MOLECULAR AND BIOINFORMATICS
CHARACTERIZATION OF FRUIT BROMELAIN
FROM ANANAS COMOSUS

TUAN NORSYALIEZA BINTI TUAN AZNAN

Master of Science

UNIVERSITI MALAYSIA PAHANG
SUPERVISOR’S DECLARATION

We hereby declare that We have checked this thesis and in our opinion, this thesis is adequate in terms of scope and quality for the award of the degree of Master of Science.

(Supervisor’s Signature)
Full Name : DR AIZI NOR MAZILA BINTI RAMLI
Position : SENIOR LECTURER
Date : 13 AUGUST 2018

(Co-supervisor’s Signature)
Full Name : DR NOORMAZLINAH BINTI AHMAD
Position : SENIOR LECTURER
Date : 13 AUGUST 2018

(Co-supervisor’s Signature)
Full Name : EN HAIRUL BIN ABD RAHIM
Position : LECTURER
Date : 13 AUGUST 2018
STUDENT’S DECLARATION

I hereby declare that the work in this thesis is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at Universiti Malaysia Pahang or any other institutions.

________________________________________
(Student’s Signature)

Full Name : TUAN NORSYALIEZA BINTI TUAN AZNAN
ID Number : MKT 15004
Date : 13 AUGUST 2018
MOLECULAR AND BIOINFORMATICS CHARACTERIZATION OF FRUIT BROMELAIN FROM ANANAS COMOSUS

TUAN NORSYALIEZA BINTI TUAN AZNAN

Thesis submitted in fulfillment of the requirements for the award of the degree of
Master of Science

Faculty of Industrial Sciences & Technology
UNIVERSITI MALAYSIA PAHANG

AUGUST 2018
Dedicated to my parents
Assalamualaikum and greetings. All praises and thanks to Allah (S.W.T), who has guided me to this. With his blessing, I manage to accomplish my Master Degree to the end. I also learn a lot of things along the journey.

Word cannot express my gratitude towards my beloved supervisor, Dr Aizi Nor Mazila Bt Ramli. Her encouragement, assistance and guidance right from the beginning to the end enabled me to develop an understanding of this research. Her patience also means a lot to me. Besides that, I wish to deliver my appreciation to my co-supervisor, Dr. Noormazlinah Binti Ahmad and Mr. Mohd Hairul Bin Ab. Rahim for valuable insights and discussions.

Special thanks go to Faculty of Industrial Sciences & Technology (FIST) laboratory staff, Mrs. Salmah, Mr. Husaini and all of them for their kind assistance and consultation throughout this research.

I also want to convey my regards to my laboratory colleagues who helped me in many ways, particularly Nur Dini Binti Johari and Nurul Haida Bt Idrus. This project would not have been possible without support of many people especially my parents and family. They are always there to support me with their best wishes.

Last but not least, I would also like to deliver my thanks to Institute of Postgraduate Studies (IPS), Central Laboratory, Universiti Malaysia Pahang (UMP) and Lembaga Nenas Malaysia (LPNM) for providing my research study with materials, training and workshop.
ABSTRAK

