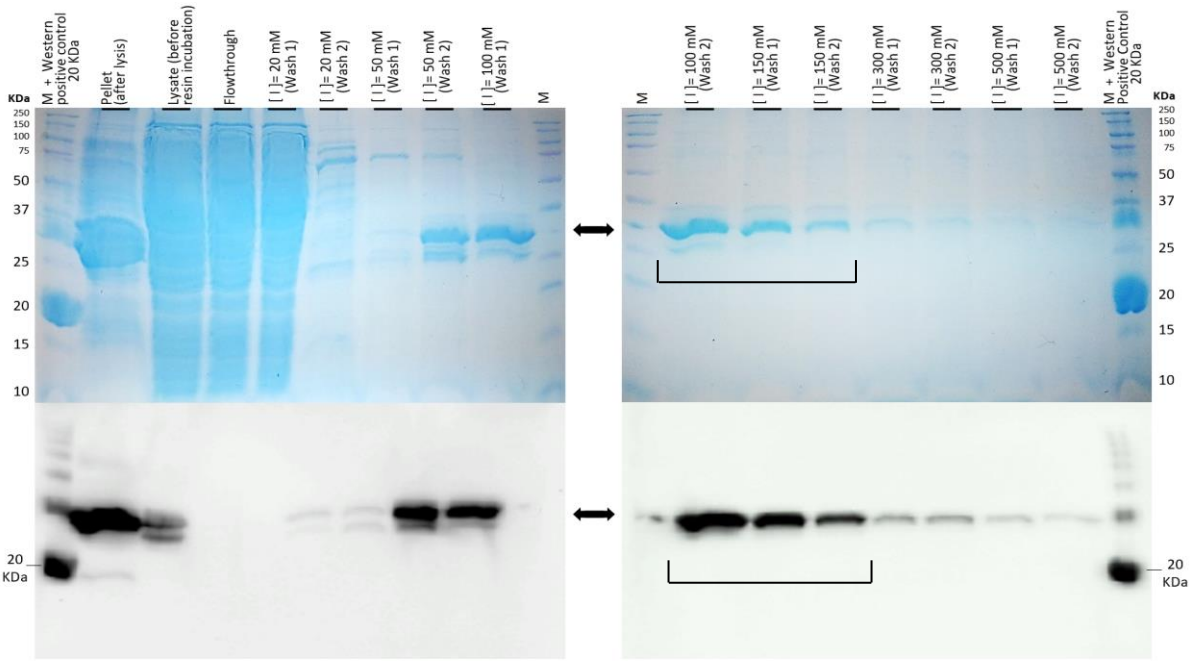


Figure 4.1 – 15% SDS-PAGE + Anti-Histag Western Blot results after 1st attempt of a batch/gravity-flow purification using Profinity™ IMAC Ni^{2+}-charged resin. [I] - Imidazole concentration. M - Precision Plus Protein WesternC™ Standards. Wash 1 is the first 15 mL elution to each [I] and the second is the following 15 mL. Arrows indicate to the molecular weight where is expected to be TEM-1 band.

Therefore, it was decided to do a second column with the following pooled fractions: 100 mM [I] Wash 2, 150 mM [I] Wash 1, and 150 mM [I] Wash 2. After finishing the size exclusion-chromatography (SEC) by using a Superose 12 12 10/300 GL with 500 µL of injected sample, a chromatogram was obtained (Figure 4.2). Purification buffer was 25 mM HEPES pH 7.5 plus 150 mM NaCl.

Results were more impure than previous purification, as shown by the SDS-PAGE in figure 4.3. The first flowthrough (F.T. 1) is made from pooled fractions after concentration (45 mL → 5 mL), and the second one (F.T. 2) is pooled fractions after concentration (5 mL → 500 µL).

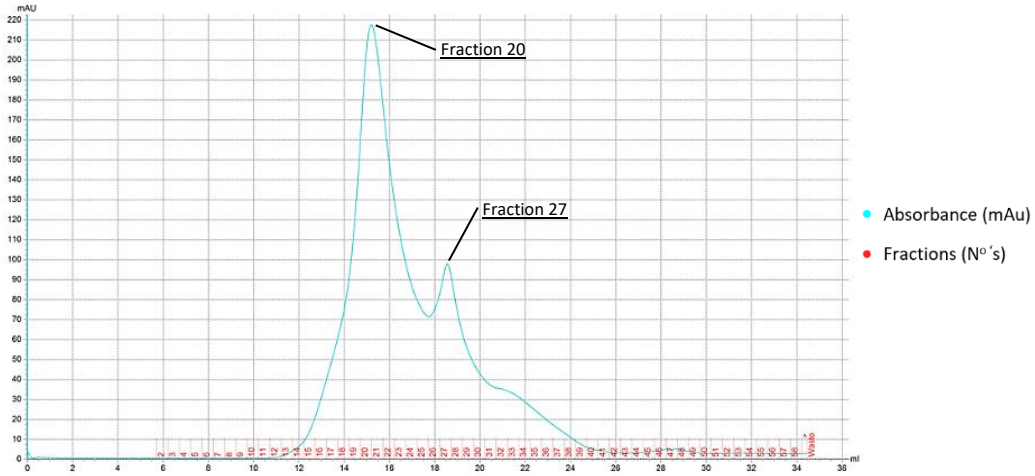


Figure 4.2 – Chromatogram from a SEC belonging to the 1st attempt of TEM-1 purification. Flow rate was 0.5 mL/min. Main peaks are highlighted with the identification of the fraction.

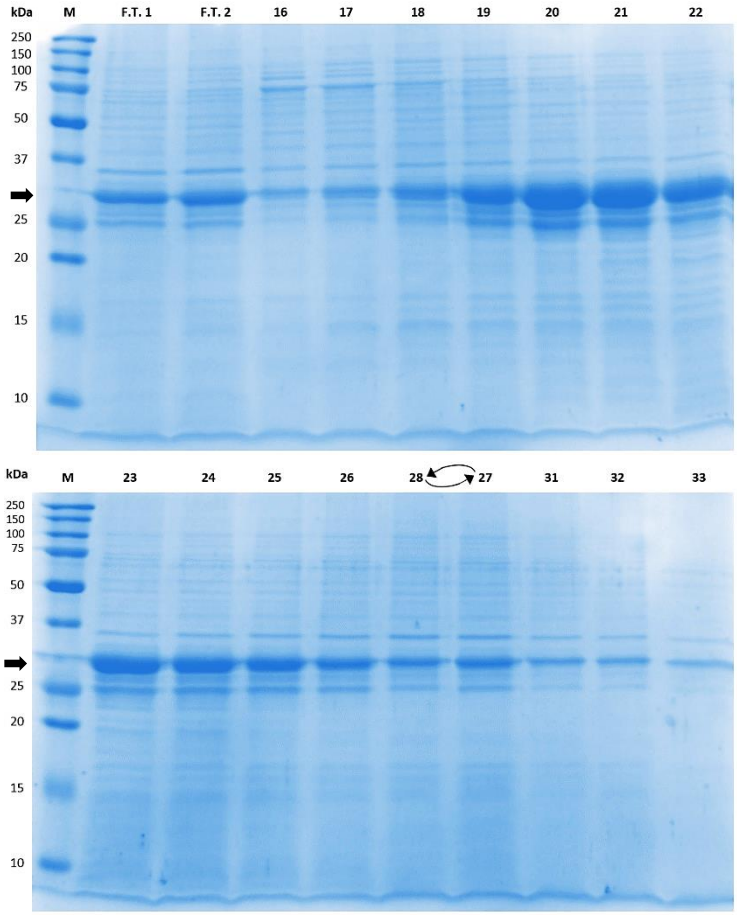


Figure 4.3 – 15% SDS-PAGE of the results from previous SEC, belonging to the 1st attempt of TEM-1 purification. M - Precision Plus Protein™ Dual Color Standards. Arrows indicate the molecular weight where is expected to be TEM-1 band.

After the first purification, the authors [26] kindly provided us more detailed information about the expression and purification procedures for TEM-1 protein. New information revealed alterations mainly in the method of purification by using a 5 mL histrap fastflow column instead of Ni^{2+}-charged resin and followed by a size-exclusion chromatography, the pH of HEPES in the lysis and purification

buffer was 7.5 instead of pH 8.0 as firstly referred, and additional details about the lysis protocol by sonication were also provided.

A second purification was made by knowing those new details. Results from the histrap purification with a 5 mL histrap fastflow column are present in figure 4.4. The buffers used were: buffer A with 20 mM HEPES pH 7.5, 500 mM NaCl plus 5 mM imidazole; buffer B with 20 mM HEPES pH 7.5, 500 mM NaCl plus 500 mM imidazole. In the first chromatogram (Figure 4.4 - C2), each fraction increased in 6.25 mM imidazole [I], and for figure 4.4 - C1 it increased 7.5 mM [I] for each fraction. Volumes of injected samples were 5 mL of sample plus 5 mL of buffer A for C1 and 10 mL of sample for C2.

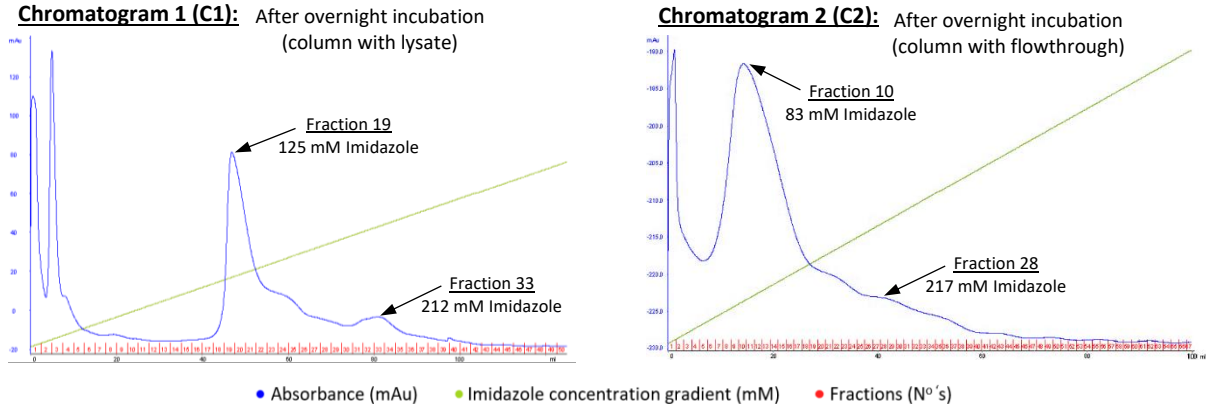


Figure 4.4 – Chromatograms of TEM-1 histrap purification. Performed with an imidazole linear gradient (0 mM → 500 mM). Main peaks are highlighted with arrows and their imidazole concentration.

Results after nickel-charged column were analyzed by 15% SDS-PAGE (Figure 4.5) to see in which fractions our protein was present, by elution with different imidazole concentrations [I].

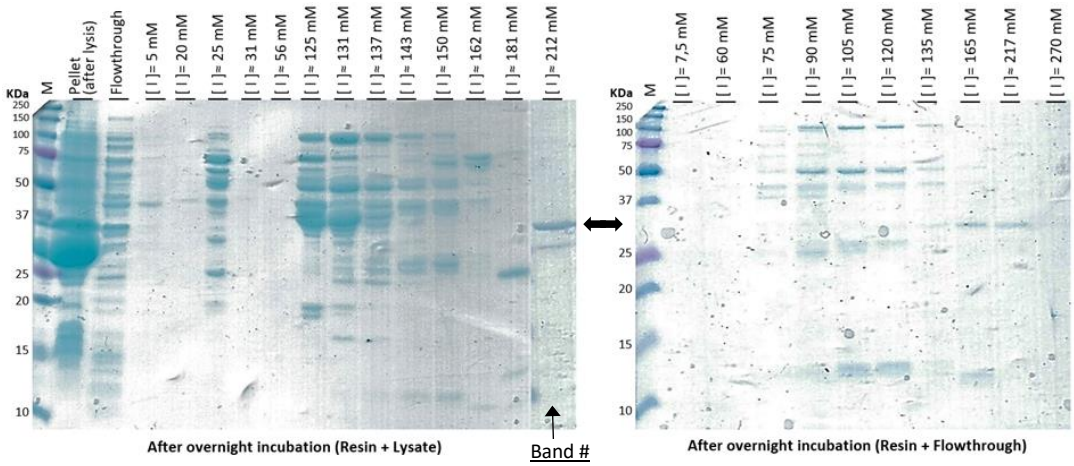


Figure 4.5 –15% SDS-PAGE after the 2nd attempt of a histrap purification. M - Precision Plus Protein™ Dual Color Standards. Band # - it belongs to the gel on the right, which means it ran in electrophoretic conditions of the gel after overnight incubation (resin + flowthrough). Arrows indicate to the molecular weight where is expected to be TEM-1 band.

This time the main issue was the lack of protein in the soluble fraction, as is seen in the gel by the slight bands where it was expected to be TEM-1, around 32 KDa. Some problems may have happened, such as: during the protein expression, the protein may not have been properly overexpressed and

went to inclusion bodies; during lysis with sonicator, e.g. due to a misuse or positioning of the sonicator tip inside the cell extract and consequently an incomplete lysis.

Next, with the valuable knowhow from the technicians from ITQB NOVA research facilities was performed a new production and purification. This third purification was done with the two previous methods: a batch/gravity-flow purification with Ni^{2+}-NTA resin and using the histrap column. Cell lysis was performed with French press and the histrap column was followed by an ionic exchange chromatography and then a size-exclusion chromatography. Results are detailed in Appendix B, from figure B.1 to figure B.6. Histrap purification proved to be more efficient than the resin, and after Q-Sepharose column the protein came out pure, as we can see in gels of figure B.4. Afterwards, pooled purer fractions were concentrated and injected in SEC column but TEM-1 became impure (showing degradation bands in gel of figure B.6).

Perhaps it happened self-proteolysis, indicating that the third purification step here may not be a good option. The main reason for we did not stop the purification after Q-Sepharose column was mostly because we wanted to be sure that TEM-1 did not have any oligomeric form that could difficult the following crystallization steps, and also because previously in the authors' e-mail with details about the purification protocol they mentioned a final polishing step with a size-exclusion chromatography using a standard superdex 200 column.

As future protocol adjustments, it will be added in the purification buffers PMSF at 0.2 mM plus glycerol at 5%, besides maintaining the protein fractions and buffers always at 4°C.

4.1.1 TEM-1 purification: 1st batch of protein

After several attempts and understanding which changes were an improvement, a new purification was achieved and allowed to stablish a TEM-1 purification protocol optimized that can be replicated to obtain a batch of pure protein to work with.

Starting with four liters of cell culture, it originated a fraction with the soluble proteins – where it should be TEM-1 β-lactamase – of about 65 mL. This sample was first injected in the histrap column, using the following buffers: buffer A (50 mM HEPES pH 7.5 + 500 mM NaCl + 5 mM Imidazole + 5% Glycerol + 0.2 mM PMSF) and buffer B (50 mM HEPES pH 7.5 + 500 mM NaCl + 500 mM Imidazole + 5% Glycerol + 0.2 mM PMSF). Purification was performed with the following gradients of imidazole ([I]): 1st with no [I] for 10 minutes; 2nd with 50 mM [I] for 10 minutes; 3rd from 50 mM to 250 mM [I] for 25 minutes; 4th a final gradient with 500 mM [I] for 15 minutes.

The main chromatogram is shown in figure 4.6 - C1, and a close-up of the region where TEM-1 was eluted is then shown in figure 4.6 – C2.

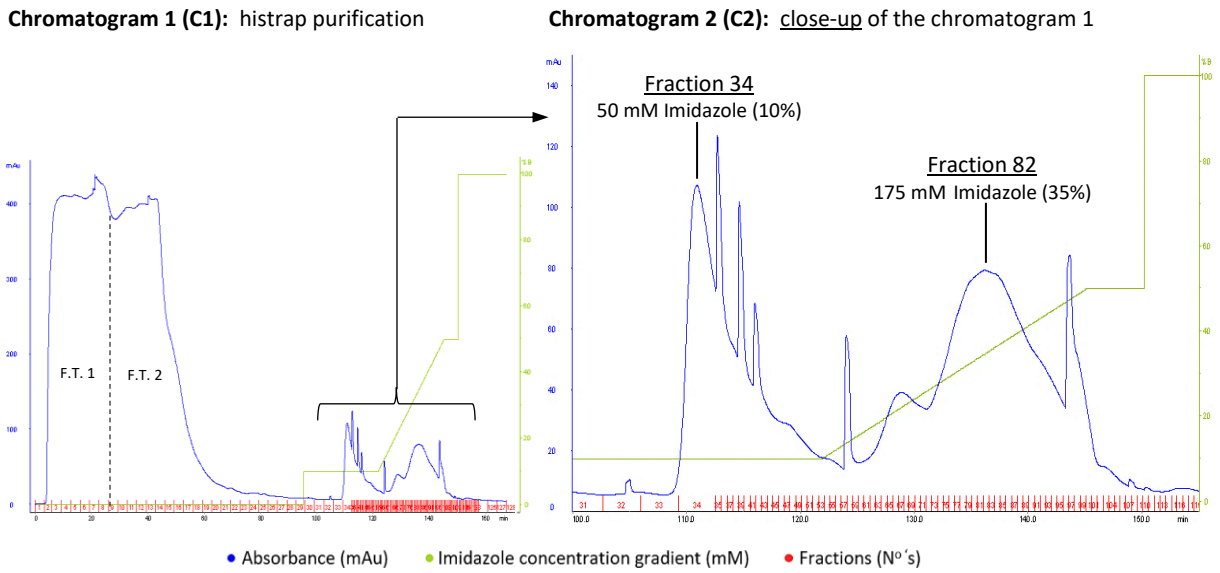


Figure 4.6 – Chromatograms for the best TEM-1 histrap purification. Flow rate was 3 mL/min. Injected sample (F.T. 1 – Flowthrough 1) was reinjected again in the column (F.T. 2 - Flowthrough 2) before starting the gradients. Main peaks are highlighted with corresponding imidazole concentration and fraction number.

A 15% SDS-PAGE (Figure 4.7) was done to evaluate the protein purity of the histrap fractions but the gels showed some problems related with the sample running and they were not able to show the bands in a proper way.

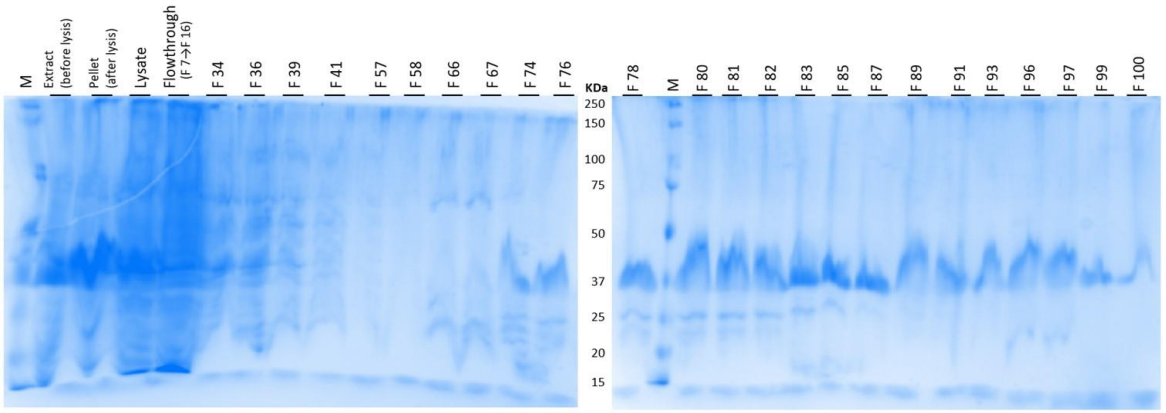


Figure 4.7 – 15% SDS-PAGE after the best purification attempt by doing a histrap purification, performed with a 5 mL histrap column. M - Precision Plus Protein™ Dual Color Standards.

Hence, based on the information gave by the highlighted peak of figure 4.6 at 175 mM imidazole, and by previous purifications that allowed to estimate where protein elutes, was made a pool of the enriched fractions to subsequently concentrate and inject in a second column. Pool (named pool +) that was made contained the fractions between number 73 and 90, concentrating it up to 5 mL with a Pierce™ protein concentrator 20K MWCO (*Thermo Scientific*) and then this pool was washed with 20 mM HEPES pH 7.5 using the concentrator itself to low as much as possible the imidazole concentration in solution so that the sample could be ready to be injected in the next column without imidazole to interfere. At this point the sample had a protein concentration of 2.4 mg/mL (in 5 mL), quantified by nanodrop measurements using the molar extinction coefficient in the oxidized form ($\varepsilon = 28\,085\ M^{-1} \cdot cm^{-1}$) at 280 nm.

Next purification step was an ionic exchange chromatography, with the following buffers: buffer A (20 mM HEPES pH 7.5 + 5% Glycerol + 0.2 mM PMSF) and buffer B (20 mM HEPES pH 7.5 + 1 M NaCl + 5% Glycerol + 0.2 mM PMSF). Previous pool + was injected with 5 mL of buffer A to make sure that no imidazole was present in the sample. Purification was done with linear gradients of sodium chloride (NaCl), varying between 0 and 1 molar. First, column was washed with no NaCl for 20 minutes and then was done a 1st gradient up to 300 mM NaCl for 1 hour. After that, a 2nd gradient between 300 mM and 1 M NaCl was performed. Figure 4.8 shows the resulting chromatogram with the 200 mM NaCl peak highlighted, where TEM-1 have eluted.

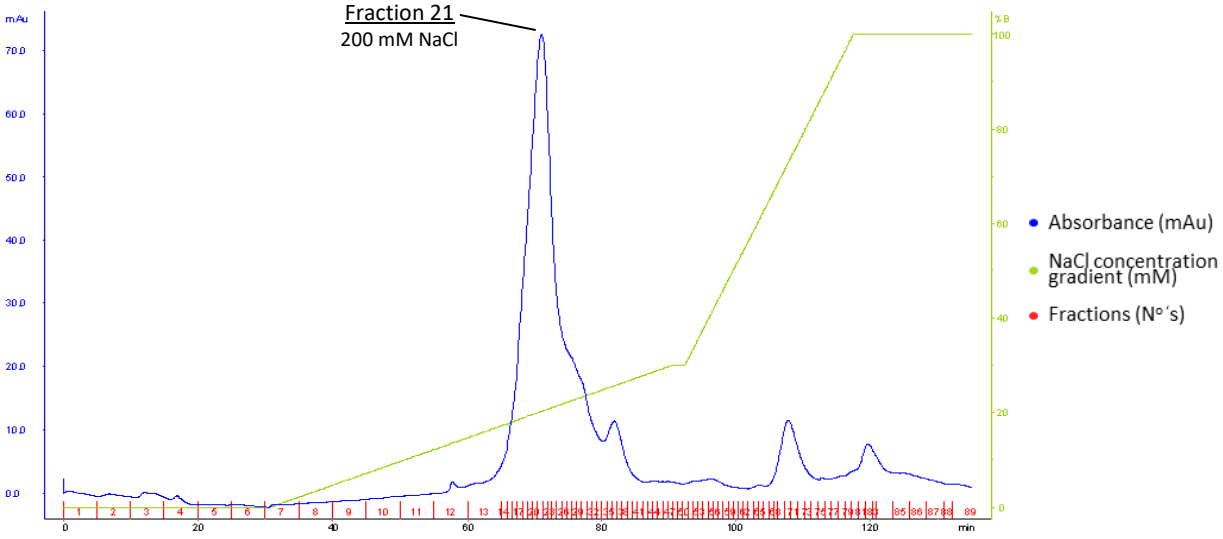


Figure 4.8 – Chromatogram for the best TEM-1 ionic exchange purification. Flow rate was 2 mL/min. Main peak is highlighted with the concentration value and fraction number.

15% SDS-PAGE was done to confirm the presence of the protein of interest and its purity along this purification, and the gels are represented in figure 4.9.

