Vanessa Cristina Tavares da Silva Almeida
Licenciatura em Química Aplicada

Structural Characterization of Serine Protease Complexes with Novel Inhibitors

Dissertação para obtenção do Grau de Mestre em
Bioquímica

Orientador: Doutora Margarida Archer Frazão,
Investigadora principal, ITQB NOVA

Co-orientador: Tânia Filipa Oliveira, Post-Doctoral Fellow, ITQB NOVA
José Artur Brito, Post-Doctoral Fellow, ITQB NOVA

Júri:

Presidente: Doutor Carlos Alberto Gomes Salgueiro
Arguente: Doutor Filipe Miguel dos Santos Freire
Vogal: Doutora Margarida Archer Frazão

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Setembro 2017

Structural Characterization of Serine Protease Complexes with Novel Inhibitors

Vanessa Almeida

2017

Structural Characterization of Serine Protease Complexes with Novel Inhibitors

Acknowledgments

During this last year I received lots of support from many people, directly and indirectly, to whom I am very grateful and happy to have been able to count on them. Trying not to forget anyone, I would like to thank you:

Firstly, to Professor Margarida Archer for allowing me to be integrated into the Membrane Protein Crystallography Laboratory (MPX), the opportunity to work on this project and all the training that has provided me. Thank you also for your supervision, support, enthusiasm, availability and friendliness.

To Professor Rui Moreira and his student Luís Carvalho from Instituto de Investigação do Medicamento, Faculdade de Farmácia, Universidade de Lisboa, for allowing me to participate in their interesting research project.

To my laboratory colleagues for the help, support, companionship and good work environment. In addition, to Tânia Oliveira for her directions, explications, patience and for believing in me since the beginning. To José Brito for his constructive criticism, his help, supervision, accompaniment, availability and good disposition. To Márcia Alves, Diogo Athayde, José Rodrigues and Ana Coxixo for the patience, sympathy, good disposition, motivation and support.

To all members of ITQB's Macromolecular Crystallography unit, for good reception, knowledge sharing and help.

To my college classmates for fellowship, joy, strength and support.

To my friends for giving me strength, confidence, encouragement, support and patience.

To my boyfriend for his patience, understanding, affection, encouragement and for believing in me.

And finally, to my family (mainly mother, father and grandmother Clara) for the opportunity to hold this master's degree, for all the unconditional love, support, strength and patience, for helping me to overcome my obstacles, for always believing in me and in what I do.

To all my sincerely and deep special thanks!

Abstract

Human Neutrophil Elastase (HNE) is a serine protease responsible for cleavage of peptide bonds conferring elasticity to the connecting tissues. For this reason, this enzyme is mainly found in the lungs, arteries and ligaments [1-2]. In case of over-expression, HNE enables the appearance of some diseases, such as Chronic Obstructive Pulmonary Disease (COPD), Rheumatoid Arthritis, Psoriasis and Arteriosclerosis [3-4]. Currently, diseases affecting the respiratory tract are one of the major causes of death in the world, so HNE is a potential drug target of considerable interest [4].

Porcine Pancreatic Elastase (PPE) is commonly used as a model for HNE, sharing 37% of amino acid sequence identity [5]. According to previous studies, the catalytic serine performs a nucleophilic attack on a carbonyl group present in the inhibitors [6].

The focus of this work was the three-dimensional structure determination of elastases (PPE and HNE) in complex with inhibitors by X-ray crystallography to characterize their interactions at atomic level. The rational is to correlate structure and function and contribute to the design of more potent and specific inhibitors. These newly synthetic compounds were provided by the group of Prof. Rui Moreira, Instituto de Investigação do Medicamento, Faculdade de Farmácia, Universidade de Lisboa.

X-ray diffraction data of PPE crystals were collected at a synchrotron source and three 3D-structures of PPE in complex with inhibitors were determined at resolutions around 1.4 Å. Analysis of the electron density maps revealed that the nucleophilic attack occurred at the sulfonyl group of the inhibitors, contrary to what was initially expected (which would be in the carbonyl group).

In silico energy minimization studies of the docked ligand structure into the active site of HNE, show no relevant structural modifications of the protein structure upon ligand binding.

Finally, crystals of HNE have already been obtained and experiments are ongoing to grow complexes of HNE with various inhibitors.

Resumo

Elastase Neutrófila Humana (HNE) é uma protease de serina responsável pela clivagem das ligações peptídicas que conferem elasticidade aos tecidos de conexão. Por esta razão, esta enzima é encontrada principalmente nos pulmões, artérias e ligamentos [1-2]. Em casos de sobre-expressão, esta permite o aparecimento de algumas doenças, como Doença Pulmonar Obstrutiva Crónica (DPOC), Artrite Reumatóide, Psoríase e Arteriosclerose [3-4]. Atualmente, as doenças que afetam o trato respiratório são uma das principais causas de morte no mundo, sendo então a HNE um potencial alvo terapêutico de considerável interesse [4].

A Elastase Pancreática Suína (PPE) é normalmente usada como modelo para HNE, compartilhando 37% de identidade de sequência primária [5]. De acordo com estudos anteriores, a serina catalítica realiza um ataque nucleofílico ao grupo carbonilo presente nos inibidores [6].

O foco deste trabalho foi a determinação por cristalografia de raios-X da estrutura tridimensional de elastases (HNE e PPE) complexadas com inibidores, de modo a caracterizar as respetivas interações a nível atómico. O racional é correlacionar a estrutura com a função e contribuir para o desenho de inibidores mais fortes e mais específicos. Estes novos compostos sintéticos foram fornecidos pelo grupo do Prof. Rui Moreira, Instituto de Investigação do Medicamento, Faculdade de Farmácia, Universidade de Lisboa.

Os dados de difração de raios-X dos cristais de PPE foram recolhidos numa fonte de sincrotrão e três estruturas 3D de três complexos da PPE com inibidores foram determinadas com resoluções em torno dos 1,4 Å. A análise dos mapas de densidade eletrónica revelaram que o ataque nucleofílico ocorreu no grupo sulfonilo dos inibidores ao contrário do que era inicialmente esperado (que seria no grupo carbonilo).

A minimização de energia *in silico* da estrutura do ligando acoplado no centro ativo da HNE não mostra modificações relevantes na estrutura da proteína após a ligação do ligando.

Finalmente, já foram obtidos cristais de HNE, estando já em curso experiencias para o crescimento de cristais de complexos de HNE com vários inibidores.

Index

List of Figures

List of Tables

List of Equation

Abbreviations

Å	Angstrom
BSA	Bovine Serum Albumin
Da	Dalton
ddH_2O	Bidestilated water
DMSO	Dimethylsulfoxide
E.C.	Enzymatic Comission
EM	Electron Microscopy
ESRF	European Synchrotron Radiation Facility
HD	Hanging drop
HNE	Human Neutrophil Elastase
HPE	Human Pancreatic Elastase
IUBMB	International Union of Biochemistry and Molecular Biology
MAD	Multiple wavelength Anomalous Dispersion
MIR	Multiple Isomorphous Replacement
MPD	2-methyl-2,4-pentanediol
MR	Molecular Replacement
NMR	Nuclear Magnetic Resonance
PDB	Protein Data Bank
PEG	Polyethylene glycol
PPE	Porcine Pancreatic Elastase
rpm	Revolutions per minute
SAD	Single wavelength Anomalous Dispersion
SAXS	Small-angle X-ray Scattering
SD	Sitting drop
SDS	Sodium Dodecyl Sulfate
SIR	Single Isomorphous Replacement
Tris	Tris(hidroxymethyl)-aminomethane
3D	Three-dimensional

1. Introduction

Since the XVIII century, with the development of the knowledge on the biological cell, biochemical reactions proved to be quite important in the functioning of the living organism. Such reactions were shown to be involved in many biological processes like regulation of the cellular growth and division, energy generation and storage, catalysis, stimuli response, transport of small biomolecules, etc.

These reactions occur through the action of macromolecules known as proteins, which comprise 20-30% of the cytoplasm mass and 60-80% of the dry weight of the cellular membrane [7]. Its production starts with transcription and translation of a gene present in the cell genome, resulting in a polypeptide chain (primary structure of the protein, Figure 1.1). In order to achieve stability, the amino acids of this chain organize themselves structurally, being this three-dimensional arrangement intimately connected with the protein's function [8].

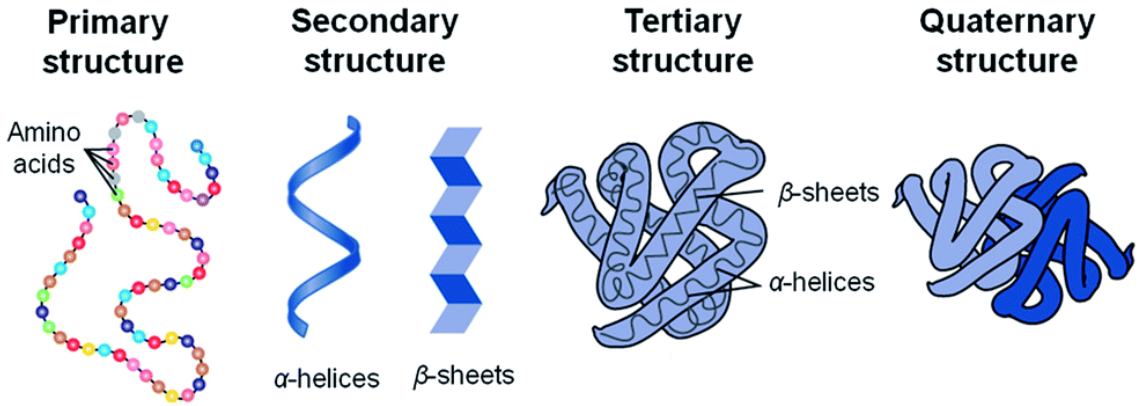


Figure 1.1.1- Structure levels of a protein [9].

Due to their great importance at the biological level, these macromolecules are often the main research target of several scientific areas (pharmaceutical, medical, agricultural, food, cosmetic and technological sciences). Lately, for the development of new products, these areas use structural biology and biochemical techniques to understand the function, activity and affinities of the target proteins [10–12]. These studies can be accomplished through some techniques, namely X-ray diffraction (macromolecular crystallography), nuclear magnetic resonance (NMR), small angle X-ray scattering (SAXS), cryo-electron microscopy (Cryo-EM), isothermal titration calorimetry (ITC), microscale thermophoresis

(MST), thermal shift assays (TSA), surface plasmon resonance (SPR) and dual polarization interferometry (DPI) [13- 15]. Typically, the most commonly used technique is X-ray diffraction which, as the name implies, results from the diffraction of incident X-ray radiation, when it interacts with the electron cloud of the atoms of a crystallized macromolecule (see details below) [15], [16].

1.1 Enzymes

As explained above, the proteins can exhibit various functions, depending on their three-dimensional structure.

Proteins that catalyze the biochemical reactions present in organisms are called enzymes. Their purpose is to accelerate the chemical reaction (by decreasing its activation energy, Figure 1.1.1), controlling its equilibrium and specifying the products that are formed, without being consumed by the reaction [17], [18].

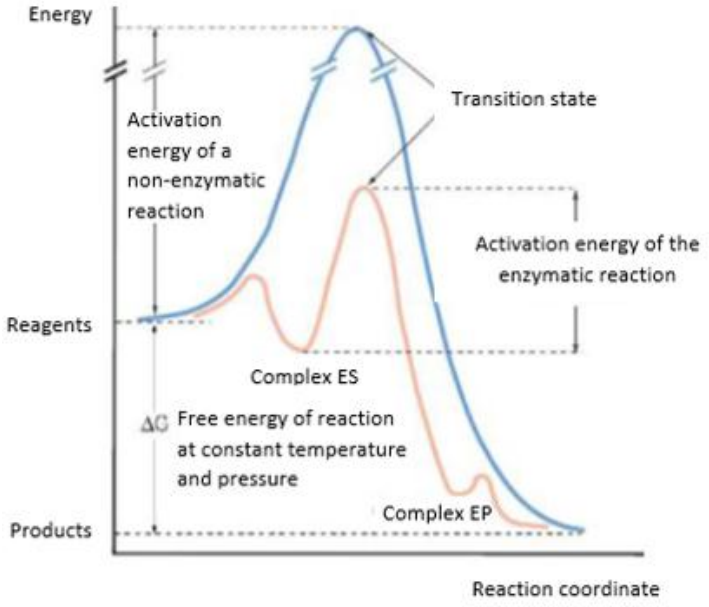


Figure 1.1.1 - Energetic profile of enzymatic and non-enzymatic reactions [17].

The blue diagram corresponds to the non-enzymatic reaction and the orange diagram to the enzymatic reaction.

The origin of its production may be intercellular [inside the cell) or extracellular (excreted into the outer environment), depending on its target. According to their mode of action, these types of proteins can be divided into two categories: Endoenzymes or Exoenzymes. Endoenzymes only cleave the chemical bonds that are found in the inner

regions of the target molecule, whereas Exoenzymes only act on chemical bonds at the extreme of the molecule of interest, generating dimers or trimers. Because of their high specificity to the substrate and the promotion of a particular reaction, each enzyme is only able to recognize and act on a specific chemical group as well as to produce a specific product without the formation of co-products [19]. The efficiency of its activity is dependent on the optimization of two major factors: Temperature and pH [17].

Initially, the name of the enzymes derived from the name of their substrate with the addition of the suffix "-ase" (eg urease, an enzyme that catalyzes the hydrolysis of urea, resulting from the name urea with the suffix "ase"). With the advancement of the enzymatic discoveries, it was necessary to create a classification system capable of differentiating even better the various types of biological enzymes. In 1961, a numerical classification system was created by the Enzymatic Commission of the International Union of Biochemistry and Molecular Biology (IUBMB). In this system, each enzyme receives a classification number (known as "E.C.") composed of four digits. The first digit refers to the number of the 6 main classes (Table 1.1.1) to which the enzyme belongs. The next two numbers are associated, respectively, with the class and subclass of the substrate. Finally, the fourth number represents the serial number of the respective enzyme [17].

Table 1.1.1 - Classes of enzymes [17].

Class	Reaction	Enzymes
1. Oxidoreductases	$A_{red} + B_{ox} \rightarrow A_{ox} + B_{red}$	Dehydrogenases, peroxidases
2. Transferases	$A - B + C \rightarrow A + B - C$	Hexokinase, transaminases
3. Hydrolases	$A - B + H_2O \rightarrow A - H + B - OH$	Alkaline phosphatase, trypsin
4. Liases (synthase)	$X - A - B - Y \rightarrow A = B + XY$	Fumarase, dehydratase
5. Isomerases	$A \rightleftharpoons isoA$	Triose phosphate isomerase, phosphoglycerone
6. Ligases (synthetases)	$A + B + ATP \rightarrow A - B + ADP + Pi$	Pyruvate Carboxylase, DNA Ligases

1.2 Proteases

Proteases, as the name infers, are enzymes that break down proteins by hydrolyzing their peptide bond. These type of enzymes can also be named as Proteolytic enzymes and its general term are coded by EC 3.4.X.X [20].

These enzymes are ubiquitous in nature because of their involvement in various physiological processes of living organisms. For this reason, they are present in a wide variety of sources (such as plants, animals and microorganisms) [20], [21].

Its action at the extracellular level essentially consists of promoting the breaking of large proteins into small molecules in order to make them more absorbable by the cell. At the intracellular level, the main goal is to regulate cell metabolism.

Like all enzymes, proteases are divided into two major groups, according to their place of action: Exopeptidases (exoenzymes) and Endopeptidases (endoenzymes).

Since exopeptidases act at the ends of proteins, they can be classified as aminopeptidases (proteases that attack the N-terminal of the protein) or carboxypeptidases (proteases that attack the C-terminal of the protein) [21].

In the aminoprotease group there are enzymes capable of removing a dipeptide or a tripeptide from the protein (EC 3.4.14) as well as enzymes capable of removing only one amino acid from the protein (EC 3.4.11).

In contrast, the carboxypeptidase group can only remove a dipeptide (EC 3.4.15) or only an amino acid from the protein. The grouping enzymes that enter into the release reactions of only one amino acid are differentiated according to the nature of their catalytic center (Serin-type EC 3.4.16; Cysteine-type EC 3.4.18 and metallocarboxypeptidase EC 3.4.17).

However, there are exopeptidases that are not specific to one end of the protein (eg: Dipeptidases EC 3.4.13 and Omega EC 3.4.99).

In the major group of endopeptidases, also known as proteinases, there is only differentiation between enzymes according to the nature of the active center of the same, composing 5 different families: Aspartic, Cysteine, Metallo, Threonine and serine endopeptidases (Table 1.2.1). Nowadays, there are still endopeptidases whose catalytic mechanism remains unknown (Unknown proteases, EC 3.4.99) [20].

There are also proteases with an unspecific activity (Mixed proteases). As the name implies, they can act as endopeptidases as well as exopeptides (eg: pronase) [22].

Table 1.2.1 - Protease classification according to their place of action [23].

Protease	Enzyme commission (EC) code	Mechanism
Aminopeptidases	3.4.11	Release N-terminal amino acid residues from polypeptides and protein
Dipeptidyl peptidase	3.4.14	Release of an N-terminal dipeptide from a polypeptide
Tripeptidyl peptidases	3.4.14	Release of an N-terminal tripeptide from a polypeptide
Peptidyl dipeptidases	3.4.15	Release of free C-terminus liberate a dipeptide
Serine-type carboxypeptidases	3.4.16	Release of a single residue C-terminal from a polypeptide and have an active center serine involved in the catalytic process
Metallocarboxypeptidases	3.4.17	Release of a single residue C-terminal from a polypeptide using a metal ion in the catalytic mechanism
Cysteine-type carboxypeptidases	3.4.18	Release of a single residue C-terminal from a polypeptide and have a cysteine in the active center
Omega peptidases	3.4.19	Remove terminal residues that are linked by isopeptide bonds
Dipeptidases	3.4.13	Exopeptidases specific for dipeptides
Aspartic endopeptidases	2.4.23	Cleave internal bonds in polypeptide chains having an aspartic acid residue for their catalytic activity

Cysteine endopeptidases	3.4.22	Cleave internal bonds in polypeptide chains. Have a cysteine in the active center
Metalloendopeptidases	3.4.24	A metal ion (often, but not always, Zn^{2+}) is involved in the catalytic mechanism for cleaving internal bonds in polypeptide chains
Threonine endopeptidases	3.4.25	Cleave internal bonds in polypeptide chains having a threonine residue for their catalytic activity
Serine endopeptidases	3.4.21	Cleave internal bonds in polypeptide chains. Have an active center serine involved in the catalytic process
Endopeptidases of unknown catalytic mechanism	3.4.99	Acting on peptide bonds (peptide hydrolases)

As there are structural similarities between proteolytic enzymes, a new classification emerged in 1993, the MEROPS classification, which takes into account the homology between proteases and their molecular structures, dividing them by families and clans (Table 1.2.2).

Each family brings together the enzymes which show homology according to a comparison of their amino acid sequence. Its MEROPS ID is initialized by a letter that represents the catalytic type (S for serine, C for cysteine, T for threonine, A for aspartic, G for glutamic, M for metallo, N for asparagine, P for mixed and U for unknown) followed of an arbitrary number.

A clan assembles the enzymes with similar three-dimensional structures (taking into account the arrangement and similarities in the amino acid sequence around the active center). The only difference in its MEROPS ID, compared to the ID of families, is that the letter that represents the active center instead of being followed by a random number is followed by a random letter [24].

Table 1.2.2 - Protease families [25].

Family	Example	Catalytic center	Favorable pH
Serine protease I	Chymotrypsin Trypsin Elastase	Asp^{102},Ser^{195},His^{57}	Neutral
Serine protease II	Subtilisin	Asp^{32},Ser^{221},His^{158}	
Cysteine protease	Papain	Cys^{25},His^{159},Asp^{158}	
Aspartic protease	Penicilopepsin Renin Chymosin	Asp^{33},Asp^{213}	Acidic
Metalloprotease I	Carboxypeptidase A bovine	Zn,Glu^{270},Try^{248}	Neutral or alkaline
Metalloprotease II	Thermolysin	Zn,Glu^{145},His^{231}	

The graphic in Figure 1.2.1 was constructed according to the total grouping of proteases identified until 2017, according to the entries in the PDB (Protein Data Bank). In the same case, serine proteases were the proteases that presented the highest number of entries in the PDB, according to the MEROPS ID classification (37%) [20].