ABSTRACT

Pineapple scientifically known as *Ananas comosus*, has several available cultivars in Malaysia, including Moris cv, N36 cv, and Sarawak cv. Bromelain has been identified as an active component and a major protease of *A. comosus* and has gained wide acceptance and compliance as a phototherapeutic drug. Although a considerable level of research has been devoted to bromelain from *A. comosus*, less attention has been paid to the fruit bromelain compared to the stem bromelain. Therefore, the purpose of this research is to reveal an in-depth information regarding fruit bromelain from *A. comosus*. Until recently, the three-dimensional (3D) structure of bromelain remained to be elucidated. A comprehensive information on the thorough structural organisation of bromelain is vital for therapeutic application and in the understanding of their role in cells and in other related molecular mechanisms. In this study, the screening of fruit bromelain from the local pineapple cultivars (Morris cv, N36 cv. and Sarawak cv) was implemented, followed by the isolation and cloning of the fruit bromelain from the best cultivar with the highest proteolytic activity for sequence analysis. Additionally, a comparison of the fruit and stem bromelain was performed using bioinformatics tools, including both amino acids and structural comparisons. From the screening results, the highest proteolytic activity (0.8220 U/mL) was observed from the fruit bromelain of Morris cv, followed by N36 cv (0.7695 U/mL) and Sarawak cv (0.6942 U/mL). A gene encoding for pineapple fruit bromelain was successfully isolated from Morris cv. using Reverse Transcription - Polymerase Chain Reaction (RT-PCR) techniques. The amino acid sequence and domain analysis of the fruit and stem bromelains demonstrated several differences and similarities of the cysteine protease family members. Additionally, an analysis of the modelled fruit (BAA21848) and stem (CAA08861) bromelains revealed the presence of unique properties of the predicted structures Cys-148, His-281, Gln-174 and Asn-275 are the catalytic residues of fruit bromelain whereas stem bromelain Cys-147, His-281, His-141 and Asn-302. This play crucial roles in chemical catalysis as general acid/base catalysts. The sequence analysis and structural prediction of the stem and fruit bromelain from *A. comosus*, along with the comparison of both structures provided a new insight on their distinct properties for industrial application. From the analysis, stem bromelain is more hydrophobic than fruit bromelain. The knowledge of the structure of these proteolytic enzymes from *A. comosus* is expected to increase the understanding of their functions and mechanism.
# TABLE OF CONTENT

DECLARATION

TITLE PAGE

DEDICATION

ACKNOWLEDGEMENTS

ABSTRAK

ABSTRACT

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

LIST OF ABBREVIATIONS

CHAPTER 1 INTRODUCTION

1.1 Introduction

1.2 Background of study

1.3 Problem Statement

1.4 Research Objectives

1.5 Scope of Study

CHAPTER 2 LITERATURE REVIEW

2.1 Introduction

2.2 Enzyme

2.3 Proteases
2.4 *Ananas comosus* (Pineapple) 6

2.5 Bromelain 8

2.5.1 Fruit Bromelain 9

2.5.2 Stem Bromelain 10

2.6 Enzymatic analysis of Bromelain 10

2.7 Bioinformatics analysis of Bromelain 11

2.8 Industrial Application of Bromelain 12

2.8.1 Clinical Use of Bromelain 13

2.8.2 Pharmaceutical Care 14

2.8.3 Industrial Use 15

CHAPTER 3 METHODOLOGY 17

3.1 Sample preparation 18

3.2 Chemical and Enzyme 18

3.3 Enzymatic Assay Buffers and Reagents Preparation 18

3.4 Bacterial Growth Preparation 18

3.5 Enzyme Assay of Crude Enzyme 19

3.6 Isolation of mRNA transcript from pineapple fruit 19

3.7 Fruit Bromelain Isolation and Cloning 20

3.7.1 Full-length amplification of fruit bromelain from *A. comosus* via reverse transcription-polymerase chain reaction (RT-PCR) 20

3.7.2 Fruit bromelain amplification via Polymerase Chain Reaction (PCR) 21

3.7.3 Agaarose Gel Electrophoresis 22

3.7.4 Insertion of the target gene into the cloning vector 23

3.7.5 Transformation of recombinant plasmid in cloning host 23

3.6.5.1 Competence cell preparation 23
3.6.5.2 Transformation of competent cell 24
3.7.6 Colony PCR of transformed colonies 24
3.8 Sequencing of DNA 25
3.9 Bioinformatics analysis of fruit bromelain 25
   3.9.1 Data mining 25
   3.9.2 Primary sequence analysis 25
   3.9.3 Secondary sequence analysis 26
   3.9.4 Tertiary sequence analysis 26
   3.9.5 Structural comparison of fruit bromelain with Stem Bromelain 26

CHAPTER 4 RESULTS AND DISCUSSION 27
4.1 Introduction 27
4.2 Sample collection and plant morphological analysis 27
4.3 Enzymatic analysis of crude bromelain from different cultivars 29
4.4 Isolation and cloning of fruit bromelain 30
   4.4.1 Total fruit RNA extraction of *A. comosus* 30
   4.4.2 Total RNA detection using gel electrophoresis 31
   4.4.3 Complementary DNA (cDNA) amplification of fruit bromelain 32
   4.4.4 Ligation of the target gene into the cloning vector 33
   4.4.5 The transformation of competent *E. coli* cells 34
   4.4.6 Screening of the recombinant plasmid DNA using the colony PCR 35
   4.4.7 Plasmid extraction of the recombinant vector and DNA sequencing analysis 36
4.5 Data mining of fruit and stem bromelain 39
4.5.1 Amino acids and predicted domain analysis of fruit and stem bromelain