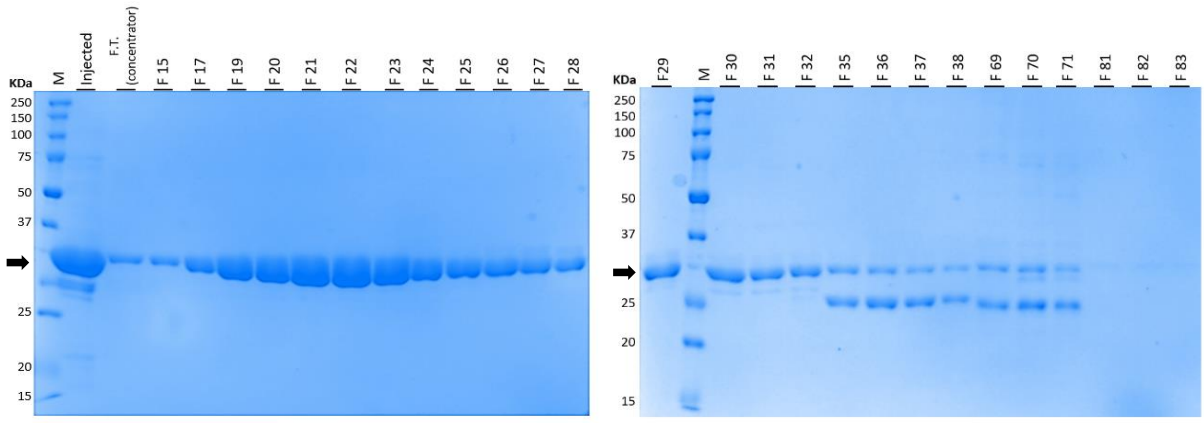


Figure 4.9 – 15% SDS-PAGE of the previous best ionic exchange chromatography. M - Precision Plus Protein™ Dual Color Standards. Injected (previously named pool +) had 9 µg protein, in 5 mL. Arrows indicate where is expected to be TEM-1 band.

Several fractions came out with very pure protein and a good quantity. A first pool (Pool 1) was made from fraction Nº 14 until Nº 24, since these were the purer fractions, and it was concentrated up to about 200 µL and after that was made a quantification by Bradford assay (sub-section 4.1.3), to ensure that was obtained enough protein concentration to the crystallization trials. Meanwhile a

second pool (Pool 2) from fraction Nº 25 to Nº 30 was made, with a considerable TEM-1 quantity but less pure. This one needed to be concentrated to about 100 µL to be injected in a gel filtration column.

As a third step of purification, a size-exclusion chromatography was performed (Figure 4.10) by injecting the pool 2 with a volume of 100 µL and an estimated concentration of about 20 mg/mL. As purification buffer, it was used 20 mM HEPES pH 7.5, 150 mM NaCl, 5% Glycerol and 0.2 mM PMSF.

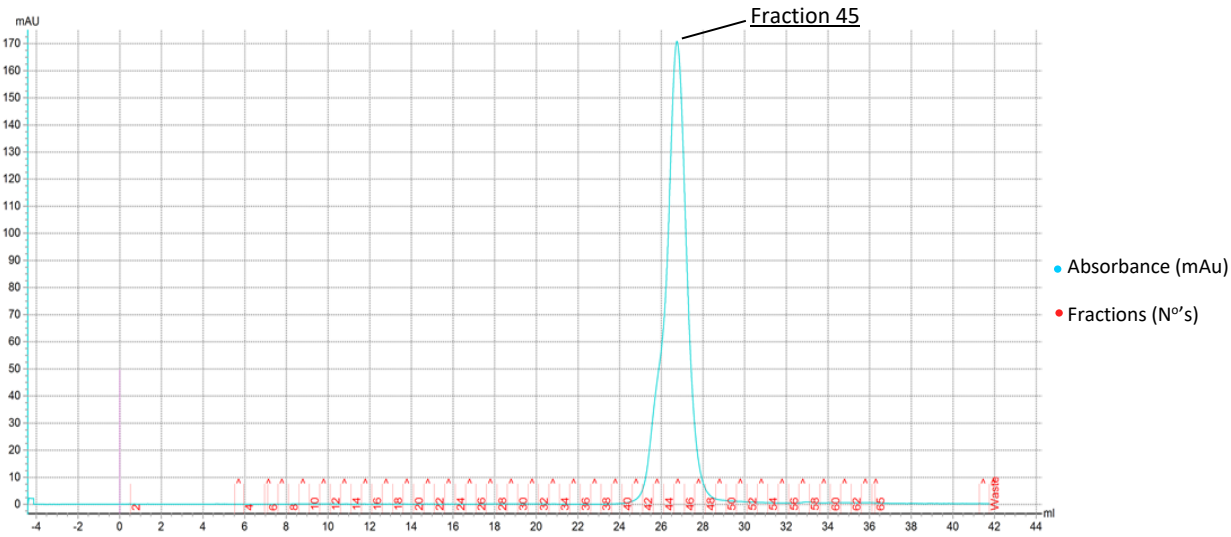


Figure 4.10 – Chromatogram for the best TEM-1 size-exclusion chromatography. Flow rate was 0.5 mL/min. The peak is highlighted with the identification of its fraction.

The fractions of this peak were loaded in a gel to see the success of the last purification step, and simultaneously both pools were injected in the same gel to assess if they were in good conditions (without sign of degradation) after concentration, especially pool 1 that should be ready to crystallization. So, the 15% SDS-PAGE can be visualized below in figure 4.11.

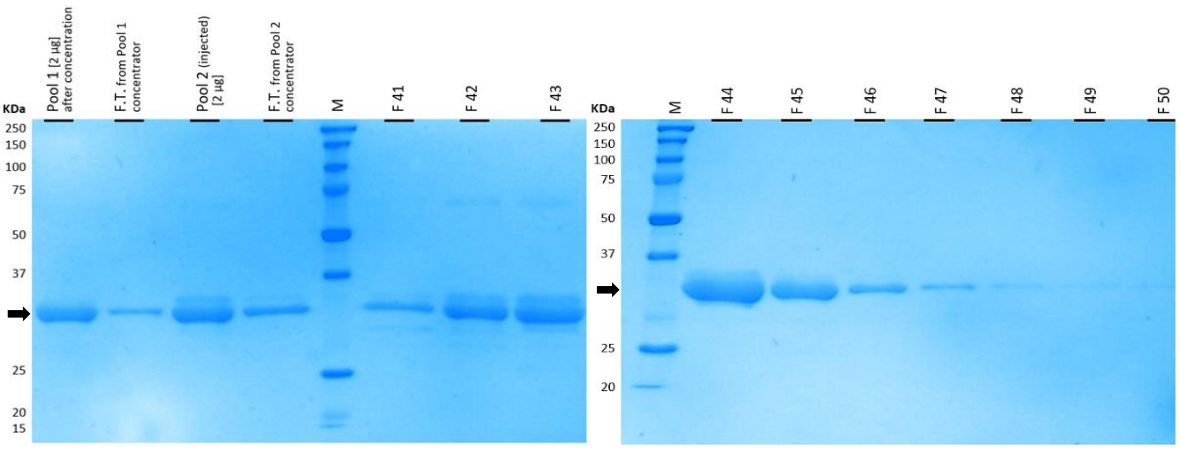


Figure 4.11 – 15% SDS-PAGE of previous size-exclusion chromatography, as well as one sample from the purer pool 1 after its concentration up to 200 µL. M - Precision Plus Protein™ Dual Color Standards. F.T. – Flowthrough after concentration. Arrows indicate where is expected to be TEM-1 band.

Then a final pool (Pool after SEC) was prepared by joining the fractions from F 44 to F 50. That was concentrated to about 40 µL, the required volume to have a concentration enough to eventually try to

do crystallization trials, and it presented a final concentration of about 13 mg/mL. This pool will remain on "*standby*" once pool 1 has more quantity of protein and purer one. In this last gel filtration chromatography, the first fractions of the peak (fractions 41 - 43) had clearly some bands of impurities at lower and higher molecular height compared to TEM-1 band at 32 KDa. From fraction F 44 forward, the protein seems to be pure (only F 44 have a slight and tiny band below to the main one). The analysis of the gel in figure 4.11 also allows to corroborate the purity of pool 1.

About the purification yield, this best attempt allowed to obtain a purer pool 1 with about 6.6 mg of TEM-1, extremely lower than expected since the authors referred to us that they got more than 100 mg of protein in 34 grams of cells. One possible explanation is that TEM-1, when present in high concentrations inside *E. coli* cells, can be toxic and the cell machinery will direct it to inclusion bodies. So, protein will accumulate there and will be available in less quantity within cytosol (the typical location of soluble proteins). Sometimes can also happen some problems related with improper protein folding – when the protein is expressed in large quantities – which will produce aggregated protein, being present in those inclusion bodies.

Yet, the authors never mentioned any troubles on this. What can also happen is that the protein is overexpressed and part of it can be trapped inside the periplasmic space – once the expression system is a gram-negative bacteria – increasing the difficulty to obtain a good quantity of soluble protein during cell lysis.

4.1.2 TEM-1 purification: 2st batch of protein

A second batch of TEM-1 was obtained (also designated as fifth purification) to have more protein stock to use in crystallization trials if necessary, always giving priority to the use of the 1st batch and this will be mostly to prevent from running out of protein and stop any crystallization experiments.

This new batch was produced using the same three purification steps of the first one, but in this attempt the difference was that both pools after the anionic exchange chromatography were injected on the SEC column, producing two chromatograms. All these purification results are present on the Appendix B, 5th attempt (from figure B.7 to figure B.12). At the end of those three columns were obtained three different pools, according to their purity after the SEC, named: pool 1.0 (C1:F21 + F22); pool 2.0 (C2:F21 + F22, and the purer one); and pool 2.1 (C2:F23 + F24). The first two pools are purer than the third. Nanodrop quantifications were made for these three pools, by order: 10.5 mg/mL (in 25 µL), 12 mg/mL (in 55 µL), and 15 mg/mL (in 25 µL).

4.1.3 TEM-1 quantification: Bradford assay

As previously mentioned, the Bradford method was used to do an accurate protein quantification of the first batch of pure TEM-1 β-Lactamase that was achieved, already mentioned as pool 1.

First, several standards were prepared with seven different BSA final concentrations. The absorbance measurements were done in triplicates for each concentration, and then the mean value was calculated using those values. The experimental data obtained is summarized in table 4.1.

Table 4.1 – Results of the absorbance measurements, at 595 nm, of the standards with several BSA final concentration values (µg/mL) used to calculate a calibration curve.

BSA calibration curve

[BSA] µg/mL	Absorbance (AU) - Mean value
1000	0,482
800	0,502
600	0,457
400	0,433
200	0,385
100	0,361
0	0,321

The graphic representation of the BSA calibration curve with its linear regression trendline, calculated by interpolation of the experimental values, is showed below in figure 4.12.

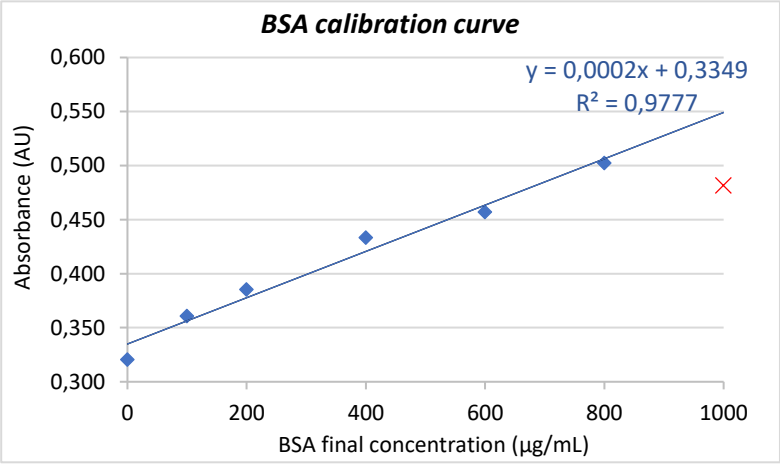


Figure 4.12 – BSA calibration curve to quantify TEM-1 from pool 1. Blue points represent the experimental mean values obtained to trace the calibration curve. Red cross represents point that was not considered to do the calculations (outlier). Blue line is the trendline of the linear regression, with the corresponding equation represented at the top right.

After evaluation of the BSA calibration curve, there is one outlier (red cross) for the point with BSA concentration at 1000 µg/mL, and so it was excluded from the calculations to obtain the equation of the linear regression and the coefficient of determination (R^2). The remaining points were used to build the calibration curve with the equation $y = 0{,}0002x + 0{,}3349$, and the $R^2 = 0{,}9777$. Considering the existence of a good relation of linearity, based on the presented R^2 value that is close to 1, it is possible to move forward and apply the equation to estimate the protein concentration of our desired sample. The dilution of the protein sample was 1:100 since that was the one which had the absorbance mean value within the calibration curve ($Abs = 0{,}402$), with the others protein concentrations having values out of the range of the calibration curve. After some steps of calculations by using the determined equation, the dilution factor and corresponding conversions of protein concentration units, we obtained a final value to the protein concentration: [TEM-1, pool 1 from 1st batch] ≈ $33\ mg/mL$, in 200 µL. Thereby, it is possible to guarantee the exact protein concentration to step forward in the experimental work and perform the crystallization trials.

4.2 TEM-1 crystallization and data collection

The aim was first to obtain crystals from native protein (without any presence of ligands), to select the best conditions of crystallization and then move on to the soaking or/and co-crystallization experiments to obtain protein complexes with the synthesized ligand.

Since the Bradford method allowed an accurate protein quantification of pool 1 (the purer from the 1st batch), it was used to start the crystallization. The first TEM-1 native crystal was obtained with BCS Screen in condition 1-32 and can be seen in figure 4.13. The crystallization plate was visualized and the pictures taken by using a stereomicroscope (model MZ16 A, *Leica*), as well as from now on in all drops. The crystal appeared at a protein concentration of 16.5 mg/mL, growing to about 50 µm.

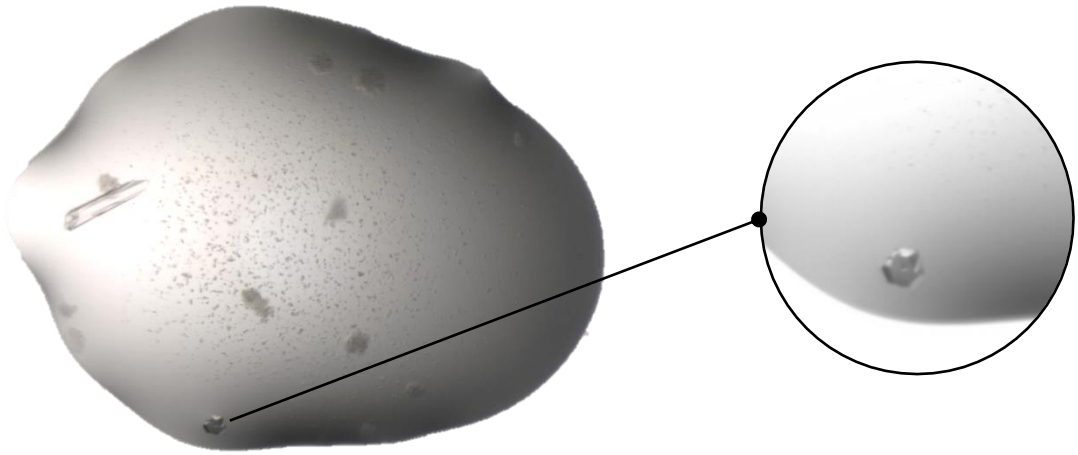


Figure 4.13 – TEM-1 native crystal with a hexagonal shape, obtained after 10 days in the condition 1-32 of BCS Screen.

A second native crystal has appeared in the well E6 of the 48-well plate made with the crystallization robots to scale-up and optimize that BCS Screen best condition, and the optimization matrix corresponding to that plate is in figure C.2 of Appendix C. The crystal appeared at a protein concentration of 16.5 mg/mL, growing to about 60 µm of length. The composition of this drop is almost the same of the previous condition 1-32 that had a crystal, with a little difference in the concentration of PEG smear broad (in this case was 26.8% instead of 28%). In the next figure 4.14 is possible to see the drop of this well.

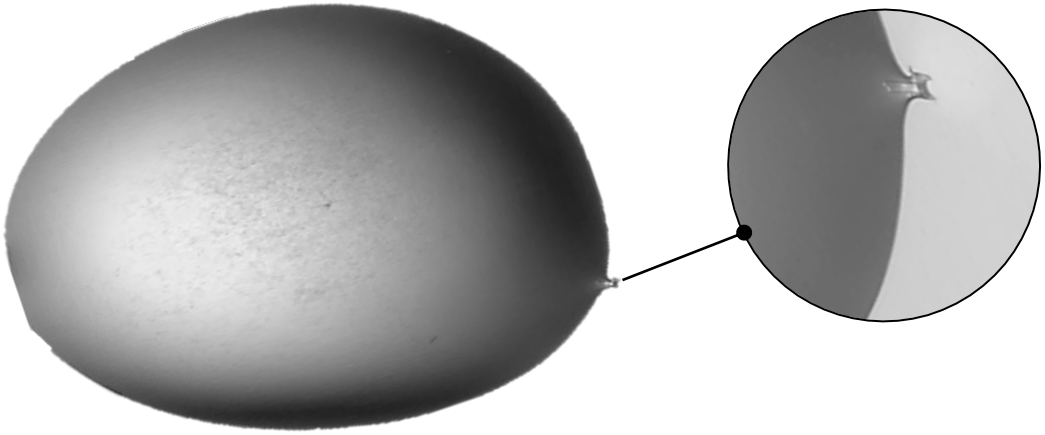


Figure 4.14 – TEM-1 native crystal obtained after 7 days in the well E6 of the optimization matrix of BCS Screen.

Related to this optimization in a 48-well plate, it is important to highlight that was made with commercial stock solutions and only the stock of PEG smear broad was the same from BCS Screen best

condition. The glycerol and arginine solution were from different suppliers, and the most significant difference is in the glutamic acid solution that was not exactly the same in terms of composition. In the laboratory, the only solution of this amino acid available was the DL-Glutamic acid monohydrate, that have a small difference from that one present in the original solution provided by *Molecular Dimensions*. These two amino acid solutions have different stereoisomers compositions, and this fact could be a possible reason for the lack of crystals in the plate as we intended to. We only obtained one drop with a possible tiny protein crystal as figure 4.14 shows, but we did not confirm if it was salt or protein.

Besides, other optimization of BCS Screen best condition was made manually in a 24-well Linbro plate and bigger crystals appeared in well A2 (Table 3.3), as shown in figure 4.15. Crystals were formed after 7 days, with a protein concentration at 8 mg/mL and having about 120 µm (0.12 mm) of length.

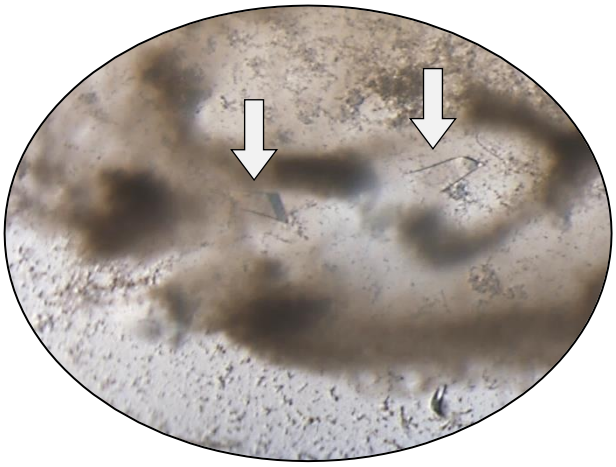


Figure 4.15 – TEM-1 native crystal obtained after optimization of the initial hit from a BCS screen condition. White arrows point to the two crystals that were tested, one *in-house* and the other in synchrotron.

About the result of the crystal tested in the diffractometer at ITQB NOVA (Oeiras, Portugal), it was not possible to collect any diffraction data. The most likely reason must have been due to the size of the crystal, that was not big enough to have diffraction with an X-ray source like the one that was used.

Besides, the other crystal from the same drop that was tested at Diamond Light Source synchrotron and did not produce good diffraction patterns. In figure 4.16 is present an image of the beamline with the loop mounted with the TEM-1 crystal centered, as well as one example of a diffraction pattern that demonstrates the absence of spots.

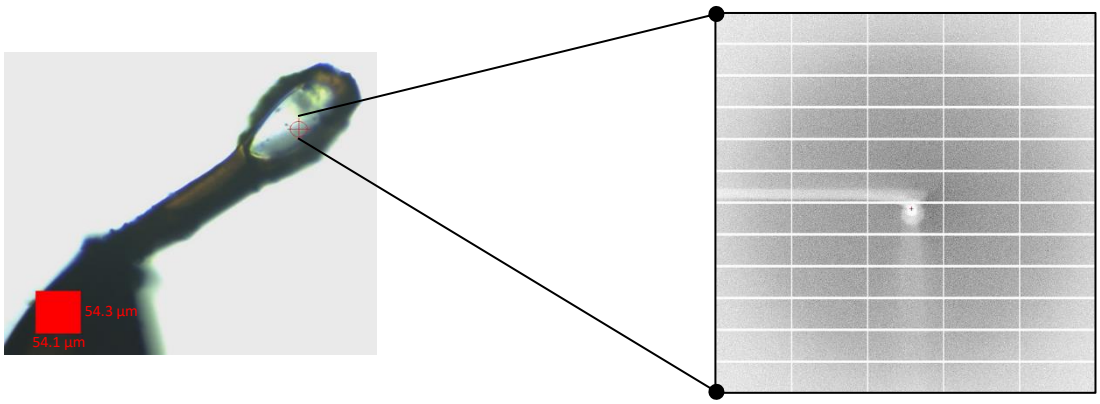


Figure 4.16 – Data collection for the TEM-1 native crystal in beamline I24 at Diamond Light Source. The loop with the crystal ready to diffraction screening is on the left, and on the right is a diffraction pattern obtained with no spots visible.

As we can observe, crystals of native TEM-1 were able to be obtained and its diffraction tested but without a diffraction pattern typical of protein crystals. The crystal does not have an inner structure sufficiently well-ordered to produce spots, so the crystallization process needs to be optimized. Also, is possible to assess that it was not a salt crystal since the diffraction pattern is not like a typical pattern of that [29], where usually are spots with strong intensity and close to others [28], neither are ice crystals since ice rings are not seen in the diffraction pattern.

Once a 2nd batch of protein had been produced, where their three different pools were tested and just one (Pool 2.0) had produced one hit in condition 1-30 of PACT premier™ screen (conditions are detailed in figure C.3 of Appendix C), with protein at 6 mg/mL. This tiny crystal in figure 4.17 was visible after 11 days and grew up to a length of about 40 µm.

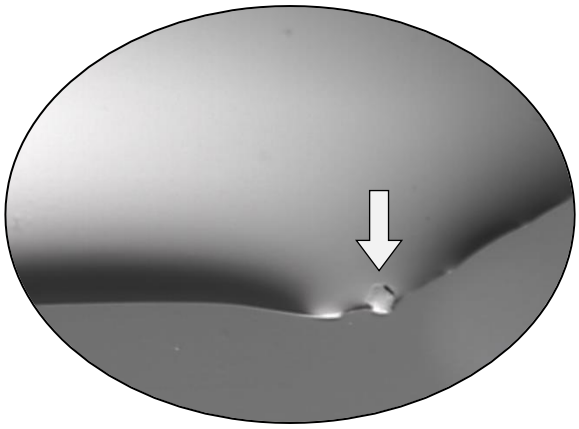


Figure 4.17 – TEM-1 native crystal obtained in condition 1-30 of PACT premier™ screen, with protein from pool 2.0. The white arrow points to the crystal formed in the edge of the drop.

We did not test this small crystal in synchrotron, since we chose to focus in the optimization of the other hits related to the 1st batch of protein, so is not possible to affirm if was a salt or protein crystal.

As an overview of the β-Lactamase optimization plates, many drops resulted in a lot of precipitates (light and heavy ones), indicating that predominantly the equilibrium of the drops had complications and neither the nucleation process nor growth were easily reached. Some examples happened in several drops of the three screen plates (Table 3.2), also in the 48-well optimization plate, and even in the seeding related plates. Without success in the optimization process and in obtaining drops with native crystals (which we could be sure they would be protein crystals) was not possible to go further in the studies with complexes of TEM-1 by soaking the protein with ligands, testing if they would be good candidates to be β-Lactamase inhibitors and how they could interact with the enzyme.