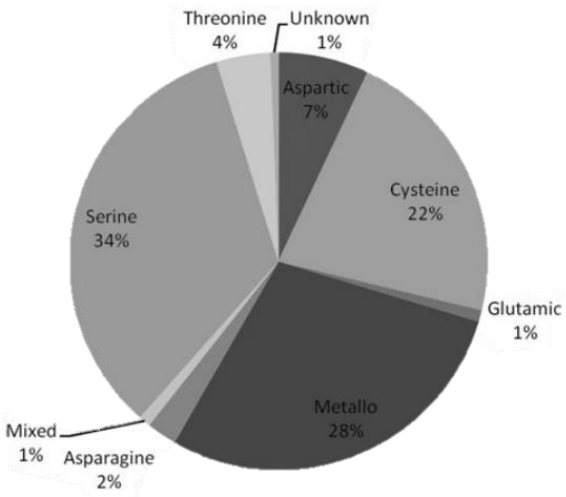


Figure 1.2.1 - Statistical results of the identification of each protease family in the total number of proteases, according to the entries of the PDB (Protein Data Bank) by 2017 [20].

1.3 Serine proteases

Serine proteases are a set of enzymes present in various biological processes, which presents an active site with an aspartic acid, a histidine and a serine (catalytic triad) [17], [20]. After several inhibition studies, it was found that, out of the three amino acid residues of the active site, serine would be the most reactive, naming this type of enzymes as serine proteases[26].

With the approach of the carbonyl group of the peptide bond to the serine of the catalytic center of the enzyme, the catalytic mechanism of the enzyme is started (Figure 1.3.1).

This mechanism is divided into two stages: acylation and deacylation.

The first stage (acylation) occurs through a nucleophilic attack of the catalytic serine oxygen on the carbon of the carbonyl group of the peptide bond. Through it, a covalent bond is formed between both elements which in turn causes the breaking of the peptide bond with the release of an amino acid or a small peptide. The final complex of this stage is called acyl-enzyme intermediate.

After the breakdown of the peptide bond, to separate the substrate from the enzyme, the second phase (deacylation) is carried out. This results from a nucleophilic attack, this time by a water molecule on carbon which at this stage is bound to the serine. In this way the hydrolysis of the peptide bond is terminated.

As can be seen in Figure 1.3.1, all reactions of the mechanism are reversible, so they can occur in both directions [20], [27].

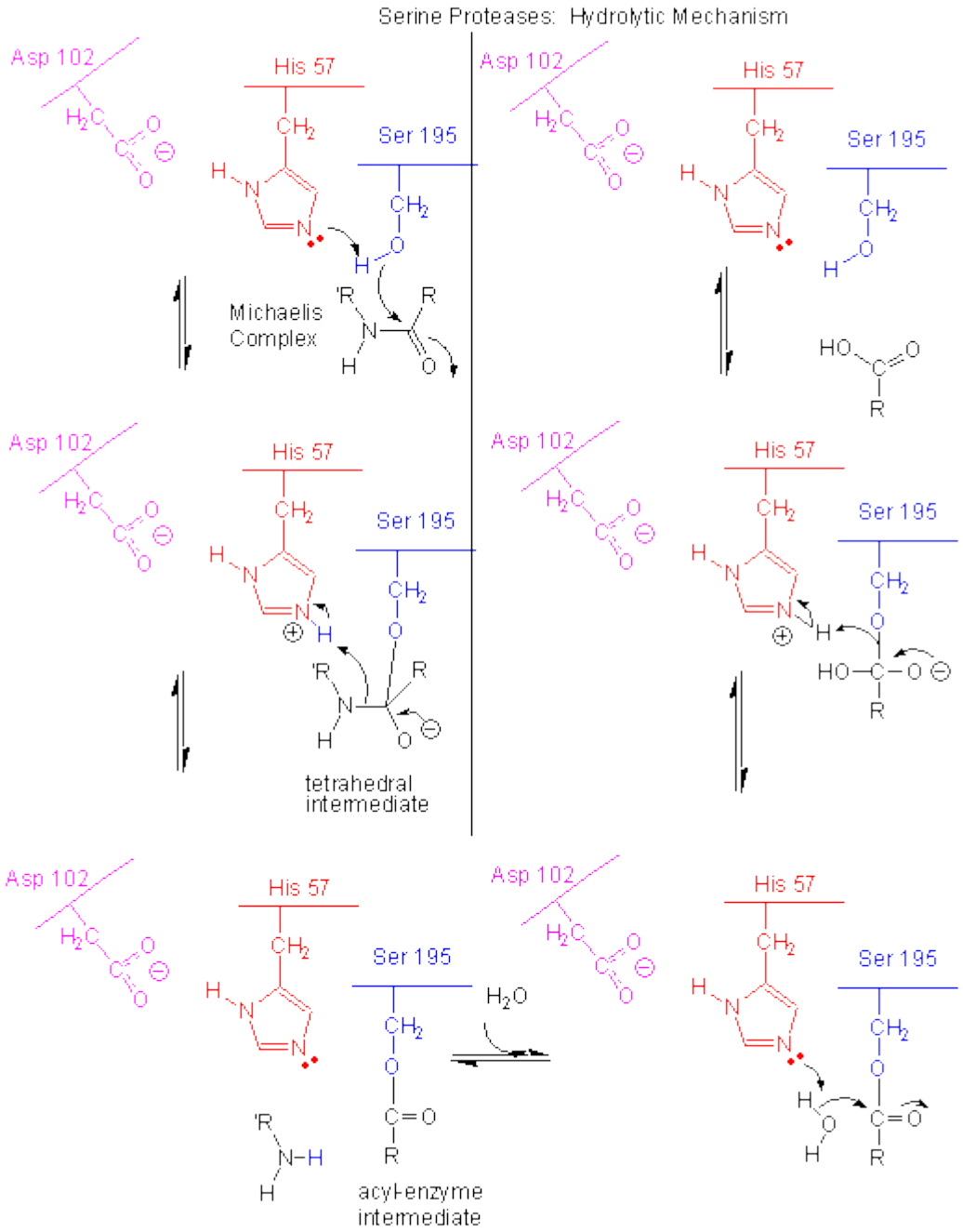


Figure 1.3.1 - Catalytic mechanism of serine proteases [27].

1.4 Elastases

Elastases are Serine proteases whose main function is the cleavage of peptide bonds of many proteins (Figure 1.4.1) such as elastin, which is responsible for the elasticity of the connective tissues, being mainly located in the lungs, arteries and ligaments [1], [2], [28].

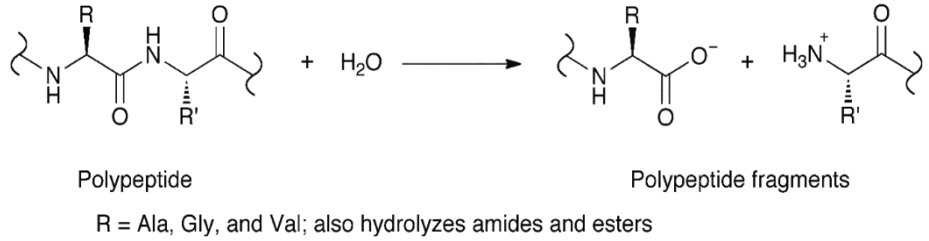


Figure 1.4.1 - Action of catalytic serine on peptide bonds [29].

Neutrophilic elastases (HNE) and pancreatic elastases (HPE) are the most abundant serine proteases in humans. As the names indicate, pancreatic elastases are stored in the pancreas in their inactive zymogenic form, being only activated by the action of trypsin when they are released into the intestine, improving the digestive process; neutrophil elastases are produced in neutrophils with the main objective of defending the organism against invasive microorganisms that cause infections, facilitating phagocytosis [28].

Under normal conditions, both elastases are well regulated through specific inhibitors present in plasma (α1-antitrypsin for HPE and α2-macroglobin for HNE). However, in cases of deregulation by over-expression, severe permanent damages are observed like liver failure, rheumatoid arthritis, psoriasis, arteriosclerosis, emphysema, cystic fibrosis and asthma [11], [26], [27]. At present, several diseases affecting the respiratory tract are one of the major causes of death in the world, so HNE is a therapeutic target with considerable interest [30].

Great efforts have been applied over the last three decades to the development of innovative elastase inhibitors. As can be seen from the various structures of elastase-inhibitor complexes present in PDB, a number of novel synthesized compounds (peptidic and non-peptidic derivatives) have been studied. However, there is currently only one non-peptidic drug available in Japan and Korea used for the treatment of acute lung injury (ALI) and adult respiratory distress syndrome (ARDS). As promising drugs, AZD9668 (Alveltat, Astra Zeneca, Cambrige, UK) and Bay 85-8501 (Bayer HealthCare, Leverkusen, Germany) are

already in Phase II of the clinical trials for patients with bronchiectasis, cystic fibrosis, COPS and lung diseases [31]. However, the synthesis of new compounds that have more specificity and more efficacy for elastase is still being studied.

Because of its difficult purification and crystallization processes, it is usual to use Porcine Pancreatic Elastase (PPE) as a model protein for HNE. PPE is a serine protease with 240 amino acids that shares 37% amino acid sequence identity (Figure 1.4.2) with HNE (218 amino acids) [5]. As explained above, since they are from the same protease family (serine proteases), they both share the same catalytic center constituted by Ser-195, His-57 and Asp-102 (catalytic triad, Figure 1.4.4). Despite the similarity, HNE is more hydrophobic and more basic than PPE (Figure 1.4.3). This is because its amino acid sequence has a greater number of hydrophobic components (> 40% compared to 30% of PPE) and its surface is constituted by a greater number of basic than acidic amino acids (19 arginines versus 9 acidic residues). One of the structural advantages of PPE is that its catalytic center is more accessible than HNE, which is surrounded by 18 arginines, making it difficult for non-linear ligands to enter [28].

```
3EST:A|PDBID|CH..      1   VVGGTEAQRNSWPSQISLQYRSGSSWAHTCGGTLIRQNWVMTAAHCVDRE--LTFRVVVG   58
3Q76:A|PDBID|CH..      1   IVGGRRARPHAWPFMVSLQLRGG----HFCGATLIAPNFVMSAAHCVANVNVRAVRVVLG   56
                           :*** .*: ::**  :*** *.*    * **.***  *:**:***** .    :.***:*

3EST:A|PDBID|CH..     59   EHNLNQNNGTEQYVGVQKIVVHPYWNTDDVAAGYDIALLRLAQSVTLNSYVQLGVLPRAG  118
3Q76:A|PDBID|CH..     57   AHNLSRREPTRQVFAVQRIFENGYD---PVNLLNDIVILQLNGSATINANVQVAQLPAQG  113
                            ***.:.: *.* ..**:*. : *     *    **.:*:*  *.*:*: **:. **  *

3EST:A|PDBID|CH..    119   TILANNSPCYITGWGLTRTNGQLAQTLQQAYLPTVDYAICSSSSYWGSTVKNSMVCAGGD  178
3Q76:A|PDBID|CH..    114   RRLGNGVQCLAMGWGLLGRNRGIASVLQELNVTVV-TSLCRR----------SNVCTLVR  162
                             *.*.  *   ****   *  :*..**:  : .*  ::*          * **:

3EST:A|PDBID|CH..    179   G-VRSGCQGDSGGPLHCLVNGQYAVHGVTSFVSRLGCNVTRKPTVFTRVSAYISWINNVI  237
3Q76:A|PDBID|CH..    163   GRQAGVCFGDSGSPLVCNG----LIHGIASFV-RGGCASGLYPDAFAPVAQFVNWIDSII  217
                           *   . * **** .** *       :**::*** * **     * .*: *: ::.**:.:*

3EST:A|PDBID|CH..    238   ASN                                                        240
3Q76:A|PDBID|CH..    218   Q--                                                        218
```

Figure 1.4.2 - Amino acid sequence alingment: PPE (PDB:3EST) vs. HNE (PDB: 3Q76) [5].

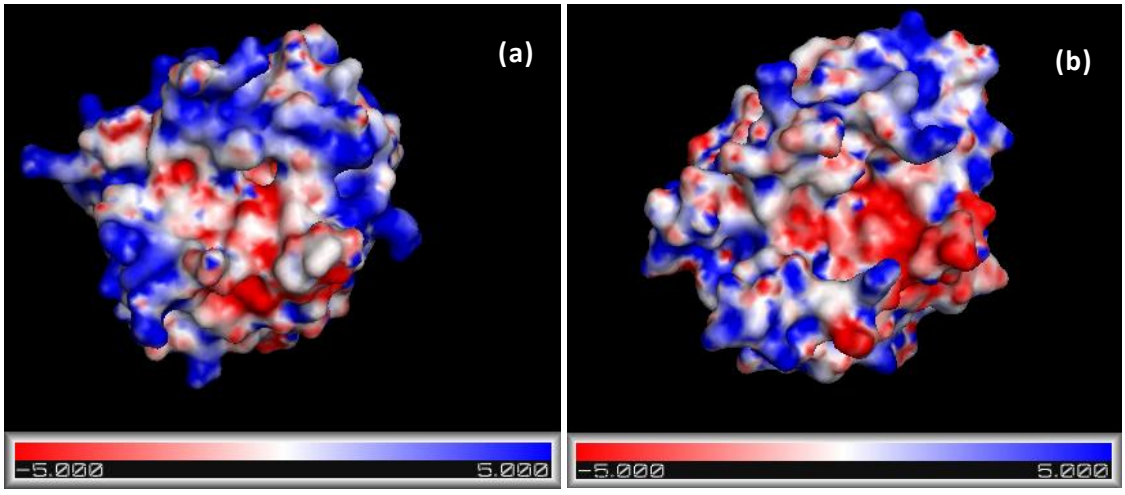


Figure 1.4.3 -Electrostatic potential map of HNE (Figure a, PDB: 3EST) vs PPE (Figure b, PDB:3EST)

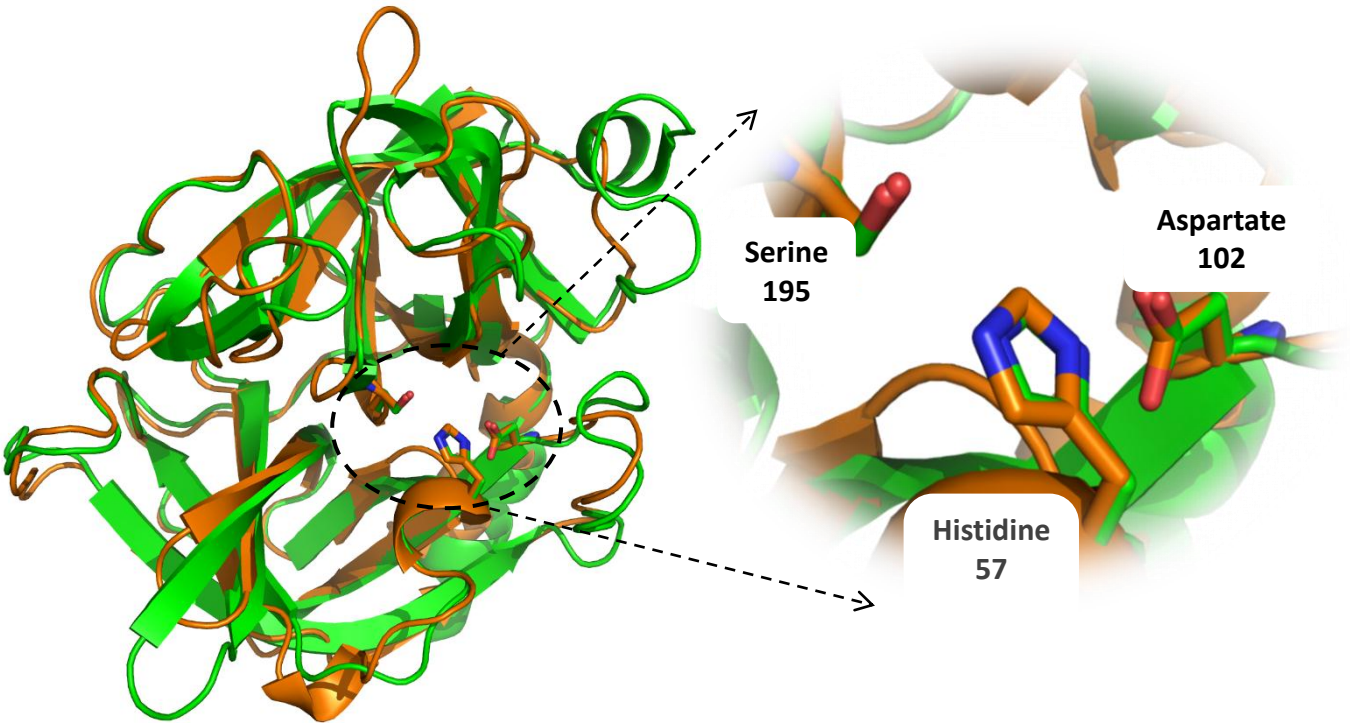


Figure 1.4.4 - Structure superposition of PPE (PDB: 3EST, green) and HNE (PDB: 3Q76, orange), with an expansion of the active sites (rmsd=1.2774Å), [32], [33].

1.5 Crystallography

X-ray diffraction analysis by macromolecular crystallography requires the presence of a vast number of molecules of the protein under study in the crystalline state in order to amplify the generated signal [15].

The crystallization of a protein is based on the slow precipitation of each individual molecule, in order to facilitate its ordering into a crystalline lattice. During this procedure, there is the formation of hydrogen bonds between the side chain of surface amino acids, and sometimes with water molecules/ions present in the solvent, keeping most of the molecules (>10^{15}) in the same orientation. By presenting 20-80% of the volume in the form of solvent channels, the crystal is shown to be a sensitive structure but, on the other hand, it allows the interaction between the protein and small molecules by the soaking technique (see details below) [15], [17], [18].

Of all the crystallization methods (vapour diffusion, microbatch, free-interface diffusion and microdialysis), vapour diffusion is the most commonly used. It consists in the equilibrium of concentrations between two solutions in a closed system, through evaporation of water from the less concentrated solution (drop) to the more concentrated solution (reservoir). In this method, there are various possible techniques having as a variant the shape of the drop. This variation will alter the drop's superficial area and tension, which will ultimately affect the equilibrium of the system and the number of nuclei formed. Hanging and sitting drop are two of the most widely used systems of the Vapor Diffusion method (Figure 1.5.1) [15].

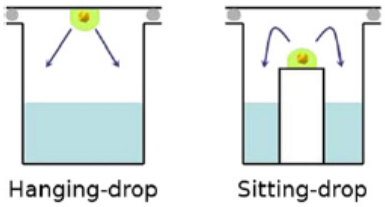


Figure 1.5.1 - Representation of the vapour diffusion hanging drop and sitting drop methods [15].

For the formation of good quality crystals (large, single and with well-defined faces), it is necessary to optimize the crystallization condition, in order to reach the Supersaturation State. This state is divided into three zones: the Metastable Zone, the Nucleation Zone and the Precipitation Zone (Figure 1.5.2). In the crystallization process, the drop needs to form the first ordered aggregates in the Nucleation Zone (also known as Labile Zone), by the increase of protein and precipitate concentration in the drop, during the evaporation. Then, by the decrease of free protein concentration in the drop solution, crystal growth occurs in the Metastable Zone [34].

At an initial stage, a crystallization screening is performed, varying the parameters of the crystallization solutions that influence the appearance of crystals like: protein concentration, precipitant type and concentration, pH and temperature. Commonly used solutions as precipitants are salts, low molecular weight alcohols and polyethylene glycol molecules (PEG) with molecular weight varying between 400 and 10000 g/mol.

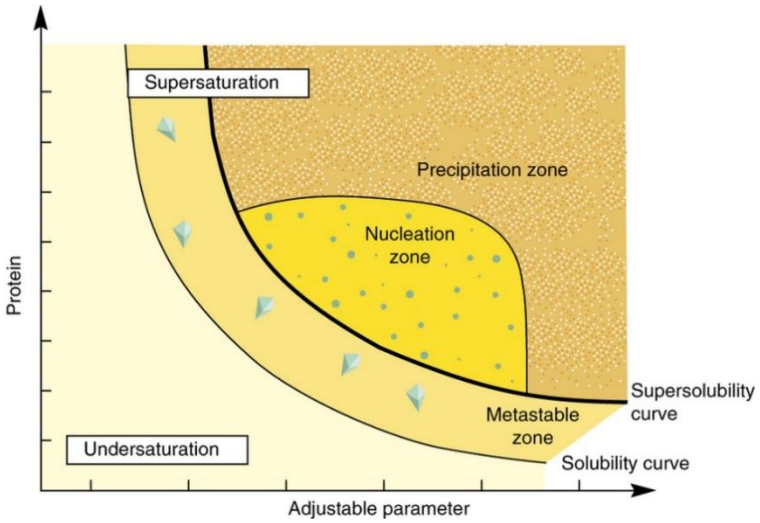


Figure 1.5.2- Phase diagram for crystallization [35].

