

**CRYSTALLOGRAPHIC STUDIES OF  
COAGULATION FACTOR XII**

**Rosa Manna**

**Thesis submitted to the University of Nottingham for the  
degree of Doctor of Philosophy**

**September 2015**

## DECLARATIONS

This thesis is submitted to the University of Nottingham in support of my application for the degree of Doctor of Philosophy in Pharmacy. It has not been submitted in any previous application for any degree and it has been composed by myself.

The cloning of  $\beta$ FXIIG570R mutant in pMT-PURO vector and its expression in *Drosophila* Expression System (DES) was performed by Dr. Monika Pathak. The kinetic characterization of all the constructs used in this thesis was performed with the help of Kareem Hamad Badaldrin, PhD student in the group of Dr L. Dekker.

## ABSTRACT

Coagulation Factor XII (FXII) is an important protein involved in the initiation of the intrinsic pathway of the coagulation cascade. Recent studies suggested that FXII may play a role in pathological thrombosis without compromising physiological hemostasis. For this reason, the inhibition of FXII could represent a new and selective strategy for preventing stroke and other thromboembolic diseases. Since there is no structure available for the protease domain of FXII, the knowledge of the overall domain structure can be significantly helpful in the development of inhibitors with specific selectivity for the target.

Three different constructs were made:  $\beta$ FXIIa,  $\beta$ FXIIG570R, a missense mutations that causes a reduced activity of FXIIa, and  $\beta$ FXIIa fused with Maltose Binding Protein (MBP) at the N-terminus. An expression and purification protocol was established for all the proteins. In order to assess the activity of the recombinant protein, the hydrolysis of chromogenic substrate S2302 was measured and a comparison with commercial  $\alpha$ FXIIa and  $\beta$ FXIIa was performed. The kinetic parameters ( $k_{\text{cat}}$ ,  $K_{\text{M}}$ ,  $k_{\text{cat}}/K_{\text{M}}$ ,  $V_{\text{max}}$ ) indicated that the catalytic behaviour of recombinant MBP- $\beta$ FXIIa and  $\beta$ FXIIa is essentially identical to that of commercial  $\beta$ FXIIa purified from plasma. Therefore, the protein expressed and purified from insect cell system is catalytically competent.

Large efforts were devoted to the identification of possible crystallization conditions for the expressed and purified constructs, in complex with different

inhibitors, both small molecules and macromolecular inhibitors. Diffracting crystals were obtained for MBP- $\beta$ FXIIa in complex with D-Phe-Pro-Arg-chloromethylketone (PPACK). A complete data set was collected at 4 Å. Notwithstanding the low resolution of the data, the structural analysis of key elements and the comparison with the zymogen confirmed that recombinant protein is in the active conformation. Moreover, a first analysis of the structure and of the lattice packing enabled to understand some key differences between the interactions that PPACK forms when in complex to FXIIa and with thrombin, thus possible explaining the lower affinity of PPACK for FXIIa. This is the first crystallographic structure of FXIIa and it represents a first step in the study of protein-inhibitor interactions from structural consideration.

$\beta$ FXII construct was also cloned in a different vector for the expression in *E. coli*. The cleavage of TF tag was performed using three different enzymes (HRV-3C protease, FXa and thrombin). The best results were obtained, using the thrombin and setting the reaction overnight at room temperature. However, it was not possible to continue the process, since the protein was lost after the cleavage. This could be due to a poor stability of the protein. For that reason, a different strategy was developed, consisting of co-purification of  $\beta$ FXII with Ecotin. The presence of the inhibitor was thought to stabilize the protein during the cleavage. However, the gel filtration revealed that the complex was not formed.

The possibility to investigate at atomic level the interactions of the inhibitors with FXII could allow an understanding as to how the substrate-inhibitor recognition mechanism work. This could therefore represent a good starting

point for the development of new inhibitors that have higher specificity for FXII.

## **Acknowledgments**

I would like to express my gratitude to my main supervisor, Prof. Jonas Emsley for the possibility to complete my PhD and so widen my knowledge and skills.

I am immensely grateful to him for his support over the past four years. I would also like to thank my other two supervisors, Prof P. Fisher and Dr. L Dekker for their advices and guidance.

I would like to thanks all the people working in C84 lab. Anna, Buba and Susan: you've been great friends and supporters, I will always remember you.

I am grateful to my friends Luke, Ula, Victoria, Paolo and Danilo: I never felt alone in the last months.

I would also like to acknowledge the British Heart Foundation (BHF) for the financial support to complete my PhD.

I would like to dedicate this thesis to several people: first of all to my family, my mum Raffaella, my dad Feliciano and my sisters Lina and Serena for all their great love and support. I also dedicate this thesis to Franky, for constantly believing in me and helping me always, reading the thesis and giving me tips and advice, for sharing this and other adventures with me, to his family, and to our son Raffaele, my best experiment ever, who made the past years an incredible journey.

# TABLE OF CONTENTS

<b>Declaration.....</b>	<b>ii</b>
<b>Abstract.....</b>	<b>iii</b>
<b>Acknowledgments.....</b>	<b>v</b>
<b>Table of Contents.....</b>	<b>vi</b>
<b>List of Figures.....</b>	<b>xii</b>
<b>List of Tables.....</b>	<b>xvi</b>
<b>Abbreviations and Symbols.....</b>	<b>xviii</b>
<b>CHAPTER 1 INTRODUCTION.....</b>	<b>1</b>
1.1 Hemostasis and Coagulation.....	2
1.1.1 Endothelium.....	3
1.1.2 Platelets.....	5
1.2 Coagulation.....	6
1.2.1 Coagulation Cascade.....	6
1.3 Immunothrombosis.....	11
1.4 Regulation of the coagulation cascade.....	14
1.5 Coagulation factors deficiencies, thrombosis and pharmacological strategies.....	17
1.6 Coagulation Factor FXII: primary structure and function.....	20
1.6.1 Structure of FXII.....	22
1.6.2 <i>In vitro</i> and <i>in vivo</i> activators of FXII.....	30
1.6.3 FXII and thrombosis.....	32
1.7 Mutation in F12 gene and their effect on FXII activity.....	32
1.8 Inhibition of Coagulation Factor FXII.....	34
1.8.1 Inhibition of FXII by Cloro Methyl Ketone (CMK) peptides: PPACK and PCK.....	35
1.8.2 Inhibition of FXII by macromolecular biological inhibitors: Ecotin.....	37
1.8.3 Inhibition of FXII by macromolecular biological inhibitors: CTI.....	39

1.9 Objectives.....	40
<b>CHAPTER 2: THE THEORY OF X-RAY CRYSTALLOGRAPHY.....</b>	<b>42</b>
2.1 Introduction.....	42
2.2 Protein crystallization.....	43
2.2.1 Parameters affecting the crystallization process.....	44
2.2.2 Crystallization methods.....	46
2.2.3 Approaches to find crystallization conditions.....	47
2.3 Crystal structure.....	49
2.4 Principle of X-ray diffraction by single crystal.....	50
2.4.1 Scattering by two electrons.....	51
2.4.2 Scattering by an atom.....	52
2.4.3 Scattering by a unit cell.....	52
2.4.4 Scattering by a crystal.....	53
2.4.5 Bragg's Law.....	54
2.4.6 The reciprocal lattice.....	56
2.4.7 Bragg's law in the reciprocal space and the Ewald construction.....	57
2.5 Instrumentation.....	58
2.6 Data collection.....	59
2.6.1 Indexing.....	60
2.6.2 Scaling.....	60
2.7 The phase problem.....	60
2.7.1 The Patterson function.....	62
2.7.2 The Molecular Replacement (MR).....	63
2.8 Model building, structure refinement.....	64
2.8.1 Electron density modification.....	64
2.8.2 Model building.....	65
2.8.3 Model refinement.....	66

2.9 Validation.....	67
2.9.1 Ramachandran plot.....	68
2.9.2 Stereochemistry validation.....	68
<b>CHAPTER 3: MATERIAL AND METHODS.....</b>	<b>70</b>
3.1 General methods.....	70
3.1.1 Bacterial transformation.....	70
3.1.2 Recovery of the plasmid DNA by Miniprep and sequencing.....	71
3.1.3 Agarose gel electrophoresis.....	71
3.1.4 SDS polyacrylamide gel electrophoresis.....	72
3.1.5 Western Blot.....	74
3.1.6 Determination of DNA and protein concentration.....	74
3.2 Vectors.....	75
3.2.1 pMT-PURO vector.....	75
3.2.2 pCold-TF vector.....	76
3.2.3 pET-26b(+) vector.....	78
3.3 Tandem fixed arm MBP/SER strategy.....	79
3.4 Polymerase Chain Reaction (PCR).....	81
3.4.1 PCR for the tandem fixed-arm MBP- $\beta$ FXII system.....	81
3.4.2 PCR for $\beta$ FXIIa in pMT-PURO vector.....	84
3.4.3 Ligation of MBP- $\beta$ FXIIa in pMT-PURO vector.....	86
3.4.4 Ligation of $\beta$ FXIIa in pMT-PURO vector.....	87
3.5 Expression of the constructs.....	88
3.5.1 Expression of MBP- $\beta$ FXIIa (AAAAS) and $\beta$ FXIIa cloned in pMT-PURO vector into insect cells.....	88
3.5.2 Expression of $\beta$ FXII and Ecotin in Origami2 cells for <i>E. coli</i> expression.....	89
3.6 Purification.....	91
3.6.1 Purification of MBP- $\beta$ FXIIa from <i>Drosophila</i> Expression System (DES).....	91

3.6.2 Purification of $\beta$ FXIIa from DES.....	91
3.6.3 Purification of $\beta$ FXIIG570R from DES.....	92
3.6.4 Expression and purification of $\beta$ FXII in pCold-TF vector.....	93
3.6.5 Expression and purification of Ecotin in pET26b(+) vector.....	94
3.6.6 Co-purification of $\beta$ FXII and Ecotin from <i>E. coli</i> .....	94
3.6.7 Cleavage of Trigger Factor from $\beta$ FXII in pCOLD-TF using HRV-3C protease and FXa.....	98
3.6.8 Cleavage of Trigger Factor from $\beta$ FXII in pCOLD-TF using thrombin.....	98
3.7 Preparation of the protein-inhibitor complex.....	99
3.7.1 Complex with the CMK peptides: PPACK and PCK.....	99
3.7.2 Complex with serine protease inhibitor: Ecotin.....	100
3.7.3 Complex with serine protease inhibitor: CTI.....	100
3.8 Crystallization.....	100
3.8.1 Search for crystallization conditions.....	100
3.8.2 Crystal optimization for (MBP- $\beta$ FXIIa)-PPACK using the incomplete factorial method and the micro-seeding technique.....	101
3.8.3 Crystal optimization for (MBP- $\beta$ FXIIa)-Ecotin using the incomplete factorial method.....	107
3.9 Kinetic assay.....	108
3.10 Data collection and structure determination for (MBP- $\beta$ FXIIa)-PPACK.....	109
<b>CHAPTER 4: EXPRESSION AND CHARACTERIZATION OF MBP-<math>\beta</math>FXIIa.....</b>	<b>111</b>
4.1 Cloning of MBP- $\beta$ FXIIa in pMT-PURO vector and expression in DES system.....	111
4.2 Purification of MBP- $\beta$ FXIIa.....	113
4.3 Kinetic characterization of MBP- $\beta$ FXIIa.....	117
4.4 Crystallization of MBP- $\beta$ FXIIa.....	120
4.5 Co-crystallization with CMK inhibitors: PCK and PPACK.....	123
4.5.1 Co-crystallization with PCK.....	123

4.5.2 Co-crystallization with PPACK.....	123
4.6 Crystal structure of (MBP- $\beta$ FXIIa)-PPACK.....	128
4.6.1 Crystal diffraction and data collection.....	128
4.6.2 Data processing and scaling.....	130
4.6.3 Structure determination.....	131
4.6.4 Model building and refinement.....	132
4.6.5 (MBP- $\beta$ FXIIa)-PPACK structure.....	134
4.6.6 Crystal structure of $\beta$ FXIIa-PPACK.....	137
4.7 PPACK-induced dimerization.....	147
4.8 Expression and Purification of Ecotin.....	148
4.9 Complex with serine protease inhibitors: Ecotin and CTI.....	150
4.9.1 Complex of MBP- $\beta$ FXIIa with Ecotin.....	150
4.9.2 Complex of MBP- $\beta$ FXIIa with CTI.....	152
4.10 Crystallization of MBP- $\beta$ FXIIa with serine protease inhibitors.....	154
4.10.1 Crystallization of MBP- $\beta$ FXIIa in complex with Ecotin.....	154
4.10.2 Crystallization of MBP- $\beta$ FXIIa in complex with CTI.....	154
4.11 Summary.....	154
<b>CHAPTER 5: ENZYME KINETICS CHARACTERIZATION AND CRYSTALLIZATION OF <math>\beta</math>FXIIa AND <math>\beta</math>FXIIG570R.....</b>	<b>161</b>
5.1 Cloning of $\beta$ FXIIa in pMT-PURO vector and expression in DES system.....	161
5.2 Purification.....	162
5.2.1 Purification of $\beta$ FXIIa from DES.....	162
5.2.2 Purification of $\beta$ FXIIG570R from DES.....	165
5.3 Kinetic characterization of $\beta$ FXIIa and $\beta$ FXIIG570R.....	169
5.4 Crystallization of $\beta$ FXIIa and $\beta$ FXIIG570R.....	172
5.5 Co-crystallization of $\beta$ FXIIa and $\beta$ FXIIG570R with CMK inhibitor PPACK.....	174
5.6 Complex of $\beta$ FXIIG570R with serine protease inhibitor Ecotin.....	174

5.7 Crystallization of $\beta$ FXIIG570R-Ecotin complex.....	178
5.8 Summary.....	179
<b>CHAPTER 6: EXPRESSION AND PURIFICATION OF <math>\beta</math>FXII FROM <i>E.coli</i>.....</b>	<b>185</b>
6.1 Expression of $\beta$ FXII and purification.....	185
6.2 Cleaving the Trigger Factor tag using HRV-3C protease, FXa and Thrombin.....	191
6.3 Co-purification.....	194
6.4 Summary.....	196
<b>CHAPTER 7: CONCLUDING REMARKS AND PERSPECTIVE.....</b>	<b>199</b>
<b>CHAPTER 8: BIBLIOGRAPHY.....</b>	<b>208</b>

## List of Figures

1.1 Endothelial hemostatic function.....	4
1.2 Platelet plug formation.....	5
1.3 The contact activation pathway and the coagulation cascade.....	8
1.4 Immunothrombosis promoted by NETs.....	14
1.5 Regulation of the coagulation cascade.....	17
1.6 Different forms of human FXII.....	22
1.7 Enzymatic mechanism for serine proteases.....	23
1.8 FXII domains.....	25
1.9 Cartoon diagram of the protease domain of FXII.....	29
1.10 FXII-driven contact pathways.....	31
1.11 Binding mode of PPACK.....	35
1.12 Chemical structure of PCK.....	36
1.13 Cartoon representation of Ecotin structure.....	39
1.14 Cartoon representation of CTI.....	40
2.1 Crystallization phase diagram.....	44
2.2 Vapour diffusion method of protein crystallization.....	47
2.3 Unit cell in the crystal lattice.....	50
2.4 Scattering by two electrons.....	52
2.5 Bragg diffraction.....	56
2.6 The Ewald sphere construction.....	58
3.1 Map of pMT-PURO .....	76
3.2 Map of pCold-TF....	77
3.3 Map of pET-26b(+) vector.....	78

<b>3.4</b> Cartoon representation of the crystal structure of Maltose Binding Protein (PDB code: 1HSJ).....	80
<b>3.5</b> Schematic representation of the linker region.....	81
<b>3.6</b> Schematic representation of the constructs of $\beta$ FXIIa having an MBP fusion tag at the N-terminus.....	82
<b>3.7</b> Schematic representation of $\beta$ FXIIa construct.....	85
<b>4.1</b> 1% Agarose gel showing PCR products from PCR amplification.....	111
<b>4.2</b> Western Blot of MBP- $\beta$ FXIIa with linker AAAAS using Anti 6X-His tag primary antibody.....	112
<b>4.3</b> Purification of MBP- $\beta$ FXIIa by anion exchange chromatography.....	114
<b>4.4</b> Purification of MBP- $\beta$ FXIIa by affinity chromatography.....	115
<b>4.5</b> Schematic representation of the Cys340-Cys467 disulfide bridge.....	116
<b>4.6</b> Purification of MBP- $\beta$ FXIIa by size exclusion chromatography.....	117
<b>4.7</b> S2302 time depend assay for MBP- $\beta$ FXIIa.....	118
<b>4.8</b> Activity of MBP- $\beta$ FXIIa against various concentration of S2302.....	119
<b>4.9</b> Crystals of (MBP- $\beta$ FXIIa)-PPACK.....	125
<b>4.10</b> Crystals of (MBP- $\beta$ FXIIa)-PPACK grown from the incomplete factorial grid.....	127
<b>4.11</b> Crystals of (MBP- $\beta$ FXIIa)-PPACK grown from the finer grid.....	128
<b>4.12</b> Needle crystal of (MBP- $\beta$ FXIIa)-PPACK frozen in a loop.....	129
<b>4.13</b> Crystal structure of (MBP- $\beta$ FXIIa)-PPACK.....	134
<b>4.14</b> $F_O-F_C$ electron density map for maltose residue.....	135
<b>4.15</b> Intermolecular interactions of the MBP- $\beta$ FXIIa unit.....	136
<b>4.16</b> Crystal lattice interactions.....	137
<b>4.17</b> $\beta$ FXIIa protease structure.....	138
<b>4.18</b> $F_O-F_C$ electron density map for GlcNAc residues.....	139
<b>4.19</b> $F_O-F_C$ electron density map for PPACK.....	140
<b>4.20</b> Dimerization by PPACK.....	141

<b>4.21</b> Comparison between the structure of PPACK in a) $\beta$ FXIIa and b) thrombin.....	142
<b>4.22</b> Structural comparison of MBP- $\beta$ FXIIa with FXIIc and FXIIac.....	144
<b>4.23</b> Structural comparison of MBP- $\beta$ FXIIa with FXIIc and thrombin.....	146
<b>4.24</b> H1 pocket.....	147
<b>4.25</b> Analytical gel filtration analysis to verify the possible dimerization operated by PPACK.....	148
<b>4.26</b> Purification of Ecotin by affinity chromatography.....	149
<b>4.27</b> Analytical gel filtration analysis of complex formation between MBP- $\beta$ FXIIa and Ecotin in 1:1 ratio.....	151
<b>4.28</b> Analytical gel filtration analysis of complex formation between MBP- $\beta$ FXIIa and Ecotin in 1:2 ratio.....	152
<b>4.29</b> Analytical gel filtration analysis of complex formation between MBP- $\beta$ FXIIa and CTI in 1:4 ratio.....	153
<b>4.30</b> Stabilizing interactions between HCR and the activation loop.....	158
<b>5.1</b> 1% Agarose gel showing amplified $\beta$ FXIIa PCR fragment.....	161
<b>5.2</b> Western Blot of $\beta$ FXIIa using Anti 6X-His tag primary antibody.....	162
<b>5.3</b> Purification of $\beta$ FXIIa by affinity chromatography.....	163
<b>5.4</b> Purification of $\beta$ FXIIa by anion exchange chromatography.....	164
<b>5.5</b> Purification of $\beta$ FXIIa by size exclusion chromatography.....	165
<b>5.6</b> Purification of $\beta$ FXIIG570R by affinity chromatography.....	166
<b>5.7</b> Purification of $\beta$ FXIIG570R by affinity chromatography.....	167
<b>5.8</b> Purification of $\beta$ FXIIG570R by size exclusion chromatography.....	168
<b>5.9</b> S2302 time dependent assay for $\beta$ FXIIa.....	169
<b>5.10</b> S2302 time dependent assay for $\beta$ FXIIG570R.....	170
<b>5.11</b> Activity of $\beta$ FXIIa and $\beta$ FXIIG570R (pmol/sec) against various concentrations of S230.....	171
<b>5.12</b> Analytical gel filtration analysis of complex formation between $\beta$ FXIIG570R and Ecotin in 1:1 ratio.....	175

<b>5.13</b> Analytical gel filtration analysis of complex formation between $\beta$ FXIIG570R and Ecotin in 1:2 ratio.....	176
<b>5.14</b> Superimposition of gel filtration chromatograms for $\beta$ FXIIG570R and Ecotin in 1:1 and 1:2 ratio.....	178
<b>5.15</b> Effects of G570R mutation.....	181
<b>5.16</b> Surface interaction of Arg570 with Glu495.....	182
<b>6.1</b> Non reducing SDS-PAGE gel analysis of fractions form affinity chromatography.....	186
<b>6.2</b> Purification of TF- $\beta$ FXII by affinity chromatography.....	187
<b>6.3</b> Purification of TF- $\beta$ FXII by size exclusion chromatography.....	188
<b>6.4</b> Purification of TF- $\beta$ FXII by affinity chromatography.....	189
<b>6.5</b> Purification of TF- $\beta$ FXII by anion exchange chromatography.....	190
<b>6.6</b> HRV-3C protease cleavage.....	191
<b>6.7</b> HRV-3C and FXa protease cleavage.....	192
<b>6.8</b> Thrombin cleavage.....	193
<b>6.9</b> Co-purification of TF- $\beta$ FXII and Ecotin by affinity chromatography.....	195
<b>6.10</b> Co-purification of TF- $\beta$ FXII and Ecotin by size exclusion chromatography.....	196

## List of Tables

<b>1.1:</b> Inherited and acquired coagulation and metabolic risk factors associated with DVT.....	2
<b>1.2</b> List of the main molecules activated by NETs.....	13
<b>3.1</b> Antibiotics used for general transformation of the constructs in Novo Blue cells.....	70
<b>3.2</b> Solutions for preparing resolving and stacking gels for a 15% SDS-PAGE.....	72
<b>3.3</b> Primers used for the amplification of MBP and $\beta$ FXIIa.....	83
<b>3.4</b> PCR reaction mixture.....	83
<b>3.5</b> Cycling parameters used in order to perform PCR reaction.....	84
<b>3.6</b> Primers used to produce $\beta$ FXIIa construct. The forward and the reverse primers contain restriction site for BglII (AGATCT) and MluI (ACGCGT) restriction endonucleases respectively.....	85
<b>3.7</b> Reaction conditions for ligation.....	86
<b>3.8</b> Expression conditions for FXII in pCOLD-TF vector and for Ecotin in pET-26(+) vector.....	90
<b>3.9</b> Summary of the chromatographic methods, columns and buffers used for the purification of the various expressed and purified proteins with DES system.....	96
<b>3.10</b> Summary of the chromatographic methods, columns and buffers used for the purification of the various expressed and purified proteins with <i>E. coli</i> system.....	97
<b>3.11</b> Experimental reagents influencing (MBP- $\beta$ FXIIa)-PPACK crystallization.....	102
<b>3.12</b> Initial search conditions for the incomplete factorial method. All the experiments were performed at 20 °C.....	103
<b>3.13</b> Experimental reagents influencing the (MBP- $\beta$ FXIIa)-PPACK crystallizations. The list has been expanded adding new salts and precipitants similar to the ones already tested.....	104
<b>3.14</b> Initial search conditions for the new incomplete factorial grid. All the experiments were performed at 20 °C.....	105
<b>3.15</b> Experimental reagents influencing the crystallization of the (MBP- $\beta$ FXIIa)-Ecotin complex.....	107

<b>3.16</b> Initial search conditions for(MBP- $\beta$ FXIIa)-Ecotin. All the experiments were performed at 20 °C.....	108
<b>4.1</b> Enzyme kinetic parameters for MBP- $\beta$ FXIIa.....	120
<b>4.2</b> Crystallization conditions for MBP- $\beta$ FXIIa from the PEG suite.....	121
<b>4.3</b> Crystallization conditions for MBP- $\beta$ FXIIa from optimization.....	122
<b>4.4</b> Crystallization conditions for (MBP- $\beta$ FXIIa)-PPACK from the ProComplex suite.....	123
<b>4.5</b> Crystallization conditions for (MBP- $\beta$ FXIIa)-PPACK after initial optimization.....	125
<b>4.6</b> Crystallization conditions for (MBP- $\beta$ FXIIa)-PPACK obtained from incomplete factorial grid.....	126
<b>4.7</b> Crystallization conditions for (MBP- $\beta$ FXIIa)-PPACK obtained from finer incomplete factorial grid.....	127
<b>4.8</b> Data collection statistics for (MBP- $\beta$ FXIIa)-PPACK. Values in parentheses refer to the highest resolution shell.....	131
<b>4.9</b> Refinement statistics for (MBP- $\beta$ FXIIa)-PPACK. Values in parentheses refer to the highest resolution shell.....	133
<b>5.1</b> Enzyme kinetic parameters for $\beta$ FXIIa and $\beta$ FXIIG570R.....	171
<b>5.2</b> Crystallization conditions for $\beta$ FXIIa from the AmSO <sub>4</sub> suite.....	173
<b>5.3</b> Crystallization conditions for $\beta$ FXIIG570R from the ProComplex suite.....	179

## Abbreviations and Symbols

APC	Activated Protein C
APS	Ammonium persulphate
aPTTs	Activated Partial Thromboplastin Time
ATIII	Antithrombin III
BiP	<i>Drosophila</i> secretion signal
BK	Bradykinin
BSA	Bovine Serine Albumin
C1NH	C1 esterase inhibitor
CCD	Charge- Couple Device
CMK	Chloro Methyl Ketone
<i>cspA</i>	Cold shock protein A
CTI	Corn Trypsin Inhibitor
CVD	Cardiovascular disease
DES	<i>Drosophila</i> Expression System
DVT	Deep vein thrombosis
EDTA	Ethylenediaminetetraacetic acid
EGF-like	Epidermal growth factor like
EGFR	Epidermal growth factor receptor
EPCR	Endothelial Protein C receptor
EPI	Extrinsic pathway inhibitor
FIX	Coagulation factor IX
FnI	Fibronectin-type I
FnII	Fibronectin-type II
FV	Coagulation factor V
FVa	Activated coagulation factor V
FVII	Coagulation factor VII

FVIIa	Activated coagulation factor VII
FVIII	Coagulation factor VIII
FVIIIa	Activated coagulation factor VIII
FXa	Activated coagulation factor X
FXa	Activated coagulation factor X
FXI	Coagulation factor XI
FXIa	Activated coagulation factor XI
FXII	Coagulation factor XII
FXIIa	Activated coagulation factor XII
FXIIac	Recombinant coagulation factor XII zymogen
FXIIc	Recombinant coagulation factor XII zymogen-like
FXIII	Coagulation factor XIII
FXIIIa	Activated coagulation factor XIII
gC1qR	Complement Component 1 Subcomponent-binding Protein Mitochondrial
GlcNAc	N-Acetyl-D-glucosamine
GST	Glutathione S-transferase
HAE	Hereditary Angioedema
HA EI	Hereditary Angioedema type I
HA EII	Hereditary Angioedema type II
HA EIII	Hereditary Angioedema type III
HCR	Heavy Chain Remnant
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HK/HMWK	High molecular weight kininogen
HRV-3C Protease	Human Rhinovirus 3C Protease
HS	Heparan Sulphate
IMAC	Metal Ion Affinity Chromatography
IPTG	Isopropyl $\beta$ -D-1 thiogalactopyranoside
<i>lac</i>	Lactose operon

LACI	Lipoprotein-associated coagulation inhibitor
LB	Lysogeny Broth
MBP	Maltose Binding Protein
MCS	Multiple Cloning Site
MES	2-( <i>N</i> -morpholino)ethanesulfonic acid
NETs	Neutrophil Extracellular Traps
OD	Optical Density
PBS	Phosphate Buffer Saline
PCK	D-prolyl-L-phenylalanyl-L-arginine chloromethyl ketone
PCOLD-TF	pCOLD-Trigger Factor
PCR	Polymerase Chain Reaction
PK	Activated plasma kallikrein
PPACK	D-phenylalanyl-L-prolyl-L-arginine chloromethyl ketone
PPK	Plasma pre-kallikrein
S2	<i>Drosophila Schneider 2</i>
S2302	H-D-Pro-Phe-Arg-pNA•2HCl
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEC	Size Exclusion Chromatography
SER	Surface Entropy Reduction
TAFI	Thrombin activable fibrinolysis inhibitor
TBE	Tris/Borate/EDTA
TBS	Tris Buffer Saline
TBST	Tris Buffer Saline Tween
TEMED	N, N, N', N'-tetramethylethylenediamine
TF	Tissue Factor
TFPI	Tissue Factor Pathway Inhibitor
TF-βFXII	Recombinant Coagulation Factor βFXII with TF tag
TM	Thrombomodulin

tPA	Tissue-type Plasminogen activator
Tris	Tris(hydroxymethyl)aminomethane
uPA	Urokinase-type Plasminogen activator
vWF	von Willebrand Factor
WB	Western Blot
$\alpha$ FXIIa	Activated $\alpha$ FXII
$\beta$ FXIIa	Activated $\beta$ FXIIa
$\beta$ FXIIG570R	Mutated Coagulation Factor $\beta$ FXIIa

### CRYSTALLOGRAPHIC TERMS

$(uvw)$	Coordinates in Patterson map
$(x, y, z)$	Real space coordinates
$ F_{calc} $	Calculated Structure Factor amplitudes
$ F_{obs} $	Observed Structure Factor amplitudes
$\mathbf{a}^*, \mathbf{b}^*, \mathbf{c}^*$	Vectors in the reciprocal space
$ F(h, k, l) $	Structure Factor amplitude for the $(h, k, l)$ reflection
$F_M(\mathbf{r})$	Molecular Structure Factor
$F_{calc}$	Calculated Structure Factor
$F_{obs}$	Observed Structure Factor
$R_{factor}$	Crystallographic R factor
$R_{factor}^{free}$	Free R factor
$d_{h,k,l}$	Interplanar distance
$d_{h,k,l}^*$	Reciprocal lattice vector
Å	Angstrom
ASU	Asymmetric Unit
DM	Density Modification
$f(\mathbf{r}^*)$	Atomic structure factor

FT	Fourier Transform
$h, k, l$	Miller Indices
I	Intensity
LLG	Log-likelihood gain
MAD	Multiple wavelength Anomalous Diffraction
MIR	Multiple Isomorphous Replacement Method
MR	Molecular Replacement
NCS	Non Crystallographic Symmetry
SAD	Single wavelength Anomalous Diffraction
SIR	Single Isomorphous Replacement Method
TFZ	Translational function Z-score
$\alpha, \beta, \gamma$	Unit cell angles
$\sigma_I$	Error associated with the Intensity
<b>a, b, c</b>	Unit cell vectors
$K(\mathbf{r}^*)$	Structure Factor for the crystal
$\alpha(h, k, l)$	Phase angle
$\rho(\mathbf{r})$	Electron density over the entire space $\mathbf{r}$
$\rho(x, y, z)$	Electron Density at position $(x, y, z)$

# CHAPTER 1

## INTRODUCTION

Cardiovascular disease (CVD) is a class of diseases involving heart, blood vessels or both[1] and it represents the main cause of mortality worldwide in both developed and developing countries with a deleterious economic impact.[2, 3] Data submitted by individual countries to the World Health Organization suggested a dramatic increase of death for CVD (nearly 30%) and CVD related death (more than 80%) in the low and middle income countries.[4] Epidemiological studies indicate the industrialization, urbanization and globalization as the cause of the proliferation of the main CVD risk factors (age, gender, high blood pressure, diabetes, tobacco consumption).

Among all the cardiovascular diseases, venous thrombosis, the improper formation of a blood clot in the vein, represents the third most common cardiovascular disease, after myocardial infarction and stroke.[5] Data suggest that venous thromboembolism episodes occur with annual incidence of 1 or 2 per 1000 adults.[6, 7] To date, both acquired and inherited coagulation risk factors, that can increase the likelihood of any one individual developing a thrombosis episode, are identified (Table 1.1).[8, 9]

**Table 1.1:** Inherited and acquired coagulation and metabolic risk factors associated with DVT.[9-11]

<b>Inherited</b>	<b>Acquired</b>
Antithrombin deficiency	Antiphospholipid syndrome
Protein C deficiency	Prostate cancer
Protein S deficiency	Pregnancy
Factor V Leiden	Oral contraceptives/hormone replacement therapy
Prothrombin G20210A	

Today's therapeutic strategies to treat the different forms of thromboembolism are based on the identification of targets that block the thrombus formation. However, even if the available anticoagulant drugs [such as heparin and/or direct Factor Xa (FXa) and thrombin inhibitors ] are effective, they can cause dangerous bleeding as adverse effect.[2] For this reason, it becomes necessary to identify new target molecules that block thrombosis without compromising the physiological hemostasis.[12] In order to fulfill this goal, we need an accurate knowledge of the molecular mechanism associated with thrombosis.

## **1.1 Hemostasis and Coagulation**

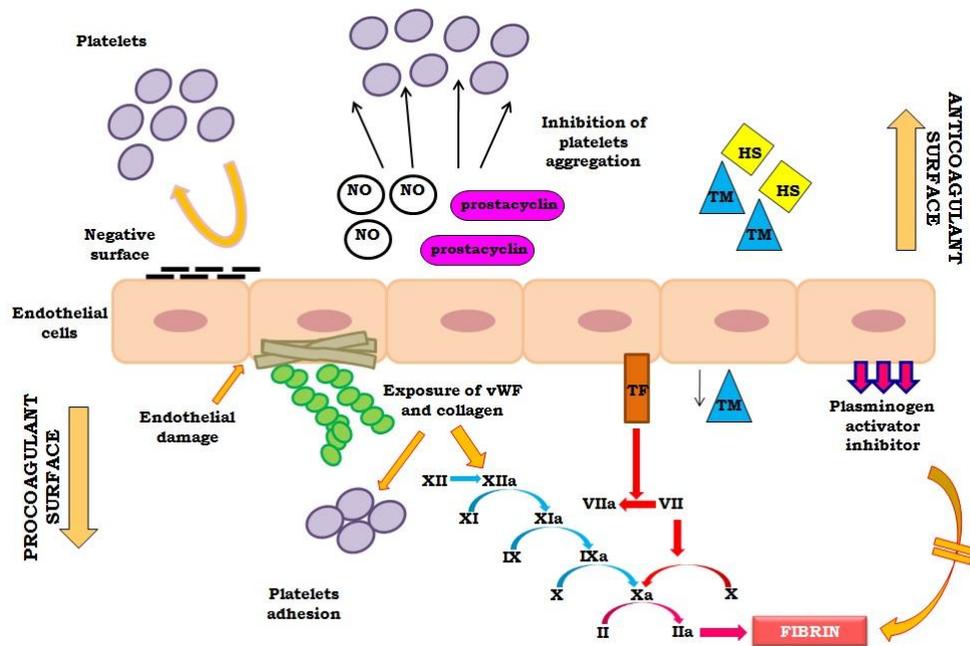
Hemostasis is the arrest of bleeding from a damaged blood vessel. The hemostatic equilibrium in the organism is usually maintained by three distinct but closely related mechanisms: vascular endothelium, platelet aggregation and coagulation pathways.[13]

### **1.1.1 Vascular endothelium**

Vascular endothelium is composed by a thin monolayer of cells, named vascular endothelial cells that constitute the interior surface of blood vessels. Vascular endothelium has multiple roles. Its primary function is to act as a selective but static physical barrier[14] and it regulates the flow of nutrients, biologically active molecules and blood cells. The capability of endothelium to act as a gate-keeper is due to the presence, on its surface, of different membrane-bound receptors for several molecules like growth factors, coagulant and anticoagulant proteins.[15]

The endothelium also plays a pivotal role in hemostasis, since it promotes the expression of coagulation inhibitors and platelet aggregation. In normal conditions, the vascular endothelium inhibits platelet function and it acts as anticoagulant (Figure 1.1). In fact, the presence of negative charges on the upper surface repels platelets, while production of nitric oxide and prostacyclin inhibits the platelet function. Moreover, the expression of thrombomodulin and heparan sulphate enhances the anticoagulant properties.[16] When a tissue damage occurs, the endothelium stops acting as anticoagulant creating a prothrombotic and anti-fibrinolytic microenvironment (lower surface, Figure 1.1).[15] This results in the exposure of blood components to prothrombotic substances such as collagen, von Willebrand Factor and fibronectin (all responsible for platelet aggregation) and tissue factor (which initiates the coagulation pathway). At the same time, anticoagulant properties are modulated by reduced expression of thrombomodulin and heparan sulphate. When the hemostatic balance is restored, the vascular endothelium returns to be anticoagulant.

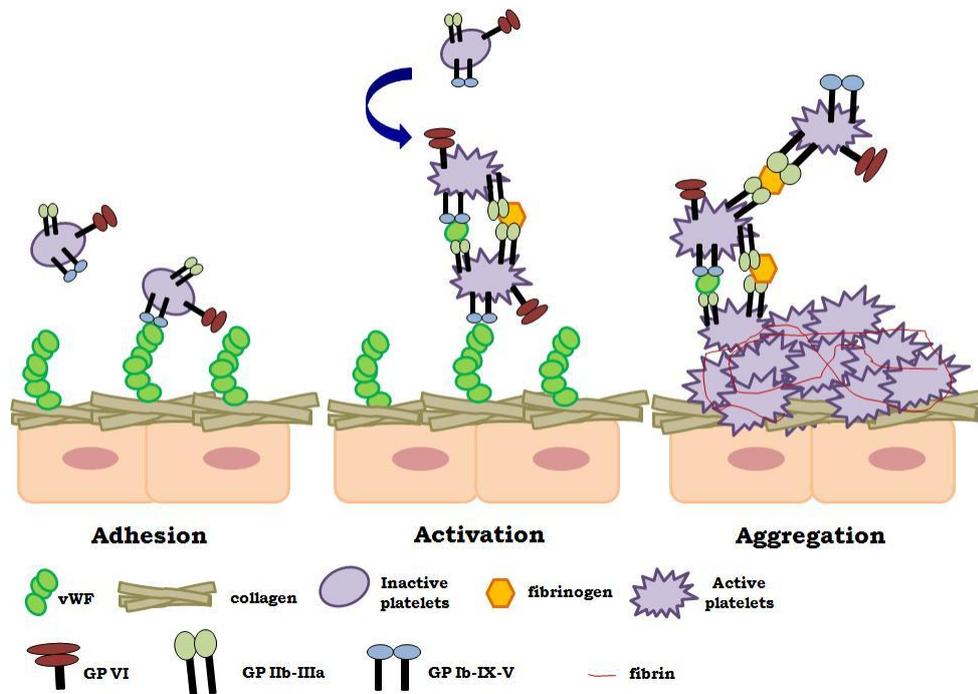
The procoagulant/anticoagulant activity of the vascular endothelium is itself regulated through different levels of expression of the involved substances.[15]



**Figure 1.1: Endothelial hemostatic function.** Schematic illustration of the function played by the vascular endothelium in the hemostasis. The upper surface shows the anticoagulant function of endothelium in normal conditions. Its negative surface repels platelets, while nitric oxide and prostacyclin inhibit platelet aggregation. Furthermore, expression of thrombomodulin (TM) and heparan sulphate (HS) enhances the anticoagulant function. In the presence of a tissue damage, instead, the endothelium becomes procoagulant (lower surface). Exposure of von Willebrand Factor (vWF) and collagen promotes platelets adhesion. Endothelial cells promote the secretion of tissue factor (TF) that in turn initiates fibrin generation, while fibrinolysis is inhibited by secretion of plasminogen activator inhibitor. This figure is adapted from ref [16].

### 1.1.2 Platelets

Platelets also contribute to the maintenance of vascular integrity and hemostasis.[16] Platelets circulate in close contact to the endothelial cells of blood vessel, in order to act promptly when the vessel wall gets injured. In these cases, platelets form a plug (primary hemostasis).[17, 18] The formation of the platelet plug occurs in three different steps: (i) platelet adhesion, (ii) platelet activation and (iii) platelet aggregation (Figure 1.2).[13]



**Figure 1.2: Platelet plug formation.** The figure shows the main events that occur in the generation of a platelet plug, as a result of vascular injury. After the vascular damage, circulating platelets adhere to the endothelial surface thanks to the interaction with exposed collagen and vWF. The adhesion to endothelial surface promotes the activation of the platelets and their aggregation. In this way, platelets form a plug that is then stabilized by the fibrin generated in the coagulation step. The figure is adapted from ref [19].

In the first event of primary hemostasis, platelets adhere to different components of the endothelium, depending on the blood flow rate. The adhesion is favored by interactions with different components, mainly collagen, fibronectin and von Willebrand factor.[20, 21] When the platelets adhere to the endothelium they become activated.[16, 17] As a result, the platelets undergo shape change becoming spherical with protuberant pseudopodia, cytoskeleton rearrangement and organelle centralization. Platelet activation also promotes the expression of several molecules that are necessary for the successive steps.[22-24] In the aggregation step[18, 25], once the first layer of platelets has been formed, additional platelets participate in the formation of the plug. The aggregation is regulated by the binding of fibrinogen to GPIIb-IIIa[24], that in turn recruits new platelets to be activated. The formed platelet plug is then stabilized by the fibrin network. The latter is released in the last step of the hemostatic process: the coagulation cascade.[13]

## **1.2 Coagulation**

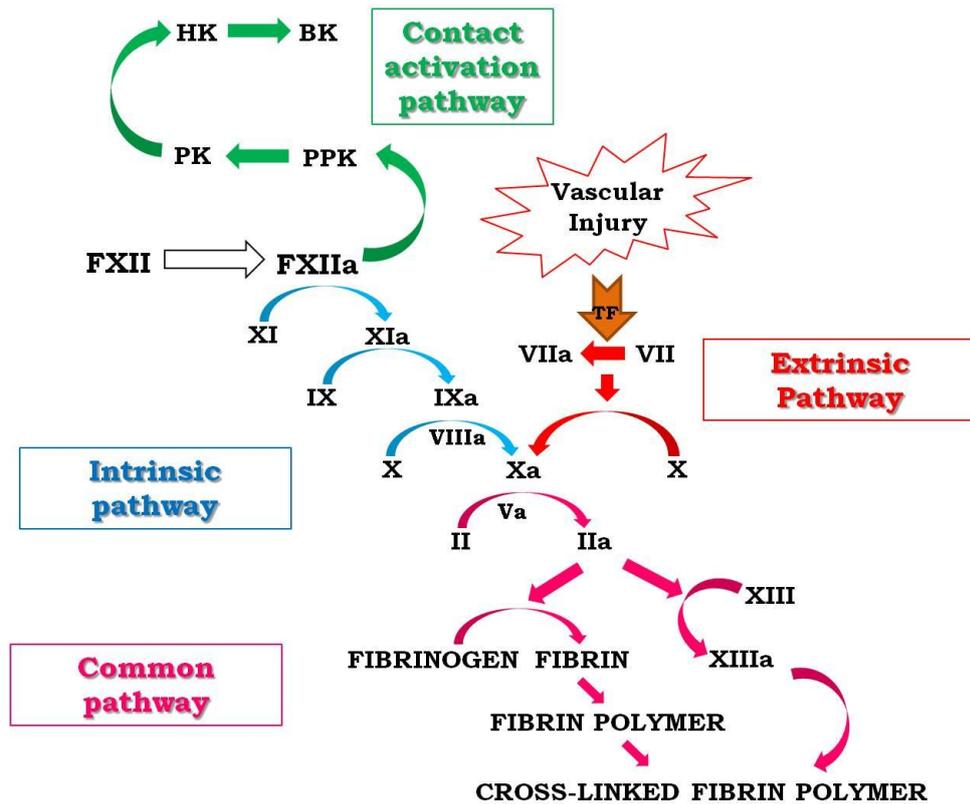
The last step in the hemostatic process is the activation of a fibrin network that clots the blood and forms an insoluble blockage: this process is called coagulation.

### **1.2.1 Coagulation cascade**

In 1964 Davie & Ratnoff[26] and McFarlane[27] proposed two similar models to explain the formation of a fibrin clot in correspondence of a tissue damage. Both models suggested that coagulation can be viewed as a cascade

of enzymatic reactions which terminates with the activation of thrombin, the enzyme responsible of promoting the conversion of soluble fibrinogen into the insoluble fibrin. The sequence of reactions was called intrinsic pathway[26] since no substances extrinsic to blood were involved. Plasma proteins implicated in this process are called coagulation factors. They circulate in blood as inactive enzyme precursors. Their main feature is to act in sequence: the zymogen of a specific coagulation factor and its glycoprotein co-factor are activated. They hence catalyze the subsequent reaction of the cascade, which ends with formation of cross-linked fibrin.[26, 27]

In more recent years, studies suggest that the intrinsic pathway was more important for the growth and maintenance of the fibrin clot[28] while another pathway, named extrinsic pathway, was the primary initiator of the coagulation.[28-30] To date, the coagulation cascade is based on the concept of two distinct pathways, the intrinsic and extrinsic pathway merging into a final common pathway, characterized by the generation of thrombin.[30] This is illustrated in Figure 1.3.



**Figure 1.3: The contact activation pathway and the coagulation cascade.** The green arrows show the reactions of the contact activation pathway. The blue arrows indicate the reactions of the intrinsic pathway (left side). The red arrows indicate the reactions of the extrinsic pathway (right side). The two pathways merge into the final common pathway, shown in pink arrows. Inactivated Coagulation Factors are indicated with roman numerals, activated coagulation factors are indicated by a lower case “a”, PPK is plasma-prekallikrein, PK is plasma-kallikrein, HK is high molecular weight kininogen, and BK is bradykinin.

The extrinsic pathway, or tissue factor pathway, plays a fundamental role in the initiation of fibrin formation.[30] Its main role is the release of thrombin through the generation of a “thrombin burst”.

The extrinsic pathway is triggered when tissue factor (TF), present on cell membranes in sub endothelial layers of blood vessels, is exposed to blood at the site of the injury. TF is a glycoprotein [31, 32] with high affinity for the blood circulating factor VII (FVII).[33] In presence of a vascular injury, TF is a co-factor for the active form of FVII, FVIIa, in the activation of

coagulation factor FX. Factor VIIa cleaves FX for the active FXa.[34, 35] The activation of FVII occurs through the action of different proteases, among which are thrombin, FXa, factor IXa (FIXa) and factor XIIa (FXIIa).[34, 35] The TF-FVIIa complex is also able to activate FIX.[36] Both reactions are catalyzed by  $\text{Ca}^{2+}$  ions and traces of other proteases circulating in the blood. Once activated, FXa forms a 1:1 complex with the activated Factor V (FVa). The reaction occurs in presence of  $\text{Ca}^{2+}$  and phospholipids.[37] This complex, called prothrombinase, is responsible for the activation of Prothrombin into thrombin.[38] FVa participates as co-factor in the generation of thrombin. At an initial stage, FVa activation is determined by FXa[39]; however, once thrombin has been generated, it also contributes to the activation of the coagulation Factor V (FV).[40] The generation of thrombin is followed by a limited proteolysis of fibrinogen to form fibrin. Thrombin is also involved in the activation of FXIII (FXIIIa), the enzyme responsible of the formation of covalent cross-links between two adjacent fibrin monomers. In presence of  $\text{Ca}^{2+}$  ions, fibrin contributes to accelerate the activation of this coagulation factor by thrombin.[41] When the concentration of TF released near the injury is low, the activation of FXa proceeds through the activation of FIX (FIXa) and factor VIII (FVIIIa).[42] In fact, the TF-FVIIa complex activates FIX; FIXa forms a complex with FVIIIa; the newly generated FIXa-FVIIIa complex activates FXa. Again, all these reactions occur in presence of  $\text{Ca}^{2+}$  and phospholipids.[43] FVIIIa also participates to the reaction as co-factor. The extrinsic pathway is short-lived due to the presence of the tissue factor pathway inhibitor (TFPI),[44, 45]

and anti-thrombin III (ATIII).[46] In this case, the generation of FXa occurs through FIXa, following the intrinsic pathway.

The intrinsic pathway of blood coagulation starts after the activation of coagulation factor XI (FXI). The main activators of FXI are thrombin[47] and FXII in its active form ( $\alpha$ FXIIa).[48] When blood comes in contact with a negatively charged surface, activation of FXII is operated by a process called “contact activation” (Figure 1.3).[28] The contact system is composed by three inactive enzymes [(FXI, FXII and plasma prekallikrein (PPK)] and the non-enzymatic pro-cofactor high molecular weight kininogen (HK).[49] When FXII zymogen is exposed to a negative surface, it undergoes a conformational change resulting in its auto activation. Activation of FXII can also be operated by plasma kallikrein (PK).[50] Once active, FXIIa leads two different events: (i) release of the peptide hormone bradykinin (BK) through kallikrein-kinin system; (ii) fibrin formation through the intrinsic pathway.[51]

BK generation starts with activation of PPK to PK mediated by FXIIa; in turn, PK cleaves HK and produce BK, an important mediator of vascular permeability, inflammation and pain.[52]

Regarding the classic intrinsic pathway (Figure 1.3), the activation of FXII is followed by the generation of active factor XI (FXIa) and, subsequently, by the activation of FIXa.[53] The intrinsic and extrinsic pathways converge at the level of FX activation into the final common pathway, characterized by the generation of thrombin and, subsequently, the conversion of fibrinogen into fibrin.[30]

The contribution of Davie & Ratnoff[26] and MacFarlane[27] in the understanding of the biochemical interactions regulating the coagulation was significant but the model cannot explain some events. For example, it cannot explain why patients deficient in one of the components of the contact activation pathway (FXII, PK, HK) do not show any bleeding tendency but prolonged aPTTs (activated partial thromboplastin time). Moreover, patients lacking one of the components of the intrinsic pathway (FVIII or FIX) have serious bleeding tendency even if the extrinsic pathway is intact. Similarly, patients with FVII deficiency show serious bleeding even if the intrinsic pathway is intact. Clearly intrinsic and extrinsic pathway are not independent from each other and more complex interactions take place *in vivo*. [16, 28, 54]

### **1.3 Immunothrombosis**

When activated, the coagulation system has to provide a quick response through a rapid and highly localized activation of mechanisms that lead to the generation of fibrin. [55, 56] The same immediate and local response is a feature of the innate immune system. [57]

The innate immune system comprises several mechanisms that can limit pathogen dissemination. The innate immunity acts through recognition of pathogens and recruitment of cells and molecules able to identify and remove the foreign host. [58]

Evolutionary studies showed a connection between coagulation and innate immunity. Organisms, like *Drosophila Melanogaster* and the Horseshoe crab, use clotting factors to initiate intravascular coagulation to respond to

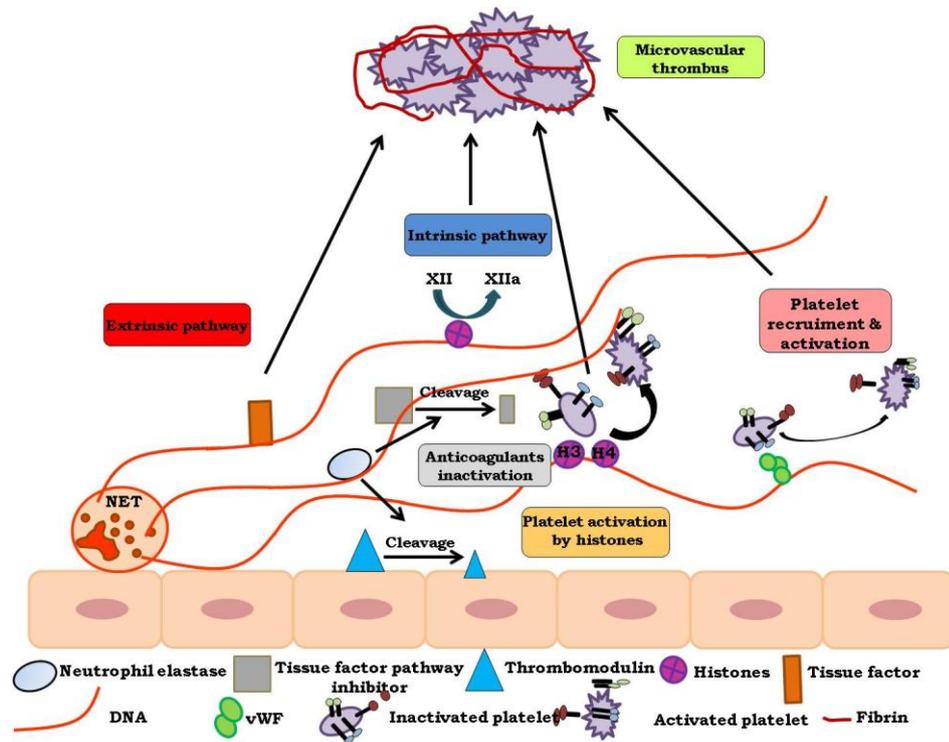
pathogenic infection and protect them.[59, 60] Studies have also shown that fibrin is able to bind and activate innate immune cells at the site of infections.[61, 62] Thus, coagulation is clearly involved in the host response[63, 64]. Microorganisms, such as monocytes and neutrophils, are able to activate the coagulation cascade in response to pathogen dissemination.[65] This led to the concept of “immunothrombosis”, coined for the first time by Engelmann and Massberg,[66] which is defined as the capability of the organism to activate intravascular coagulation in response to blood pathogen invasion. Hence, activation of blood coagulation culminating in fibrin formation helps the detection, containment and elimination of pathogens. In order to limit pathogen dissemination, specific cells and molecules are able to capture and retain pathogens within the fibrin network through the formation of thrombi inside the blood micro vessel. Micro thrombi formation is followed by local accumulation of innate immune cells that deliver active tissue factor, which then initiates the extrinsic pathway.[66]

One of the major components of immunothrombosis is represented by Neutrophil Extracellular Traps (NETs), which are fibers produced by neutrophils that are mainly composed of DNA and histones. Several studies showed that NETs support immunothrombosis in several ways (Figure 1.4), mainly acting as a surface which promotes and compartmentalizes the coagulation system.[67] NETs can explicate their procoagulant activity in many ways, as summarized in Table 1.2:

**Table 1.2:** List of the main molecules activated by NETs.

<b>vWF[67]</b>	Enhancement of platelet recruitment
<b>Platelets[67]</b>	Histone molecules on the NET surface activate platelets and stimulate platelet adhesion
<b>Neutrophil elastase[68]</b>	Extrinsic pathway of thrombin generation through the degradation and inactivation of TFPI and thrombomodulin
<b>FXII[69]</b>	Intrinsic pathway of blood coagulation

Immunothrombosis seems to be an important player in the host defense mechanisms at intravascular level. Nevertheless, immunothrombosis can promote thrombosis-associated pathologies.[66]



**Figure 1.4: Immunothrombosis promoted by NETs.** The figure shows how NETs promotes the formation of thrombi in the microvessel. They can explicate their procoagulant activity in many ways: (i) extrinsic pathway by activating TF; (ii) contact pathway by activation of FXII mediated by histones; (iii) anticoagulant inactivation by neutrophil elastase; (iv) platelets recruits and activation operated by histones and vWF. Figure adapted from ref. [68].

#### 1.4 Regulation of the coagulation cascade

Since the coagulation pathway involves various proteins and biochemical interactions, its regulation occurs at different levels, mainly inhibiting the generation of fibrin.[70]

Fibrin formation starts with exposure of plasma to the damaged vessel wall and subsequent exposure of coagulation factor VII to TF. TF is expressed by vessel wall cells that are not exposed to blood. The presence of endothelial cells creates a barrier that physically separates blood clotting factors from TF in intact tissue, to ensure that coagulation occurs only when the vessel

wall is damaged.[66] The TF coagulation activity is mainly regulated by TFPI. TFPI is a multivalent Kunitz-type plasma protease that inhibits TF, by binding FXa to or near to the active site.[71] Inhibition mainly involves the second Kunitz-type domain, however it seems that also the other domains contribute to the inhibition of the protease.[44] Once the TFPI-FXa complex is formed, TFPI binds to FVIIa in complex with TF using the first kunitz-type domain.[71, 72] Two different pathways are proposed to explain the formation of the quaternary complex. In one case, it is suggested that TFPI binds to FXa in solution and after that TFPI-FXa complex binds to VIIa-TF complex; a second mechanism proposes the binding of TFPI to a preformed ternary complex FXa-VIIa-TF. In TFPI knockout mice, an embryonic lethality due to hemorrhage has been observed thus underlining the importance of TFPI as hemostatic regulator.[73]

The activity of thrombin is regulated by several plasma proteins, such as ATIII and Protein C, through direct or indirect mechanisms.[70]

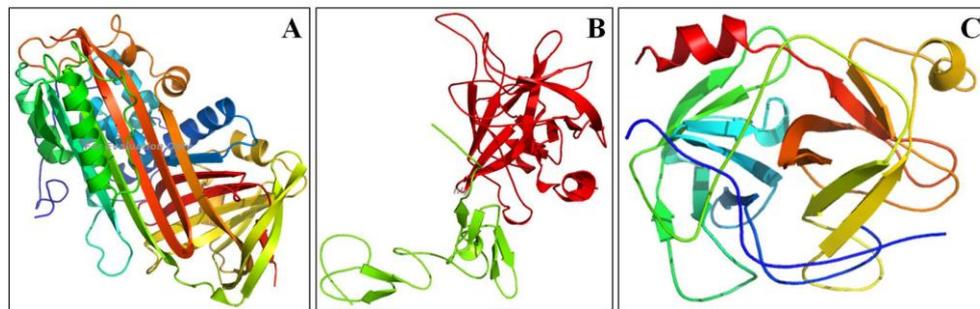
ATIII (Figure 1.5 A), a 432 residue single chain glycoprotein, mainly inhibits the activity of thrombin [46] but also of other coagulation factors, such as FIXa, FXa, FXIa and FXIIa.[30] Once thrombin cleaves the reactive site loop of Antithrombin in Arg393-Ser394 position, a covalent complex is formed.[74] ATIII function is regulated by heparin, a glycosaminoglycan that binds ATIII, inducing a conformational change, which enhances the formation of thrombin-ATIII complex of 1000-fold.[75, 76] The inactivated thrombin-ATIII complex dissociates from heparin, so that it can bind

another ATIII molecule and catalyze the inactivation of other molecules of thrombin.[70]

Activated Protein C (APC) (Figure 1.5 B) is a vitamin K-dependent glycoprotein circulating in blood plasma as a zymogen. It acts in conjunction with its co-factor, thrombomodulin, by proteolytically inactivating FVa and FVIIIa.[77] Thrombin binds thrombomodulin located on the endothelial surface via an anion binding exosite.[78] The formation of this complex enhances the affinity of thrombin for APC instead of fibrinogen.[79] The binding of protein C to the complex is mediated by the protein C receptor (EPCR) on endothelial cells. APC released from EPCR can also bind Protein S, another vitamin K-dependent factor, and proteolytically inactivates FVa and FVIIIa, in presence of a negatively charged surface and calcium ion, thus limiting further coagulation.[80]

Regulation of the coagulation cascade occurs also through the degradation of fibrin. This is performed by plasmin (Figure 1.5 C), a serine protease involved in dissolving blood clots, cleaving fibrin into degradation products that inhibit excessive fibrin formation. It is present in the liver as a zymogen: plasminogen. In the circulatory system, plasminogen has a closed conformation in which the access to the activation bond (R561/V562) targeted for cleavage by tPA and uPA is blocked through the O-linked sugar on T34.[81] When plasminogen binds fibrin clots or cell surface, it changes its conformation (closed→open) and can be converted into active plasmin by a variety of enzymes, like tissue-type plasminogen activator (tPA), urokinase-type plasminogen activator (uPA), kallikrein, and FXIIa.[82, 83]

Endothelial cells are also involved in the regulation of coagulation cascade. In fact, they are indirectly involved in thrombin activity regulation since heparin sulfate and thrombomodulin are located on their surface. In addition, endothelium can inhibit platelet function through the release of nitric oxide and prostacyclin that prevents platelets aggregation.[16]



**Figure 1.5 Regulation of the coagulation cascade.** The figure shows the crystal structure of the main regulators of the coagulation cascade. A) Cartoon representation of bovine ATIII(PDB code: 1ATT, [84]); B) cartoon representation of human APC, with highlighted the light chain (green) and the heavy chain (red) (PDB code: 1AUT, [85]); C) cartoon representation of the catalytic domain of human plasmin (PDB code: 1BML, [86]).

