



**CHARACTERIZATION, CRYSTALLIZATION AND STRUCTURE
PREDICTION OF RECOMBINANT PROTEASE FROM *BACILLUS PUMILUS*
115B**

By

IZATUL AZIRA NOR AZMAN

**Thesis Submitted to the School of Graduate Studies, Universiti Putra
Malaysia, in Fulfilment of the Requirements for the Degree of Master of
Science**

February 2015

FBSB 2015 33

All material contained within the thesis, including without limitation text, logos, icons, photographs and all other artwork, is copyright material of Universiti Putra Malaysia unless otherwise stated. Use may be made of any material contained within the thesis for non-commercial purposes from the copyright holder. Commercial use of material may only be made with the express, prior, written permission of Universiti Putra Malaysia.

Copyright © Universiti Putra Malaysia



Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfillment of the requirement for the degree of Master of Science

**CHARACTERIZATION, CRYSTALLIZATION AND STRUCTURE
PREDICTION OF RECOMBINANT PROTEASE FROM *BACILLUS PUMILUS*
115B**

By

IZATUL AZIRA NOR AZMAN

February 2015

Chair : Professor Raja Noor Zaliha Raja Abd Rahman, PhD
Faculty : Biotechnology and Biomolecular Sciences

Proteases are widely used in industry as biocatalyst. Being able to synthesize peptide bonds in microaqueous environment while hydrolyse it in aqueous environment, this enzyme plays very important roles in food, pharmaceutical, detergent and leather industry. Protease from *Bacillus pumilus* 115B was capable to withstand in moderate temperature and proven to be stable in wide range of pH. To get better understanding of this enzyme as well as to enhance its productivity, molecular cloning, characterization and expression of the enzyme is compulsory. The full sequence of 1,149bp encoded a polypeptide of 383 amino acid residues from the organic solvent tolerant protease was successfully cloned into several vectors. Above all, pET 32b vector in BL21(DE3) host showed the highest expression. 115B protease was successfully purified and a single band of 33kDA mature protein was detected at the final step purification using ion exchange chromatography method. Further study on characterization, crystallization and structure prediction were carried out since the structure of this enzyme will give the new insight into organic solvent tolerant properties on a molecular level. The optimum temperature of 115B protease was found to be at 50°C and was stable in temperature range of 30°C to 45°C. The protease activity decrease rapidly at temperature higher than 55°C. 115B protease was stable in pH ranging from pH 7.0-pH 11.0 and the optimum pH was pH 8.0. The protease activity was 91% inhibited by PMSF suggesting that this protease belongs to the serine protease superfamily. Optimization on crystallization condition showed that 115B protease need the use of microseeding to kick start the nucleation process. Crystallization screening showed best crystal grew in formulation 22 from Molecular Dimension II. However, X-ray diffraction studies failed to give good diffraction spot that may result from the bad quality crystal. Homology modelling study helps to give a structural insight of the enzyme. 115B protease was modelled using crystallized

structure of subtilisin BPN complex (PDB id: 1YJB) as template for homology modelling. The predicted structure gave the overall quality Z-score of -0.305 which considered a good quality of modelled structure. The model appeared to have an Aspartic, Serine and Histidine catalytic triad, like all the subtilisin family. The predicted structure shown that 115B protease contains a calcium binding site that located at similar place in the subtilisin BPN' complex structure which is believe to attributes to the inefficient folding of subtilisin BPN' mature enzyme. Experimental result showed that expression of 115B protease was improved compared from previous recombinant 115B/pQE30 by Mahammad (2007). 115B protease was successfully crystallized with the size of 80 μ m, however due to the size factor, X-ray diffraction failed to give sufficient data to construct the electron density map to build the structure. Hence, homology modelling was performed and successfully predicted the structure of 115B protease. This provides preliminary overview about the structure and function of the enzyme.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia
sebagai memenuhi keperluan untuk ijazah Master Sains

**PENCIRIAN, PENGHABLURAN DAN RAMALAN STRUKTUR
REKOMBINAN PROTEASE DARI *BACILLUS PUMILUS* 115B**

Oleh

IZATUL AZIRA NOR AZMAN

Februari 2015

Pengerusi : Profesor Raja Noor Zaliha Raja Abd Rahman, PhD
Fakulti : Bioteknologi dan Sains Biomolekul

Protease digunakan secara meluas di dalam industri biomangkin. Kebolehan untuk mensintesis bon peptida dalam persekitaran microaqueous dan juga didalam persekitaran akueus, enzim ini memainkan peranan yang penting dalam industry makanan, farmaseutikal, bahan pencuci dan kulit. Protease daripada *Bacillus pumilus* 115B mampu bertahan pada suhu sederhana dan terbukti stabil dalam julat pH yang besar. Untuk mendapatkan pemahaman yang lebih baik daripada enzim ini dan juga untuk meningkatkan produktiviti, pengklonan molekul, pencirian dan ungkapan enzim wajib dilakukan. Urutan penuh dengan 1,149 bp polipeptida daripada 383 residu asid amino dari protease rintang pelarut organik telah berjaya diklon ke dalam beberapa vektor. pET vektor 32b di BL21 (DE3) telah menunjukkan eksperisi protin tertinggi. 115B protease telah berjaya dituliskan dan satu jalur 33kDA protein matang dikesan di langkah terakhir penulenan menggunakan kaedah kromatografi pertukaran ion. Kajian lanjut mengenai pencirian, penghabluran dan ramalan struktur telah untuk memberikan lebih pemahaman tentang sifat-sifat enzim ini di peringkat molekul. Suhu optimum 115B protease didapati pada 50 ° C dan stabil dalam julat suhu 30 ° C hingga 45 ° C. Aktiviti protease menurun dengan cepat pada suhu yang lebih tinggi daripada 55 ° C. 115B protease stabil dalam julat pH antara pH 7.0- pH 11.0 dan pH optimum adalah pH 8.0. Aktiviti protease di rencat 91% oleh PMSF, mencadangkan bahawa protease ini berada di dalam superfamili serine protease. Pencarian kondisi optimum kristal menunjukkan 115B protease memerlukan penggunaan pembenihan mikro untuk memulakan proses penukleusan. Pemeriksaan penghabluran menunjukkan kristal terbaik berkembang dalam formulasi 22 dari Molekul Dimensi II. Walau bagaimanapun, kajian pembelauan sinar-X gagal untuk memberi keputusan pembelauan yang baik. Hal ini disebabkan kualiti kristal yang tidak memuaskan. Kajian pemodelan digunakan untuk membantu memberi gambaran struktur enzim. 115B protease telah dimodelkan menggunakan struktur kristal daripada subtilisin BPN