When the crystallization conditions are not optimal, the time required for the equilibrium between concentrations will be different than ideal. Lower equilibration times tend to form amorphous precipitate, while longer equilibration times tend to form micro-crystals[36].

In cases of clear drops (where the nucleation does not occur spontaneously), it is necessary to promote nucleation by adding nuclei from protein crystals of previous assays under similar conditions, that are already in the metastable state [15]. This technique is called seeding and can be performed in three different ways:

- **Streak seeding:** Based on touching the crystal already formed with a cat mustache, passing it quickly by at least three protein:precipitant clear drops (without crystals or precipitate), promoting this way nucleation (Figure 1.5.3)[15].

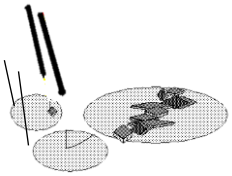


Figure 1.5.3 - Illustration of the streak seeding technique [37].

- **Microseeding:** For this technique (Figure 1.5.4), previous grown crystals are crushed in a crystallization solution, to generate a solution of microseeds. Then, this same solution, or one of its dilutions (1:10, 1:100, 1:1000 or even 1:10000), is added to a protein:precipitant clear drop (in different volume ratios) to promote nucleation.

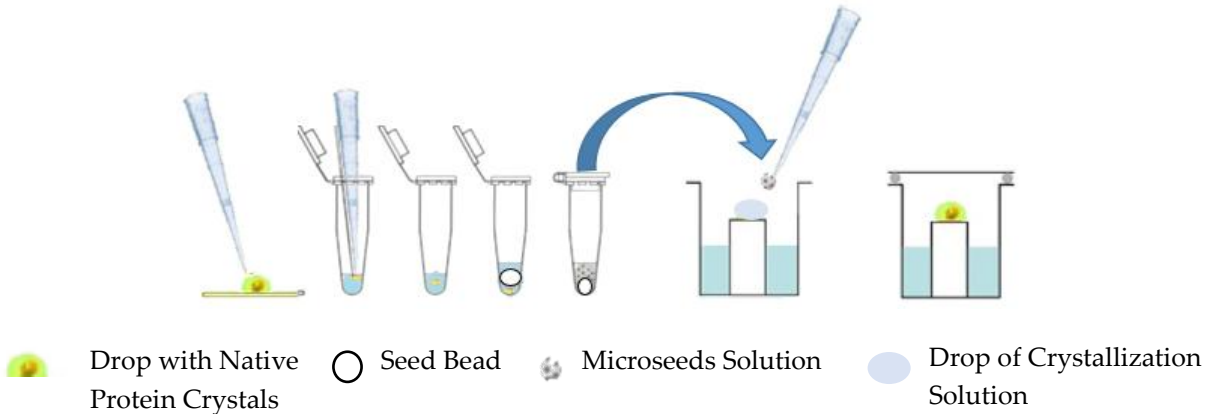


Figure 1.5.4 - Illustration of the Microseeding technique by Seed Bead®.

- **Macroseeding:** Although it is a technique to promote the nucleation of the drop, the main objective of this one is the increase of the crystal size. For that, a protein crystal is collect and directly added to a new clear drop (Figure 1.5.5).

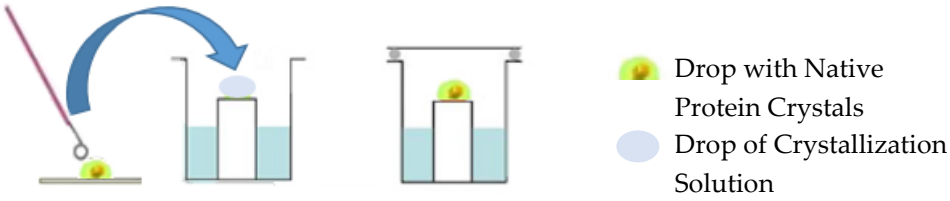


Figure 1.5.5 - Illustration of the Macroseeding technique.

After knowing the optimized crystallization conditions of a protein, it is possible to advance with studies of complexes. There are two possible techniques for the crystallization of complexes:

- **Soaking:** It is based on the diffusion of small ligands (size usually varies between 20 and 100 Å) through the solvent channels in the crystal of the native protein (Figure 1.5.6) [15].

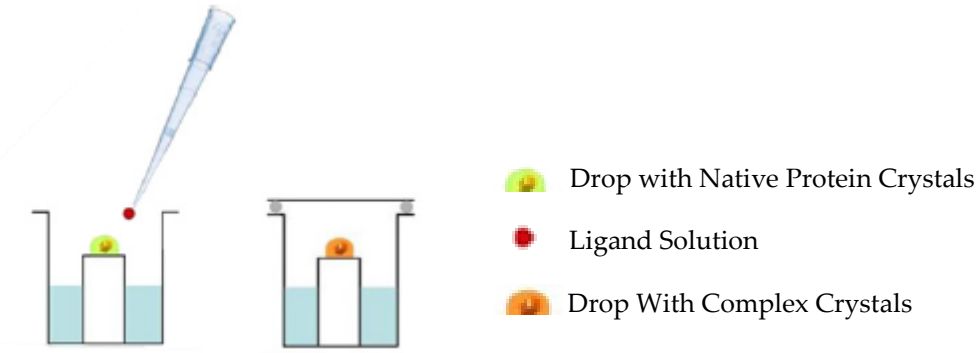


Figure 1.5.6 - Illustration of the Soaking technique.

- **Co-crystallization:** This technique requires the pre-incubation of the protein with the ligand, in order to promote the formation of the complex prior to crystal formation (Figure 1.5.7) [15]. This technique is most useful when the ligand of interest is too big to diffuse through the solvent channels and its binding promotes structural arrangements that might disrupt the crystals lattice or access to the active site is limited by crystals contacts.

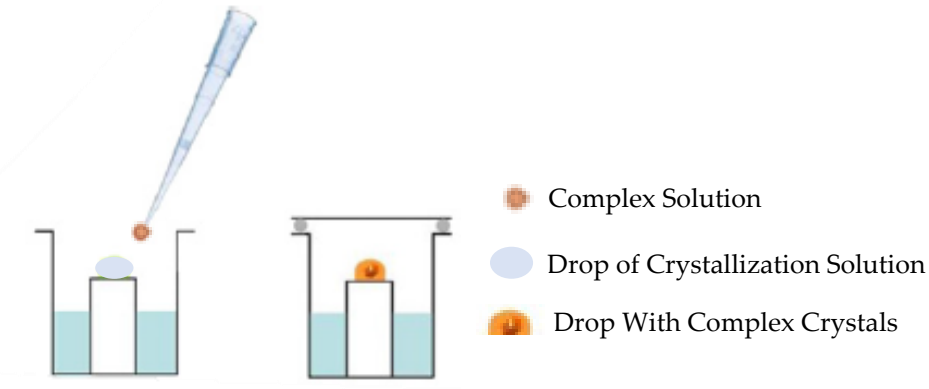


Figure 1.5.7 - Illustration of the Co-crystallization technique.

With the growth of a crystal from the macromolecule of interest, one essential step for its structural analysis by X-ray diffraction is overcome.

The drying of the crystal causes the loss of the three-dimensional structure of the macromolecule, which in turn affects its diffraction. There are several ways to prevent crystal drying by keeping it surrounded with crystallization solution, right after its acquisition: in quartz capillary (at room temperature), with a flash-cooled in liquid nitrogen (77K), in liquid propane (150K) or with cryogenic nitrogen gas stream (100K) [15].

Keeping the crystal at low temperatures is one way to prevent the degradation by the free radicals formed by the interaction of the X-ray beam with the crystal in the data collection, which makes freezing techniques more advantageous [15].

To avoid the destruction of the internal order of the crystal by freezing the water molecules present in the buffer solution, the crystal needs to be involved in a solution containing cryo-protectors such as glycerol, PEG, sucrose or salts in an appropriate concentration, so that a vitrification, rather than freezing, process occurs [13, [34].

1.6 X-ray diffraction

Electromagnetic radiation is diffracted when it intersects with an object larger than its wavelength. This is what happens when X-rays (0.1-100Å) interact with electron clouds from the atoms of a molecule (1 Å) in their crystalline state, resulting in a diffraction pattern (Figure 1.6.1).

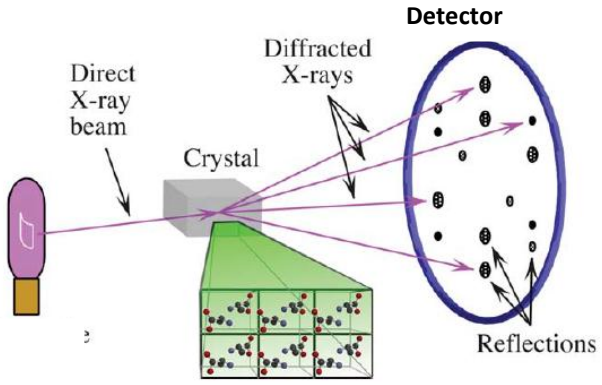


Figure 1.6.1 - Illustration of a data collection [34].

All spots in a diffraction pattern (reciprocal space lattice) are characterized by an intensity (I_{hkl}), a direction (Miller Indices: *h,k,l*) and a phase (α). Each spot corresponds to a Fourier summation of the scattered waves with constructive interference, in phase with each other (Bragg's Law, Figure 1.6.2). Within this mathematical formulation, only a few directions of the diffracted beams are detected, which makes the diffraction pattern dependent on the orientation of the crystal and the unit cell dimensions [34].

The unit cell (Figure 1.6.3) is the subunit that repeats along the crystal through translational processes, maintaining the number and the arrangements of the asymmetric unit (smaller fraction of the crystal) [15].

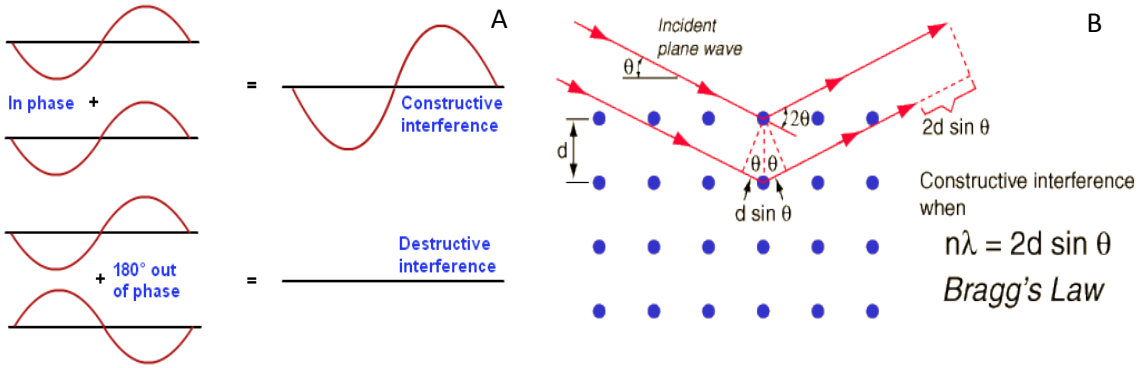


Figure 1.6.2 - Conditions for diffraction [67]-[68].
(A - Constructive interference between radiation. B- Application of Bragg's Law.)

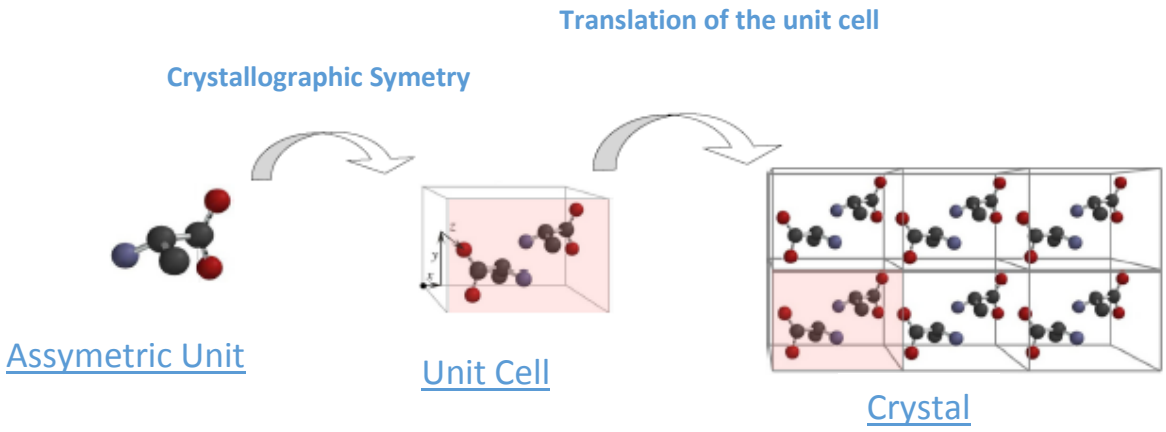


Figure 1.6.3 - Crystal Packing [34].

This way, to start a data collection it is necessary to first characterize an orientation matrix, by the determination of the parameters and the direction of the reciprocal unit cell axes. For this, two or more images are indexed at 0 ° and 90°, relative to the axis perpendicular to the X-ray beam. Through the resulting information, it is possible to identify the Laue symmetry and the space group, which determine the three-dimensional arrangement of the crystal (Bravais Lattice). After obtaining the orientation matrix, it is possible to calculate the best strategy to collect a complete dataset with the smallest crystal rotation avoiding radiation damage by overexposing the crystal.

Once calculated, the best strategy obtained is applied (Figure 1.6.4). The resulting images from the diffraction patterns of all the unique orientations of the crystal are then integrated and scaled together (data processing). Through the following parameters obtained in this step, it is possible to perform a first analysis of data quality:

- Resolution: Associated with the level of detail reached in the electron density maps.
- Completeness: Percentage of the unique reflections (theoretically estimated) that were obtained experimentally.
- Multiplicity: Estimation of the number of independent measurements for each reflection.
- Signal to noise ratio ($\frac{I}{\sigma(I)}$): Percentage of intensity obtained above noise level.
- Merging R-factor (R_{merge}): Agreement between the several independent observations of the same reflection.

At this stage, we can obtain the information about the directions, the intensities and hence calculate the amplitude of the structure factors (Equation 1.6.1).

$$|F_{obs}|^2 = I_{hkl}$$

Equation 1.6.1 - Mathematical equation to obtain the structural factors amplitude from the intensities [34].

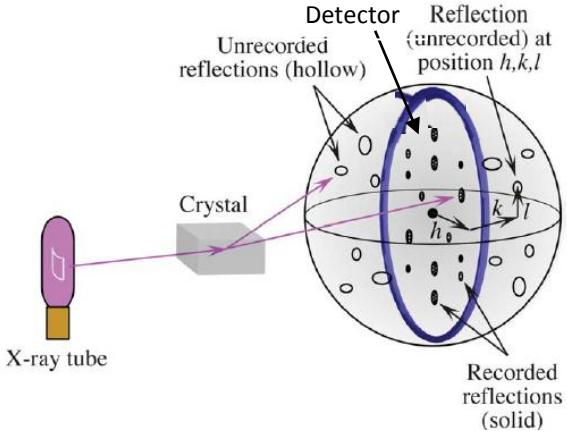


Figure 1.6.4 - Expected result in a crystallographic data collection for a three-dimensional analysis [34].

By measuring the diffraction in more than one orientation, it is possible to obtain a sphere of results, obtaining three-dimensional coordinates for each reflection (h, k and l).

The equation to calculate an electronic density map is obtained by applying a Fourier Transform to the structure factors equation (Equations 1.6.2). But for this, in addition to the intensities, it is necessary to measure the phases of individual diffracted X-ray waves, information which is lost during the data collection process (the so-called "Phase Problem") [34].

$$\rho(x,y,z) = \frac{1}{V}\sum_{hkl}|F_{hkl}| \times e^{-2\pi i(hx+ky+lz-\alpha_{hkl})} \Leftrightarrow$$

$$\Leftrightarrow \rho(x,y,z) = \frac{1}{V}\sum_{hkl}\sqrt{I_{hkl}} \times e^{-2\pi i(hx+ky+lz-\alpha_{hkl})}$$

Equation 1.6.2 - Mathematical equation used to calculate the electronic density map [34].

In order to solve this problem, there are several methods for the estimation, or close inference, of the phases. The choice of method to be used depends on the protein under study:

- In cases where an available structure of a homologous or similar protein already exists in the PDB, a Molecular Replacement (MR) procedure may be used; this method is based on the initial phase estimation of a new structure from a known structure model. For this, it is necessary to perform the best match between observed diffraction and the calculated diffraction, by testing all possible positions and orientations of the known structure model. Since, in the asymmetric unit, each

molecule is defined by six parameters (a, b and c related to rotation and d, and related to the translation), in order to make this search less exhaustive, most of the programs divides this process into two steps: first, the best solutions for the rotation function are search and then with these results, the search for solutions to the translation function. After obtaining the best possible parameters, the resulting phases of the known structure model and the amplitudes obtained in the data collection are used to calculate the initial electronic density map.

- In cases where the phases of some reflections already have starting values associated or are already known, *ab initio* phase determination is used to deduce the phases of the remaining reflections. For this method, widely used as a complement to other methods to find the atomic substructures of heavy atoms, it is necessary to obtain an atomic resolution better than 1.2 Å.
- For the cases of proteins without known struturally homologous proteins, the Multiple Isomorphous Replacement (MIR) and the Single Isomorphous Replacement (SIR) are methods that can be used. These ones require the isomorphic addition of heavy metals (Hg, Pt, Au, Pb or Ag) to the protein of interest, that is, without interfering with neither its three-dimensional structure nor the dimensions of the unit cell. Through the difference in the intensity of the diffracted beams in the presence of heavy atoms compared to the native protein, it is possible to perform a phase angles estimation.
- In the case of metalloproteins or proteins with added metals (e.g. Fe, Cu, Mo, Zn, or Ni), the methods of phase determination through anomalous scattering are the most advisable. These include the single wavelength anomalous dispersion (SAD) and the multiple wavelength anomalous dispersion (MAD). These use a radiation with one (SAD) or more (MAD) wavelengths near the absorption edges of the metal present in the protein, which causes a breakdown in Friedel's law (same intensity for reflections with symmetric miller indices, hkl and –h-k-l), but also an anomalous X-ray diffraction (with different phase and amplitude). From these differences it is possible to determine the atomic substructure, from which the phase can be estimated computationally for all the amplitude factors of the whole structure. Since they only need to collect data from a single crystal, the methods of

anomalous dispersion can overcome the problems presented by the isomorphous replacement methods. But on the other hand these have as a disadvantage the decay of the crystal due to radiation damage. After obtaining the phases, the electron density map is calculated (through Equation 1.6.3), followed by the model building process.

$$\rho\,(x,y,z) = \frac{1}{V}\sum_{h}\sum_{k}\sum_{l} F_{hkl} e^{-2\pi i(hx+ky+lz)}$$

Equation 1.6.3 - Equation for the calculation of the electron density maps [34].

In the case of proteins whose phase was determined by the MR method, a previous model is already available and model building is carried out to adjust/fit it into the electron density maps.

If the obtained phases were obtained by one of the other methods (Anomalous Scattering or Isomorphous replacement), the model has to start from scratch, according to the electronic density map, since it has no model structure to follow.

After its construction by either methods, the model needs to be refined in order to make it more consistent with the experimental data.

The refinement programs serve to optimize the agreement between the observed and calculated structural factors amplitudes by the following parameters: three dimensional coordinates, scaling factor, atomic occupation (fraction of crystal molecules in which a certain atom occupies the position determined by the model) and atomic displacement parameters, or B-factor, which correlates with the degree of mobility of an atom.

After each refinement cycle several validation parameters are obtained and used to analyze the results:

- R_{Work}: percentage of the general relative discrepancy between the structure factor amplitudes (observed and calculated).
- R_{free}: percentage of the relative discrepancy between the structure factor amplitudes (observed and calculated), in a set that have never participated in the calculation of the refinement (test set).
- Model geometry: error identification related to stereochemistry, chemical environments, bond angles and distances, chirality and planarity restraints, torsion

angles (Ramachandran plot), rotamer collection, among others, based on comparison with dictionaries of standard geometrical data.

With the refinement process we intend to see an improvement of the model, which in turn improves the phases and consequently the electronic density map. When, after the application of several iterative refinement cycles, there are no longer significant changes for the improvement of the model and the parameters of validation are of good quality (e.g. by benchmarking with statistics of known structures with similar resolution), the three-dimensional structure of the protein under study is terminated and then the structural analysis of the protein can be started.

1.7 Synchrotron

There are currently several sources of X-rays: the sealed tube, rotating anode, liquid anode, micro-source and synchrotron radiation source [7].