### **1.5 Coagulation factors deficiencies, thrombosis and pharmacological strategies**

Physiological hemostasis is responsible for regulating normal blood fluidity and, when it is necessary, to induce a prompt formation of a fibrin clot. It is fundamental that the human body maintains an appropriate hemostatic equilibrium; in fact every alteration of this process can cause several bleeding disorders like hemophilia, leukemia, anemia and vitamin K deficiency.

Thrombosis is considered to be the pathological version of hemostasis and involves platelet activation and coagulation.[28, 66, 87] Thrombosis is

defined as the formation of an occlusive blood clot in the vascular system and it can affect both arteries and veins.[88] While in arteries platelets generate a thrombus, in veins the thrombus is fibrin and red blood cells rich.[89]

Thrombosis represents the most frequent cause of mortality worldwide. In particular, deep vein thrombosis (DVT), which is the formation of a blood clot within the deep vein, is a common condition further complicated by detachment and embolization of the thrombi in the lung, causing pulmonary embolism.[65] The formation of venous thrombi can be ascribed to any alteration of blood flow (venous stasis), hypercoagulability and change in blood vessel wall. To date, the mechanisms by which coagulation factors contribute to the formation of a venous thrombosis are better understood.[89] For this reason, the current anticoagulant therapy is mainly based on blocking the initiation step of coagulation, to prevent thrombin generation or inhibition of thrombin to drastically reduce fibrin formation. In particular thrombin and coagulation factor Xa are useful pharmacological targets for venous thromboembolism.[90], [91]

New FXa inhibitors can block the protein indirectly or directly. Indirect FXa inhibitors mimic the heparin action in mediating its binding to ATIII. Fondaparinux[92-94] is the only licensed alternative to heparin. Direct FXa inhibitors, instead, bind to its active site, thus blocking any further interaction. These kind of inhibitors inhibit both free FXa molecules and FXa molecules bound to the platelets in prothrombinase complex.[95] Rivaroxaban[96] is the first orally active direct factor Xa inhibitor. Rivaroxaban is capable of inhibiting both free factor Xa and factor Xa

bound in the prothrombinase complex. Similarly to warfarin, Rivaroxaban can prevent non-hemorrhagic strokes and embolic events.[97]

Direct thrombin inhibitors have better biological and pharmacokinetic properties.[98] Dabigatran [99] and warfarin[100] are direct oral anticoagulant (DOAC) and are strong inhibitors of thrombin, as proved by clinical studies.[101]

A huge effort has been put in the past years in the development of new and safer anticoagulants that can represent a valid alternative to warfarin, the most widely prescribed oral anticoagulant drug in North America and United Kingdom.[102] However, since these anticoagulant drugs target molecules that are also fundamental for physiological hemostasis, their therapeutic use is also correlated with the risk of bleeding, that can be fatal in some cases.[2, 28] Thus, it is necessary to identify alternative targets. Recent studies suggest that other coagulation proteins can be used as therapeutic targets in the treatment of thrombosis, mainly FIX and FXII.

Great interest has been focused on FXII, due to new studies showing that FXII deficient mice models are protected from experimental ischemic stroke and pulmonary embolism. Also FXII deficiency is not associated with bleeding disorder.[103, 104] These findings indicate that FXII contributes to the fibrin formation and propagation in pathological thrombosis but it is not essential for normal hemostasis, thus suggesting that FXII can be a useful target for a novel antithrombotic therapy.[105]

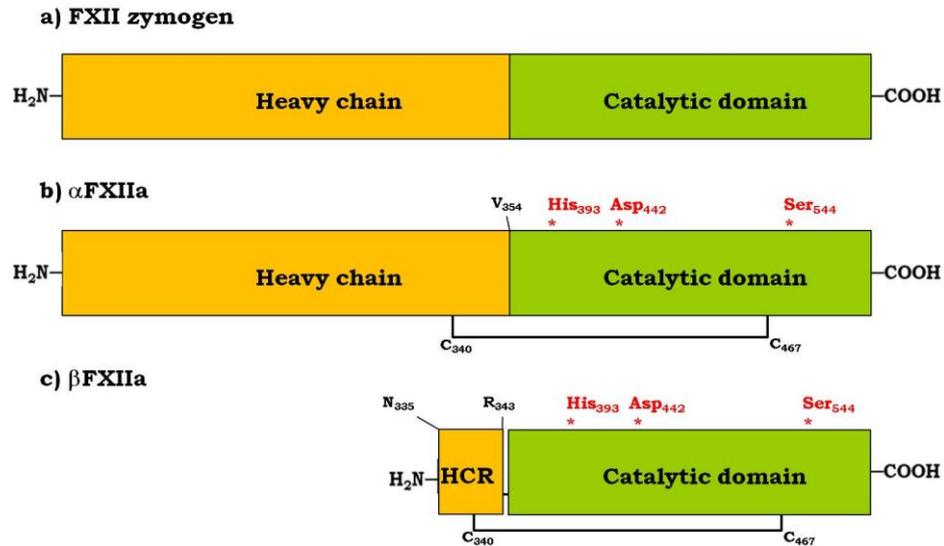
## 1.6 Coagulation factor XII

Coagulation factor XII, FXII, plays an important role in the coagulation cascade, being involved in the initiation of the intrinsic pathway. It is also known as Hageman factor, from the name of the 37-years old patient, John Hageman, who first showed a prolonged clotting time even if he had no protracted bleeding. In 1955, Ratnoff and coworkers suggested that the patient was lacking of an unknown coagulation protein, which was able to initiate the clotting in a glass test tube. Later, Ratnoff connected the Hageman factor with the recently discovered FXII and they also established that FXII circulated in blood as a zymogen before being activated upon contact with a negatively charged surface (contact activation).[106] The discovery of coagulation FXII allowed to describe for the first time the intrinsic pathway of coagulation.[26, 27] Since FXII deficiency is not associated with a bleeding disorder, FXII is not indispensable for physiological hemostasis *in vivo*. [103]

The F12 gene, whose size is 12 kilobase pairs, encodes FXII and it is organized in 13 introns and 14 exons. The intron/exon disposition is found to be similar to the genes of plasminogen activators family.[107]

FXII is secreted by the liver as glycoprotein (it has eight glycosylation sites) and circulates in blood as zymogen with a concentration of 375 nM and half-life of 50-70 h.[108-110] The zymogen consists of a single chain of 596 amino-acid residues, with a molecular weight of 80 kDa (Figure 1.6 a), and different proteases, such as kallikrein or trypsin, can activate it by limited proteolysis by In fact, a single cleavage between Arg353 and Val354 generates the active  $\alpha$ FXIIa.[111]

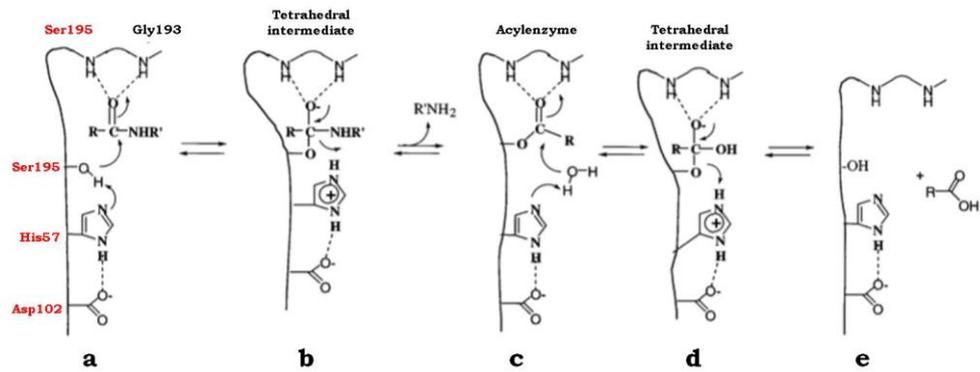
$\alpha$ FXIIa (Figure 1.6 b) consists of two chains, the N-terminal heavy chain composed of 353 residues ( $M_w$  50 kDa) and the C-terminal light chain of 243 residues ( $M_w$  28 kDa), connected by a disulfide bridge between Cys340 and Cys467.[112] Once  $\alpha$ FXIIa is formed, a subsequent reduction of the disulfide bridge releases the light chain of FXII[113, 114], containing the entire catalytic domain. A further cleavage of  $\alpha$ FXIIa in position Arg334 and Arg343 yields another FXII fragment of 28-30 kDa, named  $\beta$ FXIIa (Figure 1.6 c), consisting of the entire light chain plus a short fragment of nine residues of the heavy chain. While  $\alpha$ FXIIa can bind negatively charged surfaces and is capable of activating FXI,  $\beta$ FXIIa can only activate prekallikrein in solution.[115]



**Figure 1.6: Different forms of human FXII.** a) FXII zymogen; b) activated  $\alpha$ FXIIa generated by proteolytic cleavage between Arg353 and Val354.  $\alpha$ FXIIa is composed of a heavy chain and a light chain (catalytic domain) held together by a disulfide bridge. The catalytic triad (His 393, Asp442, Ser594) is indicated with asterisks; c) a further proteolytic cleavage generates  $\beta$ FXIIa consisting of the entire light chain linked by the disulfide bridge to a short fragment [9 residues, heavy chain remnant (HCR)] of the heavy chain.

### 1.6.1 Structure of FXII

FXII belongs to the class of enzymes, known as serine proteases, so called because of the presence of a nucleophilic Ser residue in the active site.[116, 117] Their main function is the cleavage of peptide bonds in proteins. They are characterized by a distinctive structure consisting of two six-strand  $\beta$ -barrel domains and a catalytic triad (His57, Asp102, Ser195) located at their junction.[117, 118]



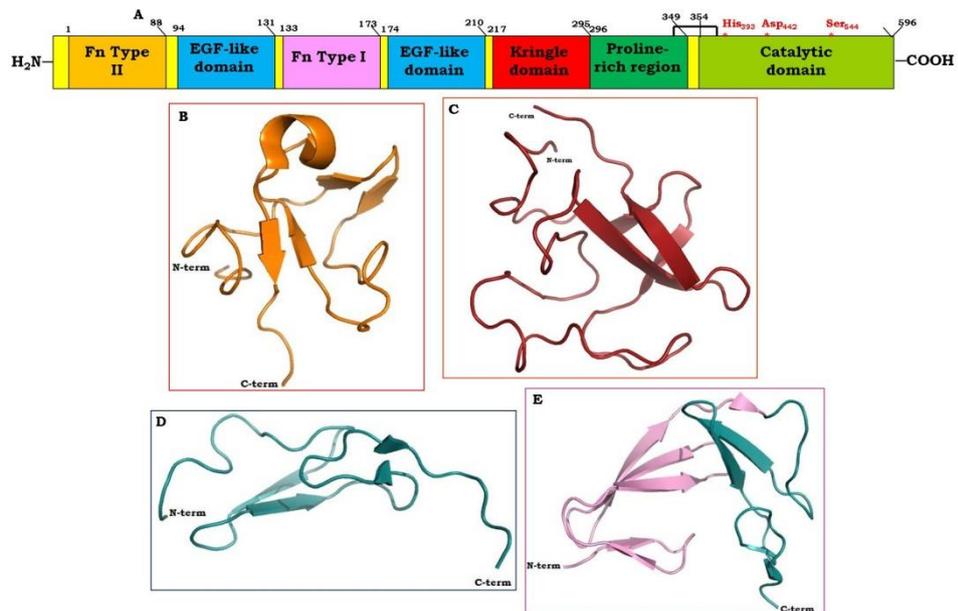
**Figure 1.7: Enzymatic mechanism for serine proteases.** The substrate binds the surface of the serine protease with the scissile bond inserted in the active site and the carbonyl carbon of the scissile bond positioned near to Ser195; a) nucleophilic attack of Ser195 and formation of a tetrahedral intermediate; b) release of the N-terminus peptide c) nucleophilic attack of water and d) formation of a second tetrahedral complex; e) release of the C-terminus peptide and regeneration of the active site. Figure adapted from ref. [116]. See text for detailed description of the mechanism.

Once the substrate has bound the surface of the serine protease, the enzymatic mechanism features the interaction between the OH group of the serine protease and the carbonyl carbon of the scissile peptide bond of the substrate. This, in conjunction with the proton transfer from Ser195 to His57 and the electron transfer from the substrate carbonyl double bond to the oxygen, brings to the formation of an intermediate with distorted tetrahedral geometry (Figure 1.7 a). This intermediate is stabilized by Gly193 and Ser195, which form the so-called oxyanion hole. Gly193 and Ser195 donate hydrogen bonds to the negatively charged oxyanion, lowering the activation energy of the reaction. The C-N bond of the substrate is hence broken, bringing the release of the N-terminus of the target protein and the formation of an acyl-enzyme intermediate (Figure 1.7 b). This further interacts with a water molecule (Figure 1.7 c), which is involved in a proton transfer to His57 and interaction with the carbon of the acyl-enzyme

intermediate, with the formation of a second tetrahedral complex (Figure 1.7 d). In the final step of the reaction the C-terminus of the substrate is released and the active site is regenerated (Figure 1.7 d). The role of Asp102 throughout the enzymatic reaction is the acceptance of a hydrogen bond from His57, thus increasing the electronegativity of the latter. This favors the acceptance of a hydrogen bond from the OH of Ser195.[116, 119]

Serine proteases are classified according to the properties of the pocket accommodating the substrate (the specificity pocket), which can favor the interaction of the serine protease with specific substrates.[120, 121] There are four classes of serine proteases: trypsin-like, chymotrypsin-like, elastase-like and subtilisin.[122] FXII belongs to the chymotrypsin-like class of serine proteases, whose specificity pocket is hydrophobic.

The primary structure of  $\alpha$ FXIIa consists of several domains. The heavy chain is composed by the signal peptide (residues 1-19), a fibronectin type II domain (residues 1-88), an epidermal-growth-factor-like domain (EGF-like, residues 94-131), a fibronectin domain type I (residues 133-173), another EGF-like domain (residues 174-219), a kringle domain (residues 215-295) and the Proline-rich region domain (residues 296-349), while the light chain contains the protease domain (residues 354-596) (Figure 1.8).



**Figure 1.8: FXII domains.** A) Representation of the domains of human FXII (different colors for different domains of the protein). B) Cartoon representation of FnII domain (PDB code: 1KS0, [123]). C) Cartoon representation of the Kringle Domain (PDB Code: 1PK4, [124]). D) Cartoon representation of EGF-like domain (PDB Code: 1F7M, [125]). E) Cartoon representation of FnI (pink) and EGF-like domain (blue) (PDB code: 4BDX, [126]).

The fibronectin type II domain (Figure 1.8 B) is eighty-eight residues long and exhibits around 40% of sequence homology with type II domain of fibronectin, together with the four conserved cysteines.[127] The fibronectin type II domain is composed by different sub-domains of known functional activity. Amino acid sequence 1-28 of this domain is a putative surface-binding site [128, 129] for negatively charged substances. Among this region, residues 3-19 play a significant role in the activation of FXI, through a 2:1 FXI:FXII interaction.[130] Further studies suggest that active fibronectin type II domain of  $\alpha$ FXII interacts with endothelial cells and platelets. On endothelial cells, FXII interacts with urokinase plasminogen activator receptor (uPAR) through residues 39-47. [131] This interaction is

found to be dependent on the  $Zn^{2+}$  ion concentration, and, in fact, it was possible to identify two Zn-binding sites (40-44, 78-82) in this N-terminal region of coagulation factor XII.[132, 133] Bradford and co-workers demonstrated that  $\alpha FXIIa$  also binds GPIIb $\alpha$  on platelet surface.[134]

The FnII domain is followed by the EGF-like domain (Figure 1.8 D). EGF-like domain is an evolutionary conserved domain, which has been found in other proteins, like t-PA[135] and coagulation factors FX[136], FIX[137] and FVII.[125] Proteins belonging to the EGF-Family show similar structural and functional characteristics. A characteristic feature of EGF-like domain is the presence of a fifty amino acid region, that is very well conserved, including nine invariant cysteines and glycine residues.[111] The 3D structure of the domain is formed by two sub-domains, N-terminal and C-terminal respectively, that folds as two-stranded  $\beta$ -sheets connected by a loop.[138] EGF is involved in the growth, proliferation and differentiation of cells by binding to the receptor EGFR.[139] Studies on FXII indicate that it enhances cell proliferation, even if it is still unknown if this activity is modulated by EGF-like domains.[140, 141]

The two EGF domains in FXII are separated by the fibronectin type I domain, a region of 43 amino acid residues; this region is so called because of the limited homology sequence with fibronectin type I region.[112] A typical aspect of this domain is the existence of two disulfide bridges allowing the formation of a finger domain. In 2013, Beringer *et al.*[126] published the first crystallographic structure of FnI-EGF-like tandem domain (Figure 1.8 E) and they proposed that these two domains are implicated in the binding with  $\beta$ -amyloid fibrils via electrostatic

interactions, due to the presence of five positively charged residues. They also hypothesized that the FnII-EGF-like domain could participate in the FXII binding with negatively charged surfaces throughout the same mechanism.

The primary structure of FXII is also characterized by the presence of the Kringle domain. Kringle domains (Figure 1.8 C) are structured as typical triple loop folds, which are stabilized by the presence of three disulfide bridges and a certain number of hydrogen bonds leading to short anti-parallel  $\beta$ -sheet structures.[142] Kringle domains of different proteins such as plasmin, urokinase-type plasminogen activator (uPA), tissue-type plasminogen activator (tPA), have been extensively studied and it has been shown that kringles are mediators for the interaction with integrins, furthermore suggesting that a kringle-integrin interaction is a common mechanism in kringle-containing proteins.[143]

The C-terminal domain of the heavy chain is the proline rich region, which owes its name to its high content in proline (33% -12 of 53 residues). This domain is unique to FXII, since it doesn't share any homology with other known proteins.[112] The function of this region is uncertain. However, this domain is interesting, due to a series of mutations responsible for the Hereditary Angioedema type III (HAEIII), which will be discussed in details in section 1.6.

The catalytic domain of FXII represents the entire primary sequence of the light chain. This region presents a high percentage of homology with

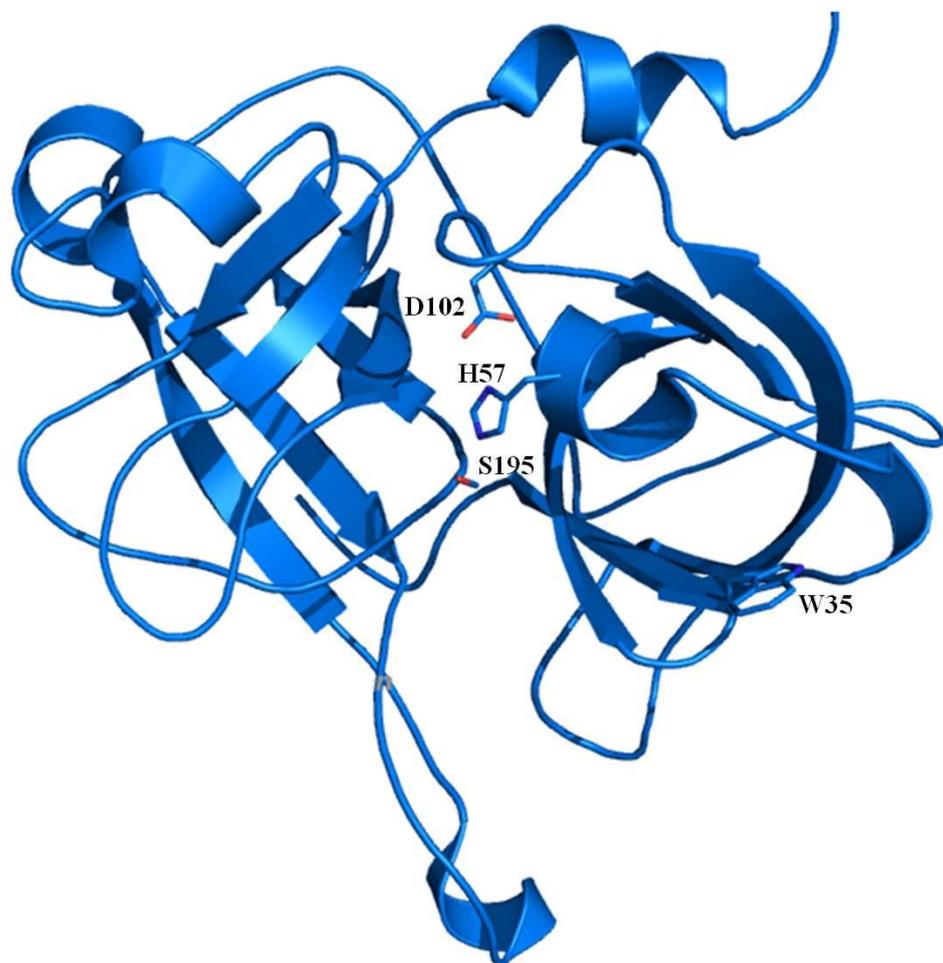
plasmin and urokinase, indicating that they all belong to the same protease subfamily.[142]

The crystallographic structure of the light chain of FXII was recently solved in the group of Prof J. Emsley.[144] The proteins investigated were FXIIac which has the native N-terminus (PDB code: 4XE4), and FXIIc (PDB code: 4XDE), a zymogen-like protease in which the N-terminal Val16 (chymotrypsin residue numbering is used in the description of the crystallographic structure of FXIIc and FXIIac) is blocked due to the presence of two additional residues, Arg and Ser.

The global folding of both structures is very similar to those observed in other serine proteases,[145] with two interacting six-stranded  $\beta$ -barrel domains and the catalytic triad (His57, Asp102 and Ser195) located at the junction between the two domains (Figure 1.9).

A detailed analysis of the crystallographic structure suggested that both models are the zymogen form of FXII. In particular, FXIIac is characterized by the absence of the oxyanion hole (Gly193, Ser195). Differences are also in the unusual organization of the loops 180 and 220 in the S1 pocket area: residues 182-183 and 217-221 are not visible in the electron density suggesting the flexibility of the mentioned loops. It has been already reported that, in the zymogen, the loops of interest are usually disordered to prevent any substrate recognition. They become partially ordered only after the cleavage bringing to the active form. At variance, the S1 pocket is clearly defined in FXIIc. In FXIIc and FXIIac, the conformation of the S34 pocket is retained. Nevertheless, a 90° rotation of the Trp35 residue side chain and a shift in the main chain results in a repositioning of Trp35 in the

H1 pocket. Therefore, a lid is formed and the pocket becomes partially buried. The analysis of the conformation of some key residues allowed to further confirm the zymogen nature of the solved structures. In particular, the spatial organization of Asp194 was analyzed. In the active protease, this residue makes a salt bridge interaction with N-terminal Val16, so that the active site cleft is exposed to the substrate.[146] In the zymogen form, instead, the same residue pairs with His40.[28] In the case of FXIIc and FXIIac, the Asp194 is far away and cannot make any interaction with N-terminal Val16.



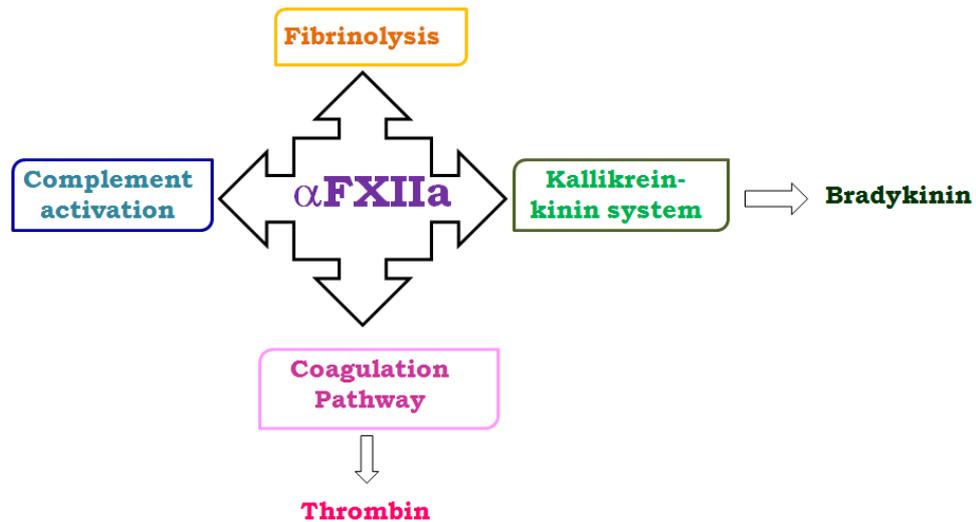
**Figure 1.7: Cartoon diagram of the protease domain of FXII.** The catalytic triad (His57, D102, S195) is shown as sticks. (PDB code: 4XDE)

### **1.6.2 *In vitro* and *in vivo* activators of FXII**

While *in vivo* functions of FXII are just beginning to emerge, its *in vitro* activation is well established.[147] Activation of FXII can occur through two different mechanisms:

- the zymogen gets (auto)activated upon exposure to a negatively charged surface, such as kaolin. The interaction with a negative surface induces a conformational change causing FXII to be more susceptible to the proteolytic cleavage. In presence of polyphosphate surfaces, the activation of FXII is kinetically favored and totally dependent from the concentration of  $Zn^{2+}$  ions.[148]
- FXII is activated being proteolytically cleaved by other proteases like PK (contact activation).[149]

Once activated,  $\alpha$ FXIIa is able to trigger both proinflammatory and procoagulant responses that culminate with the generation of BK and thrombin, respectively, together with the complement system and fibrinolysis pathway (Figure 1.10).



**Figure 1.10: FXII-driven contact pathways.** The figure shows the pathways that FXIIa is able to trigger after it has been activated by contact with a negative surface.

While several substances have been known to be FXII activators *in vitro* for a long time, physiological activators of FXII are emerging only recently. Several studies have shown that FXII activation is linked to platelets. In fact, Back and coworkers[150] demonstrated that, under physiological conditions, activated platelets are able to bind FXII and trigger the contact activation pathway and, most importantly, FXII activation occurs on platelets surface with implications for inflammation and clotting. NETs can trigger the activation of FXII due to their electrochemical properties.[68] Most recently it has been found that polyphosphates with chain length greater than 75 subunits are able to initiate the plasma contact activation.[151] *In vivo* activation of FXII can also occur on the endothelial surface: for example, HS as well as other vascular endothelial proteins, like gC1qR,[152] can provide a useful site for FXII, thus facilitating the assembly of the contact system on endothelial cells.[153]

### **1.6.3 FXII and thrombosis**

For decades, it was widely accepted the idea that FXII does not contribute to normal hemostasis and, hence, TF is the physiological initiator of blood coagulation.[103, 112] However, recent studies have challenged this concept. In particular, it has been shown that FXII-deficient mice model are protected from thrombosis both in arterial and venous beds.[103] Kleinschnitz *et al* also showed that FXII-deficient mice model are protected from ischemic brain injury.[104] Similar results have also been obtained in other animal models (rabbit and baboon).[154] Studies carried out by van der Meijden *et al.* showed that formation of thrombi can be reduced operating on the inhibition of FXII activity through the inhibition of platelets activity. While these studies confirmed that FXII is not necessary for hemostasis, they also clearly indicated that FXII can play a role in pathological thrombosis.[155] The mechanism featuring FXII in the generation of thrombi is still unknown; however, a novel antithrombotic strategy could, in principle, aim to inhibit FXII.[2, 66, 104, 156]

### **1.7 Mutations in F12 gene and their effect on FXII activity**

In addition to its role as enzyme in thrombosis formation, altered FXII activity is also related to other deficiencies. In some cases, the altered activity of FXII can be ascribed to mutations in the gene encoding for it. In particular, mutations in F12 gene have been related with hereditary angioedema type III (HAEIII).[157-159]

HAE is a rare autosomal dominant blood disorder. It causes episodic attacks of swelling, affecting face, extremities, genitals, gastrointestinal tract and

upper ways.[160] In the last case, it can cause asphyxiation and is, therefore, an important clinical feature of the disease.[161]

The majority of HAE cases are caused by a deficiency of the C1 esterase inhibitor (C1NH), a protein of the complement system involved in the regulation of the coagulation, complement and contact cascade through the inhibition of coagulation factors (FXI and FXII), complement and contact proteases. When C1NH is deficient, there is an uncontrolled production of BK with a consequent vascular permeability increase.[162]

We can distinguish three different types of HAE, I, II and III. Type I and II are caused by a deficiency of C1NH; the C1NH deficiency can be quantitative or qualitative leading to HAEI and HAEII, respectively.[163]

At variance, HAEIII is associated with mutations of the F12 gene.[164] In particular, two missense mutations in the F12 gene were found: c1032CrA[164] and c1032CrG.[165] These two point mutations translate the Thr328 residue, located in the C-terminal proline rich region into Thr328Lys and Thr328Arg, respectively. This causes an increased  $\alpha$ FXII activity. Bork *et al.* recently reported a novel mutation consisting of a deletion of 72 base pairs in the same region of the F12 gene where the two mutations are located, confirming an association between the F12 gene mutations and HAE III.[166] Moreover, a 18 base pair duplication was found in F12 gene in a woman and her daughter with diagnosed HAEIII. The duplication translates for six additional residues in 298-303 positions.[167] The molecular mechanisms defining HAEIII associated with FXII mutations are still unknown; it is possible that the presence of these

mutations cause alterations in FXII activation or in kallikrein/kinin system.[168]

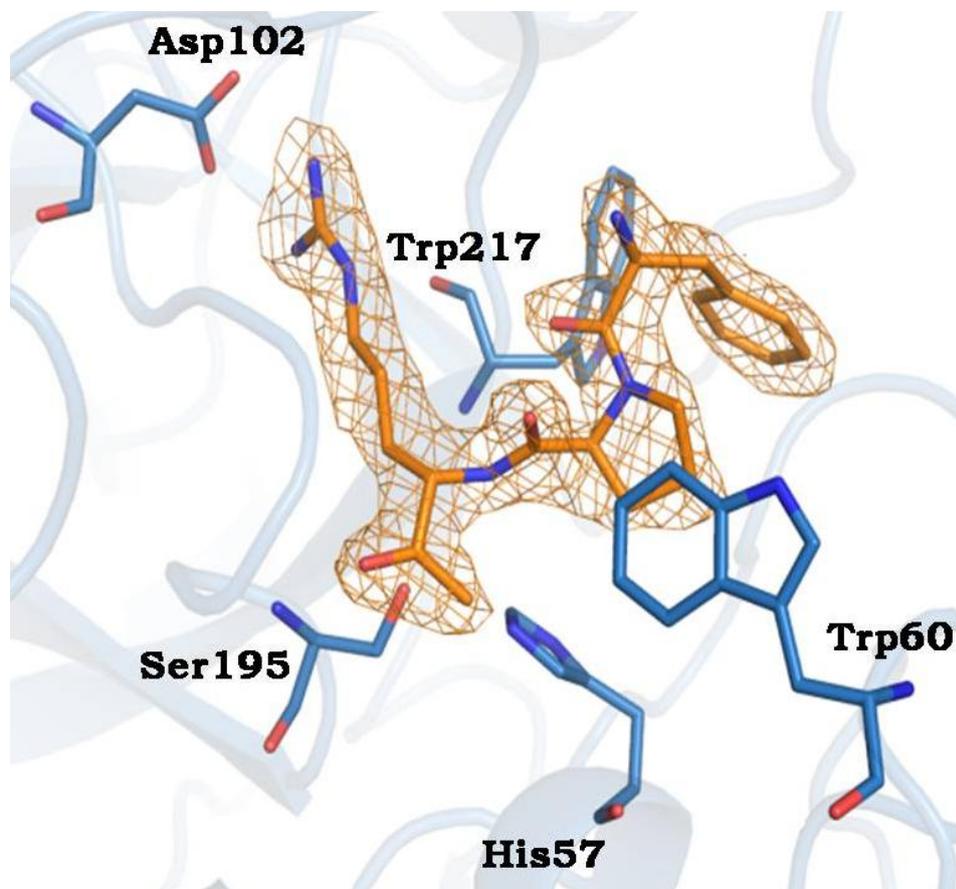
Another interesting missense mutation that causes a reduction in the activity of FXII is the G570R mutation. The G570R is caused by a mutation of a GGC codon to CGC.[169] The glycine residue is next to the Cys571 and the surrounding sequence is well conserved in other serine proteases with the exception of the chymotrypsin.[107, 170, 171] The presence of the positively charged Arg instead of Gly can inhibit the formation of the essential disulfide bridge with Cys540, with a consequent alteration of conformation integrity and drop of activity.[169]

### **1.8 Inhibition of coagulation factor FXII**

Alterations in the blood coagulation cascade are associated with life-threatening conditions, such as arterial and venous thrombosis.[28] Even if drugs currently used in the treatment of pathological thrombosis are efficacious, they have serious limitations, due to the risk of fatal bleeding.[2] For its implication in development of thrombosis and because its deficiency/inhibition is not related with any bleeding, the possibility of adopting FXII as therapeutic target could represent a major advancement, thus providing a valid alternative to the limited number of treatments in the field of antithrombotic therapies.[48, 104, 149] The knowledge of its three dimensional structure represents the starting point for the rational design of new antithrombotic inhibitors.

### 1.8.1: Inhibition of FXII by Chloro Methyl Ketone (CMK) peptides: PPACK and PCK.

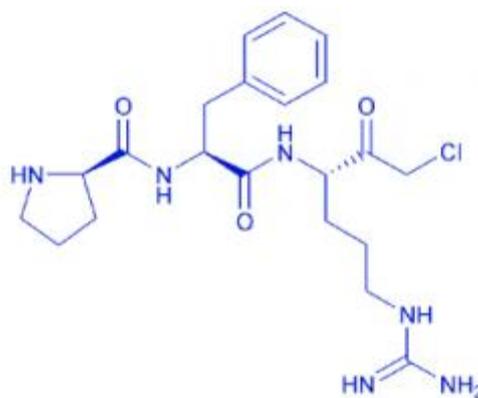
PPACK (D-Phe-Pro-Arg chloromethylketone) and PCK (D-Pro-Phe-Arg chloromethylketone) are small covalent inhibitors consisting of a tripeptide fragment, D-Phe-Pro-Arg and D-Pro-Phe-Arg respectively, modified at the C-terminus with a methylketone group.



**Figure 1.11: Binding mode of PPACK.** OMIT map for PPACK showed in orange. The figure also shows the residues of the catalytic triad (His57, Ser195, Asp102) and the residues of the hydrophobic pocket in which the Pro of the inhibitor is located (PDB Code: 1PPB).[172]

The amino-acid sequence of PPACK reproduces the cleavage site of the natural substrate fibrinogen and it shows a very high affinity for

thrombin.[173] Nevertheless, it has been shown that PPACK can also inhibit other serine proteases.[174-176] The crystal structure of PPACK in complex with thrombin[172] and other proteases is reported.[175, 177] The PPACK irreversibly inhibits the proteases alkylating the His57.[174] Figure 1.11 shows the binding mode of PPACK in the active site cleft of serine proteases. However, studies *in vivo* showed a rapid loss of activity due to the reactivity of the chloromethylketone group with other plasma components.[178] It has been shown that PPACK also inhibits FXIIa.[176] PCK (Figure 1.12) irreversibly inhibits the amidolytic activity of FXIIa and plasma-kallikrein-mediated activation of FXII.[179, 180] In mice model, it has been shown that treatment with PCK protect them from ischemic brain injury, without affecting normal hemostasis.[104].



**Figure 1.12: Chemical structure of PCK.**

Even if PPACK and PCK are not highly specific for FXII, they could be both used as templates, in order to elucidate the binding mode of the inhibitor on the target protease.

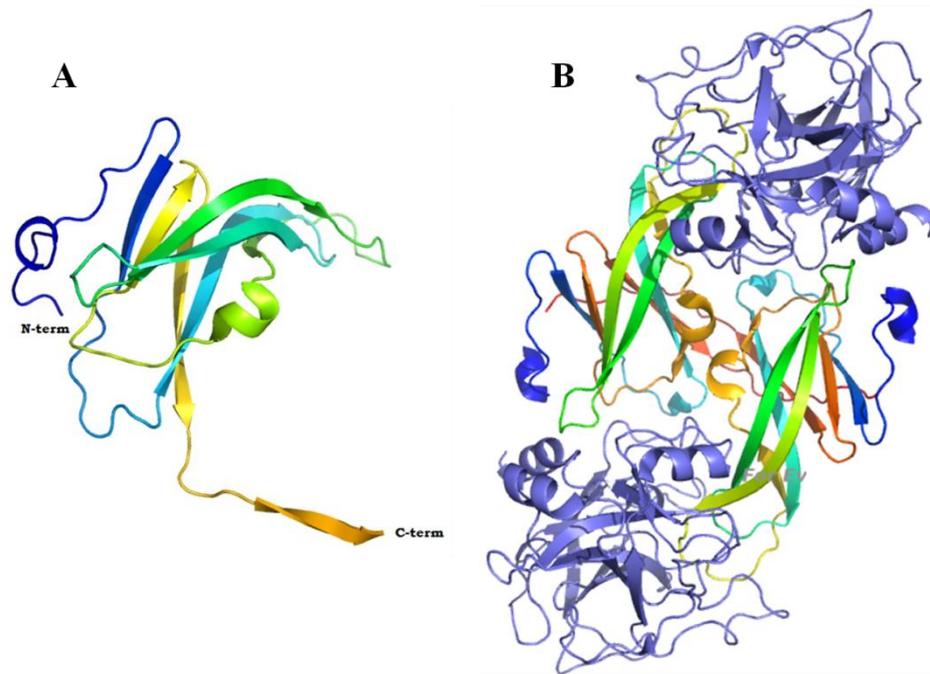
### **1.8.2 Inhibition of FXII by macromolecular protease inhibitors: Ecotin**

Ecotin is a serine protease inhibitor localized in the periplasm of *E. coli* and it is known to inhibit several serine proteases like trypsin and chymotrypsin with comparable potencies.[181] The primary sequence of Ecotin is composed by 142 amino-acid residues (Mw 16 kDa) and the monomer tends to dimerize in solution at sub-micromolar concentrations.[182] The crystal structure of Ecotin alone (Figure 1.13 A) and in complex with other serine proteases (Figure 1.13 B) shows that one dimeric inhibitor binds two protease molecules forming a heterotetrameric complex with a two-fold symmetric axis.[183-185] The interaction with the protease domain involves two distinct regions of Ecotin. The first region is called primary binding site. It involves the 80s loop (residues 81-86) and 50s loop (residues 52-54). The second region is known as secondary binding site and includes two loops of Ecotin, loop 60s (residues 67-70) and loop 100s (residues 108-113).[186] The primary binding site is known to interact with the active site of the molecule and it follows the same mode of action of other macromolecular inhibitors. The secondary binding site, instead, binds a relatively flat surface that is distant from the active site of the protease.[185] These interactions, together with the Ecotin dimer interface at the C-terminus, are all presumed to be important for determining the quaternary structure of the tetramer. In fact, Ecotin dimerization not only increases the interaction surface with the protease, but it also allows to adopt the right orientation of the molecular hinge, in order to bind different proteases. Studies showed that if the hinge region is destroyed in some way, the

position of the two binding sites may change, with a significant loss of affinity.[186]

FXIa and thrombin have been successfully co-crystallized with Ecotin[184, 185], but no structure is available for the complex with FXIIa. Ulmer *et al.* showed that Ecotin is a potent reversible tight-binding inhibitor of FXIIa with  $K_i=89$  pM.[187]

The role of Ecotin is unknown. Since it inhibits FXIIa as well as other proteases of the contact activation pathway, the knowledge of the crystal structure of FXIIa-Ecotin complex could be useful for several reasons. In fact, the analysis of the crystal structure and a comparison with already solved structures could help to better understand why Ecotin interacts with the target using three distinct areas. While Ecotin is not specific to FXIIa, a model of FXIIa-Ecotin complex can be used to design new inhibitors that not only interact with the active site but that also possess a long tail thus allowing interactions with other regions of the target, in order to enhance the specificity and selectivity of the inhibitor.

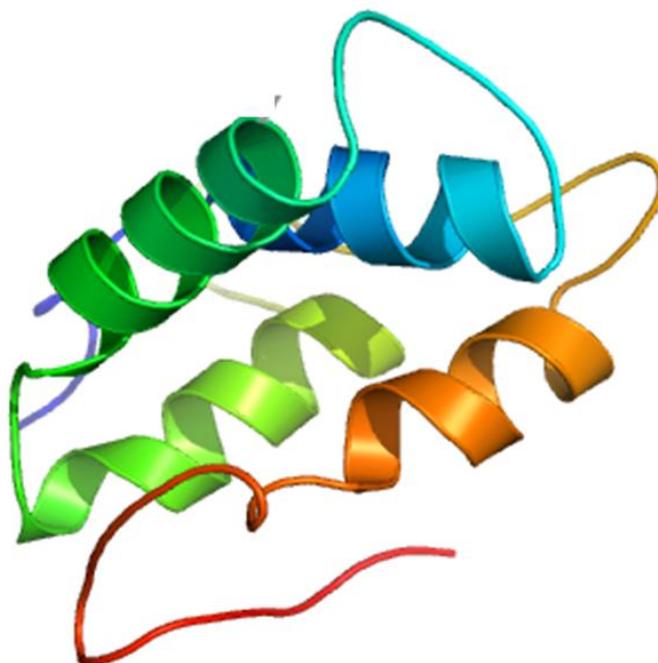


**Figure 1.13:** A) Cartoon representation of Ecotin structure (PDB code: 1ECY, [188]); B) Cartoon representation of the hetero-tetrameric complex Ecotin-thrombin. Ecotin is shown in rainbow colors while thrombin in blue (PDB code: 1ID5, [185]).

### 1.8.3 Inhibition of FXII by macromolecular protease inhibitors: Corn Trypsin Inhibitor (CTI)

CTI (Figure 1.14) is a 13.6 kDa protein isolated from sweet corn seeds and it is known to be a highly selective inhibitor of FXIIa.[189] CTI forms a 1:1 molar complex with FXIIa or trypsin and prolongs the activated partial thromboplastin time of human plasma. Studies reports that CTI has only little inhibitory effect on other plasma proteases.[190] For this reason, CTI is the most common used reagent to inhibit FXIIa in *in vitro* experiments.[191] Because CTI is unique in its strong inhibition of FXIIa, a detailed understanding of the interaction taking place between FXIIa and CTI could provide considerable insight into the highly specific interaction between the inhibitor and the target. Therefore, understanding the

interactions between CTI and FXIIa at atomistic level could contribute to the design of specific inhibitors of FXIIa.[192, 193]



**Figure 1.14.** Cartoon representation of CTI (PDB code: 1BFA, [194])

## 1.9 Objectives

FXII plays a pivotal role in several processes that can be related to thrombosis, ischemia, myocardial infarction, [103], recurrent pregnancy loss [195] and HAE III [157]. Since inhibition of FXIIa results in the reduction of thrombus formation without compromising physiological hemostasis, inhibition of the active form of FXII represents a new and selective strategy for preventing stroke and other thromboembolic diseases [196].

In this context, the main objective of this PhD project is the structural characterization of the protease domain of FXIIa in complex with different inhibitors (PPACK, PCK, Ecotin, CTI), using X-ray crystallography

technique. In this way, we aim to describe the atomic details of the substrate-inhibitor recognition. The complex structure can be used as a starting point for rational drug design to the synthesis of strong and selective antithrombotic inhibitors targeted toward FXII.

## CHAPTER 2

### THE THEORY OF X-RAY CRYSTALLOGRAPHY

#### 2.1 Introduction

It is well established that the knowledge of the tertiary structure of macromolecules, such as proteins, is an essential prerequisite to understand its function.[197] Moreover, a detailed description of protein-ligand interactions is required in order to develop new therapeutic drugs. In this framework, X-ray crystallography is able to provide accurate models of large molecular structures and complexes at high resolution, theoretically with no limits in terms of molecular weight.[198, 199]

A diffraction experiment is carried out as follows: a single crystal of the macromolecule of interest is mounted on a goniostat and exposed to a collimated, intense X-ray beam by obtaining diffraction images that result in electron density maps. The electron density map, which represents the atomic structure of the molecule, is obtained by the Fourier Transform (FT) of the structure factors, vector entities, whose module is proportional to the intensity of the diffraction spots in the collected data. Since the phase information is lost during the data collection, complex mathematical procedures are required to get this information (the so-called phase problem). Once the initial electron density map has been obtained, the model has to be refined, in order to obtain the best fit between the experimental and calculated data.[200]

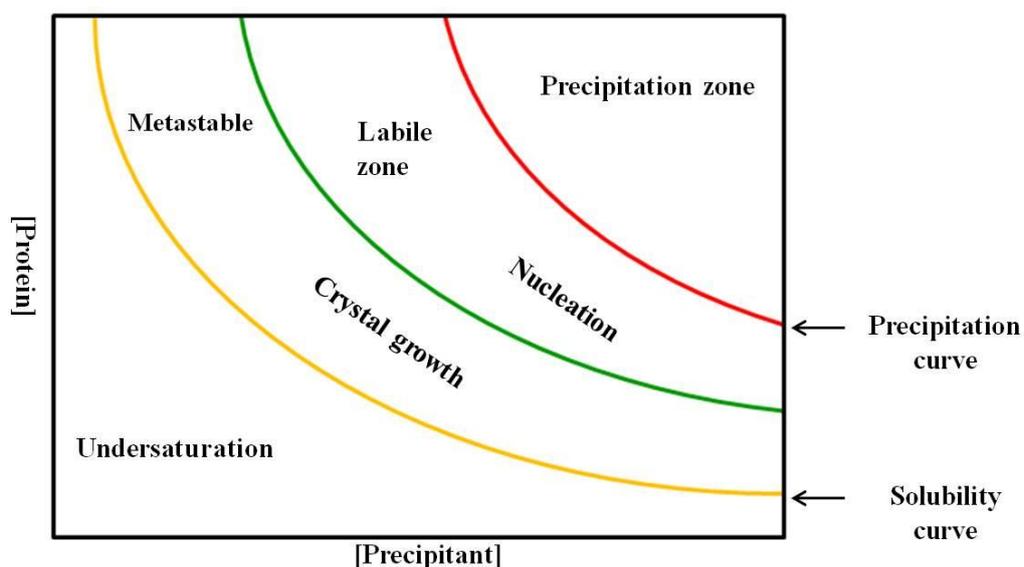
In this Chapter, the theory of protein crystallography is described. First, the principles of protein crystal growth are introduced (section 2.2). Then, the

crystal structure (section 2.3) and the geometric principles of single crystal diffraction are discussed (section 2.4). Hence we describe the instrumentation (section 2.5), the data collection (section 2.6), the phase problem (section 2.7), the structure refinement (section 2.8) and the validation methods (section 2.9).

## **2.2 Protein crystallization**

Crystallization is a physical process which brings to the ordered state of the matter called crystal state, where a molecular entity is repeated by following symmetry rules. Protein crystals form molecular crystals, where the repeated molecular entities are protein molecules. The obtainment of high-quality crystals of proteins is usually hampered by their large size and flexibility, and by the weakness of the interactions between them in the molecular crystal.[201, 202]

In order to obtain a protein crystal from a solution, it is necessary to reduce its solubility, in order to allow the protein molecules to reach a supersaturation state (Figure 2.1) and to assemble into crystals. The supersaturation state is characterized by more solute than the quantity that could be dissolved by the solvent under normal circumstances. This state is not stable and in presence of small perturbations leads to the precipitation of the excess of solute and to crystals, if the conditions are suitable for crystallization. This process can be explained through the protein crystallization phase diagram in Figure 2.1, where protein concentration is reported as a function of the precipitant concentration.



**Figure 2.1: Crystallization phase diagram.** A schematic phase diagram showing the solubility of a protein solution as function of precipitant concentration.

Crystal formation usually occurs in two steps: (i) nucleation (labile zone) and (ii) crystal growth (metastable zone). During the nucleation step, the formation of molecular clusters, called nuclei, occurs. that provide the surface for the subsequent crystal growth. Not all nuclei are stable and quickly dissolve, whilst the stable one grow forming crystals.[203]

### 2.2.1 Parameters affecting the crystallization process

Usually, protein crystals are obtained by reducing protein solubility, so that the supersaturation becomes accessible. Unfortunately, supersaturation is necessary but not sufficient to get crystals and several attempts are usually needed to find supersaturation conditions that lead to crystals. This time consuming process is performed by a trial and error approach.

The crystallizability of a given protein depends on its chemical and physical properties. The strategy employed to carry out the crystallization process is to guide the system towards the supersaturation state.

This can be done in different ways:

a) change of pH: proteins are polyelectrolytes and their solubility is strictly related to the pH of the system. Alterations of the pH of the system can have two effects: (i) variation in the ionization state of the amino-acids located on the surface of the protein; (ii) alteration of the nature of the interactions between the protein molecules and the solvent.[204] For many years, it was thought that the optimal pH for protein crystallization was close to the protein isoelectric point (pI). However, studies indicated that the pH at the isoelectric point favours the formation of amorphous precipitate.[201] Nowadays, the best approach is the crystallization near to physiological pH, in order to minimize the risk of denaturation.[205]

b) Use of salts: at high salt concentration, the solubility of the protein is reduced due to the salting out phenomenon. The solubility reduction is due to the alteration of the protein/solvent interactions resulting in an increased association degree of protein molecules;[204]

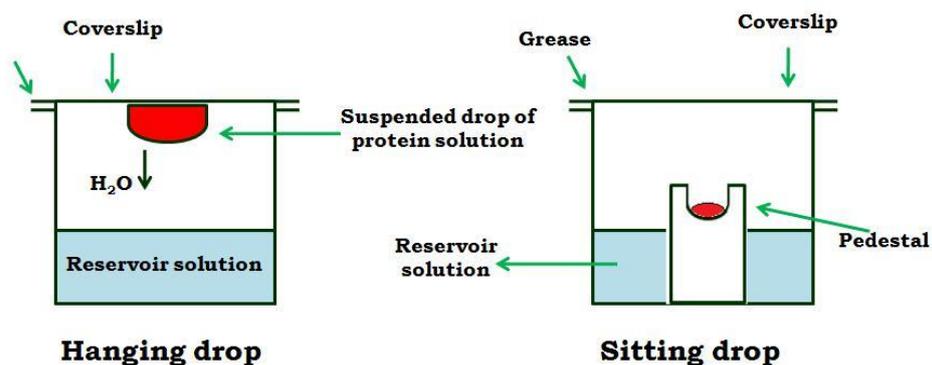
c) Use of polymers: polymers like PEG induce reorganization of solvent molecules. Protein molecules are excluded, producing a phase separation.[206]

The crystallization of a protein is largely affected by protein concentration, purity and homogeneity. The sample should be concentrated to the highest possible concentration avoiding aspecific aggregation.[207] Moreover, it has to be pure and homogeneous enough to obtain well diffracting crystals,

and to increase the reproducibility of the crystallization experiment.[204] Crystallization has been reported to occur from 0 °C to 40°C.[208] However, 4 °C and 20 °C are the most used temperatures, since higher and lower temperatures can cause the denaturation of the protein.[204]

### **2.2.2 Crystallization methods**

Different methods have been developed to grow protein crystals. The most common one is the vapour diffusion method.[209] In the vapour diffusion method, a drop containing a mixture of precipitant and protein solution is placed on a siliconized microscope glass cover and then sealed in a chamber with a reservoir solution. The latter contains the same buffer and precipitant of the drop, but in higher concentrations. Water vapour then diffuses out of the drop, but in higher concentrations. Water vapour then diffuses out of the drop. This dehydration process causes a reduction of the solvent in the drop and a consequent increase of concentration of both protein and precipitant. Equilibrium is achieved when the concentration of precipitant in the drop is equal to the concentration of precipitant in the reservoir solution. When equilibrium is reached in the nucleation zone of the phase diagram, crystal nuclei are obtained.[207] Vapour diffusion can be performed in either hanging-drop or sitting-drop.[210] Hanging-drop apparatus (Figure 2.2, left panel) involves a drop of protein solution placed on an inverted cover slip, which is then suspended above the reservoir. At variance, in sitting-drop crystallization apparatus (Figure 2.2, right panel), the drop is placed on a pedestal that is separated from the reservoir.[210]



**Figure 2.2: Vapour diffusion method of protein crystallization.** A drop containing a mixture of protein sample and reservoir solution is suspended on a glass cover slip (hanging drop, left panel) or placed on a pedestal (sitting drop, right panel) in vapor equilibration with the reservoir solution.

### 2.2.3 Approaches to find crystallization conditions

Usually, crystallization conditions for a particular protein are screened using different commercially available crystallization suites (QIAGEN, Hampton Research, Molecular Dimension). Using these kits, it is possible to test the solubility of a protein with respect to a wide range of different precipitating agents and buffers.[211] The crystallization trays are performed on a microscopic scale using the sitting or hanging drop vapor diffusion technique. The hits from the initial screening are used as a starting point for a subsequent optimization, in order to obtain a large single crystal that is suitable for diffracting experiments.

There are different approaches used for the optimization of a crystallization condition obtained from a screen. A very simple approach consists of changing a single parameter (i.e. salt concentration, pH) while keeping the others constant. This method is called “star search”[212] and it allows to monitor whether a variation of a certain parameter produces an

improvement in the crystallization condition or not. If there is no or negligible effect on the crystallization from the modification of a parameter, that variation can be omitted in the subsequent round of optimization. If the modification produces improvements in the crystallization, in the second round the modified parameter is kept fixed while the other are screened. This is repeated until the procedure is converged on the optimal value for all the parameters.

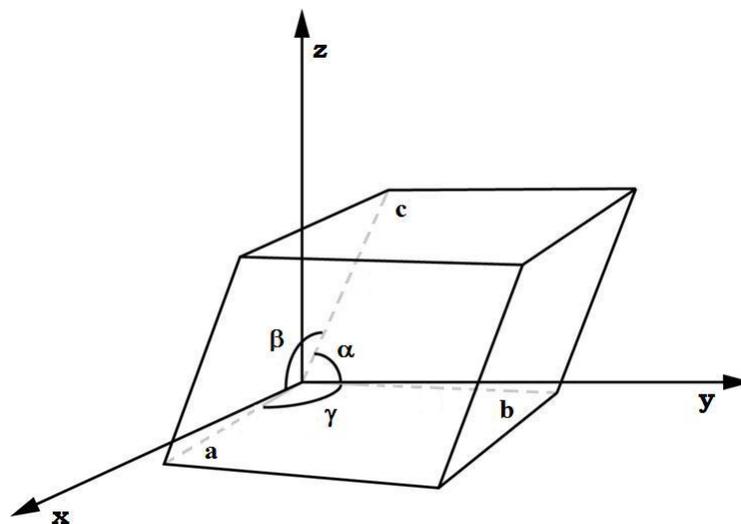
An alternative approach in the design of the crystallization experiment is the so called incomplete factorial method.[213] This technique features the sampling of a multidimensional space in which each dimension is associated with a given parameter (factor). Each factor is a variable that can assume different values for each experiment involving the given factor. There are several advantages deriving from the application of the incomplete factorial approach to protein crystallization: the number of trials can be significantly reduced and a large number of crystallization conditions can be assessed with few experiments. This is accomplished by varying more than one parameter simultaneously in a given experiment; this allows to readily determine the factors that are critical for crystallization.[214]

Once initial hits have been obtained, optimization techniques can be used to improve their quality.[215] Seeding is one of the most used optimization techniques. In seeding, nuclei or small crystals are introduced in a solution in metastable conditions, allowing to separate the nucleation step from the crystal growth. In macro-seeding, a 5-50  $\mu\text{m}$  crystal is transferred from the initial drop into the new crystallization drop. At variance, in the micro-seeding technique, a suspension of microcrystal seeds is prepared by

crushing a bigger crystal in a stabilizing solution.[216] In streak seeding, animal hairs are used to crush and transfer micro crystals in a fresh drop.[212]

### 2.3 Crystal structure

A crystal is a periodic array of atoms, ions or molecules in three dimensions.[217] Each repeated entity in the crystal lattice is described by a node. In this regular packing, we can identify the unit cell. A unit cell represents the smallest unit of volume, whose repetition in the three dimensions results in the entire crystalline lattice. The unit cell can be specified by three vectors, which define the length of the cell edges ( $\mathbf{a}$ ,  $\mathbf{b}$ ,  $\mathbf{c}$ ) and the angles between them ( $\alpha$ ,  $\beta$ ,  $\gamma$ ) (Figure 2.3).



**Figure 2.3: Unit cell in the crystal lattice.** The figure is adapted from ref. [218]

Three nodes of the crystal lattice form a crystallographic plane, which can be described by three integers,  $h$ ,  $k$ , and  $l$ , called Miller indices, which

define a family of planes that cut the **a**, **b** and **c** edges  $h, k$  and  $l$ , times respectively. Miller indices also describe a reciprocal space vector perpendicular to a set of parallel planes in the real space (direct space), where the reciprocal space is the Fourier Transformation of the real space.

The asymmetric unit is the smallest unit of volume that, upon application of the symmetry operation, can reproduce the unit cell.[219] Crystallographic space groups, that arise from the combination of 32 possible point groups and 14 Bravais space lattices, are used to describe the crystal symmetry. Only 65 space group of the 230 possible space groups are compatible with protein crystals, since no inversion and reflection symmetry operations are possible, due to the chirality of the amino acids.[220]

Multiple copies of a molecule or subunit within the asymmetric unit are related by local symmetry, which is not compatible with the periodicity of the crystal pattern, and this is referred to as non crystallographic symmetry (NCS).[221] NCS operates only at local level and the superimposition may be inexact because subunits in different environments are never identical. The presence of NCS can be advantageous in the process of structure determination because it gives information on the relationships between density in different parts of the crystal and between coordinates in different parts of the structure.[222]

## **2.4 Principles of X-ray diffraction by a single crystal**

When an electromagnetic wave encounters an obstacle, it is deviated from its trajectory. This phenomenon is named scattering and, in the case of obstacles having size comparable to the wavelength of the hitting beam, it is

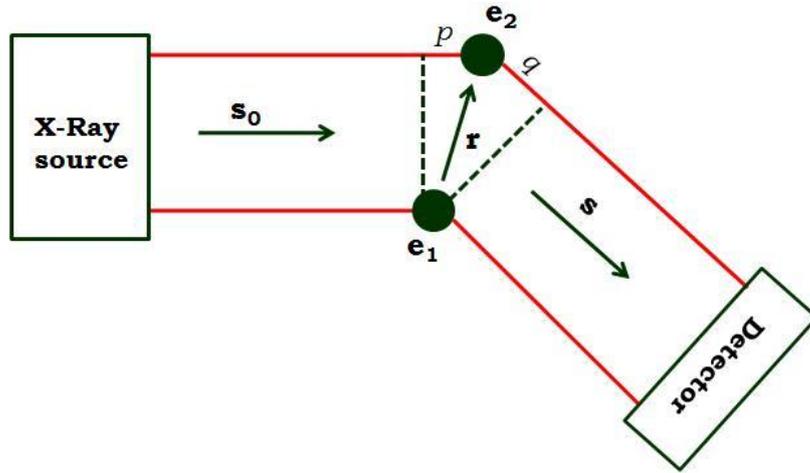
named diffraction. As a result of the collision, the emitted wavelength is diffracted in another direction that will depend on the angle between the incident and diffracted beam and the particular disposition of the object in the space.