4.6 Structural comparison of fruit and stem bromelain

4.6.1 Secondary and tertiary structure analysis

4.6.2 Structural analysis of fruit and stem bromelain

4.6.3 Structural characteristics comparison of BAA21848 and CAA08861

CHAPTER 5 CONCLUSION

5.1 Conclusions

5.2 Future prospect

REFERENCES

APPENDIX A

APPENDIX B

APPENDIX C
### LIST OF TABLES

<table>
<thead>
<tr>
<th>Table 3.1</th>
<th>Reagents for the PCR amplification reactions of fruit bromelain</th>
<th>21</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 3.2</td>
<td>The PCR amplification reaction conditions of fruit bromelain</td>
<td>22</td>
</tr>
<tr>
<td>Table 3.3</td>
<td>Reagents for the ligation reaction of fruit bromelain</td>
<td>23</td>
</tr>
<tr>
<td>Table 3.4</td>
<td>Reagents for Colony PCR amplification reactions of selected fruit bromelain transformants</td>
<td>25</td>
</tr>
<tr>
<td>Table 3.5</td>
<td>Colony PCR amplification reaction condition of fruit bromelain transformants</td>
<td>25</td>
</tr>
<tr>
<td>Table 4.1</td>
<td>Morphological features of <em>A. comosus</em> cultivars</td>
<td>28</td>
</tr>
<tr>
<td>Table 4.2</td>
<td>Weight and volume of fruit from each pineapple cultivar with final mass used to get the concentrated juice, 300g</td>
<td>29</td>
</tr>
<tr>
<td>Table 4.3</td>
<td>The mean absorbance (A) and enzyme activity (U/ml) enzyme value of fruit bromelain in different cultivar; Morris, N36 and Sarawak</td>
<td>30</td>
</tr>
<tr>
<td>Table 4.4</td>
<td>The concentration and purity of fruit bromelain, total RNA from sample; F1, F2, F3, F4</td>
<td>31</td>
</tr>
<tr>
<td>Table 4.5</td>
<td>Summary of fruit bromelain sequence from Genbank NCBI</td>
<td>41</td>
</tr>
<tr>
<td>Table 4.6</td>
<td>The clans and family of cysteine protease</td>
<td>45</td>
</tr>
<tr>
<td>Table 4.7</td>
<td>Physicochemical properties of bromelain from <em>A. comosus</em></td>
<td>46</td>
</tr>
<tr>
<td>Table 4.8</td>
<td>Amino acid composition of fruit bromelain</td>
<td>49</td>
</tr>
<tr>
<td>Table 4.9</td>
<td>Amino acid composition of stem bromelain</td>
<td>50</td>
</tr>
<tr>
<td>Table 4.10</td>
<td>Predicted templates for fruit and stem bromelain obtained from different servers along with their percentages of identity and query coverage. 1PCI appeared in the all of the results given and was chosen as the model template for BAA21848 and CAA08861.</td>
<td>58</td>
</tr>
<tr>
<td>Table 4.11</td>
<td>Energy levels; ERRAT plot, VERIFY3D and RAMPAGE score of fruit bromelain and stem bromelain models at different stages</td>
<td>62</td>
</tr>
<tr>
<td>Table A1</td>
<td>The absorbance value of L-Tyrosine at different L-Tyrosine concentration for standard curve</td>
<td>86</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>2.1</td>
<td>Pineapple plant grown in plantation area</td>
<td>8</td>
</tr>
<tr>
<td>3.1</td>
<td>The flowchart of overall experimental procedure</td>
<td>17</td>
</tr>
<tr>
<td>3.2</td>
<td>The powdered fruit bromelain of pineapple, <em>A. comosus</em> from Morris cv. after grounded in liquid nitrogen</td>
<td>20</td>
</tr>
<tr>
<td>4.1</td>
<td>The crown (leafy part on top) and the fruit cross section of pineapple cultivars (<em>A. comosus</em>) (a) Morris (b) N36 and (c) Sarawak found in Malaysia.</td>
<td>28</td>
</tr>
<tr>
<td>4.2</td>
<td>Agarose gel 0.8% (w/v) of fruit bromelain, RNA fragment; F3 (Lane 1) and F4 (Lane 2)</td>
<td>32</td>
</tr>
<tr>
<td>4.3</td>
<td>Determination of the cDNA amplified fragment of fruit bromelain; Lane 1, M: 1kb DNA Ladder, Lane 2 and Lane 3 represent F3 and F4 amplified cDNA, respectively</td>
<td>33</td>
</tr>
<tr>
<td>4.4</td>
<td>Blue-white screening of the fruit bromelain transformed colonies on LB agar plate supplemented with ampicilin, IPTG and X-Gal. Colony; White: Successful transformation, Blue: Unsuccessful transformation</td>
<td>35</td>
</tr>
<tr>
<td>4.5</td>
<td>Determination of the plasmid integration with fruit bromelain via Colony PCR in pGEM-T Easy vector. Lane 1, M: 1kb DNA ladder, PCR colony amplified products (Lane 2: F1, Lane 3: F2, Lane 4: F3, Lane 5: F4, Lane 6: F5 and Lane 7: F6).</td>
<td>36</td>
</tr>
<tr>
<td>4.6</td>
<td>The multiple sequence alignment of the isolated fruit bromelain from moris cv. with the nucleotide sequence (D14058) obtained from the NCBI database. The coding sequence substitution of fruit bromelain from Morris cv was indicated in the box, where adenine (A) was replaced by guanine (G) at the same location.</td>
<td>38</td>
</tr>
<tr>
<td>4.7</td>
<td>Domain organization of (a) fruit bromelain and (b) stem bromelain</td>
<td>51</td>
</tr>
<tr>
<td>4.8</td>
<td>Alignments of amino acid sequence of fruit bromelain and stem bromelain. The conserved motif residues of GXNXFXD are indicated with box</td>
<td>53</td>
</tr>
</tbody>
</table>
The sequence alignment of stem (ADY68475.1, P14518.1, BAA21849.1 and CAA08861.1) and fruit bromelain (OAY62650, OAY65848, OAY71019, OAY68894, OAY68854, OAY85858, AEH26024, AGS78388, OAY68270, BAA21929 and BAA21848) sequences from the NCBI Genbank Database. The catalytic residues of Asn (N), Cys (C), His (H) and Gln (Q) were boxed.