In comparison to the other sources, the synchrotron source (Figure 1.7.1) uses a very small beam in an ultra-high vacuum environment with less divergence, more intensity, highest polimerization and more brightness [38]. With these characteristics, the data collection process is faster and more efficient, which makes it possible to analyze small and/or weakly scattering crystals. Another advantage of this source is the possibility of selecting a specific wavelength of the incident X-rays, which allows the application of more specific studies such as anomalous dispersion experiments. As a disadvantage, by using a more intense X-ray beam one has a greater chance of causing radiation damage in the sample [15].

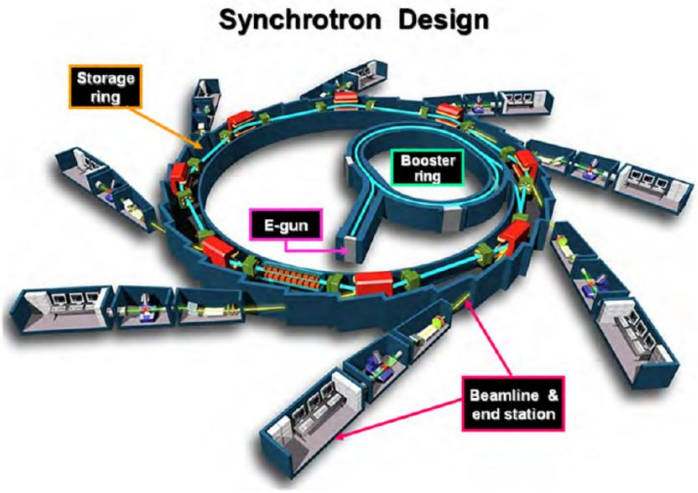


Figure 1.7.1 – General scheme of operation of a Synchrotron [39].

E-gun – *Linear accelerator of emission electrons from a cathode. In this the electrons are accelerated up to a speed close to the speed of light;* ***Booster Ring -*** *Circular accelerator that realizes a boost of energy of the electrons so that they arrive to a Giga electron volts (GeV);* ***Storage Ring –*** *Uses the path shift of the high-speed electrons (by magnetic fields) for the production of sinchrotron light;* ***Beamline -*** *Where the samples under study are analyzed* [40].

Figure 1.7.2 - Photographs of synchrotrons from where data were collected for this work [69], [70].

2. Materials and Methods

2.1 Crystallization of PPE in the native state

Before initiating the structural characterization of elastase with inhibitors, it is necessary to optimize the crystallization conditions in order to ensure that good quality crystals of protein are obtained. For this, six different crystallization conditions, based on previous reports ([30], [41]–[47]), were tested with the protein in the native state.

PPE (lyophilized) was purchased from SERVA Electrophoresis GmbH (Heidelberg, Germany), and dissolved in double-distilled water to a concentration of 40 mg/mL, without further purification. Since not all the reported conditions used the same concentrations of protein, it was necessary to dilute the sample to the concentrations mentioned in Table 2.1.1.

All conditions were tested with both techniques (hanging and siting drop). In each well 500 µL of crystallization solution were added and a drop of 1 µL of PPE mixed with 1 µL of reservoir solution was dispensed using the hanging and sitting drop techniques.

Table 2.1.1 - Summary of Porcine Pancreatic Elastase (PPE) crystallization conditions.

Technique	Hanging drop and Siting drop Vapour diffusion		
Temperature (°C)	20		
Protein	**Protein Buffer**	ddH$_2$O (Bidestilated water)	
Reservoir Volume (µL)	500		
Crystallization Solution	**1**	100 mM Sodium Acetate pH 5.2 200 mM Sodium Sulfate	
		Protein concentration (mg/mL)	20, 30 and 40
	2	100 mM Sodium Acetate pH 5.2 50 mM Sodium Citrate 5 mM Calcium Chloride	
		Protein concentration (mg/mL)	12 and 30
	3	50 mM Sodium Citrate Buffer pH 6	
		Protein concentration	40
	4	300 mM Sodium Chloride 50 mM Tris-HCl pH 7	
		Protein concentration (mg/mL)	30 and 40
	5	70% (v/v) 2-Methyl-2,4-pentanediol (MPD) 10 mM Sodium Phosphate Buffer pH 5.9	
		Protein concentration (mg/mL)	20
	6	17% (w/v) PEG 3350 200 mM Bicine pH 8.1 60 mM Sodium Citrate	
		Protein concentration (mg/mL)	12, 20 and 30
Drop ratio (µL)	1 protein + 1 reservoir		

2.2 Crystallization of HNE in the native state

Similar to PPE, five crystallization conditions, inferred from previous reports ([6], [48]–[50]), were tested with the protein in its native state, in order to obtain an initial crystallization condition. HNE (lyophilized) purchased from Elastin Products Company (Owensville, Missouri, USA) was dissolved in 20 mM Tris-HCl (pH 7.5) with 50 mM Sodium Chloride (again based on the same literature information), for a concentration of 20 mg/mL, without further purification. During this test, the concentrations of protein used (10-20 mg /mL), the ratios of the drops (1:1 and 2:1) and the technique used (hanging or sitting drop), were varied as mentioned in Table 2.2.1.

Table 2.2.1 - Summary of Human Neutrophilic Elastase (HNE) crystallization conditions.

Temperature (°C)			20
Protein	**Protein Buffer**		20 mM Trid-HCl (pH 7.5) 50 mM Sodium Chloride
Reservoir Volume (µL)			500
Crystallization Solution	**1**	1.5 M Ammonium Phosphate (pH 7)	
		Protein concentration (mg/mL)	10, 15 and 20
		Technique	HD
		Drop Ratio (µL) : 0.5 protein + 0.5 reservoir and 1 protein + 0.5 reservoir	
	2	20% (w/v) PEG8000 0.1 M HEPES (pH 7.5)	
		Protein concentration (mg/mL)	10, 15 and 20
		Technique	HD
		Drop Ratio (µL) : 0.5 protein + 0.5 reservoir and 1 protein + 0.5 reservoir	
	3	2 M Sodium Formate (pH 4.5)	
		Protein concentration (mg/mL)	10, 15 and 20
		Technique	HD
		Drop Ratio (µL) : 0.5 protein + 0.5 reservoir and 1 protein + 0.5 reservoir	
	4	28% (w/v) PEG4K 0.1 M Tris-HCl (pH 8.2) 0.7 M Lithium chloride	
		Protein concentration (mg/mL)	20
		Technique	HD and SD
		Drop Ratio (µL) : 1 protein + 0.5 reservoir	
	5	**Drop:** 0.6 M Sodium Phosphate (pH 5) 70 mM Sodium Chloride **Reservoir:** 1.5 M Sodium Phosphate (pH 5)	
		Protein concentration (mg/mL)	20
		Technique	HD
		Drop Ratio (µL) : 1 protein + 0.5 reservoir	

In order to increase the stability of the protein, reducing the possibility of undergoing inactivation by autolysis and / or proteolysis, another protein batch was dissolved in 20 mM Bis-Tris (pH 6) with 50 mM sodium chloride, to a concentration of 15 mg /ml. With this sample, four crystallization screens were carried out: Salt Rx (from Hampton Research), ShotGun, PACTPremier and BCS (all from Molecular Dimention).

2.3 Biochemical characterization

Determination of the final concentration of HNE was performed by the Bradford method and analysis of its purity by SDS-PAGE.

2.3.1 Bradford method

This procedure was started with the determination of a calibration curve. For this, several dilutions of a standard protein (Bovine Serum Albumin, BSA) were performed in the buffer of the protein under study. For each of them triplicates were prepared in a 96-well immune flat-bottom plate, each with 150 µl Coomassie plus (Bradford) assay reagent (from Thermo Fisher Scientific) and 5 µl of the respective dilution. After 30 minutes of incubation, the absorbances of the resulting samples from each well were measured at 595 nm using a BMG FluoSTAR Optima plate reader. Data were processed in the equipment's data analysis software (MARS Data Analyse Software version 2.10), obtaining, with the averages of the triplicates of each solution, the calibration curve.

At the same time, similar procedure was applied to the dilutions of the protein under study (dilution of 1:5, 1:10 and 1:100). Using the average of the three results obtained for each dilution, it was possible to estimate, from the calibration curve, the total concentration of the protein in the sample.

2.3.2 Electrophoresis (SDS-PAGE)

In the sample preparation process, the HNE solution was centrifuged for 15 minutes at 15000 rpm in a 5424 R centrifuge (Eppendorf). As a result the formation of a pellet was observed.

For the SDS-PAGE procedure it was prepared: one sample of 20 µL with 0.15 mg/ml of HNE in 1x Loadding Buffer solution (50 mM Tris-HCL, pH 6.8; 2% SDS; 0.025% Bromophenol Blue; 10% Glycerol and 12.5% β-mercaptoethanol) and one sample of 20 µL with 1uL of the pellet dissolution (in 50 µL of protein buffer) in 1x Loading buffer solution. Before being loaded onto Amersham ECL Gel 8-16%, (of 10 wells), both solutions were subjected to 6 minutes of incubation at 95° C.

In addition, the same gel was also loaded with 5 µl of PageRuler Plus Prestained Protein Ladder (marker from Figure 2.3.2.1).

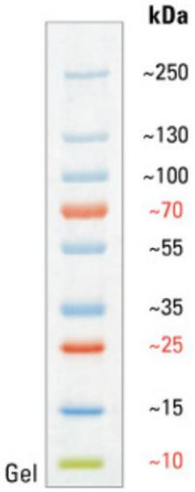


Figure 2.3.2.1 - PageRuler Plus Prestained Protein Ladder (marker) [51].

After loading, the gel was run at a constant voltage of 160 V. Then, two gel washes of the 10 minute in bi-distilled water were performed, followed by a 30 minute with Coomassie (from Biorad) staining and another two washes of 10 minutes in bi-distilled water. Finally, a scanner was run through the ImageScanner III, using the Coomassie filter of LabScan 6.0 program.

2.4 Micro-seeding with native crystals

Some native crystals of PPE were collected from crystallization dropplets (Figure 3.1.1.3(a) and 3.1.1.8(a)) with a micro-pipette into 50 µl of crystallization solution in an eppendorf with a Seed Bead® (Molecular Dimensions), in order to prepare the seeds for microseeding. Then the solution was vortexed twice for 40 seconds, with an interval of 10 seconds of incubation on ice between them. From the resulting solution, various dilutions were prepared (1:10^2, 1:10^3, 1:10^4 and 1:10^5). New drops were prepared in a sitting drop vapor diffusion experiment, using 1 µL of PPE, 0.8 µL crystallization solution and 0.2 µL of microseeds dilution solution.

2.5 Soaking of the native crystals with the Inhibitors

The best native crystals of PPE were obtained in condition 5 composed of 70% MPD and 10 mM Sodium Phosphate Buffer pH 5.9. To some of these drops 1 µL of solution with 100 mM of Ligand (Figure 2.5.1) dissolved in 100% (v/v) Dimethylsulfoxide (DMSO) was added, to obtain around 5X of molar excess ([PPE] ~$7{,}69 \times 10^{-4}$ M). After the ligand addition, the drop was left overnight at room temperature, in order to allow the ligand to diffuse through the solvent channels present in the crystal.

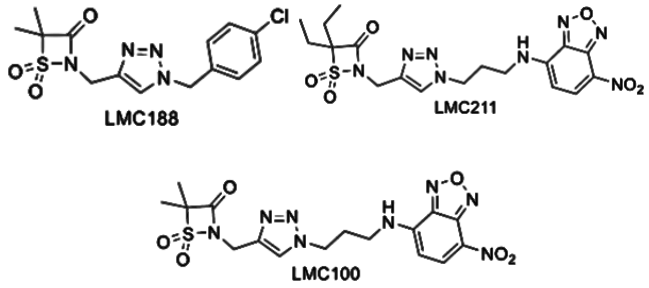


Figure 2.5.1 - Ligands synthesized by the group of Rui Moreira, Faculdade de Farmácia, for Pancreatic Porcine Elastase, used in the technique of soaking.

2.6 Co-crystallization with the inhibitors

In parallel with the soaking experiments, several co-crystallization tests were carried out for PPE. During these assays two different molar excess (4X and 6X) of ligands (Figure 2.6.1) were incubated with the protein for 45 min at a room temperature (with a final DMSO concentration of 5%) prior to the crystallization experiments.

For the assays with HNE, only a molar excess of 7X was tested with the ligand displayed in Figure 2.6.2. Here, the ligand solution (final concentration around 5 mM in 5% DMSO) was added to 20 mg/mL HNE solution.

After an incubation of approximately one hour, at room temperature, the crystallization drops were prepared with several ratios between the incubated sample (protein + ligand) and crystallization solutions (1:1, 1:2 and 2:1) for a hanging and a sitting drop vapor diffusion experiments.

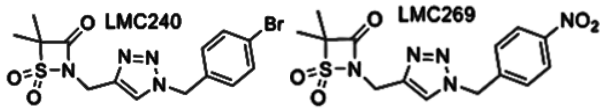


Figure 2.6.1 - Ligands synthesized by the group of Rui Moreira, Faculdade de Farmácia, for Pancreatic Porcine Elastase, used in the method of co-crystallization.

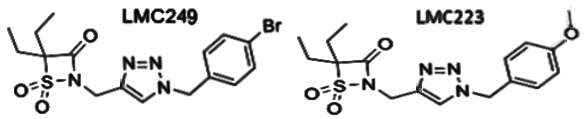


Figure 2.6.2 - All the ligands synthesized by the group of Rui Moreira, Faculdade de Farmácia, for Human Neutrophil Elastase.

2.7 X-ray diffraction Data Collection and Processing

In this work, the X-ray diffraction experiments were only performed with the crystals of the PPE complexes.

In order to carry out the data collection process, the crystals were removed from the crystallization drop with a loop and then frozen and stored in liquid nitrogen (100K). For the crystals of PPE crystallized under condition 5, no cryoprotective solution was required since 70% MPD is already a cryoprotectant. However, the crystals of condition 1 (100 mM Sodium Acetate pH 5.2 and 200 mM Sodium Sulfate) required the supplementation of the crystallization solution with 30% glycerol.

X-ray data for this study were collected on ID23-1 and ID30A-3 beamlines at the European Synchrotron Radiation Facility (ESRF) in Grenoble (France) and beamline I03 of Diamond Light Source in Didcot (Oxforshire - United Kingdom).

The data collected at ESRF were processed through the *autoPROC* pipeline [52] using the programs *XDS*, for the indexing and integration [53], *SCALA*/*AIMLESS* for the scaling [54] and *POINTLESS* for space-group determination [55]. The data collected at Diamond were processed with the *Xia2 3dii* pipeline [56] which also uses the program XDS to index each measured reflections, but for scaling uses the XSCALE program [57].

The structures were determined by molecular replacement performed with *PHASER* as implemented in the *PHENIX* suite of programs [58], using the coordinates of the complex PPE-JM102 (PDB entry 4YM9 [59]) devoid of any solvent or ligand molecules. Iterative cycles of model building and refinement with *COOT* [60], [61] and phenix.refine, also from the same suite of programs, were performed until convergence. *MolProbity* [47], as implemented in *PHENIX*, together with *PROCHECK* [62], and *WHATCHECK* [63], from the *CCP4* suite programs[64], were used for model validation.

3. Results and Discussion

3.1 Structural analysis of PPE complexes

3.1.1 Crystallography, Data collection and Processing

The crystallization process, as already mentioned in the working methods, started with the attempt to reproduce six crystallization conditions that had already been reported in the literature for this protein. Several drops were made under these conditions varying the method used (hanging or sitting drop), protein concentration (ranging from 12-40 mg/mL), maintaining the drop ratio (1:1) and the crystallization temperature (20 °C).

After 5 days of incubation, it was possible to observe a crystalline material in condition 5 (70% MPD with 10 mM Sodium Phosphate Buffer pH 5.9, in a sitting drop format using PPE at 20 mg/mL - Figure 3.1.1.1(a)) and, after 14 days in condition 2 (100 mM Sodium Acetate pH 5.2 with 50 mM Sodium Citrate and 5 mM Calcium Chloride in Sitting drop with 30 mg/mL of protein – Figure 3.1.1.1(b)).

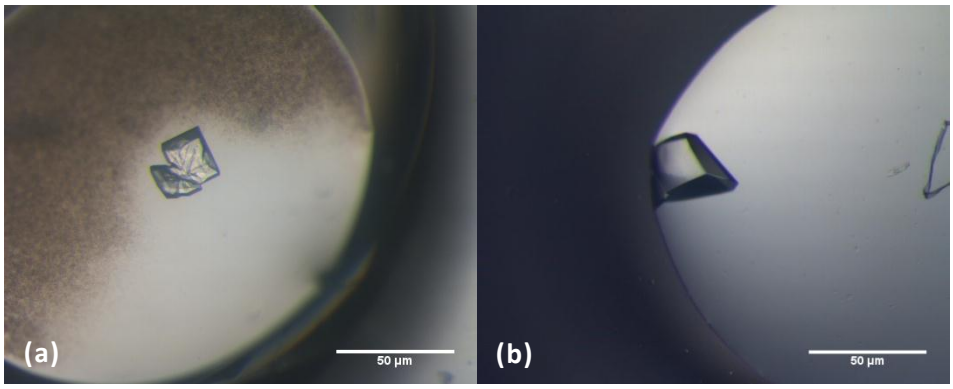


Figure 3.1.1.1 - Native PPE crystals obtained in a screen of conditions (Vapour Diffusion Sitting Drop).

(a) - Crystal of condition 5 with a drop ratio of 1:1, 20 mg/mL of protein, at 20°C; (b) - Crystal of condition 2 with a drop ratio of 1:1, 30 mg/mL of protein, at 20°C.

After careful analysis, it was decided to purpose the work with condition 5, being the least aqueous solution of all, facilitating the solubility of the compounds in question (hydrophobic compounds) for the complex crystallization. Since the parameters of the protein concentration and the crystallization method had already been optimized with the first assay, only the optimization of the precipitant percentage of the condition was required.

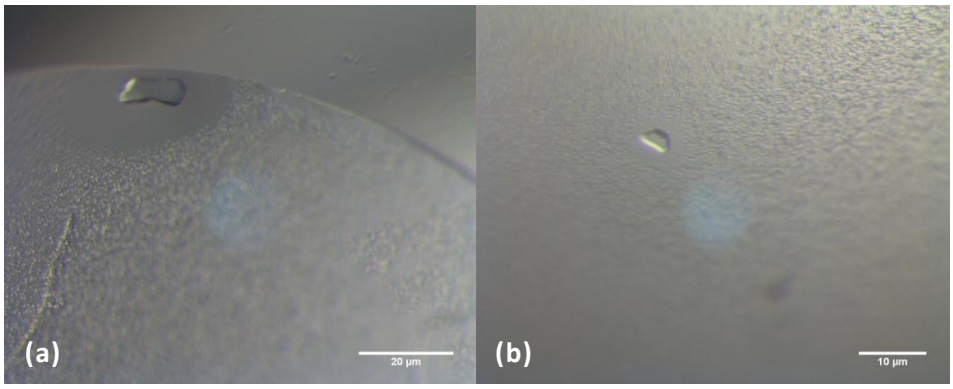


Figure 3.1.1.2 - Precipitant percentage optimization assay.
(a) - Condition 5 with 65% of MPD; (b) - Condition 5 with 75% MPD.

After crystals of appearance (Figure 3.1.1.2), it was possible to verify that the optimal percentage of MPD for the crystallization of this protein is the one initially used (70%).

Streak seeding was performed on the remaining drops of the same condition using a seed the crystal of Figure 3.1.1.1(a), in order to promote nucleation. The grown crystals (Figure 3.1.1.3) are intended to apply the soaking technique or to serve as the seed source for the following tests.

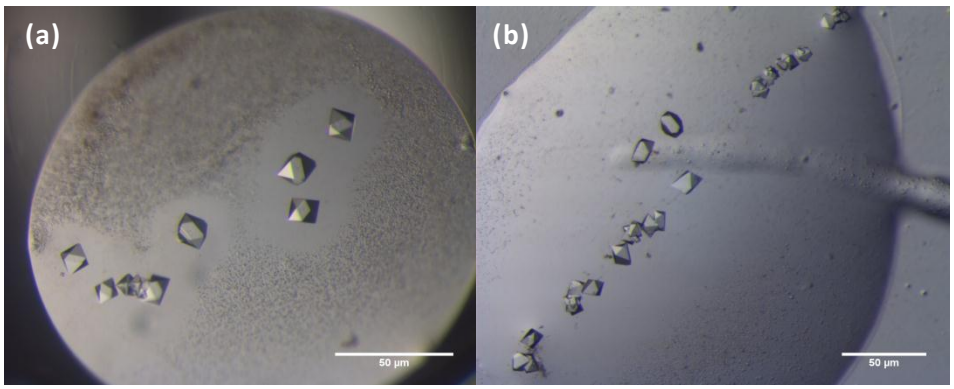


Figure 3.1.1.3 – Native crystals obtain in Condition 5 by Streak seeding.