#### 2.4.1 Scattering by two electrons

Let us consider a system of two electrons,  $e_1$  and  $e_2$ , whose position is stable in the time. A planar X-ray wave, having wavelength  $\lambda$  and represented by the vector  $s_0$  with length  $1/\lambda$ , hits the system and it is diffracted in the direction of the vector  $s$  having the same length  $1/\lambda$ . The difference in position between the two electrons is  $\mathbf{r}$ . Both the incident and diffracted beams will follow a longer path for electron  $e_2$ . The phase difference between the two diffused waves is represented by  $p + q$ . The phase difference is:

$$(p + q) = 2\pi\mathbf{r} \cdot \mathbf{r}^* \quad (2.1)$$

where  $\mathbf{r}^* = \frac{s-s_0}{\lambda}$ .



**Figure 2.4: Scattering by two electrons.** X-ray photons leave the X-ray source at the same moment, are each scattered by an electron, and then meet a distant X-ray detector. This figure is adapted from ref. [202].

### 2.4.2 Scattering by an atom

If we consider an atom, the total scattering depends on the number of its electrons and their relative position. If we assume that the origin of the system  $O$  is in the nucleus, then the amplitude of the scattered wave, also called the atomic structure factor  $f(\mathbf{r}^*)$ , is:

$$f(\mathbf{r}^*) = \int_r \rho(\mathbf{r}) \exp[2\pi i \mathbf{r} \cdot \mathbf{r}^*] dr \quad (2.2)$$

where  $\rho(\mathbf{r})$  is integrated over the entire space  $\mathbf{r}$ .

### 2.4.3 Scattering by a unit cell

In a unit cell with  $n$  atoms at position  $r_j$  with respect to the origin, the whole electron density is described by:

$$\rho_M(\mathbf{r}) = \sum_{j=1}^N \rho_j(\mathbf{r} - \mathbf{r}_j) \quad (2.3)$$

And the total scattering is:

$$\begin{aligned} F_M(\mathbf{r}^*) &= \int_r \sum_{j=1}^N \rho_j(\mathbf{r} - \mathbf{r}_j) \exp[2\pi i \mathbf{r} \cdot \mathbf{r}^*] dr = \\ &= \sum_{j=1}^N \int_r \rho_j(\mathbf{r} - \mathbf{r}_j) \exp[2\pi i \mathbf{r} \cdot \mathbf{r}^*] dr = \\ &= \sum_{j=1}^N f_j(\mathbf{r}^*) \exp[2\pi i \mathbf{r}_j \cdot \mathbf{r}^*] \end{aligned} \quad (2.4)$$

$F_M(\mathbf{r}^*)$  is called structure factor and it depends on the arrangement of the atoms in the unit cell.

#### 2.4.4 Scattering by a crystal

The scattering for a whole crystal can be calculated as the sum of the structure factors of all the unit cells that constitute the crystal with respect to a common origin  $O$ . Suppose that the crystal contains a large number of unit cells,  $n_1$  in  $\mathbf{a}$  direction,  $n_2$  in  $\mathbf{b}$  direction and  $n_3$  in  $\mathbf{c}$  direction. If the position of a certain unit cell respect to the common origin  $O$  is specified by the position  $t \cdot \mathbf{a} + u \cdot \mathbf{b} + v \cdot \mathbf{c}$ , with  $t, u$  and  $v$  whole numbers, the total structure factor for the crystal,  $K(\mathbf{r}^*)$ , is:

$$K(\mathbf{r}^*) = F_M(\mathbf{r}^*) \times \sum_{t=0}^{n_1} \exp[2\pi i t \mathbf{a} \cdot \mathbf{r}^*] \times \sum_{u=0}^{n_2} \exp[2\pi i u \mathbf{b} \cdot \mathbf{r}^*] \times \sum_{v=0}^{n_3} \exp[2\pi i v \mathbf{c} \cdot \mathbf{r}^*] \quad (2.5)$$

where the summations are not zero if the scalar products  $\mathbf{a} \cdot \mathbf{r}^*$ ,  $\mathbf{b} \cdot \mathbf{r}^*$ , and  $\mathbf{c} \cdot \mathbf{r}^*$  are integers.

In conclusion, a crystal will scatter X-rays if the following conditions are satisfied:

$$\mathbf{a} \cdot \mathbf{r}^* = h \quad (2.6)$$

$$\mathbf{b} \cdot \mathbf{r}^* = k \quad (2.7)$$

$$\mathbf{c} \cdot \mathbf{r}^* = l \quad (2.8)$$

Equations 2.6-2.8 are known as Laue conditions and  $h, k$  and  $l$  are integers and they can be positive, negative or zero. If the Laue conditions are satisfied, all the unit cells scatter in phase and the amplitude of the total scattered wave is proportional to the structure factor  $F_M(\mathbf{r}^*)$  and to the number of unit cells in the crystal.

#### 2.4.5 Bragg's Law

The values  $h, k$  and  $l$  of the Laue conditions represent the Miller indices for a certain family of planes with an interplanar distance  $d_{hkl}$ . In 1912, Sir William Lawrence Bragg [223] interpreted the diffraction as a reflection phenomenon. Particularly, he showed that a set of parallel planes with Miller indices  $(hkl)$  and interplanar spacing  $d_{hkl}$  produces a diffracted beam when the incident ray of wavelength  $\lambda$  meets the planes at angle  $\theta$  and is reflected at the same angle only if the following condition is satisfied:

$$n\lambda = 2d_{hkl}\sin\theta \quad (2.9)$$

where  $n$  is an integer.

Let us consider two monochromatic waves,  $R_1$  and  $R_2$ , that are reflected by a set of parallel planes with Miller indices  $(hkl)$  and interplanar distance  $d_{hkl}$  at angle  $\theta$  (Figure 2.5).  $2BC$  is the path difference between the two rays. We consider now the  $ABC$  triangle.  $AB$  is perpendicular to the atomic plane and  $AC$  is perpendicular to  $R_2$ . Therefore, the angle  $C\hat{A}B = \theta$ . Since  $ABC$  is a right triangle, we have:

$$\sin \theta = \frac{BC}{AB} \quad (2.10)$$

$$AB = d_{hkl} \quad (2.11)$$

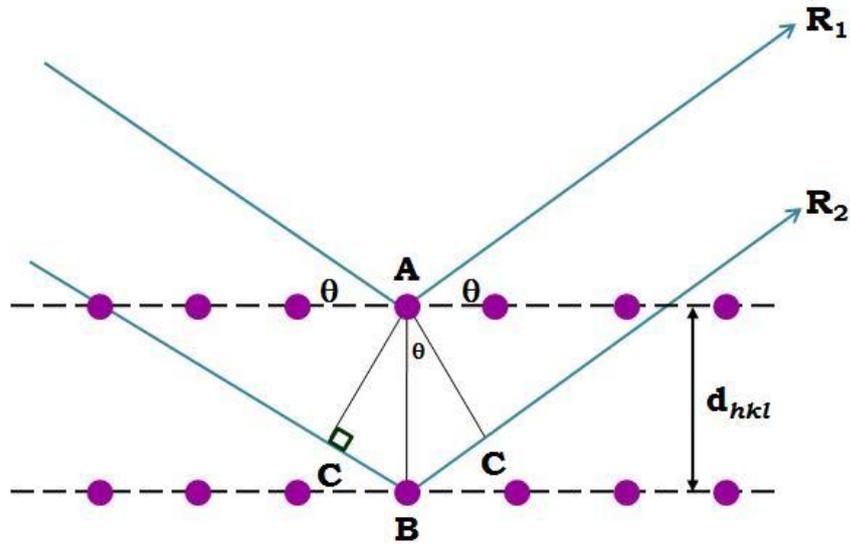
$$\sin \theta = \frac{BC}{d_{hkl}} \quad (2.12)$$

It follows that:

$$BC = \sin \theta d_{hkl} \quad (2.13)$$

The additional distance  $2BC$  travelled by  $R_2$  is equal to  $2d_{hkl} \sin \theta$ . If the path length difference is equal to an integer number of wavelength  $n\lambda$  a constructive interference occurs. In this case, we can write:

$$n\lambda = 2d_{hkl} \sin \theta \quad (2.14)$$



**Figure 2.5: Bragg diffraction.** Two beams  $R_1$  and  $R_2$  of identical wavelength approach a family of planes with interplanar distance  $d_{hkl}$ . Beam  $R_2$  needs to traverse an extra distance  $2BC = 2d_{hkl} \sin \theta$ . Constructive interference occurs when the extra distance is equal to an integer multiple of the wavelength radiation.

#### 2.4.6 The reciprocal lattice

A useful concept in crystallography is the reciprocal lattice. The reciprocal lattice is a mathematical construct developed by Ewald [224] which allows to rebuild the diffraction pattern in a simple way. While the real lattice describes the periodic structure of the lattice, the reciprocal space determines how the periodic structure interacts with the wave. It represents the space occupied by the reflections. As for the real lattice, we can speak of unit cell in the reciprocal lattice, defined by three vectors,  $\mathbf{a}^*$ ,  $\mathbf{b}^*$  and  $\mathbf{c}^*$ , whose lengths are the reciprocals of the length of the corresponding real cell edges  $\mathbf{a}$ ,  $\mathbf{b}$  and  $\mathbf{c}$ :

$$\mathbf{a}^* = 1/\mathbf{a} \quad (2.15)$$

$$\mathbf{b}^* = 1/\mathbf{b} \quad (2.16)$$

$$\mathbf{c}^* = 1/\mathbf{c} \quad (2.17)$$

Because of the reciprocity relation between real and reciprocal cell, if the real cell is small the reciprocal cell will be large and *vice versa*.

#### 2.4.7 Bragg's law in the reciprocal space and the Ewald construction

The reciprocal lattice allows to determine the direction of a diffracted beam from a set of planes. A useful way to understand the occurrence of diffraction spots is the Ewald construction or sphere of reflections, conceived by Paul Peter Ewald.[224]

Consider an  $\mathbf{a}^*\mathbf{b}^*$  plane in the two-dimensional reciprocal lattice. Let us assume that an X-ray beam meets this plane. We choose  $O$  as the origin of the reciprocal space. We draw a circle passing for  $O$  with a of  $1/\lambda$  radius and with center  $C$  (Figure 2.6). If we assume to rotate the crystal about  $O$ , also the reciprocal lattice will be rotated about  $O$  bringing the lattice point  $P$  in contact with the circle. We now consider the triangle  $OPB$ . The angle in  $B$  is  $\theta$  and

$$\sin \theta = \frac{OP}{OB} = \frac{OP}{2/\lambda} \quad (2.18)$$

Rearranging the equation 2.18:

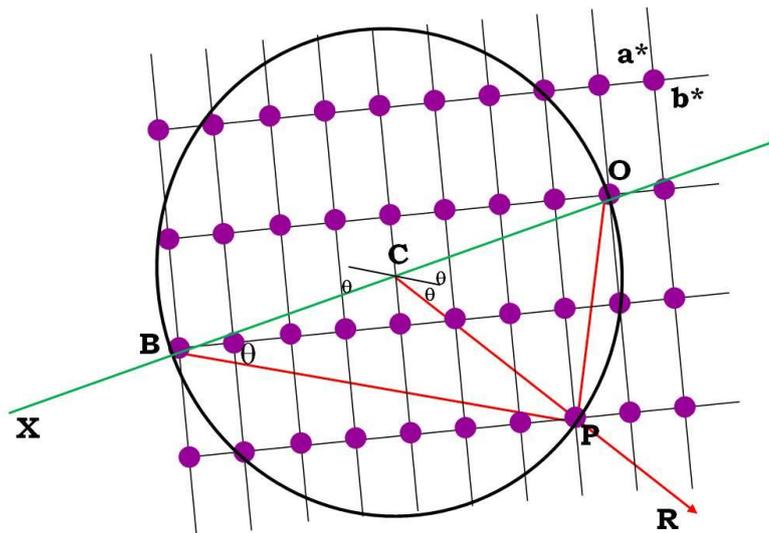
$$2 \frac{1}{OP} \sin \theta = \lambda \quad (2.19)$$

Since  $OP = 1/d_{hkl}$ , equation 2.19 can be rewritten as follows:

$$2d_{hkl} \sin \theta = \lambda \quad (2.20)$$

Equation 2.20 represents the Bragg's law for  $n = 1$

Whenever the crystal is rotated clockwise about the origin  $O$ , a point of the reciprocal lattice comes in contact with the circle of radius  $1/\lambda$ , the Bragg's law is satisfied and a reflection occurs.



**Figure 2.6: The Ewald sphere construction.** The ray R emerges from the crystal when the point P intersects the circle. This figure is adapted from ref.[210]

## 2.5 Instrumentation

A diffractometer measures the position and the intensity of the diffracted X-ray beams. The main components needed for the collection of diffraction

data are: an intense X-ray source, a monochromator, a goniostat to orient the crystal in the X-ray beam and a detector to collect the diffracted spots.

A diffractometer includes also a cryo cooling system to keep the crystal at cryogenic temperatures during X-ray exposure, to reduce the radiation damage.

Three different types of X-ray generators are mainly used: X-ray tubes, rotating anode tubes and particles storage rings (synchrotrons).

Synchrotrons are a source of intense electromagnetic radiations emitted by charged particles brought at high velocities. Electrons are pre-accelerated in the booster ring and then injected into the storage ring. Third generation synchrotrons uses special arrays of magnets called insertion devices to change the direction of the electrons, allowing the emission of an intense electro-magnetic radiation. The generated light can then be channelled out of the storage ring and into the experimental stations, called beamlines.

Diffracted rays are collected by a detector. Modern detectors use charge-coupled device (CCD) technology to transform X-ray photons into an electrical signal, which is sent to a computer.[200]

## **2.6 Data collection**

The main aim of data collection is to obtain a set of measured and indexed intensities for as many reflections as possible. The crystal sample is mounted on the goniometer head and then fixed on the goniostat of the diffractometer. In presence of X-ray exposure, diffraction images can be collected by rotating the crystal. These images represent the starting point for the structure determination process.[225]

### **2.6.1 Indexing**

The indexing process allows to determine the crystal lattice parameters ( $\mathbf{a}, \mathbf{b}, \mathbf{c}, \alpha, \beta, \gamma,$ ) and the relative crystal orientation, using the relative positions of the spots on one or more images. Then, an integration of the intensity of the spots compatible with the previously found unit cell is performed. A list of reflections along with the relative Miller indices, intensities ( $I$ ) and associated error ( $\sigma I$ ) is made. These procedures are usually performed by XDS,[30] Mosflm,[226] and HKL2000[227] softwares.

### **2.6.2 Scaling**

The integrated intensities from different images collected during the diffraction experiments are not all on the same scale, because of a certain number of physical factors (detector properties, detector settings). Scaling is a procedure that align all the integrated values of the different images on a common scale, producing a unique data set which is internally consistent.[228]

## **2.7 The phase problem.**

The main goal of protein crystallography is to reconstruct the electron density  $\rho(xyz)$  at every position  $(xyz)$  in the unit cell.  $\rho(xyz)$  is expressed as the FT of the structure factor  $F(hkl)$ :

$$\rho(xyz) = \frac{1}{V} \sum_{hkl} |F(hkl)| \exp[-2\pi i(hx + ky + lz) + i\alpha(hkl)] \quad (2.21)$$

where  $|F(hkl)|$  is the structure factor amplitude of  $(hkl)$  reflection and  $\alpha(hkl)$  is the phase angle. Equation 2.21 shows that the calculation of  $\rho(xyz)$  requires two terms: the structure-factor amplitudes and the relative phase angles. However, when waves are diffracted by a crystal, while the amplitude of the wave  $F(hkl)$  can be easily calculated from the spot intensities ( $I(hkl) = |F(hkl)|^2$ ), all the information relative to the phase are lost. Therefore, in order to rebuild the electron density we need to recover the phase information.

So far, four different techniques have been developed to solve the phase problem in X-ray crystallography:

a) Isomorphous Replacement Method (SIR, MIR), in which information about unknown phases can be obtained by making changes in the unit cell and measuring the effect on the diffraction pattern. In practice, this can be done by soaking the native crystal with heavy-atoms, making sure that their presence does not disturb the structure of the protein.[229]

b) Multiple wavelength Anomalous Diffraction method (MAD). In this case, the phase problem is solved using strong anomalous signals from atoms in the protein structure itself. Every atom has few ‘absorption edges’. Around these edges, the scattering varies rapidly in amplitude and phase with wavelength. By varying the wavelength around the absorption edge for an atomic type, the contribution to the total scattering from the atom changes. In practice, different wavelengths can be used to have better phase estimates. In SAD (single-wavelength anomalous diffraction) a single

dataset at appropriate wavelength is used. The time spent by the crystal in the beam is minimized. Therefore the chances of potential radiation damages are largely reduced.[230]

c) Direct methods. These methods are valuable for small and intermediate-size molecules where the atoms are usually well ordered.[231]

d) Molecular Replacement (MR). In this technique, phases are calculated using the phases of a known structure with high sequence identity.[232]

MR will be discussed in more detail in section 2.7.2 as this method has been adopted to solve the phase problem for the protein studied in this doctoral dissertation. Nevertheless, in the next section, we first introduce the Patterson function because of its importance as a valid tool in the early stage of phase determination.

### 2.7.1 The Patterson function

The Patterson function, introduced for the first time in 1935 by A. L. Patterson [233], is defined as the Fourier sum of the squared amplitudes whose values are proportional to the diffracted intensities, and thus there is no need to provide any information about the phases:

$$P(uvw) = \frac{1}{V} \sum_{hkl} |F(hkl)|^2 \cos[2\pi(hu + kv + lw)] \quad (2.22)$$

where the coordinates  $(uvw)$  locate a point in the Patterson map and  $V$  is the volume of the unit cell. The resulting Patterson map is a map that shows peaks at every position corresponding to an inter-atomic vector in the

structure. For proteins with  $N$  atoms in the unit cell, the Patterson map will have  $N(N - 1)$  interatomic vectors, many of them being overlapped. Typically, Patterson map provides a great advantage in detecting heavy atoms in the crystal structure.

### 2.7.2 The Molecular Replacement (MR)

The Molecular Replacement, developed by Rossmann and Blow[232] in 1972, represents a valid alternative to SIR and MAD. In this technique, the phases from a known protein can be used to build an initial model for the protein under examination, if the phasing model and the target protein have high homology sequence.

When the phasing and the target protein are isomorphous, the electron density map can be calculated according to the equation (2.23):

$$\rho(xyz) = \frac{1}{V} \sum_{hkl} |F(hkl)_{target}| \exp [-2\pi i(hx + ky + lz) + i\alpha(hkl)_{model}] \quad (2.23)$$

where  $|F(hkl)_{target}|$  are calculated from the native intensities of the new protein, and  $i\alpha(hkl)_{model}$  are the phases of the phasing model.

If the phasing model is not isomorphous with the target protein, the MR is obtained by superimposition of the phasing model to the target protein in the target unit cell.

A Patterson map is calculated for both the unknown and known structure: if they are oriented in the same way in the unit cell of the same dimension and with the same symmetry, the two maps will be similar. The procedure is

simplified if the calculation of the solution is carried out in two steps: (i) the search for the best orientation and (ii) the search for the best position. This is done by the application of the rotation and translation function, respectively.

A rotation function is used for the comparison of orientations. The Patterson maps for the phasing and the target model are evaluated in various orientations. If the two Patterson maps have many coincident peaks, the function will exhibit a maximum. The position search is carried out using a translation function. In this way, the correctly oriented model is translated to the correct coordinates within the asymmetric unit. In practice, the model is moved, a new Patterson map is evaluated and it is compared to the target protein Patterson map until convergence.

## **2.8 Model building, structure refinement**

### **2.8.1 Electron density modification**

Once the initial phases have been obtained, the electron density map can be calculated. The quality of the initial map will depend on the quality of the estimated phases: in fact, large errors in the density map can be found if the calculated phases are not good enough. However, the initial electron density map can be improved using the density modification techniques.[234] These techniques are based on the different physical properties of the protein and of the disordered solvent.

Among the density modification methods developed, the solvent flattening[235] is a powerful tool for improving crystallographic phases

obtained at medium resolution. Solvent flattening is an iterative method based on the fact that in the disordered solvent region the electron density is constant ( $0.33 \text{ e}^{-1}/\text{\AA}^3$ ) while for proteins the average density is around  $0.44 \text{ e}^{-1}/\text{\AA}^3$ . The first step to improve the quality of the map by solvent flattening is to define the solvent regions (or masked region). A 3D grid is superimposed to the electron map. At each point, the electron density is computed in a sphere of  $10 \text{ \AA}$  radius. If the value of the electron density is above a given cut-off, it is belonging to the protein; if the electron density is below the cut-off value, it belongs to the solvent. Hence, new structure factors are calculated and new phases are generated. The new structure factors and phases are used to calculate a new map and process can be continued until convergence has been reached.

If the asymmetric unit contain more than one identical molecule, this information can be used for NCS averaging. NCS averaging consists in a superimposition of the electron densities corresponding to the molecules of the asymmetric unit, followed by summation of the maps within the superimposed molecules. Averaging significantly improves the signal to noise ratio of a typical electron density map. This can be extremely useful in the early stages of molecular replacement as a method to reduce phase bias.[222]

### **2.8.2 Model building**

The following step in determining a crystal structure is the model building, the process in which the electron density is interpreted in terms of a set of

atomic coordinates. First, the protein backbone is fitted in the electron density and then, if the resolution is sufficiently high, the amino-acid sequence is inserted.

### 2.8.3 Model refinement

Refinement is the process of adjusting the model, in order to find the maximum agreement between the calculated and observed structural data. This agreement is usually represented by the work  $R_{factor}$  index defined as:

$$R_{factor} = \frac{\sum_{hkl} ||F_{obs}| - k|F_{calc}||}{\sum_{hkl} |F_{obs}|} \times 100\% \quad (2.24)$$

Where  $|F_{obs}|$  are the observed structure factors and  $|F_{calc}|$  are the calculated structure factors for a certain reflection with Miller indices ( $hkl$ ).

The  $R_{factor}$  is calculated using only 5% of the reflections that are excluded from refinement process; therefore, it gives an independent measure of the refinement progress. The drop of the  $R_{factor}$  during the refinement is important as it indicates that the refinement is successful. Final values of  $R_{factor}$  lie between 15% and 25%. However, it has been shown that  $R_{factor}$  can reach low values in the refinement stage for structures that appear to be incorrect. The reason beyond this is that the number of parameters taken into consideration is too high. For this reason, Brünger introduced a new index, named  $R_{factor}^{free}$ , which is unbiased from the refinement.[236]

In this case, the reflections are divided into two sets: a test set of unique reflections and a working set. The test set is composed of 5-10% of

observed reflections that are excluded from the structure refinement, while the working set is given by the remaining reflections.

The  $R_{factor}^{free}$  is defined as follow:

$$R_{factor}^{free} = \frac{\sum_{hkl \in T} ||F_{obs}| - k|F_{calc}||}{\sum_{hkl \in T} |F_{obs}|} \quad (2.25)$$

Where  $hkl \in T$  means all the reflections belong to the test set.

If during the refinement stage the quality of the model improves, then both the  $R_{factor}^{free}$  and  $R_{factor}$  will decrease; if, instead, the  $R_{factor}^{free}$  increases, it means that there is a problem in the refinement.

The parameters to be refined against the experimental data are the atomic coordinates and B-factors for each atom, together with overall parameters, such as bulk solvent and anisotropy corrections. The B-factor, or temperature factor,[237] is a factor that describe the attenuation of X-ray scattering caused by the thermal motion of the atoms. B-factors of incorrectly built atoms tend to be higher if compared to the ones of correctly built atoms.

Most of the modern refinement programs are based on the maximum likelihood method, which imposes restraints on the bond distances, angles, torsions and temperature factors of the atomic model.

## 2.9 Validation

After the molecular model of the protein has been refined, a wrong interpretation of the electron density map could still lead to some errors

(particularly in those regions where the density is weak). A qualitative and quantitative analysis of the final model is performed in the last step: the validation. Validation methods detect inconsistencies in the final model, based on information that were not used during the refinement process.

### **2.9.1 Ramachandran plot**

The stereochemistry of the main chain can be checked through the Ramachandran plot,[238] that allows to visualize the dihedral angles for each residue. The Ramachandran plot is a square matrix where the horizontal axis on the plot shows  $\varphi$  value, while the vertical shows  $\psi$  values.  $\varphi$  and  $\psi$  values for each residue in a protein are represented as dots in the plot. Due to steric reasons, only certain values of  $\varphi$  and  $\psi$  are allowed. The Ramachandran plot is thus a good indicator of the quality of a structure.

### **2.9.2 Stereochemistry validation**

Several programs have been developed for the validation of molecular structures. These programs evaluate the quality of the final model in different ways:

a) they assess the validity of the backbone conformation as well as side chains, by comparing the bond lengths, bond angles and planarities against a library of ideal values with their root-mean-square deviations;

b) they use global indicators, like  $R_{factor}$  and  $R_{factor}^{free}$  values, to signal major errors, to prevent over fitting and to monitor the progress of the rebuilding and refinement process.

Among the different validation softwares, PRO\_CHECK[239] is the most commonly used. The PRO\_CHECK suite provides an exhaustive analysis of the stereochemistry of the protein structure. Its output consists of a detailed analysis performed residue-by-residue, but also an assessment of the overall quality of the structure by comparison with other refined structures deposited in PDB.

## CHAPTER 3

### MATERIAL AND METHODS

#### 3.1 General methods

##### 3.1.1 Bacterial transformation

For every bacterial transformation, 50-100 ng of recombinant plasmid DNA were added to 50  $\mu$ L of competent cells and incubated on ice for 30 min. Afterwards, the cells were heat shocked for 1 min at 42 °C and chilled on ice for 15 min. 600  $\mu$ L of Lysogeny Broth (LB) media were added and incubated at 37 °C, shaking at 180 rpm for 1 h. Lastly, 50-100  $\mu$ L of cells were spread onto LB agar plates containing the proper antibiotic (cf. Table 3.1) and incubated at 37 °C overnight.

**Table 3.1:** Antibiotics used for general transformation of the constructs in Novo Blue cells

Construct	Antibiotic
MBP- $\beta$ FXIIa in pMT-PURO vector	100 $\mu$ g/ml ampicillin
$\beta$ FXIIa in pMT-PURO vector	100 $\mu$ g/ml ampicillin
$\beta$ FXIIa in pCOLD-TF vector	100 $\mu$ g/ml ampicillin
Ecotin in pET-26b(+) vector	50 $\mu$ g/ml kanamycin

### **3.1.2 Recovery of the plasmid DNA by Miniprep and sequencing**

Single colonies were picked from LB agar plates and transferred in 5 mL of LB medium containing the proper antibiotic and incubated at 37 °C, at 180 rpm overnight. Plasmid DNA was recovered using the “GenElute Plasma MiniPrep Kit” (Sigma). The sequence was verified by DNA sequencing (Sanger sequencing) performed at the DNA sequencing facility of the University of Nottingham.

### **3.1.3 Agarose gel electrophoresis**

DNA fragments were separated according to their molecular weight by agarose gel electrophoresis. 1 g of agarose powder (SIGMA) was added to 100 mL of 1X TBE (89 mM Tris base, 89 mM boric acid, 2 mM EDTA) buffer, mixed and heated in a microwave until the powder was completely dissolved. The prepared agarose solution (1%) was then poured into an appropriate gel casting tray with a previously inserted comb, to allow the formation of the wells where the sample is loaded into, and let the solution solidify at room temperature. The electrophoresis chamber was then filled with 1X TBE to cover the gel. 20 µL of DNA sample were mixed with 4 µL of 6X Gel Loading Dye (NEB) and 4 µL of 1X Syber Green and then loaded onto the gel, along with 10 µL of 1 kb DNA ladder [New England Biolabs (NEB)] in the first well. The gel was run at 100 V for 60 min and it was then visualized under UV light.

### 3.1.4 Sodium Dodecyl Sulphate PolyAcrylamide Gel Electrophoresis (SDS-PAGE)

The separation degree and purity of every protein that has been expressed and purified during this PhD project has been assessed by SDS-PAGE in both reducing and non-reducing conditions, by staining with Instant Blue Coomassie (EXPEDEON).

A 15% SDS-PAGE gel was prepared as shown in Table 3.2:

**Table 3.2:** Solutions for preparing resolving and stacking gels for a 15% SDS-PAGE

RESOLVING GEL	
CHEMICAL COMPONENTS	VOLUME (mL)
H <sub>2</sub> O	3.7
Acrylamide/bis-acrylamide (30%/0.8% w/v)	8
1.5 M TrisHCl pH 8.8	4
10% (w/v) SDS	0.16
10% (w/v) ammonium persulphate	0.16
TEMED	16 * 10 <sup>-3</sup>

<b>STACKING GEL</b>	
<b>CHEMICAL COMPONENTS</b>	<b>VOLUME (mL)</b>
H <sub>2</sub> O	5.3
Acrylamide/bis-acrylamide (30%/0.8% w/v)	2
0.5 M TrisHCl pH 6.8	2.5
10% (w/v) SDS	0.10
10% (w/v) APS	0.10
TEMED	10 * 10 <sup>-3</sup>

Each sample (20 µL) was mixed with 6X loading buffer (375 mM Tris-HCl pH 6.8, 6% SDS, 48% glycerol, 0.03% bromophenol blue; 9% 2-mercaptoethanol for reducing conditions) , incubated at 95 °C for 5 min, in order to induce the denaturation of the protein, and then run for 60 min at 200 V. After 1 h, the gel was stained with Instant blue Coomassie stain on a shaker at room temperature for 30 min.

### **3.1.5 Western Blot**

Protein expression was confirmed by Western Blot analysis using anti-6X His tag antibody. 15  $\mu$ L of sample was mixed with 6X loading SDS Sample Buffer and incubated at 95 °C for 10 min and run on a 15% SDS gel for 1 h at 200 V. The gel was transferred onto a nitrocellulose membrane using iBlot (Invitrogen) at 20 V for 6 min. The membrane was then blocked with 1X tris-saline-buffer (TBS, 50 mM Tris, 150 mM NaCl, pH 8.0) and 1% bovine serum albumine (BSA) and placed on a shaker for 2 h at room temperature. After that, the membrane was washed with 1X tris-buffer-saline tween (TBST, 50 mM Tris, 150 mM NaCl, 0.05% Tween 20, pH 8.0), the primary antibody “Anti-6X His tag” (Catalog #MAB050) added and left overnight at 4 °C. After the primary antibody incubation, 1X TBST buffer was used to wash the membrane three times and then the secondary antibody (Goat Anti-Mouse IgG, Catalog #F0107) was added and left for incubation for 90 min. At the end of the incubation with the secondary antibody, the membrane was washed again with 1X TBST. Western blot was treated with a solution containing luminol (250 mM), p-coumaric acid (90 mM) and hydrogen peroxide (1%). WB was visualized using Luminescence image analyzer “LAS 3000”.

### **3.1.6 Determination of DNA and protein concentration**

Both DNA and protein concentrations were determined measuring the absorbance at 260 nm and 280 nm, respectively, using Thermo Scientific NanoDrop.

Protein concentration was calculated using the Lambert-Beer law:

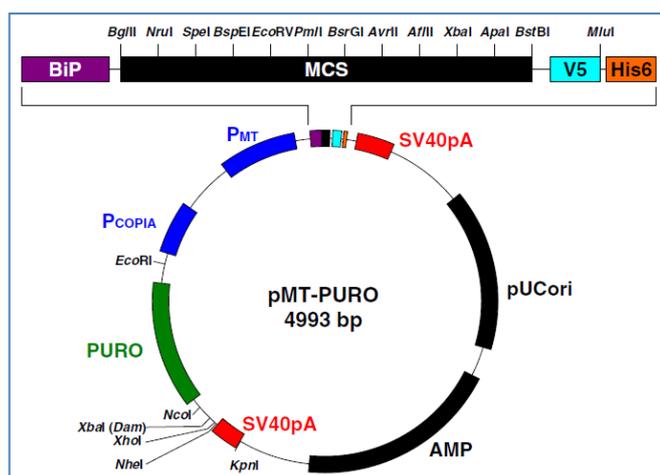
$$A = \epsilon lc \quad (3.1)$$

where  $A$  is the absorbance,  $l$  is the path length,  $c$  is the molar concentration and  $\epsilon$  is the extinction coefficient for each protein, calculated using the online program ProtParam (ExpASy).

## **3.2 Vectors**

### **3.2.1 pMT-PURO vector**

pMT-PURO vector (Figure 3.1) (Addgene) was chosen for the expression of the constructs of interest in *Drosophila Schneider* S2 cells. It has been shown that pMT-PURO vector is very efficient for the expression of coagulation proteins in insect cells system.[240] The main feature of this vector is the protein secretion signal (BiP) that facilitates the proper secretion of the recombinant protein, along with puromycin resistant cassette. The fusion of the selection vector with the expression one has two main advantages, that are (i) the creation of a stable cell line in a very short period (3 days) compared to other vectors used for expression in insect cells, and (ii) an increased stability of the cell line, thus allowing long term utilization of the same cell lines for protein expression. Moreover, the vector encodes for a 6X His tag sequence located at the C-terminus that enables the purification of the protein by metal ion affinity chromatography.



**Figure 3.1: Map of pMT-PURO.** PMT, metallothionein promoter; BiP, BiP signal sequence; MCS, multiple cloning site; V5, V5 peptide tag; His6, 6× histidine; SV40 pA, SV40 late polyadenylation signal; PCOPIA, copia promoter; Amp, ampicillin resistant gene; pUCori, pUC origin; PURO, puromycin N-acetyl-transferase (pac). There are two *XbaI* site in this vector although one of them is Dam methylated. Figure is adapted from ref. [240].

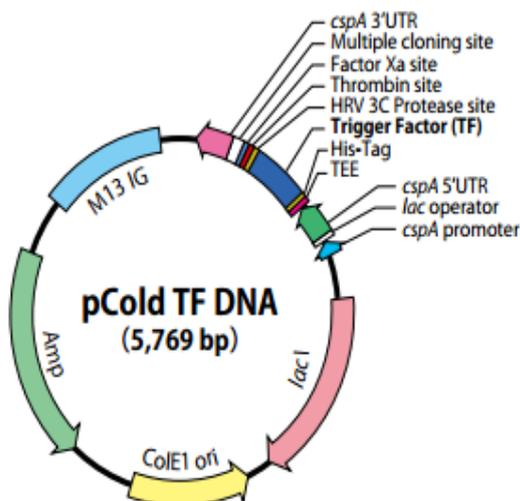
### 3.2.2 pCold-trigger factor vector

*E coli* expression system represents a very efficient tool for the production of recombinant proteins, mainly because it is easy to use and cheap. Another main feature of the *E coli* expression system is the absence of post-translational modifications, such as glycosylation. This is an advantage for crystallization purposes, since glycosylated proteins can be difficult to crystallize, due to the presence of flexible glycans. The main disadvantage is that the protein may not fold properly and deposits in the inclusion bodies. It is well established that the presence of a chaperone can help the expression of the target protein in terms of yield of protein expression and solubility.[241]

pCold-TF (Figure 3.2) is a new expression vector developed by Takara Bio which combines the advantages of the presence of a soluble tag, as trigger

factor chaperone, and the fusion cold shock expression technology. Trigger factor (TF, 48 kDa) is a prokaryotic ribosome-associated chaperone, highly expressed in *E. coli*, which acts as a guide in the initial folding step and protects the growing polypeptide chain from misfolding and aggregation.[242]

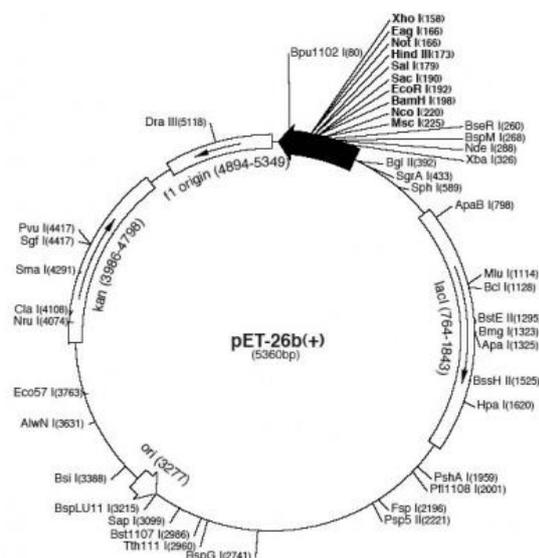
The cold shock technology increases the yield, purity and solubility of expressed protein, due to the presence of *cspA* (cold shock protein A) promoter, which regulates target protein production at lowered incubation temperatures (37 °C - 15 °C). The regulation of expression level is due to the presence of a *lac* operator inserted downstream of the *cspA* promoter. The insertion of recognition sites for HRV 3C Protease, thrombin, and FXa between trigger factor and the Multiple Cloning Site (MCS) facilitates tag removal from the expressed fusion protein.



**Figure 3.2: Map of pCold-TF vector.** M13IG, intergenic region of M13 bacteriophage.; Amp, ampicillin resistant gene; ColE1ori, ColE1 origin of replication; lacI, lac operator. Figure is adapted from ref. [243].

### 3.2.3 pET-26b(+) vector

pET vectors are widely used as bacterial expression systems, due to their ability to produce high yields of target protein. Analyzing the vector map, several elements can be found, such as the *lac* promoter that regulates the pET expression together with the *lac* operator. The latter can block transcription of T7 promoter, an ampicillin resistance cassette and a ColE1 origin of replication. IPTG (isopropyl  $\beta$ -D-1 thiogalactopyranoside) is used to induce the protein expression. When it is added to the cells, also T7 polymerase is expressed. T7 polymerase is extremely promoter-specific and transcribes only DNA downstream of a T7 promoter. Among all the series of pET vectors, pET-26b(+) (Figure 3.3) carries a N-terminal pelB signal sequence for potential periplasmic localization.



**Figure 3.3: Map for pET-26b(+) vector.** f1 origin, origin of replication; Kan, kanamycin resistant gene; ori, lacI, lac operator.

### 3.3 Tandem fixed-arm MBP/SER Strategy.

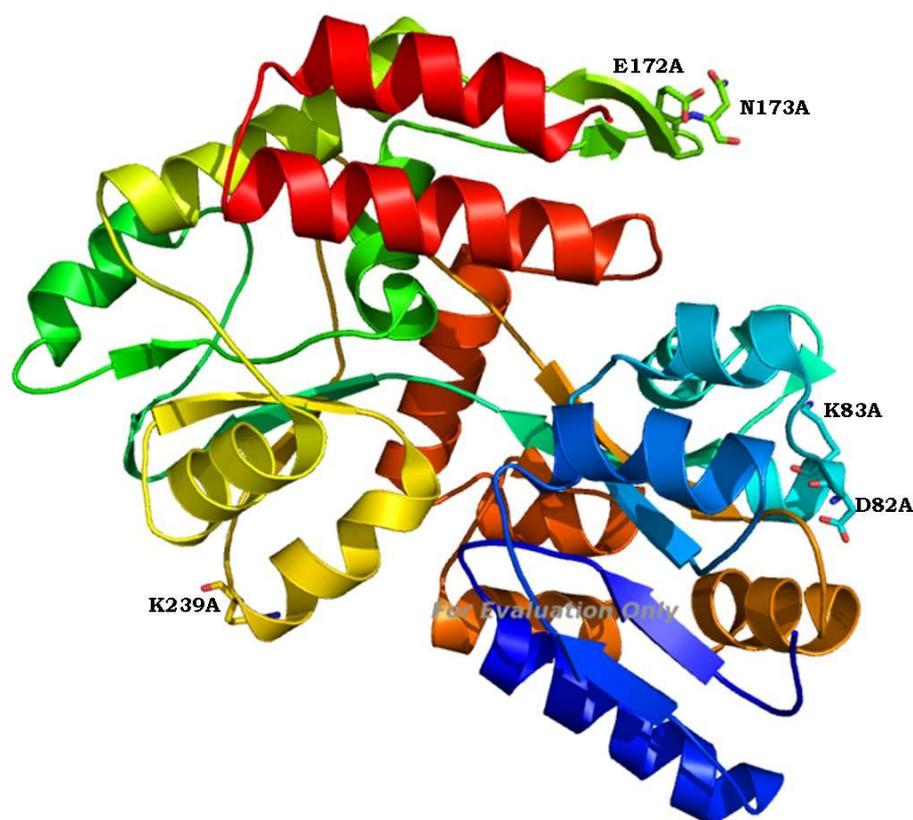
The limiting step in macromolecular crystallography is represented by the obtainment of high quality crystals that diffract at high resolution. In fact, less than 10% of proteins can generate crystals of the right size and quality, suitable for X-Ray experiments.[244] Recently, crystallographers tried to solve the crystallization problem and developed different techniques to handle difficult proteins. Moon *et al.* built up a novel approach that combines the use of a carrier protein with the surface entropy reduction (SER) technique.[244]

The presence of the carrier increases the levels of expression and solubility of the protein. The N-terminal or C-terminal of the protein of interest are provided with the carrier. The resulting effect is “chaperone-like”, meaning that the fusion protein assists the folding of the protein and also increases the yield of active product.[245, 246] Usually, the carrier protein crystallizes readily, providing the molecular surface for crystal lattice formation. Several proteins have been used as fusion proteins, primarily lysozyme, thioredoxin and maltose binding protein (MBP).

SER is a powerful technique that increases the probability of protein crystallization by reducing the entropy of the system.[247, 248] Since the molecular motions of the side chains cause a reduction of lattice contacts, because of the high entropy of the system, a selective mutation of the residues exposed on the surface allows an entropy reduction and promotes the formation of the crystal.[249]

MBP (Figure 3.4) has been chosen as carrier protein because of its high solubility and its tendency to crystallize easily. The carrier is fused to the N-

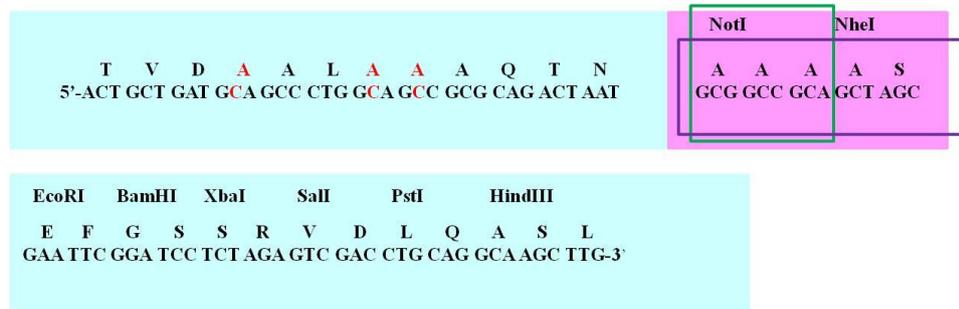
terminus of the target through a linker. It should be noted that the choice of the length of the linker is critical. In fact, it plays a role as fixed arm, merging the two proteins into an inflexible crystallization unit.[244] The tandem fixed-arm MBP/SER system was chosen as an alternative strategy to produce a large amount of  $\beta$ FXIIa, suitable for crystallization purposes.



**Figure 3.4: Cartoon representation of the crystal structure of Maltose Binding Protein (PDB code: 1HSJ, [250]).** The mutate residues (D82A, K83A, E172A, N173A, K239A) for the SER mutations are drawn in stick in rainbow colour. The figure is adapted from ref. [244].

The MBP gene was provided into pMALX vector (New England Biolabs). The amino acid sequence of MBP contains mutations (E359A, K362A, D363A) designed to increase the chances of the crystallization by reducing the entropy of the fusion system. Performing a mutation on the carrier is

advantageous, because SER mutations on each target are avoided, providing an efficient cloning protocol that requires minimal costs. The gene also contains the linker region (Figure 3.5), a sequence encoding for three Ala and the restriction sites NotI and NheI, upstream of EcoRI site.[244]

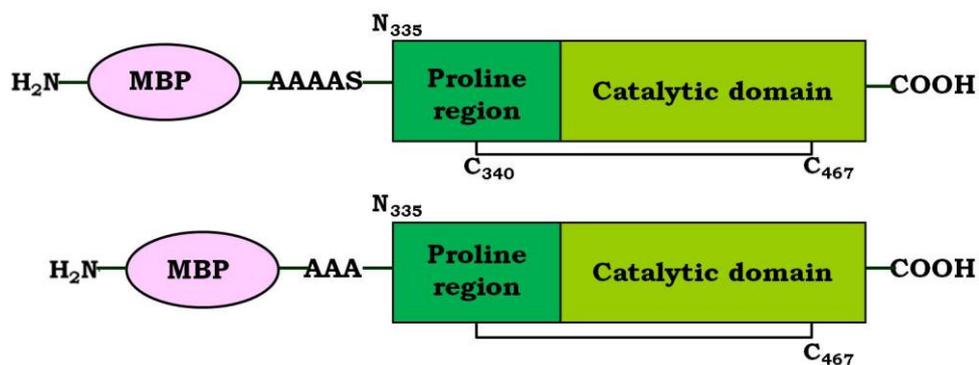


**Figure 3.5: Schematic representation of the linker region.** The MBP region is highlighted in cyan, with Ala substitution and nucleotide changes shown in red font. The pink box represents the insertion point of the restriction sites NotI and NheI. The green and purple squares highlight the linker sequences chosen for MBP- $\beta$ FXIIa. This figure is adapted from ref. [244].

### 3.4 Cloning of MBP- $\beta$ FXIIa and - $\beta$ FXIIa

#### 3.4.1 PCR for the tandem fixed-arm MBP- $\beta$ FXIIa system

Two constructs have been made, which differ for the length of the linker region. In the first construct, the linker consists of four Ala and a Ser (AAAAS), while, in the second one, is the amino acid sequence of the linker is AAA (Figure 3.6).



**Figure 3.6: Schematic representation of the constructs of  $\beta$ FXIIa having an MBP fusion tag at the N-terminus**

The cloning of MBP- $\beta$ FXIIa into pMT-PURO vector was performed using the “In-Fusion Advantage PCR Cloning KIT” protocol (Clontech), that allows joining multiple DNA fragments into any vector, avoiding the use of restriction enzymes. A truncated construct of FXII was made, encoding for the amino acid sequence starting from the residue N335. PCR primers were designed in order to share 15 bases of homology with the sequence at the ends of the linearized vector, MBP or FXII. In Table 3.3, the primers used for PCR amplification are reported:

**Table 3.3:**Primers used for the amplification of MBP and  $\beta$ FXII.

Linker	Gene amplified	Forward Primer 5'to 3'	Reverse Primer 5'to 3'
AAA	MBP	CTCGCTCGGGAGATC TAAAATCGAAGAAGG TAAACTGGTA	CGGGCCCCCGGGCGAA TTAGTCTGCGCGGCTG C'
	$\beta$ FXII	CAGACTAATGCGGCC GCAGGCCCGCTGAGC TGC'	ATGACCGGTACGCGTG CTCACGGTATGTTCGC
AAAAS	MBP	CTCGCTCGGGAGATC TAAAATCGAAGAAGG TAAACTGGA	CGGGCCCTTAAGGCTA GCTGCGGCCCGCATTAG TCTG
	$\beta$ FXII	GCGGCCGCAGCTAGC GAATTCGGCCCGCTG AGCTGC'	ATGACCGGTACGCGTG CTCACGGTATGTTCGC

PCR samples were prepared following the protocol reported in Table 3.4.

**Table 3.4:** PCR reaction mixture.

Component	Final Concentration
10X PCR Buffer	1X
5X Q-Solution	1X
dNTP mix (10 mM each)	200 mM
Primer A	500 pmol
Primer B	500 pmol
DNA template	5 ng
Enzyme	2.5U/reaction
H <sub>2</sub> O	

Each sample was brought to a final volume of 100  $\mu$ L with distilled H<sub>2</sub>O.

The PCR reactions were performed using a thermal cycler with a heated lid.

The machine was set as reported in Table 3.5:

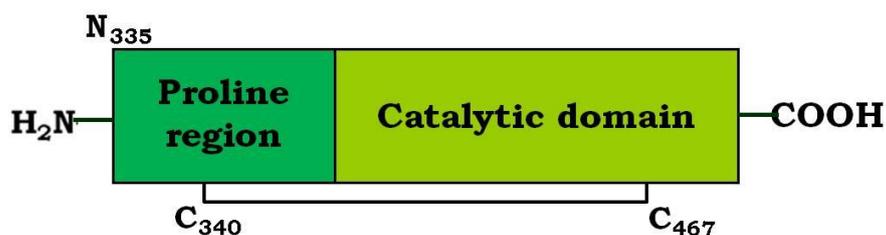
**Table 3.5:** Cycling parameters used in order to perform PCR reaction.

<b>Step</b>	<b>Time (min)</b>	<b>T (°C)</b>	<b>Number of Cycle</b>
Initial activation	5	95	
Denaturation	1	94	35
Annealing	1	55	
Extension	1	72	
Final extension	10	72	

In order to identify the optimal conditions, a gradient PCR was performed, in which the annealing temperature was varied from 50 °C to 60 °C. Moreover, the template concentration (5 ng-100 ng), as well as the DNA polymerase, were optimized, to achieve a specific PCR amplified product. The PCR products were loaded on a 1% Agarose gel and purified, using "QIAquick Gel Extraction Kit"(QIAGEN).

### **3.4.2 PCR for $\beta$ FXIIa in pMT-PURO vector**

A construct for  $\beta$ FXIIa, starting from N335 (Figure 3.7), was amplified from a DNA template containing the full-length of hFXII, by polymerase chain reaction (PCR), using the following primers (Table 3.6):



**Figure 3.7:** Schematic representation of  $\beta$ FXIIa construct.

**Table 3.6:** Primers used to produce  $\beta$ FXIIa construct. The forward and the reverse primers contain restriction site for BglII (**AGATCT**) and MluI (**ACGCGT**) restriction endonucleases respectively

Name of the construct	Forward Primer 5' to 3'	Reverse Primer 5' to 3'
$\beta$ FXIIa	GTCGAGATCTCGGAC CCCGCCTCAGTCC	GCGCACGCGTGGAAAC GGTGTGCTCCCGGA'

Each primer was designed using a general guideline, which involves:

1. A primer length ranging from 18-30 nucleotides;
2. 40-60% G/C content;
3. a melting temperature ( $T_m$ ) of 58-70 °C and the difference between each forward and reverse primers pair should be  $\leq 4$  °C, in order to allow for good amplification.
4. the last five nucleotides at the 3' end of each primer may not contain more than two C or G.

PCR samples were prepared following the protocol reported in Table 3.4. PCR reactions were performed using a thermal cycler with a heated lid, programmed as reported in Table 3.5. DNA fragments were subsequently

loaded onto 1% Agarose gel and purified using "QIAquick Gel Extraction Kit"(QIAGEN).

### 3.4.3 Ligation of MBP-βFXIIa in pMT-PURO vector

PCR products were ligated into pMT-PURO expression plasmid between BglII and MluI restriction sites. The pMT-PURO vector[240] was digested with 1 U/μL of DNA by BglII and MluI restriction endonucleases in Buffer 3 overnight at 37 °C, resulting in a linear double-stranded DNA with sticky ends.

The amplified DNA fragments were cloned into pMT-PURO vector following the procedure reported in the "In-Fusion HD Cloning Kit" protocol. Vector and insert were ligated under the reaction conditions reported in Table 3.7:

**Table 3.7:** Reaction conditions for ligation.

<b>Component</b>	<b>Amount</b>
5X InFusionHD Enzyme Premix	2 μL
Linearized vector	100 ng
Purified PCR fragment	100 ng

The total reaction volume was adjusted to 10 μl using deionized H<sub>2</sub>O. The reaction mixture was incubated for 15 min at 50°C and then placed on ice. The ligation sample was then transformed, as described in paragraph 3.1.1.

The plasmid DNA of selected colonies was purified, as previously reported in paragraph 3.1.2, and sent for sequencing.

#### **3.4.4 Ligation of $\beta$ FXIIa in pMT-PURO vector**

Before proceeding with the ligation, the PCR product was digested with the appropriate restriction enzyme (BglII and MluI), according to the NEB usage information protocol. The restriction enzyme digestion was performed in a reaction volume of 50-100  $\mu$ L containing 0.2-1.5  $\mu$ g of DNA, 10 U/ $\mu$ L of each restriction enzyme, 1X BSA (bovine serum albumin) and 1X Buffer 3 (100 mM NaCl, 50 mM Tris-HCl, 10 mM MgCl<sub>2</sub>, 1 mM dithiothreitol, pH 7.9). The digestion mixture was incubated at 37 °C overnight. The digested DNA sample was then purified using "GeneElutePCR clean-up kit" (SIGMA).

The insert and the linearized plasmid were ligated accordingly, using the "Quick Ligation Kit" protocol (NEB): 50 ng of linearized pMT-PURO vector were combined with 3-fold molar excess of insert and incubated for 5 min at room temperature. The transformation and recovery of the plasmid were performed as previously reported (paragraphs 3.1.1 and 3.1.2).

### **3.5 Expression of constructs**

#### **3.5.1 Expression of MBP-βFXIIa (AAAAS) and βFXIIa cloned in the pMT-PURO vector into insect cells**

The pMT-PURO vector containing the FXII gene was transfected into *Drosophila Schneider* 2 cells, S2 (Invitrogen) according to the following protocol. Once the cells had reached a density of 2 to 4×10<sup>6</sup> cells/ml, a transient transfection with Calcium phosphate was performed, to confirm whether the protein was expressed or not. 19 µg of recombinant plasmid DNA were mixed with a solution A containing 2 M CaCl<sub>2</sub> (SIGMA). This solution was then added drop wise to 300 µL of 2X HEPES-Buffered-Saline (50 mM Na<sub>2</sub>HPO<sub>4</sub>, 280 mM NaCl, pH 7.1, solution B) and then incubated for 30 min at room temperature to allow the formation of a fine precipitate, which is necessary for an efficient transfection. This mixture was added to the cells, swirled and stored in the incubator for 24 h. The calcium phosphate was then removed, by washing the cells twice with complete *Drosophila Schneider* medium (complete media); fresh complete media was added to the transfected cells and then replaced at 28 °C. At day 4, 2 mL of cells were induced with 500 µM CuSO<sub>4</sub>, while the remaining volume was saved for stable transfection. 200 µl of cells were harvested at day 3, 6 and 7, to run a western blot, in order to check if the protein has been expressed.

Once the Western Blot confirmed that the protein was expressed, a stable transfection was initiated, following the protocol suggested by Invitrogen for S2 cells. To select the cells, they were first centrifuged for 2 min at 1000 g and then resuspended in complete media containing 10 µg/mL of

puromycin as selecting agent. The media was replaced every 2 days until cells started growing out. When cells reached a density of  $9 \times 10^6$  cells/mL, they were passaged by diluting 1:2, in order to remove dead cells, and finally expanded for large-scale expression.

### **3.5.2 Expression of $\beta$ FXII and Ecotin in Origami2 cells for *E. coli* expression**

$\beta$ FXII gene in pCOLD-TF vector and Ecotin gene in pET-26b(+) vector were both provided by GenScript. 2  $\mu$ l of each plasmid were transformed in 50  $\mu$ l of Origami 2 cells. A single colony from a fresh plate, streaked with the right antibiotic resistance, was transferred to 5 mL of LB media containing the proper antibiotic and left growing at 37 °C overnight. 10 mL of this starter culture were then used to inoculate 1 L of LB media, containing the proper antibiotic, which was that incubated at 37 °C until an optical density (OD) of 0.5-0.8 was reached. At this point, the culture was induced by adding 0.8 mM of IPTG and then incubated at different temperatures and for different times, according to the characteristics of the used expression vector. Table 3.8 reports the information relative to the used antibiotics, the OD, the temperature set after induction, and the length of the expression for the two constructs.

**Table 3.8:** Expression conditions for  $\beta$ FXII in pCOLD-TF vector and for Ecotin in pET-26b(+) vector.

<b>Construct</b>	<b>Antibiotic</b>	<b>OD for the induction</b>	<b>T after induction and length of expression</b>
$\beta$ FXII	Ampicillin Tetracycline	0.8	10 °C-48 h
Ecotin	Tetracycline Kanamycin	0.6	15 °C-24 h

### **3.6 Purification**

The purity of a protein is a key step for crystallization as any contaminants can affect crystal growth. A combination of purification steps is generally used to assure the quantity, quality, and desired purity of the target protein. Usually, a three-step purification strategy is applied to obtain protein pure enough (>90%) for crystallization. During this PhD project, a purification protocol was developed for each construct.

Table 3.9 and 3.10 summarize all the buffers used for every purification.

### **3.6.1 Purification of MBP- $\beta$ FXIIa from *Drosophila* Expression System (DES)**

The cells were harvested by centrifugation at 5000 g for 20 min. A further centrifugation at 15000 g for 1 h was performed, to separate insoluble matter. The MBP- $\beta$ FXIIa media was filtered and diluted 1:1 with 20 mM Tris HCl pH 8.0 (Buffer A) and loaded onto Captor Q column, previously equilibrated with Buffer A. The elution was carried out at a flow rate of 1.5 mL/min, linearly increasing the concentration of NaCl to 1M. The progress of the chromatography was evaluated monitoring the absorbance at 280 nm. The fractions containing the protein of interest were identified by SDS-PAGE, and pooled. The fractions were then loaded onto HiTrap Chelating HP column (GE Healthcare) charged with 100 mM NiSO<sub>4</sub>, previously equilibrated with 50 mM TrisHCl pH 8.0, 150 mM NaCl, 20 mM imidazole (Buffer A). The column was washed with buffer A, until the absorbance had become constant, to remove non-specifically bound proteins. The elution was performed at a flow rate of 1 mL/min, increasing the concentration of imidazole from 0 to 1 M. The eluted protein was concentrated using Falcon Ultra Centrifugal filter units with a 10 kDa cut-off and then applied to the size exclusion chromatography on a Superdex S200-16/60 column (GE Healthcare), equilibrated with 20 mM Tris HCl pH 8.0, 200 mM NaCl, 40 mM D-(+)-maltose buffer. The pure protein was concentrated in 20 mM Tris HCl pH 8.0, 50 mM NaCl and 5 mM maltose.