kompleks (PDB id: 1 1YJB) sebagai acuan untuk pemodelan homologi. Kualiti keseluruhan struktur memberikan Z-skor -0,305 yang dianggap berkualiti baik bagi struktur pemodelan. Model ini mempunyai persamaan pemangkin aktif aspartik, serine dan histidine, seperti yang terdapat pada semua keluarga subtilisin itu. Analisis struktur meramalkan menunjukkan bahawa 115B protease mengandungi sebuah laman mengikat kalsium yang terletak di tempat yang sama pada struktur kompleks subtilisin 'BPN enzim yang matang. Analisis hasil eksperimen menunjukkan bahawa ekspresi protin 115B protease telah meningkat berbanding dari rekombinan sebelumnya 115B / pQE30 oleh Mahammad (2007). Pengkristalan 115B protease telah berjaya namun disebabkan oleh faktor saiz, sinar-X pembelauan gagal memberikan data yang mencukupi untuk membina peta elektron seterusnya menentukan strukturnya. Ramalan struktur oleh pemodelan homology berjaya meramalkan struktur 115B protease dengan itu membantu dalam memberikan gambaran keseluruhan secara teori mengenai enzim tersebut.

ACKNOWLEDGEMENTS

Hereby, I would like to express my profound gratitude to individuals who have played an important part in helping me undergo this project and finish this thesis. First and foremost, I would like to express my highest gratitude to Allah for the blessing and His guidance throughout this journey. Had it not to be His will, this thesis will not be completed.

At this opportunity, I would like to express my deepest thanks to my main supervisor, Professor Raja Noor Zaliha Raja Abd Rahman, for her exemplary guidance, monitoring, constant encouragement and most importantly, for believing in me. The blessing, help and guidance given by her time to time shall carry me a long way in the journey of life on which I am about to embark. I am also highly indebted to Dr. Mohd Shukuri Mohamad Ali, Dr. Adam Leow Thean Chor and Dr. Fairalniza Sharif for giving me such attention and time in providing necessary information regarding the project.

I would like to express my biggest gratitude towards my parents and my family for the constant support and encouragement. Everything that I have accomplished has been because of your help. I couldn't ask for a better team on my side. To my husband thank you for being my number one supporter and helping me to go on when I sometimes wanted to quit.

My thanks and appreciations also go to my labmates, especially Randa, Dura, Farhanie, Zarir, Menega, Iffah, Dina, Ana and Arilla who have willingly helped me out with their abilities. Thank you for your valuable assistance and friendship. Thank you, all.

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Master of Science. The members of the Supervisory Committee were as follows:

Raja Noor Zaliha Raja Abd. Rahman, D.Eng

Professor
Faculty of Biotechnology and Biomolecular Sciences
Universiti Putra Malaysia
(Chairman)

Mohd Shukuri Muhamad Ali, PhD

Associate Professor
Faculty of Biotechnology and Biomolecular Sciences
Universiti Putra Malaysia
(Member)

Adam Leow Thean Chor, PhD

Associate Professor
Faculty of Biotechnology and Biomolecular Sciences
Universiti Putra Malaysia
(Member)

BUJANG KIM HUAT, PhD

Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 25 June 2015

Declaration by the Graduate Student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any institutions;
- intellectual property from the thesis and the copyright of the thesis are fully-owned by Universiti Putra Malaysia, as stipulated in the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from the supervisor and the office of the Deputy Vice-Chancellor (Research and innovation) before the thesis is published in any written, printed or electronic form (including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials) as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld in accordance with the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2015-2016) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software

Signature: _____ Date: _____

Name and Matric No.: Izatul Azira binti Nor Azman

Declaration by Members of the Supervisory Committee

This is to confirm that:

- the research and the writing of this thesis were done under our supervision;
- supervisory responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2015-2016) are adhered to.

