

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

A Structural Investigation of Squash Aspartic Peptidase Inhibitor  
(SQAPI) using Nuclear Magnetic Resonance spectroscopy  
(NMR).

A thesis presented in partial fulfillment of the requirements for the degree of

Master of Science  
in  
Biochemistry

at Massey University, Palmerston North  
New Zealand

Ursula Kate MacAskill

2007

## Abstract

Peptidases are enzymes that hydrolyse peptide bonds. This potentially dangerous activity is regulated by post translational modification and peptidase inhibitors. The best characterized of the peptidase inhibitors are the serpins whilst the aspartic peptidase inhibitors are the least characterized. Aspartic peptidase inhibitors are rare with only nine known sources. However, they are of great interest because they play an important part in several human diseases such as metastasis of breast cancer cells, *Candida albicans* infections and HIV.

The aims of this research project were to investigate the structure of Squash Aspartic peptidase inhibitor (SQAPI), using nuclear magnetic resonance spectroscopy (NMR). This required large amounts of relatively pure and isotopically labeled protein, which was achieved by heterologously expressing His-tagged rSQAPI fusion protein in *Escherichia coli* using a rich to minimal media transfer method. The fusion protein was purified with a nickel column and the N-terminal extension containing the His<sub>6</sub>-tag was removed by cleavage of the fusion protein with enterokinase followed by nickel column purification.

Preliminary 1 dimensional NMR spectra indicated that SQAPI was folded in solution at pH 3. This was confirmed from the results of a preliminary <sup>15</sup>N-edited HSQC. These results combined justified the production of a <sup>15</sup>N <sup>13</sup>C labeled SQAPI sample for the collection of further NMR spectra. From the spectra produced with double labeled protein the backbone and the side-chain atoms of SQAPI were assigned. The chemical shifts are currently 88.89% complete and have been submitted to the biological magnetic resonance bank (BMRB). A preliminary estimate of the secondary structure of SQAPI has been calculated from the HNHA spectrum suggesting that the SQAPI structure has some similarity to the previously proposed model of the inhibitor's structure. Furthermore, the region corresponding to the putative binding loop on the model of SQAPI was found to be mobile and deuterium exchange experiments indicate that the SQAPI structure is more globular than open.

## Acknowledgements

I would like to express my appreciation and gratitude to the following people for their valuable contribution to this work:

My supervisor Dr. Peter Farley for his time, expertise and constant encouragement.

To Dr. Stephen Headey, for his generous investment of time and knowledge.

Quentin, Carole, Carel and Lin from the Mainland Lab. Jolyon, Martin and Dr. Stephen Pascal from the BioNMR Lab. Matt, Simon, Alice and others from the X-lab.

My family and friends for their support and encouragement.

The NZFGW fellowship trust and the J.P. Skipworth Plant Biology scholarship.

## Table of Contents

Abstract.....	I
Acknowledgements.....	II
Table of contents.....	III
Abbreviations.....	VII
List of figures.....	VIII
List of tables.....	X
Chapter One: Introduction.....	1
Peptidase Activity and Regulation.....	1
Classification and Naming of Peptidase Inhibitors.....	2
The Serine Peptidase Inhibitors.....	2
The Cysteine Peptidase Inhibitors .....	3
The Metallopeptidase Inhibitors.....	3
The Aspartic Peptidase Inhibitors.....	4
The Yeast Inhibitor (IA3).....	4
The Nematode Inhibitors (PI 1-4).....	5
The Potato Inhibitor.....	6
The Wheat Inhibitor.....	7
The Sea Anemone Inhibitor (7Equistatin).....	7
The Kiwifruit Inhibitor.....	8
Squash Aspartic Peptidase Inhibitor (SQAPI).....	8
<i>In vivo</i> functions of SQAPI.....	9
The <i>Cucurbita maxima</i> SQAPI gene family.....	10
SQAPI Gene Phylogeny.....	10
Protein Structure.....	12
Structural Determination Techniques.....	12
Stages of Protein Structure Determination.....	13
Aims of this project.....	14

Chapter Two: Materials and Methods.....	15
Materials.....	15
Chemicals.....	15
Enzymes.....	15
Other.....	15
Phosphate buffer preparation.....	16
Vectors.....	16
Plasmid preparation methods.....	16
Bulk plasmid preparation (Alkaline lysis method).....	16
Plasmid Preparation (Rapid Boil Method).....	17
Protein sample production methods.....	17
Expression of His-tag rSQAPI fusion protein.....	17
Optimization of glucose.....	18
Purification of His-tag rSQAPI fusion protein.....	18
Cleavage of His-tag rSQAPI Fusion protein with enterokinase.....	19
Removal of His-tag from cleaved rSQAPI fusion protein.....	19
Protein electrophoresis.....	19
Glycerol stocks.....	20
NMR Methods.....	20
General NMR methods.....	20
Deuterium exchange .....	21
Residual dipolar couplings.....	22
$\phi$ angle	
restraints.....	22
Chapter Three: Protein Sample Preparation for NMR Analysis.....	24