### **3.6.2 Purification of $\beta$ FXIIa from DES**

The cells were harvested by centrifugation at 5000 g for 20 min, while insoluble matter was separated by centrifugation at 15000 g for 1 h. The media was clarified and diluted 1:1 with 50 mM Tris HCl pH 8.0, 200 mM NaCl, 20 mM imidazole (Buffer A) and loaded onto a Ni Sepharose Excel (GE Healthcare), pre-equilibrated with Buffer A. The column was washed with buffer A until the absorbance had become constant, to remove non-specifically bound proteins. Bound proteins were then eluted by imidazole gradient from 20 mM to 500 mM at a flow rate of 1.5 mL/min. The fractions containing the protein of interest were identified by SDS-PAGE. The fractions of interest were pooled and diluted 10 times, with 20 mM Tris HCl pH 8.0, in order to reduce the concentration of imidazole, and then loaded onto Resource Q column. The elution was carried out at a flow rate of 2 mL/min, increasing linearly the concentration of NaCl up to 1 M. The fractions containing the protein were identified by SDS-PAGE, pooled and concentrated down to 1 mL for the size exclusion chromatography, using a Superdex S75-16/60 column (GE Healthcare), equilibrated with 20 mM Tris HCl pH 8.0, 150 mM NaCl buffer. The pure protein was concentrated to 20 mg/mL in 20 mM Tris HCl pH 8.0, 50 mM NaCl.

### **3.6.3 Purification of $\beta$ FXIIG570R from DES**

First, the cells were harvested and the media clarified; a 1:1 dilution with buffer A (50 mM Tris HCl pH 8.0, 200 mM NaCl) was made and the media was then loaded onto a Ni Sepharose Excel column. The column was

washed with buffer A until the absorbance was constant, to remove non-specifically bound proteins. The elution was performed at a flow rate of 1 mL/min increasing the concentration of imidazole from 0 to 500 mM. The fractions containing the protein of interest were identified by SDS-PAGE, pooled and subsequently loaded onto HiTrap Chelating HP column charged with 100 mM NiSO<sub>4</sub>. The column was washed with three column volumes of buffer A in order to remove any unspecific binding. Bound protein was eluted by imidazole gradient (20 mM to 500 mM) over 16 column volumes. The fractions containing the protein of interest were first identified by SDS-PAGE gel analysis and then concentrated to 3 mL for the last purification step, size exclusion chromatography on a Superdex S75-16/60 column. The pure protein was concentrated to 3.6 mg/mL in 20 mM Tris HCl pH 8.0, 50 mM NaCl.

#### **3.6.4 Expression and purification of $\beta$ FXII in pCOLD-TF vector**

The cells were harvested by centrifugation at 4600 g for 30 min and the supernatant was discarded. Cell pellet was re-suspended in 25 mL of 50 mM Tris HCl pH 8.0, 200mM NaCl, 1% glycerol and sonicated for 7 min for 30 sec with 10 sec rest periods. The cell lysate was then centrifuged at 15000 g for 1 h to clarify the supernatant, which was then loaded onto HiTrap Chelating HP column (GE Healthcare) charged with 100 mM NiSO<sub>4</sub>, previously equilibrated with 50 mM TrisHCl pH 8.0, 200 mM NaCl, 1% glycerol (Buffer A). The column was washed with buffer A until the absorbance was constant, to remove non-specifically bound proteins. In order to separate unbound trigger factor from the protein of interest, the

elution was performed by increasing buffer B concentration from 0% to 30% over a large number of column volumes. The fractions were identified by SDS-PAGE and then pooled together for the subsequent purification step, performed on HiTrap Q for anion exchange chromatography. The elution was run at a flow rate of 1 mL/min by increasing the concentration of NaCl up to 1M.

### **3.6.5 Expression and purification of Ecotin in pET-26b(+) vector**

The cells were harvested by centrifugation at 4600 g for 30 min and the supernatant was discarded. Cell pellet was re-suspended in 25 mL of 20 mM Tris HCl pH 7.5. the cells were lysed by sonicating for 5 min for 30 sec with 30 sec rest periods on ice. The cell lysate was then centrifuged at 15000 g for 1 h to clarify the supernatant. The supernatant was diluted 1:10 with 50 mM Tris HCl pH 8.0, 200 mM NaCl (buffer A) and loaded onto Ni column for affinity chromatography, following the same procedure adopted for the other proteins purified during this PhD project. Since the protein was already 95% pure after one purification step, the fractions of interest, previously identified by SDS-PAGE, were concentrated to 5.2 mg/mL.

### **3.6.6 Co-purification of $\beta$ FXII and Ecotin from *E. coli* lysates**

$\beta$ FXII and Ecotin were expressed separately according to the protocol previously reported. Once the cells were harvested, the pellets of the two proteins were mixed together by re-suspending them in 50 mL of 20 mM Tris HCl pH 7.5. The cells were lysed by sonicating for 6 min for 30 sec

with 30 sec rest periods on ice. The supernatant was hence clarified by centrifuging the cell lysate at 15000 g for 1 h. The supernatant was then loaded onto a HiTrap Chelating HP column (GE Healthcare) charged with 100 mM NiSO<sub>4</sub>, previously equilibrated with 50 mM TrisHCl pH 8.0, 200 mM NaCl, (Buffer A). The column was washed with buffer A until the absorbance was constant, to remove non-specifically bound proteins. The elution was performed by increasing the buffer B concentration from 0% to 30% over 50 column volumes. The fractions, identified by SDS-PAGE, were pooled together for the size exclusion chromatography, using a Superdex S200-10/300 column (GE Healthcare), equilibrated with 50 mM Tris HCl pH 7.5, 200 mM NaCl buffer. The protein was then treated with thrombin in order to cleave the trigger factor.

**Table 3.9:** Summary of the chromatographic methods, columns and buffers used for the purification of the various expressed and purified proteins with DES system.

Construct	Purification step	Column used	Buffers
MBP-βFXII	Anion exchange chromatography	Capto Q	A) 20 mM Tris HCl pH 8.0 B) 20 mM Tris HCl pH 8.0, 1 M NaCl
	Metal ion affinity chromatography	HiTrap Chelating HP column charged with 100 mM NiSO <sub>4</sub>	A) 50 mM TrisHCl pH 8.0, 200 mM NaCl, 20 mM imidazole B) 50 mM TrisHCl pH 8.0, 200 mM NaCl, 1 M imidazole
	Size exclusion chromatography	Superdex 200 16/60	20 mM TrisHCl pH 8.0, 200 mM NaCl, 40mM D-(+)-maltose
βFXII	Metal ion affinity chromatography	Ni Sepharose Excel	A) 50 mM TrisHCl pH 8.0, 200 mM NaCl, 20 mM imidazole B) 50 mM TrisHCl pH 8.0, 200 mM NaCl, 500 mM imidazole
	Anion exchange chromatography	Resource Q	A) 20 mM Tris HCl pH 8.0 B) 20 mM Tris HCl pH 8.0, 1 M NaCl
	Size exclusion chromatography	Superdex 75 16/60	20 mM TrisHCl pH 8.0, 200 mM NaCl
βFXII G570R	Metal ion affinity chromatography	Ni Sepharose Excel	A) 50 mM TrisHCl pH 8.0, 200mM NaCl, 20mM imidazole B) 50 mM TrisHCl pH 8.0, 200 mM NaCl, 500 mM imidazole
	Metal ion affinity chromatography	HiTrap Chelating HP column charged with 100 mM NiSO <sub>4</sub>	A) 50 mM TrisHCl pH 8.0, 200 mM NaCl, 20 mM imidazole B) 50 mM TrisHCl pH 8.0, 15 200 mM NaCl, 500 mM imidazole
	Size exclusion chromatography	Superdex 75 16/60	20 mM TrisHCl pH 8.0, 200 mM NaCl

**Table 3.10:** Summary of the chromatographic methods, columns and buffers used for the purification of the various expressed and purified proteins with *E. coli* system.

Construct	Purification step	Column used	Buffers
βFXII	Metal ion affinity chromatography	HiTrap Chelating HP column charged with 100 mM NiSO <sub>4</sub>	A) 50 mM TrisHCl pH 8.0, 200 mM NaCl, 20 mM imidazole 1% glycerol B) 50 mM TrisHCl pH 8.0, 15 200 mM NaCl, 500 mM imidazole  1% glycerol
	Anion exchange chromatography	HiTrap Q	A) 20 mM Tris HCl pH 8.0 B) 20 mM Tris HCl pH 8.0, 1 M NaCl
Ecotin	Metal ion affinity chromatography	HiTrap Chelating HP column charged with 100 mM NiSO <sub>4</sub>	A) 50 mM TrisHCl pH 8.0, 200 mM NaCl, 20 mM imidazole B) 50 mM TrisHCl pH 8.0, 200 mM NaCl, 500 mM imidazole
	Size exclusion chromatography	Superdex 75 16/60	20 mM TrisHCl pH 8.0, 200 mM NaCl
βFXII-Ecotin copurification	Metal ion affinity chromatography	HiTrap Chelating HP column charged with 100 mM NiSO <sub>4</sub>	A) 50 mM TrisHCl pH 8.0, 200 mM NaCl, 20 mM imidazole B) 50 mM TrisHCl pH 8.0, 200 mM NaCl, 500 mM imidazole
	Size exclusion chromatography	Superdex 200 10/300	50 mM TrisHCl pH 7.5, 200 mM NaCl

### **3.6.7 Cleavage of Trigger Factor from $\beta$ FXII in pCold-TF using HRV-3C protease and FXa**

pCold-TF vector is characterized by the presence of the cleavage site for HRV-3C protease, FXa and thrombin to easily remove the tag. In order to find the best enzyme and the optimal conditions for the cleavage of trigger factor, a small scale reaction was performed for each enzyme, following the protocol suggested by the supplier.

HRV-3C protease (Pierce) cleaves protein substrates with the recognition sequence Leu-Glu-Val-Leu-Phe-Gln-Gly-Pro between the Gln and Gly residues. A small scale reaction (25  $\mu$ l) was first carried out. Eight different enzyme (U) : substrate ( $\mu$ g) ratios were used: 1:50, 1:100; 1:200, 1:400, 2:50, 2:100; 2:200, 2:400. All the reactions were performed overnight at two different temperatures, 37 °C and room temperature.

FXa (NEB), cleaves after the Arg residue in its preferred cleavage site Ile-(Glu or Asp)-Gly-Arg. Four different FXa :substrate ratios were used: 1:50, 1:100, 1:200 and 1:400 wt:wt at 37 °C overnight. An SDS-PAGE gel was then run on all the samples, in order to identify the optimal reaction conditions.

### **3.6.8 Cleavage of Trigger Factor from $\beta$ FXII in pCold-TF using thrombin**

Thrombin recognizes the consensus sequence Leu-Val-Pro-Arg-Gly-Ser, cleaving the peptide bond between Arg and Gly. In order to find the best conditions for the cleavage of TF operated by thrombin (HaemTech), a

small scale digestion was performed. A 1:100 and 1:200 wt:wt ratio (thrombin : target protein) were tested at room temperature, taking 10  $\mu$ l samples after 10 min, 1 h, 2 h, 4 h and overnight. When a satisfactory condition was found (1:200 wt:wt, 2 h and overnight), the reaction was scaled up proportionately. After the affinity purification step, the protein was exchanged into a low-medium salt buffer (20 mM Tris-HCl, 100 mM NaCl pH 8) using a PD-10 column, to remove the imidazole, since high imidazole concentrations can inhibit the cleavage. The protein was then incubated with thrombin over night. The day after, the protein solution was first loaded onto the nickel column to remove the TF, while the eluate was kept and subsequently loaded onto benzamidine column, for thrombin removal.

### **3.7 Preparation of protein-inhibitor complexes**

#### **3.7.1 Complex with the CMK peptides PPACK and PCK.**

The formation of the complex with the CMK peptides, PPACK and PCK was carried out following the procedure reported in literature.[251] On a frozen protein sample at 1 mg/mL, the inhibitor was added in a 1:10 protein : inhibitor excess and then it was left for 2 h at 4 °C, allowing the inhibitor to diffuse into the protein solution. The complex was washed twice with the final buffer, in order to remove the excess of peptide and it was then concentrated by ultra-centrifugation at 4 °C.

### **3.7.2 Complex with serine protease inhibitors: Ecotin**

Stoichiometric quantities (1:1 and 1:2, protein:Ecotin) of MBP- $\beta$ FXIIa and  $\beta$ FXIIG570R were incubated with Ecotin and left overnight at 4 °C. The (MBP- $\beta$ FXIIa)-Ecotin and  $\beta$ FXIIG570R-Ecotin complexes were then isolated by gel filtration using Superdex 200 10/300 and Superdex 75 10/300 respectively, at a flow rate of 0.8 mL/min (running buffer, 20 mM TrisHCl, 150 mM NaCl, pH 8.0). The fractions corresponding to the complex were identified by SDS-PAGE and then concentrated, using a centricon concentrator with a 10 kDa cut-off.

### **3.7.3 Complex with serine protease inhibitors: CTI**

Stoichiometric quantities (1:4, protein:CTI) of MBP- $\beta$ FXIIa were incubated with CTI and left overnight at 4 °C. The (MBP- $\beta$ FXIIa)-CTI complex was then isolated by gel filtration using Superdex 200 10/300 at a flow rate of 0.8 mL/min (running buffer, 20 mM TrisHCl, 150 mM NaCl, pH 8.0). The fractions corresponding to the complex were identified by SDS-PAGE and then concentrated, using a centricon concentrator with a 10 kDa cut-off.

## **3.8 Crystallization**

### **3.8.1 Search for the crystallization conditions**

The search for crystallization conditions for all the proteins expressed and purified during this PhD project was achieved using different commercially available crystallization suites (QIAGEN, Hampton research). Using these kits, it is possible to test the solubility of a protein with respect to different

precipitating agents and buffers. The crystallization trays were set up on a microscopic scale, using the sitting drop vapour diffusion technique.

Several crystallization suites, (i.e. PACT, PEGs, AmSO<sub>4</sub>, Classic, JCSG+ and Morpheus), were set, using a wide range of protein concentrations between 2.7 mg/mL and 10 mg/mL. 100 nL of protein were mixed with 100 nL of reservoir solution using a Mosquito robot (TTP Labtech).

The initial hits were then used as starting point for an extensive optimization. The optimization was accomplished by changing the salt concentration and the percentage of used precipitant. In addition, the effect of pH on the nucleation was tested and different additives, mainly sugars and alcohols, were used to control nucleation.

### **3.8.2 Crystal optimization for (MBP- $\beta$ FXIIa)-PPACK using the incomplete factorial method and the micro-seeding technique**

Conditions containing hits of (MBP- $\beta$ FXIIa)-PPACK were optimized using the approaches described in Chapter 2. Initially, the “star search” system was the optimization method of choice. The parameters were first changed one by one (salt concentration, percentage of PEG, pH), monitoring any possible improvement in the crystallization. Then, two parameters were changed simultaneously.

A further optimization was performed using the incomplete factorial and micro-seeding methods. Following the procedure reported in ref. [213], the conditions leading to crystals were identified. Therefore, a list of the used precipitants, along with salts and buffers, was prepared (Table 3.11).

**Table 3.11:** Experimental reagents influencing (MBP- $\beta$ FXIIa)-PPACK crystallization.

<b>PRECIPITANT</b>	
15% PEG 4000	1
10% PEG 4000	2
25% PEG 500MME	3
20% PEG 3350	4
15% PEG 3350	5

<b>BUFFER</b>	
0.1 M Na <sub>3</sub> Citrate pH 5.5	1
0.1 M NaAc pH 4.6	2
0.1 M Hepes pH 7.5	3

<b>SALT</b>	
0.1 M MgCl <sub>2</sub>	1
0.15 M MgCl <sub>2</sub>	2
0.2 M NaAc	3

Experimental conditions were then randomized and balanced simultaneously (Table 3.12). As reported in ref. [213] the table was constructed top-to-bottom, column-by-column, left-to-right. The components reported in Table 3.11 were combined together, ensuring the coverage of all the possible interactions. In this way, the incomplete factorial design was obtained. An experiment number was assigned to uniquely specify each experiment.

**Table 3.12:** Initial search conditions for the incomplete factorial method.  
All the experiment were performed at 20 °C.

Experiment number	Precipitant	Buffer	Salt
1	1	1	1
2	2	1	1
3	3	1	1
4	4	1	1
5	5	1	1
6	1	2	1
7	2	2	1
8	3	2	1
9	4	2	1
10	5	2	1
11	1	3	1
12	2	3	1
13	3	3	1
14	4	3	1
15	5	3	1
16	1	1	2
17	2	1	2
18	3	1	2
19	4	1	2
20	5	1	2
21	1	2	2
22	2	2	2
23	3	2	2
24	4	2	2
25	5	2	2
26	1	3	2
27	2	3	2
28	3	3	2
29	4	3	2
30	5	3	2
31	1	1	3
32	2	1	3
33	3	1	3
34	4	1	3
35	5	1	3
36	1	2	3
37	2	2	3
38	3	2	3
39	4	2	3
40	5	2	3
41	1	3	3
42	2	3	3
43	3	3	3
44	4	3	3
45	5	3	3

In Table 3.12 the performed search is reported: it includes 45 trials. For each trial, we used 100 nL of (MBP- $\beta$ FXIIa)-PPACK mixed with 100 nL of reservoir in sitting-well plates, using Mosquito robot (TTP Labtech).

Hits were obtained in 0.1 M Na<sub>3</sub>Citrate pH 5.6 or 0.1 M NaAc pH 4.6 with 15% PEG 3350 and 0.1 M MgCl<sub>2</sub> and 0.2 M NaAc, respectively. A further optimization was performed, but no crystals were obtained.

A second incomplete factorial grid was built in the same way, but additional rows were added to expand the sampled range for the concentrations of precipitants, salts for pH. New salt and precipitants, previously not considered but similar to the ones already employed, were also tested. For example, since crystallization occurs in MgCl<sub>2</sub>, experiments adopting MnCl<sub>2</sub> and CoCl<sub>2</sub> were also carried out. Regarding the precipitating agents, PEGs with different molecular weight have been tested (e.g. PEG4000 and

PEG8000). In Table 3.13 the list of these new crystallization reagents is reported:

**Table 3.13:** Experimental reagents influencing the (MBP- $\beta$ FXIIa)-PPACK crystallization. The list has been expanded adding new salts and precipitants similar to the ones already tested.

<b>PRECIPITANT</b>		<b>SALT</b>	
15% PEG 3350	1	0.1 M MgCl <sub>2</sub>	1
15% PEG 4000	2	0.2 M NaAc	2
15% PEG 8000	3	0.1 M MnCl <sub>2</sub>	3
10% PEG 3350	4	0.1 M CoCl <sub>2</sub>	4
10% PEG 4000	5	0.2 M MnCl <sub>2</sub>	5
10% PEG 8000	6	0.2 M CoCl <sub>2</sub>	6
		0.2 M MgCl <sub>2</sub>	7
		0.1 M NaAc	8
<b>BUFFER</b>			
0.1 M Na <sub>3</sub> Citrate pH 5.5	1		
0.1 M NaAc pH 4.2	2		

Table 3.14 with new experimental conditions was constructed in the same way as previously reported.

**Table 3.14:** Initial search conditions for the new incomplete factorial grid.  
All the experiments were performed at 20 °C.

Experiment number	Precipitant	Buffer	Salt
1	1	1	1
2	2	1	1
3	3	1	1
4	4	1	1
5	5	1	1
6	6	1	1
7	1	2	1
8	2	2	1
9	3	2	1
10	4	2	1
11	5	2	1
12	6	2	1
13	1	1	2
14	2	1	2
15	3	1	2
16	4	1	2
17	5	1	2
18	6	1	2
19	1	2	2
20	2	2	2
21	3	2	2
22	4	2	2
23	5	2	2
24	6	2	2
25	1	1	3
26	2	1	3
27	3	1	3
28	4	1	3
29	5	1	3
30	6	1	3
31	1	2	3
32	2	2	3
33	3	2	3
34	4	2	3
35	5	2	3
36	6	2	3
37	1	1	4
38	2	1	4
39	3	1	4
49	1	1	5
50	2	1	5
51	3	1	5
52	4	1	5
53	5	1	5
54	6	1	5
55	1	2	5
56	2	2	5
57	3	2	5
58	4	2	5
59	5	2	5
60	6	2	5
61	1	1	6
62	2	1	6
63	3	1	6
64	4	1	6
65	5	1	6
66	6	1	6
67	1	2	6
68	2	2	6
69	3	2	6
70	4	2	6
71	5	2	6
72	6	2	6
73	1	1	7
74	2	1	7
75	3	1	7
76	4	1	7
77	5	1	7
78	6	1	7
79	1	2	7
80	2	2	7
81	3	2	7
82	4	2	7
83	5	2	7
84	6	2	7
85	1	1	8
86	2	1	8
87	3	1	8

40	4	1	4
41	5	1	4
42	6	1	4
43	1	2	4
44	2	2	4
45	3	2	4
46	4	2	4
47	5	2	4
48	6	2	4
88	4	1	8
89	5	1	8
90	6	1	8
91	1	2	8
92	2	2	8
93	3	2	8
94	4	2	8
95	5	2	8
96	6	2	8

The search resumed in Table 3.14 includes 96 trials, each using 100 nL of (MBP- $\beta$ FXIIa)-PPACK mixed with 100 nL of reservoir in sitting-well plates, using Mosquito robot (TTP Labtech).

Hits were obtained in 0.1 M Na<sub>3</sub>citrate pH 5.6 or 0.1 M NaAc pH 4.2 with different PEGs (PEG3350, PEG4000, PEG8000) and salts (NaAc, CoCl<sub>2</sub>, MgCl<sub>2</sub>).

Experiments with the same conditions were repeated at two different temperatures (15 °C and 10 °C), in order to monitor a possible beneficial effect of a different temperature on the crystallization.

Since all the conditions gave tiny crystals, micro-seeding was adopted, aiming to influence the nucleation events, following the procedure reported in ref. [252]. Crystals obtained from initial screens or optimized conditions were placed in 50  $\mu$ L of their respective reservoir solution and homogenized by vortexing for 3 min. Sitting-drop vapour-diffusion experiments were performed. 100 nl screening solution and 100 nl microseeds were added to 900 nl protein solution. The screening solutions used for the experiments were Index from Hampton Research, PEGs Suite from Qiagen and the 96 well incomplete factorial grid. The plates were sealed with clear plastic tape and incubated at 20 °C and 4 °C.

### 3.8.3 Crystal optimization for (MBP- $\beta$ FXIIa)-Ecotin complex using the incomplete factorial method.

The crystallization screening for the (MBP- $\beta$ FXIIa)-Ecotin complex gave initial hits only in the ProComplex suite. These conditions were then used to build an incomplete factorial screen. Table 3.15 reports a list of the precipitating agent used, as well as salt and buffer.

**Table 3.15:** Experimental reagents influencing the crystallization of the (MBP- $\beta$ FXIIa)-Ecotin complex.

PRECIPITANT		SALT	
15% PEG4000	1	0.1 M MgCl <sub>2</sub>	1
20% PEG4000	2	0.15 M AmSO <sub>4</sub>	2
20% PEG6000	3	0.2 M NaCl	3
20% PEG8000	4	0.2 M LiCl <sub>2</sub>	4
15% PEG20000	5		

BUFFER	
0.1 M Hepes pH 7	1
0.1 M Tris pH 8	2
0.1 M Tris pH 8.5	3

Experimental conditions were then randomized and balanced simultaneously (Table 3.16). The components reported in Table 3.15 were combined together ensuring the coverage of all possible interactions. In this way, the incomplete factorial design was obtained. An experiment number was assigned to uniquely specify each experiment.

**Table 3.16:** Initial search conditions for (MBP- $\beta$ FXIIa)-Ecotin. All the experiment were performed at 20 °C.

Experiment number	Precipitant	Buffer	Salt
1	1	1	1
2	2	1	1
3	3	1	1
4	4	1	1
5	5	1	1
6	1	1	2
7	2	1	2
8	3	1	2
9	4	1	2
10	5	1	2
11	1	1	3
12	2	1	3
13	3	1	3
14	4	1	3
15	5	1	3
16	1	1	4
17	2	1	4
18	3	2	4
19	4	2	4
20	5	2	4
21	1	2	1
22	2	2	1
23	3	2	1
24	4	2	1
25	5	2	1
26	1	2	2
27	2	2	2
28	3	2	2
29	4	2	2
30	5	2	2
31	1	2	3
32	2	2	3
33	3	2	3
34	4	2	3
35	5	2	3
36	1	2	4
37	2	2	4
38	3	2	4
39	4	2	4
40	5	2	4
41	1	3	1
42	2	3	1
43	3	3	1
44	4	3	1
45	5	3	1
46	1	3	2
47	2	3	2
48	3	3	2
49	4	3	2
50	5	3	2
51	1	3	3
52	2	3	3
53	3	3	3
54	4	3	3
55	5	3	3
56	1	3	4
57	2	3	4
58	3	3	4
59	4	3	4
60	5	3	4

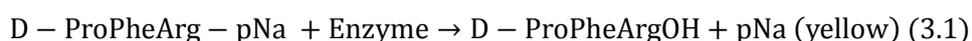
In Table 3.16 the performed search is reported: it includes 60 trials, each using 100nL of (MBP- $\beta$ FXIIa)-Ecotin mixed with 100nL of reservoir in sitting-well plates. However, this optimization did not give any crystals.

### 3.9 Kinetic assay

The enzymatic activity of the purified protein was assessed by measuring the change in absorbance at 405 nm associated with the cleavage of the

chromogenic substrate, S2302 (Chromogenix, Epsom, UK), using an ENVISION multimode plate reader.

S2302 (H-D-Pro-Phe-Arg-pNA . 2HCl) is characterized by the presence of the p-Nitroaniline chromophore (pNa). In the presence of the enzyme, the pNa is cleaved according to the following reaction:



The activity is measured from the difference in absorbance (optical density) between the formed pNA and the original substrate. The rate of pNA formation is proportional to the activity of the enzyme under examination.

The protein, at a fixed concentration of 150 nM, was tested against increasing concentrations of the chromogenic substrate S2302 (0.4 mM, 1 mM, 2 mM, 3 mM, 4 mM, 6 mM) using 96 wells plate. In each well, 10  $\mu\text{L}$  of protein and 10  $\mu\text{L}$  of substrate were diluted with 80  $\mu\text{L}$  of PBS buffer. An average of four experiments was used to determine the enzyme kinetics. For every experiment, each substrate concentration has been used in quadruplicates. The kinetic parameters were calculated from the experimental data using the Graphpad Prism software with the non-linear Michaelis-Menten algorithm.

### **3.10 Data collection and structure determination for (MBP- $\beta$ FXIIa)-PPACK**

Crystals of MBP- $\beta$ FXIIa in complex with PPACK suitable for data collection were grown in 0.1 M sodium citrate pH 5.6 with 10% PEG 4000

and 0.15 M MgCl<sub>2</sub> at 20 °C. The diffraction experiments were performed at the beamline IO4 of the Diamond Light Source in Oxford. Crystals were transferred in a cryo-solution (25% glycerol in mother liquor) for few seconds and then flash-cooled. The crystals diffracted at 4 Å resolution. Data sets were processed with XDS[253] and were scaled and merged in AIMLESS.[254] The initial model was determined by molecular replacement using PHASER from the CCP4 suite.[255] The crystallographic structure of the protease domain of FXII solved in the group of Prof. J. Emsley was used as initial model for the βFXIIa (PDB code: 4DXE), while the structure of MBP (PDB code: 1HSJ ) was used as search model for the MBP protein. The model was build with COOT[256] and refined with phenix.refine.[257]

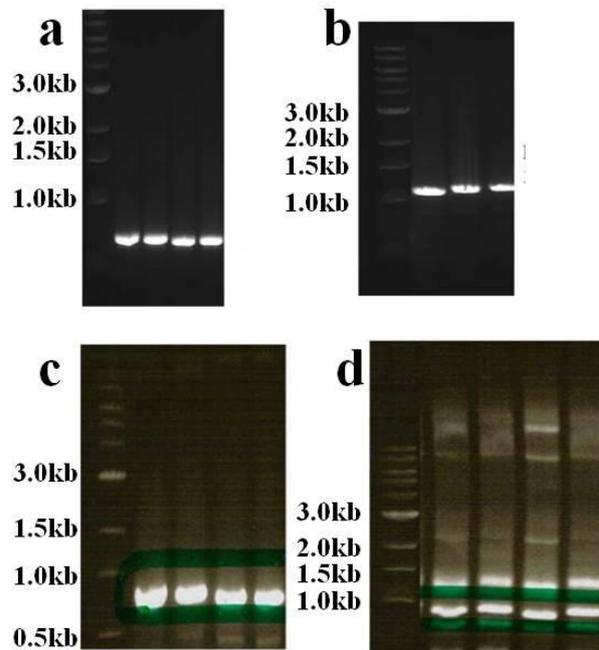
## CHAPTER 4

### EXPRESSION AND CHARACTERIZATION OF

### MBP- $\beta$ FXIIa

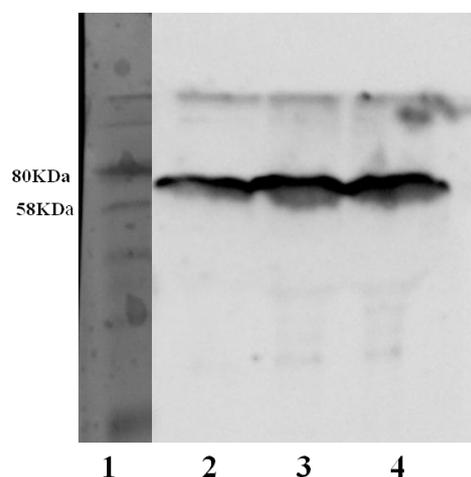
#### 4.1 Cloning of MBP- $\beta$ FXIIa in pMT-PURO vector and expression in DES

PCR fragments of the right size for MBP (1.1 kb) and  $\beta$ FXIIa (0.78 kb) suitable for cloning were obtained, as confirmed by the 1% Agarose gel (Figure 4.1). The PCR fragments were then ligated into pMT-PURO on BglII/MluI restriction site.



**Figure 4.1: 1% Agarose gel showing PCR products from PCR amplification.** a) PCR product for  $\beta$ FXIIa with linker AAA; b) PCR product for MBP with linker AAA; c) PCR product for MBP with linker AAAAS and d) PCR product for  $\beta$ FXIIa with linker AAAAS.

The insertion of MBP and  $\beta$ FXIIa into the proper vector was verified by double digestion of the DNA sample obtained from the MiniPrep of the resulting colonies, using the BglII/MluI restriction sites. The plasmid was sequenced to ensure the absence of undesirable mutations. The cloning worked only for the construct with longer linker (AAAAS), as the sequencing step for the AAA linker resulted in the formation of undesirable mutations, even if different protocols for ligation and transformation have been tested.



**Figure 4.2: Western blot of MBP- $\beta$ FXIIa with linker AAAAS using Anti 6X-His tag primary antibody (Catalog #MAB050). Lane 1: M; lanes 2, 3 and 4: media collected 3, 6 and 7 days after induction respectively.**

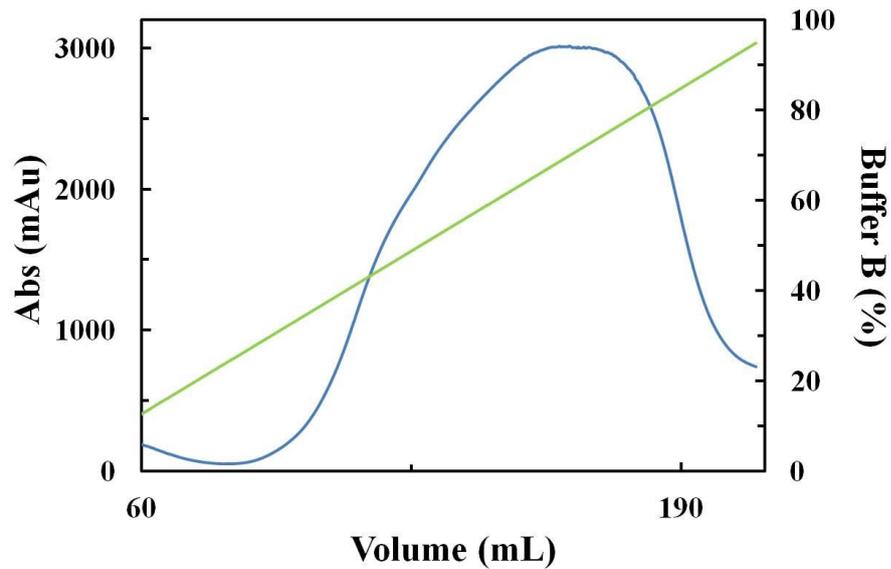
MBP- $\beta$ FXIIa with AAAAS linker was expressed in the *Drosophila Schneider* (S2) system as a fusion engineered protein with a 6X-His tag in the PMT-PURO vector. A Western Blot of the supernatant using an anti-His antibody (Figure 4.2) confirmed that the protein was expressed 3 days after induction with  $\text{CuSO}_4$  and secreted into the media. In fact, we observe a band at  $\sim 70$  kDa, corresponding to the molecular weight of the MBP-

$\beta$ FXIIa complex (30 kDa for  $\beta$ FXIIa and 40 kDa for MBP). We also observe that expression of the complex continued until day 6-7, when we decided to harvest the cells.

#### **4.2 Purification of MBP- $\beta$ FXIIa**

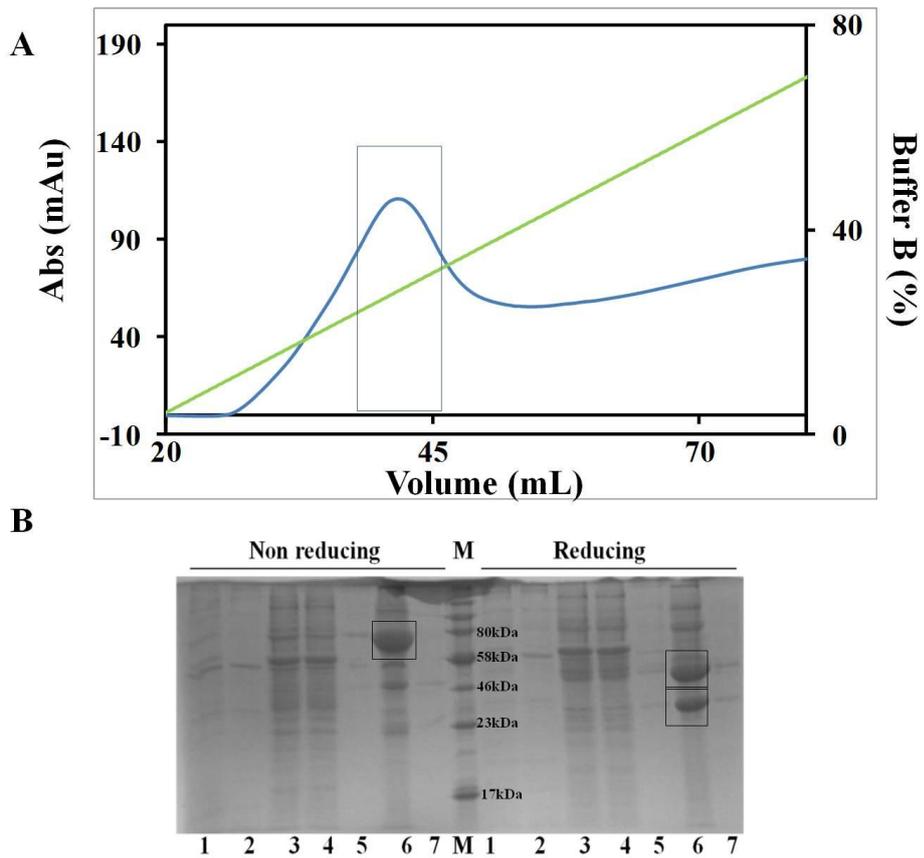
MBP- $\beta$ FXIIa was purified with a three step purification scheme, consisting of consecutive anion exchange chromatography, nickel affinity, and size exclusion chromatography experiments.

A Capto Q column was used for the anion exchange chromatography purification step with a NaCl gradient elution (Figure 4.3). The fractions corresponding to the peak in the chromatogram were analyzed by 15% SDS-PAGE in non reducing conditions (not shown) to confirm the size of the protein and to verify its degree of purity.

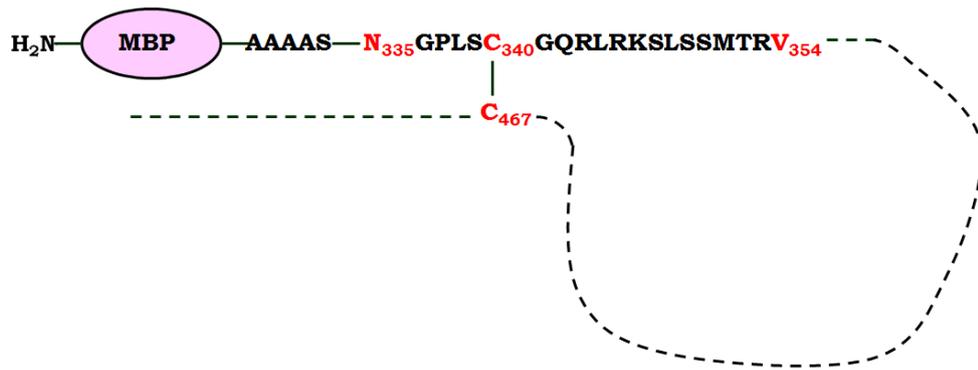


**Figure 4.3: Purification of MBP- $\beta$ FXIIa by anion exchange chromatography.** UV trace from anion exchange chromatography using Capto Q column.

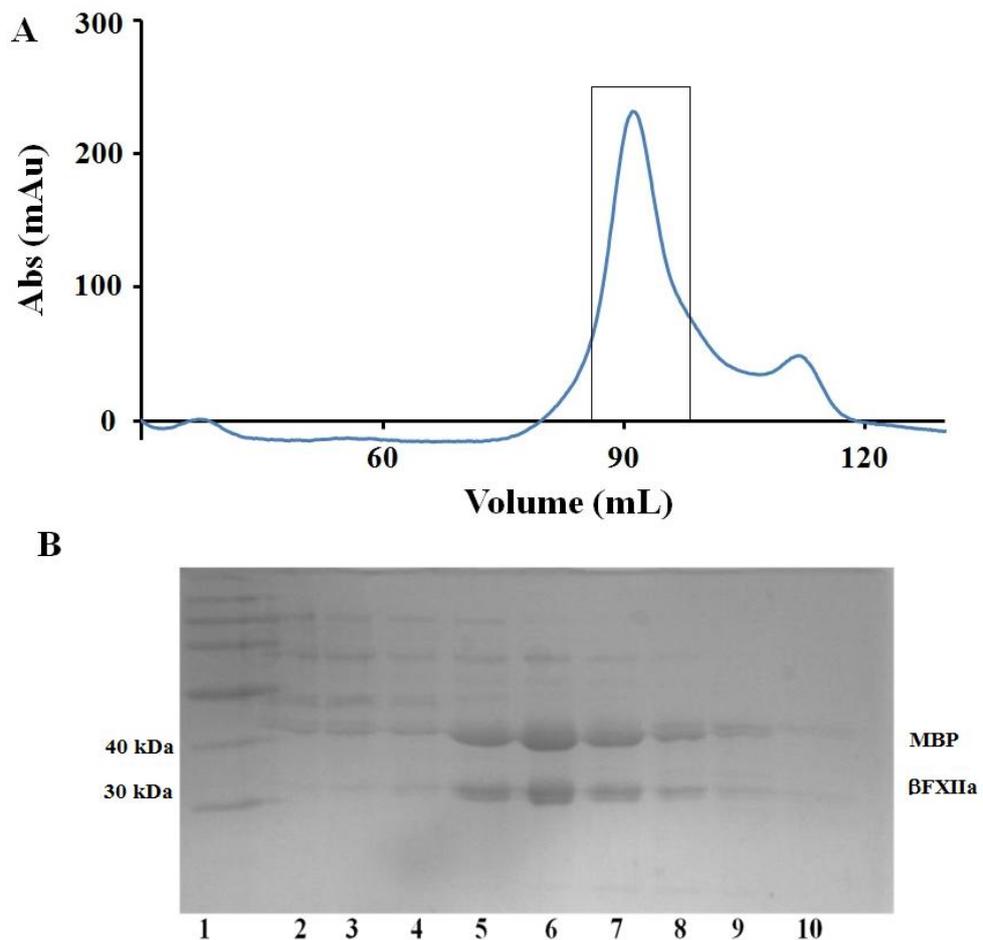
The fractions were collected together, loaded onto a Ni(II) column for affinity chromatography step through a gradient elution of imidazole (Figure 4.4 A). The fractions corresponding to the peak were then analyzed by SDS-PAGE gel (Figure 4.4 B). The SDS-PAGE gel in reducing conditions shows two bands at 30 kDa and 40 kDa corresponding to  $\beta$ FXIIa and MBP, respectively, indicating that the protein has been activated during the purification procedure. In fact, the protein in its active form is cleaved in correspondence of Val354 (cf. Fig. 4.5). In reducing conditions, the Cys340-Cys467 disulphide bridge is broken, thus separating the carrier and the target.



**Figure 4.4: Purification of MBP- $\beta$ FXIIa by affinity chromatography. A)** UV trace from affinity chromatography using nickel column. The rectangular box highlights the peak fractions corresponding to the ones shown in the gel. **B)** Non reducing and reducing 15% SDS-PAGE analysis. Lane 1: media loaded into Capto Q; lane 2: flow through from Capto Q; lane 3: sample loaded into Ni(II) column; lane 4: flow through from Ni(II); lane 5,6,7: elutions; lane M: marker. The boxes highlight the protein of interest in non reducing and reducing conditions.



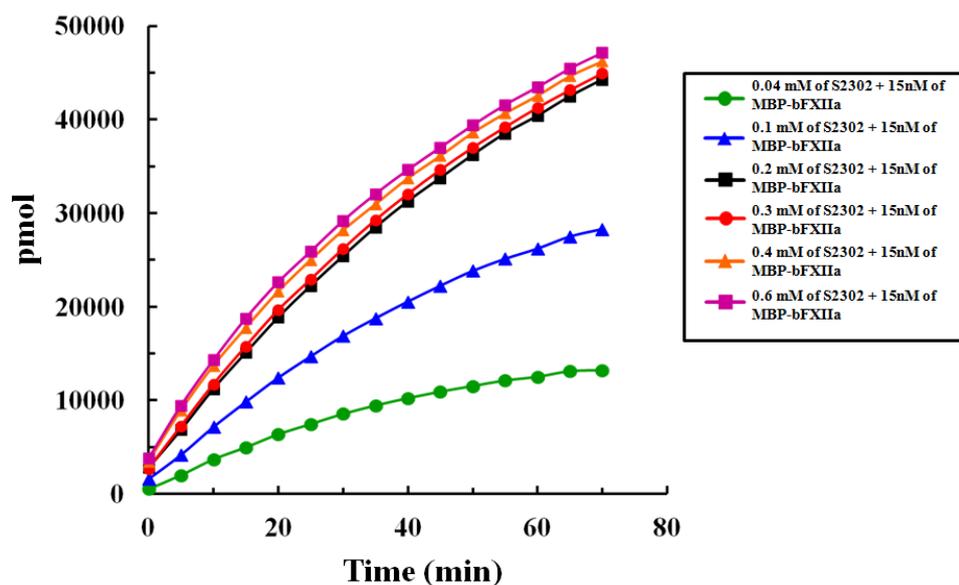
**Figure 4.5: Schematic representation of the Cys340-Cys467 disulfide bridge.** The two cystein residues involved in the disulfide bridge are shown in red along with the Val354, involved in the activation of FXII, and Asn335, which is the N-terminus of  $\beta$ FXIIa. See text for description of the mechanism.



**Figure 4.6: Purification of MBP-βFXIIa by size exclusion chromatography.** A) UV trace from size exclusion chromatography using Superdex 200 16/60 column. The rectangular box highlights the peak fractions corresponding to the ones shown in the gel. B) Reducing SDS-PAGE analysis. Lane 1: marker; lane 2-10: elutions.

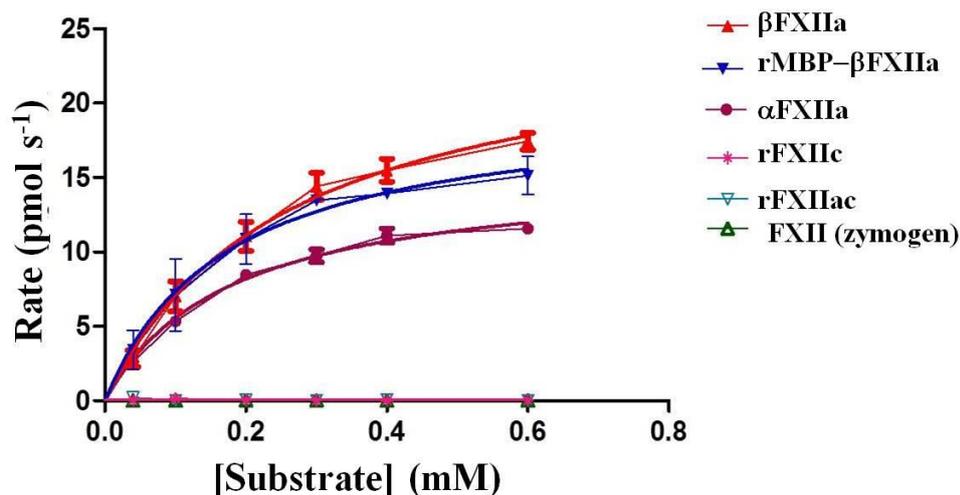
### 4.3 Kinetic characterization of MBP-βFXII

To characterize the catalytic activity of MBP-βFXIIa, the hydrolysis of chromogenic substrate S2302 was monitored at 405 nm. Figure 4.7 shows the pmols of pNitroaniline generated versus time, for different concentrations of S2302.



**Figure 4.7: S2302 time depend assay for MBP-βFXIIa.** pmols of pNitroaniline generated as a function of time, after the addition of S2302 to a 15nM MBP-βFXIIa solution. Different concentrations of S2302 have been tested: 0.04 mM (green), 0.1 mM (blue), 0.2 mM (black), 0.3 mM (red), 0.4mM (orange), 0.6 mM (purple) For every experiment, each substrate concentration has been used in quadruplicates.

The time-dependent generation of pNitroaniline, due to the cleavage of chromogenic substrate S2302 operated by MBP-βFXIIa, reaches the steady state within 70 min for all the considered concentrations of S2302 (cf. Figure 4.7). The kinetic parameters for MBP-βFXIIa (cf. Table 4.1) were calculated plotting the initial reaction rates as function of the concentration of S2302. We observe that MBP-βFXIIa obeys the Michaelis-Menten kinetics, similarly to commercial βFXIIa and αFXIIa (cf. Figure 4.8).



**Figure 4.8: Activity of MBP-βFXIIa against various concentration of S2302.** Initial rate ( $\text{pmol s}^{-1}$ ) of hydrolysis of S2302 as function of substrate concentration (mM) for recombinant MBP-βFXIIa (blue triangles), commercial βFXIIa (red triangles), commercial αFXIIa (purple circles), recombinant FXIIc (pink stars), recombinant FXIIac (cyan triangles), commercial FXII zymogen (green triangles). An average of four experiments were used to determine the enzyme kinetics. For every experiment, each substrate concentration has been used in quadruplicates.

We calculated kinetic parameters for MBP-βFXIIa, commercial βFXIIa and αFXIIa. In particular, we computed (i) the maximal velocity of reaction,  $V_{\max}$ , corresponding to saturation of the enzyme; (ii) the turnover number  $k_{\text{cat}}$  which is defined as the maximum number of chemical conversions of substrate molecules per unit of time carried out by a single catalytic site for a given enzyme concentration; (iii)  $K_M$  is the Michaelis constant defined as the substrate concentration for which the reaction rate is equal to half of  $V_{\max}$ . An estimate of the catalytic efficiency of the enzyme is hence provided by the ratio  $k_{\text{cat}}/K_M$ , also called specificity constant, which provides us with an estimate of the efficiency of the enzyme for a particular substrate. The  $V_{\max}$  and  $k_{\text{cat}}$  for recombinant MBP-βFXIIa are intermediate

between those of commercial  $\beta$ FXIIa and  $\alpha$ FXIIa. The Michaelis constant of MBP- $\beta$ FXIIa is very close to that of commercial  $\alpha$ FXIIa (within the experimental error) and slightly lower than the value calculated for commercial  $\beta$ FXIIa. These parameters show that our recombinant protein is at least as active as the commercial ones. In particular, the specificity constant is comparable to that of commercial  $\beta$ FXIIa and slightly higher than  $\alpha$ FXIIa.

**Table 4.1:** Enzyme kinetic parameters for MBP- $\beta$ FXIIa.

Sample	$V_{\max}$ (pmol/s)	$k_{\text{cat}}$ (s <sup>-1</sup> )	$K_M$ (mM)	$k_{\text{cat}}/K_M$ (Lmol <sup>-1</sup> s <sup>-1</sup> )
MBP- $\beta$ FXIIa	28.01±0.6	20.01±0.54	0.17±0.05	117705
$\beta$ FXIIa	35.65±3.34	25.47±2.37	0.25±0.05	101880
$\alpha$ FXIIa	21.72±1.04	15.52±0.74	0.17±0.02	91294

#### 4.4 Crystallization of MBP- $\beta$ FXIIa

The search for crystallization conditions for MBP- $\beta$ FXIIa with the AAAAS linker was achieved using different crystallization screens. Among all the possible screened conditions, only the PEGs suite (Molecular Dimensions) gave tiny crystals at 5.6 mg/mL of protein. Table 4.2 reports the initial conditions for MBP- $\beta$ FXIIa (AAAAS) crystals grown from the PEGs suite:

**Table 4.2:** Crystallization conditions for MBP- $\beta$ FXIIa from the PEGs suite.

<b>Protein</b>	MBP- $\beta$ FXIIa (AAAAS)
<b>Method</b>	Sitting-drop
<b>Protein concentration</b>	5.6 mg/mL
<b>Reservoir</b>	0.2 M LiCl      20% PEG3350 0.2 M NaAc      20% PEG3350 0.2 M Na <sub>3</sub> Citrate      20% PEG3350
<b>Crystallization time</b>	~10 days
<b>Crystallization temperature</b>	20 °C

The conditions listed in Table 4.2 were used as a starting point for a wide optimization aiming to identify better conditions. New conditions were generated using two different approaches, changing one parameter at time or both parameters at the same time. Therefore, in some conditions the salt concentration was changed between 0.1 M and 0.3 M, while in other trays the percentage of PEG3350 was varied (10%, 15% and 25%). Other optimizations were performed varying the salt concentration and the percentage of the precipitating agent at the same time. From this initial optimization, two new hits were found (cf. Table 4.3). Crystals grown in these conditions however were still tiny and not suitable for X-Ray diffraction experiments.

**Table 4.3:** Crystallization conditions for the MBP- $\beta$ FXIIa from optimization.

<b>Protein</b>	MBP- $\beta$ FXIIa (AAAAS)
<b>Method</b>	Sitting-drop
<b>Protein concentration</b>	5.6mg/mL
<b>Reservoir</b>	0.1 M NaAc 15%PEG3350 0.2 M NaAc 15%PEG3350
<b>Crystallization time</b>	2-3 days
<b>Crystallization temperature</b>	20 °C

Since all the hits for MBP- $\beta$ FXIIa were grown in the absence of buffer, the conditions listed in Table 4.2 and 4.3 were also repeated in the presence of buffers at three different pH (0.1 M HEPES pH 7.5; 0.1 M NaAc pH 4.5; 0.1 M MES pH 6.5) to investigate a possible improvement in the crystal quality. The same tiny crystals were obtained at pH 4.5 and 6.5. A further optimization was done adding different sugars (sucrose, glucose, maltose) and alcohols (i.e. isopropanol) as additives. The aim of these experiments was to induce the formation of few nuclei, providing more protein molecules suitable for growing crystals. All the optimizations performed for MBP- $\beta$ FXIIa did not give suitable diffracting crystals.

## 4.5 Co-crystallization with CMK inhibitors: PCK and PPACK

### 4.5.1 Co-crystallization with PCK

The search for crystallization conditions for MBP- $\beta$ FXIIa in complex with PCK inhibitor did not give any suitable hits to work with for optimization. To the best of our knowledge, no previous crystallographic structure of protein-PCK complexes has been reported.

### 4.4.2 Co-crystallization with PPACK

For the (MBP- $\beta$ FXIIa)-PPACK a wide search for crystallization conditions was done using commercially available crystallization suites. Initial hits were identified in the conditions listed in Table 4.4:

**Table 4.4:** Crystallization conditions for (MBP- $\beta$ FXIIa)-PPACK from the Pro-complex suite.

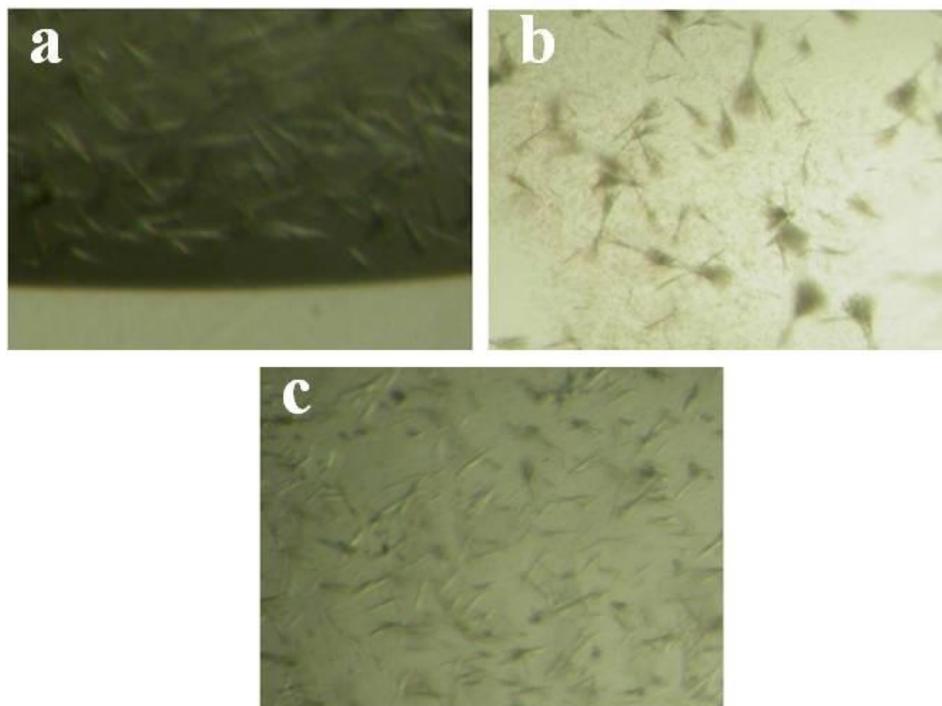
<b>Protein</b>	(MBP- $\beta$ FXIIa)-PPACK
<b>Method</b>	Sitting-drop
<b>Protein concentration</b>	8mg/mL
<b>Reservoir</b>	<ol style="list-style-type: none"><li>1. 0.1 M KCl, 0.1 M TrisHCl pH 8, 15% PEG2000MME</li><li>2. 0.1 M MgCl<sub>2</sub>, 0.1 M Na<sub>3</sub>Citrate pH 5.5, 5% PEG4000</li><li>3. 0.1 M NaCacodylate pH 5.5, 25% PEG4000</li></ol>
<b>Crystallization time</b>	~3 weeks
<b>Crystallization temperature</b>	20 °C

Initial hits from the conditions listed in Table 4.4 were very tiny and spread all over the drop. In order to grow high quality crystals, the conditions reported in Table 4.4 were used for an extensive optimization, which was performed using different approaches. Initially the “star search” method was used to get better crystals. In this case, one parameter at time was altered and then any improvement was monitored. The salt concentration was varied between 0.05 M and 0.2 M, while the percentage of PEG was changed between 5% and 20%. This simple approach allowed the identification of three new conditions, listed in Table 4.5.

**Table 4.5:** Crystallization conditions for (MBP-βFXIIa)-PPACK after initial optimization.

<b>Protein</b>	(MBP-βFXIIa)-PPACK
<b>Method</b>	Sitting-drop
<b>Protein concentration</b>	8 mg/mL
<b>Reservoir</b>	1. 0.1 M MgCl <sub>2</sub> , 0.1 M Na <sub>3</sub> Citrate pH 5, 15% PEG4000 2. 0.15 M MgCl <sub>2</sub> , 0.1 M Na <sub>3</sub> Citrate pH 5, 20% PEG4000 3. 0.15 M MgCl <sub>2</sub> , 0.1 M Na <sub>3</sub> Citrate pH 5.6, 10% PEG4000
<b>Crystallization time</b>	~1 week
<b>Crystallization temperature</b>	20 °C

The conditions listed in Table 4.5 gave tiny crystals as shown in the Figure 4.9.



**Figure 4.9: Crystals of (MBP-βFXIIa)-PPACK.** a) 0.1 M MgCl<sub>2</sub>, 0.1 M Na<sub>3</sub>Citrate pH 5, 15% PEG4000; b) 0.15 M MgCl<sub>2</sub>, 0.1 M Na<sub>3</sub>Citrate pH 5, 20% PEG4000; c) 0.15 M MgCl<sub>2</sub>, 0.1 M Na<sub>3</sub>Citrate pH 5.6, 10% PEG4000.

Since the obtained crystals were very small and clustered, a further optimization was carried out in order to reduce the nucleation. For this optimization, different salts (ZnAc<sub>2</sub>, NaCl, Am<sub>2</sub>SO<sub>4</sub>, MgAc<sub>2</sub>) were used, looking for an improvement in the crystal quality, due to a variation of the cation or the anion. The PEG percentage was also varied between 10% and 20%. However, all the trays did not provide crystals suitable for X-Ray experiments.

A different strategy featured the additive screen by Hampton Research, a kit designed to allow the rapid and convenient evaluation of 96 unique

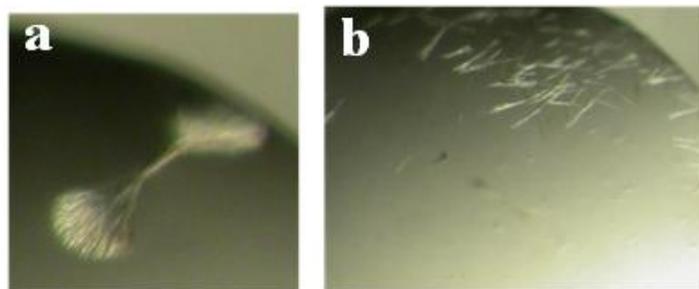
additives and their ability to influence the crystallization of the sample. In this case, single crystals were grown in the presence of spermine tetrahydrochloride. These crystals were tested at Diamond Light Source based in Oxford, but unfortunately they did not diffract.

Since no suitable crystals were obtained from this intensive optimization, the crystallization of (MBP- $\beta$ FXIIa)-PPACK was investigated through the incomplete factorial approach. The analysis of the grid from the incomplete factorial method allowed the identification of two new conditions for (MBP- $\beta$ FXIIa)-PPACK, reported in Table 4.6.