Signature: _____
Name of Chairman
of Supervisory
Committee: _____

Signature: _____
Name of Member of
Supervisory
Committee: _____

Signature: _____
Name of Member of
Supervisory
Committee: _____

TABLE OF CONTENTS

		Page
ABSTRACT		i
ABSTRAK		iii
ACKNOWLEDGEMENTS		v
APPROVAL		vii
DECLARATION		viii
LIST OF TABLES		xiii
LIST OF FIGURES		xiv
LIST OF ABBREVIATIONS		xviii
CHAPTER		
1	INTRODUCTION	1
	1.1 Problem statement	3
	1.2 Objective	3
2	LITERATURE REVIEW	4
	2.1 <i>Bacillus</i>	4
	2.1.1 <i>Bacillus Pumillus</i>	4
	2.1.2 <i>Bacillus pumillus</i> 115b	4
	2.2 Proteases	4
	2.2.1 Definition of proteases	4
	2.2.2 Physical and chemical properties of proteases	5
	2.3 Classification of proteases	5
	2.4 Application of proteases	5
	2.4.1 Detergent	5
	2.4.2 Leather industry	5
	2.5 Recombinant DNA Technology	6
	2.6 Vector DNA	6
	2.7 Plasmid DNA as a Vector	6
	2.8 Protein Expression	6
	2.9 <i>Escherichia coli</i> Expression System	7
	2.10 Cloning and Expression of Protease Gene	7
	2.11 Protein purification	7
	2.12 Protein crystallization	8
	2.12.1 Important steps in crystallization of proteins	9
	2.12.2 Methods in protein crystallization	9
	2.12.2.1 Vapour Diffusion method	9
	2.12.2.2 Batch Experiment	10
	2.13 Bioinformatics	10
	2.14 Protein Structure Prediction	10
	2.14.1 <i>Ab initio</i> protein modelling	10
	2.14.2 Homology modeling	11
	2.14.3 Protein threading	12

3	MATERIALS AND METHODS	14
3.1	Materials	14
	3.1.1 List of media and chemicals	14
	3.1.2 Equipments	15
	3.1.3 Kits	16
3.2	Preparation of media, buffers and solutions	16
3.3	Preparation of 115B protease gene	18
	3.3.1 Microorganism and culture condition	18
	3.3.2 Plasmid extraction	18
3.4	Cloning of 115B protease gene into expression vectors	18
	3.4.1 Amplification of 115B gene	18
	3.4.2 Purification of the PCR product	19
	3.4.3 Digestion of purified PCR product and plasmid	19
	3.4.4 Preparation of competent cells	19
	3.4.5 Ligation of purified digested vector DNA and PCR product	19
	3.4.6 Transformation	20
3.5	Screening and analysis of positive <i>E.coli</i> transformants	20
	3.5.1 Colony PCR	20
	3.5.2 Restriction enzyme analysis	20
	3.5.3 Sequencing of the 115B protease gene	20
3.6	Secretory expression of recombinant 115B protease gene	21
	3.6.1 Screening for clones with high-level secretion of 115B protease	21
	3.6.1.1 Expression by ptrcHis 2/115B gene construct	21
	3.6.1.2 Expression by pET 32a expression systems	21
	3.6.2 Optimization of secretory expression of 115B protease	21
	3.6.2.1 Time course analysis	22
	3.6.2.2 Effect of inducer (IPTG) concentration	22
3.7	Purification of recombinant 115B protease	22
3.8	Determination of protein concentration	22
3.9	Polyacrylamide gel electrophoresis	23
3.10	Determination of protease activity	23
3.11	Characterization of 115B protease	23
	3.11.1 Optimal temperature for enzyme activity	23
	3.11.2 Thermostability of the enzyme	23
	3.11.3 Optimal pH for 115B protease	24
	3.11.4 pH stability for 115B protease	24
	3.11.5 Enzyme inhibitor	24
3.12	Crystallization of 115B protease	24

3.12.1	Screening of crystallization condition	24
3.12.2	Seeding	25
3.12.3	Salt concentration	25
3.13	X-ray diffraction	25
3.14	Prediction of 115b protease 3D structure	25
3.15	Validation and analysis of modelled structure	26
3.15.1	Verify 3-D, Errat and Procheck	26
3.15.2	PDB Sum	26
4	RESULTS AND DISCUSSION	27
4.1	Growth of <i>E.coli</i> harboring the 115B protease gene	27
4.2	Amplification of 115B protease gene	27
4.3	Cloning of <i>Bacillus pumilus</i> 115B protease gene into expression vectors	29
4.3.1	Expression with pTrcHis	31
4.3.2	Expression with pET 32a system	33
4.4	Protein purification of 115B protease	38
4.5	Characterization of 115B protease	41
4.5.1	Effect of temperature on protease activity and stability	41
4.5.2	Effect of pH on protease activity and stability	43
4.5.3	Effect of inhibitor on protease activity	45
4.6	Crystallization of 115B protease	46
4.6.1	Crystal's optimization strategy	47
4.6.2	Effect of salt/precipitant concentration on 115B protease crystallization	50
4.7	Homology Modelling	54
4.7.1	Template searching and sequence alignment	54
4.7.2	Modelling of 115B protease	55
4.7.3	Analysis of 115B predicted model	58
4.8	Evaluation of the predicted 3D structure	59
5	CONCLUSION AND RECOMMENDATIONS	63
	REFERENCES	65
	BIODATA OF STUDENT	71

LIST OF TABLES

Table		Page
1	List of preparation of media, buffers and solution	17
2	Program and world wide web server use in evaluating the model	26
3	The corresponding Z-scores of the predicted 115B protease generated by YASARA are listed in the table below:	56