Once native crystals were obtained, the crystallization of the complexes was performed. For crystals of the complex, both possible techniques were tested (Soaking and Co-crystallization).

As previously mentioned, the crystalline matrixes are extremely sensitive, which means that in the soaking technique, when adding a solution to the drop different from the crystallization solution, the equilibrium of the system is distrurbed. As a consequence, this perturbation may affect the crystal in the drop, which can also cause the reduction of the X-ray diffracting signal. Another aspect of this technique is that the ligand can change the

conformation of the protein, which in turn can destruct the crystal upon ligand addition. Therefore, we first tested the "resistance" of the native crystals in the soaking process with two ligands (LMC188 in Figure 3.1.1.4(b) and LMC100 in Figure 3.1.1.4(d)). As hydrophobic compounds the inhibitors were dissolved in 100% of the organic solvent DMSO (dimethylsufoxide) in stock solution. To each drop with native crystals (Figure 3.1.1.4 (a) and (c)) 1uL of a 1:10 ligand dilution was added, leaving each drop with 3% DMSO and 3 mM ligand, during 6 days.

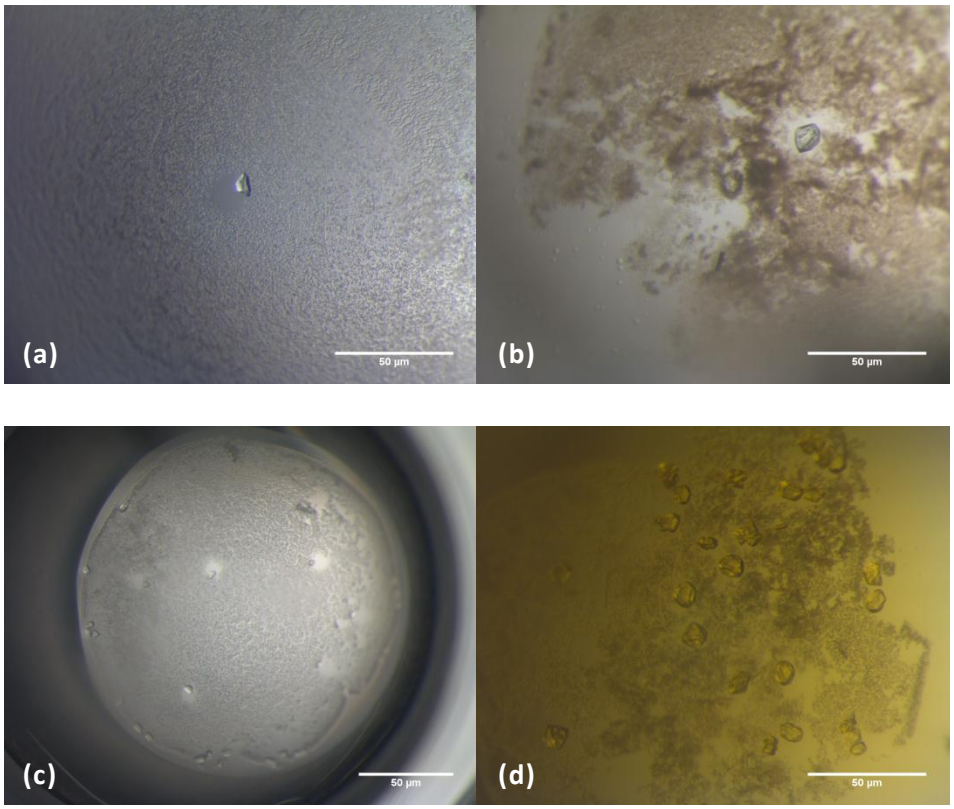


Figure 3.1.1.4 - Soaking experiments.

(a) and (c) - Crystals of native PPE after 6 days of incubation; (b) - Crystal (a) soaked with 1 µL of a 1:10 dilution of the ligand LMC188 in 85% MPD with 10 mM of Phosphate Buffer pH 5.9; (d) - Crystal (c) soaked with 1 µL of a 1:10 dilution of the ligand LMC100 in 85% MPD with 10 mM of Phosphate Buffer pH 5.9;

As we can observe, after soaking, the crystals presented a slight rounding in their edges, but no crash, which made us decide to proceed with this technique and further soak crystals of Figure 3.1.1.3(a) with LMC211, for 24 hours. The soaked crystals were frozen without any addition of cryo solution and X-ray diffraction were measured at in beamline ID23-1, ESRF.

In parallel with the soaking tests, co-crystallizations experiments were performed with the ligands as well, using the same conditions (Figure 3.1.1.5). This technique is widely used in cases where one of the two elements of the complex easily aggregates.

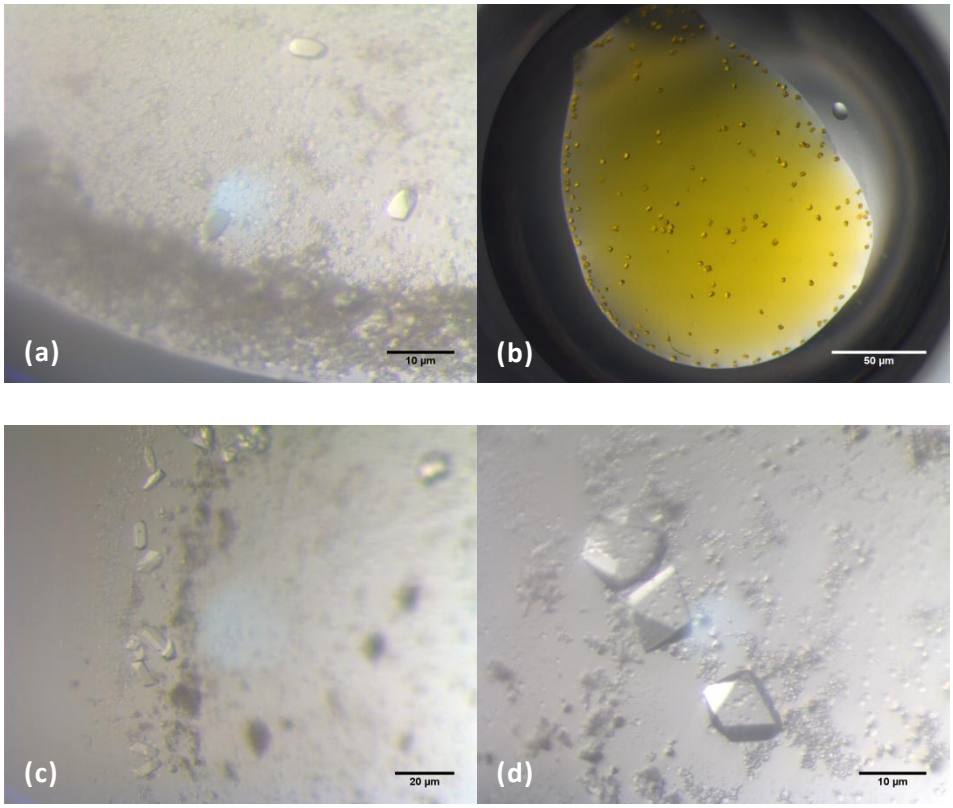


Figure 3.1.1.5 - Crystals obtained by co-crystallization (Sitting Drop).

Incubation for 15-30 minutes with a 5-fold molar excess at a room temperature. (a) – LMC188 ligand with a drop ratio of 1:2; (b) – LMC100 ligand with a microseeding of a 10^{-4} seeds dilution; (c) – LMC240 ligand with a drop ratio of 1:1; (d) – LMC269 ligand with a microseeding of a 10^{-4} seeds dilution.

Like the crystals obtained by soaking, these crystals were also directly frozen in liquid Nitrogen and studied using Synchrotron radiation (Figure 3.1.1.5(a) on the ESRF Beamline ID23, (b-c) on the Diamond Beamline IO3 and (d) on the ESRF Beamline ID30A-3).

After all beamline ID-23 data sets were processed, scaled and merged through the *auto*PROC pipeline, the molecular replacement method was applied for phase determination, having as search model the PPE structure with a PDB code of 4YM9 [59]. With these, it was possible to compute an electronic density map with a preliminary model in order to analyze possible blobs corresponding to the electronic density of the ligand. Statistics for data

collection and processing of each complex are shown in Table A.1 (in appendix) and in summary with the most relevant parameters in Table 3.1.1.1.

Table 3.1.1.1 - Data collection and processed data of ESRF (ID23-1)

Ligand	**LMC 188 (Co-Crystallization)**	**LMC188 (Soaking)**	**LMC 100 (Soaking)**	**LMC 211 (Soaking)**
Space Group	$P\ 2_1\ 2_1\ 2_1$		$P\ 2_1\ 2_1\ 2$	
Unit cell (Å)	*a*=50.56 *b*=57.56 *c*=74.60 (*α*=*β*=*γ*= 90°)	*a*=74.20 *b*=49.97 *c*=57.81 (*α*=*β*=*γ*= 90°)	*a*=57.73 *b*=74.64 *c*=50.30 (*α*=*β*=*γ*= 90°)	*a*=50.51 *b*=57.85 *c*=74.82 (*α*=*β*=*γ*= 90°)
Resolution (Å)	45.57 – 1.33 (1.40 – 1.33)	41.45 – 1.25 (1.28 – 1.25)	45.67 – 1.17 (1.19 – 1.17)	45.77 – 1.27 (1.29 – 1.27)
Completeness (%)	90.1 (58.3)	97.9(97)	97.7 (98.1)	90.6 (54.5)
<*I*/σ(*I*)>	10.2 (0.8)	11.8 (2.2)	11.1 (2.1)	13.3 (2.2)
Multiplicity	3.7 (2.6)	4.2 (4.1)	3.6 (3.3)	4.0 (3.7)
$CC_{1/2}$	1 (0.38)	1 (0.73)	1 (0.72)	1 (0.81)
Matthews coeficiente (A³/Da)	2.10	2.03	2.09	2.10
Crystal solvent (%)	41	39	41	41
Number of molecules in asymmetric unit	1	1	1	1

*Values between parentheses correspond to the values of the highest resolution

The data obtained in the Beamline ID-23 were important to choose the best method of crystallization for the study of these complexes (soaking or co-crystallization).

In the analysis of the electron density maps of the soaking complexes, it was not possible to observe any positive density blob (F_o-F_c) near the active center of the protein. One of the possible reasons for this is that the ligand is too large to enter and reach the active center of the protein through the solvent channels of the native crystal.

On the contrary, in the electron density map of the crystal obtained by co-crystallization with LMC188, it was possible to observe a positive density (F_o-F_c), near the active center corresponding to the structure of the ligand (Figure 3.1.1.6).

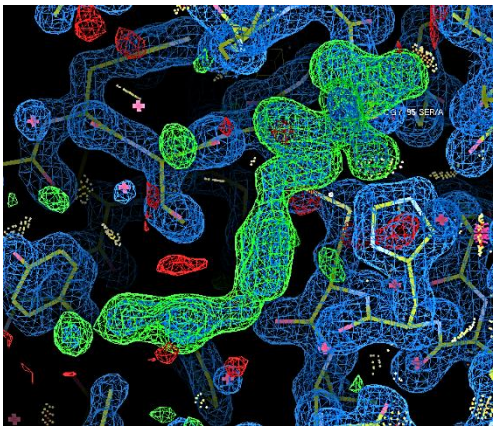


Figure 3.1.1.6 - Electronic density map around the active site with co-crystallized LMC188. $2F_O$-F_C maps are contored at 1σ level (blue) and F_o-F_c maps are contoured at 2.5σ and colored green (positive) or red (negative).

These results, show that the most successful technique, for the interaction study of these ligands with PPE is the co-crystallization, being decided to proceed similarly for the remainders assays.

All data of the other complexes were also collected in synchrotron, at a beamline I03 of Diamond and ID30A-3 of ESRF. For these data sets, the procedure as described above was followed (the data obtained in beamline I03 was processed, scaled and merged through the Xia2 3dii pipeline), obtaining as preliminary results the electronic density maps of the Figure 3.1.1.7. Statistics for the data collection and processing are shown in Table A.2 (in appendix) and in summary with the most relevant parameters in Table 3.1.1.2 and the Table 3.1.1.3.

Table 3.1.1.2 - Data collection and reprocessed synchrotron data (Diamond - Beamline I03).

Ligand	**LMC 100 (Co-crystallization)**	**LMC211 (Co-crystallization)**	**LMC 240 (Co-crystallization)**
Space Group	$P\ 2_1\ 2_1\ 2_1$		
Unit cell (Å)	*a*=51.63 *b*=57.31 *c*=74.54 (*α*=*β*=*γ*= 90°)	*a*=51.61 *b*=57.38 *c*=74.62 (*α*=*β*=*γ*= 90°)	*a*=51.10 ; *b*=57.71 ; *c*=74.64 (*α*=*β*=*γ*= 90°)
Resolution (Å)	38.36 – 1.22 (1.24 – 1.22)	37.31 – 1.30 (1.32 1.30)	31.34 – 1.28 (1.30 – 1.28)
Completeness (%)	99,4 (95.5)	99.7 (99.9)	99.6 (99.6)
<*I*/σ(*I*)>	13.2 (1)	12.2 (1.4)	14.8 (1.2)
Multiplicity	4.7 (2.8)	5.2 (5.3)	3.9 (3.6)
***CC*$_{1/2}$**	1 (0.48)	1 (0.7)	1 (0.5)
Matthews coeficiente (A^3/Da)	2.13	2.13	2.12
Crystal solvent (%)	42	42	42
Number of molecules in asymmetric unit	1	1	1

*Values between parentheses correspond to the values of the high resolution shell

Table 3.1.1.3 - Data collection and reprocessed synchrotron data (ESRF - Beamline ID30A-3).

Ligand	**LMC 269 (Co-crystallization)**		
Space Group	$P\ 2_1\ 2_1\ 2_1$		
Unit cell (Å)	*a*= 57.74 ; *b*= 57.71 ; *c*= 74.65 (*α*=*β*=*γ*= 90°)		
Resolution (Å)	30.07 – 1.38 (1.40 – 1.38)	**Matthews coeficiente (A^3/Da)**	2.40
Completeness (%)	96.8 (79.2)		
<*I*/σ(*I*)>	16.3 (2.3)	**Crystal solvent (%)**	49
Multiplicity	4.4 (3.2)	**Number of molecules in asymmetric unit**	1
***CC*$_{1/2}$**	1 (0.82)		

*Values between parentheses correspond to the values of the high resolution shell

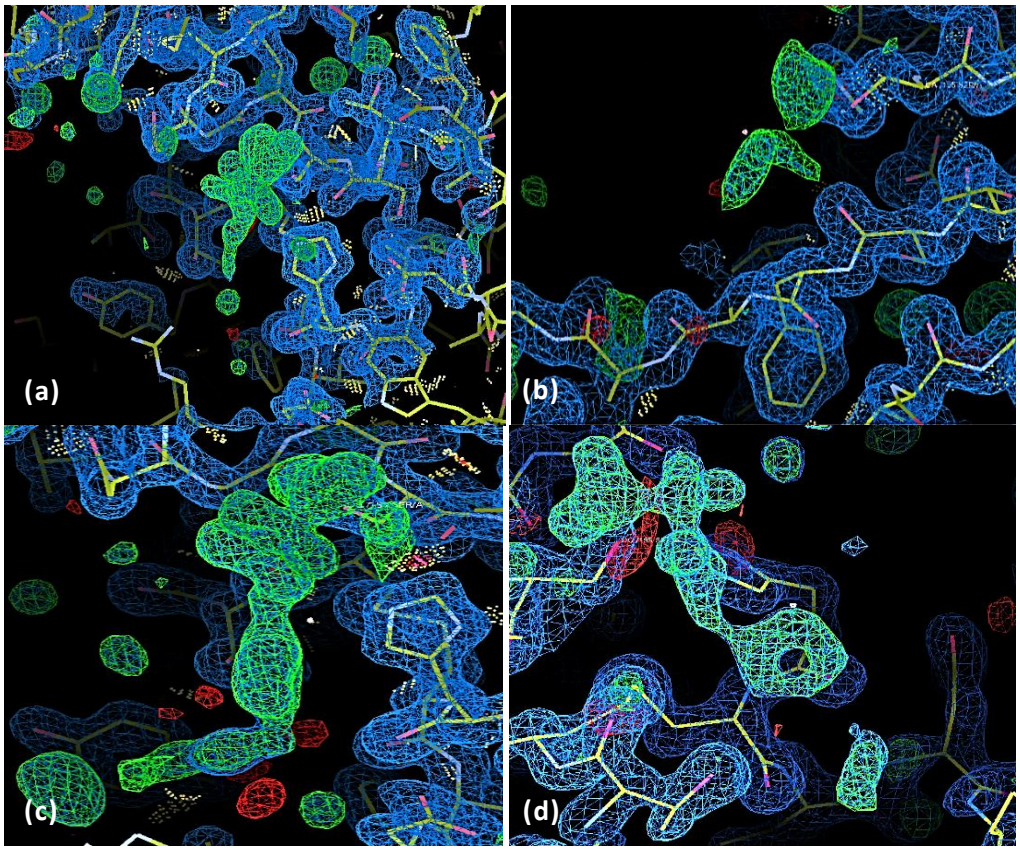


Figure 3.1.1.7 - Electronic density map around the active site of the putative complexes formed by Co-crystallization. (a) – Complex with LMC100; (b) - Complex with LMC211; (c) - Complex with LMC240; (d) - Complex with LMC269.

As observed in Figure 3.1.1.7, the electron density map of the complexes with LMC240 and LMC269 (Figure 3.1.1.7-C and 3.1.1.7-D) had a good positive density at the active site sufficient to cover the entire test compound. In contrast, the electron density maps of the LMC100 and LMC211 complexes only showed a small positive density at the serine 195 (Figure 3.1.1.7-A and 3.1.1.7-B), not sufficient to cover the compound in its entirety.

In order to solve the problem of the appearance of a weak positive density near the active center in the PPE-LMC100 and PPE-LMC211 complexes, the second promising condition of crystallization (Condition 2 – 100 mM Sodium Acetate pH 5.2 with 50 mM Sodium Citrate and 5 mM Calcium Chloride) was used in the following studies.

The crystal initially obtained of the condition 2 (Figure 3.1.1.1(b)) was used as seed source to perform streak seeding on the remaining drops of equal (condition 2) or similar (condition 1) conditions (Figure 3.1.1.8).

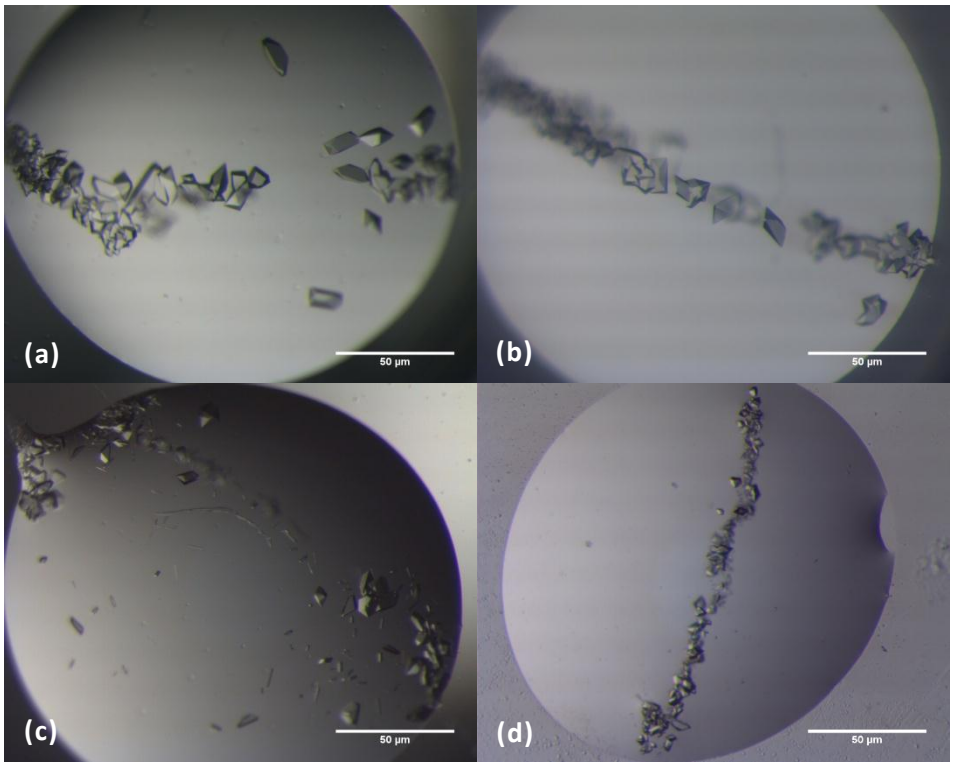
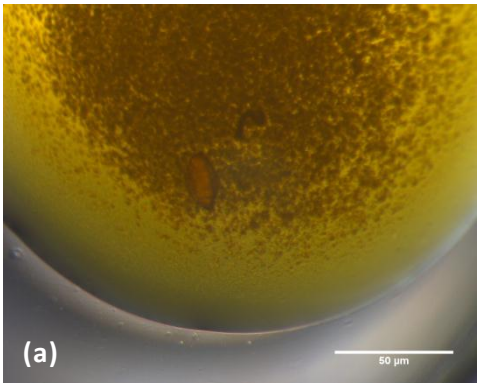


Figure 3.1.1.8 - Streak seeding in the Sodium Acetate Conditions.