**Table 4.6:** Crystallization conditions for (MBP- $\beta$ FXIIa)-PPACK obtained from incomplete factorial grid.

<b>Protein</b>	(MBP- $\beta$ FXIIa)-PPACK
<b>Method</b>	Sitting-drop
<b>Protein concentration</b>	8 mg/mL
<b>Reservoir</b>	0, 0.1 M Na <sub>3</sub> Citrate pH 5.6, 0.1 M MgCl <sub>2</sub> 15% PEG335 15% PEG3350, 0.1 M NaAc pH 4.6, 0.2 M NaAc
<b>Crystallization time</b>	~1 week
<b>Crystallization temperature</b>	20 °C

Crystals obtained in these two conditions are shown in Figure 4.10.



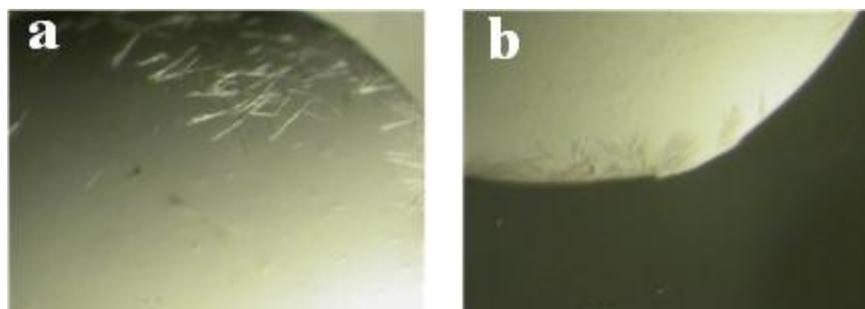
**Figure 4.10: Crystals of (MBP-βFXIIa)-PPACK grown from the incomplete factorial grid. a)** 0, 0.1 M Na<sub>3</sub>Citrate pH 5.6, 0.1 M MgCl<sub>2</sub>, 15% PEG335; **b)** 0.1 M NaAc pH 4.6, 0.2 M NaAc, 15% PEG3350.

Conditions producing the crystals shown in Figure 4.10 were further optimized, but no further improvement was obtained. A further incomplete factorial grid of the crystallization conditions was done as explained in Chapter 3. The 96 well grid built in this way allowed to identify other possible conditions, which are reported in Table 4.7.

**Table 4.7:** Crystallization conditions for (MBP-βFXIIa)-PPACK obtained from finer incomplete factorial grid.

<b>Protein</b>	(MBP-βFXIIa)-PPACK
<b>Method</b>	Sitting-drop
<b>Protein concentration</b>	8 mg/mL
<b>Reservoir</b>	10% PEG3350, 0.1 M NaAc pH 4.2, 0.2 M NaAc 15% PEG8000, 0.1 M NaAc pH 4.2, 0.1 M CoCl <sub>2</sub> 10% PEG8000, 0.1 M Na <sub>3</sub> Citrate pH 5.6, 0.2 M MgCl <sub>2</sub> 10% PEG4000, 0.1 M NaAc pH 4.2, 0.1 M NaAc 10% PEG8000, 0.1 M NaAc pH 4.2, 0.1 M NaAc
<b>Crystallization time</b>	2-10 days
<b>Crystallization temperature</b>	20 °C

Figure 4.11 shows crystals obtained for two of the conditions of the finer grid.



**Figure 4.11: Crystals of (MBP-βFXIIa)-PPACK grown from the finer grid. a) 10% PEG8000, 0.1 M Na<sub>3</sub>Citrate pH 5.6, 0.2 M MgCl<sub>2</sub>; b) 10% PEG4000, 0.1 M NaAc pH 4.2, 0.1 M NaAc.**

Experiments at the conditions listed in Table 4.7 were also repeated at two lower temperatures: 15 °C and 10 °C. At 10°C, an improvement was obtained for crystals grown in 10% PEG8000, 0.1 M NaAc pH 4.2, 0.1 M NaAc.

An additional optimization was done using the micro-seeding technique. Two different 96 well plates were made using this method. The used reservoir solutions were the same of the finer grid, since it gave the largest number of hits. The plates were tested at two different temperatures (4 °C and 20 °C). However, it was not possible to obtain crystals within experimental procedure.

## 4.6 Crystal structure of (MBP- $\beta$ FXIIa)-PPACK

### 4.6.1 Crystal diffraction and data collection

Diffracting crystals of (MBP- $\beta$ FXIIa)-PPACK (Figure 4.12) were grown in 0.15 M MgCl<sub>2</sub>, 0.1 M Na<sub>3</sub>Citrate pH 5.6, 10% PEG4000 by sitting drop vapour diffusion method. The crystals appeared between 10-15 days after the setup of the experiments. Crystals were transferred in a cryo-solution, containing 25% glycerol in mother liquor, for few seconds and then flash-frozen in liquid nitrogen. The diffraction experiments were performed at Diamond Light Source in Oxford at the beamline IO4, with a radiation of 0.9795 Å (12.658 keV). and using a Pilatus 6M-F detector. The crystals diffracted at 4 Å and a complete data-set was recorded.



**Figure 4.12: Needle crystal of (MBP- $\beta$ FXIIa)-PPACK frozen in a loop.** The crystal was grown in 0.15 M MgCl<sub>2</sub>, 0.1 M Na<sub>3</sub>Citrate pH 5.6, 10% PEG4000. The red arrow indicates the crystal.

#### 4.6.2 Data processing and scaling

The images were indexed and integrated using XDS[258] and were scaled and merged in AIMLESS.[254] The crystal was predicted to be tetragonal and belonging to the space group  $P4_12_12$ , with the following unit-cell parameters:

$a = 131.7 \text{ \AA}$ ,  $b = 131.7 \text{ \AA}$ ,  $c = 240.30 \text{ \AA}$ , and angles  $a = 90^\circ$ ,  $b = 90^\circ$ ,  $c = 90^\circ$ . Data collection statistics are summarized in Table 4.8.

**Table 4.8:** Data collection statistics for (MBP- $\beta$ FXIIa)-PPACK. Values in parentheses refer to the highest resolution shell.

DATA COLLECTION	
Resolution Range (Å)	29.46-4.00
Space group	P4 <sub>1</sub> 2 <sub>1</sub> 2
Cell dimensions	
a, b, c (Å)	131.27, 131.27, 238.2
$\alpha, \beta, \gamma$ (°)	90, 90, 90
Mosaicity	0.2
No of reflections	42458 (11910)
No of unique reflections	17015 (4810)
Multiplicity	2.5
I/ $\sigma$ (I)	3.8 (1.7)
Completeness (%)	94.8 (95.7)
$R_{merge}$ *	0.189 (0.468)
$R_{pim}$ **	0.21 (0.3)
CC 1/2***	0.95 (0.61)
Molecules per asymmetric unit	2
Matthews coefficient (Å <sup>3</sup> Da <sup>-1</sup> )	3.68
Solvent Content (%)	66.57

$$*R_{merge} = \frac{\sum_{hkl} \sum_{j=1}^N |I_{hkl}(j) - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_{j=1}^N I_{hkl}(j)}, \quad **R_{pim} = \frac{\sum_{hkl} \sqrt{\frac{1}{N-1}} \sum_{j=1}^N |I_{hkl}(j) - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_{j=1}^N I_{hkl}(j)}$$

\*\*\*CC 1/2 is the correlation coefficient value calculated with XDS to determine the resolution cut-off

### 4.6.3 Structure determination

The structure was solved by molecular replacement (MR) using PHASER[259]. The published structure of MBP (PDB code: 1HSJ, [250]) was first used and two molecules of MBP were found in the asymmetric unit. The published structure of FXII zymogen (PDB code 4DXE[144]) was then used to determine the relative position of  $\beta$ FXIIa to MBP with a second

round of MR in PHASER. The MR was successful, resulting in a log-likelihood gain (LLG) score of 1537, while the obtained translation function Z-score (TFZ) was 17.0. The initial model is characterized by the presence of two copies of MBP- $\beta$ FXIIa in the asymmetric unit, with a Matthew's coefficient[260] of  $3.68 \text{ \AA}^3/\text{Da}$  and 66.57% of solvent content.

#### **4.6.4 Model building and refinement**

Because of the medium resolution, the electron density in some parts of the model was poor. For this reason, the following strategy was adopted: first, the chains with poor electron density were removed. A first rigid body refinement cycle was hence performed. Hence, water molecules were placed, according to the electron density. Following a second cycle of refinement, the electron density was checked and water molecules were replaced with poly-Ala chains, if the electron density was good enough. A third refinement step was followed by a density modification step and a manual modification of the model in Coot. This procedure allowed to rebuild the entire chain with the exception of the loops. The latter were rebuilt, using ARP/wARP loop tool from CCP4 suite[261]. Once the entire chain was build, the NCS symmetry was applied to create the symmetric molecule. The solution was then subjected to several cycles of coordinates minimization.[257] Each run was alternated with manual model building using Coot.[256] In the final stage, two maltose molecules could be observed in electron density, together with the PPACK and two GlcNAc and they were all modelled accordingly. The final model has  $R_{work}$  and

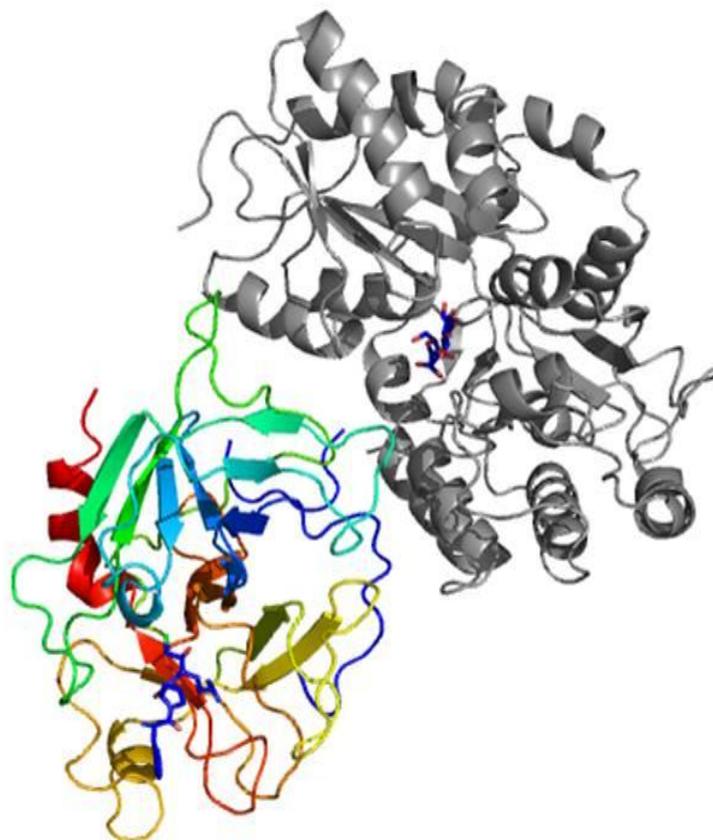
$R_{free}$  values of 27.61% and 35.95%, respectively. The refinement statistics are summarized in Table 4.9.

**Table 4.9:** Refinement statistics for (MBP- $\beta$ FXIIa)-PPACK. Values in parentheses refer to the highest resolution shell.

REFINEMENT	
$R_{work}$	0.2761 (0.3221)
$R_{free}$	0.3595 (0.3544)
Number of non-hydrogen atoms	9281
Macromolecules	9281
Ligands	6
Protein residues	1216
Overall $B$ -factor (Å)	102.94
RMS(bonds)	0.016
RMS(angles)	1.40
Ramachandran favoured (%)	85
Ramachandran allowed (%)	11.3
Ramachandran outliers (%)	3.7
Clash score	5.11
$R_{work} = \frac{\sum_{hkl}  F_{obs}  - k F_{calc} }{\sum_{hkl}  F_{obs} } \quad R_{free} = \frac{\sum_{hkl \in T}  F_{obs}  - k F_{calc} }{\sum_{hkl \in T}  F_{obs} }$	

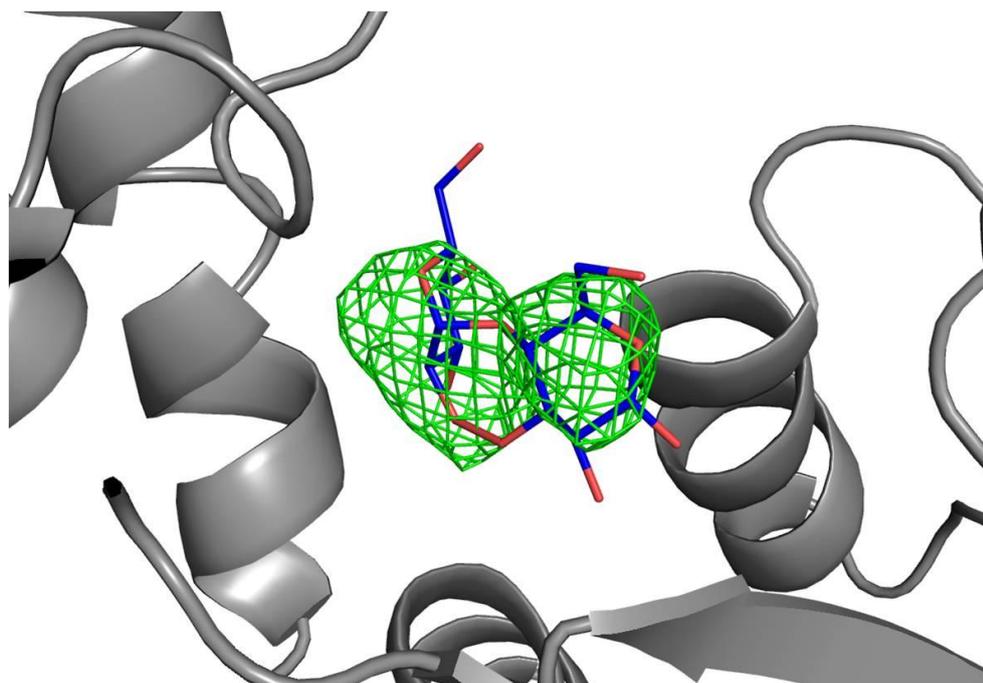
#### 4.6.5 (MBP- $\beta$ FXIIa)-PPACK structure

The structure of MBP- $\beta$ FXIIa is shown in Figure 4.13.



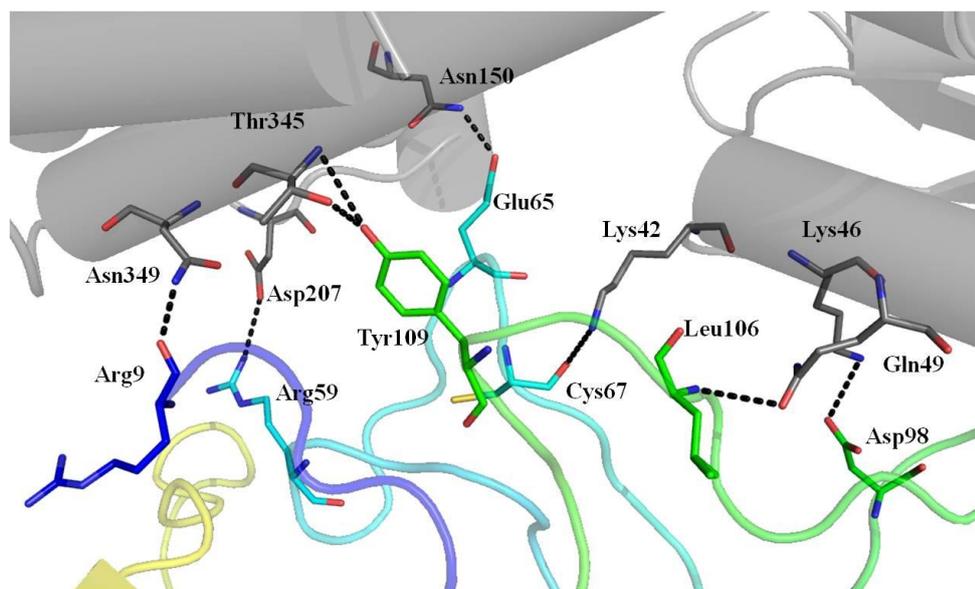
**Figure 4.13: Crystal structure of (MBP- $\beta$ FXIIa)-PPACK.** Cartoon representation of one copy of MBP- $\beta$ FXIIa with MBP shown in grey and  $\beta$ FXIIa in rainbow. The maltose bound to MBP is drawn in stick in blue. PPACK bound to  $\beta$ FXIIa is drawn in stick in blue.

The model includes two molecules of maltose, one for each copy of MBP, that were modelled in the  $F_O - F_C$  map as shown in Figure 4.14.



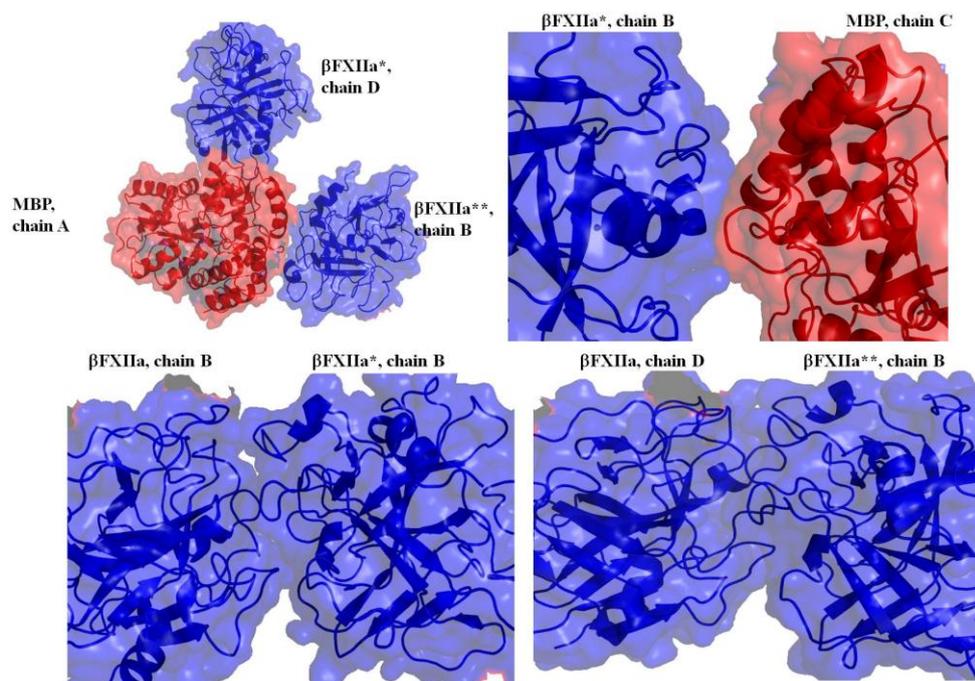
**Figure 4.14:**  $F_o - F_c$  electron density map for maltose residue. Fourier difference map, visualized at  $3.0\sigma$  level, showing the position of maltose (blue) bound to MBP (grey).

A detailed analysis of the complex structure revealed that the interactions stabilizing the MBP- $\beta$ FXIIa complex are extended over a large surface area ( $654.8 \text{ \AA}^2$ , calculated with PISA server [262]) as shown in Figure 4.15. In particular, the N-terminal Arg9 (chymotrypsin residue numbering is used in the description of the crystallographic structure of MBP- $\beta$ FXIIa), from the activation loop of  $\beta$ FXIIa is involved in an interaction with Asn349 of MBP. This interaction could influence the conformation of the activation loop.



**Figure 4.15: Intermolecular interactions of the MBP- $\beta$ FXIIa unit.** The figure shows the interactions occurring at the interface between MBP and  $\beta$ FXIIa. The interacting residues are drawn in stick in grey for MBP and in rainbow for  $\beta$ FXIIa.

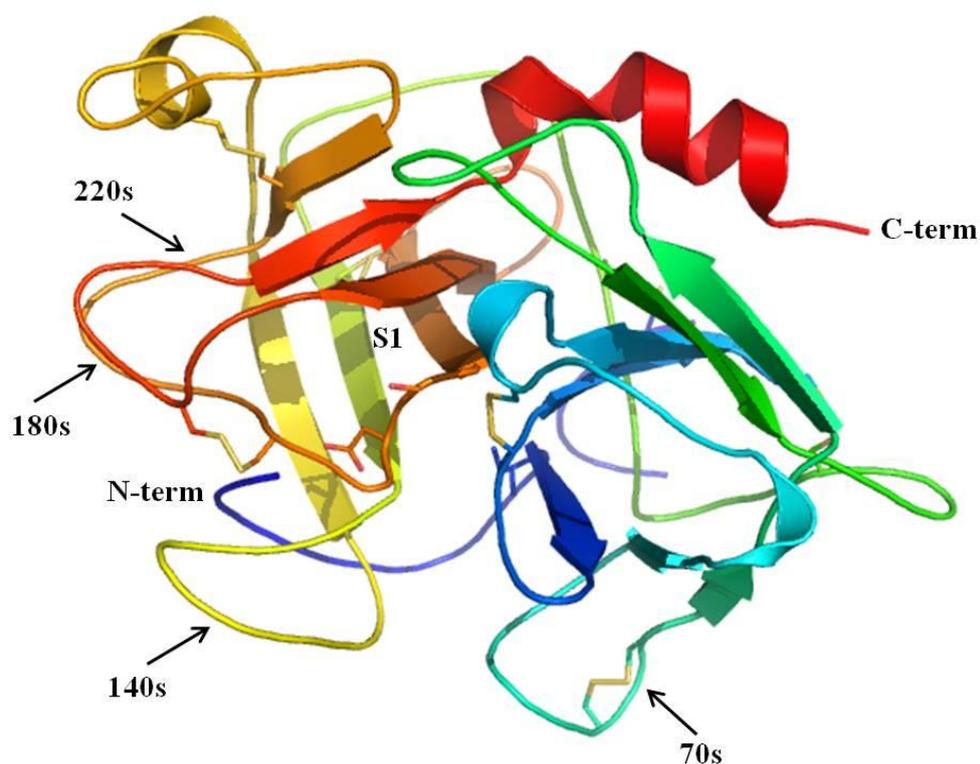
The analysis of the structure within the crystal lattice allowed to investigate the interactions involved in the formation of the crystal lattice between symmetry related molecules. In particular, we observe interaction both between MBP and  $\beta$ FXIIa and between two  $\beta$ FXIIa molecules, as reported in Figure 4.16. Interestingly, there are no MBP-MBP interactions.



**Figure 4.16: Crystal lattice interactions.** The figure illustrates the interactions forming the crystal lattice. MBP is shown in red and  $\beta$ FXIIa in blue. The symmetry related molecules are indicated with asterisks. **A)** MBP molecule (chain A) interacts with two molecules of  $\beta$ FXIIa of two symmetry related molecules (MBP- $\beta$ FXIIa<sup>\*</sup> interface: 635.9 Å<sup>2</sup>; MBP- $\beta$ FXIIa<sup>\*\*</sup> interface : 216.2 Å<sup>2</sup>); **B)** MBP (chain C) interacts with  $\beta$ FXIIa of a symmetry related molecule (MBP- $\beta$ FXIIa<sup>\*</sup> interface: 147.5 Å<sup>2</sup>); **C)**  $\beta$ FXIIa (chain B) interacts with its symmetry related ( $\beta$ FXIIa<sup>\*</sup>- $\beta$ FXIIa<sup>\*\*</sup> interface: 350.9Å<sup>2</sup>); **D)**  $\beta$ FXIIa (chain D) interacts with  $\beta$ FXII of a symmetry related molecule ( $\beta$ FXIIa<sup>\*</sup>- $\beta$ FXIIa<sup>\*\*</sup> interface: 498.8 Å<sup>2</sup>). Interface areas are calculated with the PISA server.[262]

#### 4.6.6 Crystal structure of $\beta$ FXIIa-PPACK

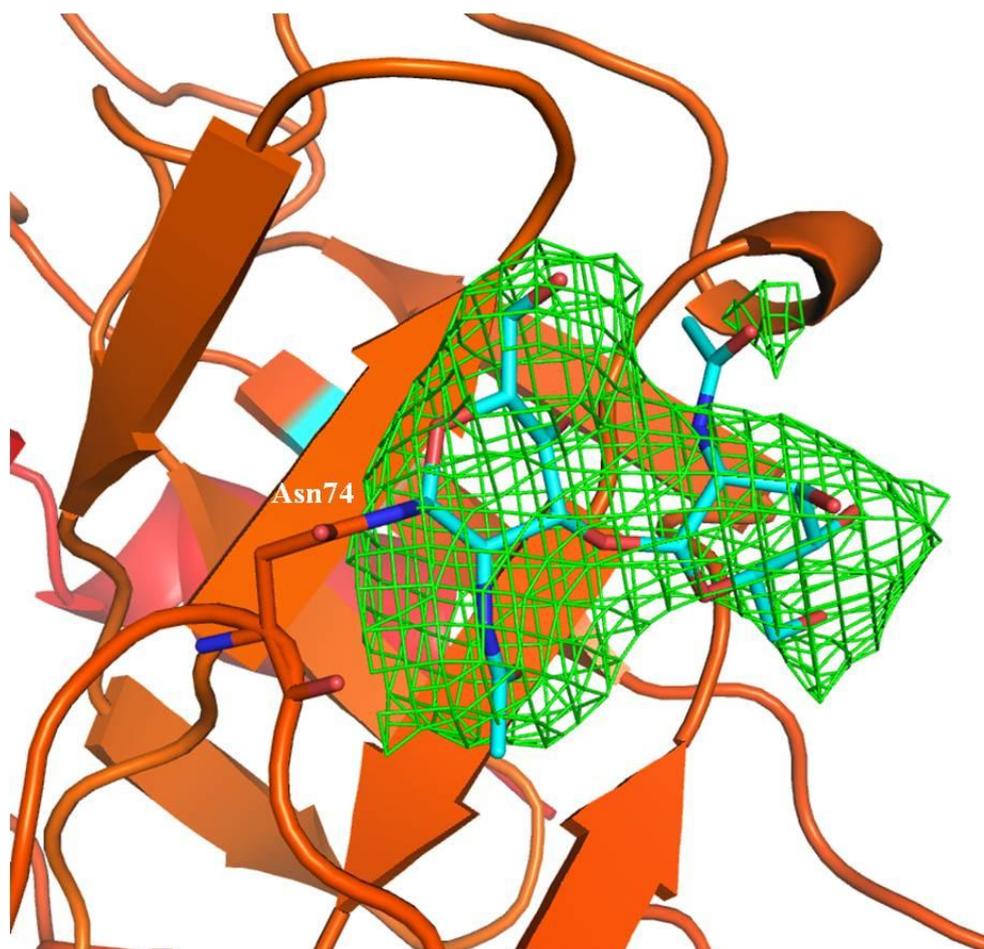
The overall folding (cf. Figure 4.17) of  $\beta$ FXIIa is similar to those observed in other serine proteases, with two interacting six-stranded  $\beta$ -barrel domains and the active site cleft containing the catalytic triad (His57, Asp102 and Ser195) located at the interface between the two domains.



**Figure 4.17:  $\beta$ FXIIa protease structure.** Cartoon representation of the  $\beta$ FXIIa topology, with the catalytic triad shown as sticks. Loops are labelled as 70s, 140s, 180s, 200s according to the chymotrypsin sequence numbering. Disulphide bridges are shown in yellow.

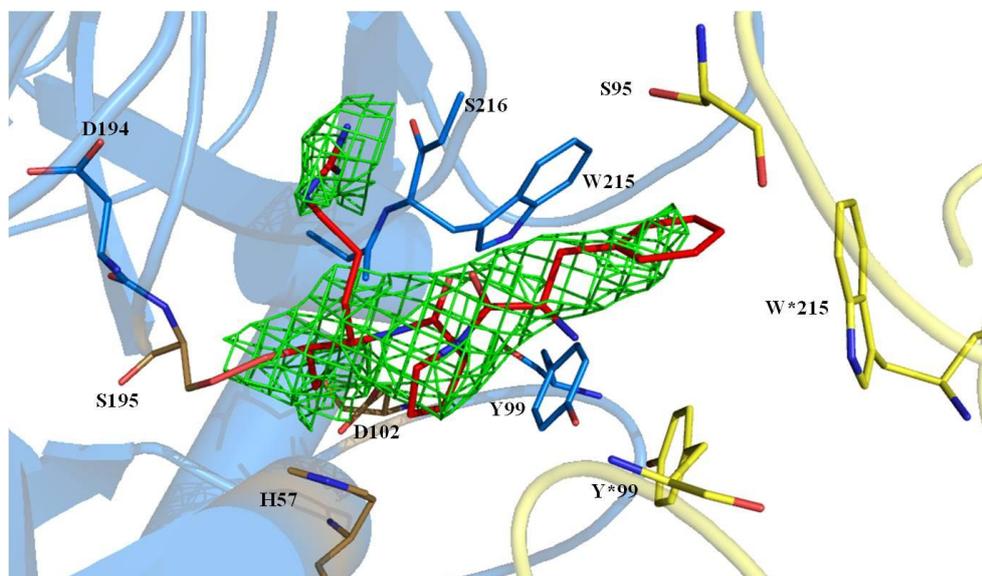
There are six disulphide bridges in the protease structure. Among these, Cys77-Cys80 is unique to FXII. Two GlcNAc residues are visible in the electron density and they are covalently linked to Asn74 (Figure 4.18).

A comparison of  $\beta$ FXIIa structure with a database of available PDBs, using the server DALI [263], indicates that the closest homologue of  $\beta$ FXIIa is Hepatocyte Growth Factor Activator (HGFA), resulting in 243 equivalent residues with an overall root mean square deviation (r.m.s.d.) of 1.3 Å (45% amino acid identity). HGFA is a serine protease that activates Hepatocyte Growth Factor (HGF), a growth factor that targets a large range of cells, in particular epithelial and endothelial cells.[264]



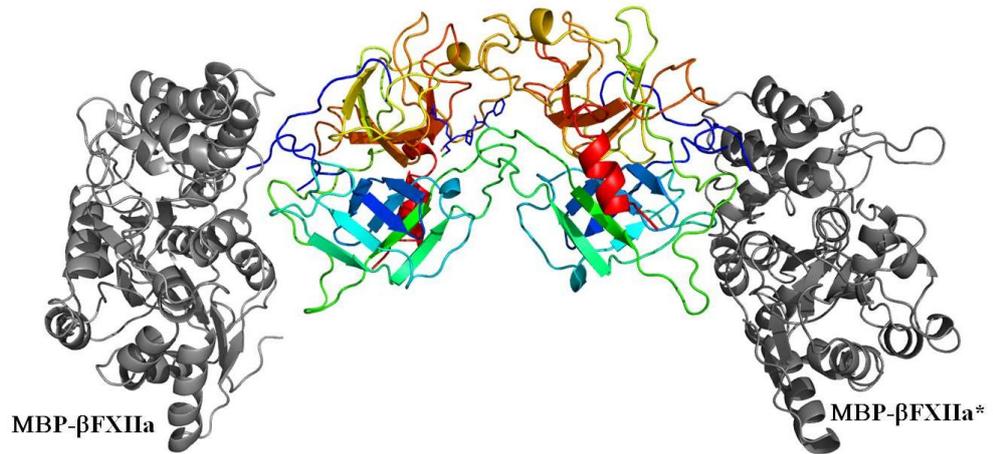
**Figure 4.18:**  $F_o - F_c$  electron density map for GlcNAc residues. Fourier difference map, visualized at  $2.5\sigma$  level, showing two GlcNAc units covalently attached to Asn74. Sugar residues are drawn in stick in cyan.

MBP- $\beta$ FXIIa crystallized only in the presence of the PPACK inhibitor. The analysis of the  $F_o - F_c$  maps confirmed the presence of PPACK. Figure 4.19 shows that the inhibitor interacts not only with the catalytic triad of the main molecule but also with residues belonging to a symmetry-related molecule.

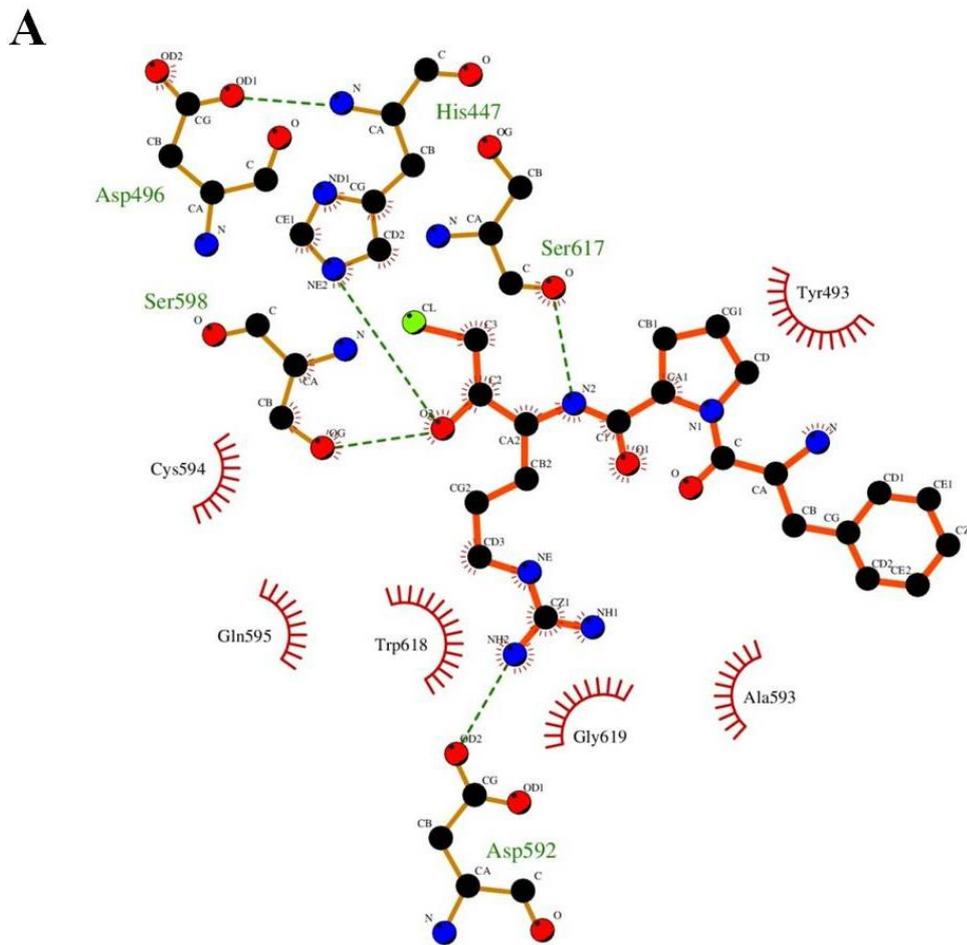


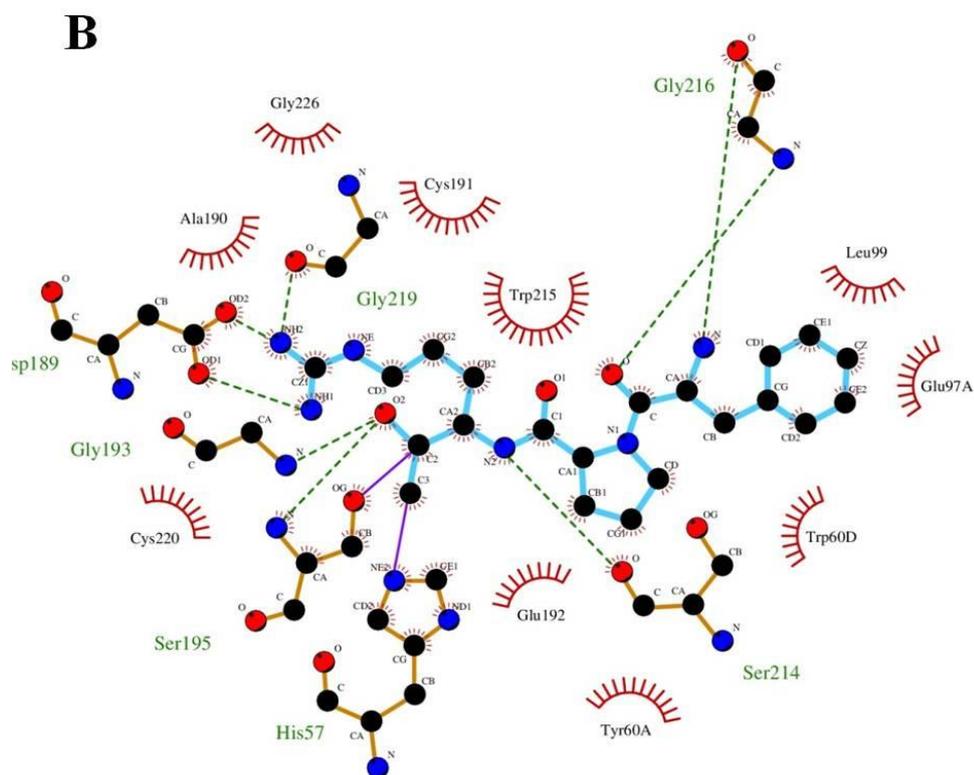
**Figure 4.19:**  $F_o - F_c$  electron density map for PPACK. Fourier difference map, visualized at  $2.0\sigma$  level, showing the position of PPACK in the active site cleft.  $\beta$ FXIIa is shown in cartoon in marine, while the symmetry related molecule is shown in cartoon in yellow. Side chains interacting with the inhibitor are shown in sticks. The catalytic triad is drawn in stick in sand (H57, Asp102, Ser195). The residues of the symmetry related molecule are drawn in stick in yellow and labelled with asterisks.

In contrast to other structures crystallized in the presence of PPACK,  $\beta$ FXIIa seems to form a bridge between two symmetry related molecules through the Phe residue (cf. Figure 4.20).



**Figure 4.20. Dimerization by PPACK.** The figure shows how PPACK forms a bridge between to symmetry-related molecules. MBP is shown in cartoon in grey and  $\beta$ FXIIa is in cartoon in rainbow. PPACK is drawn in stick in blue. The symmetry related molecule is indicated with an asterisk.



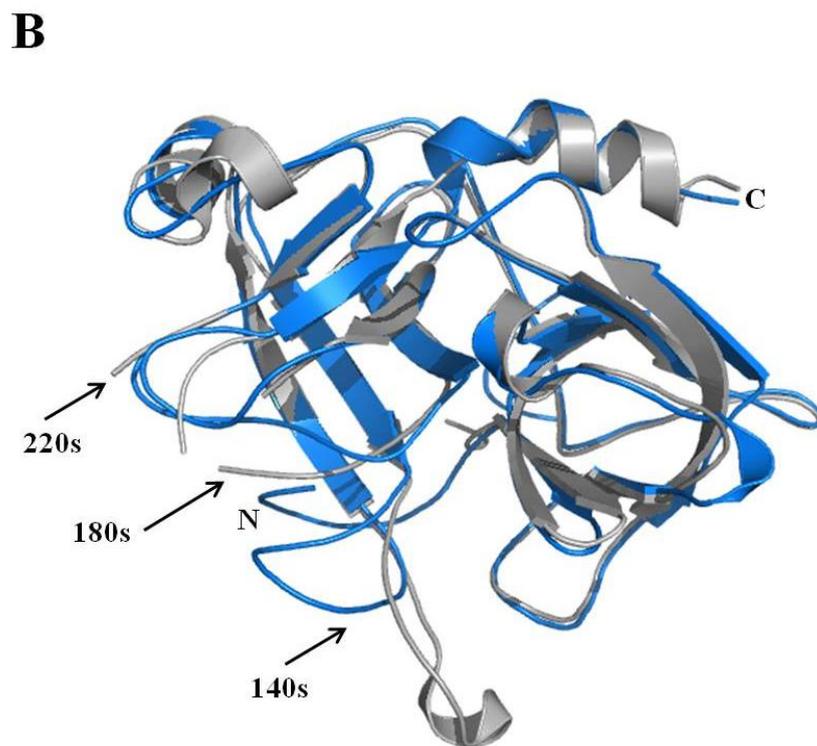
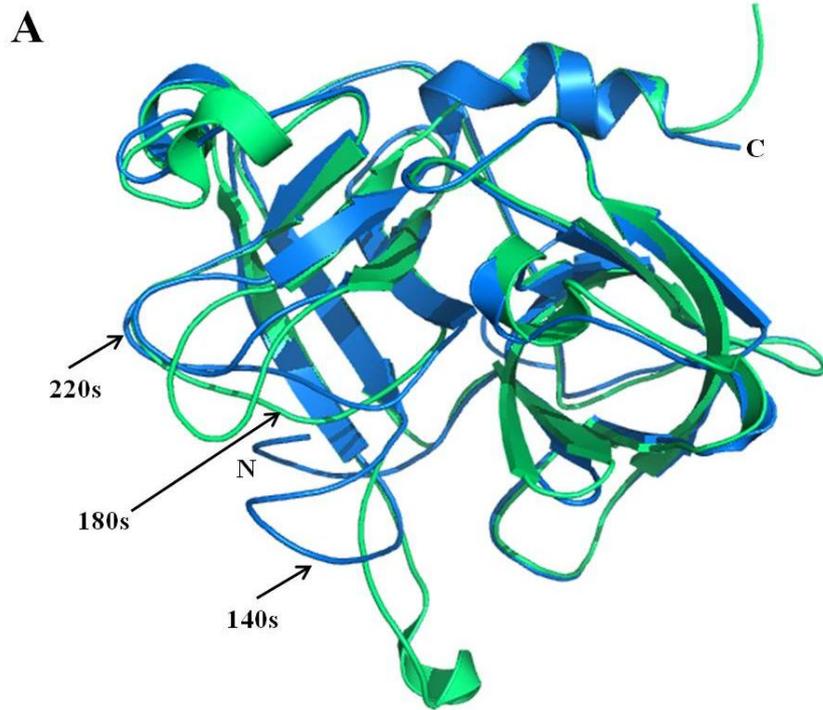


**Figure 4.21. Comparison between the structure of PPACK in A)  $\beta$ FXIIa and B) thrombin.** PPACK is shown in ball and sticks. C atoms in black, O atoms in red, N atoms in blue, Cl atoms in green. Ligand bonds are shown in orange and cyan, for PPACK in  $\beta$ FXIIa and thrombin (PDB code: 1PPB[172]), respectively.

An analysis of the conformation of PPACK in the PPACK- $\beta$ FXIIa complex has been performed using LIGPLOT (cf. Figure 21).[265] We also report the same analysis for the PPACK-thrombin, for comparison. The main differences between the two complexes are (i) the position of the Pro which is flipped of  $180^\circ$  in the PPACK- $\beta$ FXIIa complex; (ii) the Phe of PPACK is stabilized by interactions with thrombin residues. These interactions are not present in the PPACK- $\beta$ FXIIa complex and the Phe residue is stabilized by

interactions with residues belonging to the symmetry related molecule (cf. Figure.4.19).

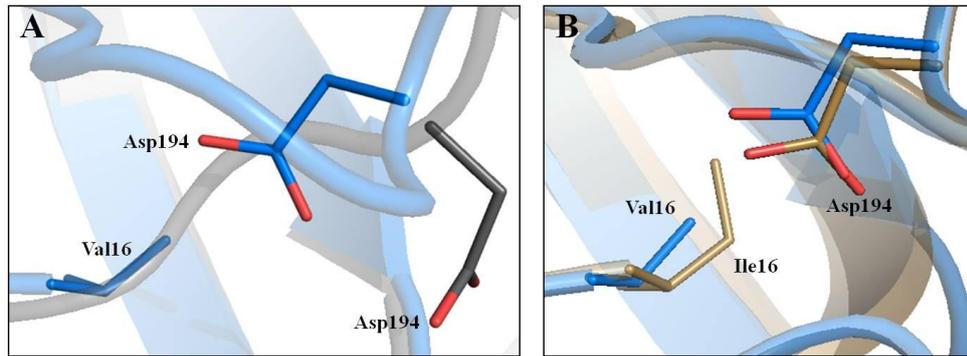
We compared our structure with (i) FXIIc (zymogen-like protease with N-terminal Val354 blocked due to the presence of additional Arg and Ser residues, PDB code: 4XDE) and (ii) FXIIac (zymogen-like protease with native N-terminus, PDB code: 4XE4). Comparison with FXIIc results in 241 equivalent residues with an overall r.m.s.d. of 1.9 Å (97% amino acid identity) while the same analysis performed on FXIIac results in 231 equivalent residues with an overall r.m.s.d. of 1.8 Å (95% amino acid identity). Superimposition of  $\beta$ FXIIa with the two structures, (cf. Figure 4.22), indicated that the main differences are in the conformation of the 140-loop and 220-loop. In particular, 140-loop forms a hairpin-like structure interacting with 180-loop. This conformation is commonly assumed by activated proteases. At variance, the 140-loop is positioned away from the body of the C-terminal in FXIIc and FXIIac structures.



**Figure 4.22: Structural comparison of MBP- $\beta$ FXIIa with FXIIc and FXIIac.** **A)** Superposed protease domain of MBP- $\beta$ FXIIa (marine) and FXIIac (green). **B)** Superposed protease domain of MBP- $\beta$ FXIIa (marine) and FXIIc (grey).

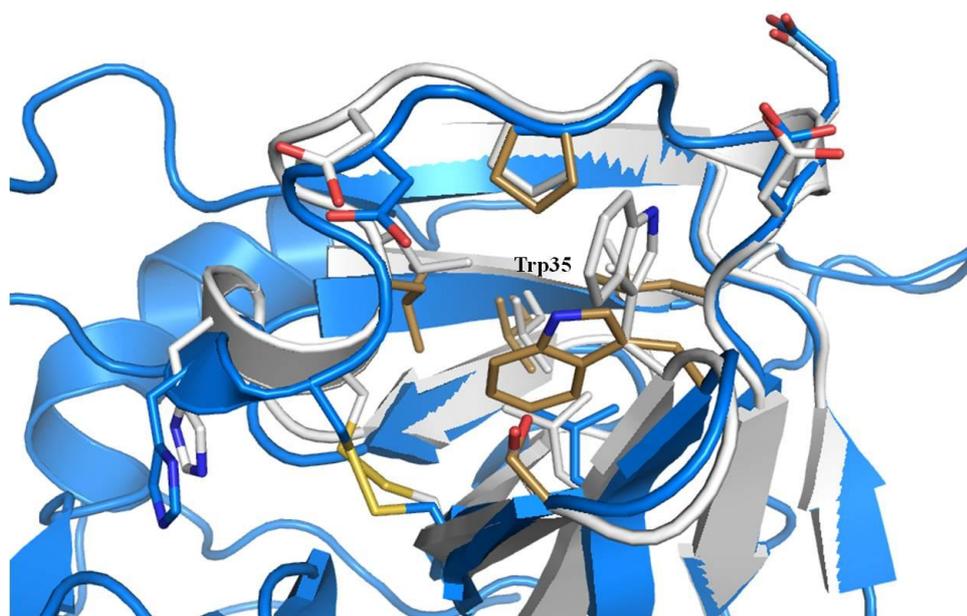
Another key feature that identifies MBP- $\beta$ FXIIa as the active form is the formation of the oxyanion hole. The oxyanion hole is a structural element of the enzyme, formed by the nitrogen backbone of Gly193 and Ser195. These two atoms form a positively charged pocket that activates the carbonyl of the scissile bond and stabilizes the negatively charged tetrahedral intermediate. In FXIIc, the different conformation of Asp194, which is flipped by 180° (Figure 4.22 A), does not allow the formation of the hole, while in MBP- $\beta$ FXIIa structure the Asp194 is in the right conformation, resulting in the oxyanion hole.

An additional feature that helps to identify MBP- $\beta$ FXIIa structure as active is the conformation of the N-terminal Val16. When the serine protease is active, the N-terminal residue Val16 is cationic and pairs with Asp194, while in the zymogen form Asp194 pairs with His40. As a consequence of the formation of this salt bridge, the active site cleft is exposed to the substrates.[146] A comparison of the crystallographic model with the zymogen structure showed that Asp194 is in the right conformation to form the salt bridge with Val16 (Figure 4.23 A). The structure was also superimposed to the structure of an active protease to confirm the position of the residue of interest (Figure 4.23 B).



**Figure 4.23: Structural comparison of MBP-βFXIIa with FXIIc zymogen and thrombin.** **A)** stick representation of the salt bridge interaction between Asp194 and Val16. The residues of MBP-βFXIIa are shown in marine, while the FXIIc is shown in grey. The figure also highlights how the Asp194 changes its conformation to make this important interaction respect to the zymogen. **B)** superimposition of Asp194 residue shown as marine stick to Asp194 of thrombin (PDB code: 1PPB, sand) as further confirmation of the active conformation of MBP-βFXIIa.

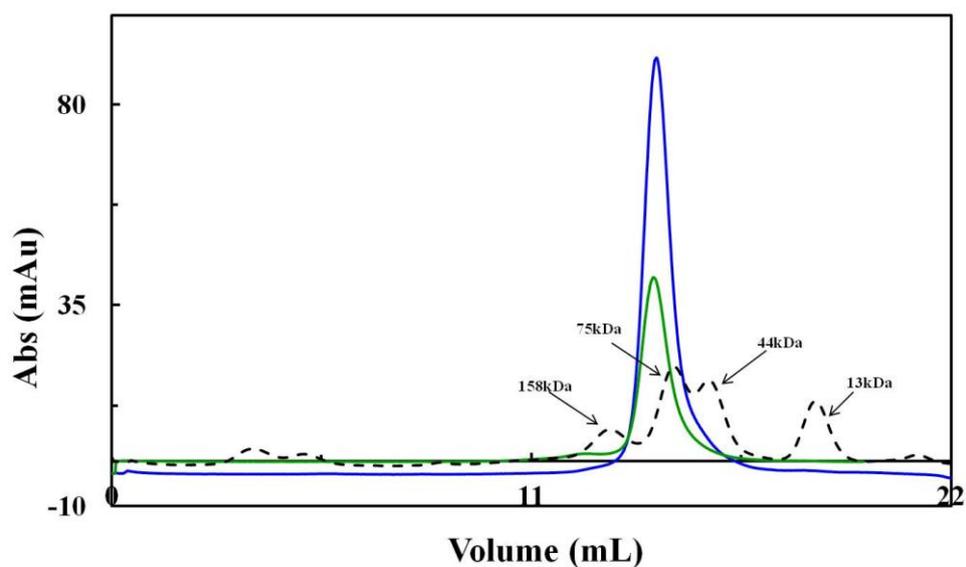
Another distinctive feature of the protein is the so-called H1 pocket, which is positioned at the center of the N-terminal β-barrel and opposite to the specificity pocket. The H1 pocket is composed by four Leu and one Trp residues, contributing to the hydrophobicity of the pocket itself. In FXIIc (PDB code: 4XDE) the region of the H1 pocket is altered, with the Trp35 rotated of 90°. This results in a shift of the main chain and repositioning of Trp35. For these reasons, the H1 pocket becomes partially buried. In MBP-βFXIIa, it assumes a closed conformation as shown in Figure 4.24.



**Figure 4.24: H1 pocket.** MBP- $\beta$ FXIIa is drawn in cartoon and coloured in marine; FXIIc is shown as cartoon in white. The figure also shows the residues of the S34 pocket drawn in white for FXIIc and in sand for MBP- $\beta$ FXIIa.

#### 4.7 PPACK-induced dimerization

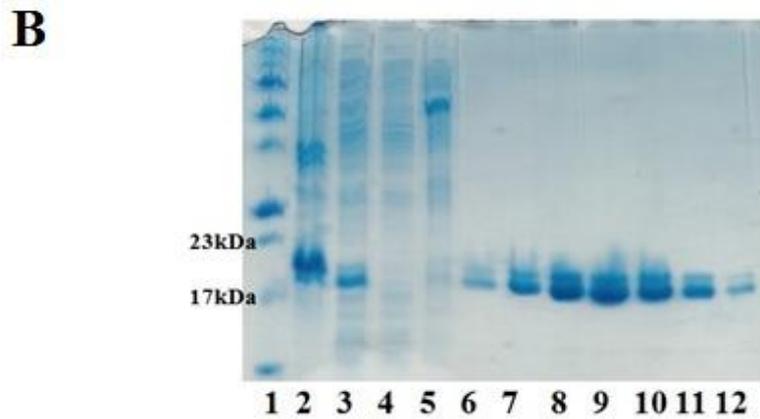
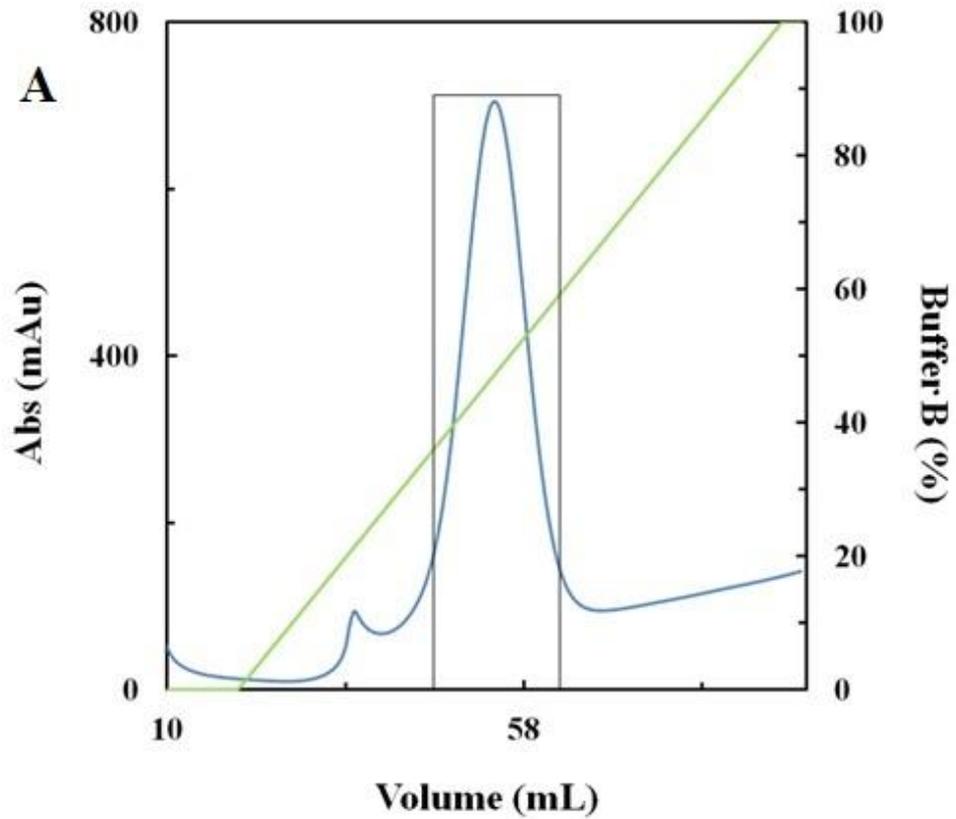
The crystallographic structure of MBP- $\beta$ FXIIa in the presence of PPACK shows that the inhibitor interacts also with a symmetry-related molecule (Figure 4.19), meaning that PPACK could induce dimerization of the protein. In order to verify that, the complex solution was applied on the gel filtration and compared with the elution profile of MBP- $\beta$ FXIIa on its own. The complex was prepared as already reported, by adding a 10X excess of PPACK and incubating the sample at 4 °C overnight. The complex was then resolved by gel filtration (Figure 4.25). Standards were run in order to better estimate the molecular mass of the formed complex. The chromatogram reported in Figure 4.25 does not show any shift with respect to the peak of the protein (the  $V_e$  for both elution profiles is 14.31 mL). Therefore, the dimerization is only an artefact of the crystal lattice.



**Figure 4.25: Analytical gel filtration analysis to verify the possible dimerization operated by PPACK.** UV traces from analytical gel filtration using Superdex 200 10/300 column. MBP- $\beta$ FXIIa: solid blue line; MBP- $\beta$ FXIIa-PPACK: solid green line. The column was calibrated with Aldolase ( $M_w$ : 158000 Da), Conalbumin ( $M_w$ : 75000 Da), Ovalbumin ( $M_w$ : 44000 Da) and RNase A ( $M_w$ : 13000 Da). The calibration curve is shown in dotted black line.

#### 4.8 Expression and purification of Ecotin

The Ecotin gene in pET-26b(+) vector was provided by GenScript. Protein expression was carried out at 37 °C in 1L of LB media. When  $OD_{600nm}=0.8$ , cells were induced with 0.5 mM IPTG, the temperature was lowered to 15 °C and the cells were grown for 24 hours. Cells were then lysed by sonication. The protein was purified by metal ion affinity chromatography through a gradient elution of imidazole (cf Figure 4.26).



**Figure 4.26: Purification of Ecotin by affinity chromatography.** A) UV trace from affinity chromatography using nickel column. The rectangular box highlights the peak fractions corresponding to the ones shown in the gel B) Non Reducing SDS-PAGE analysis. Lane 1: marker; lane 2: pellet; lane 3: supernatant loaded onto the column; lane 4 flow through; lane 5: wash; lanes 6-12: elution.

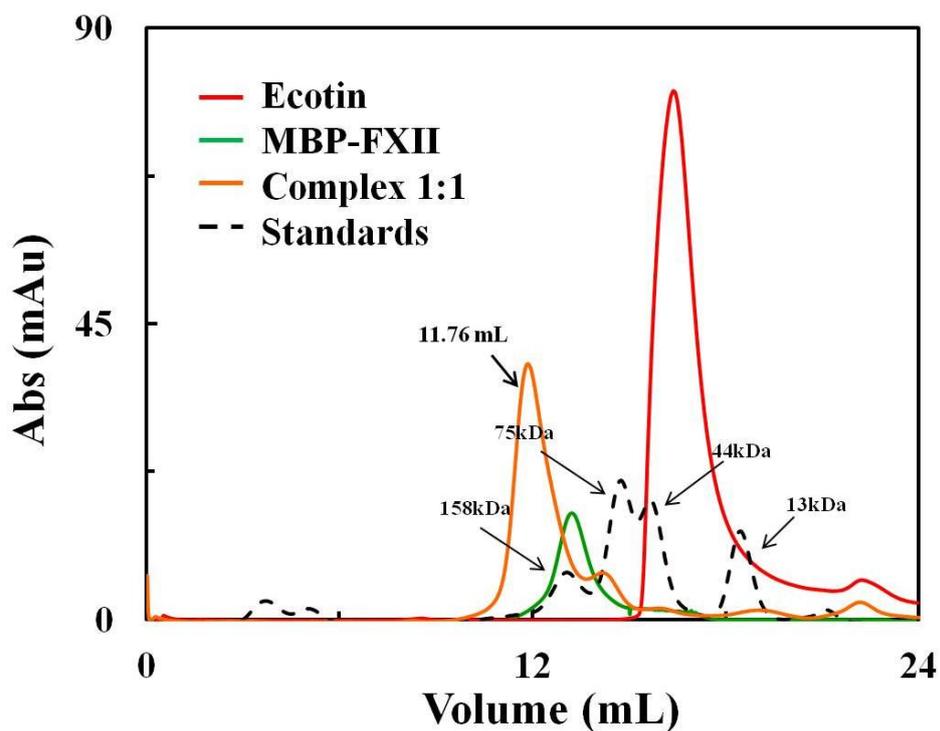
The SDS-PAGE gel shows two close bands. Neither anion exchange nor size exclusion chromatography allowed to separate these bands.