LIST OF FIGURES

Figure		Page
1	Step in comparative protein structure modeling	12
2	Plasmid extraction of recombinant pQE-30/115B. Note: The extracted plasmid pQE-30/115B was electrophoresed on 1% (w/v) of agarose gel. Lane M: Gene Ruler™ 1kb DNA ladder; Lane 2: Extracted plasmid of pQE -30/115B.	28
3	Amplification of the 115B protease gene from plasmid pQE-30/115B by PCR. Note: The PCR product was electrophoresed on 1% (w/v) of agarose gel. Lane M: Gene Ruler™ 1kb DNA ladder; Lane 2: 1.1kb amplified gene of 115B protease.	28
4	Nucleotide and deduced amino acid sequence of 115B protease gene consist of 29 putative presequence, 79 putative prosequence and 275 putative catalytic domain.	30
5	Schematic diagram of pTrcHis cloning vector	31
6	Analysis of recombinant plasmid pTrcHis/115B harbouring 115B protease gene. Note: A) Recombinant plasmid pTrcHis/115B. Lane M: Gene Ruler™ 1kb DNA ladder; Lane 2: Undigested recombinant plasmid; Lane 3: Single digested recombinant plasmid with EcoR1. B) Amplified PCR product from recombinant plasmid pTrcHis/115B. Lane 1: Gene Ruler™ 1kb DNA ladder; Lane 2: PCR product of open reading frame of 115B protease gene (~1.1kb).	32
7	Schematic diagram of pET 32 a cloning vector	33
8	Production of protease by recombinant 115B/pET32a on LB-SMA. Note: The single colony was streaked on LB-skim milk agar (1%) plate containing ampicillin (50 µg/mL). The plate was then incubated at 37 °C.	34
9	Detection of PCR product of pET32a/115B protease. Lane M: Gene Ruler™ 1kb DNA ladder; Lane 1-2: Amplified PCR product from positive colony.	35
10	Analysis of recombinant plasmid pET 32a/115B through double digestion. Lane M: Gene Ruler™ 1kb DNA ladder;	35

Lane 1: Digested plasmid; Lane 2: Digested empty recombinant plasmid.

- 11 Sequence alignment between 115B/pET32 a sequencing result and nucleotide and deduced amino acid of organic solvent protease obtain from NCBI genebank shown 99.2% identity. 36
- 12 SDS-PAGE analysis of total protein, excreted by the *E.coli* BL21 (DE3) cells harboring 115B protease gene at optimum growth condition. Note: Lane 1: Marker; Lane 2: The organic solvent tolerant 115b protease gene that was expressed by induction with 0.5mM IPTG, 8 hours after induction time and was detected by SDS-PAGE analysis with a molecular weight of around 42kDa; Lane 3: Control. 37
- 13 Affinity chromatography purification of 115B protease. A) Immobilized metal affinity chromatography of His-tagged recombinant 115B protease; B) SDS-PAGE of ~33 kDA His-tagged recombinant 115B (partially purified). Note: Lane M: Marker; Lane 2-9: 115B protease band. 39
- 14 Ion exchange chromatography purification of 115b protease. A) Ion exchange chromatography of purified recombinant 115B protease; B) SDS-PAGE of approximately 33.5 kDA recombinant 115B. Note: Lane M: Marker, Lane 1-2: Purified 115B protease; C) Protease activity staining. Note: Lane M: Marker, Lane 2 Staining activity of 115B protease. 40
- 15 Effect of temperature on enzyme activity. Note: The reaction was carried out at each temperature for 30 minutes. Error bars represent standard deviation of triplicate determinations of three independently repeated experiments (n=3). 42
- 16 Effect of temperature on protease stability. Note: The reaction was carried out at each temperature for 30 minutes. Error bars represent standard deviation of triplicate determinations of three independently repeated experiments (n=3). 42
- 17 Effect of pH on enzyme activity. Note: The substrate azocasein was dissolved in different pH of buffer as follows: (◆) 20mM sodium acetate (pH 4.0-6.0), (■) 20mM Potassium phosphate (pH 6.0-8.0), (▲) 20mM Tris-HCl (pH 8.0-9.0), (●) 20mM Glycine NaOH (pH 9.0-11.0), and (■) 20mM Disodium dihydrogen phosphate (pH 12.0- 44

13.0). Error bars represent standard deviation of triplicate determination of three independently repeated experiments (n=3)

18	pH stability profile of 115B protease. Note: The enzyme solution was pre-incubated at 50°C for 30 minutes. The residual activity was measured at pH 8.0. Error bars represent standard deviation of triplicate determination of three independently repeated experiments (n=3).	45
19	Effect of inhibitors on 115B protease activity. Note: The enzyme was preincubated with the inhibitors for 30 minutes in Tris-HCl buffer pH 8.0 and the residual activity was assayed. Error bars represent standard deviation of triplicate determination of three independently repeated experiments (n=3).	46
20	Crystal of 115B protease in formulation composed of 2% v/v PEG 400, 2M Ammonium sulphate. (Crystal screen 1, Hampton Research) with elastase seed.	48
21	Crystals of 115B protease in formulation composed of 1.0M ammonium sulphate, 0.1M Tris pH 8 with elastase seed.	49
22	Microcrystals of 115B protease in formulation composed of 0.2M Ammonium sulphate, 0.1 M Sodium cacodylate trihydrate pH 6.5, 30% w/v Polyethylene glycol 8,000 with W200R seed.	49
23	Crystals of 115B protease. Formulation composed of 0.2M ammonium sulphate, 0.1M Tris pH 8 with elastase seed.	50
24	Crystal of 115B protease. The formulation composed of 0.5M ammonium sulphate, 0.1M Tris pH 8 with elastase seed.	51
25	Crystal of 115B protease. The formulation composed of 1M ammonium sulphate, 0.1M Tris pH 8 with elastase seed	51
26	Crystals of 115B protease in formulation composed of 1.5M ammonium sulphate, 0.1M Tris pH 8 with elastase seed	52
27	Precipitation and microcrystals of 115B protease in formulation composed of 2.0M ammonium sulphate, 0.1M Tris pH 8 with elastase seed.	52

