



UNIVERSITEIT VAN PRETORIA  
UNIVERSITY OF PRETORIA  
YUNIBESITHI YA PRETORIA

**Engineering plant cysteine protease inhibitors for the transgenic control of banana weevil, *Cosmopolites sordidus* (Germar) (Coleoptera: Curculionidae) and other coleopteran insects in transgenic plants**

By

**ANDREW KIGGUNDU**

Thesis submitted in partial fulfilment of the requirements for the degree

**PHILOSOPHIAE DOCTOR**

Department of Plant Sciences and Forestry and Agricultural Biotechnology  
Institute (FABI)  
in the

Faculty of Natural and Agricultural Sciences  
University of Pretoria South Africa

Supervisor:  
PROF. K.J. KUNERT

Co-supervisors:  
PROF. A. VILJOEN  
PROF. D. MICHAUD

May 2008



## TABLE OF CONTENTS

<b>ABSTRACT .....</b>	<b>x</b>
<b>THESIS COMPOSITION .....</b>	<b>xii</b>
<b>ACKNOWLEDGEMENTS .....</b>	<b>xiii</b>
<b>ABBREVIATIONS AND SYMBOLS .....</b>	<b>xv</b>
<b>CHAPTER 1 Introduction: The banana weevil and protease inhibitors.....</b>	<b>1</b>
1.1 Plant improvement and Africa .....	2
1.2 The banana weevil.....	3
1.2.1 Weevil resistance .....	5
1.2.2 Weevil resistance screening .....	5
1.2.3 Resistance mechanisms .....	8
1.2.4 Resistance breeding.....	9
1.3 Protease/protease inhibitor system .....	12
1.3.1 Insect proteases .....	12
1.3.2 Plant protease inhibitors .....	14
1.3.3 Regulation of protease inhibitors .....	29
1.3.4 Structure of protease inhibitor genes .....	32
1.3.5 Protease inhibitors and insect control .....	34
1.3.6 Engineering of protease inhibitors.....	36
1.4 Study hypothesis aim and objectives .....	39
<b>CHAPTER 2 Characterization of the digestive proteases in the banana weevil gut and the effects of recombinant phytocystatins on early larval growth and development.....</b>	<b>40</b>
2.1 Abstract .....	41
2.2 Introduction .....	42
2.3 Material and methods .....	43



2.3.1 Reagents .....	43
2.3.2 Insect colony and maintenance .....	43
2.3.3 Gut extractions and protein concentration determination .....	44
2.3.4 Determination of pH optima .....	44
2.3.5 Fluorometric assay.....	45
2.3.6 Gelatin SDS-polyacrylamide gel electrophoresis .....	46
2.3.7 Cloning of OC-I and PC genes .....	47
2.3.8 Protein expression and purification.....	49
2.3.9 In-vitro assays with recombinant phytocystatins .....	50
2.3.10 Infiltration of banana stem with phytocystatin.....	51
2.4 Results .....	52
2.4.1 pH optima .....	52
2.4.2 Fluorometric assays .....	52
2.4.3 Gelatin SDS-polyacrylamide gel electrophoresis .....	55
2.5 Discussion .....	62
<b>CHAPTER 3 Phylogenetic and structural comparisons of phytocystatins: A</b>	
<b>bioinformatics approach.....</b>	<b>64</b>
3.1 Abstract .....	65
3.2 Introduction .....	66
3.3 Materials and methods .....	67
3.3.1 Sequence analysis .....	67
3.3.2 Protein structure modelling .....	68
3.3.3 Active site and docking.....	68
3.4 Results .....	68
3.5 Discussion .....	81