## **4.9 Complex with serine protease inhibitors: Ecotin and CTI**

### **4.9.1 Complex of MBP- $\beta$ FXIIa with Ecotin**

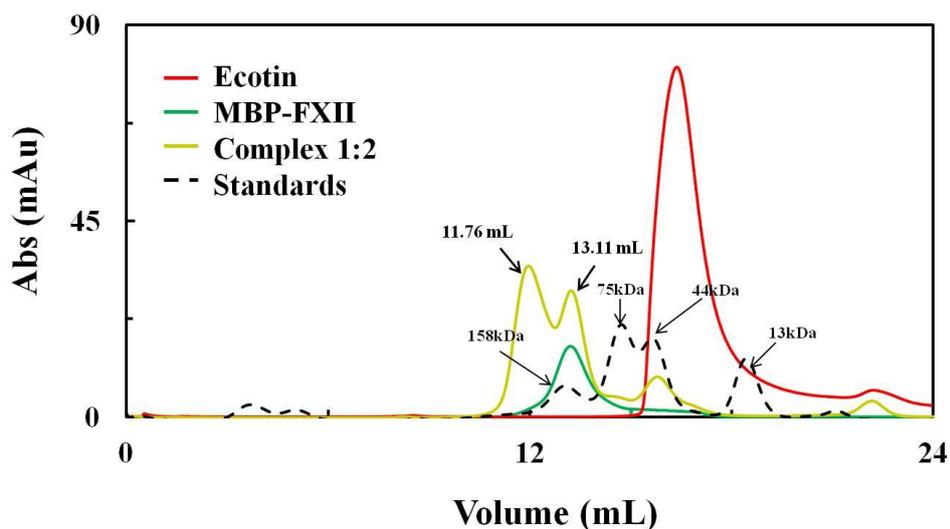
Size exclusion chromatography was used to isolate and identify the complex of recombinant MBP- $\beta$ FXIIa with Ecotin. Ecotin was present in 1X and 2X molar excess over MBP- $\beta$ FXIIa. The proteins were allowed to form a complex by incubation at 4 °C overnight. The complex was then resolved by gel filtration. A calibration curve was run in order to better estimate the molecular mass of the formed complex. Formation of the complex was verified by SDS-PAGE.

The chromatogram in Figure 4.27 shows a single peak for the 1:1 MBP- $\beta$ FXIIa - Ecotin complex corresponding to the tetramer (32 kDa for Ecotin plus 150 kDa for MBP- $\beta$ FXIIa). It should be noted that MBP runs as a dimer on the gel filtration column: this is due to the fact that the protein was not freshly prepared and so it aggregated. Therefore, the molecular weight of MBP- $\beta$ FXIIa is 158 kDa instead of 80 kDa. From the SDS-PAGE gel analysis (gel not showed) it was possible to confirm the formation of the complex: in fact, there are two bands, one for the MBP- $\beta$ FXIIa and one for Ecotin.



**Figure 4.27: Analytical gel filtration analysis of complex formation between MBP-βFXIIa and Ecotin in 1:1 ratio.** UV trace from analytical gel filtration using Superdex S200 10/300. MBP-βFXIIa-Ecotin (1:1): solid orange line; MBP-βFXIIa: solid green line; Ecotin: solid red line. The column was calibrated with aldolase ( $M_w$ : 158000 Da), conalbumin ( $M_w$ : 75000 Da), ovalbumin ( $M_w$ : 44000 Da) and RNase A ( $M_w$ : 13000 Da). The calibration curve is shown in dotted black line. . It is also shown the elution volume for the tetramer.

The elution profile for the 1:2 complex is shown in Figure 4.28. The chromatogram shows two peaks, the first of them corresponding to the tetramer. The second peak instead corresponds to the unbound MBP-βFXIIa, as confirmed by SDS-PAGE gel analysis (gel not shown).



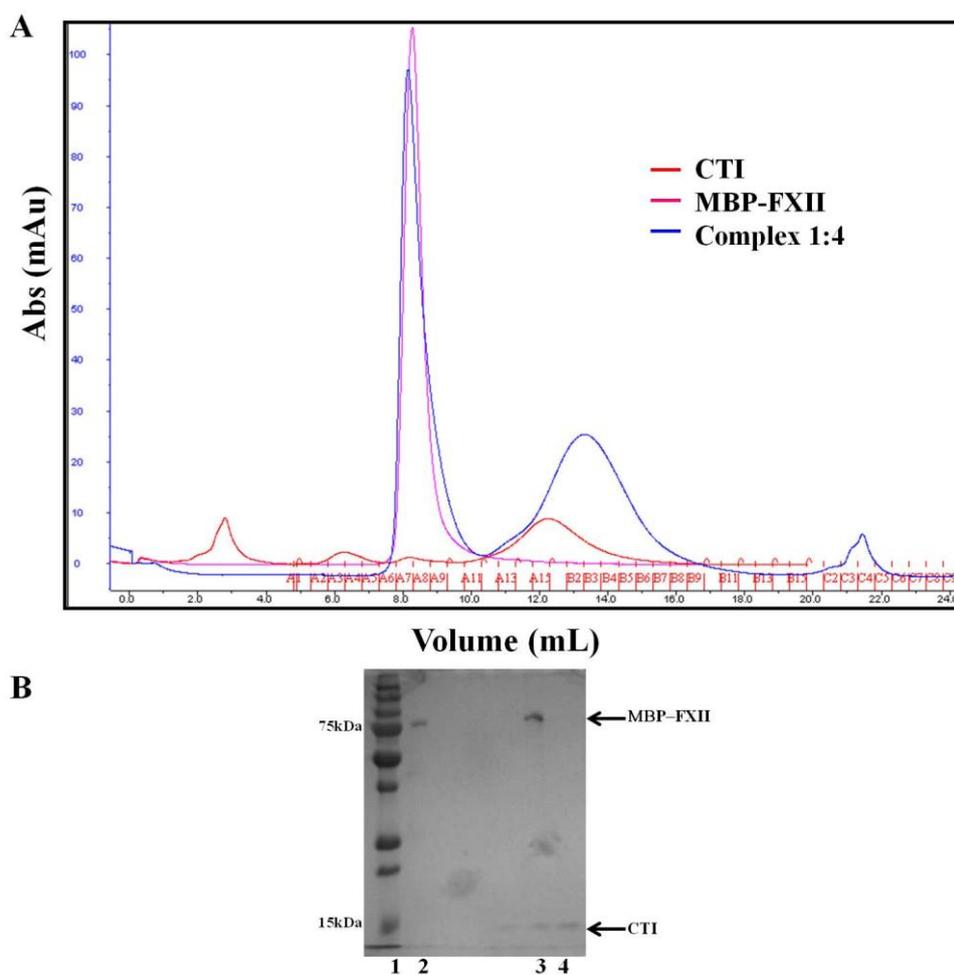
**Figure 4.28: Analytical gel filtration analysis of complex formation between MBP-βFXIIa and Ecotin in 1:2 ratio.** UV trace from analytical gel filtration using Superdex S200 10/300 column. MBP-βFXIIa-Ecotin (1:1): solid yellow line; MBP-βFXIIa: solid green line; Ecotin: solid red line. The column was calibrated with aldolase ( $M_w$ : 158000 Da), conalbumin ( $M_w$ : 75000 Da), ovalbumin ( $M_w$ : 44000 Da) and RNase A ( $M_w$ : 13000 Da). The calibration curve is shown in dotted black line. We also show the elution volumes for the tetramer and unbound MBP-βFXIIa in the complex.

The fractions corresponding to the complexes were concentrated for the crystallization trials.

#### 4.9.2 Complex of MBP-βFXIIa with CTI

Size exclusion chromatography was used to isolate and identify the complex of recombinant MBP-βFXIIa with CTI. CTI was present in 4X molar excess over MBP-βFXIIa in the initial mixture. The proteins were allowed to form a complex by incubation at 4 °C overnight. The complex was then resolved by gel filtration. Formation of the complex was verified by SDS-PAGE. The chromatogram in Figure 4.29 A shows a single peak shifted to the left, with

respect to the peak for MBP- $\beta$ FXIIa control (solid pink line). The second peak corresponds to the excess of CTI. The SDS-PAGE gel (cf. Figure 4.28 B) confirmed the formation of the complex: it is possible to visualize two bands, one for MBP- $\beta$ FXIIa and one for CTI. The fractions corresponding to the complex were concentrated for crystallization purposes.



**Figure 4.29: Analytical gel filtration analysis of complex formation between MBP- $\beta$ FXIIa and CTI in 1:4 ratio.** **A)** UV trace from analytical gel filtration using Superdex 200 10/300. MBP- $\beta$ FXIIa-CTI (1:4): solid blue line; MBP- $\beta$ FXIIa: solid pink line; CTI: solid red line. **B)** Non reducing SDS-PAGE analysis. Lane 1: marker; lane 2: MBP- $\beta$ FXIIa control; lane 3: MBP- $\beta$ FXIIa-CTI complex; lane 4:CTI control.

## **4.10 Crystallization of MBP- $\beta$ FXIIa with serine protease inhibitors**

### **4.10.1 Crystallization of MBP- $\beta$ FXIIa in complex with Ecotin**

The search for crystallization conditions for the (MBP- $\beta$ FXIIa)-Ecotin complex was performed using different crystallization screens. The ProComplex suite gave tiny crystals in different conditions. Some of the crystals were tested at the Synchrotron in Oxford, but they poorly diffracted. The initial conditions were used to generate an incomplete factorial grid, as previously reported. However, it was not possible to obtain single crystals.

### **4.10.2 Crystallization of MBP- $\beta$ FXIIa in complex with CTI**

For the (MBP- $\beta$ FXIIa)-CTI complex, only 75  $\mu$ L at 1 mg/ml were obtained. Therefore, only two crystallization screens were tested: ProComplex and MIDAS. However, no crystals were obtained.

## **4.11 Summary**

In this chapter, we reported the results obtained for MBP- $\beta$ FXIIa expressed in the DES system. The target protein was fused at the N-terminus with MBP. The presence of MBP has multiple advantages: (i) it acts as a chaperone as it assists the folding of the target protein; (ii) it increases the level of expression and the stability of soluble protein; (iii) MBP crystallizes easily providing the surface for the crystallization of the protein, thus enhancing the chances to get crystals.[244]

MBP- $\beta$ FXIIa gene with AAAAS linker was successfully cloned in pMT-PURO vector and transfected in S2 cells. The protein expression was confirmed by Western Blot using an anti-6X His tag antibody. A three step purification protocol was established, consisting of (i) anion exchange chromatography, (ii) affinity chromatography and (iii) size exclusion chromatography. The purity of the protein was checked after every step, by SDS-PAGE gel analysis in both reducing and non reducing conditions. The amount of obtained MBP- $\beta$ FXIIa varied between 4 and 10 mg for 3-6 L expression.

The kinetic characterization of MBP- $\beta$ FXIIa was performed by monitoring the change in absorbance at 405 nm following the cleavage of S2302. It revealed that the recombinant protein is active and its activity is comparable to commercial  $\beta$ FXIIa.

MBP- $\beta$ FXIIa protein was mainly used for crystallization purposes. A wide search of the conditions was performed, using commercial screens. However no hits were obtained. This could be due to the low concentration of the tested protein or because of the micro-heterogeneity of the sample.

In order to determine the crystal structure of the active form of MBP- $\beta$ FXIIa, the protein was complexed with two different peptide inhibitors: PPACK and PCK. Both of them are known to inhibit FXII, even if they are not specific. The complex with CMK peptides was prepared as reported.[251]

The research of crystallization conditions for (MBP- $\beta$ FXIIa)-PPACK was initially done using a wide range of available commercial crystallization suites, in order to first identify the biophysical conditions that yield some

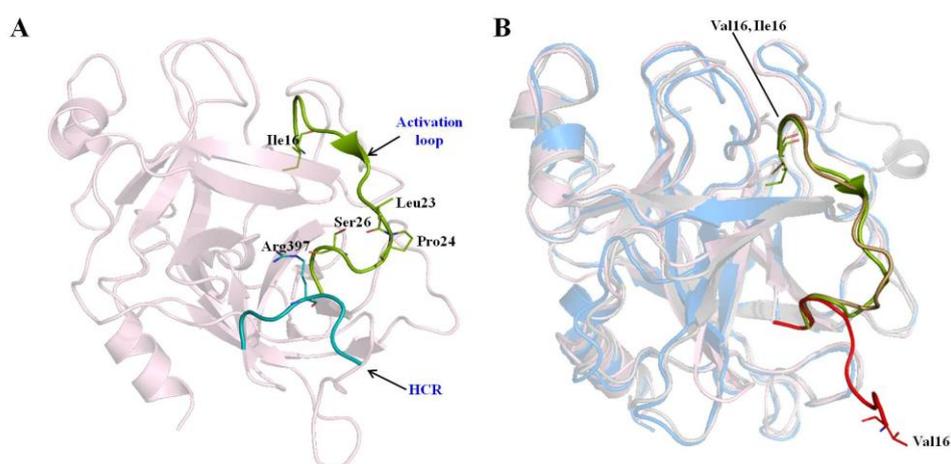
crystalline material. Initial hits were obtained for  $5.5 < \text{pH} < 8$  using PEG2000MME and PEG4000, as precipitant. The identified conditions were then used for extensive optimization. The optimization was performed through the systematic variation of one parameter at a time and monitoring any improvement in the crystallization. This optimization gave crystals suitable for X-Ray experiments. These crystals were grown in 0.1 M sodium citrate pH 5.6 with 10% PEG 4000 and 0.15 M  $\text{MgCl}_2$  at 20 °C. The crystals diffracted at a resolution of 4 Å. An optimization was done around the condition that gave diffracting crystals, but no better conditions were identified. Since the classic optimization approach (star search) did not provide satisfying results, a further optimization was attempted using the incomplete factorial method and seeding technique. A first incomplete factorial grid included 45 trials. Hits were obtained in 0.1 M NaCitrate pH 5.6 or 0.1 M NaAc pH 4.6 with 15% PEG 3350 and 0.1 M  $\text{MgCl}_2$  and 0.2 M NaAc respectively. Using these two conditions, an extensive optimization was carried out but no high-quality crystals were obtained. A second incomplete factorial grid was designed with additional rows to expand the concentration range for each precipitant and salt, and different PEGs and salts were also taken into account. This new grid gave five new conditions. However, all the obtained crystals were still too small for any diffracting experiments. Since all the optimizations were performed at 20 °C, the conditions identified through the second incomplete factorial grid were also tested at other two temperatures, 15 °C and 10 °C, to find out if a different temperature could improve the quality of the crystals. At 10 °C, an improvement was obtained for crystals grown in 10% PEG8000, 0.1 M

NaAc pH 4.2, 0.1 M NaAc. In fact, the amount of crystals in the drop was smaller than that grown at 20 °C. The lower density of crystals in the drop provides a larger surface area for crystal growth, allowing the formation of larger crystals. Another well-established technique for the screening and optimization of crystals is the seeding technique. Two micro-seeding matrices were made at 20 °C and 4 °C (representing the upper and lower limit of the range of typical crystallization temperatures[204]) with the aim to influence the nucleation event. No crystals suitable for diffraction experiments were obtained.

The crystals obtained from a first optimization were sent to Diamond Light Source in Oxford. A complete data-set was collected at 4 Å and the initial phases were solved by molecular replacement. A lot of effort was put in the refinement stage. Because the electron density was poor or missing, some parts of the  $\beta$ FXIIa chain were rebuilt, alternating refinement/density modification steps with manual modification of the model in Coot.

A preliminary analysis of the  $\beta$ FXIIa structure revealed the conformation hallmark of an active serine protease. When  $\beta$ FXIIa is superimposed to the FXII zymogen-like structures[144], the main differences are in the 140 and 220 loops. In particular, the 140-loop assumes a conformation already seen in other active serine proteases[266]: a hairpin-like structure, interacting with the 180-loop, and folded against the side of the C-terminal  $\beta$ -barrel. A further confirmation that the expressed protein is in its active conformation, comes from the comparison of some key residues of the model with the zymogen form of FXII and thrombin. In particular, the conformations of the N-terminal Val16 and Asp194 were evaluated. From this comparison, it

resulted that Val16 is positioned correctly. This conformation can be due to a direct interaction of the heavy chain remnant (HCR) with some residues of the activation loop with the consequent formation of a  $\beta$ -hairpin. This allow the proper insertion of Val16 into the protease core as reported for the hepatocyte growth factor activator HGFA,[267] which is found to be the closest homolog of FXII. For MBP- $\beta$ FXIIa, the electron density in correspondence of the HCR is very poor. However, the conformation of the activation loop is similar to that one found in HGFA (cf. Figure 4.30 A), while in the FXIIac model, the activation loop does not form the  $\beta$ -hairpin (Figure 4.30 B).



**Figure 4.30: Stabilizing interactions between HCR and the activation loop.** **A)** cartoon representation of the stabilizing interactions between the heavy chain remnant (HCR, cyan) and the activation loop (forest) in Hepatocyte Growth Factor Activator (HGFA) (PDB code: 1YC0[267]). Are also highlighted the main interactions that stabilize the b-hairpin of the activation loop and the interactions that it makes with the HCR. **B)** Comparison of MBP- $\beta$ FXIIa and FXIIac (PDB code: 4XE4[144]) with HGFA. For MBP- $\beta$ FXIIa it is not possible to visualize the HCR; however the conformation of the activation loop (sand) is similar to that of HGFA, with the formation of the  $\beta$ -hairpin. In FXIIac, instead the activation loop (red) does not form the  $\beta$ -hairpin.

The conformation of Asp194 was also assessed. In FXIIc, the residue is flipped of 180 °C and it is unable to make any interaction with N-terminal Val16; in MBP-βFXIIa, instead, the Asp194 has the proper conformation to form a salt bridge with Val16. This is confirmed when the structure is superimposed to the model of thrombin.

Crystals of MBP-βFXIIa were obtained only in presence of PPACK. The analysis of the Fourier map differences revealed the presence of the inhibitor in the active site cleft. A characteristic feature of the model, which is unique to FXIIa, is that the electron density for the inhibitor protrudes into a symmetry related molecule. In order to verify if PPACK induces a dimerization of the protein or it is an artefact of the crystal lattice, the complex was run on the Superdex 200 10/300 and the elution profile was compared to that of the protein on its own. The comparison revealed that PPACK did not interact with two molecules of FXII because of a dimerization process.

The conformation of PPACK in the PPACK-βFXIIa was compared with the one observed in the PPACK-thrombin complex. We found that the position of the Pro is flipped of 180° in the PPACK-βFXIIa complex and that the Phe of PPACK is stabilized by interactions with the symmetry related molecule, rather than with the molecule itself, as observed in thrombin.

MBP-βFXIIa was also crystallized with CTI and Ecotin, since crystallographic structures of MBP-βFXIIa in complex with different inhibitors can provide more structural information, that could be used for rational drug design. The complexes with the macromolecular inhibitors were separated by gel filtration and verified by SDS-PAGE gel. Once the

fractions of interest were identified and concentrated, several screens were tested. In the case of CTI, no hits were obtained because of the low concentration obtained. For the complex with Ecotin, the initial hits were further optimized and the obtained crystals sent for diffraction experiments. However, no diffraction pattern was obtained for any of the crystals. Different concentrations have to be tested in the future for crystallization with biological inhibitors.

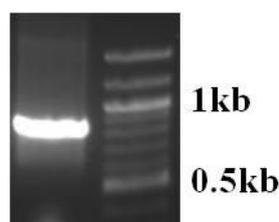
For PPACK, the difficulties encountered in reproducing the crystals could be due to the micro-heterogeneity of the sample. Different strategies can be tested to overcome this problem, i.e. addition of glycerol in every purification step, or elimination of contaminants, such as ions left from previous purification steps. Moreover, post-translational modification which negatively affect crystallization could be avoided with rational mutations. Finally, different crystallization methods, such as the micro-batch technique, could be tested.

## CHAPTER 5

### ENZYME KINETICS CHARACTERIZATION AND CRYSTALLIZATION OF $\beta$ FXIIa and $\beta$ FXIIG570R

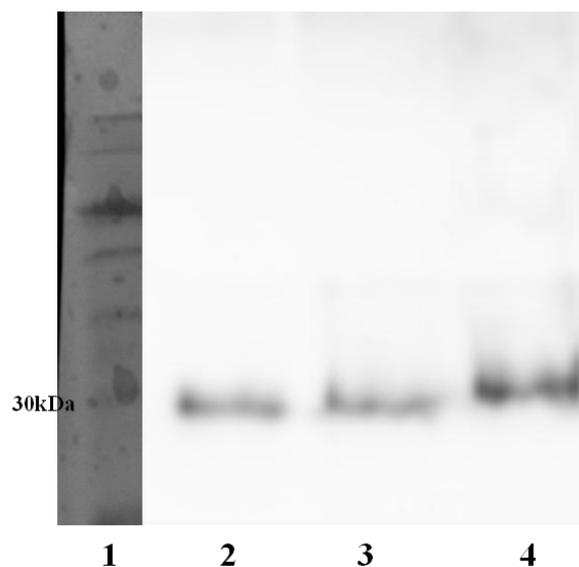
#### 5.1 Cloning of $\beta$ FXIIa in pMT-PURO vector and expression in DES system

The sequence encoding for human  $\beta$ FXIIa was amplified by PCR. PCR product of the correct size (0.75 kb) was obtained (cf. Figure 5.1). The PCR fragment was then cloned into pMT-PURO vector on BglIII/MluI restriction sites. The ligation of the insert into the vector was verified by carrying out mini-preparations of resulting colonies and then digesting the recombinant DNA for 4 h, using BglII and MluI endonuclease restriction sites.



**Figure 5.1: 1% Agarose gel showing amplified  $\beta$ FXIIa PCR fragment.**

Recombinant  $\beta$ FXIIa was expressed in *Drosophila* Schneider (S2) system with a 6X-His tag at the C-terminus. A Western Blot (cf. Figure 5.2) of the supernatant, using an anti-6X His tag antibody, confirmed that the protein was expressed three days after induction with 0.8 mM  $\text{CuSO}_4$  and secreted in the media.



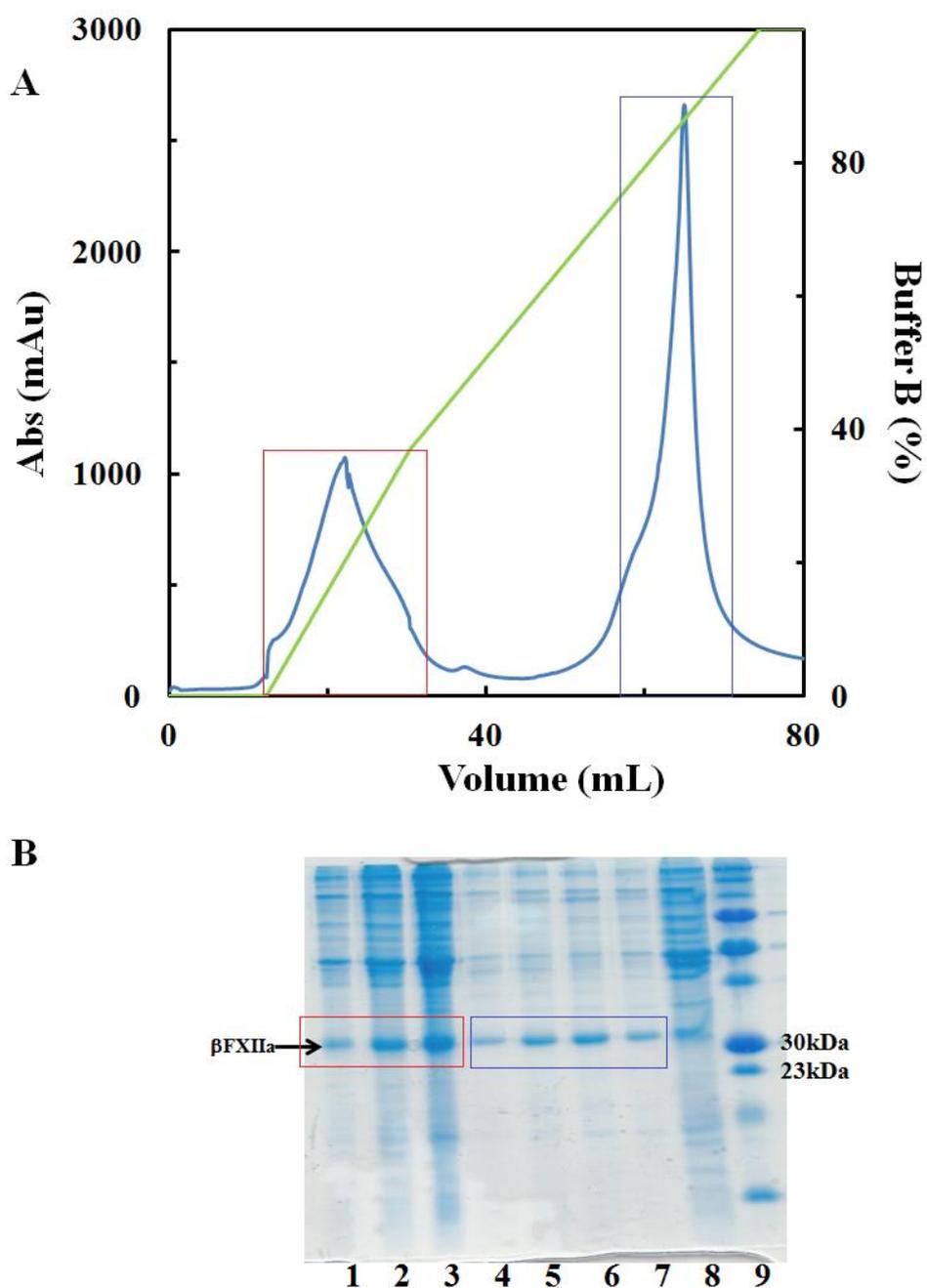
**Figure 5.2: Western blot of human  $\beta$ FXIIa using Anti-6X His tag primary antibody.** Lane 1: M; lanes 2-4: media collected 3, 6 and 7 days respectively after induction in non reducing condition.

## 5.2 Purification

### 5.2.1 Purification of $\beta$ FXIIa from DES

Recombinant  $\beta$ FXIIa was purified in three purification steps: (i) affinity chromatography, (ii) anion exchange chromatography, and (iii) size exclusion chromatography.

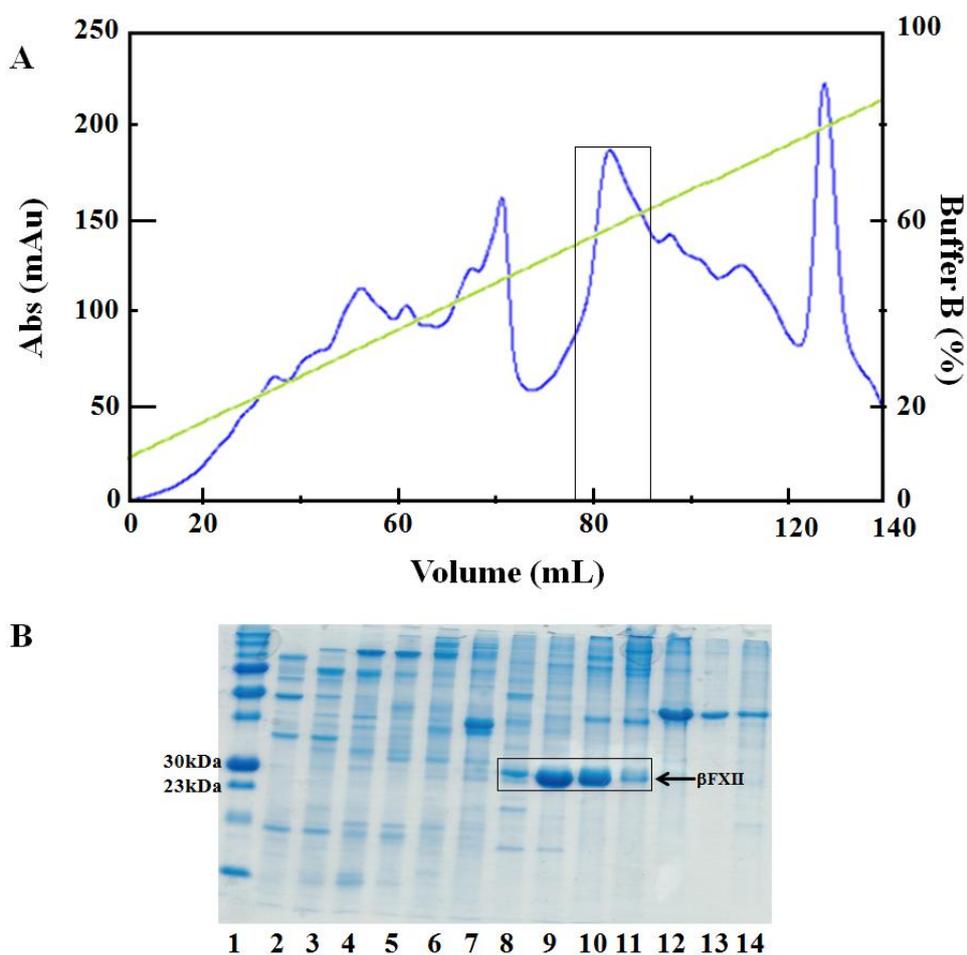
A Ni(II) excel column was used to perform affinity chromatography with an imidazole gradient elution (cf. Figure 5.3 A). A 15% SDS-PAGE gel (cf. Figure 5.3 B) under non reducing conditions confirmed the molecular weight of the protein and its degree of purity. The analysis of the gel indicated the presence of recombinant  $\beta$ FXII in the fractions corresponding to the peak in the chromatogram.



**Figure 5.3: Purification of  $\beta$ FXIIa by affinity chromatography.** **A)** UV trace from affinity chromatography using nickel excel column. **B)** Non reducing SDS-PAGE gel. lanes 1-7: elutions; lane 8: wash; lane 9 marker. Colored boxes in panel A highlight peaks that correspond to the fractions in panel B highlighted with boxes of the same colour.

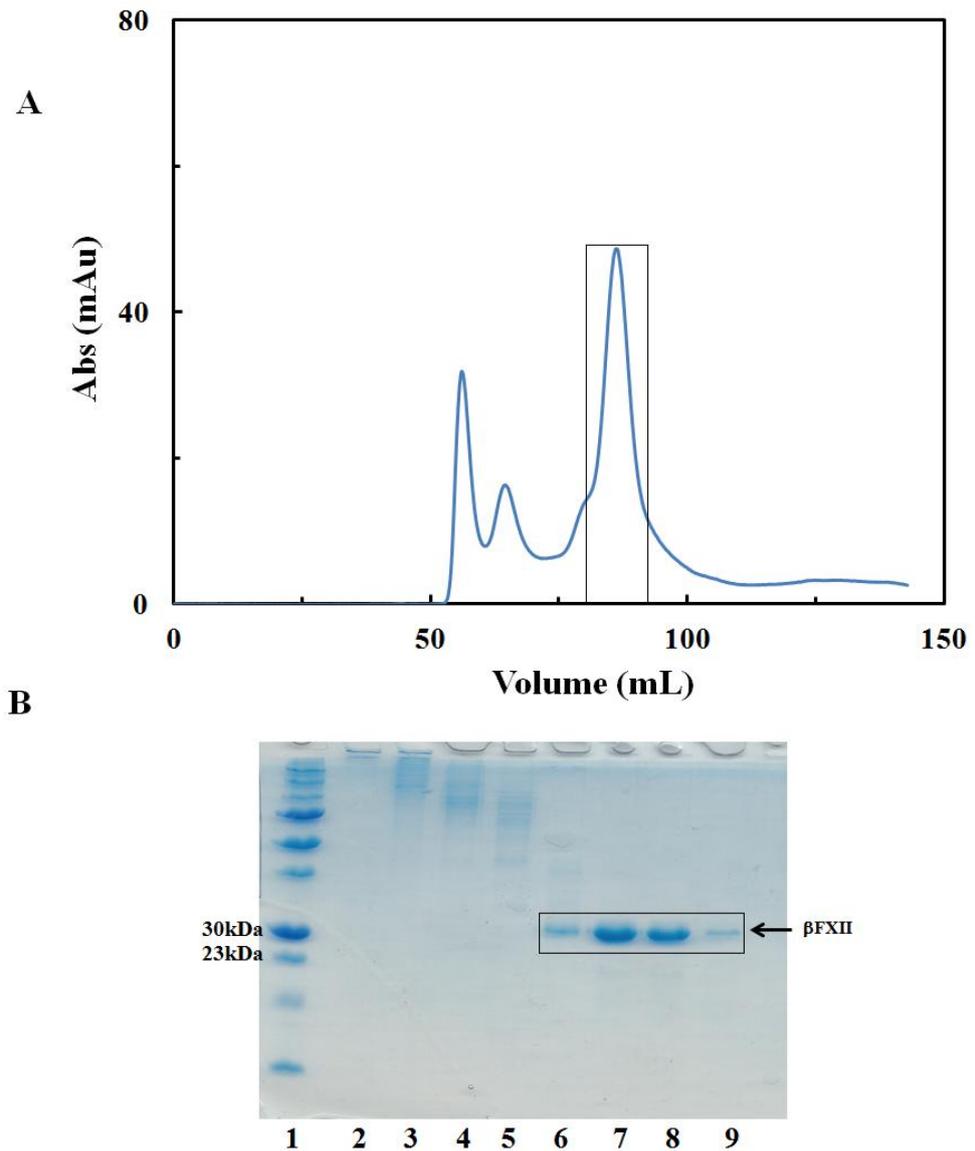
The fractions corresponding to the peak were loaded onto Resource Q column for anion exchange chromatography (cf. Figure 5.4 A) using a NaCl

gradient elution. A 15% SDS-PAGE gel (cf. Figure 5.4 B) was run to identify the peak of the elution profile corresponding to recombinant  $\beta$ FXII.



**Figure 5.4: Purification of  $\beta$ FXIIa by anion exchange chromatography.** A) UV trace from anion exchange chromatography using Resource Q column. The black box indicates the peak corresponding to  $\beta$ FXIIa. B) Non reducing SDS-PAGE. Lane 1: marker; lanes 2-14: elutions. The black box highlights the fractions corresponding to the protein of interest.

Recombinant  $\beta$ FXIIa was further purified by size exclusion chromatography (cf. Figure 5.5 A) and the fractions were analyzed by SDS-PAGE in non reducing conditions (cf. Figure 5.5 B).

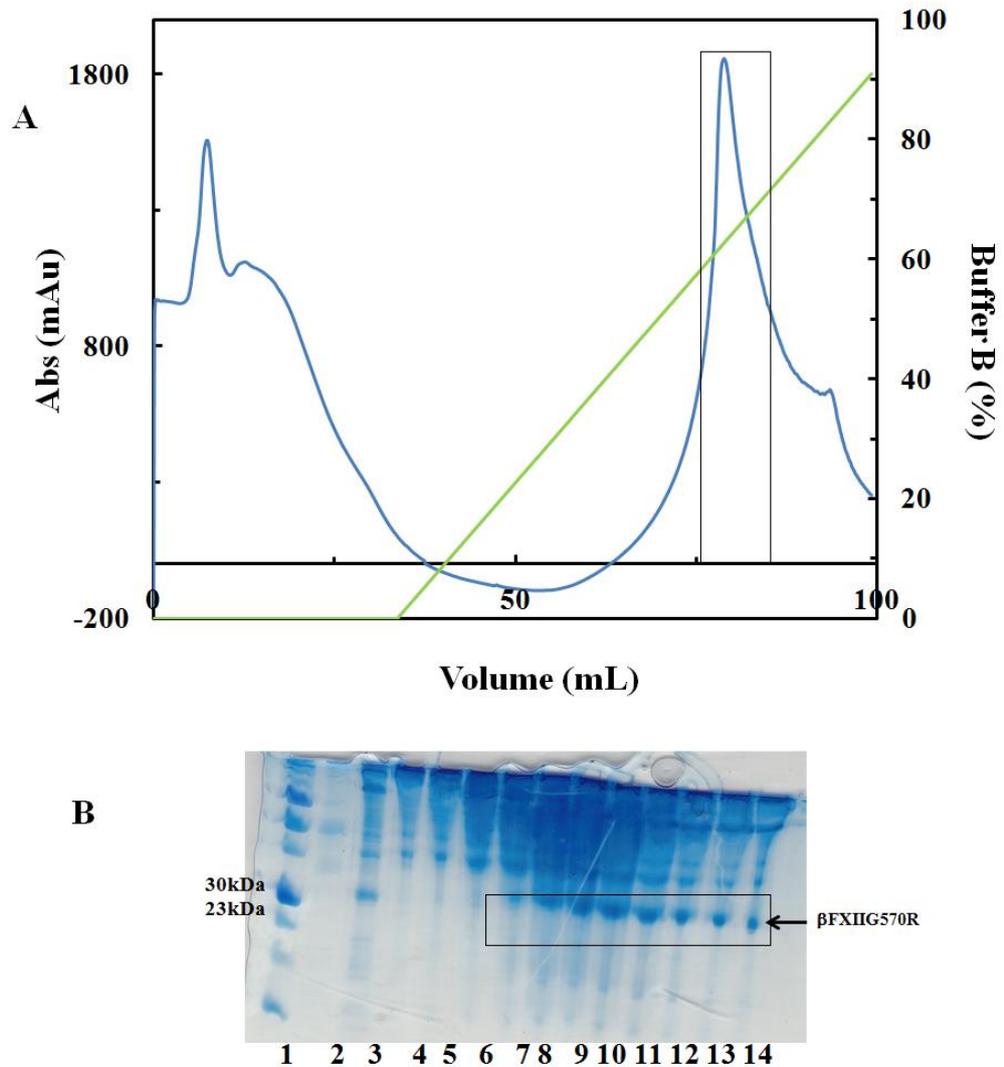


**Figure 5.5: Purification of  $\beta$ FXIIa by size exclusion chromatography.** A) UV trace from size exclusion chromatography using Superdex75 16/60. The black box indicates the peak of interest. B) Non Reducing SDS-PAGE. Lane 1: marker; lanes 2-9: elutions. The black box highlights the fractions corresponding to the protein of interest.

### 5.2.2 Purification of $\beta$ FXIIG570R from DES

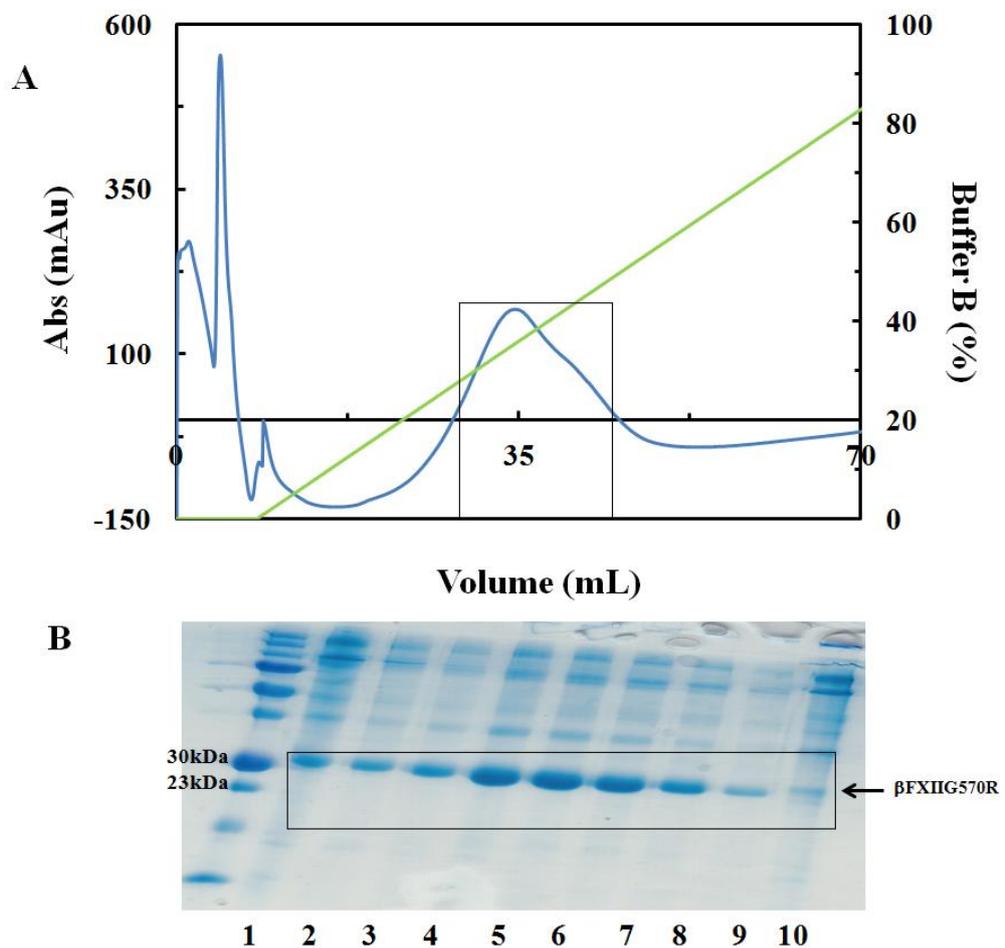
Purification of  $\beta$ FXIIG570R expressed in *Drosophila Schneider* cells was realized using ion affinity chromatography (two steps), followed by size exclusion chromatography. Ni(II) excel column (cf. Figure 5.6 A) was first

used to get rid of  $\text{CuSO}_4$ , and to remove the majority of the insect proteins, thus preparing the protein sample for the subsequent purification step. The protein of interest was identified by SDS-PAGE (cf. Figure 5.6 B).

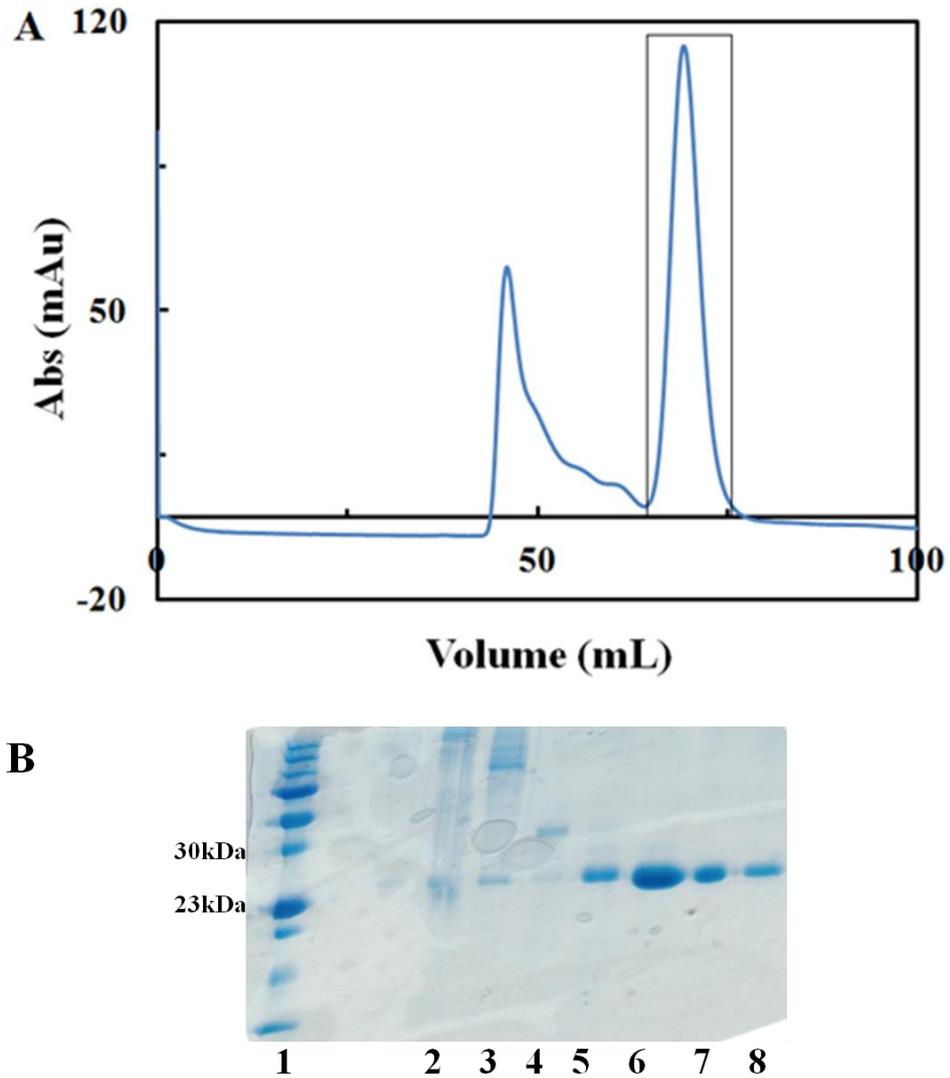


**Figure 5.6: Purification of  $\beta\text{FXIIG570R}$  by affinity chromatography. A)** UV trace from affinity chromatography using nickel excel column. The black box indicates the peak of interest. **B)** Non reducing SDS-PAGE gel analysis. Lane 1: Marker; lane 2-6: wash; lane 7-14: elution. The black box highlights the fractions corresponding to the protein of interest.

This first purification step was followed by other two purification steps, consisting of affinity chromatography and size exclusion chromatography. The UV traces for these steps are shown in Figure 5.7 A and Figure 5.8 A respectively, along with the respective SDS-PAGE gel analysis (cf. Figure 5.7 B and Figure 5.8 B, respectively). The protein was concentrated to 3.6mg/mL.



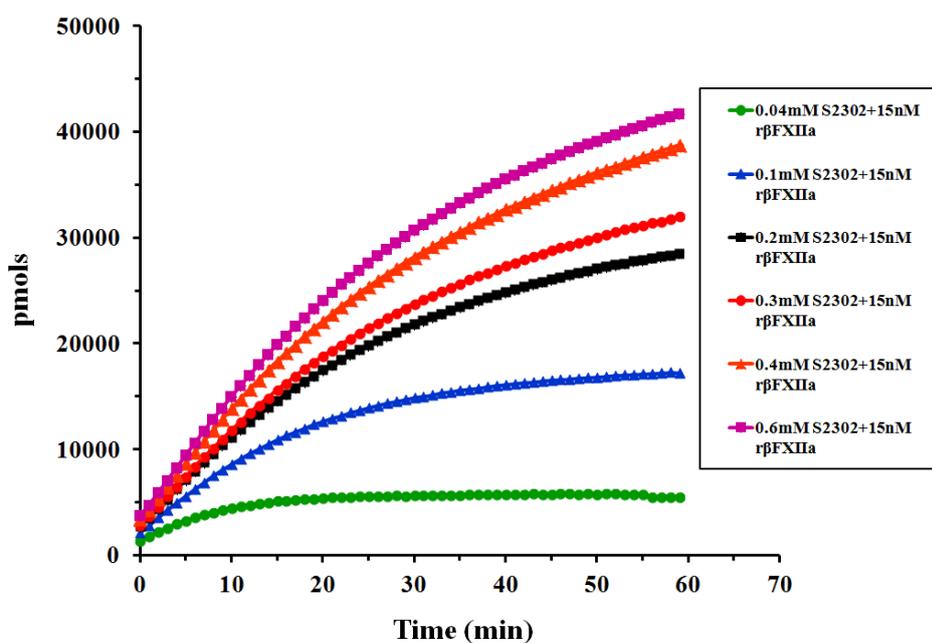
**Figure 5.7: Purification of  $\beta$ FXIIG570R by affinity chromatography. A)** UV trace from affinity chromatography using nickel column. The black box indicates the peak of interest. **B)** Non reducing SDS-PAGE gel analysis. Lane 1 marker; lanes 2-10: elution. The black box highlights the protein of interest.



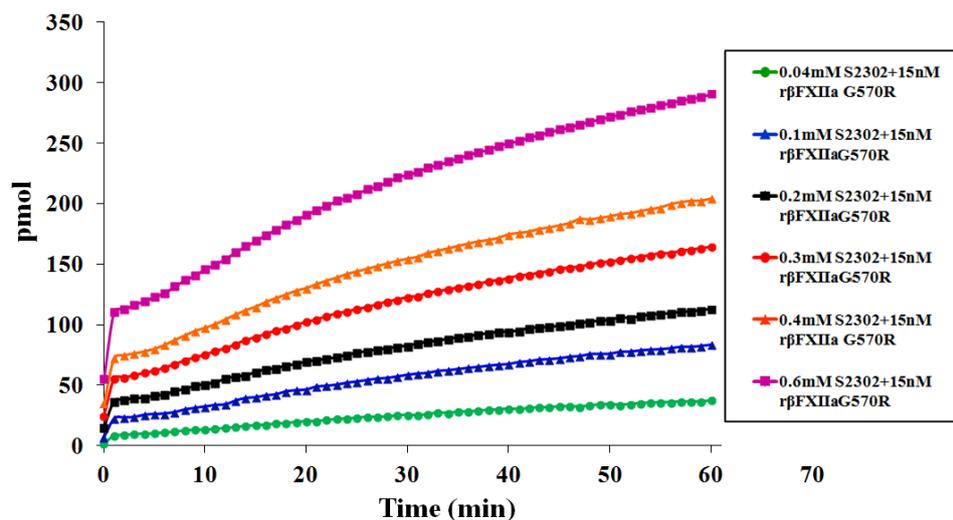
**Figure 5.8. Purification of  $\beta$ FXIIG570R by size exclusion chromatography.** **A)** UV trace from size exclusion chromatography using Superdex 75 16/60. The black box indicates the peak of interest. **B)** Non reducing SDS-PAGE gel analysis. Lane 1: marker; lanes 2-8: elution corresponding to the second peak in the chromatogram. The black box highlights the protein of interest.

### 5.3 Kinetic characterization of $\beta$ FXIIa and $\beta$ FXIIG570R

The catalytic activity of recombinant  $\beta$ FXIIa and mutant  $\beta$ FXIIG570R was determined by measuring the hydrolysis of chromogenic peptide substrate S2302. This was accomplished monitoring the change in absorbance at 405 nm. The initial rates were obtained from a time-dependent chromogenic substrate assay. 0.015  $\mu$ M of recombinant protein were incubated with different concentrations of S2302. A plot of the amount of produced p-nitroaniline versus time was obtained (cf. Figure 5.9 and Figure 5.10 for  $\beta$ FXIIa and  $\beta$ FXIIG570R, respectively).

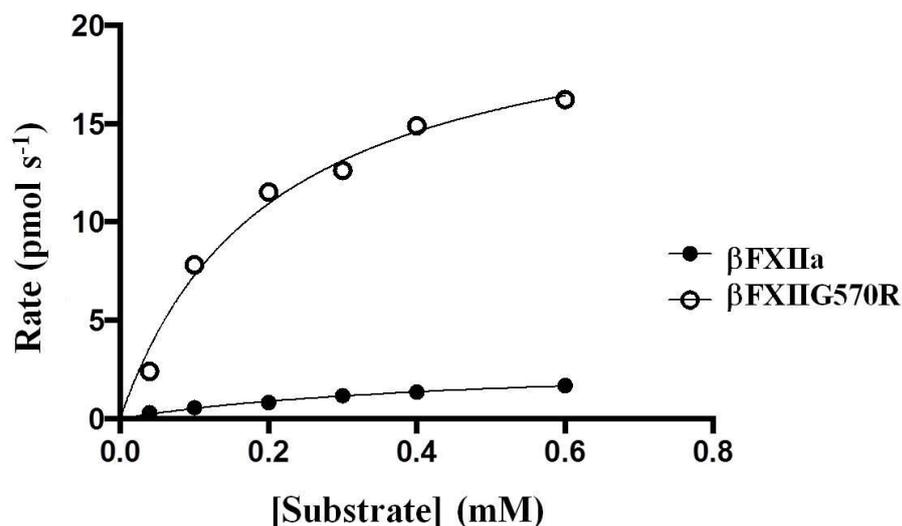


**Figure 5.9: S2302 time dependent assay for  $\beta$ FXIIa.** pmols of pNitroaniline generated as a function of time, after the addition of S2302 to a 15nM  $\beta$ FXIIa solution. Different concentrations of S2302 have been tested: 0.04 mM (green), 0.1 mM (blue), 0.2 mM (black), 0.3 mM (red), 0.4mM (orange), 0.6 mM (purple) For every experiment, each substrate concentration has been used in quadruplicates.



**Figure 5.10: S2302 time dependent assay for  $\beta$ FXIIG570R.** pmols of pNitroaniline generated as a function of time, after the addition of S2302 to a 15nM  $\beta$ FXIIG570R solution. Different concentrations of S2302 have been tested: 0.04 mM (green), 0.1 mM (blue), 0.2 mM (black), 0.3 mM (red), 0.4mM (orange), 0.6 mM (purple) For every experiment, each substrate concentration has been used in quadruplicates.

Kinetic parameters for  $\beta$ FXIIa and  $\beta$ FXIIG570R were obtained plotting the initial reaction rates as a function of the substrate concentration (cf. Figure 5.11). While  $\beta$ FXIIa obeys the Michelis-Menten kinetic, the same observation cannot be made for the mutant  $\beta$ FXIIG570R.



**Figure 5.11** Activity of  $\beta$ FXIIa and  $\beta$ FXIIG570R (pmol/sec) against various concentrations of S2302. An average of four experiments were used to determine the enzyme kinetics. For every experiment, each concentration of substrate has been used in quadruplicates.

In Table 5.1 the kinetic parameter measured for recombinant  $\beta$ FXIIa and  $\beta$ FXIIG570R mutant are reported.

**Table 5.1** Enzyme kinetic parameters for  $\beta$ FXIIa and  $\beta$ FXIIG570R.

Sample	Vmax (pmol/s)	$k_{cat}$ (s <sup>-1</sup> )	$K_M$ (mM)	$k_{cat}/K_M$ (Lmol <sup>-1</sup> s <sup>-1</sup> )
$\beta$ FXIIa	32.93±2.797	21.75±1.86	0.2±0.04	108750
$\beta$ FXIIG570R	9.37±1.5	6.25±1	0.62±0.16	10080

$K_M$  for recombinant  $\beta$ FXIIa is 0.2 mM, which is comparable to the  $K_M$  value of commercial  $\beta$ FXIIa purified from plasma (0.25 mM  $\pm$  0.05). At variance,  $K_M$  for  $\beta$ FXIIG570R is three times higher (0.62 mM  $\pm$  0.16). The  $k_{cat}$ , which represents the number of substrate molecules that each active site converts into product over time, is 21.75 s<sup>-1</sup> for  $\beta$ FXIIa, while it is 6.1 s<sup>-1</sup> for the

mutant. The catalytic efficiency, represented by  $k_{\text{cat}}/K_{\text{M}}$ , and used to compare the reactivity of two enzymes acting on the same substrate, shows that the efficiency of  $\beta\text{FXIIa}$  is one order of magnitude higher than the efficiency of  $\beta\text{FXIIG570R}$ .

#### **5.4 Crystallization of $\beta\text{FXIIa}$ and $\beta\text{FXIIG570R}$**

An extensive crystallization screening was carried out using the commercially available suite from Hampton Research and Molecular Dimension, in order to establish initial conditions to be used as a starting point.

The search for crystallization conditions for  $\beta\text{FXIIa}$  was carried out at two different concentrations, 20 mg/mL and 10 mg/mL. Only the  $\text{Am}_2\text{SO}_4$  suite gave tiny crystals in the conditions listed in Table 5.2:

**Table 5.2:** Crystallization conditions for  $\beta$ FXIIa from the AmSO<sub>4</sub> suite.

<b>Protein</b>	$\beta$ FXIIa
<b>Method</b>	Sitting drop
<b>Protein concentration</b>	10 mg/mL
<b>Reservoir</b>	1. 0.1 M Tris pH 8, 0.8 M Am <sub>2</sub> SO <sub>4</sub> 2. 0.1 M Hepes pH 7, 0.8 M Am <sub>2</sub> SO <sub>4</sub>
<b>Crystallization time</b>	3 days
<b>Crystallization temperature</b>	20 °C

The optimization of the initial hits was carried out changing the salt concentration from 0.2 M up to 2 M. The concentration of Hepes pH 7 and Tris pH 8 was varied between 0.05 M and 0.2 M. Moreover, the concentration of  $\beta$ FXIIa was lowered (5 mg/mL and 2.5 mg/mL). From this optimization procedure, only small crystals were obtained. They diffracted at a very low resolution (~8-9 Å) and it was not possible to index the diffraction data.

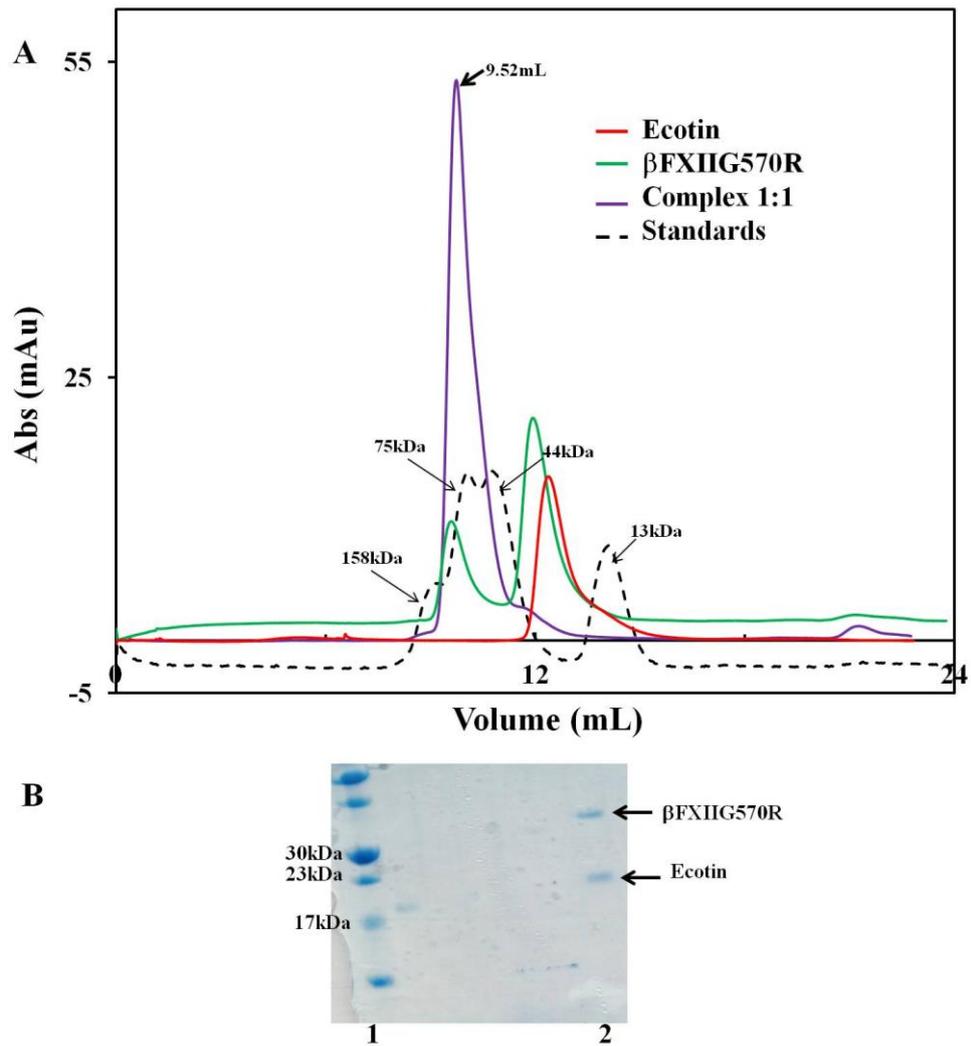
For  $\beta$ FXIIG570R instead the sparse grid was unsuccessful since no hits were obtained.