28	Crystals in the capillary using the microseeding counter-diffusion method	53
29	Crystal of 115B showed blue appearance after IZIT dye was applied indicating protein crystal. Note: The protein crystal above was the crystal that subjected to X-ray diffraction. The formulation composed of 0.5M ammonium sulphate, 0.1M Tris pH 8 with elastase seed.	53
30	X-ray diffraction pattern of 115B crystal in formulation composed of 0.5M ammonium sulphate, 0.1M Tris pH 8 with elastase seed. Note: 3D View showed the 3d pattern of diffraction spots.	54
31	Result of template searching and alignment for 115B protease using Basic Local Alignment Search Tool (BLAST). Note: The chosen template is highlighted in red.	55
32	3D model of 115B protease predicted by homology modeling approach using YASARA software based on the crystal structure of subtilisin BPN complex. Note: 115B protease contains a catalytic triad, Asp 32, His 64 and Ser 221 (stick and ball) similar with subtilisin BPN.	57
33	The proposed open topology structure of 115B protease generated by PDB Sum shows 9 α -helices and 9 β -sheets. Note: Cylinder;helix, Arrow;beta strand.	58
34	The calcium binding site present in predicted 115B structure. Note: The 'calcium A' binding site present in 115B predicted structure is located at same residues as in subtilisin BPN. The predicted ligand metal interaction for Calcium ion in 3D structure of 115B protease was generated by PDB Sum.	59
35	Ramachandran plot of the phi and psi distribution of predicted 115B protease structure produced by PROCHECK	60
36	Verify_3D server structure analysis, shown 94.27% of the residues had an average 3D-1D profile score > 0.2 in predicted 115B structure.	61
37	Evaluation of 115B protease using Errat showed an overall quality factor of 81.602%.	62

LIST OF ABBREVIATIONS

bp	Base pair
°C	Degree centrigade
dH ₂ O	Distilled water
g	Gram
g/L	Gram per liter
h	Hour
kDa	Kilo Dalton
L	Liter
M	Molar
mM	Milimolar
mL	Mililiter
min	Minute
nm	Nanometer
MW	Molecular weight
ml/min	Mililiter per minute
mg/ml	Miligram per milliliter
PAGE	Polyacrylamide gel electrophoresis
pI	Isoelectric point
rpm	Revolutions per minute
SDS	Sodium dodecyl sulphate
%	Percentage
µg	Microgram
µg/mL	Microgram per mililiter

μl	Microliter
μm	Micrometer
v/v	Volume per volume
w/v	Weight per volume



CHAPTER 1

INTRODUCTION

A protease is degradative enzyme that catalyzes hydrolysis of protein by breaking down the peptide bonds that link amino acids together in the polypeptide chain known as proteolysis. Proteases present naturally in all organisms and plays various role in physiological enzymatic reactions (Ray *et al.*, 2004).

Proteases are the most important kind of industrial enzymes (Joo *et al.*, 2002) and constitute approximately half of the total sale of industrial enzymes in the world market. Proteases were extensively use in various industrial sectors and the value in the world market is predicted to continuously grow and reached \$6 billion by 2011 (Rao *et al.*, 1998; David *et al.*, 2009).

(Gupta *et al.*, 2002) reported that detergent alkaline protease which is active and stable in the alkaline pH range dominate the enzyme market. Alkaline proteases are predominantly utilized as cleansing additive in detergents because they able to withstand the high pH conditions that occurs in washing environment. As an important industrial enzyme, alkaline protease especially subtilisin was extensively studied by scientist worldwide. Abo-Aba *et.al* (2006) reported that, there were more than half of the total industrial enzymes were extracted from genetically engineered microorganisms.

In the industrial processes, it is not practical to harvest the enzymes by isolation and purification from naturally occurring microorganism. The maintenance of wild type bacteria grow in bioprocessing fermenters is tedious and costly, as it requires special equipment suitable with its natural growth condition. The advances of molecular biology techniques have made it possible to produce enzymes in *E.coli* which is easier to maintain and handle. Genetic engineering greatly improves productivity and cost effectiveness in existing processes.

Methods for expressing large amount of protein from cloned gene introduced into new host such as *Escherichia coli* or *Bacillus* have proven invaluable in the purification, localization and also functional analysis of the proteins. For example, fusion proteins consisting of amino-terminal peptides encoded by a portion of the *E.coli* lacZ linked to eukaryotic proteins have been used to prepare polyclonal and monoclonal antibodies against these proteins. These antibodies have been used to purify proteins by immunoaffinity chromatography, in diagnostic assays to quantitate the levels of protein and to localize the proteins in organisms,tissues and individual cells by immunofluorescence.

Intact native proteins have also been produced in *E.coli* in large amount for functional studies. For example, both prokaryotic and eukaryotic DNA-binding proteins produced using *E.coli* expression vectors have been used to study the role of these proteins in gene expression.

Previously, the properties of target protein were determined by means of protein characterization. In this new era where biological system was studied at molecular level, X-ray crystallography are now used routinely by the scientist to understand functions of protein by determining their 3 dimensional structures. Protein crystallography are important in order to understand the functional inferens and the establishment of the biochemical pathway such as how a pharmaceutical drug interacts with its protein target and what changes might improve it (Scapin, 2006).

Crystal structures of proteins (which are irregular and hundreds of times larger than cholesterol) began to be solved in the late 1950s, beginning with the structure of sperm whale myoglobin by Max Perutz and Sir John Cowdery Kendrew, for which they were awarded the Nobel Prize in Chemistry in 1962 (Kendrew *et al.*,1958). Since that success, over 48970 X-ray crystal structures of proteins, nucleic acids and other biological molecules have been determined.