Alternatively, we have decided to revisit the previous work developed in our laboratory with serine proteases [56] – particularly with the Elastase – and begin new studies to try to confirm if the inhibition mechanism in PPE with ligands designed for HNE will be similar to what was earlier discovered in PPE complexes with novel inhibitors.

5. Introduction

5.1 Enzymes as important therapeutic targets

Enzymes are biomolecules essential for a wide variety of biological processes (e.g., in metabolism, cell regulation, digestion, tissue healing, or blood clotting), and their correct functioning is mandatory for a healthy organism. Almost all the enzymes are proteins; the only exception known to date is a group of catalytic RNA molecules, that act by decreasing the activation energy of a certain chemical reaction which results in faster chemical reactions with no enzyme consumption by the reaction. Almost every biochemical reaction is catalyzed by an enzyme. [28, 30]

Enzymes can be produced and recruited endogenously, the so-called endoenzymes, or can be produced and secreted to the extracellular space, the so-called exoenzymes. The high substrate specificity is an important feature of enzymes where each one only recognizes and reacts with a specific chemical group or a set of certain protein residues. Their efficiency in terms of catalytic activity is dependent of two major factors, pH and temperature, that can be optimized. [28]

Firstly, many enzymes were named by adding the suffix "-ase" to their substrate name or to a keyword or phrase that describes their activity. Still, the name of others was created based on their biological source. The enzymology was getting more complex with the time there were cases where the same enzyme had more than one name or two different enzymes with the same name, so it was necessary to create a classification system able to provide a way of differentiating these ambiguities. A numerical classification that divides all the enzymes into six main classes, and different sub-classes according to the type of catalytic reaction, created by the Enzyme Commission (EC). [28] The six classes of the EC number are summarized in figure 5.1, where is present the general mechanism of each reaction.

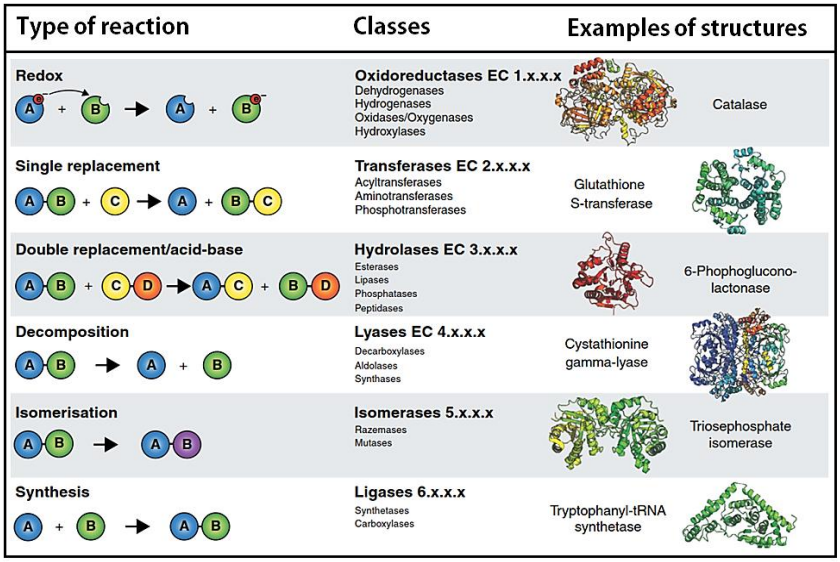


Figure 5.1 – Summary of the six main classes of enzymes, organized by their types, the general reaction, the classes with the respective EC number, and an example of one enzyme for each of them. Adapted from 36

Other distinctive feature of a reaction catalyzed by an enzyme is that it usually happens within a pocket of the enzyme called the active site; the molecule or compound that binds to the active site and reacts upon the enzyme is called substrate. In most of the cases, the active site encloses the

substrate, making a complete sequestration from the solvent. [28] In many organisms, enzymes can suffer mechanisms of negative feedback due to the presence of inhibitors. Those molecules can be produced by the own organism and modulate the enzyme function, but in some other cases the inhibitors can be drugs that are administered and act in a reversible or irreversible way. [31]

These biomolecules are often chosen as the main target on structural biology research due to their biological importance towards health problems treatment, and that will ultimately contribute to the development of new drugs. The structural biology uses numerous methods as X-ray crystallography, small-angle X-ray scattering (SAXS), nuclear magnetic resonance (NMR) or, more recently, cryo-electron microscopy (Cryo-EM). Besides that, biochemical and biophysical techniques are used to improve the knowledge about the enzymes, their function, activity, specificity or affinity. Techniques such as thermal shift assay (TSA or Thermofluor), surface plasmon resonance (SPR), dynamic light scattering (DLS), isothermal titration calorimetry (ITC) or circular dichroism (CD). [18]

5.2 Proteases

Proteases (EC 3.4.X.X) are proteolytic enzymes that, as the name suggests, break the peptide bond that connects the amino acids together in proteins. They are also known as hydrolases because their function requires the addition of a water molecule. Although the hydrolysis of a peptide bond is energetically downhill, this reaction is very slow at physiological temperature. [28, 32]

Proteases are present in several organisms as bacteria, viruses and man. Highlighting the third case, proteases are essentially useful for food digestion, cleavage of signal peptides, or in cell metabolism regulation. *In vivo*, the activity of many proteases is controlled by endogenous protein inhibitors. [32]

Proteases can be divided into exopeptidases and endopeptidases according to their preferred site of action. The former can act at the end of the proteins by degradation in the N-terminal region or in the C-terminal region of the protein. [30, 33]

It is very variable the number of amino acids that they can remove, being a single residue, or a dipeptide or tripeptide according to the nature of their catalytic center. However, there are some exopeptidases that are not so specific to one of the ends of the protein. The endopeptidases only are differentiated according to the main residue that is present in their active site. [30, 33]

Due to some cases of proteases with mixed and unspecific activity and based on the nature of their amino acid sequence homology and the triad of residues in the active site, a new classification emerged in 1993 named MEROPS [34] and it has all the detailed information for proteases and their inhibitors is grouped in an online database. It is constituted by four functional families: serine proteases, cysteine proteases, aspartic proteases, and metallo proteases. [30, 33, 34, 35]

Figure 5.2 shows the grouping of the proteases identified until 2017, according to the PDB entries.

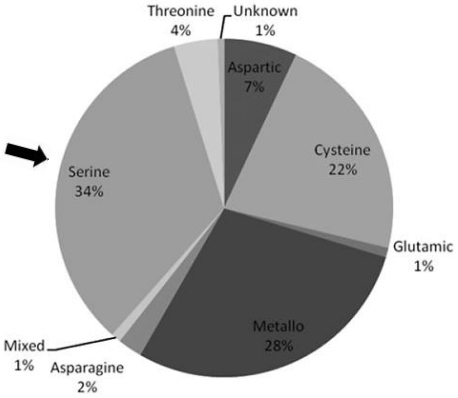


Figure 5.2 – Overview of protease families identified and with PDB entries, until 2017. Arrow highlights the highest number that belong to serine proteases. It comprises not only the types from MEROPS classification but also others. [30]

This representation shows the prevalence of the serine proteases family (34%) comparing to the rest, which is corroborated with the available data of the MEROPS classification. [34] It emphasizes that this group of proteases is currently of high importance in biological structural studies.

5.3 Serine proteases

Serine proteases (EC 3.4.21.X) are abundantly involved in our biological processes and besides the functions already described for proteases in general, this serine-based enzymes are important too in the control of the blood pressure and clotting processes. All these proteases are hydrolases, so they perform the hydrolysis of a certain substrate. [32]

They use a common catalytic triad to hydrolyze the peptide bonds, composed by a serine, a histidine and an aspartic acid (Figure 5.3). The last one is the most reactive and participates directly in the acyl-enzyme intermediate formation. [30, 32]

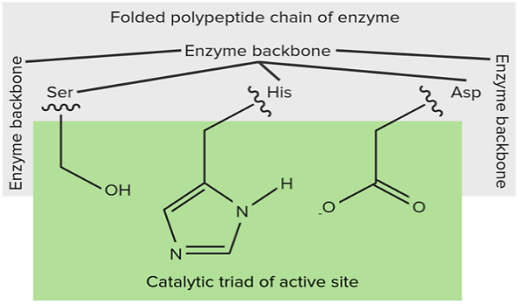


Figure 5.3 – General catalytic triad of serine proteases, formed by serine, histidine and aspartic acid. [38]

Enzymatic catalytic mechanisms are divided in two main stages - acylation and deacylation – and it is started by the approach between the serine of the active site and the carbonyl group present in the acyl group of a peptide bond. This mechanism is reversible and can be seen in more detail in figure 5.4, with the identification of the mechanism's steps, as well as the interactions of the three important residues.

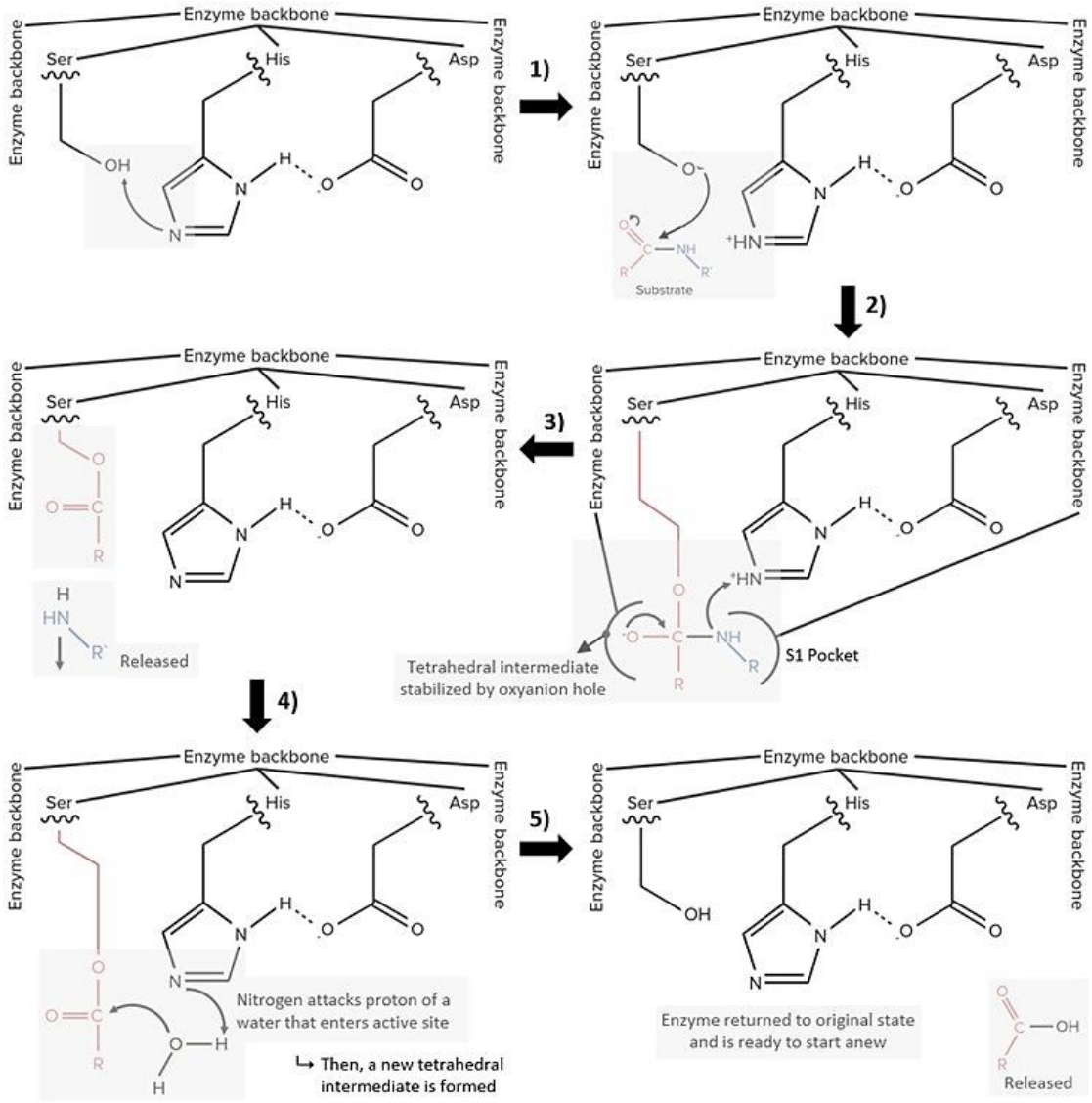


Figure 5.4 – General catalytic mechanism of serine proteases, with the five main steps identified. Adapted from 38

The acylation is triggered with the nucleophilic attack of the catalytic oxygen of serine on the carbonyl group of the substrate (**2**), and the negative charge generated on the oxygen is stabilized by the oxyanion hole (which will help next in stabilizing the tetrahedral intermediate). So, a new covalent bond is formed, causing the breakage of a peptide bond (**3**). In the end of this stage an acyl-enzyme intermediate is formed. Afterwards, the second part of the mechanism called deacylation starts with a new nucleophilic attack by one water molecule on a carbon that is bound to the serine (**4**). Lastly, the catalytic triad is regenerated, and the hydrolysis of the substrate is finished (**5**). [32, 37, 38]

5.4 Elastases

Elastases are globular proteins, with a single polypeptide chain and 4 disulfide bridges. They are also named as serine endoproteases, with a hydrophobic core due to the presence of a distinctive structure formed by two β-barrel domains that converge at the active site, and a wide sequence homology with others of the serine proteases family such as the chymotrypsin. They act by doing the cleavage of amides and esters in collagen and elastin, besides many other proteins. These two proteins are related with the elasticity of the connective tissues of our body, and elastin is mainly present in the lungs, arteries, skin, ligaments and some other types of soft tissue. [39]

In humans, the most prevalent serine proteases are the pancreatic elastases and neutrophilic elastases. Pancreatic elastases are stored inside pancreas in inactive form and it only changes by the action of trypsin during the digestion in small intestine. Regarding neutrophilic elastases, the main purpose is the defense against pathogens and facilitate the act of phagocytosis. In normal physiological conditions, both types of elastases are properly regulated by specific inhibitors present in plasma. [39-41] However, in case of deregulation, the release of these proteolytic enzymes from neutrophils will break the peptide bonds of connective tissue. That causes pathologies in liver, lungs or skin, and leads to severe inflammatory diseases as psoriasis or rheumatoid arthritis, but also cystic fibrosis and chronic obstructive pulmonary disease (COPD) where neutrophils will migrate to the lungs. The latter is currently one of the major cause of death related with the respiratory tract and so HNE (Human Neutrophil Elastase, EC 3.4.21.37) is a therapeutic target of enormous importance. [39, 40]

PPE (Porcine Pancreatic Elastase, EC 3.4.21.36) is a serine protease and a well-studied and validated structural model for HNE, both having the same catalytic triad composed of His^{57}, Asp^{102} and Ser^{195}. [39, 40] Also, they share 38% of sequence identity, based on a Protein BLAST between the two protein sequences of PPE and HNE (UniProt codes P00772 and P08246, respectively). However, a first difference lies in the fact that PPE has a less hydrophobic surface than HNE because the first elastase has just 30% of hydrophobic residues when compared to more than 40% in the case of the human elastase. Other variation is that PPE has its catalytic site more accessible to the approaching of non-linear ligands due to lower number of basic amino acids (especially arginines) in there. So, this second difference is clearly an advantage to use PPE in this attempt to obtain protein-ligands complexes. [40]

Additionally, it has been reported the need of discover low-molecular-weight inhibitors for elastases, and specifically the molecules based on the β-lactam ring appear to be good candidates since they have been reported as potent inhibitors of serine proteases. [39] In this work, the ligands that will be tested as candidates to elastase inhibitors were indeed designed based on the structure of other β-lactam molecules.

6. Materials and methods

Solutions and buffers used in this section for all the crystallization experiments were prepared using a precision balance and an analytic balance (model 440 and ADB, both from KERN), and a pH meter (pH 1100L, *VWR*). In this chapter all the protein quantifications were performed in a NanoDrop™ One spectrophotometer (*Thermo Scientific*), using 2-µL sample drops and using milli-Q water as blank. A stereomicroscope (model MZ16 A, *Leica*) was also used here to visualize all the crystallization plates.

6.1 Crystallization of native PPE

Before moving to the crystallization tests of the protein in complex with the ligands for upcoming structural studies of inhibition, it is important to know the optimized conditions for PPE native crystals. Since a lot of previous work of optimization was done in our laboratory – some results have already been published [39, 42] and the most recent ones are *in preparation* to be published [56] – in order to achieve the best crystallization condition of native PPE and good diffracting crystals, we started by replicating that best condition and see if the crystal growth would be as expected.

The porcine pancreatic elastase (PPE) used along this work is, as the name indicates, from *Sus scrofa* (pig) and it is commercially available from SERVA Electrophoresis GmbH (Heidelberg, Germany). It is a globular protein constituted by a single chain of 240 amino acids and a molecular mass of 25.9 KDa, is chromatographically purified, and sold in a lyophilized and salt-free form. The batch of protein was reconstituted in water to a concentration of $40\ mg/mL$, with no additional steps of purification (performed as mentioned in the supporting information of [39, 42]) and ready to perform the crystallization. Information about the optimized crystallization condition is summarized in table 6.1.

Table 6.1 – Crystallization details for the optimized condition that allow to obtain the best PPE native crystals.

Protein buffer	milli-Q water
Protein concentration	20 mg/mL
Optimized crystallization condition	70% (V/V) 2-Methyl-2,4-pentanediol (MPD) and 10 mM Sodium phosphate buffer pH 5.9
Temperature	20°C
Technique	Sitting – drop
V_{Drop}	(1: 1) µL
$V_{Reser.}$	500 µL

Several drops were prepared in a 24-well Linbro plate following the information of table 6.1. Crystals appeared after 5 days. The same conditions were also tested with a protein concentration of $15\ mg/mL$ with the same positive results.

6.2 Co-crystallization of PPE with inhibitor candidates

Previous work in our laboratory [56] has demonstrated that the co-crystallization technique as the best method for study the binding between this protein and the ligands, so we have decided just to focus on that technique and trying soakings only if necessary.

All the tested ligands were synthesized again by the collaboration group in Instituto de Investigação do Medicamento (*iMed.ULisboa*) of Professor Rui Moreira. Ligands LMC223 and LMC249 were designed for tests with HNE. Below in figure 6.1 are presented the structures of those two ligands designed for HNE and table 6.2 has the main chemical information of each one.

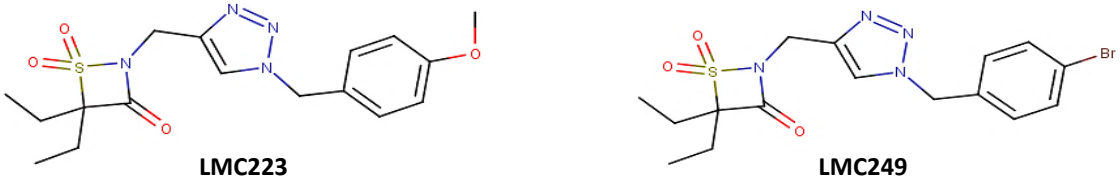


Figure 6.1 – Structures of the ligands originally designed for human neutrophil elastase (HNE), the LMC223 and LMC249, synthesized at *iMed.ULisboa*, and built with MarvinSketch® (from *ChemAxon*) open-source software.

Table 6.2 – Detailed information about HNE ligands: the exact molecular weight and formula of each ligand. Both were calculated using the MarvinSketch® (from *ChemAxon*) open-source software.

Ligand	**LMC223**	**LMC249**
Exact molecular weight	378.136 g/mol	426.036 g/mol
Formula	$C_{17}H_{22}N_4O_4S$	$C_{16}H_{19}BrN_4O_3S$

Other ligands (Figure 6.2 and table 6.3) were designed specifically for PPE and already tested. [56]

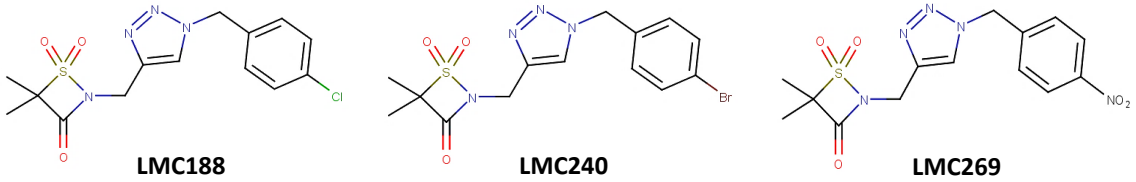


Figure 6.2 – Structures of the three ligands designed and already tested for porcine pancreatic elastase (PPE), named LMC188, LMC240 and LMC269, synthesized at *iMed.ULisboa*, and built with MarvinSketch® (from *ChemAxon*).

Table 6.3 – Detailed information about HNE ligands: the exact molecular weight and formula of each ligand. Both were calculated using the MarvinSketch® (from *ChemAxon*).

Ligand	**LMC188**	**LMC240**	**LMC269**
Exact molecular weight	354.055 g/mol	398.005 g/mol	365.079 g/mol
Formula	$C_{14}H_{15}ClN_4O_3S$	$C_{14}H_{15}BrN_4O_3S$	$C_{14}H_{15}N_5O_5S$

When is mentioned that certain ligand or set of ligands are for PPE or HNE is due to previous pharmacological studies – also done by at *iMed.ULisboa* – of inhibition kinetics by measuring the enzymatic activity of each protein incubated with the different ligands and determining the IC_{50} (the inhibitor concentration that corresponds to 50% of inhibition of the enzymatic reaction). The ligands with lower value of IC_{50} are more powerful inhibitors than those with high IC_{50}. The details of these assays are available on supporting information. [42, 43]

Mainly, we intend to obtain crystals of PPE in complex with the ligands that prior studies have shown more tendency to bind to HNE. It will be evaluated if there is any interaction of these HNE ligands with PPE in a way that could eventually corroborate the mechanism of inhibition in PPE by the nucleophilic attack on the sulphonyl group of the three ligands in figure 6.2, with this work and results from X-ray crystallography performed in our laboratory but not yet published. Otherwise, it can allow to know new information about variations in the mechanism of ligand binding with this serine proteases. This approach is rational because both enzymes are from the same protease, sharing the same catalytic triad and having a good sequence identity as it is mentioned before.

6.2.1 Co-crystallization: ligands originally designed to HNE and now tested with PPE

Thus, for co-crystallization (technique explained in figure 2.5 - b) experiments, a concentrated stock solution (100 mM) of each HNE ligand (LMC223 and LMC249) was prepared in 100% (V/V) dimethyl sulfoxide (DMSO), since they are organic compound and consequently soluble in hydrophobic (and non-polar) solvents. Here we first tested – by repeating previous work in our lab [56] and doing co-crystallizations of PPE complexes with LMC188 and LMC269 – two different molar excesses, fivefold and tenfold the protein molar concentration, and doing all the tests only in 24-well Linbro plates. The best results were obtained using the molar excess of tenfold, so this was the first value of molar excesses we chose to use in the next new co-crystallizations with the two ligands designed for HNE.

Next, the incubation of the protein with each ligand took place using a molar excess of tenfold for about 45 minutes at room temperature, and with a final concentration of 5% DMSO. The crystallization conditions were similar to those used with the protein in its native form, varying the protein concentration and the ratio of the drops, between the volume (µL) of incubated sample (protein plus ligand) and the reservoir solution. This is explained in more detailed in the next table 6.4.

Table 6.4 – Summary of the first plate with PPE co-crystallizations made at 20°C. Incubations with ligands LMC223 and LMC249 by sitting-drop, and a 500 µL reservoir with 70% (V/V) MPD and 10 mM of sodium phosphate buffer pH 5.9.

	1	2	3	4
A	Drop: PPE + **LMC223** [Protein] = 15 mg/mL Ratio: (1:1) µL	Drop: PPE + **LMC223** [Protein] = 15 mg/mL Ratio: (1:2) µL	Drop: PPE + **LMC223** [Protein] = 12.5 mg/mL Ratio: (1:1) µL	Drop: PPE + **LMC223** [Protein] = 12.5 mg/mL Ratio: (1:2) µL
B	Drop: PPE + **LMC249** [Protein] = 15 mg/mL Ratio: (1:1) µL	Drop: PPE + **LMC249** [Protein] = 15 mg/mL Ratio: (1:2) µL	Drop: PPE + **LMC249** [Protein] = 12.5 mg/mL Ratio: (1:1) µL	Drop: PPE + **LMC249** [Protein] = 12.5 mg/mL Ratio: (1:2) µL

Obs: Ratio = [(Drop) : (Reservoir)] µL

Several hits were obtained, with the best ones in drops A2 and B3 after 12 days. Conditions B2 and B4 had the drop full of tiny crystals after the same number of days.