## **5.5 Co-Crystallization of $\beta$ FXIIa and $\beta$ FXIIG570R with CMK inhibitor PPACK**

For  $\beta$ FXIIa-PPACK and  $\beta$ FXIIG570R-PPACK, a wide search for crystallization conditions was carried out with commercially available crystallization suites. Unfortunately, no crystals were obtained.

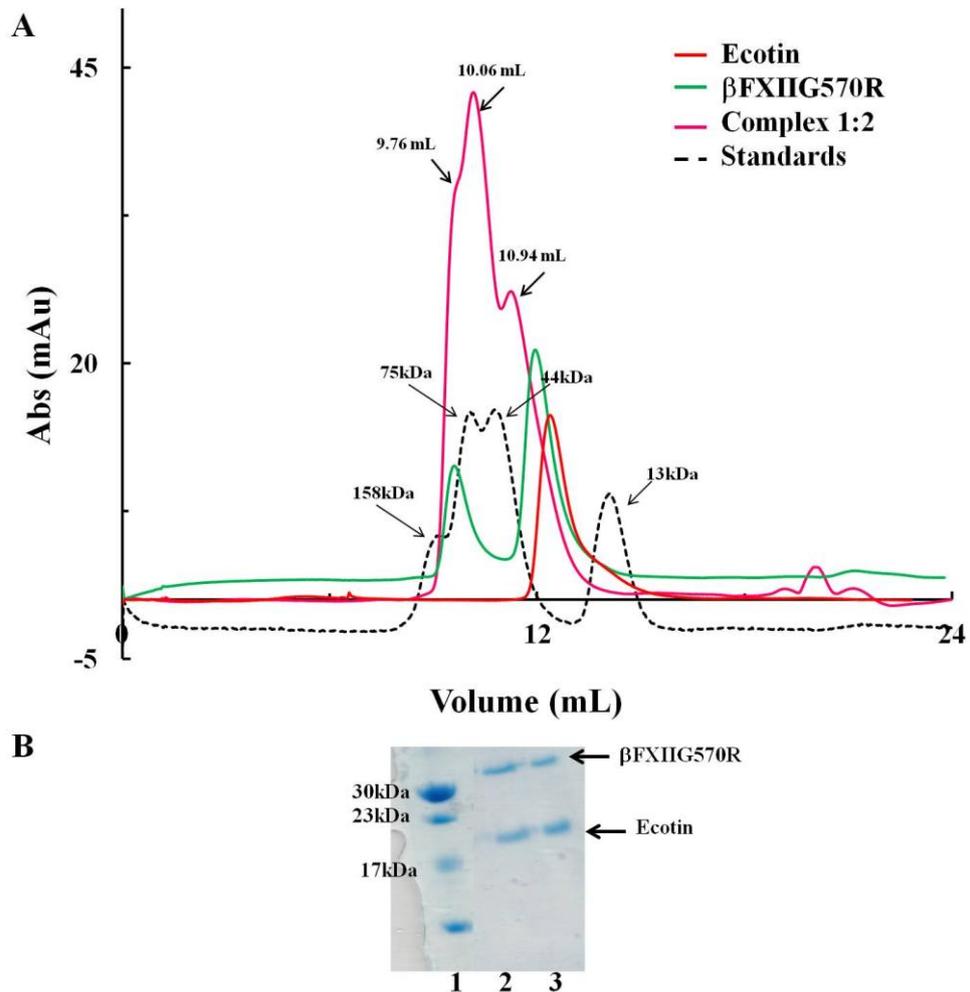
## **5.6 Complex of $\beta$ FXIIG570R with serine protease inhibitor Ecotin**

$\beta$ FXIIG570R-Ecotin complex formation was assessed by size exclusion chromatography. Ecotin was present in 1X and 2X molar excess over  $\beta$ FXIIG570R. The proteins were allowed to form a complex by incubation at 4 °C overnight. The complex was then resolved by gel filtration (CF. Figure 5.12 A and 5.13 A for 1:1 ratio and 1:2 ratio, respectively). Standards were run in order to better estimate the molecular mass of the formed complex. The chromatogram reported in figure 5.12 A shows a single sharp peak ( $V_e = 9.52$  mL) for the 1:1 complex corresponding to the tetramer (32 kDa for Ecotin dimer plus  $2 \times 30$  kDa for  $\beta$ FXIIG570R). A 15% SDS gel (Figure 5.12 B) was run, in order to verify the formation of the complex. The gel shows two bands, one for Ecotin (17 kDa, monomeric form) and one band at 30 kDa for  $\beta$ FXIIG570R.



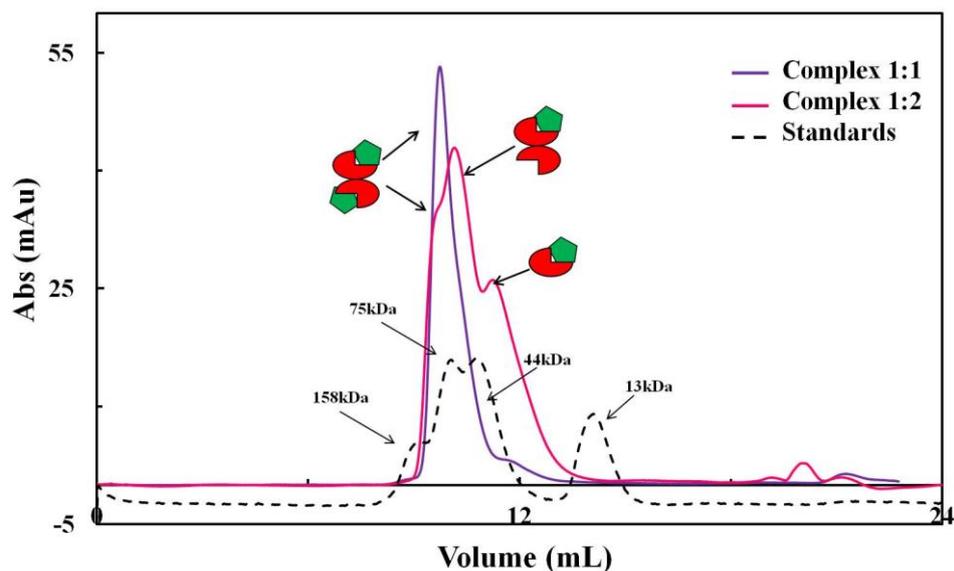
**Figure 5.12: Analytical gel filtration analysis of complex formation between  $\beta$ FXIIG570R and Ecotin in 1:1 ratio.** A) UV trace from analytical gel filtration using Superdex 75 10/300.  $\beta$ FXIIG570R-Ecotin (1:1): solid purple line;  $\beta$ FXIIG570R: solid green line; Ecotin: red solid line. The column was calibrated with aldolase ( $M_w$ : 158000 Da), conalbumin ( $M_w$ : 75000 Da), ovalbumin ( $M_w$ : 44000 Da) and RNase A ( $M_w$ : 13000 Da). The calibration curve is shown in dotted black line. The elution volume for the tetramer is also reported. B) Non reducing SDS-PAGE analysis. Lane 1: marker; lane 2: fraction containing  $\beta$ FXIIG570R-Ecotin (1:1).

For the 1:2 complex, the elution profile shows a broader peak. The analysis of the SDS gel (cf. Figure 5.13 B) indicates the presence of two bands corresponding to the monomeric form of each component, for each peak.



**Figure 5.13: Analytical gel filtration analysis of complex formation between  $\beta$ FXIIG570R and Ecotin in 1:2 ratio.** A) UV trace from analytical gel filtration using Superdex 75 10/300.  $\beta$ FXIIG570R-Ecotin (1:2): solid pink line;  $\beta$ FXIIG570R: solid green line; Ecotin: red solid line. The column was calibrated with aldolase ( $M_w$ : 158000 Da), conalbumin ( $M_w$ : 75000 Da), ovalbumin ( $M_w$ : 44000 Da) and RNase A ( $M_w$ : 13000 Da). The calibration curve is shown in dotted black line. The elution volumes for the peaks are also reported. B) Non reducing SDS-PAGE analysis. Lane 1: marker; lanes 2-3: fraction containing  $\beta$ FXIIG570R-Ecotin (1:2).

When an excess of Ecotin is added, a single peak ( $V_e = 9.76$  mL), corresponding to the tetramer, appears, as confirmed by superimposition of the two UV traces (cf. Figure 5.14). As a consequence of the excess of added Ecotin, a second peak of apparent molecular weight of 60 kDa becomes visible, corresponding to the trimer ( $V_e = 10.06$  mL). The chromatogram also shows a third peak, which could be assigned to the heterodimer ( $V_e = 10.94$  mL).



**Figure 5.14: Superimposition of gel filtration chromatograms for  $\beta$ FXIIG570R and Ecotin in 1:1 and 1:2 ratio.**  $\beta$ FXIIG570R-Ecotin (1:1): solid purple line.  $\beta$ FXIIG570R-Ecotin (1:2): solid pink line. The column was calibrated with Aldolase ( $M_w$ : 158000 Da), Conalbumin ( $M_w$ : 75000 Da), Ovalbumin ( $M_w$ : 44000 Da) and RNase A ( $M_w$ : 13000 Da). The calibration curve is showed in dotted black line. The figure also shows a schematic representation of the multimeric forms, with the Ecotin represented by a red circle and  $\beta$ FXIIG570R as a green pentagon.

The fractions of interest were concentrated to 1.7 mg/ml and 3.6 mg/mL for 1:1 complex and 1:2 complex, respectively.

### 5.7 Crystallization of $\beta$ FXIIG570R-Ecotin complexes

$\beta$ FXIIG570R in complex with Ecotin was crystallized using the sparse matrix from Hampton Research and Molecular Dimensions. Hits were obtained only for the (1:2) complex from the ProComplex screen in the conditions listed in Table 5.3:

**Table 5.3:** Crystallization conditions for  $\beta$ FXIIG570R from the ProComplex suite.

<b>Protein</b>	$\beta$ FXIIG570R-Ecotin (1:2)
<b>Method</b>	Sitting drop
<b>Protein concentration</b>	3.6 mg/mL
<b>Reservoir</b>	0.1 M MgAc <sub>2</sub> ; 0.1 M NaAc pH 4.5; 8% PEG8000
<b>Crystallization time</b>	1-3 days
<b>Crystallization temperature</b>	20 °C

The crystals were tested for diffraction at the Diamond Synchrotron but they gave only weak diffraction spots. An optimization of the conditions listed in Table 5.3 was performed, but no crystals were obtained.

## 5.8 Summary

In this chapter we discussed the results obtained for  $\beta$ FXIIa and  $\beta$ FXIIG570R mutant. The cloning of  $\beta$ FXIIa in pMT-PURO vector and the subsequent transfection in S2 cells, followed by its expression by inducing with CuSO<sub>4</sub>, was successful as confirmed by the Western Blot. The cloning and expression of  $\beta$ FXIIG570R mutant in DES was performed by Dr Monika Pathak.

Purification protocols for  $\beta$ FXIIa and  $\beta$ FXIIG570R were established. They both required a three step purification; the obtained protein was 95% pure,

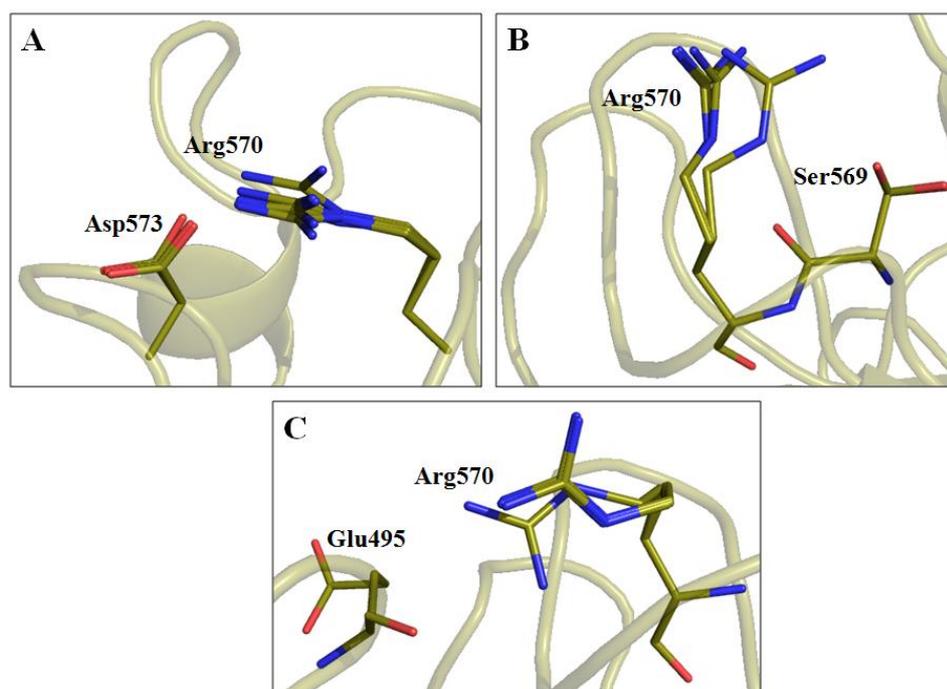
as confirmed by SDS-PAGE gel analysis, although the total yield was low (2.4 mg for  $\beta$ FXIIa and 1.8 mg for G570R from 3L of media for each construct).

Kinetic characterization of  $\beta$ FXIIa and  $\beta$ FXIIG570R was performed, monitoring the change in absorbance at 405 nm associated with the cleavage of the chromogenic substrate S2302. The experiments were performed with the assistance of Kareem Hamad Badraddin, PhD student in the group of Dr. L. Dekker, and he also performed the analysis of the data. The hydrolysis of S2302 mediated by  $\beta$ FXIIa and  $\beta$ FXIIG570R shows that the amount of p-nitroaniline, generated after the cleavage of chromogenic substrate S2302, gradually increases with the substrate concentrations, until the steady-state of the reaction is reached, after 60 min. Moreover, the S2302 time-dependent assay shows that the efficiency of  $\beta$ FXIIa (from  $k_{cat}/K_M$ ) is around one order of magnitude higher than the value calculated for the mutant. In fact, kinetic parameters were calculated from the initial rates against the substrate concentrations, and fitted with the non-linear regression Michaelis-Menten algorithm. They indicated that the catalytic behaviour of recombinant  $\beta$ FXIIa is almost identical to commercial  $\beta$ FXIIa purified from plasma. Therefore, the protein expressed and purified from insect cell system is catalytically competent. At variance,  $\beta$ FXIIG570R mutant is less efficient than the wild type.

In order to understand how a single mutation can reduce the catalytic activity of the enzyme, several side chain conformations of the mutated residue have been obtained using FoldX[268] software. FoldX calculates a series of possible rotamers for the mutated residue and for the residues in a

radius of 6 Å, in order to find a minimum in the energy landscape of the protein.

Figure 5.15 shows the most stable conformations of Arg570 in the mutant protein.

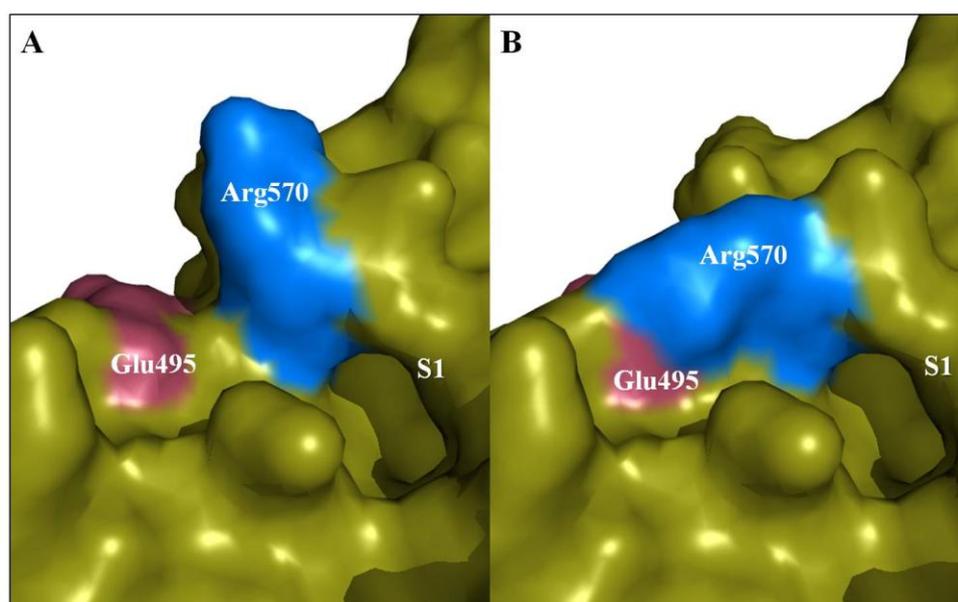


**Figure 5.15: Effects of G570R mutation.** Stick representation of the main interactions of mutated Arg570 with Asp573(panel A), Ser569 (panel B), and Glu495 (panel C).

The conformations obtained by FoldX suggest possible interactions between the mutated Arg570 and Asp573, Ser569, Glu495. According to these interactions, it is possible to group the energy-minimized structures in three clusters shown in Fig. 5.15. In Figure 5.15 A, stabilizing interactions are found (i) between the NH Arg570 (guanidinium group) and O $\delta$  of Asp573, and (ii) between N $\epsilon$  of Arg570 and O $\delta$  of Asp573. For the conformation shown in Figure 5.15 B, stabilizing interactions are found (i) between O $\gamma$  of Ser569 and NH (guanidinium group) of Arg570, (ii) between

O $\gamma$  of Ser569 and N $\epsilon$  of Arg570, (iii) between O of Ser569 and NH (guanidinium group) of Arg570, and (iv) between O of Ser569 and N $\epsilon$  of Arg570. The conformation assumed by Arg570 in Figure 5.15 C, is favoured by the interaction of the NH (guanidinium group) of Arg570 with the O of Glu595.

Interestingly, the interaction with Glu495 could in principle explain the low catalytic activity of the mutant if compared to the wild type. In fact, the Arg570-Glu495 interaction completely blocks the access to the active site cleft, as shown in Figure 5.16.



**Figure 5.16: Surface interaction of Arg570 with Glu495.** A) Surface representation of the protein in absence of the interaction between Arg570 and Glu495. The access to the S1 pocket is open. B) Surface representation of the interaction between Arg570 and Glu495. The interaction completely closes the entrance of the S1 pocket. In both panels, the protein surface is shown in deep olive colour, Arg570 is in blue and Glu495 in violet.

This empirical study suggests that the Arg570-Glu495 interaction could be a strong candidate for lowering  $\beta$ FXIIa activity. Nevertheless, molecular dynamic studies and crystallographic data are necessary to confirm this hypothesis.

In order to further characterize the structural behaviour of  $\beta$ FXIIa and its mutant, crystallization trials were sparse screened for both proteins, using commercial screens.  $\beta$ FXIIa gave initial hits, that were subsequently used for optimization. Obtained crystals were reproducible but diffracted weakly and the diffraction data were not indexable. None of the performed optimizations provided suitable crystals. In the case of the mutant, no hits were obtained. This may be due to the low concentration tested (only 3.6 mg/ml). However, it was not possible to obtain higher concentrated protein, because of the poor yields from purification.

The complex with the small covalent inhibitor PPACK, was not successful. Most of the conditions tested through the commercial screens were clear, without any precipitate.

Ecotin was the protein of choice for structural studies with macromolecular inhibitors. Deposited structures of other coagulation factors (thrombin[185] and FXIa[184]) in complex with Ecotin are available. The complex was resolved by gel filtration, following a procedure already reported in the literature.[184, 185] The formation of the complex was successful only for the mutant, and different trays were set up. Crystals were obtained in different conditions, and tested at the Diamond Light Source. However, they were poorly diffracting. An optimization around these conditions was performed but no crystals were obtained.

In the future, a new batch of  $\beta$ FXIIa will be purified and the complex with different inhibitors will be made at higher concentration, in order to verify if this can aid the production of higher quality crystals.

The knowledge of the structure of both  $\beta$ FXIIa and  $\beta$ FXIIG570R may help to understand the mechanisms behind the catalytic activity of this enzyme and to get insights about the role played by the Arg570-Glu495 interaction, which has been found from empirical studies.

Moreover, the structure of  $\beta$ FXIIa or  $\beta$ FXIIG570R in complex with the PPACK or Ecotin will allow clarifying the interaction mode of the binding pocket with the substrate. This knowledge can be used as a starting point for rational drug design.

## CHAPTER 6

### EXPRESSION AND PURIFICATION OF $\beta$ FXII FROM *E. coli*

#### 6.1 Expression of $\beta$ FXII and purification

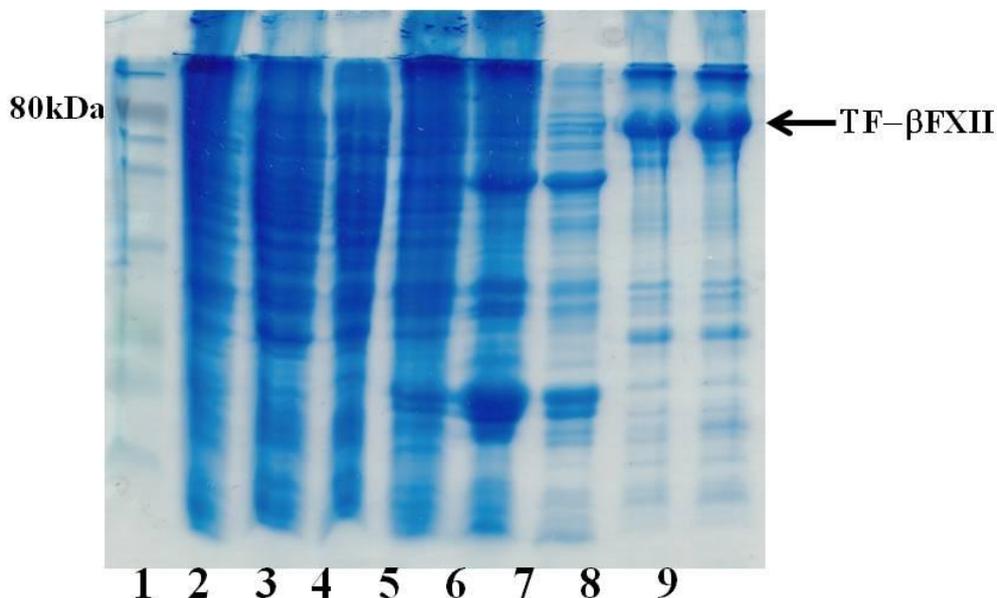
The cloning of  $\beta$ FXII construct in pCold-TF vector was ordered from GenScript. In order to increase cell density for bacterial expression and thus enhance the yield of recombinant protein, a great effort has been put in the optimization of both expression and purification protocols.

The cells were induced when the OD<sub>600</sub> was 0.3-0.8 at different IPTG concentrations: 0.2 mM, 0.5 mM and 0.8 mM. The best results were obtained inducing the cells with 0.8 mM IPTG when the OD<sub>600</sub> was 0.6-0.8.

The length of expression was also optimized. The protocol for pCold-TF vector suggested incubating for 24 h. However, it was noticed that incubating for a longer period (48 h and 72 h) resulted in a higher yield of recombinant protein. The temperature was varied between 15 °C and 10 °C; in this case, lowering the temperature down to 10 °C did not make any improvements in the expression levels. Nevertheless, the expression of other contaminants was reduced.

Cells were harvested by centrifugation and lysed by sonication in 50 mM Tris HCl pH 8.0, 200 mM NaCl, 15 mM imidazole. The lysate was further clarified by centrifugation and then purified by affinity chromatography. The elution was performed by linearly increasing the concentration of imidazole (0% to 100% buffer B) over 16 column volume. A non-reducing SDS-PAGE gel (Figure 6.1) allowed the identification of the fractions of

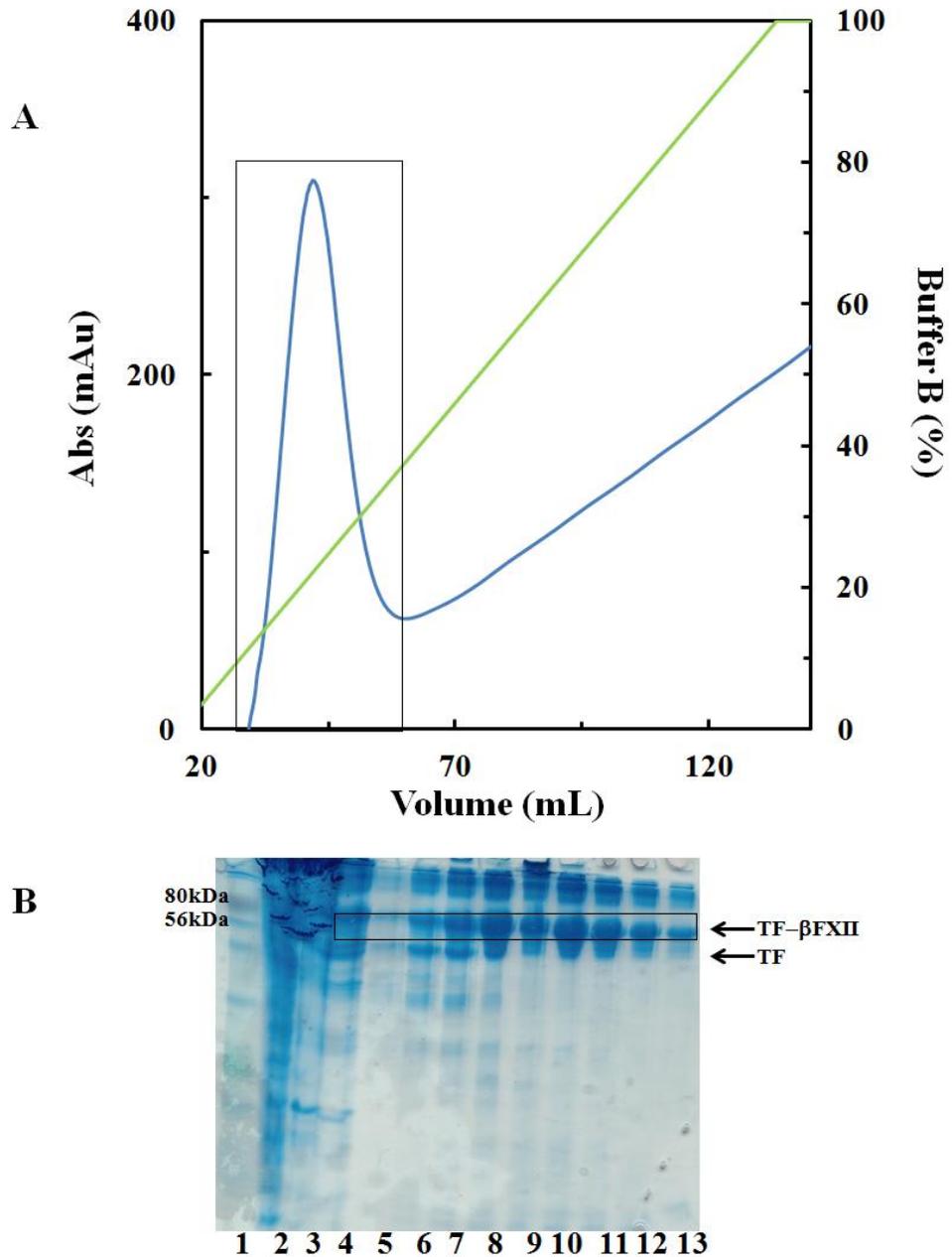
interest. The latter were concentrated down to 3 mL for gel filtration and left at 4 °C overnight.



**Figure 6.1: Non reducing SDS-PAGE gel analysis of fractions from affinity chromatography.** Lane 1: marker; lane2: clarified lysate; lane 3: flow through; lane 4-7: wash; lanes 8-9: elution.

The UV trace from size exclusion chromatography (Figure 6.1 A, solid purple line) showed a peak coming out in the void volume, indicating the aggregation of the protein.

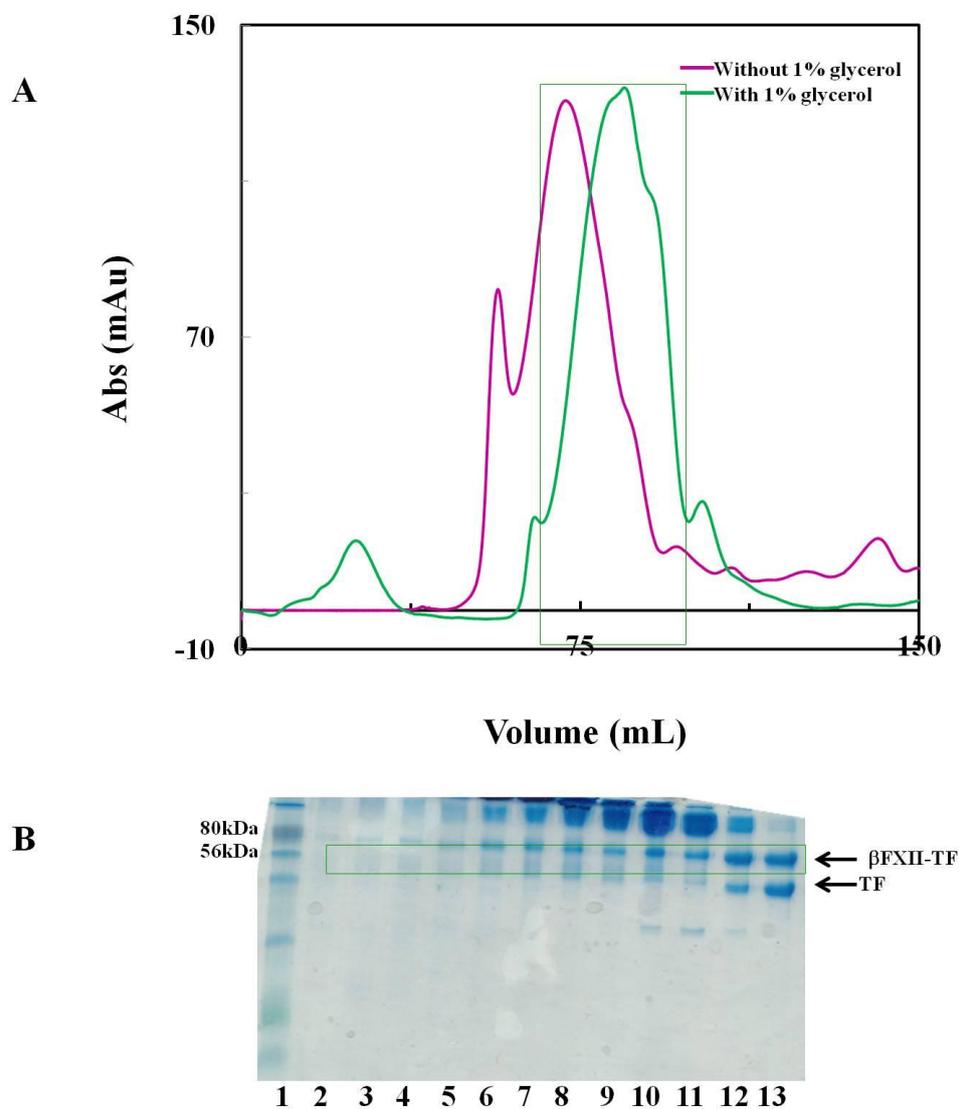
For the subsequent purification of βFXII, the expression temperature was lowered down to 10 °C. The SDS-PAGE gel from affinity chromatography step (Figure 6.2 B) indicated that fractions were cleaner, even if the presence of untagged TF in high concentrations can be noticed.



**Figure 6.2: Purification of TF-βFXII by affinity chromatography. A)** UV trace from affinity chromatography using nickel column. The black box indicates the peak of interest. **B)** Non reducing SDS-PAGE analysis. Lane 1: marker; lanes 2-3: wash; lanes 4-13: elution. The black box indicates the protein of interest.

To avoid aggregation problems, 1% glycerol was added in the gel filtration buffer and in the fractions purified by affinity chromatography. The

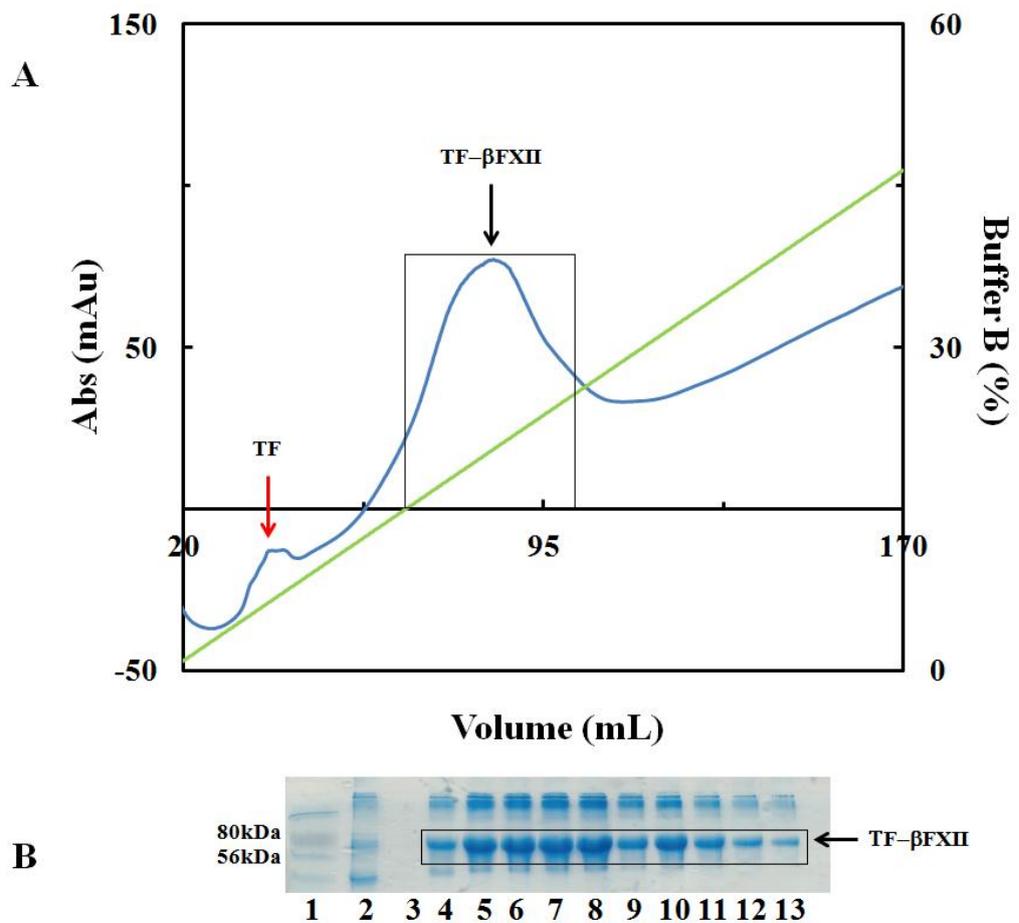
presence of glycerol helped, since the protein came out at the expected volume (Figure 6.3 A).



**Figure 6.3: Purification of TF-βFXII by size exclusion chromatography.** A) UV trace from gel filtration using Superdex 200 16/60. βFXII without 1% glycerol: solid purple line; βFXII with 1% glycerol: solid green line. The green box indicates the peak of interest when 1% glycerol is added. B) Non-reducing SDS-PAGE gel from gel filtration run in presence of 1% glycerol in the buffer. Lane 1: marker; lanes 2-13: elution. The green box indicates the protein of interest (gel filtration with 1% glycerol).

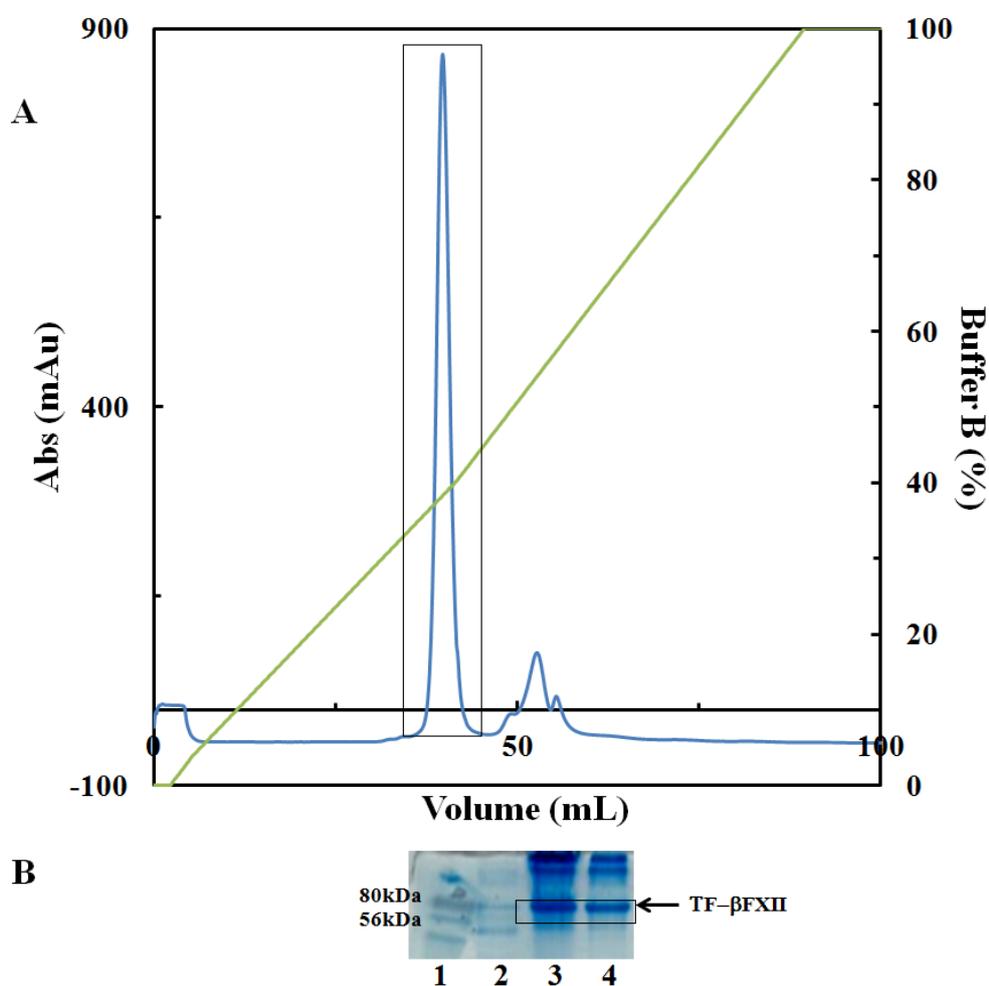
From the SDS-PAGE gel (Figure 6.3 B), the protein was 80% pure even if it was not possible to separate untagged TF from recombinant protein. The gel also shows some high molecular weight bands, that could be ascribed to a gel artifact or to aggregation.

For these two reasons, the purification protocol was further optimized by adding 1% glycerol[269, 270] in Ni(II) buffers and increasing linearly the concentration of buffer B from 0 to 60% over 40 column volume. This was performed to stabilize the protein, avoiding aggregation, and to separate TF from the protein of interest. The UV trace and SDS-PAGE gel for this step are shown in Figure 6.4.



**Figure 6.4: Purification of TF- $\beta$ FXII by affinity chromatography.** A) UV trace from affinity chromatography using nickel column. The black box indicates the peak of interest B) Non-reducing SDS-PAGE gel analysis. Lane 1: marker; lanes 2-13: second peak elution. The black box highlights the protein of interest.

The slow gradient helped to get rid of most of untagged TF, even if the gel still shows some high molecular weigh bands. For this reason, before the gel filtration, an anion exchange chromatography was performed (cf. Figure 6.5).



**Figure 6.5: Purification of TF- $\beta$ FXII by anion exchange chromatography.** A) UV trace from anion exchange chromatography using capto Q column. The black box indicates the peak of interest. B) Non-reducing SDS-PAGE gel analysis. Lane 1: marker; lanes 2-4: elution. The black box highlights the protein of interest.

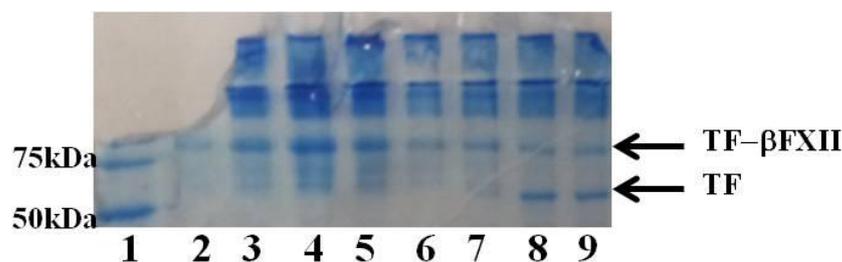
From the gel, it is still possible to observe high molecular weight bands. Before the final purification step, a thrombin cleavage was performed in order to remove TF.

## 6.2 Cleaving the Trigger Factor tag using HRV-3C protease, FXa and Thrombin

In order to find which was the best enzyme cleaving the Trigger Factor among HRV-3C protease, FXa and thrombin, a small-scale reaction (10 - 25  $\mu$ l) was set, following the protocol suggested by each supplier.

For HRV-3C protease, eight different enzyme (U) : substrate ( $\mu$ g) ratios were used: 1:50, 1:100; 1:200, 1:400, 2:50, 2:100; 2:200, 2:400.

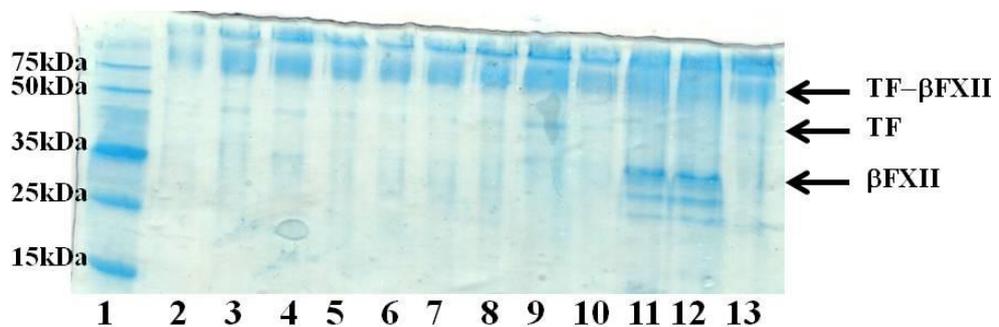
The reaction was performed at room temperature over night, but since the cleavage did not occur (cf. Figure 6.6), the experiment was repeated at 37 °C. However, also in this case, from the resulting gel (Figure 6.7),  $\beta$ FXII-TF was almost uncleaved.



**Figure 6.6: HRV-3C protease cleavage.** Cleavage reaction on small-scale to determine the optimal conditions for the cleavage of Trigger Factor. All

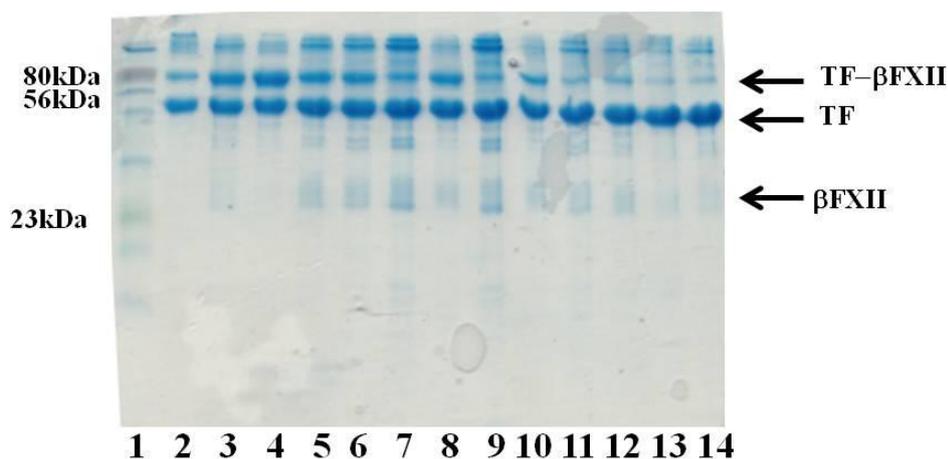
the reactions were performed overnight at room temperature. Lane 1: Marker; lanes 2-9: HRV-3C protease cleavage using different enzyme (U): substrate ( $\mu\text{g}$ ) ratios (lane 2: 1:50; lane 3: 1:100; lane 4: 1:200; lane 5: 1:400; lane 6: 2:50; lane 7: 2:100; lane 8: 2:200; lane 9: 2:400).

The cleavage reaction was performed with FXa, testing four different enzyme ( $\mu\text{g}$ ) : substrate ( $\mu\text{g}$ ) ratios at 37 °C over night. The SDS-PAGE gel analysis (cf. Figure 6.7) shows that the cleavage was not complete. In fact, it is possible to see a band ~30 kDa that could be ascribed to cleaved  $\beta\text{FXII}$ . However, there are also degradation products.



**Figure 6.7: HRV-3C protease and FXa cleavage.** Cleavage reaction on a small scale to determine the optimal conditions for the removal of Trigger Factor. All the reactions were performed overnight at 37 °C. Lane 1: Marker; lanes 2-9: HRV-3C protease cleavage using different enzyme (U) : substrate ( $\mu\text{g}$ ) ratios (lane 2: 1:50; lane 3: 1:100; lane 4: 1:200; lane 5: 1:400; lane 6: 2:50; lane 7: 2:100; lane 8: 2:200; lane 9: 2:400); lanes 10-13: FXa cleavage using different enzyme ( $\mu\text{g}$ ) : substrate ( $\mu\text{g}$ ) ratios ( lane 10: 1:50; lane 11: 1:100; lane 12: 1:200; lane 13: 1:400).

A thrombin cleavage was performed on a sample of 10  $\mu\text{l}$  to determine the optimal conditions. Two different ratios, 1:100 and 1:200 wt:wt of thrombin to the target protein, were tested at room temperature and at different times. The SDS-PAGE gel was run in non-reducing conditions (cf. Figure 6.8).



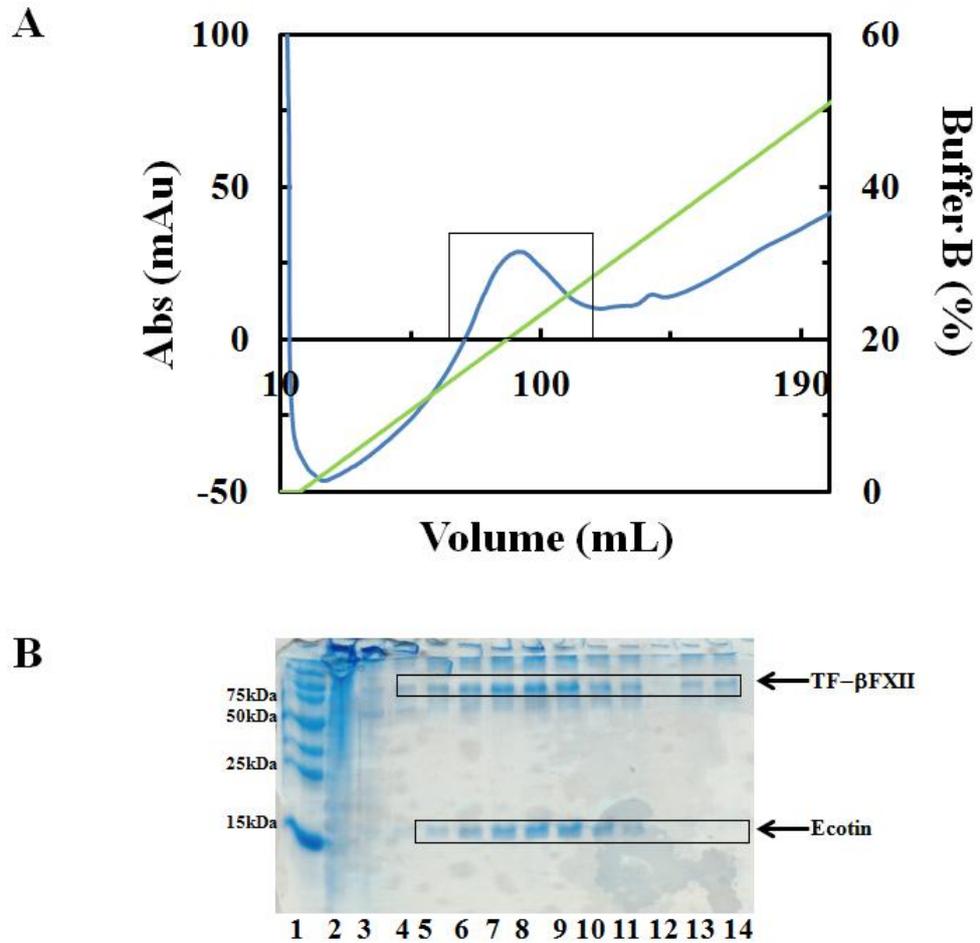
**Figure 6.8: Thrombin cleavage.** A time course was performed to establish the optimal conditions for the cleavage of TF. Lane 1: Marker; lane 2  $\beta$ FXII-TF control; lane 3: 1:100 after 10' digestion; lane 4: 1:200 after 10' digestion; lane 5: 1:100 after 30' digestion; lane 6: 1:200 after 30' digestion; lane 7: 1:100 after 1h digestion; lane 8: 1:200 after 1h digestion; lane 9: 1:100 after 2h digestion; lane 10: 1:200 after 2h digestion; lane 11: 1:100 after 4h digestion; lane 12: 1:200 after 4h digestion; lane 13: 1:100 after ON digestion; lane 14: 1:200 after ON digestion.

The gel showed that TF was mostly cleaved over night, even if the band for  $\beta$ FXII is not visible. The optimal result is achieved with 1:200 wt:wt ratio and leaving the reaction for 2h at room temperature. In fact, even the cleavage is not complete, the band for  $\beta$ FXII is very clear.

The reaction was scaled up proportionally and the two different reaction conditions were tested: 1:200 after 2h digestion and 1:200 over night. The protein sample was purified from the cleaved TF and thrombin using affinity chromatography and a benzamidine column, respectively.  $\beta$ FXII was supposed to come out in the flow through of the affinity chromatography column but no band was visible on the SDS-PAGE gel.

### 6.3 Co-purification

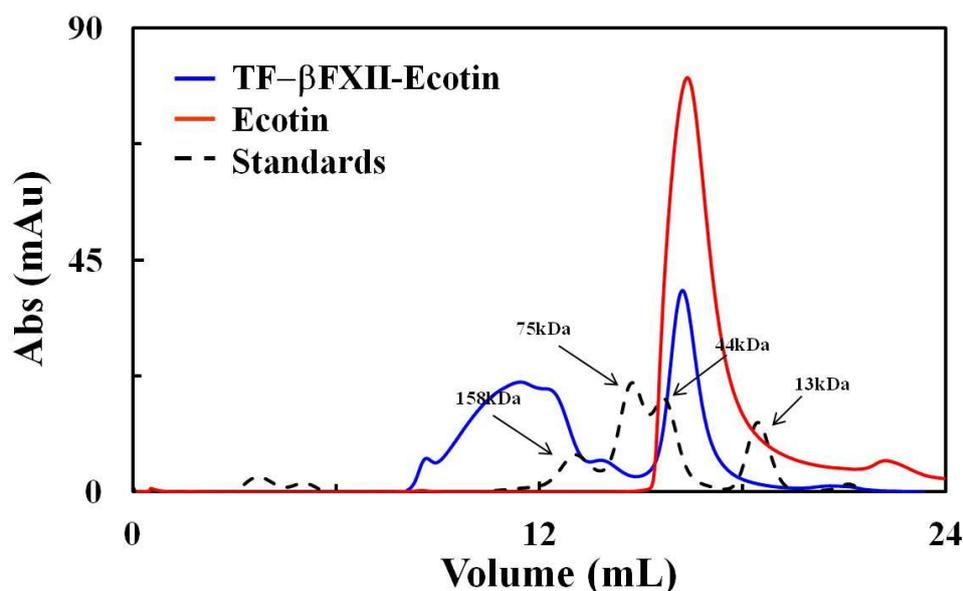
Since the cleavage of TF did not produce pure recombinant  $\beta$ FXII from *E. Coli*, a different strategy was used. It has been reported that immobilized Ecotin on a resin can be used for the purification of trypsinogen as well as other proteases as zymogens.[271] Ecotin and  $\beta$ FXII were expressed separately according to their expression protocol. Before the sonication, the pellets of the two proteins were mixed together, by suspending them in 50 mM Tris pH 8, 200 mM NaCl, 15 mM imidazole and 1% glycerol. The pellet was sonicated and the lysate clarified for 1 h at 15000 g. The supernatant was then loaded onto Ni(II) column for a first affinity purification step (cf. Figure 6.9 A). The elution was performed by linear gradient, increasing the concentration of imidazole (0 % to 60 % buffer B) over 40 column volume. A non-reducing SDS-PAGE gel (cf. Figure 6.9 B) was run to identify the fractions containing the complex.



**Figure 6.9: Co-purification of TF- $\beta$ FXII and Ecotin by affinity chromatography.** A) UV trace from affinity chromatography using nickel column. The black box indicates the peak of interest. B) Non reducing SDS-PAGE analysis. Lane 1: marker; lanes 2-3: wash; lanes 4-14: elution. The black boxes highlight the proteins of interest.

The fractions from the affinity chromatography were concentrated for the gel filtration step, performed on a Superdex 200 10/300 column. The elution profile is reported in Figure 6.10, along with the elution profiles for Ecotin and for the standards. The UV trace for the co-purified proteins shows two peaks; the first peak could correspond to the tetrameric complex, having an apparent molecular weight of 200 kDa, and it is clearly separated from

Ecotin (second peak). However, the SDS-PAGE gel analysis (gel not shown) shows that the first peak correspond to a multimeric form of  $\beta$ FXII-TF.



**Figure 6.10: Co-purification of TF- $\beta$ FXII and Ecotin by size exclusion chromatography.** UV trace from gel filtration using S200 10/300 column. TF- $\beta$ FXII-Ecotin: solid blue line; Ecotin: red solid line. The column was calibrated with aldolase ( $M_w$ : 158000 Da), conalbumin ( $M_w$ : 75000 Da), ovalbumin ( $M_w$ : 44000 Da) and RNase A ( $M_w$ : 13000 Da). The calibration curve is shown in dotted black line.

#### 6.4 Summary

In this chapter, we described the results obtained for the expression and purification of  $\beta$ FXII in *E. Coli*. pCold-TF was the vector of choice for the cloning of  $\beta$ FXII in bacterial system, because it combines the advantages deriving from the presence of the TF chaperone (helping the expression and enhancing the solubility of difficult proteins), and the fusion cold shock

expression technology. The expression in Origami 2 cells, as well as the purification, required huge efforts, in order to obtain pure protein. For the expression protocol, the best results were achieved by inducing the cells when  $OD_{600} = 0.8$  with 0.8 mM IPTG and then leaving the cell at 10 °C over the weekend, as working at low temperature may help the correct folding of the protein. The purification protocol was further optimized. The purification of  $\beta$ FXII from bacterial cells required two purification steps: affinity chromatography and anion exchange chromatography. The affinity chromatography was performed on a Ni(II) column and the elution profile was obtained by gradient increasing the concentration of buffer B over 40 column volume. This allowed to separate TF from the protein of interest. The presence of glycerol in every step, from sonication to the latest stage of purification, helped preventing the aggregation of the protein. The subsequent purification step, anion exchange chromatography, granted the production of pure protein, as confirmed by the SDS-PAGE gel analysis. Cleavage of the tag was performed before the size exclusion chromatography. Three different enzymes were tested for the cleavage. The best results were obtained using thrombin and setting the reaction overnight at room temperature. However, it was not possible to continue the process, since the protein was lost after the cleavage. This could be due to a poor stability of the protein. Therefore, a different strategy was pursued.  $\beta$ FXII was co-purified with Ecotin. The presence of the inhibitor was thought to stabilize the protein during the cleavage. However, gel filtration revealed that the complex was not formed.

In the future, it may be helpful to use a different strain, like shuffle cells that are known to help the folding of protein with disulfide bridges. Moreover, the  $\beta$ FXII plasmid could be cloned in a different vector, with no tag, or using a different tag, such as MBP: in this case, cleavage is not necessary, since MBP is known to act as a chaperone and to help the crystallization of the protein.

The possibility to obtain recombinant  $\beta$ FXII from bacterial cells is advantageous, mainly for two reasons: (i) *E. coli* is a good expression system because is fast and cheap and, hence, different strategies can be tested and the best protocol to obtain pure protein can be achieved; (ii) working with bacterial cells is useful for crystallization purposes, because there is no production of post-translational modifications, like glycosylation, that can be problematic for the crystallization.

## CHAPTER 7

### CONCLUDING REMARKS AND PERSPECTIVE

The main aim of this PhD thesis was the structural characterization of the protease domain of coagulation factor FXIIa. Since an inhibition of FXIIa results in a reduction of thrombus formation without compromising physiological hemostasis, it represents a new and selective strategy for preventing stroke and other thromboembolic diseases.[196] In this framework the comprehension of the mechanism behind the substrate-inhibitor recognition is fundamental and its molecular basis can be only addressed by crystallography.

In order to fulfil this goal, different constructs were made: (i) a construct of FXII starting from residue N335 for the expression in two different systems: *Drosophila Schneider* and *E. Coli*; (ii) an additional construct characterized by the presence of the carrier protein MBP at the N-terminus, for expression in insect cell system; (iii) the mutant  $\beta$ FXIIG570R for expression in insect cells.

MBP- $\beta$ FXIIa was successfully expressed in *Drosophila Schneider* system and a purification protocol was established. To characterize the catalytic activity of MBP- $\beta$ FXIIa, the hydrolysis of chromogenic substrate S2302 was measured and a comparison with commercial  $\alpha$ FXIIa and  $\beta$ FXIIa was performed. The MBP- $\beta$ FXIIa protein catalyzed the cleavage of S2302 substrate. The kinetic parameters ( $k_{cat}$ ,  $K_M$ ,  $k_{cat}/K_M$ ,  $V_{max}$ ) indicated that the efficiency of recombinant MBP- $\beta$ FXIIa is almost identical to that pertinent to commercial  $\beta$ FXIIa

purified from plasma, thus meaning that the protein expressed and purified from insect cell system is catalytically competent.

In order to characterize the structural behaviour of MBP- $\beta$ FXIIa, several crystallization conditions were tested with commercial screens. However, its search was not successful. MBP- $\beta$ FXIIa was also co-crystallized with different inhibitors, as crystal structure with a non specific inhibitor would aid elucidating the binding mode, enabling the rational design of new inhibitors with improved efficacy and selectivity. PPACK, PCK, Corn Trypsin Inhibitor (CTI) and Ecotin were the inhibitors of choice.

Crystals were obtained only for the complex with PPACK. The initial hits were used as a starting point for an extensive optimization. Crystals suitable for diffraction experiments were grown in 0.1 M sodium citrate pH 5.6 with 10% PEG 4000 and 0.15 M  $MgCl_2$  at 20 °C. The crystals diffracted to 4 Å of resolution and a complete data set was collected at Diamond Light Source in Oxford. The crystals are tetragonal, belong to the  $P4_12_12$  space group and there are two copies of MBP- $\beta$ FXIIa in the asymmetric unit.