<b>CHAPTER 4 Engineering of a papaya cystatin using site - directed mutagenesis to improve its activity against papain and weevil digestive cysteine proteases .....</b>	<b>85</b>
4.1 Abstract .....	86
4.2 Introduction .....	87
4.3 Materials and methods .....	89
4.3.1 <i>Phylogenetic and structural model analysis</i> .....	89
4.3.2 <i>Detection of positive selection sites in PhyCys</i> .....	89
4.3.3 <i>Construction of over-expression vector for papaya cystatin</i> .....	90
4.3.5 <i>Mutagenesis primer design</i> .....	90
4.3.6 <i>Site-directed mutagenesis</i> .....	93
4.3.7 <i>Protein expression</i> .....	95
4.3.8 <i>Purification</i> .....	96
4.3.7 <i>Enzyme Kinetics of mutants</i> .....	97
4.4 Results .....	97
4.4.1 <i>Rational of mutations</i> .....	97
4.4.2 <i>Positive selection among plant cystatin genes</i> .....	100
4.4.3 <i>Mutation and expression of recombinant mutant papaya cystatins</i> .....	103
4.4.4. <i>Inhibition activity of papaya cystatin mutants</i> .....	104
4.5 Discussion .....	108
<b>CHAPTER 5 General discussion and future outlook .....</b>	<b>110</b>
5.1 Summary .....	111
5.2 Future outlook .....	114
<b>REFERENCES .....</b>	<b>116</b>



## LIST OF FIGURES

<b>Figure 1.1</b>	The three broad methods of crop improvement compared.. .....	<b>3</b>
<b>Figure 1.2</b>	The Adult Banana Weevil ( <i>Cosmoplites sordidus</i> ).....	<b>4</b>
<b>Figure 1.3</b>	Genetic engineering strategies currently in commercially produced crops. ....	<b>12</b>
<b>Figure 1.4</b>	Substrate-like mechanism of inhibition by two serine protease inhibitors. ....	<b>21</b>
<b>Figure 1.5</b>	General classification of the cystatin super-family phytocystatins.....	<b>22</b>
<b>Figure 1.6</b>	Alignment of selected members of the 4 cystatin families illustrating the sequence conservation regions within the family members.....	<b>23</b>
<b>Figure 1.7</b>	The three dimensional structure of OC-I showing the characteristic 5 anti-parallel B strands, the single 5 turn a-helix, the N-terminal, the 1 <sup>st</sup> and 2 <sup>nd</sup> hairpin-like loops .....	<b>25</b>
<b>Figure 1.8</b>	Three-dimensional plot showing the complex between papain (blue and green) and Chicken egg white cystatin (CEW) colored light blue, red and yellow (PDB accession No. 1STF). .....	<b>26</b>
<b>Figure 2.1</b>	Schematic diagram of the construction of expression vectors pQOC-I and pQPC used in the study to express OC-I and PC in <i>E. coli</i> , respectively.....	<b>48</b>
<b>Figure 2.2</b>	Effect of pH on the hydrolysis of azocasein by banana weevil larval gut proteases.....	<b>53</b>
<b>Figure 2.3</b>	Cathepsin B, L and H like activities detected in banana weevil larval gut extracts. (B) Trypsin and chymotrypsin-like activities detected in the same extracts. ....	<b>54</b>
<b>Figure 2.4</b>	The effect of protease inhibitors on the proteolysis activity of banana weevil larval gut proteases revealed by separation in a mildly denaturing 15% SDS-PAGE co-polymerized with gelatin. ....	<b>57</b>
<b>Figure 3.2</b>	SDS-PAGE of (A) PC and (B) OC-I at different purification steps.. .....	<b>59</b>



**Figure 2.6** The effect of recombinant cysteine protease inhibitors rOC-I and papaya rPC on the cysteine protease activity of banana weevil mid-gut extracts. .... 60

**Figure 2.7** (A) Illustration of the apparatus used to vacuum infiltrate banana flower stalk disks with cystatin solution. (B) Larvae on the left after developing on cystatin-free (control) disks for 10 days, while larvae on the right developed in cystatin treated disks over the same period. (C) The growth rate of larvae that were reared on banana stem disks vacuum-infiltrated with a solution of recombinant rOC-I and rPC. .... 61