Three dimensional model of fruit bromelain, BAA21848 (a) and stem bromelain, CAA08861 (b) with the β-strands, S1 to S6 (yellow) and the α-helices, H1 to H10 (red).

Substrate-binding site (a) fruit bromelain, BAA21848 (b) stem bromelain, CAA08861 using 1PCI as a template. Green region indicates domain I29 while orange region indicates domain peptidase C1. The catalytic amino acids of both models are represented as sticks.

The 3D superimposition structure of BAA21848 and CAA08861 with 1CVZ. (a) The catalytic binding site (Cys-148, Gln-174, His-281 and Asn-302) of BAA21848 (green) with 1CVZ (purple). (b) The catalytic binding site (Cys-147, Gln-173, His-281 and Asn-302) of CAA08861 (green) with 1CVZ (purple).

The 3D superimposition structure of BAA21848 and CAA08861 with the catalytic site of 1CVZ. (a) The catalytic binding site (Cys-25, Gly-66, Trp-69, Asp-158, His-159 and Trp-177) of 1CVZ (blue) after being superimposed with the catalytic site (Cys-148, Gly-188, Asp-191, His-281, Asn-280 and Trp-304) of BAA21848 (green) (b) The catalytic binding site (Cys-25, Gly-66, Trp-69, Asp-158, His-159 and Trp-177) of 1CVZ (blue) after being superimposed with the catalytic site (Cys-147, Gly-187, His-281, Asn-190, Asn-280 and Trp-304) of CAA08861 (green).