The initial phases were solved by molecular replacement. A preliminary analysis of the crystal packing interactions revealed the presence of interactions between symmetry related molecules. In particular, we observe (i) interactions between MBP and  $\beta$ FXIIa and (ii) interactions between different  $\beta$ FXIIa molecules. The absence of interactions between different MBP molecules is somewhat surprising. In fact, these packing interactions were discussed in a previous crystallization study of proteins with MBP as fusion tag.[244] The lack of these lattice interactions for the MBP- $\beta$ FXIIa crystal could explain the low quality of the crystals and also the difficulties

in reproducing them. In this context, possible strategies to solve this problem include the attachment of the fusion tag to the C-terminus or the usage of a shorter linker. Moreover, it has been observed that the crystallized protein is glycosylated at Asn74, due to a post-translational modification which occurred in eukaryotic systems, thus affecting the crystallization.

The crystal structure of  $\beta$ FXIIa is similar to other serine proteases with two interacting six-stranded  $\beta$ -barrel domains with the catalytic triad (His57, Asp102, Ser195) located at the junction between them.

The analysis of the main structural features revealed that the recombinant protein is in the active conformation. In fact, MBP- $\beta$ FXIIa shows the conformational shape of the active protease, characterized by a well defined negatively charged S1 pocket. In particular, the 140-loop is positioned to interact with the 180-loop, a motif common to active proteases. Moreover, the conformation of Asp194 is similar to the conformation of the same residue in other active proteases, forming a salt bridge with N-terminal Val16. The position of the Val16 is comparable to the one of other active proteases. In particular, the activation loop is structured in such a way to form a  $\beta$ -hairpin. The stabilizing interactions with the HCR are fundamental to form the hairpin, thus providing a correct position to the N-terminal Val16. In the MBP- $\beta$ FXIIa structure, it was not possible to visualize the HCR, because of the low resolution. Nevertheless, the conformation of the activation loop is similar to that of HGFA, which is found to be the closest homolog of  $\beta$ FXIIa (243 equivalent residues, r.m.s.d.= 1.3 Å, 45% amino acid identity), somewhat confirming that the interactions between the HCR

and the activation loop are responsible for the correct positioning of the N-terminal Val16 inside the protease core. An interaction which also seems to play an important role in the conformation of the activation loop is a salt bridge between Arg9 of  $\beta$ FXIIa and Asn349 of MBP. Another distinctive feature of the protein is the so-called H1 pocket which assumes a closed conformation in  $\beta$ FXIIa. Finally, the presence of the oxyanion hole (not formed in the zymogen), constituted by the backbone nitrogen atoms of Gly193 and Ser195, represents a further confirmation that the crystallized protein is in the active form.

MBP- $\beta$ FXIIa crystallized with the PPACK inhibitor. Therefore, it is of interest to analyze the conformation of the inhibitor and its interactions with the protein. In particular, we observed an interaction of the Phe residue of PPACK with residues belonging to a symmetry related  $\beta$ FXIIa molecule. Therefore, the possibility that PPACK could induce a dimerization of the protein was investigated. Gel filtration experiments showed that dimerization is only an artifact of the crystal.

It is of interest to compare our results with the current literature, in particular, (i) the zymogen of FXII (FXIIc and FXIIac), whose crystallographic structure has been recently solved in the group of Professor Jonas Emsley,[144] and (ii) thrombin, which is the most studied protein of the coagulation cascade and the main therapeutic target of antithrombotic drugs. [90, 172]

Superimposition of  $\beta$ FXIIa with FXIIc (FXIIac) shows 241 (231) equivalent residues, an overall r.m.s.d. of 1.9 Å (1.8 Å) and 97% (95%) amino acid identity. There are four key differences in the structural features of  $\beta$ FXIIa

and the zymogen structures: (i) the conformation of 140-loop and 220-loop, which are more flexible in the zymogen (in FXIIc the electron density of the 220-loop is not visible), while in  $\beta$ FXIIa the 140-loop is positioned in a conformation enabling it to interact with the 180-loop. (ii) the presence of the oxyanion hole in  $\beta$ FXIIa. (iii) The H1 pocket is accessible in  $\beta$ FXIIa, while it is partially buried in the zymogen, due to a 90° rotation of Trp35. (iv) Asp194 in  $\beta$ FXIIa can interact with the N-terminal Val16, this interaction is not possible in the zymogen, due to a 180° rotation of Asp194. Superimposition of  $\beta$ FXIIa with thrombin shows 252 equivalent residues, an overall r.m.s.d. of 1.7 Å and 32% amino acid identity. The two proteins possess very similar structures. Nevertheless, it is of interest to compare the conformation of the PPACK inhibitor in the two proteins. We find that (i) the position of the Pro is flipped of 180° in the PPACK- $\beta$ FXIIa complex; (ii) the Phe of PPACK is stabilized by interactions with thrombin residues, while these interactions are not present in the PPACK- $\beta$ FXIIa complex and the Phe residue is involved only in crystal packing interactions. Therefore, PPACK interacts with  $\beta$ FXIIa only with two residues (Arg and Pro), thus possibly explaining the lower efficiency of PPACK in inhibiting  $\beta$ FXIIa.

An extensive optimization was done using the incomplete factorial method and seeding techniques, but no high quality crystals were obtained. All the crystals obtained were too small for any diffracting experiments. Since all the optimizations were performed at 20 °C, all the identified conditions were also tested at two temperatures, 15 °C and 10 °C. The complex with macromolecular inhibitors, CTI and Ecotin, was prepared as reported in the

literature, and then isolated by gel filtration. The formation of the complex was successful in both cases. However the concentration of the protein obtained was too low to grow crystals suitable for diffraction experiments. For the complex with CTI, only 1 mg/ml of protein could be obtained. In the case of the complex with Ecotin, initial hits were obtained from ProComplex suite. An optimization of the obtained conditions was done by incomplete factorial method but no better crystals were produced.

Two other constructs were made for expression in insect cell system,  $\beta$ FXIIa starting from N335, and  $\beta$ FXIIG570R. The cloning, expression and purification of both constructs was achieved and a protocol for each step was established.

The catalytic activity of recombinant  $\beta$ FXIIa and the mutant  $\beta$ FXIIG570R was determined measuring the hydrolysis of chromogenic peptide substrate S2302. The hydrolysis of S2302 mediated by  $\beta$ FXIIa and  $\beta$ FXIIG570R showed that  $\beta$ FXIIa cleaves the substrate more efficiently than the mutant. The kinetic parameters ( $K_M$ ,  $k_{cat}$ ), calculated from the initial rates against the substrate concentrations, indicated that recombinant  $\beta$ FXIIa behaves as the commercial  $\beta$ FXIIa. Calculated value of the catalytic efficiency,  $k_{cat}/K_M$ , showed that the efficiency of  $\beta$ FXIIa is substantially higher than  $\beta$ FXIIG570R. Schloesser *et al.* supposed that a reason for the low catalytic activity of  $\beta$ FXIIG570R can be due to the positively charged Arg570, inhibiting the formation of the Cys570-Cys540 disulfide bridge, with a consequent alteration of conformation integrity and drop of activity.[169] To verify this hypothesis, several conformations of the mutated Arg570 were identified, using the FoldX software. Interestingly, the interaction of

Arg570 with Glu495 could in principle explain the low catalytic activity of the mutant, if compared to the wild type. In fact, the Arg570-Glu495 interaction completely blocks the access to the active site cleft.

The screening of the crystallization conditions of the two proteins gave initial hits only for the wild type; they were subsequently used for optimization. However, even if the crystals obtained were reproducible, they were not well diffracting. Moreover, none of the performed optimizations provided suitable crystals. In the case of the  $\beta$ FXIIG570R mutant, no hits were obtained. This may be due to the low concentration tested (only 3.6 mg/ml). However, it was not possible to obtain higher concentrated protein.

The two constructs were also co-crystallized with PPACK and PCK, but the majority of the drops were clear. The complex with Ecotin was successful only for the mutant. Different trays were set up and hits were obtained in ProComplex suite. The same screen was set up for Ecotin only as a control and no crystals were obtained. The obtained crystals were further optimized and were tested at the Diamond Light Source but they were weakly diffracting.

The knowledge of the structure of both  $\beta$ FXIIa and  $\beta$ FXIIG570R will help to understand the mechanisms behind the catalytic activity of this enzyme, and to gain insights about the role played by the Arg570-Glu495 interaction. Moreover, the structure of  $\beta$ FXIIa or  $\beta$ FXIIG570R in complex with the PPACK or Ecotin will allow to clarify the interaction mode of the binding pocket with the substrate. In particular, the model of the complex with Ecotin can be used to design new inhibitors that not only interact with the active site but they also have a long tail that can make interactions with

other regions of the target, in order to enhance the specificity and selectivity of the inhibitor.

$\beta$ FXII was also cloned in the pCold-TF vector for expression in bacteria. The expression in Origami 2 cells, as well as the purification, resulted in poor yields of pure protein. Cleavage of the tag was performed. Three different enzymes were tested for the cleavage. The best results were obtained using thrombin and incubating the reaction for two hours at room temperature. However, it was not possible to continue the procedure, since the protein was lost after the cleavage. In fact, after the cleavage of trigger factor, the protein cannot be found in the flow through. This could be due to a low amount of expressed protein. This problem could be solved using a different strain (i.e. shuffle cells) or a different vector with a different tag or a vector without a tag.

$\beta$ FXII was also co-purified with Ecotin. The presence of the inhibitor was thought to stabilize the protein during the cleavage. However, gel filtration revealed that the complex was not formed.

In conclusion, the results presented in this work of thesis indicated that it is possible to produce recombinant  $\beta$ FXIIa from eukaryotic cell system and that its activity is comparable to commercial  $\beta$ FXIIa, as confirmed by enzyme kinetic assays. We also reported the first crystallographic structure of FXIIa in complex with the small peptide inhibitor PPACK. Even if the resolution was low (4Å) a preliminary structural analysis was performed and it showed that MBP- $\beta$ FXIIa is in the active conformation. Nevertheless, in order to obtain full insights of the interaction between an inhibitor and the

catalytic site of FXIIa, it is necessary to work on higher resolution data sets. For this reason, many efforts will be focused in the optimization of the quality of the crystal already obtained, and on an extensive investigation of new crystallization conditions.

## CHAPTER 8

### BIBLIOGRAPHY

1. Anthea, M., et al., *Human biology and health*. Englewood Cliffs, New Jersey, USA, 1993: p. 76-1.
2. Renné, T., *Safe (r) anticoagulation*. *Blood*, 2010. **116**(22): p. 4390-4391.
3. Kelly, B.B. and V. Fuster, *Promoting Cardiovascular Health in the Developing World:: A Critical Challenge to Achieve Global Health*. 2010: National Academies Press.
4. Mathers, C., D.M. Fat, and J. Boerma, *The global burden of disease: 2004 update*. 2008: World Health Organization.
5. Naess, I., et al., *Incidence and mortality of venous thrombosis: a population-based study*. *Journal of Thrombosis and Haemostasis*, 2007. **5**(4): p. 692-699.
6. White, R.H., *The epidemiology of venous thromboembolism*. *Circulation*, 2003. **107**(23 suppl 1): p. I-4-I-8.
7. Silverstein, M.D., et al., *Trends in the incidence of deep vein thrombosis and pulmonary embolism: a 25-year population-based study*. *Archives of internal medicine*, 1998. **158**(6): p. 585-593.
8. Cushman, M., et al., *Deep vein thrombosis and pulmonary embolism in two cohorts: the longitudinal investigation of thromboembolism etiology*. *The American journal of medicine*, 2004. **117**(1): p. 19-25.
9. Previtali, E., et al., *Risk factors for venous and arterial thrombosis*. *Blood Transfusion*, 2011. **9**(2): p. 120.
10. Nickel, K.F., et al., *The polyphosphate-factor XII pathway drives coagulation in prostate cancer-associated thrombosis*. *Blood*, 2015. **126**(11): p. 1379-1389.
11. Solomon, C.G. and I.A. Greer, *Pregnancy complicated by venous thrombosis*. *New England Journal of Medicine*, 2015. **373**(6): p. 540-547.
12. Schousboe, I., *Pharmacological regulation of factor XII activation may be a new target to control pathological coagulation*. *Biochemical pharmacology*, 2008. **75**(5): p. 1007-1013.
13. Boon, G.D., *An overview of hemostasis*. *Toxicologic pathology*, 1993. **21**(2): p. 170-179.
14. Fishman, A.P., *Endothelium: a distributed organ of diverse capabilities*. *Annals of the New York Academy of Sciences*, 1982. **401**(1): p. 1-8.
15. Cines, D.B., et al., *Endothelial cells in physiology and in the pathophysiology of vascular disorders*. *Blood*, 1998. **91**(10): p. 3527-3561.
16. Austin, S.K., *Haemostasis*. *Medicine*, 2009. **37**(3): p. 133-136.
17. Clemetson, K.J., *Platelets and primary haemostasis*. *Thrombosis research*, 2012. **129**(3): p. 220-224.
18. Sixma, J. and J. Wester. *The hemostatic plug*. in *Seminars in hematology*. 1977.
19. Jennings, L.K., *Mechanisms of platelet activation: need for new strategies to protect against platelet-mediated atherothrombosis*. *Thromb Haemost*, 2009. **102**(2): p. 248-257.
20. Piotrowicz, R.S., et al., *Glycoprotein Ic-IIa functions as an activation-independent fibronectin receptor on human platelets*. *The Journal of cell biology*, 1988. **106**(4): p. 1359-1364.

21. Turitto, V.T., et al., *Factor VIII/von Willebrand factor in subendothelium mediates platelet adhesion*. Blood, 1985. **65**(4): p. 823-831.
22. Luo, G.-P., et al., *von Willebrand factor: more than a regulator of hemostasis and thrombosis*. Acta haematologica, 2012. **128**(3): p. 158-169.
23. Behnke, O., *The morphology of blood platelet membrane systems*. Series haematologica, 1969. **3**(4): p. 3-16.
24. Yip, J., et al., *Primary platelet adhesion receptors*. IUBMB life, 2005. **57**(2): p. 103-108.
25. Triplett, D., et al., *Platelet function*. Laboratory Evaluation and Clinical Application. ASCP, 1978.
26. Davie, E.W. and O.D. Ratnoff, *Waterfall sequence for intrinsic blood clotting*. Science, 1964. **145**(3638): p. 1310-1312.
27. Macfarlane, R., *An enzyme cascade in the blood clotting mechanism, and its function as a biochemical amplifier*. 1964.
28. Gailani, D. and T. Renné, *Intrinsic pathway of coagulation and arterial thrombosis*. Arteriosclerosis, thrombosis, and vascular biology, 2007. **27**(12): p. 2507-2513.
29. Gailani, D. and T. Renne, *The intrinsic pathway of coagulation: a target for treating thromboembolic disease?* J Thromb Haemost, 2007. **5**(6): p. 1106-12.
30. Davie, E.W., K. Fujikawa, and W. Kisiel, *The coagulation cascade: initiation, maintenance, and regulation*. Biochemistry, 1991. **30**(43): p. 10363-70.
31. Bach, R., Y. Nemerson, and W. Konigsberg, *Purification and characterization of bovine tissue factor*. J Biol Chem, 1981. **256**(16): p. 8324-31.
32. Broze, G.J., Jr., et al., *Purification of human brain tissue factor*. J Biol Chem, 1985. **260**(20): p. 10917-20.
33. Sakai, T., et al., *Binding of human factors VII and VIIa to a human bladder carcinoma cell line (J82). Implications for the initiation of the extrinsic pathway of blood coagulation*. J Biol Chem, 1989. **264**(17): p. 9980-8.
34. Radcliffe, R. and Y. Nemerson, *Mechanism of activation of bovine factor VII. Products of cleavage by factor Xa*. J Biol Chem, 1976. **251**(16): p. 4749-802.
35. Masys, D.R., S.P. Bajaj, and S.I. Rapaport, *Activation of human factor VII by activated factors IX and X*. Blood, 1982. **60**(5): p. 1143-50.
36. Sheehan, J.P. and T.M. Phan, *Phosphorothioate oligonucleotides inhibit the intrinsic tenase complex by an allosteric mechanism*. Biochemistry, 2001. **40**(16): p. 4980-4989.
37. Tracy, P.B., M.E. Nesheim, and K.G. Mann, *Coordinate binding of factor Va and factor Xa to the unstimulated platelet*. J Biol Chem, 1981. **256**(2): p. 743-51.
38. Mann, K.G., et al., *Surface-dependent reactions of the vitamin K-dependent enzyme complexes*. Blood, 1990. **76**(1): p. 1-16.
39. Monkovic, D.D. and P.B. Tracy, *Activation of human factor V by factor Xa and thrombin*. Biochemistry, 1990. **29**(5): p. 1118-28.
40. Nesheim, M.E. and K.G. Mann, *Thrombin-catalyzed activation of single chain bovine factor V*. J Biol Chem, 1979. **254**(4): p. 1326-34.
41. Lorand, L. and K. Konishi, *Activation of the Fibrin Stabilizing Factor of Plasma by Thrombin*. Arch Biochem Biophys, 1964. **105**: p. 58-67.
42. Bauer, K.A., et al., *Factor IX is activated in vivo by the tissue factor mechanism*. Blood, 1990. **76**(4): p. 731-6.
43. van Dieijen, G., et al., *The role of phospholipid and factor VIIIa in the activation of bovine factor X*. J Biol Chem, 1981. **256**(7): p. 3433-42.
44. Broze, G.J., Jr., T.J. Girard, and W.F. Novotny, *Regulation of coagulation by a multivalent Kunitz-type inhibitor*. Biochemistry, 1990. **29**(33): p. 7539-46.

45. Rapaport, S.I., *Inhibition of factor VIIa/tissue factor-induced blood coagulation: with particular emphasis upon a factor Xa-dependent inhibitory mechanism*. *Blood*, 1989. **73**(2): p. 359-65.
46. Rosenberg, R.D. and P.S. Damus, *The purification and mechanism of action of human antithrombin-heparin cofactor*. *J Biol Chem*, 1973. **248**(18): p. 6490-505.
47. Naito, K. and K. Fujikawa, *Activation of human blood coagulation factor XI independent of factor XII. Factor XI is activated by thrombin and factor XIa in the presence of negatively charged surfaces*. *J Biol Chem*, 1991. **266**(12): p. 7353-8.
48. Renne, T. and D. Gailani, *Role of Factor XII in hemostasis and thrombosis: clinical implications*. *Expert Rev Cardiovasc Ther*, 2007. **5**(4): p. 733-41.
49. Maas, C. and T. Renné, *Regulatory mechanisms of the plasma contact system*. *Thrombosis research*, 2012. **129**: p. S73-S76.
50. Björkqvist, J., A. Jämsä, and T. Renné, *Plasma kallikrein: the bradykinin-producing enzyme*. *Thromb Haemost*, 2013. **110**(3): p. 399-407.
51. Schmaier, A. and K. McCrae, *The plasma kallikrein–kinin system: its evolution from contact activation*. *Journal of Thrombosis and Haemostasis*, 2007. **5**(12): p. 2323-2329.
52. Renne, T., et al., *Characterization of the H-kininogen-binding site on factor XI: a comparison of factor XI and plasma prekallikrein*. *J Biol Chem*, 2002. **277**(7): p. 4892-9.
53. Flier, J.S., et al., *Molecular and cellular biology of blood coagulation*. *New England Journal of Medicine*, 1992. **326**(12): p. 800-806.
54. Hoffman, M., *Remodeling the blood coagulation cascade*. *Journal of thrombosis and thrombolysis*, 2003. **16**(1-2): p. 17-20.
55. Dahlbäck, B., *Blood coagulation*. *The Lancet*, 2000. **355**(9215): p. 1627-1632.
56. Mann, K.G., *Biochemistry and Physiology of Blood Coagulation*. *Thrombosis and Haemostasis*, 1999. **82**(8): p. 165-174.
57. Fearon, D.T. and R.M. Locksley, *The instructive role of innate immunity in the acquired immune response*. *Science*, 1996. **272**(5258): p. 50-54.
58. Bhavsar, A.P., J.A. Guttman, and B.B. Finlay, *Manipulation of host-cell pathways by bacterial pathogens*. *Nature*, 2007. **449**(7164): p. 827-834.
59. Hoffmann, J.A. and J.-M. Reichhart, *Drosophila innate immunity: an evolutionary perspective*. *Nature immunology*, 2002. **3**(2): p. 121-126.
60. Iwanaga, S., *The molecular basis of innate immunity in the horseshoe crab*. *Current opinion in immunology*, 2002. **14**(1): p. 87-95.
61. Flick, M.J., et al., *Leukocyte engagement of fibrin (ogen) via the integrin receptor  $\alpha$  M  $\beta$  2/Mac-1 is critical for host inflammatory response in vivo*. *The Journal of clinical investigation*, 2004. **113**(11): p. 1596-1606.
62. Delvaeye, M. and E.M. Conway, *Coagulation and innate immune responses: can we view them separately?* *Blood*, 2009. **114**(12): p. 2367-2374.
63. Esmon, C.T., *The interactions between inflammation and coagulation*. *British journal of haematology*, 2005. **131**(4): p. 417-430.
64. Levi, M., T. van der Poll, and H.R. Büller, *Bidirectional relation between inflammation and coagulation*. *Circulation*, 2004. **109**(22): p. 2698-2704.
65. Schulz, C., B. Engelmann, and S. Massberg, *Crossroads of coagulation and innate immunity: the case of deep vein thrombosis*. *Journal of Thrombosis and Haemostasis*, 2013. **11**(s1): p. 233-241.
66. Engelmann, B. and S. Massberg, *Thrombosis as an intravascular effector of innate immunity*. *Nature Reviews Immunology*, 2012. **13**(1): p. 34-45.

67. Brill, A., et al., *Neutrophil extracellular traps promote deep vein thrombosis in mice*. Journal of Thrombosis and Haemostasis, 2012. **10**(1): p. 136-144.
68. Massberg, S., et al., *Reciprocal coupling of coagulation and innate immunity via neutrophil serine proteases*. Nature medicine, 2010. **16**(8): p. 887-896.
69. von Brühl, M.-L., et al., *Monocytes, neutrophils, and platelets cooperate to initiate and propagate venous thrombosis in mice in vivo*. The Journal of experimental medicine, 2012. **209**(4): p. 819-835.
70. Kottke-Marchant, K., *The Role of Coagulation in Arterial and Venous Thrombosis, in Antithrombotic Drug Therapy in Cardiovascular Disease*. 2010, Springer. p. 19-38.
71. Broze, G.J., et al., *The lipoprotein-associated coagulation inhibitor that inhibits the factor VII-tissue factor complex also inhibits factor Xa: insight into its possible mechanism of action*. Blood, 1988. **71**(2): p. 335-343.
72. Girard, T.J., et al., *Functional significance of the Kunitz-type inhibitory domains of lipoprotein-associated coagulation inhibitor*. 1989.
73. Broze Jr, M., George J, *Tissue factor pathway inhibitor and the revised theory of coagulation*. Annual review of medicine, 1995. **46**(1): p. 103-112.
74. Carrell, R.W. and D.J. Perry, *The unhinged antithrombins*. British journal of haematology, 1996. **93**(2): p. 253-257.
75. Jin, L., et al., *The anticoagulant activation of antithrombin by heparin*. Proc Natl Acad Sci U S A, 1997. **94**(26): p. 14683-8.
76. Skinner, R., et al., *The 2.6 Å structure of antithrombin indicates a conformational change at the heparin binding site*. J Mol Biol, 1997. **266**(3): p. 601-9.
77. Lane, D.A., H. Philippou, and J.A. Huntington, *Directing thrombin*. Blood, 2005. **106**(8): p. 2605-2612.
78. Sadler, J., et al., *Structure-function relationships of the thrombin-thrombomodulin interaction*. Pathophysiology of Haemostasis and Thrombosis, 1993. **23**(Suppl. 1): p. 183-193.
79. Ye, J., et al., *The active site of thrombin is altered upon binding to thrombomodulin. Two distinct structural changes are detected by fluorescence, but only one correlates with protein C activation*. Journal of Biological Chemistry, 1991. **266**(34): p. 23016-23021.
80. Walker, F.J., *Regulation of activated protein C by a new protein. A possible function for bovine protein S*. J Biol Chem, 1980. **255**(12): p. 5521-4.
81. Law, R.H., et al., *The X-ray crystal structure of full-length human plasminogen*. Cell reports, 2012. **1**(3): p. 185-190.
82. Forsgren, M., et al., *Molecular cloning and characterization of a full-length cDNA clone for human plasminogen*. FEBS Lett, 1987. **213**(2): p. 254-60.
83. Bezerra, J.A., et al., *Plasminogen deficiency leads to impaired remodeling after a toxic injury to the liver*. Proc Natl Acad Sci U S A, 1999. **96**(26): p. 15143-8.
84. Mourey, L., et al., *Crystal Structure of Cleaved Bovine Antithrombin III at 3· 2 Å Resolution*. Journal of molecular biology, 1993. **232**(1): p. 223-241.
85. Mather, T., et al., *The 2.8 Å crystal structure of Gla-domainless activated protein C*. The EMBO journal, 1996. **15**(24): p. 6822.
86. Wang, X., et al., *Crystal structure of the catalytic domain of human plasmin complexed with streptokinase*. Science, 1998. **281**(5383): p. 1662-1665.
87. Colman, R.W., *Are hemostasis and thrombosis two sides of the same coin?* The Journal of experimental medicine, 2006. **203**(3): p. 493-495.
88. Furie, B. and B.C. Furie, *Mechanisms of thrombus formation*. New England Journal of Medicine, 2008. **359**(9): p. 938-949.

89. López, J.A. and J. Chen, *Pathophysiology of venous thrombosis*. Thrombosis research, 2009. **123**: p. S30-S34.
90. Gross, P.L. and J.I. Weitz, *New anticoagulants for treatment of venous thromboembolism*. Arteriosclerosis, thrombosis, and vascular biology, 2008. **28**(3): p. 380-386.
91. Hirsh, J., *Current anticoagulant therapy—unmet clinical needs*. Thrombosis Research, 2003. **109**: p. S1-S8.
92. Bauer, K.A., et al., *Fondaparinux, a synthetic pentasaccharide: the first in a new class of antithrombotic agents—the selective factor Xa inhibitors*. Cardiovascular drug reviews, 2002. **20**(1): p. 37-52.
93. Walenga, J.M., et al., *Fondaparinux: a synthetic heparin pentasaccharide as a new antithrombotic agent*. Expert opinion on investigational drugs, 2002. **11**(3): p. 397-407.
94. Cheng, J.W., *Fondaparinux: a new antithrombotic agent*. Clinical therapeutics, 2002. **24**(11): p. 1757-1769.
95. Eisenberg, P.R., et al., *Importance of factor Xa in determining the procoagulant activity of whole-blood clots*. Journal of Clinical Investigation, 1993. **91**(5): p. 1877.
96. Bayer Pharma, A., *Xarelto®(rivaroxaban) summary of product characteristics*. Available at: Accessed, 2013. **28**.
97. Gómez-Outes, A., et al., *Dabigatran, rivaroxaban, or apixaban versus warfarin in patients with nonvalvular atrial fibrillation: a systematic review and meta-analysis of subgroups*. Thrombosis, 2013. **2013**.
98. Weitz, J. and S. Bates, *New anticoagulants*. Journal of Thrombosis and Haemostasis, 2005. **3**(8): p. 1843-1853.
99. Nar, H., *The role of structural information in the discovery of direct thrombin and factor Xa inhibitors*. Trends Pharmacol Sci, 2012. **33**(5): p. 279-88.
100. Holford, N.H., *Clinical pharmacokinetics and pharmacodynamics of warfarin*. Clinical pharmacokinetics, 1986. **11**(6): p. 483-504.
101. Investigators, S.P.i.A.F., *Adjusted-dose warfarin versus low-intensity, fixed-dose warfarin plus aspirin for high-risk patients with atrial fibrillation: Stroke Prevention in Atrial Fibrillation III randomised clinical trial*. The Lancet, 1996. **348**(9028): p. 633-638.
102. Holbrook, A.M., et al., *Systematic overview of warfarin and its drug and food interactions*. Archives of internal medicine, 2005. **165**(10): p. 1095-1106.
103. Renne, T., et al., *Defective thrombus formation in mice lacking coagulation factor XII*. J Exp Med, 2005. **202**(2): p. 271-81.
104. Kleinschnitz, C., et al., *Targeting coagulation factor XII provides protection from pathological thrombosis in cerebral ischemia without interfering with hemostasis*. J Exp Med, 2006. **203**(3): p. 513-8.
105. Schmaier, A.H., *Why do we want to know how factor XII levels are modulated?* Thrombosis research, 2010. **125**(2): p. 105.
106. Ratnoff, O.D. and J.E. Colopy, *A familial hemorrhagic trait associated with a deficiency of a clot-promoting fraction of plasma*. J Clin Invest, 1955. **34**(4): p. 602-13.
107. Cool, D.E. and R.T. MacGillivray, *Characterization of the human blood coagulation factor XII gene. Intron/exon gene organization and analysis of the 5'-flanking region*. J Biol Chem, 1987. **262**(28): p. 13662-73.
108. Veltkamp, J., E. Loeliger, and H. Hemker, *THE BIOLOGICAL HALF-TIME OF HAGEMAN FACTOR*. Thrombosis et diathesis haemorrhagica, 1965. **13**: p. 1-7.

109. Griffin, J.H. and C.G. Cochrane, *Human factor XII (Hageman factor)*. *Methods Enzymol*, 1976. **45**: p. 56-65.
110. Revak, S.D., et al., *Structural changes accompanying enzymatic activation of human Hageman factor*. *J Clin Invest*, 1974. **54**(3): p. 619-27.
111. Cool, D.E., et al., *Characterization of human blood coagulation factor XII cDNA. Prediction of the primary structure of factor XII and the tertiary structure of beta-factor XIIa*. *J Biol Chem*, 1985. **260**(25): p. 13666-76.
112. Stavrou, E. and A.H. Schmaier, *Factor XII: what does it contribute to our understanding of the physiology and pathophysiology of hemostasis & thrombosis*. *Thromb Res*, 2010. **125**(3): p. 210-5.
113. Revak, S.D., C.G. Cochrane, and J.H. Griffin, *The binding and cleavage characteristics of human Hageman factor during contact activation. A comparison of normal plasma with plasmas deficient in factor XI, prekallikrein, or high molecular weight kininogen*. *J Clin Invest*, 1977. **59**(6): p. 1167-75.
114. Fujikawa, K. and B.A. McMullen, *Amino acid sequence of human beta-factor XIIa*. *J Biol Chem*, 1983. **258**(18): p. 10924-33.
115. Revak, S.D., et al., *Surface and fluid phase activities of two forms of activated Hageman factor produced during contact activation of plasma*. *The Journal of experimental medicine*, 1978. **147**(3): p. 719-729.
116. Hedstrom, L., *Serine protease mechanism and specificity*. *Chemical reviews*, 2002. **102**(12): p. 4501-4524.
117. Carter, P. and J.A. Wells, *Dissecting the catalytic triad of a serine protease*. *Nature*, 1988. **332**(6164): p. 564-568.
118. Dodson, G. and A. Wlodawer, *Catalytic triads and their relatives*. *Trends in biochemical sciences*, 1998. **23**(9): p. 347-352.
119. Henderson, R., *Structure of crystalline  $\alpha$ -chymotrypsin: IV. The structure of indoleacryloyl- $\alpha$ -chymotrypsin and its relevance to the hydrolytic mechanism of the enzyme*. *Journal of molecular biology*, 1970. **54**(2): p. 341-354.
120. Huber, R., et al., *Structure of the complex formed by bovine trypsin and bovine pancreatic trypsin inhibitor: II. Crystallographic refinement at 1.9 Å resolution*. *Journal of molecular biology*, 1974. **89**(1): p. 73-101.
121. Krieger, M., L.M. Kay, and R. Stroud, *Structure and specific binding of trypsin: comparison of inhibited derivatives and a model for substrate binding*. *Journal of molecular biology*, 1974. **83**(2): p. 209-230.
122. Barrett, A.J., N.D. Rawlings, and E.A. O'Brien, *The MEROPS database as a protease information system*. *Journal of structural biology*, 2001. **134**(2): p. 95-102.
123. Gehrmann, M., et al., *The col-1 module of human matrix metalloproteinase-2 (MMP-2): structural/functional relatedness between gelatin-binding fibronectin type II modules and lysine-binding kringle domains*. *Biological chemistry*, 2002. **383**(1): p. 137-148.
124. Mulichak, A.M., A. Tulinsky, and K. Ravichandran, *Crystal and molecular structure of human plasminogen kringle 4 refined at 1.9-Å resolution*. *Biochemistry*, 1991. **30**(43): p. 10576-10588.
125. Kao, Y.-H., et al., *The effect of O-fucosylation on the first EGF-like domain from human blood coagulation factor VII*. *Biochemistry*, 1999. **38**(22): p. 7097-7110.
126. Beringer, D. and L. Kroon-Batenburg, *The structure of the FnI-EGF-like tandem domain of coagulation factor XII solved using SIRAS*. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*, 2013. **69**(2): p. 94-102.
127. Petersen, T.E., et al., *Partial primary structure of bovine plasma fibronectin: three types of internal homology*. *Proc Natl Acad Sci U S A*, 1983. **80**(1): p. 137-41.

128. Clarke, B.J., et al., *Mapping of a putative surface-binding site of human coagulation factor XII*. J Biol Chem, 1989. **264**(19): p. 11497-502.
129. Samuel, M., E. Samuel, and G.B. Villanueva, *Histidine residues are essential for the surface binding and autoactivation of human coagulation factor XII*. Biochem Biophys Res Commun, 1993. **191**(1): p. 110-7.
130. Citarella, F., et al., *The second exon-encoded factor XII region is involved in the interaction of factor XII with factor XI and does not contribute to the binding site for negatively charged surfaces*. Blood, 1998. **92**(11): p. 4198-206.
131. Mahdi, F., et al., *Factor XII interacts with the multiprotein assembly of urokinase plasminogen activator receptor, gC1qR, and cytokeratin 1 on endothelial cell membranes*. Blood, 2002. **99**(10): p. 3585-96.
132. Schousboe, I., *Contact activation in human plasma is triggered by zinc ion modulation of factor XII (Hageman factor)*. Blood Coagul Fibrinolysis, 1993. **4**(5): p. 671-8.
133. Rojkaer, R. and I. Schousboe, *Partial identification of the Zn<sup>2+</sup>-binding sites in factor XII and its activation derivatives*. Eur J Biochem, 1997. **247**(2): p. 491-6.
134. Bradford, H.N., R.A. Pixley, and R.W. Colman, *Human factor XII binding to the glycoprotein Ib-IX-V complex inhibits thrombin-induced platelet aggregation*. J Biol Chem, 2000. **275**(30): p. 22756-63.
135. Pietrucha, T., et al., *The epidermal growth factor-like domain from tissue plasminogen activator. Cloning in E. coli, purification and ESR studies of its interaction with human blood platelets*. ACTA BIOCHIMICA POLONICA-ENGLISH EDITION-, 1994. **41**: p. 25-25.
136. Selander-Sunnerhagen, M., et al., *How an epidermal growth factor (EGF)-like domain binds calcium. High resolution NMR structure of the calcium form of the NH<sub>2</sub>-terminal EGF-like domain in coagulation factor X*. Journal of Biological Chemistry, 1992. **267**(27): p. 19642-19649.
137. Handford, P., et al., *The first EGF-like domain from human factor IX contains a high-affinity calcium binding site*. The EMBO journal, 1990. **9**(2): p. 475.
138. Bork, P., et al., *Structure and distribution of modules in extracellular proteins*. Q Rev Biophys, 1996. **29**(2): p. 119-67.
139. Cohen, S., *Epidermal growth factor*. In Vitro Cell Dev Biol, 1987. **23**(4): p. 239-46.
140. Chien, P., et al., *Modulation of the human monocyte binding site for monomeric immunoglobulin G by activated Hageman factor*. J Clin Invest, 1988. **82**(5): p. 1554-9.
141. Gordon, E.M., et al., *Factor XII-induced mitogenesis is mediated via a distinct signal transduction pathway that activates a mitogen-activated protein kinase*. Proc Natl Acad Sci U S A, 1996. **93**(5): p. 2174-9.
142. McMullen, B. and K. Fujikawa, *Amino acid sequence of the heavy chain of human alpha-factor XIIIa (activated Hageman factor)*. Journal of Biological Chemistry, 1985. **260**(9): p. 5328-5341.
143. Takada, Y., *Potential role of kringle-integrin interaction in plasmin and uPA actions (a hypothesis)*. BioMed Research International, 2012. **2012**.
144. Pathak, M., et al., *Coagulation Factor XII protease domain crystal structure*. Journal of Thrombosis and Haemostasis, 2015.
145. Kraut, J., *Serine proteases: structure and mechanism of catalysis*. Annual review of biochemistry, 1977. **46**(1): p. 331-358.
146. Birktoft, J.J., J. Kraut, and S.T. Freer, *A detailed structural comparison between the charge relay system in chymotrypsinogen and in  $\alpha$ -chymotrypsin*. Biochemistry, 1976. **15**(20): p. 4481-4485.

147. Renné, T., et al., *In vivo roles of factor XII*. Blood, 2012. **120**(22): p. 4296-4303.
148. MATSUMOTO, K., et al., *Pathogenesis of Serratial Infection: Activation of the Hageman Factor-Prekallikrein Cascade by Serratial Protease1*. Journal of biochemistry, 1984. **96**(3): p. 739-749.
149. Kenne, E. and T. Renné, *Factor XII: a drug target for safe interference with thrombosis and inflammation*. Drug discovery today, 2014. **19**(9): p. 1459-1464.
150. Bäck, J., et al., *Activated human platelets induce factor XIIa-mediated contact activation*. Biochemical and biophysical research communications, 2010. **391**(1): p. 11-17.
151. Mackman, N. and A. Gruber, *Platelet polyphosphate: an endogenous activator of coagulation factor XII*. Journal of Thrombosis and Haemostasis, 2010. **8**(5): p. 865-867.
152. Joseph, K., et al., *Identification of the zinc-dependent endothelial cell binding protein for high molecular weight kininogen and factor XII: identity with the receptor that binds to the globular "heads" of C1q (gC1q-R)*. Proceedings of the National Academy of Sciences, 1996. **93**(16): p. 8552-8557.
153. Shariat-Madar, Z., F. Mahdi, and A.H. Schmaier, *Assembly and activation of the plasma kallikrein/kinin system: a new interpretation*. International immunopharmacology, 2002. **2**(13): p. 1841-1849.
154. Müller, F., D. Gailani, and T. Renne, *Factor XI and XII as antithrombotic targets*. Current opinion in hematology, 2011. **18**(5): p. 349-355.
155. van der Meijden, P.E., et al., *Dual role of collagen in factor XII-dependent thrombus formation*. Blood, 2009. **114**(4): p. 881-890.
156. Pauer, H.-U., et al., *Targeted deletion of murine coagulation factor XII gene-a model for contact phase activation in vivo*. THROMBOSIS AND HAEMOSTASIS-STUTTGART-, 2004. **92**: p. 503-508.
157. Dewald, G. and K. Bork, *Missense mutations in the coagulation factor XII (Hageman factor) gene in hereditary angioedema with normal C1 inhibitor*. Biochem Biophys Res Commun, 2006. **343**(4): p. 1286-9.
158. Bork, K., D. Gul, and G. Dewald, *Hereditary angio-oedema with normal C1 inhibitor in a family with affected women and men*. Br J Dermatol, 2006. **154**(3): p. 542-5.
159. Bork, K., et al., *A novel mutation in the coagulation factor 12 gene in subjects with hereditary angioedema and normal C1-inhibitor*. Clin Immunol, 2011. **141**(1): p. 31-5.
160. Lumry, W.R., *Overview of epidemiology, pathophysiology, and disease progression in hereditary angioedema*. Am J Manag Care, 2013. **19**(7 Suppl): p. s103-10.
161. Bork, K. and N. Ressel, *Sudden upper airway obstruction in patients with hereditary angioedema*. Transfusion and apheresis science, 2003. **29**(3): p. 235-238.
162. Tosi, M., *Molecular genetics of C1 inhibitor*. Immunobiology, 1998. **199**(2): p. 358-365.
163. Zuraw, B.L., *Hereditary angioedema*. New England Journal of Medicine, 2008. **359**(10): p. 1027-1036.
164. Cichon, S., et al., *Increased activity of coagulation factor XII (Hageman factor) causes hereditary angioedema type III*. The American Journal of Human Genetics, 2006. **79**(6): p. 1098-1104.
165. Dewald, G. and K. Bork, *Missense mutations in the coagulation factor XII (Hageman factor) gene in hereditary angioedema with normal C1 inhibitor*. Biochemical and biophysical research communications, 2006. **343**(4): p. 1286-1289.

166. Bork, K., et al., *A novel mutation in the coagulation factor 12 gene in subjects with hereditary angioedema and normal C1-inhibitor*. Clinical Immunology, 2011. **141**(1): p. 31-35.
167. Kiss, N., et al., *Novel duplication in the *F12* gene in a patient with recurrent angioedema*. Clinical Immunology, 2013. **149**(1): p. 142-145.
168. Moreno, A. *Structural and Molecular Changes Caused By Mutations Thr328Lys and Thr328Arg In FXII Associated With Hereditary Angioedema With Normal C1 Inhibitor*. in *2014 AAAAI Annual Meeting*. 2014. Aaaaai.
169. Schloesser, M., et al., *Mutations in the human factor XII gene*. Blood, 1997. **90**(10): p. 3967-3977.
170. Semba, U., et al., *Primary structure of guinea-pig Hageman factor: sequence around the cleavage site differs from the human molecule*. Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology, 1992. **1159**(2): p. 113-121.
171. Shibuya, Y., et al., *Primary structure of bovine Hageman factor (blood coagulation factor XII): comparison with human and guinea pig molecules*. Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology, 1994. **1206**(1): p. 63-70.
172. Bode, W., D. Turk, and A. Karshikov, *The refined 1.9-Å X-ray crystal structure of d-Phe-Pro-Arg chloromethylketone-inhibited human  $\alpha$ -thrombin: Structure analysis, overall structure, electrostatic properties, detailed active-site geometry, and structure-function relationships*. Protein Science, 1992. **1**(4): p. 426-471.
173. Powers, J.C. and C.-M. Kam, *Synthetic substrates and inhibitors of thrombin*, in *Thrombin*. 1992, Springer. p. 117-158.
174. Lefkovits, J. and E.J. Topol, *Direct thrombin inhibitors in cardiovascular medicine*. Circulation, 1994. **90**(3): p. 1522-1536.
175. Brandstetter, H., et al., *X-ray structure of clotting factor IXa: active site and module structure related to Xase activity and hemophilia B*. Proceedings of the National Academy of Sciences, 1995. **92**(21): p. 9796-9800.
176. de Maat, S., et al., *A nanobody-based method for tracking factor XII activation in plasma*. Thromb Haemost, 2013. **109**(4).
177. Martin, P.D., et al., *New insights into the regulation of the blood clotting cascade derived from the X-ray crystal structure of bovine meizothrombin des F1 in complex with PPACK*. Structure, 1997. **5**(12): p. 1681-1693.
178. Hauptmann, J. and F. Markwardt. *Pharmacologic aspects of the development of selective synthetic thrombin inhibitors as anticoagulants*. in *Seminars in thrombosis and hemostasis*. 1992. Thieme.
179. TANS, G., et al., *Studies on the effect of serine protease inhibitors on activated contact factors*. European Journal of Biochemistry, 1987. **164**(3): p. 637-642.
180. Ghebrehwet, B., et al., *Mechanisms of activation of the classical pathway of complement by Hageman factor fragment*. Journal of Clinical Investigation, 1983. **71**(5): p. 1450.
181. Chung, C., et al., *Purification from Escherichia coli of a periplasmic protein that is a potent inhibitor of pancreatic proteases*. Journal of Biological Chemistry, 1983. **258**(18): p. 11032-11038.
182. McGrath, M.E., et al., *Expression of the protease inhibitor ecotin and its co-crystallization with trypsin*. Journal of molecular biology, 1991. **222**(2): p. 139-142.
183. McGrath, M.E., et al., *Macromolecular chelation as an improved mechanism of protease inhibition: structure of the ecotin-trypsin complex*. The EMBO journal, 1994. **13**(7): p. 1502.

184. Jin, L., et al., *Crystal structures of the FXIa catalytic domain in complex with ecotin mutants reveal substrate-like interactions*. Journal of Biological Chemistry, 2005. **280**(6): p. 4704-4712.
185. Wang, S.X., C.T. Esmon, and R.J. Fletterick, *Crystal structure of thrombin-ecotin reveals conformational changes and extended interactions*. Biochemistry, 2001. **40**(34): p. 10038-10046.
186. Yang, S.Q., et al., *Ecotin: a serine protease inhibitor with two distinct and interacting binding sites*. Journal of molecular biology, 1998. **279**(4): p. 945-957.
187. Ulmer, J.S., et al., *Ecotin is a potent inhibitor of the contact system proteases factor XIIa and plasma kallikrein*. FEBS letters, 1995. **365**(2): p. 159-163.
188. Shin, D.H., et al., *Crystal structure analyses of uncomplexed ecotin in two crystal forms: implications for its function and stability*. protein Science, 1996. **5**(11): p. 2236-2247.
189. Wen, L., et al., *Nucleotide sequence of a cDNA clone that encodes the maize inhibitor of trypsin and activated Hageman factor*. Plant molecular biology, 1992. **18**(4): p. 813-814.
190. Hojima, Y., J.V. Pierce, and J.J. Pisano, *Hageman factor fragment inhibitor in corn seeds: purification and characterization*. Thrombosis research, 1980. **20**(2): p. 149-162.
191. Toossi, Z., et al., *Induction of expression of monocyte interleukin 1 by Hageman factor (factor XII)*. Proceedings of the National Academy of Sciences, 1992. **89**(24): p. 11969-11972.
192. Pedersen, L.C., et al., *The corn inhibitor of blood coagulation factor XIIa: Crystallization and preliminary crystallographic analysis*. Journal of molecular biology, 1994. **236**(1): p. 385-387.
193. Hazegh-Azam, M., et al., *The corn inhibitor of activated Hageman factor: purification and properties of two recombinant forms of the protein*. Protein expression and purification, 1998. **13**(2): p. 143-149.
194. Behnke, C.A., et al., *Structural determinants of the bifunctional corn Hageman factor inhibitor: X-ray crystal structure at 1.95 Å resolution*. Biochemistry, 1998. **37**(44): p. 15277-15288.
195. Inomo, A., et al., *The antigenic binding sites of autoantibodies to factor XII in patients with recurrent pregnancy losses*. Thromb Haemost, 2008. **99**(2): p. 316-23.
196. Schousboe, I., *Pharmacological regulation of factor XII activation may be a new target to control pathological coagulation*. Biochem Pharmacol, 2008. **75**(5): p. 1007-13.
197. Epstein, C.J., R.F. Goldberger, and C.B. Anfinsen. *The genetic control of tertiary protein structure: studies with model systems*. in *Cold Spring Harbor symposia on quantitative biology*. 1963. Cold Spring Harbor Laboratory Press.
198. Fersht, A.R., *From the first protein structures to our current knowledge of protein folding: delights and scepticisms*. Nature Reviews Molecular Cell Biology, 2008. **9**(8): p. 650-654.
199. Blundell, T.L. and S. Patel, *High-throughput X-ray crystallography for drug discovery*. Current opinion in pharmacology, 2004. **4**(5): p. 490-496.
200. Rupp, B., *Biomolecular crystallography*. 2010: Garland Science New York.
201. McPherson, A., *Introduction to protein crystallization*. Methods, 2004. **34**(3): p. 254-265.
202. Drenth, J., *Principles of Protein X-ray Crystallography, 1994*. Ref Type: Serial (Book, Monograph).

203. Asherie, N., *Protein crystallization and phase diagrams*. Methods, 2004. **34**(3): p. 266-272.
204. McPherson, A., *Crystallization of biological macromolecules*. 1999.
205. Gilliland, G.L., et al., *Biological Macromolecule Crystallization Database, Version 3.0: new features, data and the NASA archive for protein crystal growth data*. Acta Crystallographica Section D: Biological Crystallography, 1994. **50**(4): p. 408-413.
206. McPherson, A., *Crystallization of proteins from polyethylene glycol*. Journal of Biological Chemistry, 1976. **251**(20): p. 6300-6303.
207. Moshe A, D. and M. Yorgo, *Protein crystallization for X-ray crystallography*. Journal of Visualized Experiments, 2011(47).
208. Durbin, S. and G. Feher, *Protein crystallization*. Annual review of physical chemistry, 1996. **47**(1): p. 171-204.
209. Weber, P.C., *Overview of protein crystallization methods*. Methods in enzymology, 1997. **276**: p. 13-22.
210. Rhodes, G., *Crystallography made crystal clear: a guide for users of macromolecular models*. 2010: Academic press.
211. Wooh, J.W., et al., *Comparison of three commercial sparse-matrix crystallization screens*. Acta Crystallographica Section D: Biological Crystallography, 2003. **59**(4): p. 769-772.
212. Kundrot, C.E., *Which strategy for a protein crystallization project?* Cell Mol Life Sci, 2004. **61**(5): p. 525-36.
213. Carter, C.W., Jr. and C.W. Carter, *Protein crystallization using incomplete factorial experiments*. J Biol Chem, 1979. **254**(23): p. 12219-23.
214. Abergel, C., et al., *Systematic use of the incomplete factorial approach in the design of protein crystallization experiments*. J Biol Chem, 1991. **266**(30): p. 20131-8.
215. Bergfors, T., *Seeds to crystals*. J Struct Biol, 2003. **142**(1): p. 66-76.
216. Luft, J.R. and G.T. DeTitta, *A method to produce microseed stock for use in the crystallization of biological macromolecules*. Acta Crystallogr D Biol Crystallogr, 1999. **55**(Pt 5): p. 988-93.
217. Hook, J.R. and H.E. Hall, *Solid State Physics (The Manchester Physics Series)*. 2000, John Wiley and Sons Ltd.
218. Klug, H.P. and L.E. Alexander, *X-ray diffraction procedures*. 1954.
219. Rossmann, M.G. and D.M. Blow, *The detection of sub-units within the crystallographic asymmetric unit*. Acta Crystallographica, 1962. **15**(1): p. 24-31.
220. Wilson, A.J.C., *International Tables for Crystallography: Mathematical, physical, and chemical tables*. Vol. 3. 1992: Published for the International Union of Crystallography by Kluwer Academic Publishers.
221. Nespolo, M., B. Souvignier, and D.B. Litvin, *About the concept and definition of "noncrystallographic symmetry"*. Zeitschrift für Kristallographie International journal for structural, physical, and chemical aspects of crystalline materials, 2008. **223**(9): p. 605-606.
222. Terwilliger, T.C., *Finding non-crystallographic symmetry in density maps of macromolecular structures*. Journal of structural and functional genomics, 2013. **14**(3): p. 91-95.
223. Bragg, W. and W. Bragg, *The reflection of X-rays by crystals*. Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character, 1913: p. 428-438.
224. Ewald, P., *Introduction to the dynamical theory of X-ray diffraction*. Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography, 1969. **25**(1): p. 103-108.

225. Rossmann, M.G. and C.G. van Beek, *Data processing*. Acta Crystallographica Section D: Biological Crystallography, 1999. **55**(10): p. 1631-1640.
226. Leslie, A., *Mosfilm*. Int CCP4 and ESF-EACMB Newslett, 1992.
227. Minor, W. and Z. Otwinowski, *HKL2000 (Denzo-SMN) Software Package. Processing of X-ray Diffraction Data Collected in Oscillation Mode*. Methods in Enzymology, Macromolecular Crystallography, Academic Press, New York, 1997.
228. Evans, P., *Scaling and assessment of data quality*. Acta Crystallographica Section D: Biological Crystallography, 2006. **62**(1): p. 72-82.
229. Blow, D. and M. Rossmann, *The single isomorphous replacement method*. Acta Crystallographica, 1961. **14**(11): p. 1195-1202.
230. Hendrickson, W.A., et al., *Crystallographic structure analysis of lamprey hemoglobin from anomalous dispersion of synchrotron radiation*. Proteins: Structure, Function, and Bioinformatics, 1988. **4**(2): p. 77-88.
231. Cowtan, K., *Phase Problem in X-ray Crystallography, and Its Solution*. eLS.
232. Crowther, R., *The Molecular Replacement Method*, edited by MG Rossmann. New York: Gordon and Breach, 1972: p. 173-178.
233. Patterson, A., *Zeit. f. Krist*, 1935. **90**: p. 517.
234. Terwilliger, T.C., *Automated structure solution, density modification and model building*. Acta Crystallographica Section D: Biological Crystallography, 2002. **58**(11): p. 1937-1940.
235. Wang, B.-C., *Resolution of phase ambiguity in macromolecular crystallography*. Methods in enzymology, 1985. **115**: p. 90-112.
236. Brünger, A.T., *Assessment of phase accuracy by cross validation: the free R value. Methods and applications*. Acta Crystallographica Section D: Biological Crystallography, 1993. **49**(1): p. 24-36.
237. Debye, P., *Interference of x rays and heat movement*. Ann. Phys, 1913. **43**: p. 49-95.
238. Ramachandran, G., C. Ramakrishnan, and V. Sasisekharan, *Stereochemistry of polypeptide chain configurations*. Journal of molecular biology, 1963. **7**(1): p. 95-99.
239. Laskowski, R.A., et al., *PROCHECK: a program to check the stereochemical quality of protein structures*. Journal of applied crystallography, 1993. **26**(2): p. 283-291.
240. Iwaki, T. and F.J. Castellino, *A single plasmid transfection that offers a significant advantage associated with puromycin selection in Drosophila Schneider S2 cells expressing heterologous proteins*. Cytotechnology, 2008. **57**(1): p. 45-49.
241. Sørensen, H.P. and K.K. Mortensen, *Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli*. Microbial cell factories, 2005. **4**(1): p. 1.
242. Hesterkamp, T., et al., *Escherichia coli trigger factor is a prolyl isomerase that associates with nascent polypeptide chains*. Proceedings of the National Academy of Sciences, 1996. **93**(9): p. 4437-4441.
243. Qing, G., et al., *Cold-shock induced high-yield protein production in Escherichia coli*. Nature biotechnology, 2004. **22**(7): p. 877-882.
244. Moon, A.F., et al., *A synergistic approach to protein crystallization: combination of a fixed-arm carrier with surface entropy reduction*. Protein Sci, 2010. **19**(5): p. 901-13.
245. Kapust, R.B. and D.S. Waugh, *Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused*. Protein Science, 1999. **8**(08): p. 1668-1674.
246. Sachdev, D. and J.M. Chirgwin, *[20] Fusions to maltose-binding protein: Control of folding and solubility in protein purification*. Methods in enzymology, 2000. **326**: p. 312-321.
247. Derewenda, Z.S., *Rational protein crystallization by mutational surface engineering*. Structure, 2004. **12**(4): p. 529-35.

248. Derewenda, Z.S. and P.G. Vekilov, *Entropy and surface engineering in protein crystallization*. Acta Crystallogr D Biol Crystallogr, 2006. **62**(Pt 1): p. 116-24.
249. Avbelj, F. and L. Fele, *Role of main-chain electrostatics, hydrophobic effect and side-chain conformational entropy in determining the secondary structure of proteins*. J Mol Biol, 1998. **279**(3): p. 665-84.
250. Liu, Y., et al., *Crystal structure of the SarR protein from Staphylococcus aureus*. Proceedings of the National Academy of Sciences, 2001. **98**(12): p. 6877-6882.
251. Skrzypczak-Jankun, E., et al., *Structure of the hirugen and hirulog 1 complexes of alpha-thrombin*. J Mol Biol, 1991. **221**(4): p. 1379-93.
252. D'Arcy, A., F. Villard, and M. Marsh, *An automated microseed matrix-screening method for protein crystallization*. Acta Crystallogr D Biol Crystallogr, 2007. **63**(Pt 4): p. 550-4.
253. Kabsch, W., *Xds*. Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 125-32.
254. Winn, M.D., et al., *Overview of the CCP4 suite and current developments*. Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 4): p. 235-42.
255. McCoy, A.J., et al., *Phaser crystallographic software*. J Appl Crystallogr, 2007. **40**(Pt 4): p. 658-674.
256. Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics*. Acta Crystallographica Section D: Biological Crystallography, 2004. **60**(12): p. 2126-2132.
257. Afonine, P.V., et al., *Towards automated crystallographic structure refinement with phenix.refine*. Acta Crystallographica Section D: Biological Crystallography, 2012. **68**(4): p. 352-367.
258. Kabsch, W., *Xds*. Acta Crystallographica Section D: Biological Crystallography, 2010. **66**(2): p. 125-132.
259. McCoy, A.J., et al., *Phaser crystallographic software*. Journal of applied crystallography, 2007. **40**(4): p. 658-674.
260. Matthews, B.W., *Solvent content of protein crystals*. Journal of molecular biology, 1968. **33**(2): p. 491-497.
261. Langer, G., et al., *Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7*. Nature protocols, 2008. **3**(7): p. 1171-1179.
262. Krissinel, E. and K. Henrick, *Inference of macromolecular assemblies from crystalline state*. Journal of molecular biology, 2007. **372**(3): p. 774-797.
263. Holm, L. and P. Rosenström, *Dali server: conservation mapping in 3D*. Nucleic acids research, 2010. **38**(suppl 2): p. W545-W549.
264. Miyazawa, K., *Hepatocyte growth factor activator (HGFA): a serine protease that links tissue injury to activation of hepatocyte growth factor*. FEBS Journal, 2010. **277**(10): p. 2208-2214.
265. Wallace, A.C., R.A. Laskowski, and J.M. Thornton, *LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions*. Protein engineering, 1995. **8**(2): p. 127-134.
266. Ganesan, R., C. Eigenbrot, and D. Kirchhofer, *Structural and mechanistic insight into how antibodies inhibit serine proteases*. Biochem. J, 2010. **430**: p. 179-189.
267. Shia, S., et al., *Conformational Lability in Serine Protease Active Sites: Structures of Hepatocyte Growth Factor Activator (HGFA) Alone and with the Inhibitory Domain from HGFA Inhibitor-1B*. Journal of Molecular Biology, 2005. **346**(5): p. 1335-1349.
268. Schymkowitz, J., et al., *The FoldX web server: an online force field*. Nucleic acids research, 2005. **33**(suppl 2): p. W382-W388.

269. Bondos, S.E. and A. Bicknell, *Detection and prevention of protein aggregation before, during, and after purification*. Analytical biochemistry, 2003. **316**(2): p. 223-231.
270. Vagenende, V., M.G. Yap, and B.L. Trout, *Mechanisms of protein stabilization and prevention of protein aggregation by glycerol*. Biochemistry, 2009. **48**(46): p. 11084-11096.
271. Lengyel, Z., G. Pál, and M. Sahin-Tóth, *Affinity purification of recombinant trypsinogen using immobilized ecotin*. Protein expression and purification, 1998. **12**(2): p. 291-294.