(a – b) – Condition 2: 100 mM Sodium Acetate pH 5.2 with 50 mM Sodium Citrate and 5 mM Calcium Chloride;
(c – d) – Condition 1: 100 mM Sodium Acetate pH 5.2 with 50 mM Sodium Sulfate.

Based on previous results, co-crystallization droplets were performed with ligands LMC211 and LMC100. In order to achieve the nucleation of the drop more quickly, after overnight equilibration of the drop, microseeding and streak seeding were applied, obtaining the co-crystals shown in Figure 3.1.1.9.



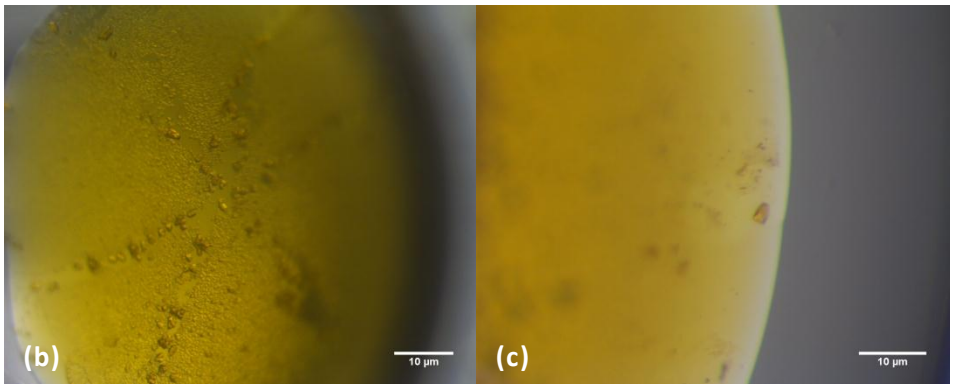


Figure 3.1.1.9 - Crystals of condition 1 obtained by co-crystallization. (a) – Microseeding nd (b-c)- Streakseeding.

Since condition 1 (100 mM Sodium Acetate pH 5.2 with 200 mM Sodium Sulfate) did not show characteristics of a cryo solution, the crystals had to be first cryo-protected with 0.1 M Sodium Acetate, 0.25 M Sodium Sulphate and 30% Glycerol, before being flash-cooled in liquid nitrogen. However, these crystals did not diffract at Diamond I03 beamline, which may be related with the choice of the cryo solution, since they were not previously tested.

3.1.2 Model building and refinement

According to the initial electron density map, changes were made (with the program COOT [60], [61]) to the final model used in the molecular replacement procedure, in order to make it more in accordance with the results of observed X-ray diffraction.

After model building, it was carried out a refinement cycle that computed an improved electronic density map. This procedure was repeated until convergence between the model and the map was obtained (Figures 3.1.2.7 (a-c) and Table 3.1.2.1 (a-c)).

As previously mentioned, in all co-crystallization experiments, the protein had a positive density corresponding to the density of the ligand at its catalytic center. This density, in all the refined models obtained, presented two blobs near the oxygen of the catalytic serine (Ser195), which did not agree with the expected results (Figure 3.1.2.3).

According to previous studies [6], the protease would break the inhibitor β-Sultam ring (Figure 3.1.2.1) through a serine 195 nucleophilic attack on its carbonyl group (Figure 3.1.2.2) [6]. Analyzing the resulting positive density (Figure 3.1.2.3), it is found that, contrary to what was expected under this mechanism, there are two blobs instead of just one next to the catalytic serine (Figure 3.1.2.4).

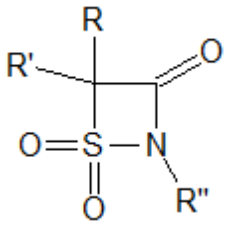


Figure 3.1.2.1 - Inhibitor β-Sultam ring [65].

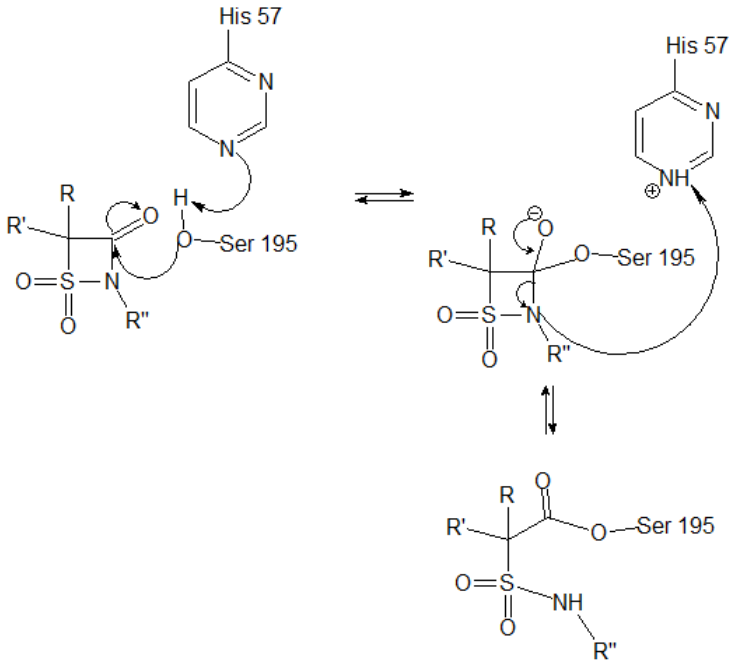


Figure 3.1.2.2 - Mechanism of action 1 - With nucleophilic attack on the carbonyl group.

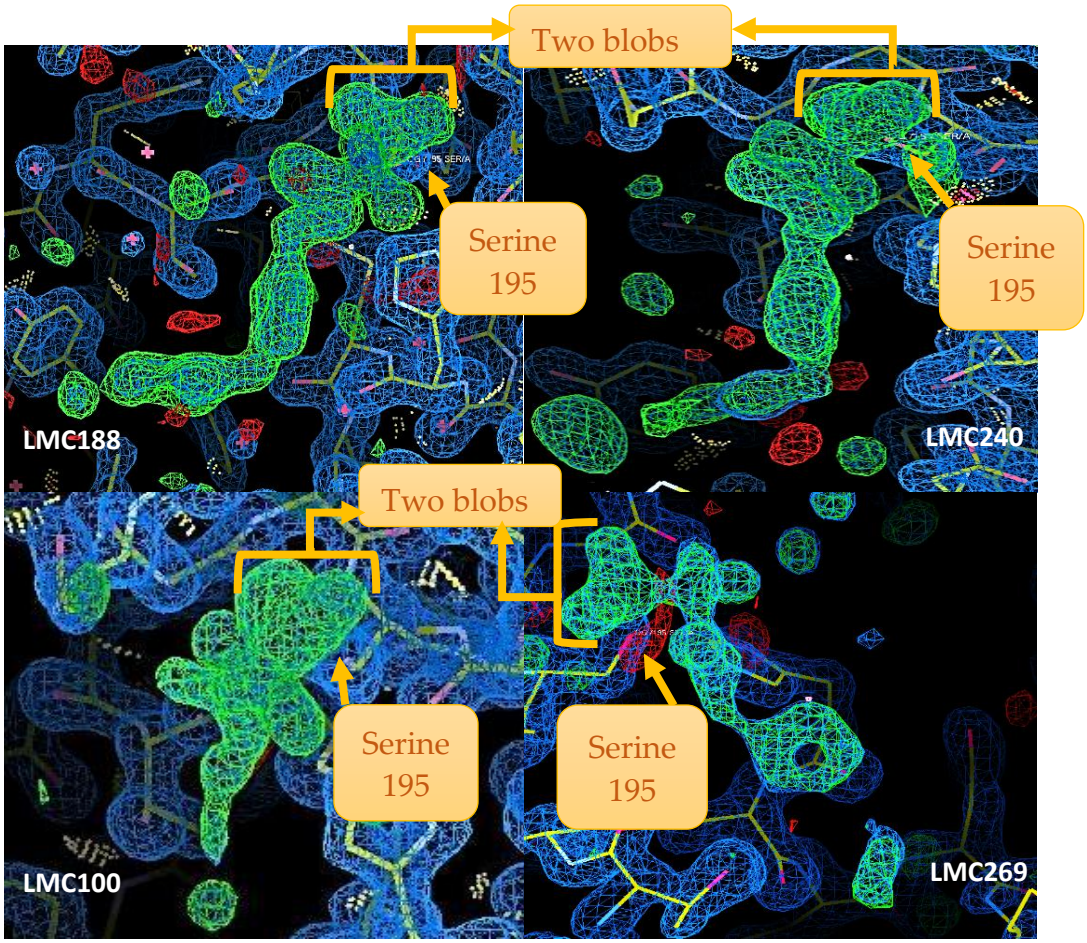


Figure 3.1.2.3 - Illustration of the blobs obtained around the catalytic center.

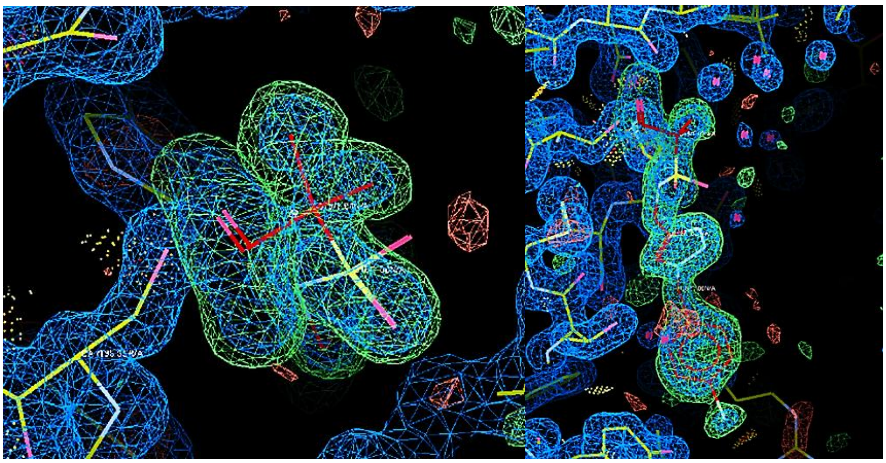


Figure 3.1.2.4- Fit of the ligand resulting from the mechanism of action 1 at the resulting F_o-F_c density, near the catalytic center of the enzyme.

Given this, it was found that the resulting positive density would only be fully justified if the inhibitor β-Sultam ring were opened through a nucleophilic attack on the sulfonyl group by the catalytic serine of PPE (Figure 3.1.2.5), as described in previous studies with β-Sultam inhibitors without the carbonyl group [65]. As a result a covalent bond is formed between the O_γ of serine 195 and the sulfur atom of the inhibitor.

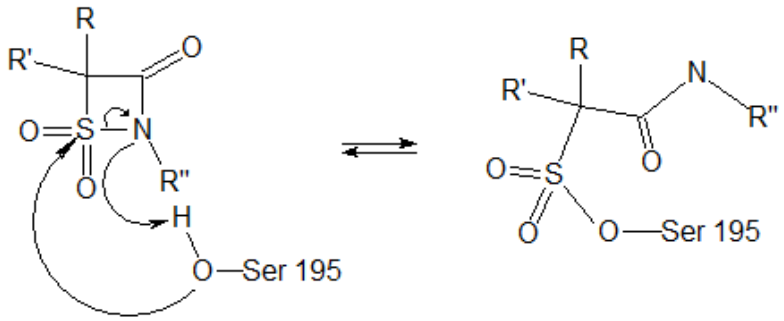


Figure 3.1.2.5 - Mechanism of action 2 - With nucleophilic attack on the sulfur group.

After this analysis, the ligands resulting from the nucleophilic attack on the sulfonyl group were drawn through the JLigand of *COOT*. The refinement of the designed ligands covalently bound into PPE active site showed good agreement with the electron density maps (Figure 3.1.2.6), confirming the mechanism 2 (Figure 3.1.2.5) predominance for this type of ligands.

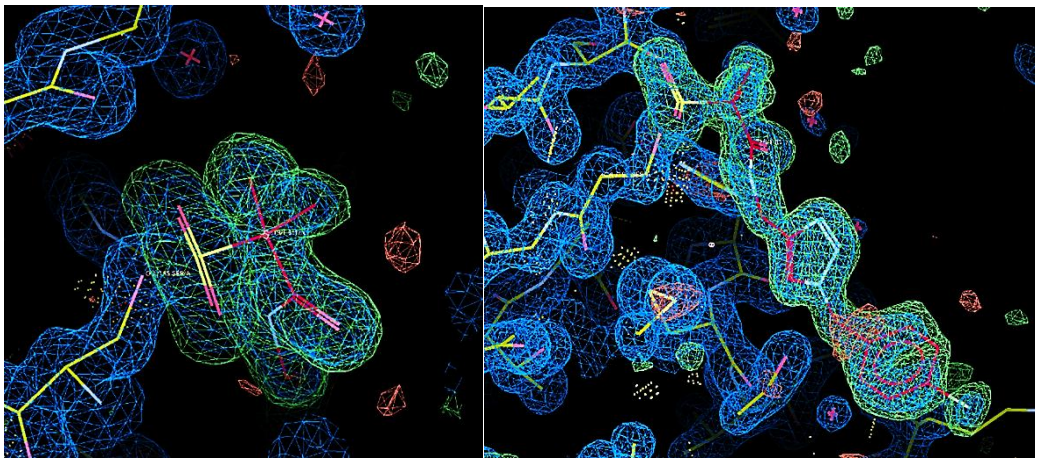


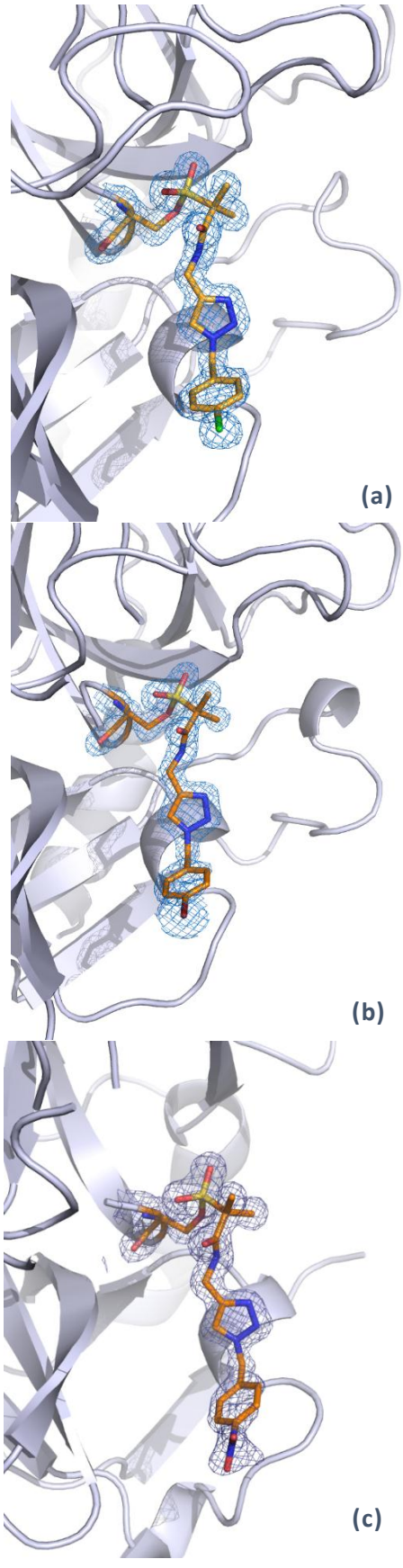
Figure 3.1.2.6 - Fit of the ligand resulting from the mechanism of action 2 at the resulting Fo-Fc density, near the catalytic center of the enzyme.

In smaller compounds (LMC240 and LMC188), only incomplete occupancy at the terminal atom (bromine with an occupation of 0.73 and chlorine with one of 0.57) removes the appearance of negative density. Low occupancies indicates that not all molecules present in the crystal contains the respective atom. Because they are atoms present at the end of the ligand, one of the possible reasons for the low occupation of bromine and chlorine will be radiation damage.

For larger compounds (LMC269 and LMC100), electronic density is not observed for the whole ligand, being most emphasized in the PPE-LMC100 complex, with only density at the interface between the protein and the ligand (Figure 3.1.1.7(a)). That is, as the ligand increases in size, its mobility also increases, causing a fading of the observed electron density.

The validation parameters of the refined models are presented in the Tables 3.1.2.1 (a-c). Comparing the structure of PPE in its "apo" and complexed forms show that the binding of the ligands under study did not cause significant conformational changes in the protein.

Table 3.1.2.1 - Validation parameters obtained in the final refinement.



(a) PPE-LMC188 Refinement (cycle number 43)

R-work (%)	13.75	**No. of molecules:**	
R-free (%)	16.25	**- Water**	213
Angles (°)	1.002	**-DMSO**	0
Bonds (Å)	0.008	**- Phosphate**	2
Ramachandran:		**- MPD**	3
-Prefered Regions (%)	97.30	**No. of solvent atoms:**	
-Allowed Regions (%)	2.3	**- Sodium**	1
-Outliers (%)	0.4		

(b) PPE-LMC240 Refinement (cycle number 19)

R-work (%)	14.57	**No. of molecules:**	
R-free (%)	16.23	**- Water**	236
Angles (°)	1.087	**-DMSO**	0
Bonds (Å)	0.009	**- Phosphate**	1
Ramachandran:		**- MPD**	1
-Prefered Regions (%)	96.60	**No. of solvent atoms:**	
-Allowed Regions (%)	2.98	**- Sodium**	0
-Outliers (%)	0.43		

(c) PPE-LMC269 Refinement (cycle number 19)

R-work (%)	12.77	**No. of molecules:**	
R-free (%)	15.77	**- Water**	268
Angles (°)	1.000	**-DMSO**	0
Bonds (Å)	0.007	**- Phosphate**	2
Ramachandran:		**- MPD**	2
-Prefered Regions (%)	97.00	**No. of solvent atoms:**	
-Allowed Regions (%)	2.6	**- Sodium**	1
-Outliers (%)	0.4		

Figure 3.1.2.7 - Electron density ($2F_o$-F_c) around the ligands after refinement at 1 sigma level.

3.2 Structural analysis of HNE complexes

3.2.1 Energy Minimization

In order to analyse the structure of the complexes between HNE and the ligands under study, the technique of energy minimization was used. For this, through the tools available in *COOT* program, structures of the native HNE and the PPE-inhibitor complexes were superimposed based on their highly similar fold (Figure 3.2.1.1).

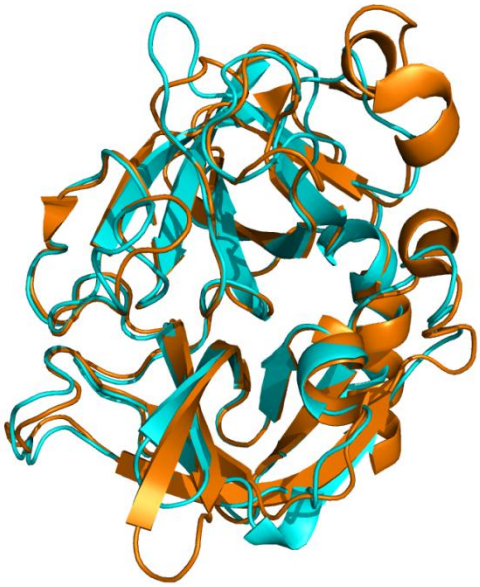


Figure 3.2.1.1 - PPE and HNE structure alignment [5].

After superposition, the inhibitor was fitted into HNE active site. Before proceeding for the energy minimization itself, the ligand was bound to serine 195, forming a covalent complex. Through the application of the *REFMAC5* [66] "energy minimization" module, within the *CCP4* suit programs [64], it is possible to obtain an energy minimized model of this complex that simulates the actual model.