After those results, we decided to perform a few more trials of PPE co-crystallizations based on the previous results of table 6.4 to obtain fresher crystals to send to diffraction and optimize some conditions with LMC223 given that less crystals appeared with that ligand.

A second plate was prepared according to table 6.5, with each incubation done in triplicate, again with a molar excess of tenfold and using both techniques of sitting and hanging-drop. Volumes and composition of the reservoir were the same of the first co-crystallization plate, with the ratio defined between the volume of the drop and the reservoir.

Table 6.5 – Summary of the second plate with PPE co-crystallizations made at 20°C. Incubations with ligands LMC223 and LMC249, with a 500 µL reservoir composed by 70% (V/V) MPD and 10 mM of sodium phosphate buffer pH 5.9.

	1	2	3	4	5	6
A	Drop: PPE + **LMC223** [Protein] = 7.5 mg/mL Ratio: (1:1) µL			Drop: PPE + **LMC223** [Protein] = 7.5 mg/mL Ratio: (1:1) µL		
B	Drop: PPE + **LMC223** [Protein] = 15 mg/mL Ratio: (1:2) µL			Drop: PPE + **LMC223** [Protein] = 15 mg/mL Ratio: (1:2) µL		
C	Drop: PPE + **LMC249** [Protein] = 12.5 mg/mL Ratio: (1:1) µL			Drop: PPE + **LMC249** [Protein] = 12.5 mg/mL Ratio: (1:1) µL		
	↳ **Sitting-drop**			↳ **Hanging-drop**		

Crystals were formed after 10 days, in drops B1, B2, C1 and C4 (drop C5 also had crystals like C4).

6.3 X-ray diffraction: data collection at synchrotron

Several crystals of PPE complexes were frozen by plunging into liquid nitrogen without any additional cryoprotectant solution since MPD at 70% already has that property. The data collection was done in the same shift of the TEM-1 crystal described in the chapter 1, at Diamond synchrotron in beamline I24 (*microfocus MX*), under a nitrogen-gas stream at 100 K.

Five crystals were screened for diffraction but only one dataset was collected, belonging to one crystal of drop B3 of the 1st plate (protein at 12.5 mg/mL, incubation of PPE with LMC249, ratio 1:1 µL), and the details concerning all the crystals fished are present in section 7 of the results and discussion.

6.4 X-ray diffraction: structure determination and refinement process

The data obtained were processed with the *autoPROC* software bundle [44], which uses *XDS* program package [45] to do the indexing of each reflection measured (peaks found in all images collected) and *XSCALE* program for the scaling. It also uses the *STARANISO* estimation server [46] and the *AIMLESS* and *POINTLESS*, two supported programs of the *CCP4* program suite [47].

The structure was obtained by applying the indirect method of the molecular replacement to obtain the phase information using *PHASER* [48], a program that is incorporated into *PHENIX* [49] software package, using the coordinates of the PPE-JM102 complex as search model (PDB entry 4YM9 [39]) devoid of solvent molecules or ligands. The result was a final structure after several iterative cycles of model building and refinement with the programs *COOT* [50] and *phenix.refine* [51]. For the model validation was used the *MolProbity* [52], implemented in *PHENIX*. The maps with electronic density plus the atomic model were produced in *COOT* software.

7. Results and Discussion

7.1 PPE crystallization and data collection

First, commercial elastase from porcine pancreas easily crystallized and native crystals grew in 70% MPD and 10 mM sodium phosphate buffer pH 5.9, with a protein concentration of 20 mg/mL as described in table 6.1. The whole drop was filled with crystals upon 5 days, with a length between 120 and 150 µm. Two drops are observed in the next figure 7.1.

Figure 7.1 – Native PPE crystals present in two different drops, in 70% (V/V) MPD with 10 mM sodium phosphate buffer pH 5.9, by using the sitting-drop method and at 20°C. Bigger crystals are present in the drop on the left.

These results could confirm that the PPE batch at laboratory was still in good conditions and crystallized in the same condition that was previous optimized and so we proceeded with the co-crystallization trials. Ligands LMC223 and LMC249 were tested with two protein and ratios varying as described in table 6.4. In next figure 7.2, are present several drops with crystals produced in the first plate of PPE co-crystallizations by using the sitting-drop method, the incubation with ligands was tenfold of molar excess for 45 minutes and at room temperature. All crystals appeared after 12 days with the same reservoir solution of native crystals.

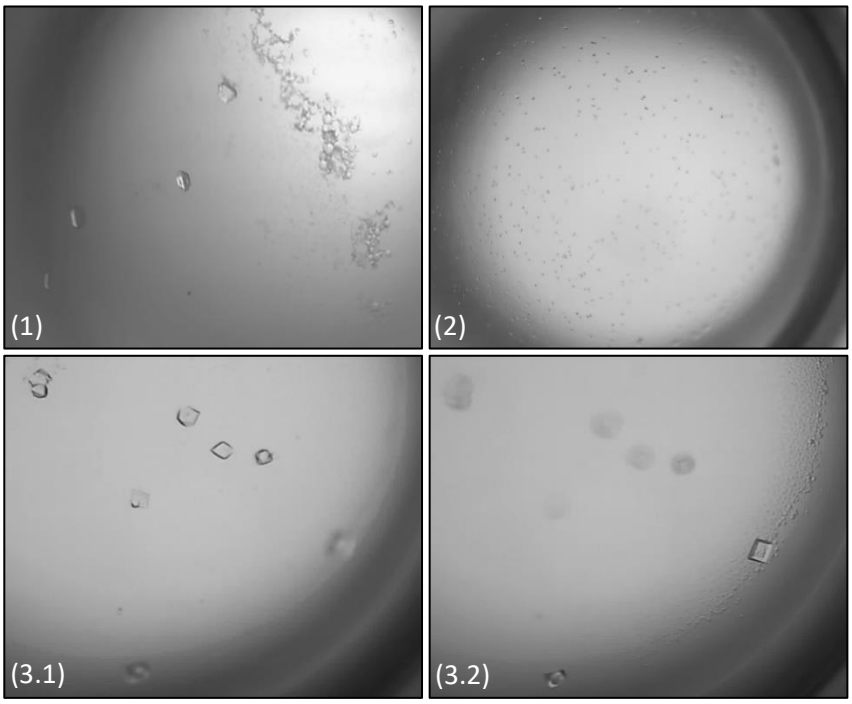


Figure 7.2 – Results from the first plate with PPE co-crystallizations. **(1)** Drop A2 – LMC223, ratio of 1:2 µL, [P] = 15 mg/mL; **(2)** Drop B4 – LMC249, ratio of 1:2 µL, [P] = 12.5 mg/mL; **(3.1)** and **(3.2)** Drop B3 – LMC249, ratio of 1:1 µL, [P] = 12.5 mg/mL.

The best crystals appeared in drop B3 of PPE co-crystallized with LMC249. Drop A2 had also at least four good crystals that could possibly be fished and measured. Crystals of those two conditions grew up to 150 µm. Also, drop B4 (Figure 7.2 – 2) helps to understand an example of a trial where many crystals were formed but with small dimensions, not bigger than 50 µm after the same number of days, as well as it happened in drop B2. That means that the crystal nucleation was well succeeded but probably the transition to the growth zone was too quick and the crystals did not have enough time to grow a little further.

A second plate (Figure 7.3) with more PPE co-crystallization experiments was done using the same ligands, incubation times and volume ratios of the previous plate, but this time trying both methods of sitting and hanging-drop. Also, another protein concentration was additionally tested to try to optimize and have more crystals of PPE incubated with LMC223 as described in table 6.5.

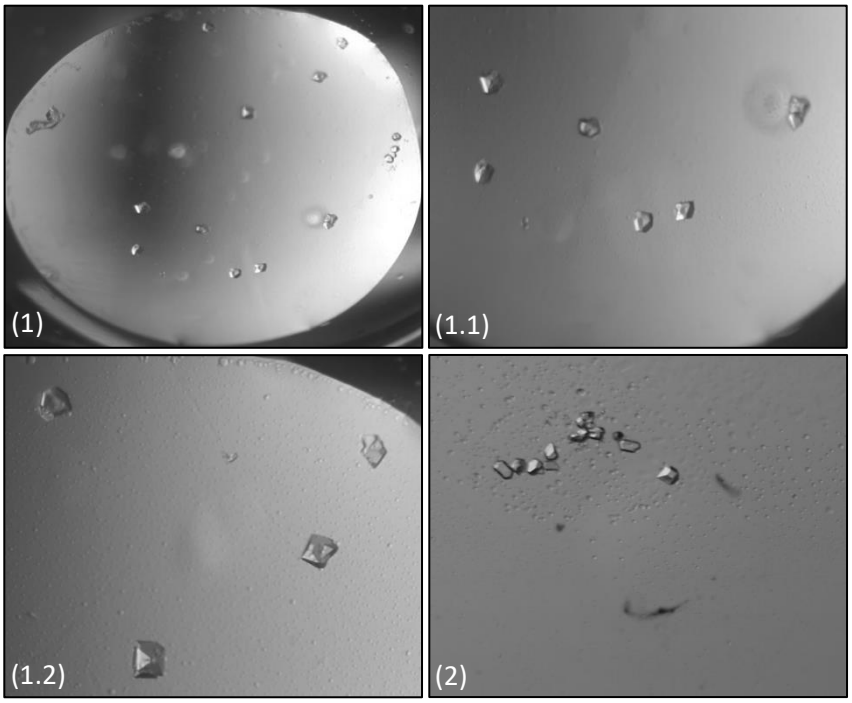


Figure 7.3 – Results from the second plate with PPE co-crystallizations, with the crystals visible after 10 days. **(1)** Drop C1 – LMC249, ratio of 1:1 µL, [P] = 12.5 mg/mL, sitting-drop; **(1.1)** and **(1.2)** different close-up's of drop C1 to highlight the best crystals; **(2)** Drop C4 – LMC249, ratio of 1:1 µL, [P] = 12.5 mg/mL, hanging-drop.

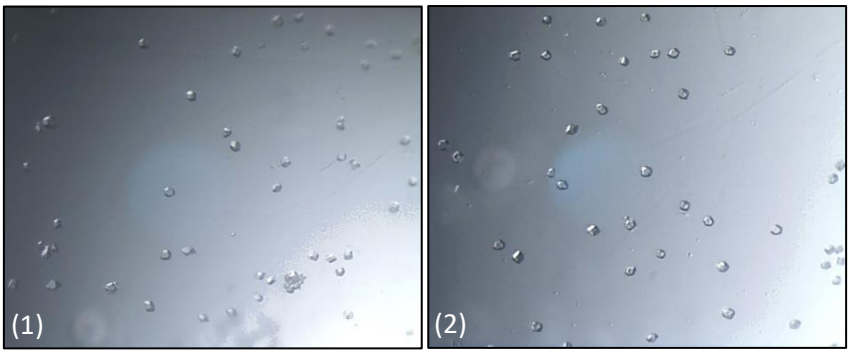


Figure 7.4 – Results also from the second plate of PPE co-crystallizations, showing the trials where appeared smaller crystals after 10 days. **(1)** drop B1 and **(2)** drop B2. Both for LMC223, ratio of 1:2 µL, [P] = 15 mg/mL, and sitting drop.

Results from figure 7.3 show again the success of this co-crystallization technique in obtaining crystals of PPE complexes. Still, it should be denoted that bigger crystals (a length of 100 µm) appeared in the condition of drop C1 and C4, with the drops B1 and B2 (Figure 7.4) having half of the size.

Also, it is important to mention that the protein concentration in which the best crystals from the co-crystallization trials appeared was mostly at 12.5 mg/mL and not at 20 or 15 mg/mL as it was firstly tested in the native form. This fact allows to conclude that the crystallization conditions, when it comes to the concentration of the protein in the drop, are different between native and complex form.

We chose five good crystals (in size and shape) from different drops with the co-crystallization experiments and they were directly frozen in liquid nitrogen and sent to the synchrotron facility to do the data collection. Namely: **1)** crystal from drop A2 of the 1st plate (Figure 7.2 – 1); **2)** and **3)** both from drop B3 of the 1st plate (Figure 42 – 3.1 and 3.2, one of each); **4)** and **5)** both from drop C1 of the 2nd plate (Figure 7.3 – 1.1 and 1.2, one of each). The four last crystals could be called redundant because were all from co-crystallizations with the ligand LMC249, however the two crystals from the 2nd plate were more recent.

Initially, it was tested the crystal **1)** but did not diffracted. The second crystal tested was the **2)**, that belonged to the incubation of PPE with LMC249, the protein at 12.5 mg/mL and ratio 1:1 µL, and this produced diffraction data with resolution around 1.4 Å. The next crystal measured was crystal **4)**, which contained the same ligand and incubation conditions but was a fresher crystal, producing a dataset with similar resolution of the previous one. Once these two crystals diffracted very well, and the remain crystals **3)** and **5)** were also from PPE complexes in the same conditions, they were not tested.

The DLS beamline platform provides a fast data processing service that was used to make a quick molecular replacement phasing step and assess the electronic density and check the presence of the ligand in the active site cleft. For both datasets the electron density yielded no electron density that could correspond to the LMC249 structure, only a small blob of density around Ser^{195}.

Nevertheless, in order to proceed with the complete pipeline to solve a structure by X-ray crystallography we proceeded with the data collected for crystal **2)** and figure 7.5 shows the respective loop and one of the images containing a diffraction pattern obtained with that PPE crystal.

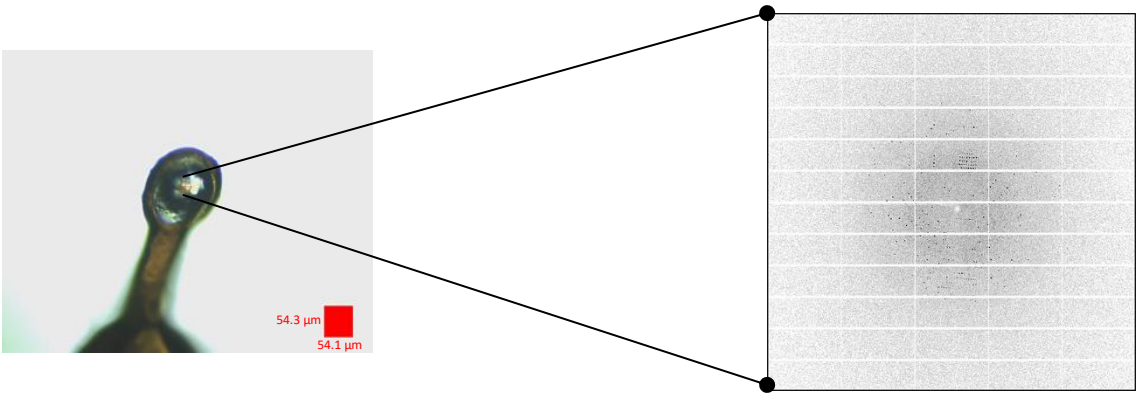


Figure 7.5 – Data collection for the crystal **2)**, PPE-LMC249 complex, in beamline I24 at Diamond Light Source. On the left is an image of the loop mounted with the crystal centered, and on the right is a diffraction pattern with several spots visible.

Next step is to process the data, phase, build and refine a model to determine that crystal's structure and perhaps to have an improvement in the density of a part of the ligand.

7.2 Data processing, model building and refinement process

The chosen dataset was processed through the *autoPROC* pipeline that works fast in processing, scale and merge the data and the following phase problem solving was achieved by using the molecular replacement method. These steps allow to calculate the preliminary electronic density ($2F_{obs}$-F_{calc} map) and structure model to analyze all the present density blobs. The statistics related with the data collection and processing of the collected dataset are summarized in next table 7.1, after choosing the most relevant parameters.

Table 7.1 – Data collection and processing statistics of the PPE crystal in complex with the ligand LMC249. Data collected at Diamond Light Source synchrotron in beamline I24.

Data collection and processing	
Wavelength of measurements (Å)	0.96862
Resolution range (Å)	33.94 – 1.35 (1.40 – 1.35)
Space group	$P\,2_1\,2_1\,2_1$ (Orthorhombic)
Unit cell parameters: *a*, *b*, *c*, *α*, *β*, *γ*	50.72 Å 57.69 Å 74.70 Å 90.0 ° 90.0 ° 90.0 °
No. of total reflections	316657 (29951)
No. of unique reflections	48645 (4767)
Completeness (%)	99.47 (98.57)
Multiplicity	6.5 (6.3)
$< I/\sigma(I) >$	9.86 (0.73)
CC 1/2 (%)	99.9 (55.9)
Wilson B-factor ($Å^2$)	18.42
R_{merge} (%)	7.9 (1.1)
R_{meas} (%)	8.6 (1.2)
Matthews coefficient ($Å^3/Da$)	2.07
Solvent content (%)	40.5
No. of molecules in the asymmetric unit	1

Obs: statistics shown in parentheses are for the highest-resolution shell.

The first electron density map produced was analyzed with the *COOT* software [50] and consecutive improvements were made to the model, for example corrections in the 3D positions of residues, rotamers adjustment and the addition of waters, in order to get a better fitting between the calculated and the observed density. After an iterative process of refinements, the convergence was reached and consequently a final map was obtained, then it was validated using a structure-validation service named *MolProbity* [52]. The statistics of the final refinement of PPE-LMC249 complex and its parameters of the model validation can be consulted in table 7.2 below.

Table 7.2 – Statistics of final refinement and the model validation parameters obtained in this experimental work.

Refinement statistics	
Resolution limits (Å)	37.35 – 1.35 (1.40 – 1.35)
No. of reflections used in refinement	48632 (4767)
No. of reflections used for R_{free}	2384 (248)
R_{work} (%)	16.25
R_{free} (%)	19.71
No. of non-hydrogen atoms:	2006
- Macromolecules	1843
- Ligands	6 (1 phosphate and 1 sodium)

- Waters	157
Protein residues	240
R.M.S.D. of bond lengths (Å)	0.006
R.M.S.D. of bond angles (°)	1.17
Ramachandran favored (%)	98.32
Ramachandran allowed (%)	1.68
Ramachandran outliers (%)	0
Rotamer outliers (%)	0
C_β outliers (%)	0
Clashscore	1.93
MolProbity overall score	0.96
Average B-factor ($Å^2$)	25.40

Obs: statistics shown in parentheses are for the highest-resolution shell.

Regarding the refinement strategy followed before and the results obtained, during the last few stages, we tested to refine the individual B-factors anisotropically and indeed it was obtained a lower R_{work}/R_{free} ratio (Table 7.2), when compared with other types of refinements as the isotropic individual B-factors or TLS groups. Inside a crystal structure the atoms are not static, vibrating about their equilibrium positions, and the anisotropic vibrations are a more realistic estimation since they consider that the vibration is not equal in all directions and the atoms move with a harmonic potential. Also, when we have data at 1.4 Å resolution or better, it is frequently done an anisotropic refinement of heavy atoms (C, N, O or heavier) and the water isotropically. [53] So, this supports and clarifies why that type of refinement was the best option in this case.

Now, it is possible to show the final atomic model with its refined density map (Figure 7.6), highlighting the region of the active site and one blob of positive electronic density where it was expected to find the ligand.

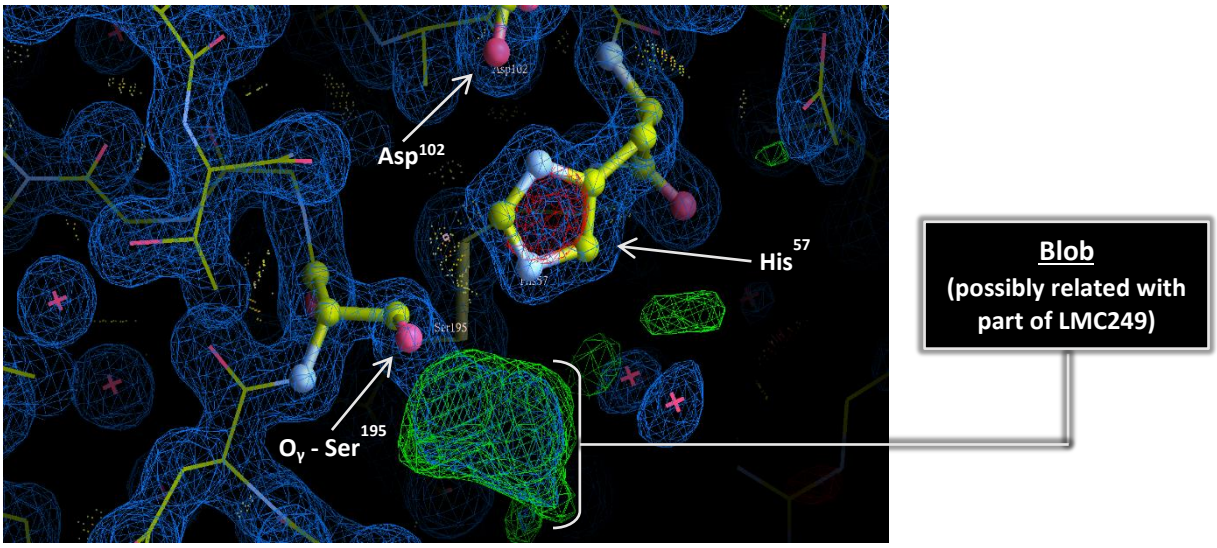


Figure 7.6 – Electronic density map of PPE complex with LMC249 focused on the active site region, obtained after the final refinement. The $2F_{obs}$-F_{calc} difference map is in blue at a contour level of 1.2 δ. The F_{obs}-F_{calc} map is in green for positive density and red for the negative, both contoured at 3.6 δ. Residues His^{57}, Ser^{195} and Asp^{102} of the catalytic triad of the enzyme are in ball & stick representation, and the data resolution of the crystal is 1.35 Å. Image produced with *COOT* software. [50]

The analysis of the previous figure shows mostly an improvement of the density in the active site vicinity with the presence of one main blob, a little bigger and more defined in terms of electronic density contiguous to the *gamma* oxygen of Ser^{195}, when compared with the first model built after the fast data processing of the DLS platform. Also, is visible the relative proximity of this blob to the other two residues – His^{57} and Asp^{102} – of the PPE catalytic triad. Though this information is not enough to elucidate the type of inhibition mechanism, if it happens a nucleophilic attack on the carbonyl or on the sulfonyl group of the ligand (mechanism present in figure 1.8). The previous structural studies in our laboratory [56] with the ligands designed for PPE (Figure 6.2) showed that the mechanism of action behind the inhibition is the attack of the *gamma* oxygen of Ser^{195} on the sulfonyl.

The main conclusion is the absence of electronic density for the ligand in the region of enzyme pocket. This outcome can be due to the high mobility that ligand LMC249 presented in the crystalline matrix and this structural biology technique studies biomolecules based on the most probable positions that they adopt.

A possible explanation can be because the ligand is too large to do a complete approach to the catalytic site in a right three-dimensional conformation that allows the nucleophilic attack of the ligand during the inhibition reaction, with emphasis in the two ethyl groups of the LMC249 that are present between the carbonyl group and the sulfonyl group attached in the ring (the main and most reactive part) of the compound.

What might be happened was also that the nucleophilic attack was stablished but the following steps of the inhibition reaction did not happen properly and the final covalent binding between protein and ligand was in some way destabilized by the longer chains of the ethyl groups present in the LMC249. This could be discussed as a reason because a previous structure of PPE in complex with LMC240 (Figure 6.2) was achieved in our laboratory and the structure of that compound is practically the same of LMC249 tested in this present work, the only difference is the type of alkyl group present. As already mentioned, the LMC249 has two ethyl groups and LMC240 has two methyl groups, making the second molecule less long.