**Figure 2.1** (A) Amino acid sequence alignment of known phytocystatins showing residue conservation across the different cystatins studied. A consensus sequence was also generated. Identical amino acids are highlighted in black while similar ones are in grey. (B) Cartoon of the generalized secondary structural elements of phytocystatins.. .... 73

**Figure 3.2** Phylogenetic tree for known phytocystatins based on the neighbour-joining method using PROTDIST and NEIGHBOR programs available in the PHYLIP (Phylogeny Inference Package) Version 3.57.. .... 75

**Figure 3.3** Predicted three-dimensional structures of selected phytocystatins representing the major phytocystatin phylogenetic groups..... 79

**Figure 3.4** Modelled complex between OC-I (top) and papain (bottom) in front and side views..... 80



<b>Figure 4.1</b>	Schematic representation of recombinant protein expression vector pQE31PC-I created to express papaya cystatin and in which site directed mutagenesis was performed. ....	<b>90</b>
<b>Figure 4.2</b>	Schematic representation of the site-directed mutagenesis protocol used (modified from QuickChange® Site-Directed Mutagenesis Kit) .....	<b>93</b>
<b>Figure 4.3</b>	Consensus sequence from a multiple alignment (see Chapter 3).To illustrate residues subjected to mutations.....	<b>100</b>
<b>Figure 4.4</b>	Location of positively selected codon sites (with Bayesian posterior probabilities greater than 60% under model M3) in Poaceae and Solanaceae cystatins.....	<b>103</b>
<b>Figure 4.5</b>	SDS-PAGE (12%) of the purified fractions of selected papaya cystatin mutants CYSI07D, CYSA53P CYSA32V and CYSW78P to establish purity of the purification. ....	<b>104</b>
<b>Figure 4.6</b>	Comparison of inhibition activity between wild-type papaya cystatin (red bar) and 18 mutants of the papaya cystatin gene.....	<b>106</b>
<b>Figure 4.7</b>	Inhibition activities between wild-type papaya cystatin (red bar) and 18 mutants of the papaya cystatin. Inhibitors were tested by monitoring change in reaction rates of banana weevil (A) and black maize beetle (B) gut extracts.	<b>107</b>



## LIST OF TABLES

<b>Table 1.1</b>	Suggested sources of banana weevil resistance in <i>Musa</i> . .....	<b>7</b>
<b>Table 1.2</b>	Mechanistic classes of proteases, amino acid residues constituting their active site, their optimum pH ranges and examples of the protease enzymes.....	<b>16</b>
<b>Table 1.3</b>	Transgenic crop plants reported to express serine protease inhibitor genes from plants with improved resistance to respective pests.....	<b>19</b>
<b>Table 1.4</b>	Insect pests with reported susceptibility to phytocystatins, either <i>in-vitro</i> , in artificial diet or in transgenic plants.....	<b>27</b>
<b>Table 2.1</b>	Inhibition of banana weevil gut proteases by cysteine (A) and serine (B) protease inhibitors.....	<b>56</b>
<b>Table 3.1</b>	Known phytocystatins obtained from sequence databases: EMBL= European Molecular Biology Laboratory, PIR=Protein information Resource, SP=SwissProt, GB=GeneBank and NCBI=National Centre for Biotechnology Information.....	<b>70</b>
<b>Table 3.2</b>	Percentage identity matrix of phytocystatins .....	<b>76</b>
<b>Table 4.1</b>	Sequence information of the mutagenic primer pairs used for the mutations. The mismatched bases are underlined.....	<b>92</b>
<b>Table 4.2</b>	Mutations performed on native papaya cystatin, the amino acid changes made and the respective rationale.....	<b>98</b>
<b>Table 4.3</b>	Evidence for positive selection events among codon sites of Poaceae and Solanaceae cystatins.....	<b>102</b>