For comparison, the nearest competing method, nuclear magnetic resonance (NMR) spectroscopy has produced 7806 structures. Moreover, crystallography can solve structures of arbitrarily large molecules, whereas solution-state NMR is restricted to relatively small ones (less than 70 kDa). However, intrinsic membrane proteins remain challenging to crystallize because they require detergents or other means to solubilize them in isolation, and such detergents often interfere with crystallization. Such membrane proteins are a large component of the genome and include many proteins of great physiological importance, such as ion channels and receptors (Lundstrom K, 2006).

Protein structure prediction is divided into three parts depending on the similarity of the target to proteins of known structure. In comparative modeling, one or more template proteins of known structure with high sequence homology to the target sequence are identified. The target and template sequences are aligned, and a three-dimensional structure of the target protein is generated from the coordinates of the aligned residues of the template protein, combined with models for loop regions and other unaligned segments. Ideally, this three-dimensional model would then be refined to bring it closer to the true structure of the target protein.

Second, if no reliable template protein can be identified from sequence homology alone, the prediction problem is denoted as a fold recognition problem. Here, the primary goal is to identify one or more template protein structures that are consistent with the target sequence, that is, template folds that the target sequence might plausibly adopt. The subsequent protocol is similar to that of

comparative modeling: align the sequences and compose a three-dimensional model from the alignment.

Third, if no template structure can be identified with confidence, the target sequence may be modeled using *de novo* (or new fold) prediction methods. An outstanding question in *de novo* prediction is whether such methods can predict structures to a resolution useful for biochemical applications (Schonbrun et al., 2002).

This project was initiated to get better understanding of the structure and function of the enzyme thus molecular cloning, expression, characterization, crystallization and homology modelling of the enzyme is compulsory.

1.1 Problem statement

1. The extensive application of protease as industrial biocatalysts requires a further effort to understand the real potential of 115B protease.

1.2 Objective

1. To clone and express gene encoding organic solvent tolerant protease in suitable vectors
2. To purify and characterize 115B protease
3. To crystallize 115B protease using sitting drop method
4. To predict the structure of 115B protease.

REFERENCES

- Ahmed H., 2005. Principles and reactions of protein extraction, purification and characterization. CRC Press.
- Abo-Aba SE, Soliman EA, Nivien A, 2006. Enhanced production of extracellular alkaline protease in *Bacillus circulans* through plasmid transfer. *Research journal of Agriculture and Biological Sciences*, **2(6)**: 526-530.
- Abrahmsen, L., Tom, J., Burnier, J., Butcher, K. A., Kossiakof, A. and Wells, J.A., 1991. Engineering subtilisin and its substrates for efficient ligation of peptide bonds in aqueous solution. *Biochemistry* **30**:4151-4159.
- Ali, A., Majeed, H., and Abdulhusain, N. 2011. Molecular Cloning and Expression of *Bacillus stearothermophilus* Protease Gene in *Escherichia coli* *Journal of Biology and Life Sciences* **2** :26-31.
- Altschul, SF., Madden TL., Schaffer AA., Zhang J., Miller W. and Lipman DJ, 1997 Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* **25**:3389-3402.
- Aono, R., Itoh, M., Inoue, A. and Horikoshi, K, 1992. Isolation of novel toluene tolerant strain *Pseudomonas aeruginosa*. *Bioscience Biotechnology and Biochemistry* **56**; 145-146.
- Becker, T. and Lawlis V., 1991. Subtilisin crystallization process. United States patent, 5041337.
- Berman, Helen M., et al. "The Protein Data Bank." *Nucleic Acids Research*. **28**: 235-242 (2000).
- Bergfors, T. 2003. Seed to crystals. *Journal of Structural Biology*, **142**; 66-76.
- Bradford, M.M., 1976. A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein dye binding. *Analytical Biochemistry* **72**:248-254.
- Branden, Carl and John Tooze., 1999. Introduction to Protein Structure. New York: Garland (pp. 374-376).
- Cappucino, J., G and Sherman, N, 1992. Microbiology, A laboratory manual. Benjamin/Cumming Publishing Co.
- Chayen, N.E. *Structure* (1997)**5**,1269-1274.
- Chayen, N. E., Shaw Stewart, P. D., and Blow, D. M. 1992. Microbatch crystallization under oil--a new technique allowing many small-volume crystallization trials. *Journal of crystal growth* **122**:176-180.