The entire molecule should be important to maintain it stabilized and connected to the enzyme and the ligands are designed to bind the enzyme pocket in a best conformation possible with different contributions through the Van der Waals interactions, hydrogen bonding, hydrophobic interactions, between others. Thus, those ethyl groups in the LMC249 may consequently disturb this set of molecular interactions and leaving that the ligand does not stay permanently bound to the enzyme as would be desired. Using alkyl groups in the compounds, depending on their size or length, they can act as a “steric shield” in part of the enzyme pocket and interfere with the binding because those groups can protect nearby reactive groups from contact with regions that would normally be involved in the reaction.

The final model (final PDB file after the last refinement) obtained for the co-crystallization of PPE with LMC249 is represented in figure 7.7 as surface representation for the vacuum electrostatics of the model, colored by the electrostatic potential (red is for the negative charges and blue is for the positive ones. Also, in green and by cartoon representation is the main chain of the protein and according to its secondary structure. It was produced in *PyMol* [55] software and is important to mention this representation is a rough approximation of the local contact potential and partial charges of the whole protein and not a precise and careful representation.

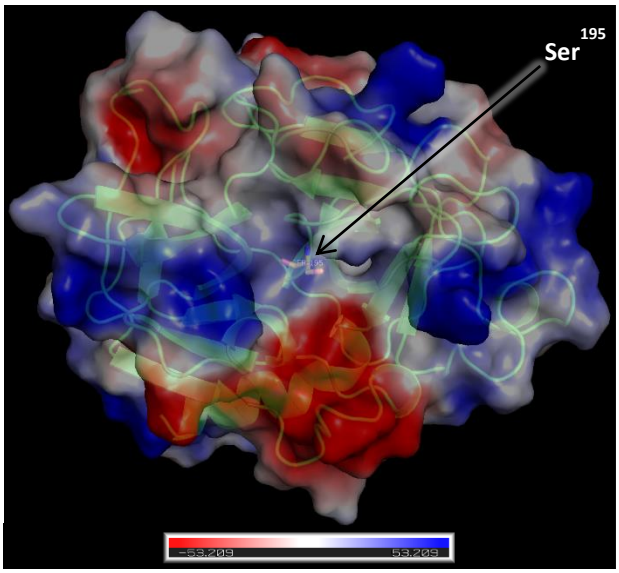


Figure 7.7 – Representation of the electrostatic surface for the final model of PPE with LMC249. Ser^{195} of the active site is properly highlighted at the center and is shown in sticks representation.

In general, the electrostatic potential maps are useful to predict the behavior of molecules that may interact with the proteins and eventually form protein-ligand complexes. Besides, it can be used to verify or search for the reactive site of a molecule, that is, a certain charged region of the molecule that can have affinity to bind with positive charged particles of certain organic molecules or designed ligands.[54] Figure 7.7 shows that Ser^{195} in the enzyme pocket is near to an area with a predominant negative electrostatic potential, and between two sites with positive electrostatic potential. These regions can be important in stabilizing the ligands that possibly can bind the active site.

Additionally, after making another analysis of structural comparison between several native PPE structures deposited in PDB database – e.g. the PPE native structure with PDB entry 3EST – and the final structure obtained in this work, there are no significant differences in PPE conformation and folding when it has a ligand in its proximity (with no covalent binding between them).

7. Conclusions

The work developed in collaboration with the group of Prof. Rui Moreira at *iMed.ULisboa* had as main aim the structural studies about the inhibition mechanism of two serine-based enzymes: TEM-1 β-lactamase and Porcine Pancreatic Elastase (PPE).

For the first protein in study, protein expression and purification proved to be challenging to obtain a good quantity of soluble protein and avoid protein degradation along the steps. The crystallization was not straightforward, with a few native crystals obtained but they did not diffract in synchrotron. So, it was not possible to proceed with the studies of the protein-ligand complexes.

In the second part of the experimental work, there were performed trials with other serine enzyme and a well-studied protein in our laboratory. Using PPE, we intended to confirm if its inhibition would occur by using ligands LMC223 and LMC249, which were designed for the human elastase (HNE), and know if the mechanism would be by a nucleophilic attack of Ser^{195} on the carbonyl group of the β-lactam derived ring present in the ligands, or on the sulfonyl group. It is important to highlight that still do not exist structures for HNE complexes with those ligands.

At the end, we managed to obtain a 3D-structure model of PPE in complex with LMC249 by X-ray diffraction, at Diamond Light Source, with a resolution of 1.35 Å. Despite of this, the electron density maps were not able to show clear density for the ligand in the region of the catalytic triad (His^{57}-Asp^{102}-Ser^{195}) to prove any covalent interaction established.

8. Future perspectives

As future work for β-lactamase, it is important to continue the crystallization trials by soaking or co-crystallized with their designed ligands and find the best crystallization conditions directly for the complexes. Also, a biophysical characterization of the protein would be useful in order to do a pH and buffer formulation screening to find better conditions for protein stabilization and consequently an improvement in the crystallization process. Techniques as the thermal shift assays (TSA) or dynamic light scattering (DLS) can be applied.

Related with the PPE studies, for the interaction between elastase and ligands LMC223 and LMC249, possibly is important to do some inhibition assays to know the inhibition constants (K_i) specifically for these complexes. Additionally, to improve the conditions and obtain complexes with the ligand bound, we can do TSA studies with different incubation conditions, varying the incubation time and the molar excess. This is still can be used to know protein-ligand affinities by determine the binding constant (K_D).

9. References

1 - O'Neil, J. Review on Antibiotic resistance. Antimicrobial Resistance: Tackling a crisis for the health and wealth of nations. Heal. Wealth Nations (2014). Available on: https://amr-review.org/Publications.html

2 - O'Neill, J. Tackling drug-resistant infections globally: final report and recommendations. Rev. Antimicrob. Resist. (2016). doi:10.1016/j.jpha.2015.11.005. Available on: https://amr-review.org/Publications.html

3 - Expresso online: "Resistência aos antibióticos vai matar mais do que o cancro". Jornal Expresso (2018) Available on: http://expresso.sapo.pt/sociedade/2018-07-02-Resistencia-aos-antibioticos-vai-matar-mais-do-que-o-cancro (consulted on Jul. 10th, 2018)

4 - Naas, T. et al. Beta-lactamase database (BLDB)–structure and function. J. Enzyme Inhib. Med. Chem. 32, 917–919 (2017)

5 - Gupta, V. An update on newer beta-lactamases. Indian J. Med. Res. 126, 417–427 (2007)

6 - Tärnberg, M. Extended-spectrum beta-lactamase producing Enterobacteriaceae: aspects on detection, epidemiology and multi-drug resistance. Linköping University medical dissertations (2012)

7 - Hall, B. G., Barlow, M. Revised Ambler classification of β-lactamases. J. Antimicrobial Chemotherapy. 55, 1050–1051 (2005).

8 - Wax, R. G. et al. Bacterial Resistance to Antimicrobials. Chapter 6. 2nd edition. CRC Press, 2008.

9 - Patrick, G. L. An Introduction to Medicinal Chemistry. Chapter 19. 5th edition. Oxford University Press, 2013.

10 - Wang, X., Minasov, G., Shoichet, B. K. Noncovalent interaction energies in covalent complexes: TEM-1 β-lactamase and β-lactams. Proteins Struct. Funct. Genet. 47, 86–96 (2002).

11 - Lejeune, A. et al. Quantitative analysis of the stabilization by substrate of *Staphylococcus aureus* PC1 β-lactamase. Chemistry & Biology 8, 831-842 (2001).

12 - Bush, K., Jacoby, G. A. Updated functional classification of β-lactamases. Antimicrob. Agents Chemother. 54, 969–976 (2010).

13 - Drawz, S. M., Bonomo, R. A. Three decades of β-lactamase inhibitors. Clin. Microbiol. Rev. 23, 160–201 (2010).

14 - Damblon, C. et al. The catalytic mechanism of beta-lactamases: NMR titration of an active-site lysine residue of the TEM-1 enzyme. Proc. Natl. Acad. Sci. 93, 1747–1752 (1996).

15 - Wang, X., Minasov, G., Shoichet, B. K. Noncovalent interaction energies in covalent complexes: TEM-1 β-lactamase and β-lactams. Proteins Struct. Funct. Genet. 47, 86–96 (2002).

16 - Medeiros, A. A. Beta-lactamases. British Med. Bulletin. 40, 18-27 (1984).

17 - Beardsell, M. et al. β-Sultams - A novel class of serine protease inhibitors. Chem. Commun. 497–498 (2001).

18 - Brito, J. A., Archer, M. Chapter 9: X-ray Crystallography, in Practical Approaches to Biological Inorganic Chemistry (edited by Crichton, R.R. and Louro, R. O.). 1st Edition. 217-255. Elsevier, 2013.

19 - Seed Bead User Guide - *Hampton Research* (HR4-780). Available on: https://hamptonresearch.com/documents/product/hr009595_4-780_seed_bead_steel.pdf (consulted on Jul. 24th, 2018)

20 - Patrick, M. H. Crystallography made crystal clear: A guide for users of macromolecular models. 3rd edition. Biochem Mol Biol Educ, 2007.

21 - Romão, M. J. Cristalografia de Proteínas: metodologia e aplicações em Bioquímica. Volume 53, 18-36. Boletim de Biotecnologia, 1996.

22 - Seeding Tool User Guide - *Hampton Research* (HR8-133). Available on: https://hamptonresearch.com/documents/product/hr009609_8-133_-_user_guide.pdf (consulted on Aug. 31th, 2018)

23 - Moore, J. W. et al. ChemPRIME - Chapter 10.3: Crystal Systems. Available online on: https://chem.libretexts.org/Textbook_Maps/General_Chemistry/Book%3A_ChemPRIME_(Moore_et_al.)/10Solids%2C_Liquids_and_Solutions/10.03%3A_Crystal_Systems (consulted on Sep. 10th, 2018)

24 - Translational Symmetry: I. Lattices and Unit Cells - Available online on: http://pd.chem.ucl.ac.uk/pdnn/symm1/trans1.htm (consulted on Sep. 7th, 2018)

25 - Garman, E. F. Developments in X-ray crystallographic structure determination of biological macromolecules. Science (80-.). 343, 1102–1108 (2014)

26 - Brem, J. et al. Structural basis of metallo-β-lactamase, serine-β-lactamase and penicillin-binding protein inhibition by cyclic boronates. Nat. Commun. 7, 1–8 (2016).

27 - Inglis, S. R., Strieker, M., Rydzik, A. M., Dessen, A., Schofield, C. J. A boronic-acid-based probe for fluorescence polarization assays with penicillin binding proteins and β-lactamases. Anal. Biochem. 420, 41–47 (2012).

28 - Nelson D.L., Cox M. M., Lehninger: Principles of Biochemistry. Chapter 6. 5th edition. New York: W H Freeman and Company, 2008.

29 - Salt or Protein Crystals? – *Hampton Research* Guide. Available online on: https://hamptonresearch.com/documents/product/hr007641_cg101_salt_or_protein_crystals.pdf (consulted on Nov. 4th, 2018).

30 - Contesini, F. J., Melo, R. R. de, Sato, H. H. An overview of Bacillus proteases: from production to application. Crit. Rev. Biotechnol. 38, 321–334 (2018).

31 - Berg, J. M., Tymoczko, J. L., Stryer, L. Biochemistry. 5th edition. New York: W H Freeman and Company, 2002. Available on: https://www.ncbi.nlm.nih.gov/books/NBK21154/

32 - Neitzel, J. J. Enzyme Catalysis: The Serine Proteases. Nat. Educ. 3, 21 (2010).

33 - Rani, K., Rana, R., Datt, S. Review on Latest Overview of Proteases. Int. J. Curr. Life Sci. 2, 12–18 (2012).

34 - MEROPS: Online Database for Peptidases. Available on: https://www.ebi.ac.uk/merops/about/classification.shtml (consulted on Nov. 25th, 2018).

35 - Neurath, H. Evolution of proteolytic enzymes. Science (80-.). 224, 350–357 (1984).

36 - Keller M.A., Piedrafita G., Ralser M. The widespread role of non-enzymatic reactions in cellular metabolism. Current Opinion in Biotechnology. 2015; 34; 153-161. 10.1016/j.copbio.2014.12.020

37 - Biochemistry Online - B: Mechanisms of enzyme catalysis. Available on: http://biochem-vivek.tripod.com/id45.html (consulted on Nov. 25th, 2018).

38 - Lecturio's e-learning platform – "Serine Protease: Background & Catalytic Mechanism" from the Biochemistry course. Available on: https://app.lecturio.com/medical-courses/serine-proteases-enzyme-catalysis.lecture (consulted on Nov. 28th, 2018).

39 - Hofbauer, S. et al. Stabilization of porcine pancreatic elastase crystals by glutaraldehyde cross-linking. Acta Crystallogr. Sect. Struct. Biol. Commun. 71, 1346–1351 (2015).

40 - Bode, W., Meyer, E., Powers, J. C. Human Leukocyte and Porcine Pancreatic Elastase: X-ray Crystal Structures, Mechanism, Substrate Specificity, and Mechanism-Based Inhibitors. Biochemistry 28, 1951–1963 (1989).

41 - Huang, W. et al. X-ray snapshot of the mechanism of inactivation of human neutrophil elastase by 1,2,5-thiadiazolidin-3-one 1,1-dioxide derivatives. J. Med. Chem. 51, 2003–2008 (2008).

42 - Ruivo, E. F. P. et al. Clickable 4-Oxo-β-lactam-Based Selective Probing for Human Neutrophil Elastase Related Proteomes. ChemMedChem 2037–2042 (2016). doi:10.1002/cmdc.201600258

43 - Areias, L. R. P. et al. A unified approach toward the rational design of selective low nanomolar human neutrophil elastase inhibitors. RSC Adv. 5, 51717–51721 (2015).

44 - Vonrhein, C. et al. Data processing and analysis with the autoPROC toolbox. Acta Crystallogr. Sect. D Biol. Crystallogr. 67, 293–302 (2011).

45 - Kabsch, W. XDS. Acta Cryst. Sect D Biol Crystallogr. 125–132 (2010). doi:10.1107/S0907444909047337

46 - Tickle I.J., Flensburg C., Keller P., et al. STARANISO. Cambridge, United Kingdom: Global Phasing Ltd (2017).

47 - Winn, M. D. et al. Overview of the CCP4 suite and current developments. Acta Crystallogr. Sect. D Biol. Crystallogr. 67, 235–242 (2011).

48 - McCoy, A. J. et al. Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674 (2007).

49 - Echols, N. et al. Graphical tools for macromolecular crystallography in PHENIX. J. Appl. Crystallogr. 45, 581–586 (2012).

50 - Emsley, P., Lohkamp, B., Scott, W. G., Cowtan, K. Features and development of Coot. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 486–501 (2010).

51 - Afonine, P. V. et al. Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr. Sect. D Biol. Crystallogr. 68, 352–367 (2012).

52 - Chen, V. B. et al. MolProbity: All-atom structure validation for macromolecular crystallography. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 12–21 (2010).

53 - Phenix's FAQs. Available on: https://www.phenix-online.org/documentation/faqs/refine.html#b-factors-adps-tls (consulted on Nov. 22th, 2018).

54 - *LibreTexts™* - Physical & Theoretical Chemistry, Electrostatic Potential maps. Available on: https://chem.libretexts.org/Textbook_Maps/Physical_and_Theoretical_Chemistry_Textbook_Maps/Supplemental_Modules_(Physical_and_Theoretical_Chemistry)/Chemical_Bonding/Fundamentals_of_Chemical_Bonding/Electrostatic_Potential_maps (consulted on Nov. 26th, 2018).

55 - The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.

56 - Carvalho, L.A.R., Almeida, V.T., Brito, J.A., Lum, K., Goncalves, L.M., Lucas, S.D.L., Archer, M., Cravatt, B., Moreira, R. 3-Oxo-beta-Sultams as a Versatile Sulfonylating Chemotype for Activity-Based Profiling of Serine Hydrolases, *in preparation*.

Appendix A

UniProt P62593 - sequence (without signal peptide, length = 263 aa):

HPETLVKVKDAEDQLGARVGYIELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRVDAGQEQLGRRIHYSQNDL
VEYSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRLDRWEPELNEAIPNDER
DTTMPAAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPLLRSALPAGWFIADKSGAGERGSRGIIAALGPDGK
PSRIVVIYTTGSQATMDERNRQIAEIGASLIKHW

pQE30-TEM1 plasmid - Sanger sequencing by *GATC Biotech* in October 2017 (length = 288 aa):

KHHHHHHHSDYDIPTTEXLYFQGXHPETLVKVKDAEDQLGARVGYIELDLNSGKILESFRPEERFPMMSTFKVLLCG
AVLSRIDAGQEQLGRRIHYSQNDLVEYSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNM
GDHVTRLDRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPLLRSALPAGWF
IADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIAEIGASLIKHW**C**

SDS-PAGE stock solutions:

Table A.1 – Recipes for SDS-PAGE gels, at 12% and 15%.

	12 %	**15 %**
	Separating gel	
Distilled H_2O	3.4 mL	2.4 mL
1.5 M Tris-HCl pH 8.8	2.5 mL	2.5 mL
20% (w/v) SDS	50 µL	50 µL
Acrylamide : Bis-acrylamide (30% : 0.8% w/v)	4.0 mL	5.0 mL
10% (w/v) Ammonium persulfate (APS)	100 µL	100 µL
TEMED	5 µL	5 µL
Total volume	10.055 mL	
	Stacking gel	
Distilled H_2O	3.075 mL	
0.5 M Tris-HCl pH 6.8	1.25 mL	
20% (w/v) SDS	25 µL	
Acrylamide : Bis-acrylamide (30% : 0.8% w/v)	0.67 mL	
10% (w/v) Ammonium persulfate (APS)	50 µL	
TEMED	5 µL	
Total volume	5.075 mL	

- **10x SDS-PAGE running buffer (1 L):** 30 g Tris-Base; 144 g Glycine; 5 g SDS.

Western blot stock solutions:

- **10x Transfer buffer (1 L):** 30 g Tris-Base; 144 g Glycine. For blotting (200 mL), prepare 20 mL 10x Transfer buffer, 20 mL methanol and 160 mL dd H_2O.
- **20x TBS (1 L):** 48.6 g Tris-Base; 175.3 g NaCl at pH 7.6.
- **1x TBS-T (1 L):** 50 mL 20x TBS; 1 mL Tween-20; 949 mL dd H_2O.
- **Blocking buffer 3% (w/v):** 0.75 g of low fat powder milk in 25 mL of 1x TBS-T.

Appendix B

Supplementary materials: results from TEM-1 β-lactamase purification attempts in order to optimize the protocol

- 3rd attempt: results of the histrap column and a batch/gravity-flow purification

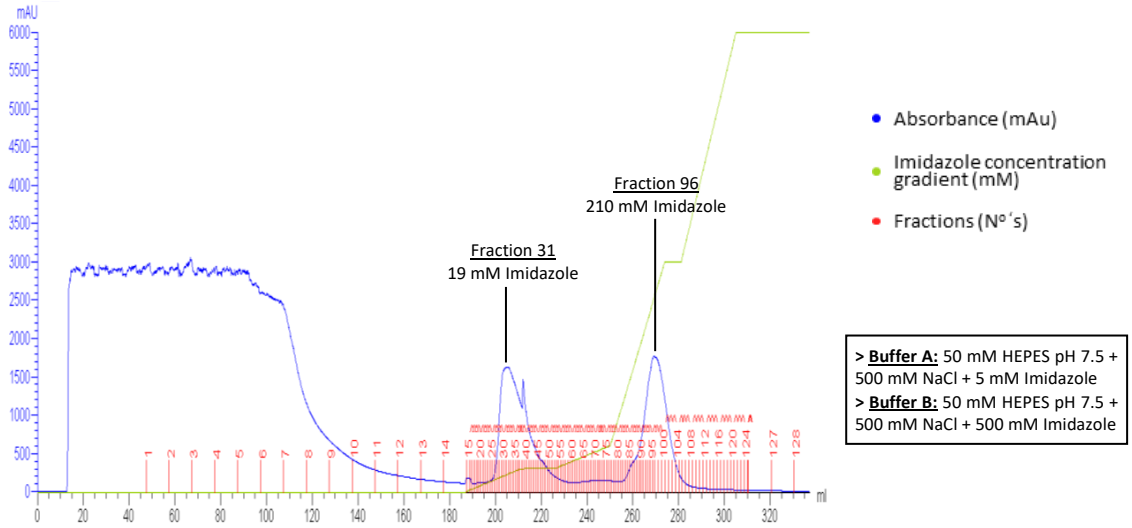


Figure B.1 – Histrap purification of 3rd attempt of TEM-1 purification. Elution was done with linear gradients of [I] imidazole concentration: 1st with no [I] for 30 mL; 2nd up to 25 mM [I] for 30 mL; 3rd up to 50 mM [I] for 30 mL; 4th from 50 mM to 250 mM [I] for 75 mL; 5th a final elution with 500 mM [I] for 30 mL. Flow rate was 3 mL/min. Total volume of injected sample was 38 mL and it was reinjected in column before starting the gradients. Main peaks are highlighted with [I] corresponding values.

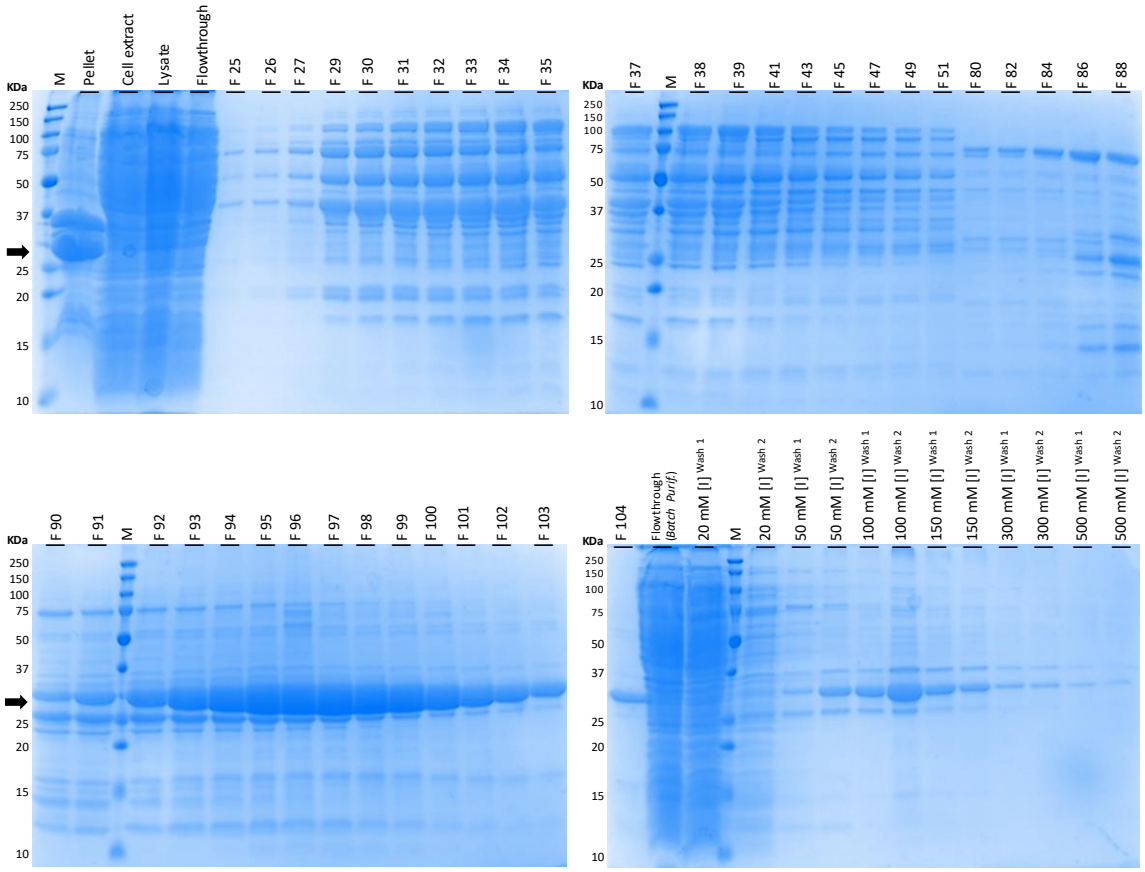


Figure B.2 – 15% SDS-PAGE after 3rd attempt of affinity chromatography, performed by two methods: a histrap purification and with a batch/gravity-flow purification, with different [I] imidazole concentrations. M - Precision Plus Protein™ Dual Color Standards. Arrows indicate where is expected to be TEM-1 band. Pooled fractions: **Pool ++**: F 90→F 104; **Pool (from batch purification)**: 50 mM [I] Wash2 → 300 mM [I] Wash2.