**Engineering plant cysteine protease inhibitors for the transgenic control of banana weevil, *Cosmopolites sordidus* (Germar) (Coleoptera: Curculionidae) and other coleopteran insects in transgenic plants**

**Andrew Kiggundu**

National Banana Research Programme, Kawanda Agricultural Research Institute,  
National Agricultural Research Organisation,  
P. O. Box 7065, Kampala, Uganda

Department of Plant Science and the Forestry and Agricultural Biotechnology Institute  
University of Pretoria, 74 Lunnon Road, Hillcrest,  
Pretoria, 0002. South Africa

Supervisor:

**Karl Kunert**

Department of Plant Science and the Forestry and Agricultural Biotechnology Institute  
University of Pretoria, 74 Lunnon Road,  
Hillcrest, Pretoria, 0002. South Africa

Co-supervisors:

**Altus Viljoen**

Department of Microbiology and Plant Pathology, Forestry and Agricultural Biotechnology  
Institute University of Pretoria, 74 Lunnon Road,  
Hillcrest, Pretoria, 0002. South Africa

*Currently at:* Department of Plant Pathology, University of Stellenboch, Private Bag X1,  
Matieland 7602, South Africa.

**Dominique Michaud**

Département de Phytologie, Pavillon Paul-Comtois,  
Université Laval, Sainte-Foy (Québec),  
Canada G1K 7P4



## ABSTRACT

Cysteine protease inhibitors (cystatins) are expressed in plants in response to wounding and insect herbivory and they form part of the native host-plant defence system. Cysteine proteases are enzymes important in the break down of dietary proteins mainly in the mid gut of coleopteran insects such as the banana weevil. The inhibition of these proteases has a direct effect on the digestive activity of the insect resulting in protein deficiency. This significantly affects insect development and survival. Based on these observations, strategies have been designed involving expression of cysteine protease inhibitors for the transgenic control of insect pests of several crop plants. For this study, it was hypothesized that the major proteases in banana weevil are cysteine proteases and can be effectively targeted by plant cystatins. It was further hypothesised that since plant cystatins are defense related, certain amino acid residues may have undergone positive selection. This provides an opportunity to increase their inhibitory potential to the weevil gut proteases via protein engineering. To prove the hypotheses, both *in-vitro* and *in-vivo* assays were set up thus allowing us to demonstrate the presence of cysteine type proteases banana weevil as well as the effect of cystatins on the weevil proteases and early development. Initial *in-vitro* experiments were able to characterize the proteolytic activity of the banana weevil gut proteases, which are mostly of the cysteine type, and in particular cathepsin B and L like. Two recombinant phytocystatins were further successfully produced using a 6xHis-tagged affinity chromatography system in *Escherichia coli* bacteria. The recombinant phytocystatins were used in a newly developed vacuum infiltration assay system using banana stems. Young weevil larvae were allowed to develop on phytocystatin-treated stems for up to 10 days. They had a 60% reduction in body weight and rate of growth compared to those that grew in untreated stems. By carrying out site-directed mutagenesis to improve the inhibition efficiency of a model papaya cystatin, more



than 8 amino acid residues were found to be subjected to positive selection. Mutation of amino acids yielded improved the inhibition potential of papaya cystatin against the model cysteine protease papain. Increased inhibition was greatest when amino acids were changed in the highly variable regions of the amino acid sequence very closely to the conserved regions.

This study has been able to show for the first time that banana weevils use cysteine protease as major protein hydrolysis enzymes and that these can be effectively targeted by plant cystatins. It has also created novel phytocystatins using engineering of single amino acid sites following an evolutionary approach to modulate them for improved activity and targeting specific proteases.