- Colovos, C. and Yeates, T.O. 1992. verification of protein structures: patterns of nonbonded atomic interaction. *Protein Science* **2**:1511-1519.
- D'Arcy, A.; Villard, F.; Marsh, M. *Acta Crystallogr*, **2007**, *D63*, 550-554.
- David, L., Vierros, M., Hamon, G., Arico, S. and Monagie, C., 2009. Marine genetic resources: a review of scientific and commercial interest. *Marine Policy* **33**:183-194.
- Davidson, V. and Sittiman (1999). *Enzymes In Biochemistry*, 4th edition. Davidson, V.I, and Sittiman, D.B., Lippincott-Raven. Publishers, Philadelphia, PA, pp:59-76.
- DeSantis. and Jones JB. (1999). Chemical modification of enzymes for enhanced functionality. *Current Opinion in Biotechnology*. **13**:324-330.
- Douguet, D., Bolla, J., Munier-Lehmann, H. and Labesse, G. 2002. From sequence to structure to function : a case study. *Enzyme and Microbial Technology* **30**:6659-6667.
- Drenth J., 1999. Principles of Protein X-Ray Crystallography. Second Edition, Springer-Verlag, New York.
- Edwards, A.M., Arrowsmith, C.H., Christendat, D., Dharamsi, A., Friesen, J.D., Greenblatt, J.F., and Vedadi, M. 2000. Protein production: Feeding the crystallographers and NMR spectroscopists. *Natural Structural Biology*. **7**: 970–972.
- Falzon, L., Patel, S., Chen, Y.J. and Inouye, M., 2007. Autotomic behavior of the propeptide in propeptide-mediated folding of prosubtilisin E. *Journal of Molecular Biology* **366**:494-503.
- Fiser, A. , Sanchez R., Melo F. and Sali A. (2001). Comparative protein structure modeling. *Computational Biochemistry and Biophysics*. 275-312.
- Georgiou, G. 1996. Expression of proteins in bacteria. In Protein Engineering: Principles and Practice, eds. J.L. Cleland and C.S. Craik, pp 101-127. New York: A John Wiley and Sons, Inc.
- Gilliland, G.L., Howard, A.J., Winbornes, E.L., Poulos, T.L., Stewart, D.B. and Durham, D.R., 1987. Crystallization and preliminary X-ray diffraction studies of Subtilisin GX from *Bacillus* sp. GX 6444. *The Journal of Biological Chemistry* **262**:4280-4283.
- Gold, A.M and Fahrney, D., 1964. Sulfonyl fluorides as inhibitors of esterase II. Formation and reactions of phenylmethanesulfonyl α -chymotrypsin. *Journal of Biochemistry* **3**: 783-791.
- Gupta, A. and Khare, S.K., 2006. A protease stable in organic solvents from solvent tolerant strain of *Pseudomonas aeruginosa*. *Bioresource Technology* **97**: 1788-1793.

- Gupta, A. Roy, I., Khare, S.K. and Gupta, M.N. 2005. Purification and characterization of a solvent stable protease from solvent tolerant *Pseudomonas aeruginosa* PseA. *Enzyme and Microbial Technology* **42**; 11-16.
- Gupta, R., Beg, Q.K. and Lorenz, P. 2002. Bacterial Alkaline Proteases: molecular approaches and industrial applications. *Applied Microbiology and Biotechnology* **32**: 18-27
- Huang, Q., Peng, Y., Li, X., Wang, H., and Zhang, Y. (2003). *Current Microbiology*, **46**,169-173).
- Illanes, A., 2008. Enzyme biocatalysis: Principles and applications. Lightning Source UK Ltd, United Kingdom.
- Jacobs, M, Elliasson, M., Uhlen, M. and Flock, J.I., 1985. Cloning and sequencing and expression of subtilisin Carlsberg from *Bacillus licheniformis*. *Nucleic Acids Research* **13**: 8913-8926.
- Jaoudi, B., Chaabouni, S. E., Rhimi, M. and Bejar, S., 2008. Biochemical and molecular characterization of a detergent-stable serine alkaline protease from *Bacillus pumillus* CBS with high catalytic efficiency. *Biochimie* **90**-1291-1305.
- John G. Holt (1986). "Gram-positive *Bacteria* other than *Actinomycetes*". *Bergey's Manual of Systematic Bacteriology* (1st Edition). Baltimore MD.
- Joo, HS., Kumar CG., Park GC., Kim KT., Paik SR. and Chan CS. (2002). Optimization of the production of an extracellular alkaline protease from *Bacillus horikoshii* *Process Biochem.* **37**:139-144
- Kendrew J. C. (1958). "A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis". *Nature* **662**.
- Klibanov, A.M, 1997. Why are enzymes less active in organic solvents than in water? *Trends in Biotechnology* **15**; 97-101.
- Koops, B. C, H. M Verheij, A. J Slotboom and M.R Egmond, 1999. Effect of chemical modification on the activity of lipases in organic solvents. *Enzyme and Microbiol Technology*, **25**: 622-631.
- Laskowski, R.A., MacArthur, M.W., Smith, D.K., Jones, D.T., Hutchinson, E.G., Morris, A.L., Naylor, D., Moss, D. and Thornton, J.M. 1994. Procheck v.3.5.4: Operating Manual.
- Li, Y., Hu, Z., Jordan, F. and Inouye, M., 1995. Functional analysis of propeptide of subtilisin E as an intermolecular chaperone for protein folding. Refolding and inhibitory abilities of propeptide mutants. *Journal of Biological Chemistry* **270**: 25127-25132.