- 3rd attempt: results of ionic (anionic) exchange chromatography

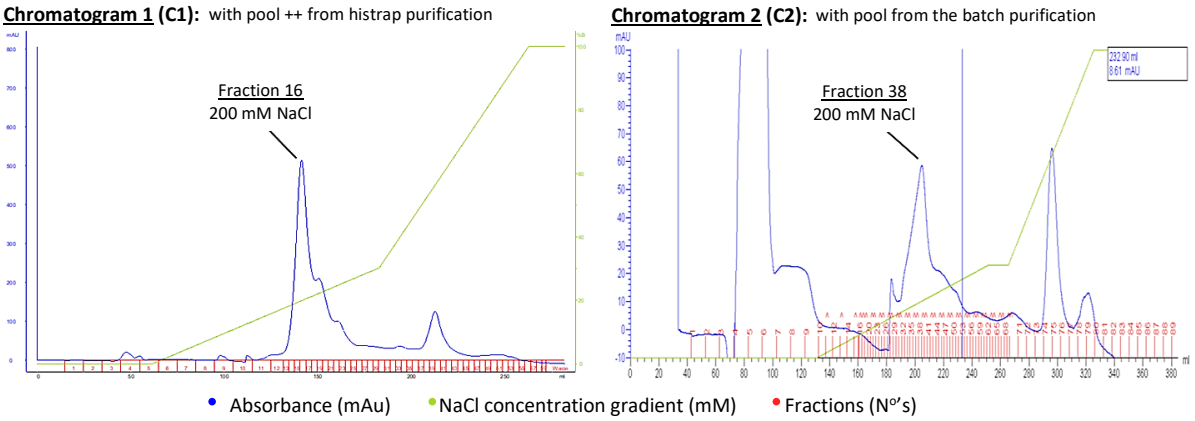


Figure B.3 – Chromatograms of the 3rd attempt of TEM-1 purification with a HiLoad 16/10 Q-Sepharose High Performance column and the buffers used were: buffer A with 20 mM HEPES pH 7.5; buffer B with 20 mM HEPES pH 7.5 + 1 M NaCl. Final volumes of injected samples were: 5 mL Sample + 5 mL buffer A in **C1**, and 10 mL Sample in **C2**. Purification was done with linear gradients of sodium chloride (NaCl) concentration (0 → 1 molar). First, the column was washed with washed with buffer A for 20 minutes and then was done a 1st gradient up to 300 mM [NaCl] for 1 hour. After that, a 2nd gradient between 300 mM and 1 M [NaCl] was performed for 25 minutes. Flow rate was 2 mL/min. Main peak is highlighted with the NaCl concentration and fraction number.

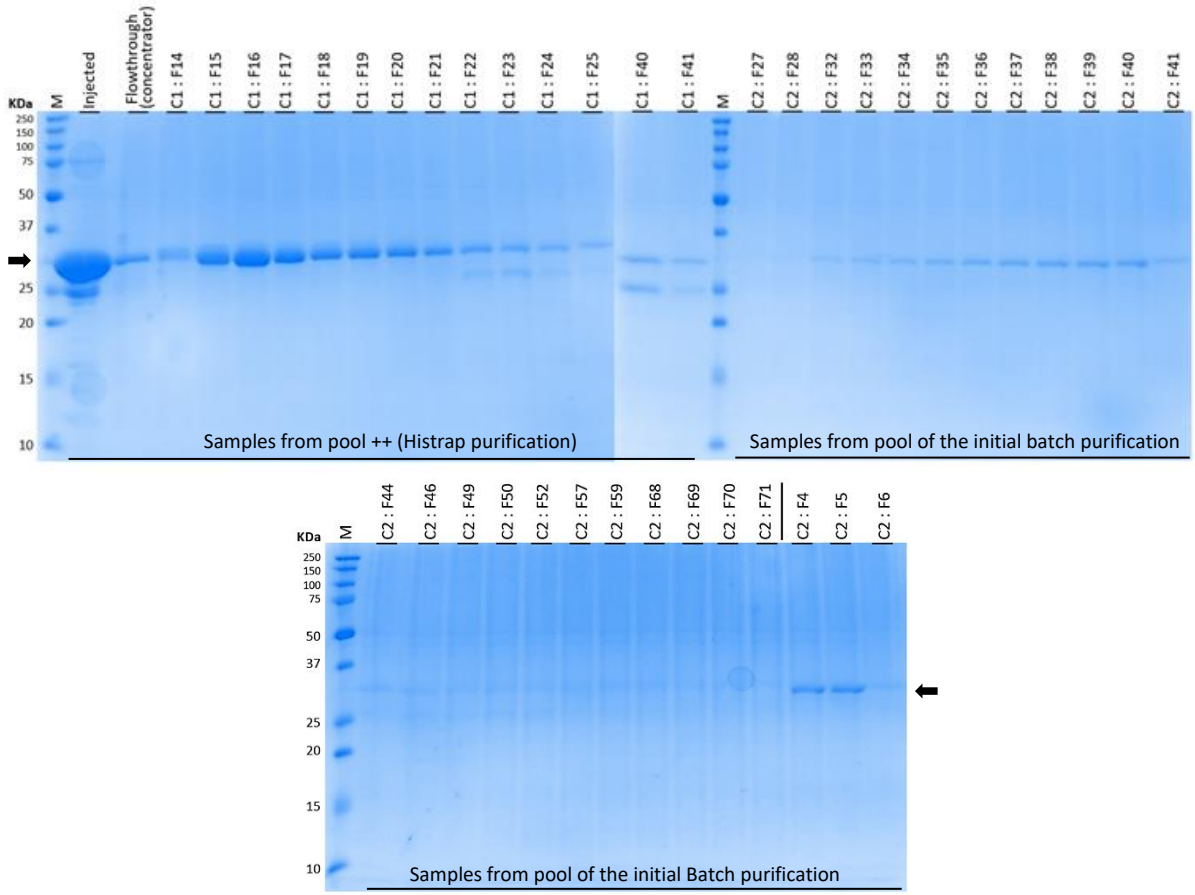


Figure B.4 – 15% SDS-PAGE with fraction from ionic exchange purification after the 3rd attempt of protein purification. M - Precision Plus Protein™ Dual Color Standards. The arrows indicate the molecular weight where is expected to be TEM-1 band. Pooled fractions: **Pool +** (from initial pool ++ of histrap purification) - F14 → F21; **Pool -** (from pool of the initial batch purification) - F32 → F41 + F4 + F5.

- 3rd attempt: results of size exclusion chromatography (SEC)

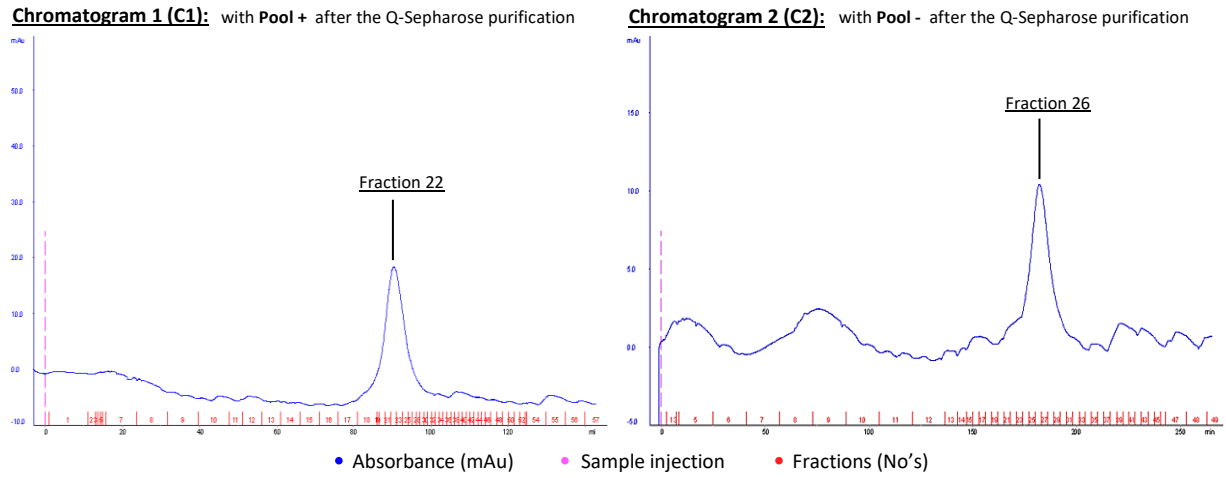


Figure B.5 – Chromatograms of 3rd attempt of TEM-1 purification, after two SEC's with a HiLoad 16/10 Sephadex 200 pg column and the buffer used was 20 mM HEPES pH 7.5 plus 150 mM NaCl. Injected sample volume was 1.5 mL for both runs. Flow rate was 0.5 mL/min. Main peaks are highlighted with the identification of the fraction.

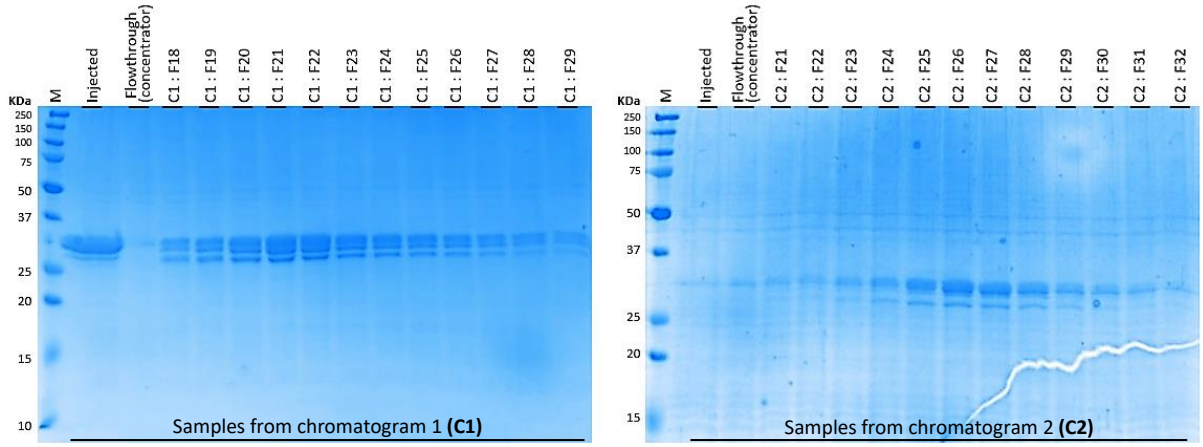


Figure B.6 – 15% SDS-PAGE to analyze the two SEC's of 3rd purification attempt. Gel on the left refers to samples obtained after purification of previous **pool +** and gel on the right has the samples after purification of **pool –**.
M - Precision Plus Protein™ Dual Color Standards.

- 4th attempt: this allowed to have the **1st batch of pure TEM-1**, and all the purification details and results are present in results and discussion section.

- 5th attempt [**2nd batch of TEM-1**]: results of the histrap purification

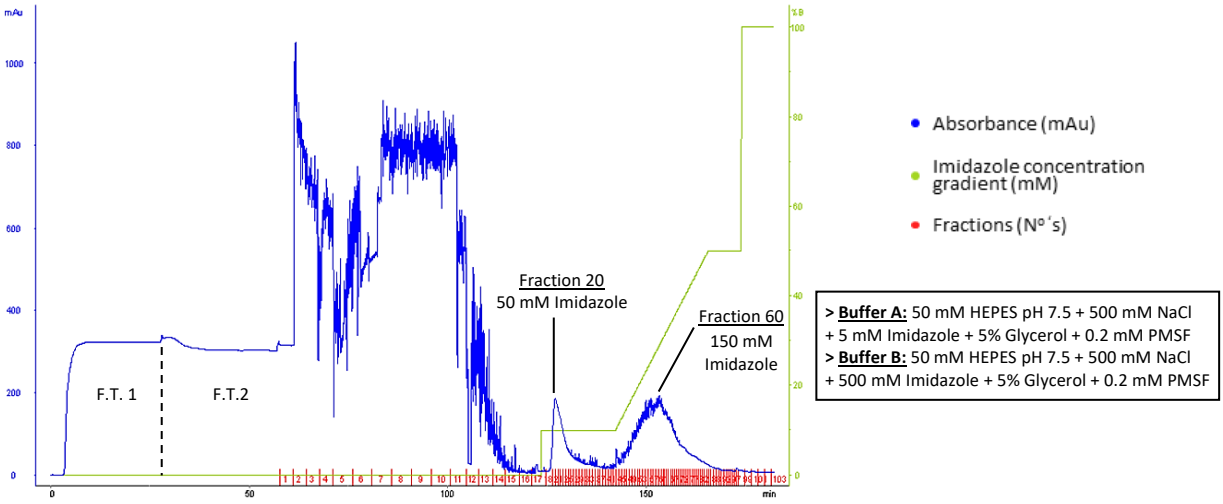


Figure B.7 – Chromatogram of 5th attempt of TEM-1 purification with a 5 mL histrap fastflow column. Elution was performed with linear gradients of imidazole ([I]): 1st with no [I] for 10 minutes; 2nd with 50 mM [I] for 10 minutes; 3rd from 50 mM to 250 mM [I] for 25 minutes; 4th a final elution with 500 mM [I] for 5 minutes. Flow rate was 3 mL/min. Total volume of injected sample (F.T. 1 - Flowthrough 1) was 85 mL and reinjected in the column (F.T. 2 - Flowthrough 2) before starting the gradients. Main chromatogram peaks are highlighted with imidazole concentration and corresponding fraction number.

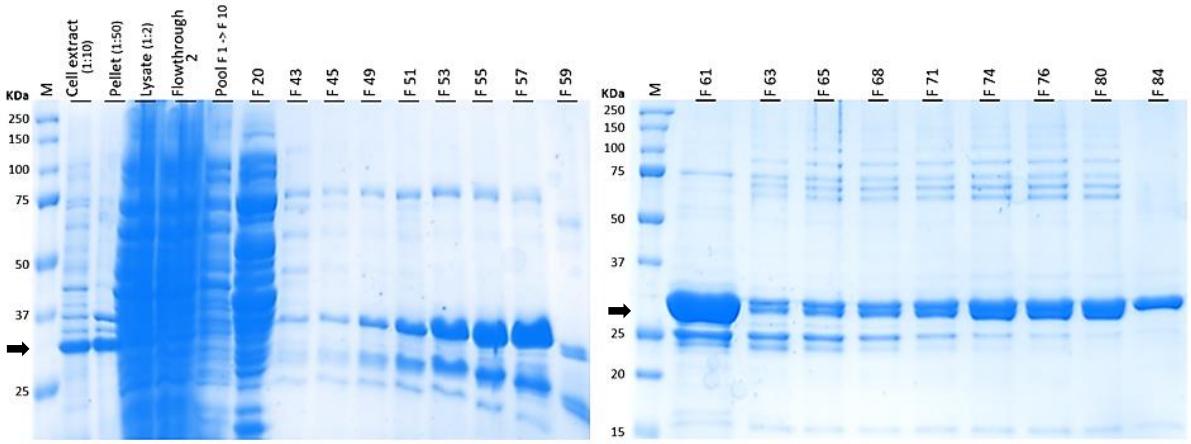


Figure B.8 – 15% SDS-PAGE after 5th attempt of the histrap purification, performed with a 5 mL histrap fastflow crude column. M - Precision Plus Protein™ Dual Color Standards. Arrows indicate the molecular weight where is expected to be the TEM-1 band. Pooled fractions: F 60 → F 86.

- 5th attempt: results of ionic exchange chromatography

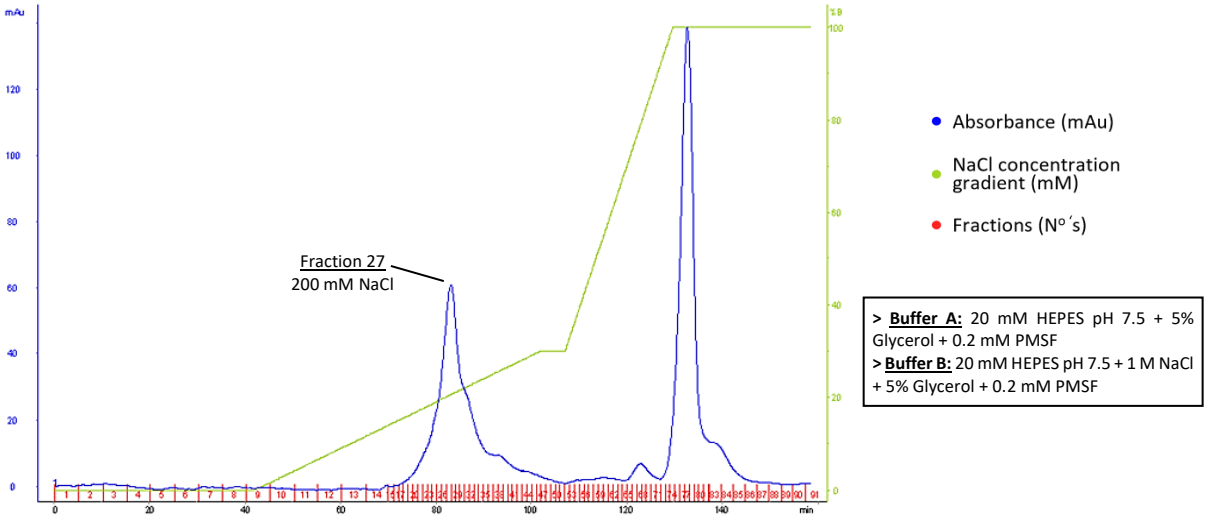


Figure B.9 – Chromatogram of 5th attempt of TEM-1 purification with a HiLoad 16/10 Q-Sepharose High Performance column. Volume of injected sample (pool made after the previous histrap purification) was: 5 mL sample + 5 mL buffer A. Purification was done with linear gradients of sodium chloride (NaCl) concentration (0 → 1 molar). First, the column was washed with buffer A for 20 minutes and then was done a 1st gradient up to 300 mM [NaCl] for 1 hour. After that, a 2nd gradient between 300 mM and 1 M [NaCl] was performed for 25 minutes. Flow rate was 2 mL/min. Main peak is highlighted with the NaCl concentration and fraction number.

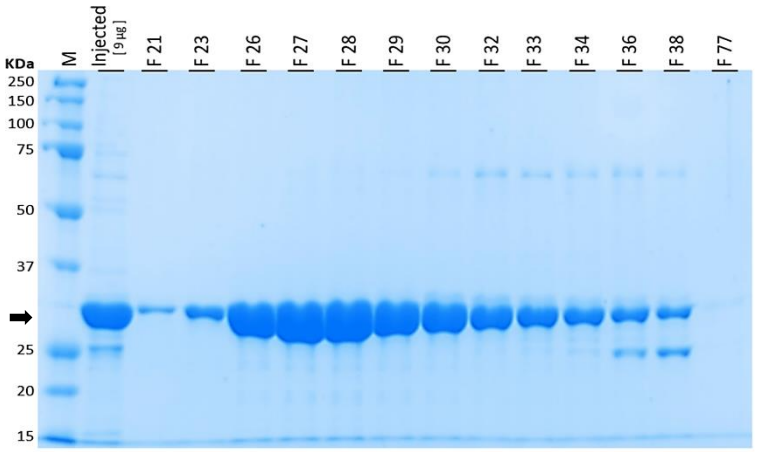


Figure B.10 – 15% SDS-PAGE of the results from ionic exchange purification after 5th attempt of protein purification.
M - Precision Plus Protein™ Dual Color Standards. Arrow indicates the molecular weight where is expected to be TEM-1 band.
Pooled fractions: **Pool +** (F 18→F 30); **Pool -** (F 31→F 34).

- ♦ 5th attempt: results of size exclusion chromatography (SEC)

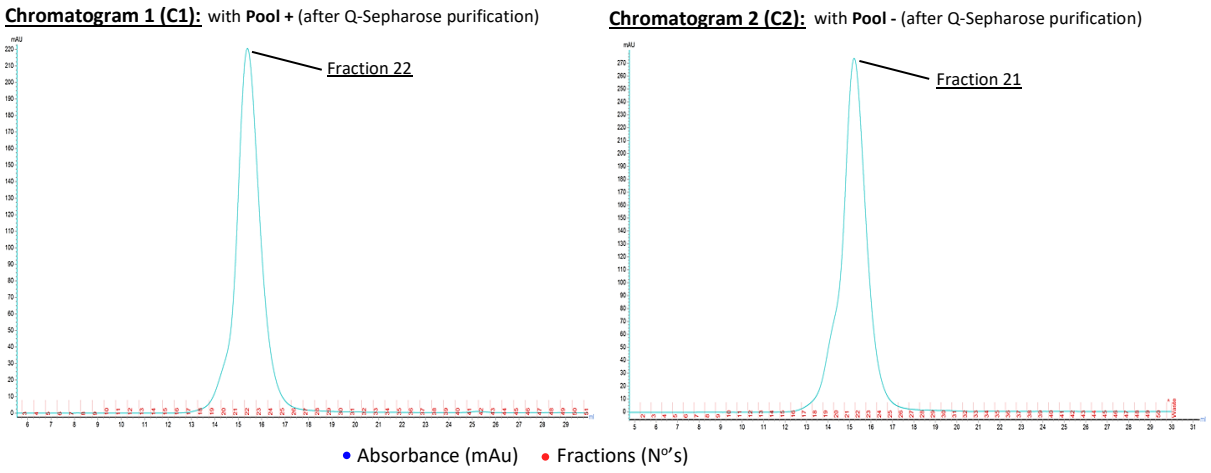


Figure B.11 – Chromatograms of the 5th attempt of TEM-1 purification, after two SEC's with a Superose 12 10/300 GL column and the buffer used in both was 20 mM HEPES pH 7.5, 150 mM NaCl, 5% glycerol and 0.2 mM PMSF. For **Pool +** the injected volume was 500 µL with a final concentration of 9 mg/mL, and the **Pool –** was injected with 6.8 mg/mL and a volume of 1 mL. Flow rate was 0.5 mL/min. Main peaks are highlighted with the identification of the fraction.