In order to compare HNE and PPE complexes, the resulting minimization models were overlaid with the electronic density maps of the final refinement of the PPE complexes (Figure 3.2.1.2).

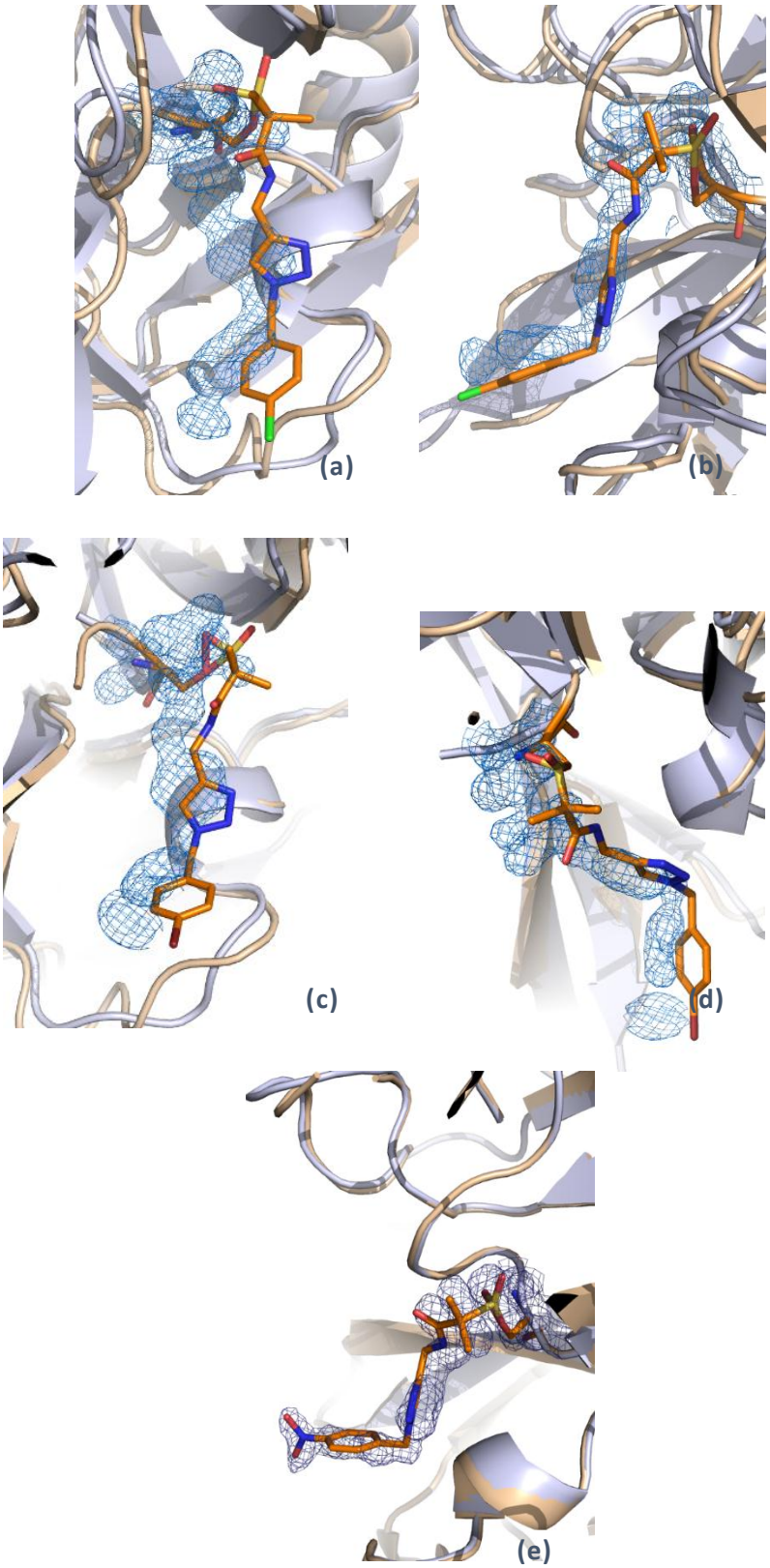


Figure 3.2.1.2 - Minimization models overlaid with the electronic density maps of the final refinement of the PPE complexes.

(a - b) – With LMC188; (c - d) – With LMC240; (e) – With LMC269.

As can be seen in the figures above, according to this prediction, the orientation of the ligand in the HNE-inhibitor complexes will be quite similar to the orientation present in the PPE-inhibitor complexes, as expected.

3.2.2 Protein Crystallization

Like the previous study, the crystallization process for HNE protein started with the attempt to reproduce five crystallization conditions, already described in the literature. For this, the lyophilized protein was dissolved in a buffer most used therein, composed of 20 mM Tris-HCl (pH7.5) with 50 mM NaCl.

Several drops of the native protein were made under these conditions varying the protein concentration (between 10 and 20 mg/mL), the drop volume (2 μL or 1.5 μL) and the drop ratio (1:1 or 2:1), using the Tableing drop method at 20 °C.

After 1 month, it was possible to observe a crystalline material in condition 4 (Figure 3.2.2.1) obtained through the hanging drop method with 20 mg/mL of HNE.

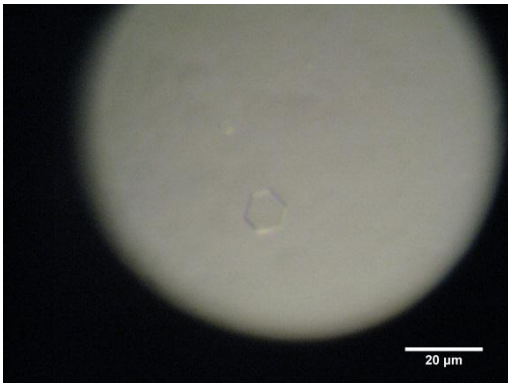


Figure 3.2.2.1 - HNE crystal obtained in condition 4, by the hanging drop method, with 20 mg/mL, at 20^{o}C.

Since these were small and few crystals, it was tried to reproduce the condition with its original composition, as well as the same condition with a small increase in the amount of precipitant (from 28% of PEG4000 to 30 %).

At the same time, both conditions were used for the co-crystallization of HNE with ligands LMC223 and LMC249 (ligands with best inhibition results in the activity assays performed for HNE inhibition). For this purpose, the protein was incubated with each of the ligands for 45 minutes at room temperature and hanging and sitting drop drops were prepared with a drop ratio, incubated protein / reservoir, of 2:1.

After one day, it was decided to proceed to the microseeding and streak seeding of the drops in question, using the crystals initially obtained, in order to try to promote nucleation. During one month of equilibrium, no crystals appeared in these drops. It was also decided to test the macroseeding technique in drops of this condition, using crystals of PPE, which resulted in the dissolution of the crystal in the drop.

Because elastase seems to be more stable in buffers with pH close to 5 (previously the protein buffer was at pH 7.5). A new protein batch was prepared in 20 mM Bis-Tris (pH 6) and 50 mM NaCl.

The Bradford method was used to confirm the protein concentration, with a calibration curve with an $R^2$of 0.9938, obtaining a concentration of 15.3 mg/mL.

Then, to check the protein purity, an SDS-Page was performed. As shown in Figure 3.2.2.2 – well 1, a band was found around 30 kDa, corresponding to the molecular weight of HNE (29.5 kDa) and small bands of higher molecular weight and low intensity, corresponding to contaminations. Through these data it is verified that the sample is in good conditions for crystallization trials. The dissolution of the pellet resulting from the centrifugation (15000 rpm during 15 min) carried out in the sample preparation step in protein buffer was also analyzed by SDS-PAGE on this gel. As shown in Figure 3.2.2.2-2, the same bands appeared, slightly more intense, which means that the protein is also present in the pellet.

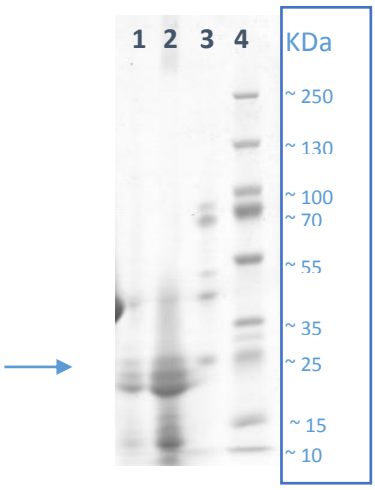


Figure 3.2.2.2 - SDS-PAGE for the HNE

(1 – Protein solution; 2 – Pellet dilution; 3 – Another protein; 4 – Marker)

After verification of the protein purity, two crystallization screens were done: Salt Rx (from Hampton) and ShotGun (from Molecular Dimensions). Of both screens, crystals only appeared in two conditions of the ShotGun screen (Figure 3.2.2.3).

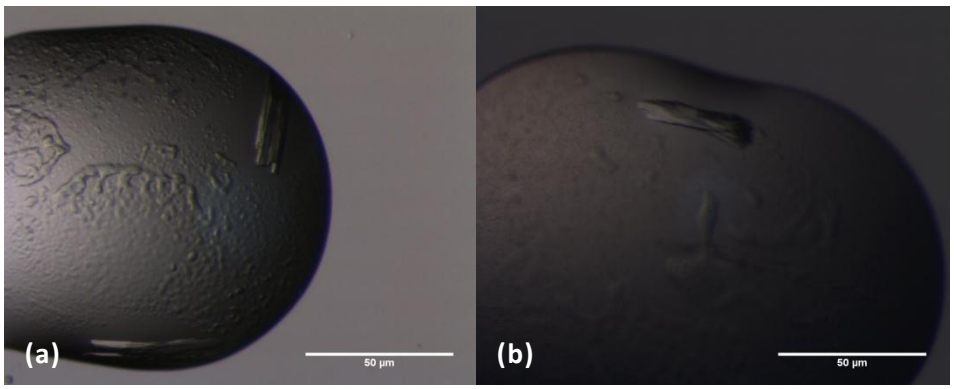
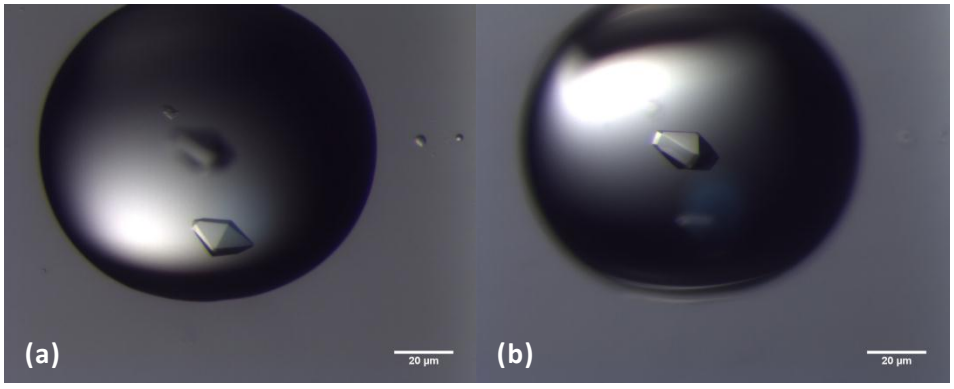


Figure 3.2.2.3 - Crystals of HNE obtained in ShotGun screen.

((a) – 0.2 M Sodium Sulfate with 20% PEG 3350 (ShotGun G2); (b) – 0.2 M Potassium/Sodium Tartrare Tetrahydrate with 20% PEG 3350 (ShotGun C4))

Through these, it was observed that the protein had a preference for conditions whose precipitant are PEGs, followed by the study of two other screens (PACT Premier and BCS, Molecular Dimension). These screens were chosen because they are specific to PEG and because they have a large number of conditions involving PEG3350 (precipitant of the conditions that obtained crystals with ShotGun. After six days of equilibration, it was possible to observe crystals under some conditions (Figure 3.2.2.4).



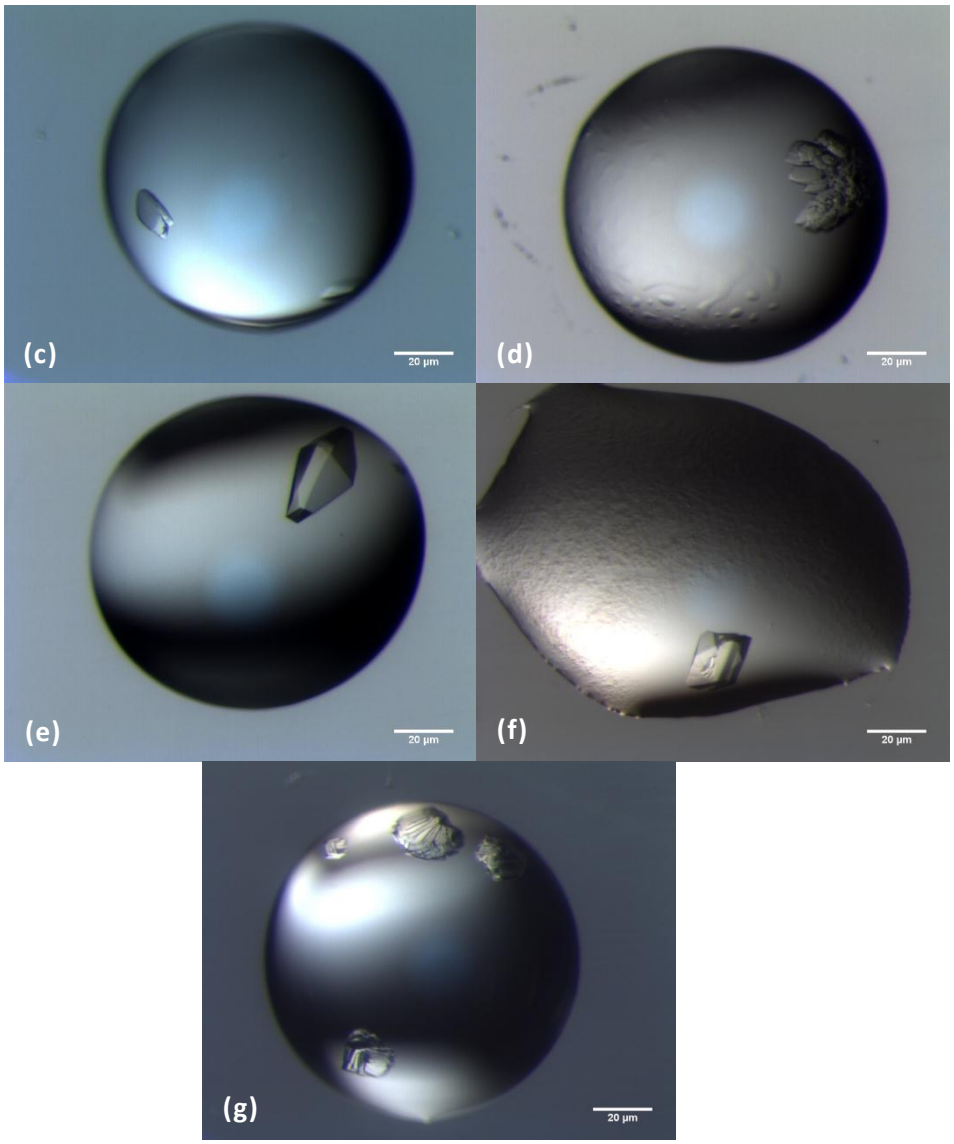


Figure 3.2.2.4 - Crystals of HNE obtained in the crystallization screens

((a – b) – 0.1 M Sodium Citrate (pH 4.5) with 20 % PEG smear high (BCS A7); (c) – 0.1 M Sodium Acetate (pH 4.5) with 22% PEG smear broad (BCS A10); (d) – 0.1 M Sodium Citrate (pH 5), 15% PEG Smear high with 0.15 M Ammonium Acetate (BCS C7); (e) – 0.1 M PIPES (pH 7), 20% PEG smear medium, 0.1 M Magnesium Chloride Hexa hydrate with 0.1 M Potassium Chloride (BCS E4); (f) – 0.1 M PIPES (pH 7), 25% PEG smear high, 0.1 M Magnesium Formate Dihydrate with 0.1 M Rubidium Chloride (BCS E7); (g) – 0.1M Bis-Tris Propane (pH 7.5), 20% PEG 3350 with 0.2 M Potassium Thiocyanate (Premier F4)).

4. Conclusions and future perspectives

The work carried out in this thesis had as main objective the structural analysis of the interaction between elastase (PPE and HNE) and several inhibitors, synthesized by group of Prof. Rui Moreira, Instituto de Investigação do Medicamento, Faculdade de Farmácia, Universidade de Lisboa.

In a first step, crystals of the PPE-inhibitor complexes were obtained under two different conditions. Of these, only the crystals grown in 70% MPD and 10 mM Sodium Phosphate Buffer (pH 5.9) diffracted, close to atomic resolution. Co-crystallization technique was more suitable for the analysis of these complexes, obtaining crystals whose electron density maps, resulting from X-ray diffraction, presented density $2F_o$-F_c near the catalytic center corresponding to the inhibitor. As this technique promotes the formation of the complex before the drop equilibrium, the probability of the ligand remaining bound in the catalytic center of the protein increases. Three-dimensional structures of PPE in complex with inhibitors LMC188, LMC240 and LMC269, were obtained by X-ray Crystallography. Analysis of the electron density around the active center of the protein, it was possible to conclude that the nucleophilic attack performed by the catalytic serine is not on the carbonyl group present in the inhibitor β-Sultam ring, but rather on the group Sulfonyl.

In a second phase, computational studies of HNE-inhibitor complexes were carried out to dock PPE inhibitors into the HNE active site, followed by energy minimization. No significant differences with the structure of PPE-inhibitor complexes was found, corroborating the use of porcine elastase as a model for the human one. Crystallization tests were started with the protein in its native state. It was concluded that the protein crystallizes more easily when dissolved in acid buffer (pH 6) than in basic buffer (pH 7.5). Several hits appeared in crystallization screens of HNE, so the optimization of these initial conditions is on-going.

As future work we intend to conclude the structural studies of HNE-inhibitor complexes, as well as to carry out structural analyzes of complexes between other serine proteases with this type of ligands, in order to verify if the mechanism of nucleophilic attack is maintained.

5. Bibliography

[1] A. Thomson and S. B. Kapadia, "The specificity of the S1 and S2 subsites of elastase," *Eur. J. Biochem.*, vol. 102, pp. 111–116, 1979.

[2] Z. Werb, M. J. Banda, J. H. McKerrow, and R. A. Sandhaus, "Elastases and elastin degradation.," *J. Invest. Dermatol.*, vol. 79 Suppl 1, p. 154s–159s, Jul. 1982.

[3] E. F. P. Ruivo, L. M. Gonåalves, L. A. R. Carvalho, R. C. Guedes, S. Hofbauer, J. A. Brito, M. Archer, R. Moreira, and S. D. Lucas, "Clickable 4-Oxo- b -lactam-Based Selective Probing for Human Neutrophil Elastase Related Proteomes," pp. 1–7, 2016.

[4] L. R. P. Areias, E. F. P. Ruivo, M. T. Duarte, R. Moreira, S. D. Lucas, and R. C. Guedes, "RSC Advances PAPER A uni fi ed approach toward the rational design of selective low nanomolar human neutrophil elastase," pp. 51717–51721, 2015.

[5] "Uniprot." [Online]. Available: http://www.uniprot.org/. [Accessed: 16-Aug-2017].

[6] W. Huang, Y. Yamamoto, Y. Li, D. Dou, K. R. Alliston, R. P. Hanzlik, T. D. Williams, and W. C. Groutas, "X-ray Snapshot of the Mechanism of Inactivation of Human Neutrophil Elastase by 1,2,5-Thiadiazolidin-3-one 1,1-Dioxide Derivatives.," *Society*, pp. 2003–2008, 2008.

[7] J. Kyte, *Structure in Protein Chemistry*, 2nd ed. Garland Science, 2006.

[8] D. Mrozek, "Formal Model of 3D Protein Structures for Functional Genomics, Comparative Bioinformatics, and Molecular Modeling," in *High-Performance Computational Solutions in Protein Bioinformatics*, 1st ed., Springer International Publishing, 2014, pp. 1–23.

[9] N. H. C. S. Silva, C. Vilela, I. M. Marrucho, C. S. R. Freire, C. Pascoal Neto, and A. J. D. Silvestre, "Protein-based materials: from sources to innovative sustainable materials for biomedical applications," *J. Mater. Chem. B*, vol. 2, no. 24, p. 3715, 2014.

[10] A. Lehninger, D. L. Nelson, and M. M. Cox, *Lehninger: Princípios de Bioquímica*, 4th ed. 2002.

[11] D. Voet and J. Voet, *Fundamentos de Bioquímica - A Vida em Nível Molecular*, 2nd ed. 2008.

[12] L. Stryer, J. L. Tymoczko, and J. M. Berg, *Bioquímica*, 5th ed. 2004.