- Lovell, S.C., Davis, I.W., Arendall, W.B., de Bakker, P.I., Word, J.M., Prisant, M.G. and Richardson, D.C. 2003. Structure validation by C-alpha geometry: phi, psi and C-beta deviation. *Proteins* **50**:437-450
- Lundstrom K (2006). "Structural genomics for membrane proteins". *Cell. Mol. Life Science* (**22**): 2597.
- Luthy R, Bowie JU, Eisenberg D. 1992. Assessment of protein models with three-dimensional profiles. *Nature* **356**:83-85.
- Madigan M., Martinko J (editors). (2005). Brock Biology of Microorganisms 11th edition. Prentice Hall.
- Marti-Renom MA, Stuart AC, Fiser A, Sanchez R, Melo F, Sali A. (2000). Comparative protein structure modeling of genes and genomes. *Annu Rev Biophys Biomol Struct* **29**: 291-325.
- McRee, Duncan E. *Practical Protein Crystallography*. San Diego: Academic Press, 1993 (pp. 1-23).
- Miyaji, T., Otta, Y., Nakagawa, T., Watanabe, T., Niimura, Y. and Tomizuka, N., 2006. Purification and molecular characterization of subtilisin-like alkaline protease BPP-A from *Bacillus pumillus* strain MS-1. *Letter of application Microbiology* **42**:242-247.
- Morihara, K. (1987). Using proteases in peptide synthesis. *Trends in Biotechnology* **5**:164-170.
- Mulligan, M. E., Brosius, J., and Clure, W.R(1985). Characterization *in vitro* of the Effect of Spacer Length on the activity of *Escherichia coli* RNA polymerase at the tac Promoter. *J.Biol.Chem.* **260**, 3539-3538.
- Neurath, H., 1989. The diversity of proteolytic enzymes. In Beynon, R.J. and Bond, J. S. (eds), *Proteolytic Enzymes: A Practical Approach*. Oxford University Press, New York.
- Rahman RNZA,,Mahamad S Basri M, Salleh AB (2007) A new organic solvent tolerant protease from *Bacillus pumilus* 115b. *Journal of Industrial Microbiology & Biotechnology* **34**:509-517..
- Rao, M., Tankasale, A., Ghatge, M. and Desphande, V. 1998. Molecular and biotechnological aspects of microbial proteases. *Microbiology and Molecular Biology Review* **62**:597-634.
- Ray and Sahelian. M.D. (2004). Protease enzyme. *Health benefit of protease enzyme.* **4**:365-372.
- Reddy, C.S., Vijayasarathy, K., Srinivas, E., Sastry, G.M and Sastry, G.N. 2006. Homology modeling of membrane proteins: A critical assessment. *Computational Biology and Chemistry* **30**:120-126.

- Rosenberger, F., Vekilov, P. G., Muschol, M. & Thomas, B. R. (1996). Nucleation and crystallization of globular proteins-what we know and what is missing. *Journal of Crystal Growth*, **168**, 1-27.
- Rhodes, Gale. *Crystallography Made Crystal Clear*. San Diego: Academic Press, 1993 (pp. 8-10, 29-38).
- S.Damodaran, X.Q.Han. Detergen stable alkaline protease from *B.pumillus* , United States Patent Number 5976859 (1999).
- Sheng Y, Sali A, Herzog H, Lahnstein J, Krilis SA. 1996. Site directed mutagenesis of recombinant human beta 2-glycoprotein I identifies a cluster of lysine residues that are critical for phospholipid binding and anti- cardiolipin antibody activity. *Journal of Immunology* **157**:3744-3751.
- Shinde, U., u, X. and Inouye, M., 1999. A pathway for conformational diversity in proteins mediated by intermolecular chaperones. *Journal of Biological Chemistry* **55**:2251-2258.
- Tan PS (1999) Screening and isolation of polycyclic aromatic hydrocarbons (PAHs) degrading bacteria from contaminated sites. BS Thesis, Universiti Putra Malaysia, Malaysia.
- Tanaka, S., Takeuchi, Y., Matsumura, H., Koga, Y., Takano, K. and Kanaya, S., 2008. Crystal structure of Tk-subtilisin folded without propeptide: Requirement of propeptide for acceleration of folding. *FEBS Letters* **582**:3875-3878.
- Toogood, H.S., Smith C.A ., Baker, E.N and Roy, M.D., 2000. Purification and Characterization of Ak1 protease, a thermostable subtilisin with a disulphide bond in the substrate-binding cleft. *Biochemical Journal* **350**:321-328.
- Ogino, H., Watanabe, ., Yamada, M., Nakagawa, S., Hirose, T., Noguchi, A., Yasuda, M. and Ishikawa. H., 1999. Purification and characterization of organic solvent *Pseudomonas aeruginosa* PST 01. *Journal of Bioscience and Bioengineering* **87**:61-68.
- Porath, J., Carlsson, J., Olsson, I., and Belrage, G., Metal chelate affinity chromatography, a new approach to protein fractionation, *Nature*, 258, 1975.
- Scapin G (2006). "Structural biology and drug discovery". *Curr. Pharm. Des.* (17): 2087.
- Siezen, R.J and Leunissen, J.A.M., 1997. Subtilases: the superfamily of subtilisin-like serine proteases. *Protein Science* **6**:501-523.
- Takahashi, M., Sekine, T., Kuba, N., Nakamori, S., Yasuda, M., and Takagi, H., 2004. The production of recombinant APRP, an alkaline protease

derived from *Bacillus pumilus* TYO-67, by in vitro refolding of proenzyme fixed on a solid surface. *Journal of Biochemistry* **136**:549-556.

Takami, H., Akiba, T. and Horikoshi, K., 1989. Production of extremely thermostable alkaline protease from *Bacillus* sp. *Applied Microbiology and Biotechnology* **30**:120-124.

Tulinsky, A. "The Protein Structure Project, 1950-1959: First Concerted Effort Of a Protein Structure Determination In the U.S." *The Rigaku Journal*. **16** (1999).

Wang, F., Hao, J., Yang, C., and Sun, M. 2010. Cloning, Expression, and Identification of a Novel Extracellular Cold-Adapted Alkaline Protease Gene of the Marine Bacterium Strain YS-80-122. *Applied Biochemistry and Biotechnology* **162** (5):1497-1505.

Zhang, Y. and Skolnick J. (2005). The protein structure prediction problem could be solved using the current PDB library. *Proc Natl Acad Sci USA*. **102**:1029-1034

Zhu, W., D., Cheng, G., Peng, Q. and Shen, P. 2007. Purification and characterization of a thermostable protease from a newly isolated *Geobacillus* sp. YMTC 1049. *Enzyme and Microbial Technology* **40**: 1592-1597.