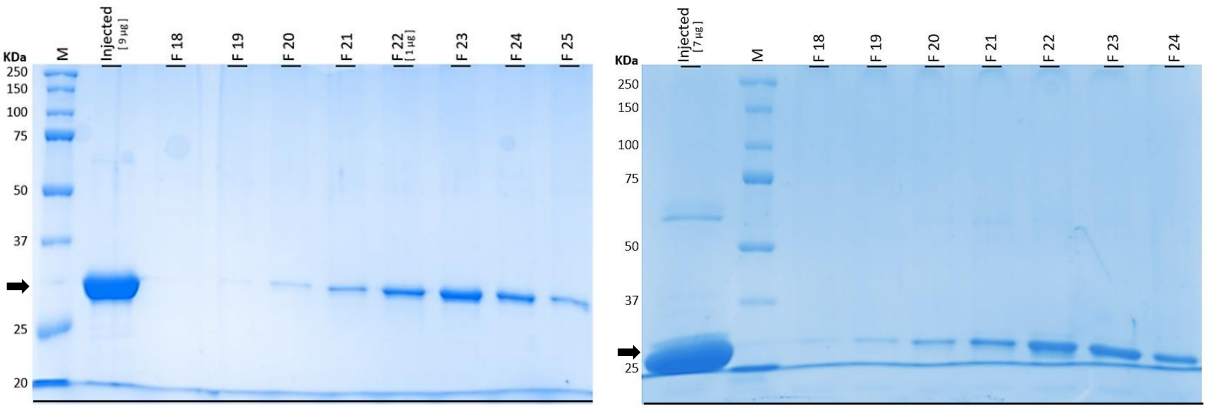
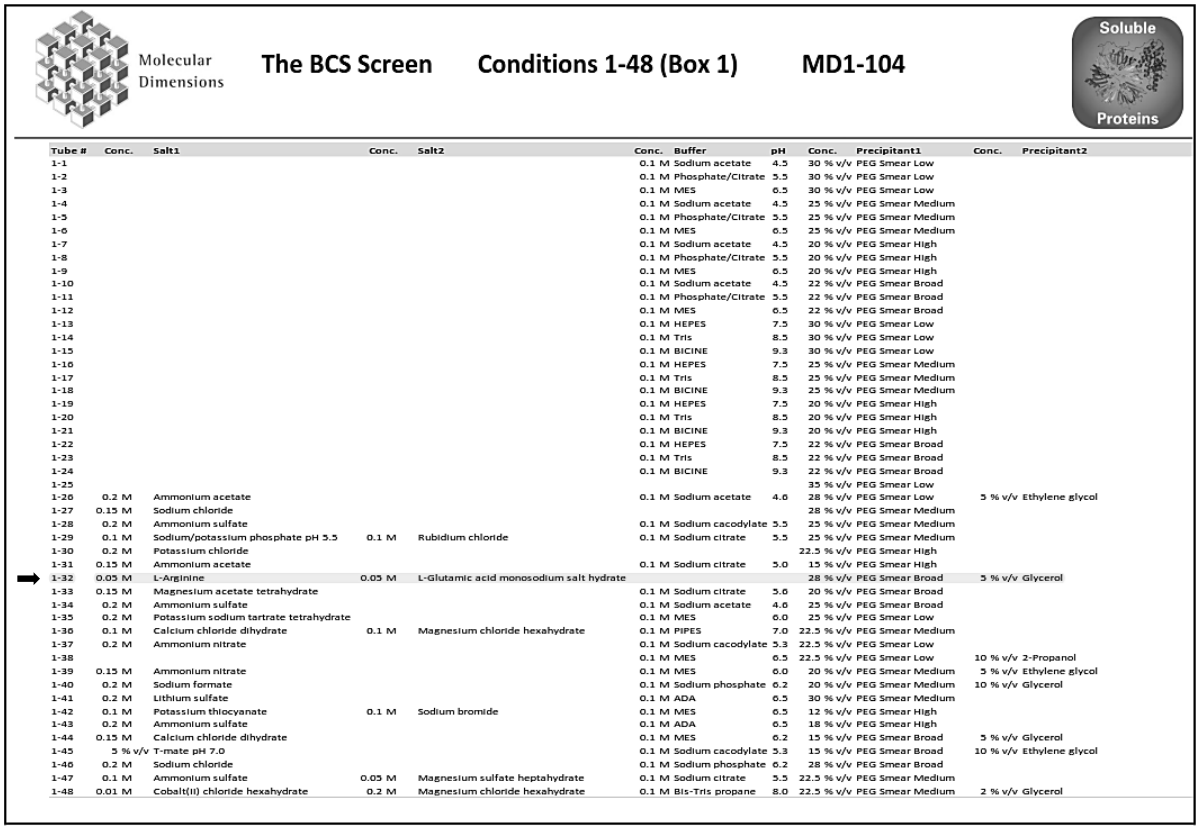


Figure B.12 – 12% SDS-PAGE gels with the results from the two SEC's performed in the 5th purification attempt. M - Precision Plus Protein™ Dual Color Standards. Arrows indicate the molecular weight where is expected to be TEM-1 band.
Final pooled fractions: **Pool 1.0** (C1: F21 + F22); **Pool 2.0** (C2: F21 + F22); **Pool 2.1** (C2: F23 + F24).

Appendix C



Molecular Dimensions

The BCS Screen **Conditions 1-48 (Box 1)** **MD1-104**

Soluble Proteins

Tube #	Conc.	Salt1	Conc.	Salt2	Conc.	Buffer	pH	Conc.	Precipitant1	Conc.	Precipitant2
1-1					0.1 M	Sodium acetate	4.5	30 % v/v	PEG Smear Low		
1-2					0.1 M	Phosphate/Citrate	5.5	30 % v/v	PEG Smear Low		
1-3					0.1 M	MES	6.5	30 % v/v	PEG Smear Low		
1-4					0.1 M	Sodium acetate	4.5	25 % v/v	PEG Smear Medium		
1-5					0.1 M	Phosphate/Citrate	5.5	25 % v/v	PEG Smear Medium		
1-6					0.1 M	MES	6.5	25 % v/v	PEG Smear Medium		
1-7					0.1 M	Sodium acetate	4.5	20 % v/v	PEG Smear High		
1-8					0.1 M	Phosphate/Citrate	5.5	20 % v/v	PEG Smear High		
1-9					0.1 M	MES	6.5	20 % v/v	PEG Smear High		
1-10					0.1 M	Sodium acetate	4.5	22 % v/v	PEG Smear Broad		
1-11					0.1 M	Phosphate/Citrate	5.5	22 % v/v	PEG Smear Broad		
1-12					0.1 M	MES	6.5	22 % v/v	PEG Smear Broad		
1-13					0.1 M	HEPES	7.5	30 % v/v	PEG Smear Low		
1-14					0.1 M	Tris	8.5	30 % v/v	PEG Smear Low		
1-15					0.1 M	BICINE	9.3	30 % v/v	PEG Smear Low		
1-16					0.1 M	HEPES	7.5	25 % v/v	PEG Smear Medium		
1-17					0.1 M	Tris	8.5	25 % v/v	PEG Smear Medium		
1-18					0.1 M	BICINE	9.3	25 % v/v	PEG Smear Medium		
1-19					0.1 M	HEPES	7.5	20 % v/v	PEG Smear High		
1-20					0.1 M	Tris	8.5	20 % v/v	PEG Smear High		
1-21					0.1 M	BICINE	9.3	20 % v/v	PEG Smear High		
1-22					0.1 M	HEPES	7.5	22 % v/v	PEG Smear Broad		
1-23					0.1 M	Tris	8.5	22 % v/v	PEG Smear Broad		
1-24					0.1 M	BICINE	9.3	22 % v/v	PEG Smear Broad		
1-25								35 % v/v	PEG Smear Low		
1-26	0.2 M	Ammonium acetate			0.1 M	Sodium acetate	4.6	28 % v/v	PEG Smear Low	5 % v/v	Ethylene glycol
1-27	0.15 M	Sodium chloride						28 % v/v	PEG Smear Medium		
1-28	0.2 M	Ammonium sulfate			0.1 M	Sodium cacodylate	5.5	25 % v/v	PEG Smear Medium		
1-29	0.1 M	Sodium/potassium phosphate pH 5.5	0.1 M	Rubidium chloride	0.1 M	Sodium citrate	5.5	25 % v/v	PEG Smear Medium		
1-30	0.2 M	Potassium chloride						22.5 % v/v	PEG Smear High		
1-31	0.15 M	Ammonium acetate			0.1 M	Sodium citrate	5.0	15 % v/v	PEG Smear High		
1-32	0.05 M	L-Arginine	0.05 M	L-Glutamic acid monosodium salt hydrate				28 % v/v	PEG Smear Broad	5 % v/v	Glycerol
1-33	0.15 M	Magnesium acetate tetrahydrate			0.1 M	Sodium citrate	5.6	20 % v/v	PEG Smear Broad		
1-34	0.2 M	Ammonium sulfate			0.1 M	Sodium acetate	4.6	25 % v/v	PEG Smear Broad		
1-35	0.2 M	Potassium sodium tartrate tetrahydrate			0.1 M	MES	6.0	25 % v/v	PEG Smear Low		
1-36	0.1 M	Calcium chloride dihydrate	0.1 M	Magnesium chloride hexahydrate	0.1 M	PIPES	7.0	22.5 % v/v	PEG Smear Medium		
1-37	0.2 M	Ammonium nitrate			0.1 M	Sodium cacodylate	5.3	22.5 % v/v	PEG Smear Low		
1-38					0.1 M	MES	6.5	22.5 % v/v	PEG Smear Low	10 % v/v	2-Propanol
1-39	0.15 M	Ammonium nitrate			0.1 M	MES	6.0	20 % v/v	PEG Smear Medium	5 % v/v	Ethylene glycol
1-40	0.2 M	Sodium formate			0.1 M	Sodium phosphate	6.2	20 % v/v	PEG Smear Medium	10 % v/v	Glycerol
1-41	0.2 M	Lithium sulfate			0.1 M	ADA	6.5	30 % v/v	PEG Smear Medium		
1-42	0.1 M	Potassium thiocyanate	0.1 M	Sodium bromide	0.1 M	MES	6.5	12 % v/v	PEG Smear High		
1-43	0.2 M	Ammonium sulfate			0.1 M	ADA	6.5	18 % v/v	PEG Smear High		
1-44	0.15 M	Calcium chloride dihydrate			0.1 M	MES	6.2	15 % v/v	PEG Smear Broad	5 % v/v	Glycerol
1-45	5 % v/v	T-mate pH 7.0			0.1 M	Sodium cacodylate	5.3	15 % v/v	PEG Smear Broad	10 % v/v	Ethylene glycol
1-46	0.2 M	Sodium chloride			0.1 M	Sodium phosphate	6.2	28 % v/v	PEG Smear Broad		
1-47	0.1 M	Ammonium sulfate	0.05 M	Magnesium sulfate heptahydrate	0.1 M	Sodium citrate	5.5	22.5 % v/v	PEG Smear Medium		
1-48	0.01 M	Cobalt(II) chloride hexahydrate	0.2 M	Magnesium chloride hexahydrate	0.1 M	Bis-Tris propane	8.0	22.5 % v/v	PEG Smear Medium	2 % v/v	Glycerol

Molecular Dimensions

The BCS Screen **Conditions 1-48 (Box 2)** **MD1-104**

Soluble Proteins

Tube #	Conc.	Salt1	Conc.	Salt2	Conc.	Buffer	pH	Conc.	Precipitant1	Conc.	Precipitant2
2-1	0.08 M	Magnesium acetate tetrahydrate	0.02 M	Magnesium chloride hexahydrate	0.1 M	MES	6.5	25 % v/v	PEG Smear Low		
2-2	0.1 M	Potassium chloride			0.1 M	HEPES	7.5	18 % v/v	PEG Smear Low	5 % v/v	Ethylene glycol
2-3	0.1 M	Zinc acetate dihydrate	0.1 M	Zinc chloride	0.1 M	Bis-tris	7.5	20 % v/v	PEG Smear Medium		
2-4	0.1 M	Magnesium chloride hexahydrate	0.1 M	Potassium chloride	0.1 M	PIPES	7.0	20 % v/v	PEG Smear Medium		
2-5	0.05 M	Magnesium sulfate heptahydrate			0.1 M	HEPES	7.5	28 % v/v	PEG Smear Medium		
2-6	0.1 M	Sodium/potassium phosphate pH 7.5			0.1 M	HEPES	7.5	15 % v/v	PEG Smear High	10 % v/v	Ethylene glycol
2-7	0.1 M	Magnesium formate dihydrate	0.1 M	Rubidium chloride	0.1 M	PIPES	7.0	25 % v/v	PEG Smear High		
2-8	0.2 M	Lithium sulfate			0.1 M	HEPES	7.2	25 % v/v	PEG Smear Broad		
2-9	0.2 M	Ammonium nitrate			0.1 M	HEPES	7.5	20 % v/v	PEG Smear Broad		
2-10	0.1 M	Magnesium chloride hexahydrate	0.1 M	Rubidium chloride	0.1 M	HEPES	7.5	30 % v/v	PEG Smear Broad		
2-11	0.05 M	Magnesium chloride hexahydrate	0.05 M	Sodium citrate tribasic dihydrate	0.1 M	Bis-Tris propane	7.8	22.5 % v/v	PEG Smear High		
2-12	7 % v/v	T-mate pH 7.0			0.1 M	BICINE	9.0	22.5 % v/v	PEG Smear High	10 % v/v	Ethylene glycol
2-13	0.15 M	Sodium citrate tribasic dihydrate			0.1 M	HEPES	7.8	25 % v/v	PEG Smear Low		
2-14	0.2 M	Sodium chloride			0.1 M	Tris	8.5	28 % v/v	PEG Smear Low	5 % v/v	Glycerol
2-15	0.08 M	Sodium acetate trihydrate	0.15 M	Sodium chloride	0.1 M	Tris	8.0	15 % v/v	PEG Smear Medium		
2-16	0.1 M	Sodium chloride	0.1 M	Sodium formate	0.1 M	Bis-Tris propane	8.5	25 % v/v	PEG Smear Medium		
2-17	0.2 M	Ammonium sulfate	0.05 M	Magnesium sulfate heptahydrate	0.1 M	BICINE	9.0	20 % v/v	PEG Smear Medium		
2-18	0.2 M	Ammonium nitrate			0.1 M	Bis-Tris propane	8.5	18 % v/v	PEG Smear High		
2-19	0.2 M	Magnesium chloride hexahydrate			0.1 M	Tris	8.0	25 % v/v	PEG Smear High	10 % v/v	Glycerol
2-20	0.15 M	Ammonium acetate	0.01 M	Calcium chloride dihydrate	0.1 M	Tris	8.5	28 % v/v	PEG Smear Broad		
2-21					0.1 M	BICINE	9.0	25 % v/v	PEG Smear Broad	10 % v/v	2-Propanol
2-22	0.2 M	Ammonium sulfate			0.1 M	Tris	8.0	20 % v/v	PEG Smear Broad		
2-23	0.02 M	Magnesium sulfate heptahydrate	0.2 M	Potassium chloride	0.1 M	BICINE	8.8	22.5 % v/v	PEG Smear Broad		
2-24	0.1 M	Potassium sodium tartrate tetrahydrate			0.1 M	Sodium cacodylate	5.5	22.5 % v/v	PEG Smear Broad	10 % v/v	Ethylene glycol
2-25	0.01 M	Cobalt(II) chloride hexahydrate	0.1 M	Magnesium formate dihydrate	0.1 M	MES	6.2	14 % v/v	PEG Smear Low		
2-26	0.15 M	Lithium sulfate	0.05 M	Magnesium chloride hexahydrate	0.1 M	Bis-Tris	6.8	25 % v/v	PEG Smear Low		
2-27	0.2 M	Ammonium sulfate	0.01 M	Cadmium chloride hemi(pentahydrate)	0.1 M	HEPES	7.5	25 % v/v	PEG Smear Medium		
2-28	0.1 M	Potassium chloride	0.1 M	Magnesium chloride hexahydrate				18 % v/v	PEG Smear Medium	10 % v/v	Ethylene glycol
2-29	0.1 M	Magnesium acetate tetrahydrate			0.1 M	MES	6.5	12 % v/v	PEG Smear Medium	10 % v/v	Ethylene glycol
2-30	0.1 M	Magnesium acetate tetrahydrate	0.1 M	Potassium chloride	0.1 M	MES	6.2	12 % v/v	PEG Smear High		
2-31	0.04 M	Calcium chloride dihydrate	0.04 M	Sodium formate	0.1 M	PIPES	7.0	8 % v/v	PEG Smear High		
2-32	0.08 M	Magnesium chloride hexahydrate	0.08 M	Sodium citrate tribasic dihydrate	0.1 M	Bis-Tris	6.0	18 % v/v	PEG Smear Broad		
2-33	0.1 M	Magnesium chloride hexahydrate	0.1 M	Sodium acetate trihydrate	0.1 M	Bis-Tris	6.5	15 % v/v	PEG Smear Broad		
2-34	0.1 M	Ammonium sulfate	0.1 M	Sodium formate	0.1 M	HEPES	7.0	25 % v/v	PEG Smear Broad		
2-35	0.2 M	Sodium/potassium phosphate pH 7.5			0.1 M	HEPES	7.5	22.5 % v/v	PEG Smear Medium	10 % v/v	Glycerol
2-36	0.3 M	Sodium chloride	0.05 M	L-Arginine	0.1 M	Tris	7.5	22.5 % v/v	PEG Smear Broad	0.05 M	L-Glutamic acid monosodium salt hydrate
2-37	0.04 M	Calcium chloride dihydrate	0.04 M	Sodium formate	0.1 M	Tris	8.0	25 % v/v	PEG Smear Low		
2-38	0.1 M	Magnesium chloride hexahydrate	0.1 M	Rubidium chloride	0.1 M	PIPES	7.0	20 % v/v	PEG Smear Low		
2-39	0.2 M	Magnesium chloride hexahydrate	10 % v/v	Ethylene glycol	0.1 M	HEPES	7.5	15 % v/v	PEG Smear Medium	5 % v/v	2-Propanol
2-40	0.05 M	Ammonium acetate	0.15 M	Magnesium sulfate heptahydrate	0.1 M	HEPES	7.0	12 % v/v	PEG Smear Medium		
2-41	7 % v/v	T-mate pH 7.0			0.1 M	HEPES	7.2	20 % v/v	PEG Smear Medium		
2-42	0.1 M	Ammonium acetate	0.1 M	Zinc chloride	0.1 M	Bis-Tris	7.2	15 % v/v	PEG Smear High		
2-43	0.15 M	Lithium sulfate	0.05 M	Magnesium chloride hexahydrate	0.1 M	HEPES	7.8	20 % v/v	PEG Smear High		
2-44	0.1 M	Potassium thiocyanate	0.1 M	Sodium bromide	0.1 M	Tris	7.8	25 % v/v	PEG Smear Broad		
2-45	0.05 M	Ammonium sulfate	0.05 M	Lithium sulfate	0.1 M	Bis-Tris propane	8.5	28 % v/v	PEG Smear Broad		
2-46	0.2 M	Ammonium sulfate	0.01 M	Cadmium chloride hemi(pentahydrate)	0.1 M	PIPES	7.0	15 % v/v	PEG Smear Broad	10 % v/v	Ethylene glycol
2-47	0.2 M	Lithium sulfate	0.05 M	Zinc acetate dihydrate	0.1 M	Bis-Tris	7.5	22.5 % v/v	PEG Smear Broad		
2-48	0.08 M	Sodium bromide	0.05 M	Sodium fluoride	0.1 M	HEPES	7.8	22.5 % v/v	PEG Smear Broad	0.08 M	Sodium iodide

Figure C.1 – List of conditions of the BCS (Basic Chemical Space) Screen, from Molecular Dimensions. Available on: https://www.moleculardimensions.com/applications/upload/MD1-104%20The%20BCS%20Screen.pdf.

dragonfly

Well Volume: 160uL

	1	2	3	4	5	6
A	PEG Smear Broad: 70.40uL, 22.00%w/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 70.40uL, 22.00%w/v Glycerol: 3.20uL, 2.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 70.40uL, 22.00%w/v Glycerol: 6.40uL, 4.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 70.40uL, 22.00%w/v Glycerol: 9.60uL, 6.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 70.40uL, 22.00%w/v Glycerol: 12.80uL, 8.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 70.40uL, 22.00%w/v Glycerol: 16.00uL, 10.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M
B	PEG Smear Broad: 74.24uL, 23.20%w/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 74.24uL, 23.20%w/v Glycerol: 3.20uL, 2.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 74.24uL, 23.20%w/v Glycerol: 6.40uL, 4.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 74.24uL, 23.20%w/v Glycerol: 9.60uL, 6.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 74.24uL, 23.20%w/v Glycerol: 12.80uL, 8.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 74.24uL, 23.20%w/v Glycerol: 16.00uL, 10.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M
C	PEG Smear Broad: 78.08uL, 24.40%w/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 78.08uL, 24.40%w/v Glycerol: 3.20uL, 2.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 78.08uL, 24.40%w/v Glycerol: 6.40uL, 4.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 78.08uL, 24.40%w/v Glycerol: 9.60uL, 6.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 78.08uL, 24.40%w/v Glycerol: 12.80uL, 8.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 78.08uL, 24.40%w/v Glycerol: 16.00uL, 10.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M
D	PEG Smear Broad: 81.92uL, 25.60%w/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 81.92uL, 25.60%w/v Glycerol: 3.20uL, 2.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 81.92uL, 25.60%w/v Glycerol: 6.40uL, 4.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 81.92uL, 25.60%w/v Glycerol: 9.60uL, 6.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 81.92uL, 25.60%w/v Glycerol: 12.80uL, 8.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 81.92uL, 25.60%w/v Glycerol: 16.00uL, 10.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M
E	PEG Smear Broad: 85.76uL, 26.80%w/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 85.76uL, 26.80%w/v Glycerol: 3.20uL, 2.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 85.76uL, 26.80%w/v Glycerol: 6.40uL, 4.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 85.76uL, 26.80%w/v Glycerol: 9.60uL, 6.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 85.76uL, 26.80%w/v Glycerol: 12.80uL, 8.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 85.76uL, 26.80%w/v Glycerol: 16.00uL, 10.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M
F	PEG Smear Broad: 89.60uL, 28.00%w/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 89.60uL, 28.00%w/v Glycerol: 3.20uL, 2.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 89.60uL, 28.00%w/v Glycerol: 6.40uL, 4.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 89.60uL, 28.00%w/v Glycerol: 9.60uL, 6.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 89.60uL, 28.00%w/v Glycerol: 12.80uL, 8.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 89.60uL, 28.00%w/v Glycerol: 16.00uL, 10.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 40.00uL, 0.05M
G	PEG Smear Broad: 80.00uL, 25.00%w/v Glycerol: 8.00uL, 5.00%v/v Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 80.00uL, 25.00%w/v Glycerol: 8.00uL, 5.00%v/v L-Arginine: 3.20uL, 0.02M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 80.00uL, 25.00%w/v Glycerol: 8.00uL, 5.00%v/v L-Arginine: 6.40uL, 0.04M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 80.00uL, 25.00%w/v Glycerol: 8.00uL, 5.00%v/v L-Arginine: 9.60uL, 0.06M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 80.00uL, 25.00%w/v Glycerol: 8.00uL, 5.00%v/v L-Arginine: 12.80uL, 0.08M Glutamate: 40.00uL, 0.05M	PEG Smear Broad: 80.00uL, 25.00%w/v Glycerol: 8.00uL, 5.00%v/v L-Arginine: 16.00uL, 0.10M Glutamate: 40.00uL, 0.05M
H	PEG Smear Broad: 80.00uL, 25.00%w/v Glycerol: 8.00uL, 5.00%v/v L-Arginine: 8.00uL, 0.05M	PEG Smear Broad: 80.00uL, 25.00%w/v Glycerol: 8.00uL, 5.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 9.60uL, 0.01M	PEG Smear Broad: 80.00uL, 25.00%w/v Glycerol: 8.00uL, 5.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 19.20uL, 0.02M	PEG Smear Broad: 80.00uL, 25.00%w/v Glycerol: 8.00uL, 5.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 28.80uL, 0.04M	PEG Smear Broad: 80.00uL, 25.00%w/v Glycerol: 8.00uL, 5.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 38.40uL, 0.05M	PEG Smear Broad: 80.00uL, 25.00%w/v Glycerol: 8.00uL, 5.00%v/v L-Arginine: 8.00uL, 0.05M Glutamate: 48.00uL, 0.06M

% PEG Smear Broad

#1

#2

1 [L-Arginine] = 0.0 – 0.1 M

2 [Glutamate] = 0.0 – 0.6 M

Figure C.2 – Matrix of optimization for TEM-1: from one hit in BCS Screen, condition #1-32.