Hydrophobic amino acid of papain in space-fill model (a) Fruit bromelain (blue) (b) Stem bromelain (yellow). Colored black in
a space-fill model are the backbone oxygen and nitrogen of the residues with the hydrophobic side chain.

Figure 4.15 Protscale output for hydropathy analysis (a) fruit bromelain (b) stem bromelain

Figure A1 The standard curve of enzyme assay using different concentration of L-Tyrosine

Figure C1 The selected colony of Dh5α with fruit bromelain gene; F1, F2, F3, F4, F5 and F6
LIST OF SYMBOLS

$R^2$  Coefficient of Determination

Kg  Kilogram

kDa  kilo-Dalton

$\mu$mol  micromoles

$\mu$L  Microliter

$\mu$M  Micrometer

mg  milligram

mL  millilitre

nm  nanometer

Ng  Nanogram

%  Percentage

°C  Celcius

$\alpha$  Alpha

$\beta$  Beta

Å  Ångström

pI  Isoelectric point
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Absorbance</td>
</tr>
<tr>
<td>A. comosus</td>
<td><em>Ananas comosus</em></td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic Local Alignment Tool</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary Deoxyribonucleic Acid</td>
</tr>
<tr>
<td>cv.</td>
<td>Cultivar</td>
</tr>
<tr>
<td>DEPC</td>
<td>Diethyl pyrocarbonate</td>
</tr>
<tr>
<td>DDBJ</td>
<td>DNA Data Bank of Japan</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic Acid</td>
</tr>
<tr>
<td>dNTP</td>
<td>Deoxynucleotide</td>
</tr>
<tr>
<td>ENA</td>
<td>European Nucleotide Archive</td>
</tr>
<tr>
<td>F-C</td>
<td>Folin Ciocalteu</td>
</tr>
<tr>
<td>GRAS</td>
<td>Generally Regarded As Safe</td>
</tr>
<tr>
<td>IDT</td>
<td>Integrated DNA Technology</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl β-D-1-thiogalactopyranoside</td>
</tr>
<tr>
<td>LB</td>
<td>Luria Bertani</td>
</tr>
<tr>
<td>LPNM</td>
<td>Lembaga Perindustrian Nanas Malaysia</td>
</tr>
<tr>
<td>MPIB</td>
<td>Malaysian Pineapple Industrial Board</td>
</tr>
<tr>
<td>Min</td>
<td>Minutes</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger Ribonucleic Acid</td>
</tr>
<tr>
<td>NCBI</td>
<td>National Center for Biotechnology Information</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear Magnetic Resonance</td>
</tr>
<tr>
<td>NSAID</td>
<td>Nonsteroidal Anti-inflammatory Drugs</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>PDB</td>
<td>Protein Database Bank</td>
</tr>
<tr>
<td>pNA</td>
<td>p-nitroalanine</td>
</tr>
<tr>
<td>RIN</td>
<td>RNA Integrity Number</td>
</tr>
<tr>
<td>RMSD</td>
<td>root-mean-square deviation</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>rpm</td>
<td>Revolutions per minute</td>
</tr>
<tr>
<td>rRNA</td>
<td>Ribosomal ribonucleic acid</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>Reverse transcription polymerase chain reaction</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>SPDBV</td>
<td>Swiss Protein Database Bank Viewer</td>
</tr>
<tr>
<td>S</td>
<td>Seconds</td>
</tr>
<tr>
<td>TA</td>
<td>Thymine Adenine</td>
</tr>
<tr>
<td>TAE</td>
<td>Tris-acetate</td>
</tr>
<tr>
<td>TCA</td>
<td>Trichloroacetic Acid</td>
</tr>
<tr>
<td>V</td>
<td>Volts</td>
</tr>
<tr>
<td>X-Gal</td>
<td>5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside</td>
</tr>
</tbody>
</table>
REFERENCES


Ramli, A. N. M. (2012). Molecular, Enzymatic and Structural Analysis of a Novel Chitinase from Glaciozyma Antarctica PI12 for Cold Adaptation. Universiti Teknologi Malaysia.

psychrophilic α-amylase from *Glaciozyma antarctica* PI12 for cold-adaptation analysis. *Journal of Molecular Modeling, 19*(8), 3369-3383.