Introduction.....	24
Expression and purification of His-tagged rSQAPI fusion protein.....	24
Glucose optimization.....	28
Optimization of induction time.....	31
Optimization of enterokinase digestion.....	32
Removal of enterokinase and cleaved N-terminal extension from the protein sample.....	34
Chapter Four: Solution structure of SQAPI at pH3.....	36
Introduction.....	36
Preliminary spectra.....	36
Preliminary spectra acquired on unlabelled protein.....	36
Preliminary spectra acquired on labelled protein.....	42
Assignments.....	51
Backbone Assignments.....	51
Vector encoded leader sequence assignments.....	55
Side-chain assignments.....	55
Chemical shifts.....	56
$\phi$ angle restraints.....	58
Deuterium exchange.....	60
Residual dipolar couplings.....	64
Discussion.....	68
Future work.....	70
Literature cited.....	72
Appendices.....	77

## Abbreviations

SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
NMR	Nuclear magnetic resonance spectroscopy
SQAPI	Squash aspartic peptidase inhibitor
mM	Milli mole
mL	Milli litre
g	Gram
µg	Micro gram
µl	Micro litre
ppm	Parts per million
1D	One dimensional
MHz	Mega hertz
IPAP	In-phase/Anti-phase
LB	Luria broth
OD <sub>600</sub>	Optical density (at a wavelength of 600 nano-metres)
IPTG	Isopropyl-b-D-thiogalactopyranoside
Da	Dalton
kDa	Kilo-Dalton
HIV	Human immunodeficiency virus
N	Nitrogen
C	Carbon
DNA	Deoxyribonucleic acid
PI	Peptidase inhibitor
SAP	Secreted aspartic peptidase
PCR	Polymerase chain reaction
HSQC	Hetero-nuclear single quantum correlation
EDTA	Ethylenediaminetetraacetic acid

## List of Figures

Fig. 1	Superimposition of SQAPI onto the rice oryzacystatin structure	11
Fig. 2	SDS-PAGE of fractions from Nickel column purification of His-tagged rSQAPI fusion protein.	27
Fig. 3	Growth of induced cells at three concentrations of glucose.	28
Fig. 4	Nickel column purification of the lysate from cells induced in minimal media containing 2 g/l (A), 4 g/l (B) and 8 g/l (C) glucose concentration.	30
Fig. 5	Expression profile of SQAPI in cells incubated in minimal media containing 4 g/l glucose.	31
Fig. 6	Optimization of the amount of enterokinase for the digestion of His-tagged rSQAPI fusion protein.	33
Fig. 7	Removal of the His-tag from the cleaved rSQAPI fusion protein.	35
Fig. 8	A one dimensional $^1\text{H}$ spectrum of His-tagged rSQAPI fusion protein in a glycine –HCl buffer at pH 3	39
Fig. 9	A one dimensional $^1\text{H}$ spectra of 0.06 mM His-tagged rSQAPI fusion protein in 10 mM phosphate buffer pH 3.	41
Fig. 10	$^{15}\text{N}$ hetero-nuclear quantum correlation (HSQC) spectrum of the 0.2 mM His-tagged rSQAPI fusion protein in phosphate buffer.	45
Fig. 11	A $^{15}\text{N}$ - HSQC ran at 37 $^{\circ}\text{C}$ on a 0.2 mM rSQAPI sample in phosphate buffer pH 3 after the N-terminal extension had been removed with enterokinase.	47
Fig. 12	An overlay of the HSQCs shown in Fig. 10 and Fig 11. This clearly shows the improvement of the spectra after the N-terminal extension was removed.	49

Fig. 13 54

An overlay of 4 HSQCs run at 10<sup>0</sup>C, 25<sup>0</sup>C, 40<sup>0</sup>C and 50<sup>0</sup>C.

Fig. 14 62 & 63

Panel A: HSQC of lyophilised <sup>15</sup>N rSQAPI fusion protein that has been dissolved in <sup>2</sup>H<sub>2</sub>O. Panel B: An overlay of the HSQC shown in panel A and the HSQC shown in figure 10.

Fig. 15 66 & 67

The [in-phase plus anti-phase] (A) and [in-phase minus anti-phase] (B) HSQCs collected on SQAPI both out of alignment media (isotropic) and in alignment media (anisotropic).

Technical Note: Due to technical issues the figure legends in chapter four are printed on the pages preceding their respective figures.

## List of Tables

Table 1	29
Table of yields of His-tagged rSQAPI fusion protein from glucose optimization trials.	
Table 2	50
Spectra acquired for structural determination of SQAPI	
Table 3	57
Abridged CANDID output of missing assignments	
Table 4	59
$\phi$ angle restraints as calculated from the HNHA	