Molecular Dimensions

Soluble Proteins

PACT *premier*™ Conditions 1-48 (Box 1) MD1-29

Tube #	Conc.	Salt	Conc.	Buffer	pH	Conc.	Precipitant
1-1			0.1 M	SPG	4.0	25 % w/v	PEG 1500
1-2			0.1 M	SPG	5.0	25 % w/v	PEG 1500
1-3			0.1 M	SPG	6.0	25 % w/v	PEG 1500
1-4			0.1 M	SPG	7.0	25 % w/v	PEG 1500
1-5			0.1 M	SPG	8.0	25 % w/v	PEG 1500
1-6			0.1 M	SPG	9.0	25 % w/v	PEG 1500
1-7	0.2 M	Sodium chloride	0.1 M	Sodium acetate	5.0	20 % w/v	PEG 6000
1-8	0.2 M	Ammonium chloride	0.1 M	Sodium acetate	5.0	20 % w/v	PEG 6000
1-9	0.2 M	Lithium chloride	0.1 M	Sodium acetate	5.0	20 % w/v	PEG 6000
1-10	0.2 M	Magnesium chloride hexahydrate	0.1 M	Sodium acetate	5.0	20 % w/v	PEG 6000
1-11	0.2 M	Calcium chloride dihydrate	0.1 M	Sodium acetate	5.0	20 % w/v	PEG 6000
1-12	0.01 M	Zinc chloride	0.1 M	Sodium acetate	5.0	20 % w/v	PEG 6000
1-13			0.1 M	MB	4.0	25 % w/v	PEG 1500
1-14			0.1 M	MB	5.0	25 % w/v	PEG 1500
1-15			0.1 M	MB	6.0	25 % w/v	PEG 1500
1-16			0.1 M	MB	7.0	25 % w/v	PEG 1500
1-17			0.1 M	MB	8.0	25 % w/v	PEG 1500
1-18			0.1 M	MB	9.0	25 % w/v	PEG 1500
1-19	0.2 M	Sodium chloride	0.1 M	MES	6.0	20 % w/v	PEG 6000
1-20	0.2 M	Ammonium chloride	0.1 M	MES	6.0	20 % w/v	PEG 6000
1-21	0.2 M	Lithium chloride	0.1 M	MES	6.0	20 % w/v	PEG 6000
1-22	0.2 M	Magnesium chloride hexahydrate	0.1 M	MES	6.0	20 % w/v	PEG 6000
1-23	0.2 M	Calcium chloride dihydrate	0.1 M	MES	6.0	20 % w/v	PEG 6000
1-24	0.01 M	Zinc chloride	0.1 M	MES	6.0	20 % w/v	PEG 6000
1-25			0.1 M	PCTP	4.0	25 % w/v	PEG 1500
1-26			0.1 M	PCTP	5.0	25 % w/v	PEG 1500
1-27			0.1 M	PCTP	6.0	25 % w/v	PEG 1500
1-28			0.1 M	PCTP	7.0	25 % w/v	PEG 1500
1-29			0.1 M	PCTP	8.0	25 % w/v	PEG 1500
1-30			0.1 M	PCTP	9.0	25 % w/v	PEG 1500
1-31	0.2 M	Sodium chloride	0.1 M	HEPES	7.0	20 % w/v	PEG 6000
1-32	0.2 M	Ammonium chloride	0.1 M	HEPES	7.0	20 % w/v	PEG 6000
1-33	0.2 M	Lithium chloride	0.1 M	HEPES	7.0	20 % w/v	PEG 6000
1-34	0.2 M	Magnesium chloride hexahydrate	0.1 M	HEPES	7.0	20 % w/v	PEG 6000
1-35	0.2 M	Calcium chloride dihydrate	0.1 M	HEPES	7.0	20 % w/v	PEG 6000
1-36	0.01 M	Zinc chloride	0.1 M	HEPES	7.0	20 % w/v	PEG 6000
1-37			0.1 M	MMT	4.0	25 % w/v	PEG 1500
1-38			0.1 M	MMT	5.0	25 % w/v	PEG 1500
1-39			0.1 M	MMT	6.0	25 % w/v	PEG 1500
1-40			0.1 M	MMT	7.0	25 % w/v	PEG 1500
1-41			0.1 M	MMT	8.0	25 % w/v	PEG 1500
1-42			0.1 M	MMT	9.0	25 % w/v	PEG 1500
1-43	0.2 M	Sodium chloride	0.1 M	Tris	8.0	20 % w/v	PEG 6000
1-44	0.2 M	Ammonium chloride	0.1 M	Tris	8.0	20 % w/v	PEG 6000
1-45	0.2 M	Lithium chloride	0.1 M	Tris	8.0	20 % w/v	PEG 6000
1-46	0.2 M	Magnesium chloride hexahydrate	0.1 M	Tris	8.0	20 % w/v	PEG 6000
1-47	0.2 M	Calcium chloride dihydrate	0.1 M	Tris	8.0	20 % w/v	PEG 6000
1-48	0.002 M	Zinc chloride	0.1 M	Tris	8.0	20 % w/v	PEG 6000

3

Molecular Dimensions

Soluble Proteins

PACT *premier*™ Conditions 1-48 (Box 2) MD1-29

Tube #	Conc.	Salt	Conc.	Buffer	pH	Conc.	Precipitant
2-1	0.2 M	Sodium fluoride				20 % w/v	PEG 3350
2-2	0.2 M	Sodium bromide				20 % w/v	PEG 3350
2-3	0.2 M	Sodium iodide				20 % w/v	PEG 3350
2-4	0.2 M	Potassium thiocyanate				20 % w/v	PEG 3350
2-5	0.2 M	Sodium nitrate				20 % w/v	PEG 3350
2-6	0.2 M	Sodium formate				20 % w/v	PEG 3350
2-7	0.2 M	Sodium acetate trihydrate				20 % w/v	PEG 3350
2-8	0.2 M	Sodium sulfate				20 % w/v	PEG 3350
2-9	0.2 M	Potassium sodium tartrate tetrahydrate				20 % w/v	PEG 3350
2-10	0.02 M	Sodium/potassium phosphate				20 % w/v	PEG 3350
2-11	0.2 M	Sodium citrate tribasic dihydrate				20 % w/v	PEG 3350
2-12	0.2 M	Sodium malonate dibasic monohydrate				20 % w/v	PEG 3350
2-13	0.2 M	Sodium fluoride	0.1 M	Bis-Tris propane	6.5	20 % w/v	PEG 3350
2-14	0.2 M	Sodium bromide	0.1 M	Bis-Tris propane	6.5	20 % w/v	PEG 3350
2-15	0.2 M	Sodium iodide	0.1 M	Bis-Tris propane	6.5	20 % w/v	PEG 3350
2-16	0.2 M	Potassium thiocyanate	0.1 M	Bis-Tris propane	6.5	20 % w/v	PEG 3350
2-17	0.2 M	Sodium nitrate	0.1 M	Bis-Tris propane	6.5	20 % w/v	PEG 3350
2-18	0.2 M	Sodium formate	0.1 M	Bis-Tris propane	6.5	20 % w/v	PEG 3350
2-19	0.2 M	Sodium acetate trihydrate	0.1 M	Bis-Tris propane	6.5	20 % w/v	PEG 3350
2-20	0.2 M	Sodium sulfate	0.1 M	Bis-Tris propane	6.5	20 % w/v	PEG 3350
2-21	0.2 M	Potassium sodium tartrate tetrahydrate	0.1 M	Bis-Tris propane	6.5	20 % w/v	PEG 3350
2-22	0.02 M	Sodium/potassium phosphate	0.1 M	Bis-Tris propane	6.5	20 % w/v	PEG 3350
2-23	0.2 M	Sodium citrate tribasic dihydrate	0.1 M	Bis-Tris propane	6.5	20 % w/v	PEG 3350
2-24	0.2 M	Sodium malonate dibasic monohydrate	0.1 M	Bis-Tris propane	6.5	20 % w/v	PEG 3350
2-25	0.2 M	Sodium fluoride	0.1 M	Bis-Tris propane	7.5	20 % w/v	PEG 3350
2-26	0.2 M	Sodium bromide	0.1 M	Bis-Tris propane	7.5	20 % w/v	PEG 3350
2-27	0.2 M	Sodium iodide	0.1 M	Bis-Tris propane	7.5	20 % w/v	PEG 3350
2-28	0.2 M	Potassium thiocyanate	0.1 M	Bis-Tris propane	7.5	20 % w/v	PEG 3350
2-29	0.2 M	Sodium nitrate	0.1 M	Bis-Tris propane	7.5	20 % w/v	PEG 3350
2-30	0.2 M	Sodium formate	0.1 M	Bis-Tris propane	7.5	20 % w/v	PEG 3350
2-31	0.2 M	Sodium acetate trihydrate	0.1 M	Bis-Tris propane	7.5	20 % w/v	PEG 3350
2-32	0.2 M	Sodium sulfate	0.1 M	Bis-Tris propane	7.5	20 % w/v	PEG 3350
2-33	0.2 M	Potassium sodium tartrate tetrahydrate	0.1 M	Bis-Tris propane	7.5	20 % w/v	PEG 3350
2-34	0.02 M	Sodium/potassium phosphate	0.1 M	Bis-Tris propane	7.5	20 % w/v	PEG 3350
2-35	0.2 M	Sodium citrate tribasic dihydrate	0.1 M	Bis-Tris propane	7.5	20 % w/v	PEG 3350
2-36	0.2 M	Sodium malonate dibasic monohydrate	0.1 M	Bis-Tris propane	7.5	20 % w/v	PEG 3350
2-37	0.2 M	Sodium fluoride	0.1 M	Bis-Tris propane	8.5	20 % w/v	PEG 3350
2-38	0.2 M	Sodium bromide	0.1 M	Bis-Tris propane	8.5	20 % w/v	PEG 3350
2-39	0.2 M	Sodium iodide	0.1 M	Bis-Tris propane	8.5	20 % w/v	PEG 3350
2-40	0.2 M	Potassium thiocyanate	0.1 M	Bis-Tris propane	8.5	20 % w/v	PEG 3350
2-41	0.2 M	Sodium nitrate	0.1 M	Bis-Tris propane	8.5	20 % w/v	PEG 3350
2-42	0.2 M	Sodium formate	0.1 M	Bis-Tris propane	8.5	20 % w/v	PEG 3350
2-43	0.2 M	Sodium acetate trihydrate	0.1 M	Bis-Tris propane	8.5	20 % w/v	PEG 3350
2-44	0.2 M	Sodium sulfate	0.1 M	Bis-Tris propane	8.5	20 % w/v	PEG 3350
2-45	0.2 M	Potassium sodium tartrate tetrahydrate	0.1 M	Bis-Tris propane	8.5	20 % w/v	PEG 3350
2-46	0.02 M	Sodium/potassium phosphate	0.1 M	Bis-Tris propane	8.5	20 % w/v	PEG 3350
2-47	0.2 M	Sodium citrate tribasic dihydrate	0.1 M	Bis-Tris propane	8.5	20 % w/v	PEG 3350
2-48	0.2 M	Sodium malonate dibasic monohydrate	0.1 M	Bis-Tris propane	8.5	20 % w/v	PEG 3350

4

Figure C.3 – List of conditions of PACT premier™ screen (a pH, Anion, Cation crystallization screen), from *Molecular Dimensions*. Available on: https://www.moleculardimensions.com/applications/upload/MD1-29%20PACT%20premier%20v2.pdf

Appendix D

Table D.1 – Summary of Protein BLAST with TEM β-lactamases (available on https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins) with data filtered by Query cover ≥ 99%. Database chosen was Protein Data Bank proteins (PDB). Made on February 20th, 2018 and reviewed on August 20th, 2018. To all the entries, the expression system was E. coli and none of them have the signal peptide. Order of results: First is the native structures, and then the complexes. Inside each one, it is ordered by year, from the oldest to the most recent.

				Crystal growth procedure							Protein				
	Year	PDB entry	TEM family (Mutant or wild type)	Precipitant, Buffer, pH	T (ºC)	Method	Ratio	Crystallization time	Space group	Unit cell length (Å)	Source	Concentration (mg/mL)	Buffer	Sequence length (# Residues)	Identity (%)
Native	1995	1ESU	TEM-1 (S235A mutant)	Imidazole 0.1 M pH 7.0, Ammonium sulfate 43-48% + 10 mM Sodium azide	19	Hanging drop; Microseeding, w/ crystals of the TEM1 wt (microcrystals obtained from seeding were used to repeat experiments to dilute the effect of heterogenous seeds)	-	1-4 weeks	P 21 21 21	a = 41.93; b = 63.24; c = 88.75	E. coli	20	-	263	262/263 (99%)
	1999	1CK3	TEM-1 (N276D mutant)	100 mM sodium-potassium phosphate buffer, 42% Ammonium sulfate + 4% Acetone, pH 7.8	4	Vapor diffusion; Seeding w/microcrystal of wt TEM-1, washed in 47% saturated ammonium sulfate + 100 mM Na KPi pH 7.8	2,5 uL [P]	4 days	P 21 21 21	a = 41.82; b = 60.36; c = 88.65	E. coli	11.6	70 mM Na KPi buffer pH 7.8 + 8.3% Ammonium sulfate		260/263 (99%)
	2001	1HTZ	TEM-52 (wt)	16% PEG 8000, 125 mM MgOAc + 25mM Na Citrate, pH 6.5	20	Hanging drop	-	3 days	P 43 21 2	a = 88.38; b = 88.38; c = 500.39	Klebsiella pneumoniae	14	-	263	260/263 (99%)
	2002	1JWP	TEM-1 (M182T mutant)	1.5 M sodium-potassium phosphate buffer pH 8.0	22	Hanging drop (8-uL droplet, seeded with microcrystals of TEM mutant M182T)	-	-	P 21 21 21	a = 41.47; b = 61.69; c = 89.21	E. coli	-			262/263 (99%)
		1LHY	TEM-30 (N132A mutant)	Sodium-Potassium Phosphate buffer (Na KPi) 1.4 M, pH 8.0						a = 41.53; b = 59.67; c = 88.1		5	0.65 M Na KPi buffer pH 8.0		262/263 (99%)
		1LI0	TEM-32 (2 mutations)							a = 41.22; b = 60.5; c = 88.71		6	200 mM potassium phosphate pH 7.0		262/263 (99%)
		1LI9	TEM-34 (1 mutation)							a = 41.37; b = 59.6; c = 88.29					262/263 (99%)
	2005	1YT4	TEM-76 (S130G mutant)	1.4 M Sodium-potassium phosphate buffer pH 8.3	23	Hanging drop (Seeding with microcrystals of TEM M182T)	2:2	5 weeks	P 21 21 21	a = 41.6; b = 59.71; c = 88.36		20	-		262/263 (99%)
		1ZG4	TEM-1 (wt)	30% (w/v) PEG 4000, 0.2 M sodium acetate + 0.1 M Tris–HCl pH 5	20	Hanging drop; After 2 weeks the drops were streakseeded using a microcrystal	1:1	1 week		a = 42.02; b = 62.84; c = 89.46		30	10 mM Tris buffer pH 7.0	286	261/263 (99%)
		1ZG6	TEM-1 (S70G mutant)							a = 41.99; b = 62.65; c = 89.49					260/263 (99%)
	2008	3CMZ	TEM-1 (L201P mutant)	1.6 M potassium phosphate pH 8.7		Hanging drop	-	-		a = 41.22; b = 59.14; c = 87.8		-	-	263	262/263 (99%)
		3DTM	TEM-1 (9 mutations)	5% PEG 400, 0.1 M Tris + 1.6-1.8 M ammonium citrate, pH 7.5	15	Sitting drop	1:1			a = 39.82; b = 79.68; c = 87.07					254/263 (97%)
	2009	3JYI	TEM-1 (N170G mutant)	20% PEG 6000, 0.1 M HEPES + 0.2 M Lithium chloride, pH 7.0	25	Hanging drop (using TTP LabTech Mosquito)		24 hours	P 43 21 2	a = 88.1; b = 88.1; c = 500.35		30	60 mM sodium phosphate buffer pH 7.8		262/263 (99%)
	2011	2V1Z	TEM-1 (8 mutations)	25% PEG 6000, 0.1 M Bis-Tris + 0.3 M NaCl + 0.02% sodium azide; 0.1 M MES buffer + 0.2 M ammonium sulfate, pH 6.2	18	Hanging drop		A few days	P 21 21 21	a = 47.11; b = 70.61; c = 78.31		-	-	291	254/268 (94%)
	2013	4IBR	TEM-1 variant.13 (G238S/E104K double mutant)	11% (w/v) PEG 8000, 100 mM MES + 200 mM Ca(OAc)2 + 10 µM ZnCl2, pH 6.7	20	Hanging drop (drops were seeded with microcrystals to obtain high-quality diffracting crystals)	-		C 1 2 1	a = 136.24; b = 46.6; c = 39 α = γ = 90º; β = 93.87º		60	25 mM Tris pH 8.4 + 100 mM NaCl		254/263 (97%)
		4IBX	TEM-1 variant.13 (wt)	9% (w/v) PEG 8000, 100 mM MES + 200 mM Ca(OAc)2, pH 6.2			1:1		P 2 21 21	a = 39.9; b = 239; c = 159.91					256/263 (97%)
	2014	4ID4	cTEM-17m (wt)	26% PEG 4000, 100 mM Tris-HCl + 0.25 M MgCl2, pH 8.0	22	Hanging drop			P 21 21 21	a = 36.73; b = 58.72; c = 109.3		25	50 mM Tris-HCl, pH 7.0		246/263 (94%)
		4MEZ	TEM-1 (M68L/M69T double mutant)	1.5 M ammonium sulfate, 0.1 M HEPES pH 7.5			-		P 1	a = 34.33; b = 55.07; c = 77.68		-	-		261/263 (99%)
	2015	4OP5	TEM-63 (R164S mutation)	9% PEG 8000, 100 mM MES, 200 mM Ca(OAc)2, pH 8.4	20	Hanging drop; Microseeding was performed to obtain larger crystals	1:1	-	C 1 2 1	a = 153.6; b = 46.1; c = 34.5 α = γ = 90º; β = 93.3º	E. coli	60	25 mM Tris pH 8.4 + 100 mM NaCl	263	255/263 (97%)
		4OP8	TEM-94 (G238S mutation)	9% PEG 8000, 100 mM MES, 200 mM Ca(OAc)2, pH 6.2						a = 154.36; b = 47.04; c = 34.92 α = γ = 90º; β = 92.6º					255/263 (97%)
		4OPQ	TEM-94 (R164S/G238S double mutant)							a = 154.9; b = 47.2; c = 34.89 α = γ = 90º; β = 92.8º					254/263 (97%)
		4QY5	cTEM-17m (wt)	30% PEG 4000, 100 mM Tris-HCl + 0.2 M MgCl2, pH 8.5	22	Sitting drop	-		P 21 21 21	a = 36.52; b = 58.21; c = 109.5		-	-		244/263 (93%)
		4RVA	TEM-1 (W165Y/E166Y/ P167G triple mutant)	25% (w/v) PEG 3350, 0.24 M tri-sodium citrate, pH 5.5	25	Hanging drop				a = 59.04; b = 59.51; c = 86.26					259/263 (98%)
		4RX2	TEM-1 (4 mutations)	30% (w/v) PEG 6000, 0.1 M HEPES pH 6.5		Hanging Drop			P 1	a = 60.32; b = 83.3; c = 95.92					257/263 (98%)
		4RX3	TEM-1 (4 mutations)	25% (w/v) PEG 4000, 0.24 M tri-sodium citrate pH 5.5		Hanging drop: apo-crystals; Co-crystallization and soaking with ceftazidime were attempted			P 21 21 21	a = 58.92; b = 60.14; c = 88.77					257/263 (98%)
		4ZJ1	TEM-1 (V216AcrF mutant)	15% PEG 20K + 15% PEG 550MME, 0.1 M MES pH 6.5, final pH 7.8; (Soaking of the crystals) 0.1 M MES pH 6.5, 17% PEG 20K + 17% PEG550MME, with 40 mM Cephalexin solution	22	Crystal growth optimization by Hanging drop after screen with robot; Streakseeding/Macroseeding from wt crystals	1:1	1 - 7 days		a = 45.47; b = 47.24; c = 128.33		12	10 mM Tris pH 7.8 + 100 mM NaCl + 10% glycerol	297	260/263 (99%)
		4ZJ2	TEM-1 (E166N mutant)							a = 46.46; b = 47.04; c = 127.58					259/263 (98%)
		4ZJ3	TEM-1 (E166N/V216AcrF double mutant)							a = 46.29; b = 47.03; c = 128.31					258/263 (98%)
	2017	5I52	TEM-1 (I263N mutant)	20% (w/v) PEG 3350, 0.2 M sodium formate, pH 7.0	21	Sitting drop	-	-	P 1 21 1	a = 59.99; b = 83.51; c = 95.78		-	-	263	261/263 (99%)
		5IQ8	TEM-1 (A224C/G283C double mutant)	20% (w/v) PEG 3350, 0.2 M ammonium tartrate dibasic, pH 7.0						a = 60.86; b = 84.71; c = 96.34					260/263 (99%)
		5KKF	TEM-1 (I263L mutant)	20% (w/v) PEG 3350, 0.1 M BIS-TRIS pH 6.5 + 2% (v/v) Tacsimate pH 6.0		Hanging drop				a = 60.5; b = 84.03; c = 94.92					261/263 (99%)
		6B2N	TEM-1 (M182N mutant)	20% PEG 1000, 0.1 M sodium phosphate dibasic/Citric acid + 0.1 M lithium sulfate, pH 4.2	18	Screenings performed with Mosquito robot; Optimized crystals were grown by Hanging drop	1:1			a = 81.91; b = 49.66; c = 122.16		25			262/263 (99%)
Complex	1998	1BT5	TEM-1 (2 mutations)	50% Ammonium sulfate, 20 mM Sodium-potassium phosphate, pH 7.8	4	Soaking in Imipenem solution	-		P 21 21 21	a = 41.42; b = 62.01; c = 89.33					261/263 (99%)
	2000	1ERM	TEM-1 (wt)	Phosphate buffer 1.5 M, pH 8.0	25	Hanging drop; Soaking with inhibitors into the previously described native crystals	2:1	6h	P 21 21 21	a = 63 ; b = 89; c = 42		-			263/263 (100%)
		1ERO		Phosphate buffer 1.8 M, pH 8.0						a = 61 ; b = 89; c = 42					263/263 (100%)
		1ERQ								a = 63 ; b = 89; c = 42					263/263 (100%)
	2001	1JTD	TEM-1 (wt)	24% (w/v) PEG 8000, 20 mM calcium acetate + 67 mM sodium cacodylate, pH 6.4	20	Hanging drop	1:1	1 - 2 weeks	P 21 21 21	a = 47.53; b = 84.92; c = 147.69	E. coli; Streptomyces exfoliatus	10	50 mM Tris + 150 mM NaCl pH 8.0	263	262/263 (99%)
	2002	1JVJ	TEM-1 (N132A mutant)	1.4 M sodium-potassium phosphate buffer, pH 8.0	22	Hanging drop (initially seeded w/microcrystals); After, crystals were soaked with Imipenem	-	-		a = 41.3; b = 61.69; c = 89.14	E. coli	5	0.65 M NaKPi buffer pH 8.0		262/263 (99%)
		1JWV	TEM-1 (G238A mutant)	1.5 M sodium-potassium phosphate buffer pH 8.0 + 2.5 mM transition-state analog (compound 2)		Hanging drop				a = 41.37; b = 61.67; c = 89.4		-	-		262/263 (99%)
		1JWZ	TEM-64 (3 mutations)	1.5 M sodium-potassium phosphate buffer pH 8.0 + 2.5 mM transition-state analog (compound 1)						a = 72; b = 34.55; c = 105.49					262/263 (99%)
	2004	1PZO	TEM-1 (1 mutation)	1.4 M potassium phosphate pH 8.3		Hanging drop; Soaking of wt crystals with compound 1				a = 41.89; b = 61.04; c = 88.68					262/263 (99%)
		1PZP				Hanging drop; Soaking of TEM-1 M182T crystals with compound 2				a = 41.55; b = 60.67; c = 89.09					262/263 (99%)
	2005	1S0W	TEM-1 (wt)	30% (w/v) PEG 8000, 2% (v/v) Dioxane, 100 mM HEPES, pH 7.5	25	Microbatch under oil				a = 45.78; b = 125.51; c = 158.76	E. coli; Streptomyces clavuligerus	12			no information
		1XXM		10% PEG 6000, 50 mM Sodium Acetate + 0.1 M LiCl, pH 5.0						a = 45.71; b = 124.47; c = 156.95					260/263 (99%)
	2006	2B5R	TEM-1 (wt)	26% (w/v) PEG 4000, 0.1 M Tris + 0.2 M Lithium chloride, pH 8.5		Hanging drop				a = 45.5; b = 123.9; c = 156.6		20 (for the protein complex)	50 mM Tris pH 8		260/263 (99%)
	2009	3GMW	TEM-1 (wt)	20% PEG 3350, 0.2 M Potassium phosphate pH 4.7	21		1:1		P 1	a = 45.42; b = 48.76; c = 106.14 α = 103.24º; β = 91.29º; γ = 90.18º	Escherichia sp. Sflu5; Streptomyces exfoliatus	10 (of each protein)	50 mM Tris-HCl pH 8.0 + 0.05 M NaCl		260/261 (99%)
	2017	5NPO	TEM-1 (wt)	7.5% PEG 6000, 50 mM Sodium acetate + 100 mM MgCl2, pH 5.5	20	Hanging drop	-	-	P 1 21 1	a = 66.91; b = 52.73; c = 67.94	E. coli	10	-	264	262/263 (99%)