[13] A. G. Kikhney and D. I. Svergun, "A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins," *FEBS Lett.*, vol. 589, no. 19, pp. 2570–2577, 2015.

[14] M. K. Grøftehauge, N. R. Hajizadeh, M. J. Swann, and E. Pohl, "Protein-ligand interactions investigated by thermal shift assays (TSA) and dual polarization interferometry (DPI)," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 71, pp. 36–44, 2015.

[15] J. Brito and M. Archer, "X-ray Crystallography," in *Practical Approaches to Biological Inorganic Chemistry*, 1st ed., Elsevier, Ed. 2013, pp. 217–255.

[16] B. Alberts, D. Bray, K. Hopkin, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Fundamentos da Biologia Celular*, 3rd ed. 2011.

[17] M. H. Dominiczak and J. Baynes, *Bioquímica Médica*, 2nd ed. 2000.

[18] O. Aaltonen, "Enzyme catalysis," *Chemical synthesis using supercritical fluids*. 1999.

[19] M. A. Z. Coelho, A. M. Salgado, and B. D. Ribeiro, *Tecnologia Enzimática*, 1st ed. 2008.

[20] F. J. Contesini, R. R. de Melo, and H. H. Sato, "An overview of Bacillus proteases: from production to application," *Crit. Rev. Biotechnol.*, pp. 1–14, Aug. 2017.

[21] K. Rani, R. Rana, and S. Datt, "Review On Latest Overview of Proteases," *Int. J. Curr. Life Sci.*, vol. 2, no. 1, pp. 12–18, 2012.

[22] "Sigma-Aldrich." [Online]. Available: http://www.sigmaaldrich.com/catalog/product/roche/pronro?lang=pt&region=PT. [Accessed: 23-Sep-2017].

[23] T. G. Villa and M. Vinas, *New Weapons to Control Bacterial Growth*. Cham: Springer International Publishing, 2016.

[24] "MEROPS." [Online]. Available: https://www.ebi.ac.uk/merops/about/classification.shtml. [Accessed: 23-Sep-2017].

[25] H. Neurath, "Evolution of proteolytic enzymes," *Science (80-.).*, vol. 224, no. 4647, pp. 350–357, Apr. 1984.

[26] J. J. Neitzel, "Enzyme Catalysis: The Serine Proteases," *Nat. Educ.*, vol. 3, no. 9, p. 21, 2010.

[27] "BIOCHEMISTRY ONLINE." [Online]. Available: http://biochem-vivek.tripod.com/id45.html. [Accessed: 23-Sep-2017].

[28] W. Bode, E. Meyer, and J. C. Powers, "Human leukocyte and porcine pancreatic elastase: X-ray crystal structures, mechanism, substrate specificity, and mechanism-based inhibitors.," *Biochemistry*, vol. 28, no. 5, pp. 1951–63, 1989.

[29] V. Worthington, K., and Worthington, "Worthington Enzyme Manual," *Worthington Biochemical Corporation*, 2011. [Online]. Available: http://www.worthington-biochem.com/pap/default.html. [Accessed: 17-Apr-2017].

[30] M. S. Weiss, S. Panjikar, E. Nowak, and P. A. Tucker, "Metal binding to porcine pancreatic elastase: Calcium or not calcium," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 58, no. 9, pp. 1407–1412, 2002.

[31] C. Vergelli, I. A. Schepetkin, L. Crocetti, A. Iacovone, M. P. Giovannoni, G. Guerrini, A. I. Khlebnikov, S. Ciattini, G. Ciciani, and M. T. Quinn, "Isoxazol-5(2H)-one: a new scaffold for potent human neutrophil elastase (HNE) inhibitors," *J. Enzyme Inhib. Med. Chem.*, vol. 32, no. 1, pp. 821–831, 2017.

[32] E. Meyer, G. Cole, R. Radhakrishnan, and O. Epp, "Structure of native porcine

pancreatic elastase at 1.65 A resolutions.," *Acta Crystallogr. B.*, vol. 44, pp. 26–38, 1988.

[33] G. Hansen, H. Gielen-Haertwig, P. Reinemer, D. Schomburg, A. Harrenga, and K. Niefind, "Unexpected active-site flexibility in the structure of human neutrophil elastase in complex with a new dihydropyrimidone inhibitor," *J. Mol. Biol.*, vol. 409, no. 5, pp. 681–691, 2011.

[34] G. Rhodes, *Crystallography Made Crystal Clear*, 3rd ed. Academic Press, 2006.

[35] S. Khurshid, E. Saridakis, L. Govada, and N. E. Chayen, "Porous nucleating agents for protein crystallization," *Nat. Protoc.*, vol. 9, no. 7, pp. 1621–1633, 2014.

[36] M. J. Romão, "Cristalografia de Proteinas: metodologias e aplicações em Bioquímica," *Boletim de Biotecnologia*, vol. 53. pp. 18–36, 1996.

[37] "School of Chemistry University of Glasgow." [Online]. Available: www.chem.gla.ac.uk. [Accessed: 19-Jun-2017].

[38] S. Mobilio, F. Boscherini, and C. Meneghini, "Synchrotron radiation: Basics, methods and applications," *Synchrotron Radiat. Basics, Methods Appl.*, pp. 1–799, 2015.

[39] "Research Gate." [Online]. Available: https://www.researchgate.net/figure/221916990_fig3_Fig-4-Schematic-diagram-of-Synchrotron-Soleil-The-circular-ring-is-the-synchrotron-a. [Accessed: 26-Jun-2017].

[40] L. N. D. L. SINCROTRON, "O QUE É UMA FONTE DE LUZ SÍNCROTRON?" [Online]. Available: http://lnls.cnpem.br/o-lnls/o-que-e-uma-fonte-de-luz-sincrotron/. [Accessed: 22-Aug-2017].

[41] M. Würtele, M. Hahn, K. Hilpert, and W. Höhne, "Atomic resolution structure of native porcine pancreatic elastase at 1.1Å," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 56, no. 4, pp. 520–523, Apr. 2000.

[42] T. F. Oliveira, J. Mulchande, R. Moreira, J. Iley, and M. Archer, "Crystallization and preliminary diffraction studies of porcine pancreatic elastase in complex with a novel inhibitor.," *Protein Pept. Lett.*, vol. 14, no. 1, pp. 93–5, 2007.

[43] G. L. Gilliland, M. Tung, D. M. Blakeslee, and J. E. Ladner, "Biological Macromolecule Crystallization Database, Version 3.0: new features, data and the NASA archive for protein crystal growth data.," *Acta Crystallogr. D. Biol. Crystallogr.*, vol. 50, no. Pt 4, pp. 408–13, 1994.

[44] K. Hilpert, H. Wessner, C. Scholz, P. Scheerer, R. Volkmer-Engert, and N. Kraubeta, "Crystallization and preliminary x-ray analysis of complexes of porcine pancreatic elastase with two natural inhibitors.," *Protein Pept. Lett.*, vol. 11, no. 4, pp. 393–399, 2004.

[45] T. Kinoshita, A. Yamaguchi, and T. Tada, "Tris(hydroxymethyl)aminomethane induces conformational change and crystal-packing contraction of porcine pancreatic elastase," *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.*, vol. 62, no. 7, pp. 623–626, 2006.

[46] M. Tsunemi, Y. Matsuura, S. Sakakibara, and Y. Katsube, "Crystal Structure of an

Elastase-Specific Inhibitor Elafin Complexed with Porcine Pancreatic Elastase Determined at 1.9 Å Resolution †," *Biochemistry*, vol. 35, no. 36, pp. 11570–11576, Jan. 1996.

[47] A. Dementiev, J. Dobó, and P. G. W. Gettins, "Active Site Distortion Is Sufficient for Proteinase Inhibition by Serpins," *J. Biol. Chem.*, vol. 281, no. 6, pp. 3452–3457, Feb. 2006.

[48] M. A. Navia, B. M. McKeever, J. P. Springer, T. Y. Lin, H. R. Williams, E. M. Fluder, C. P. Dorn, and K. Hoogsteen, "Structure of human neutrophil elastase in complex with a peptide chloromethyl ketone inhibitor at 1.84-A resolution.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 86, no. 1, pp. 7–11, 1989.

[49] M. Koizumi, A. Fujino, K. Fukushima, T. Kamimura, and M. Takimoto-Kamimura, "Complex of human neutrophil elastase with 1/2SLPI," *J. Synchrotron Radiat.*, vol. 15, no. 3, pp. 308–311, 2008.

[50] F. Von Nussbaum, V. M. Li, D. Meibom, S. Anlauf, M. Bechem, M. Delbeck, M. Gerisch, A. Harrenga, D. Karthaus, D. Lang, K. Lustig, J. Mittendorf, M. Schäfer, S. Schäfer, and J. Schamberger, "Potent and Selective Human Neutrophil Elastase Inhibitors with Novel Equatorial Ring Topology: In vivo Efficacy of the Polar Pyrimidopyridazine BAY-8040 in a Pulmonary Arterial Hypertension Rat Model," *ChemMedChem*, vol. 11, no. 2, pp. 199–206, 2016.

[51] "ThermoFisher." [Online]. Available: https://www.thermofisher.com/order/catalog/product/26619. [Accessed: 24-Sep-2017].

[52] C. Vonrhein, C. Flensburg, P. Keller, A. Sharff, O. Smart, W. Paciorek, T. Womack, and G. Bricogne, "Data processing and analysis with the autoPROC toolbox," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 67, no. 4, pp. 293–302, 2011.

[53] W. Kabsch, "Xds," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 66, no. 2, pp. 125–132, 2010.

[54] P. Evans, "Scaling and assessment of data quality," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 62, no. 1, pp. 72–82, 2006.

[55] P. R. Evans, "An introduction to data reduction: Space-group determination, scaling and intensity statistics," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 67, no. 4, pp. 282–292, 2011.

[56] Diamond Light Source, "Xia2 - Parameters," 2015. [Online]. Available: https://xia2.github.io/parameters.html. [Accessed: 22-Aug-2017].

[57] "XSCALE." [Online]. Available: http://xds.mpimf-heidelberg.mpg.de/html_doc/xscale_program.html.

[58] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, and R. J. Read, "Phaser crystallographic software," *J. Appl. Crystallogr.*, vol. 40, no. 4, pp. 658–674, 2007.

[59] S. Hofbauer, J. A. Brito, J. Mulchande, P. Nogly, M. Pessanha, R. Moreira, and M. Archer, "Crystal structure of Porcine Pancreatic Elastase (PPE) in complex with the

novel inhibitor JM102."

[60] P. Emsley and K. Cowtan, "Coot : model-building tools for molecular graphics," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 60, no. 12, pp. 2126–2132, Dec. 2004.

[61] P. Emsley, B. Lohkamp, W. G. Scott, and K. Cowtan, "Features and development of Coot," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 66, no. 4, pp. 486–501, Apr. 2010.

[62] R. a. Laskowski, M. W. MacArthur, D. S. Moss, and J. M. Thornton, "PROCHECK: a program to check the stereochemical quality of protein structures," *J. Appl. Crystallogr.*, vol. 26, no. November, pp. 283–291, 1993.

[63] R. W. W. Hooft, G. Vriend, C. Sander, and E. E. Abola, "Errors in protein structures," *Nature*, vol. 381, no. 6580. pp. 272–272, 1996.

[64] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, and K. S. Wilson, "Overview of the CCP 4 suite and current developments," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 67, no. 4, pp. 235–242, Apr. 2011.

[65] M. Beardsell, P. S. Hinchliffe, J. M. Wood, M. I. Page, R. C. Wilmouth, and C. J. Schofield, "β-Sultams—A novel class of serine protease inhibitors," *Chem. Commun.*, no. 5, pp. 497–498, 2001.

[66] A. A. Vagin, R. A. Steiner, A. A. Lebedev, L. Potterton, S. McNicholas, F. Long, and G. N. Murshudov, "REFMAC 5 dictionary: organization of prior chemical knowledge and guidelines for its use," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 60, no. 12, pp. 2184–2195, Dec. 2004.

[67] "Hyperphysics." [Online]. Available: http://hyperphysics.phy-astr.gsu.edu/hbase/quantum/bragg.html. [Accessed: 19-Jun-2017].

[68] "Tutorcircle." [Online]. Available: http://physics.tutorcircle.com/waves/wave-interference.html. [Accessed: 19-Jun-2017].

[69] "ESRF." [Online]. Available: http://www.esrf.eu/home/events/conferences/2016/summerschool-2016.html. [Accessed: 26-Jun-2017].

[70] "Diamond." [Online]. Available: http://www.diamond.ac.uk/Home/About.html. [Accessed: 26-Jun-2017].

6. Appendix

Table A. 1 - Total data collection and Processing ESRF (ID23-1)

Inhibitor	**LMC188 (Co-crystallization)**
Detector	Pilatus 6M
Beam wavelength (Å)	0.9763
Number of images	1100
Oscillation range (°)	0.1
Exposure Time (s)	0.037
Space Group	$P\ 2_1\ 2_1\ 2_1$ (16)
Unit cell	a = 50.56 Å b = 57.56 Å c= 74.60 Å α=β=γ= 90°

	Overall	**Inner Shell**	**Outer Shell**
Low resolution limit (Å)	45.570	45.570	1.400
High resolution limit (Å)	1.330	4.200	1.330
R_{merge}	0.059	0.031	0.915
R_{meas}	0.068	0.035	1.130
R_{pim}	0.032	0.016	0.649
Total number of observations	171207	6227	10971
Total number of unique reflection	45719	1599	4259
$\langle I/\sigma(I)\rangle$	10.2	34.8	0.8
Completeness (%)	90.1	93.2	58.3
Multiplicity	3.7	3.9	2.6
$CC_{1/2}$	0.998	0.998	0.375

Inhibitor	**LMC188 (Soaking)**
Detector	Pilatus 6M
Beam wavelength (Å)	0.9763
Number of images	580
Oscillation range (°)	0.2
Exposure Time (s)	0.037
Space Group	$P\ 2_1\ 2_1\ 2$ (16)
Unit cell	a = 74.20 Å b = 49.97 Å c= 57.81 Å α=β=γ= 90°

	Overall	**Inner Shell**	**Outer Shell**
Low resolution limit (Å)	41.447	41.447	1.276
High resolution limit (Å)	1.254	3.403	1.254
R_{merge}	0.059	0.034	0.707
R_{meas}	0.067	0.039	0.805
R_{pim}	0.030	0.017	0.375
Total number of observations	243845	12325	11646
Total number of unique reflection	58376	2932	2833
$\langle I/\sigma(I)\rangle$	11.8	32.3	2.2
Completeness (%)	97.9	91.5	97.0
Multiplicity	4.2	4.2	4.1
$CC_{1/2}$	0.998	0.998	0.728

Inhibitor	LMC100 (Soaking)
Detector	Pilatus 6M
Beam wavelength (Å)	0.9763
Number of images	990
Oscillation range (°)	0.1
Exposure Time (s)	0.037
Space Group	$P\ 2_1\ 2_1\ 2$ (16)
Unit cell	a = 57.73 Å b = 74.64 Å c= 50.30 Å α=β=γ= 90°

Inhibitor	LMC211 (Soaking)
Detector	Pilatus 6M
Beam wavelength (Å)	0.9763
Number of images	734
Oscillation range (°)	0.15
Exposure Time (s)	0.037
Space Group	$P\ 2_1\ 2_1\ 2$ (16)
Unit cell	a = 50.51 Å b = 57.85 Å c= 74.82 Å α=β=γ= 90°

	Overall	Inner Shell	Outer Shell
Low resolution limit (Å)	45.665	45.665	1.187
High resolution limit (Å)	1.167	3.167	1.167
R_{merge}	0.051	0.033	0.5311
R_{meas}	0.059	0.038	0.627
R_{pim}	0.028	0.018	0.324
Total number of observations	263237	13288	11995
Total number of unique reflection	72862	3661	3611
$\langle I/\sigma(I)\rangle$	11.1	30.6	2.1
Completeness (%)	97.7	91.5	98.1
Multiplicity	3.6	3.6	3.3
$CC_{1/2}$	0.999	0.998	0.722

	Overall	Inner Shell	Outer Shell
Low resolution limit (Å)	45.766	45.766	1.288
High resolution limit (Å)	1.266	3.436	1.266
R_{merge}	0.046	0.032	0.476
R_{meas}	0.052	0.037	0.552
R_{pim}	0.025	0.018	0.272
Total number of observations	217043	11273	5837
Total number of unique reflection	53621	2888	1595
$\langle I/\sigma(I)\rangle$	13.3	32.2	2.2
Completeness (%)	90.6	90.3	54.5
Multiplicity	4.0	3.9	3.7
$CC_{1/2}$	0.999	0.998	0.807

Table A. 2 -Total data collection Diamond (Beamline I03) and ESRF (ID30A-3)

Inhibitor	**LMC100 (Co-crystallization)**
Detector	Pilatus3 6M
Beam wavelength (Å)	0.9762
Number of images	2780
Oscillation range (°)	0.05
Exposure Time (s)	0.01
Space Group	$P\ 2_1\ 2_1\ 2_1$ (16)
Unit cell	a = 51.63 Å b = 57.31 Å c= 74.54 Å α=β=γ= 90°

Inhibitor	**LMC211 (Co-crystallization)**
Detector	Pilatus3 6M
Beam wavelength (Å)	0.9762
Number of images	2980
Oscillation range (°)	0.05
Exposure Time (s)	0.02
Space Group	$P\ 2_1\ 2_1\ 2_1$ (16)
Unit cell	a = 50.61 Å b = 57.38 Å c= 74.62 Å α=β=γ= 90°

	Overall	**Inner Shell**	**Outer Shell**
Low resolution limit (Å)	38.360	38.380	1.220
High resolution limit (Å)	1.220	3.310	1.240
R_{merge}			
R_{meas}	0.058	0.021	1.125
R_{pim}	0.025	0.009	0.621
Total number of observations	312466	16993	8843
Total number of unique reflection	66074	3552	3117
$\langle I/\sigma(I)\rangle$	13.2	63.4	1
Completeness (%)	99.4	99.5	95.5
Multiplicity	4.7	4.8	2.8
$CC_{1/2}$	0.999	0.999	0.484

	Overall	**Inner Shell**	**Outer Shell**
Low resolution limit (Å)	37.310	37.330	1.320
High resolution limit (Å)	1.300	3.530	1.300
R_{merge}			
R_{meas}	0.066	0.045	1.278
R_{pim}			
Total number of observations	287678	15042	14364
Total number of unique reflection	55054	2967	2722
$\langle I/\sigma(I)\rangle$	12.2	36.6	1.4
Completeness (%)	99.7	99.9	99.9
Multiplicity	5.2	5.1	5.3
$CC_{1/2}$	1	1	0.7

Inhibitor	LMC240 (Co-crystallization)
Detector	Pilatus3 6M
Beam wavelength (Å)	0.9762
Number of images	1090
Oscillation range (°)	0.1
Exposure Time (s)	0.029
Space Group	$P\ 2_1\ 2_1\ 2_1$ (16)
Unit cell	a = 51.10 Å b = 57.71 Å c= 74.64 Å α=β=γ= 90°

Inhibitor	LMC269 (Co-crystallization)
Detector	Eiger 4M
Beam wavelength (Å)	0.9677
Number of images	600
Oscillation range (°)	0.2
Exposure Time (s)	0.01
Space Group	$P\ 2_1\ 2_1\ 2_1$ (16)
Unit cell	a = 50.74 Å b = 57.71 Å c= 74.65 Å α=β=γ= 90°

	Overall	Inner Shell	Outer Shell
Low resolution limit (Å)	31.340	31.350	1.300
High resolution limit (Å)	1.280	3.470	1.280
R_{merge}			
R_{meas}	0.041	0.025	1.115
R_{pim}			
Total number of observations	221694	11400	9997
Total number of unique reflection	57349	3068	2813
$\langle I/\sigma(I)\rangle$	14.8	51.9	1.2
Completeness (%)	99.6	99.1	99.6
Multiplicity	3.9	3.7	3.6
$CC_{1/2}$	1	1	0.5

	Overall	Inner Shell	Outer Shell
Low resolution limit (Å)	30.066	30.066	1.402
High resolution limit (Å)	1.378	3.739	1.378
R_{merge}	0.042	0.022	0.451
R_{meas}	0.047	0.024	0.534
R_{pim}	0.022	0.011	0.276
Total number of observations	197062	10453	5795
Total number of unique reflection	44471	2342	1802
$\langle I/\sigma(I)\rangle$	16.3	50.1	2.3
Completeness (%)	96.8	94.2	79.2
Multiplicity	4.4	4.5	3.2
$CC_{1/2}$	0.999	0.999	0.822