## THESIS COMPOSITION

**Chapter 1** introduces the banana weevil which is a coleopteran pest of banana that barrows through the underground stem of banana plants causing considerable damage. The chapter reviews conventional efforts towards screening the banana germplasm for resistance, resistance mechanisms, and cross breeding activities targeting the banana weevil as well as protease inhibitors as one group of genes that have potential for weevil control in a transgenic approach. **Chapter 2** reports on investigations into the nature of the banana weevil gut environment *vis a vis* protease activity reveals the protease profile of the gut and bioassays are developed and conducted to test the hypothesis that banana weevil use mostly cysteine protease in protein digestion and can be targeted by cysteine protease inhibitors from plants. **Chapter 3** relates to the phylogeneic, structural and protein modelling analysis of plant cysteine protease inhibitors in an effort to understand evolutionary trends. This could assist a protein engineering strategy to improve the cystatin action against weevil and other coleopteran insects. **Chapter 4** combines evolutionary analysis to determine if positive selection has acted on the cysteine protease inhibitor amino acid residues to lead to the observed diversity. This was followed by protein engineering approaches using site-directed mutagenesis guided by evolutionary analysis to produce novel mutants of the papaya cystatin with increased inhibition capacity. Finally **Chapter 5** discusses the contributions of this thesis to our better understanding of these important plant proteins. It further discusses how best to make future use of them, not only in the improvement of resistance to banana weevil but also to other coleopteran crop pests.



## ACKNOWLEDGEMENTS

The encouragement my late father gave me to pursue science as a career is highly appreciated. Even that he is no longer with us I am sure he is happy to know that this is part of his efforts many years back. To my mother I do not have the right words to say thank you for all the patience you had with me and the toughness that ensured I do not take a wrong detour.

I am very grateful to the National Agricultural Research Organisation, Uganda for allowing me to take time off work and for supporting my studies both financially and morally. Special thanks to Drs. Wilberforce Tushemereirwe and Eldad Karamura whose encouragement and moral support were very helpful not only towards this thesis but also in ensuring that what I have achieved finds usefulness in NARO and Uganda.

Very importantly, I am indebted to the Rockefeller Foundation for the scholarship support that allowed me to undertake this study. To Dr. Joe DeVries of the foundation for his strong belief in me and awarding me this scholarship, but also very insightful were the discussions and suggestions he gave that led to the work in this thesis.

My sincere thanks go to Prof. Karl Kunert who was more than willing to work in his lab under his supervision, way back in February 2002. I thank him for the very good mentoring discussions and useful suggestions that have led to the completion of this scientific accomplishment. I thank FABI, the Plant Science Department of the University of Pretoria for proving such a wonderful and high standard environment for me to achieve what I have. Doing this work at the University of Pretoria and in South Africa has really made me

different. Dr. Altus Viljoen for his encouragement, co-supervision and untiring support that made my stay at FABI as comfortable as possible.

My sincere gratitude also to colleagues and friends at Forestry and Agricultural Biotechnology Institute (FABI), the University of Pretoria, who supported me during my study and stay at FABI



## ABBREVIATIONS AND SYMBOLS

BBTI	Bowman-Birk trypsin inhibitor
bp	Base pair
CaMV	Cauliflower Mozaic Virus
E-64	Trans-epoxysuccinyl-L-leucylamido (4-guanidino) butane
EDTA	Ethylenediaminetetraacetic acid
kDa	Killo Dalton
LB	Luria-Bertani
mL	Milliliter
nm	Nanometer
OC-I	oryzacystatin-I
PAGE	Polyacrylamide gel electrophoresis
PC	Papaya cystatin
PCR	Polymerase Chain Reaction
PI	Protease inhibitor
PMSF	Phenylmethysulphonyl fluoride
SD	Standard deviation
SDS	Sodium dodecyl sulphate
SE	Standard error
U	Unit
Z-phe-arg-AMC	Benzyloxycarbonyl-phenylalanine-arginie aminomethylcoumarin
µg	Microgram
µl	Microlitre
µM	Micromolar
%	Percentage
°C	Degree Celsius
m	Metre