Durham University

Durham E-Theses

Expression of novel insecticidal proteins in the yeast pichia pastoris

Gareth Hinchliffe

How to cite:

Hinchliffe, Gareth (2007) Expression of novel insecticidal proteins in the yeast pichia pastoris. Masters thesis, Durham University.

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a https://etheses.durham.ac.uk/id/eprint/2361/ is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the full Durham E-Theses policy for further details.

Academic Support Office, The Palatine Centre, Durham University, Stockton Road, Durham, DH1 3LE
e-mail: e-theses.admin@durham.ac.uk Tel: +44 0191 334 6107
https://etheses.durham.ac.uk

Expression of Novel Insecticidal Proteins in the Yeast *Pichia pastoris*

The copyright of this thesis rests with the author or the university to which it was submitted. No quotation from it, or information derived from it may be published without the prior written consent of the author or university, and any information derived from it should be acknowledged.

Gareth Hinchliffe

Biological and Biomedical Sciences

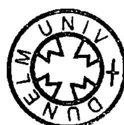
University of Durham

MSc Thesis

2007

Supervisor: Dr J A Gatehouse

0 1 SEP 2008



Abstract

The development of Fusion Protein Technology during the search for new environmentally benign pesticides has prompted further investigation into novel insecticidal proteins from predatory and parasitic insects.

Eulophus pennicornis is a parasitoid wasp used in biological pest control which causes developmental arrest in its host. *E. pennicornis* venom gland cDNAs were isolated and sequenced. Searches for sequence similarity in the global databases revealed a splicing factor, serine protease and a neprilysin-like protein as key components besides an already isolated reprolysin-like metalloprotease and juvenile hormone-inducible protein. Attempts were made to express the reprolysin-like metalloprotease in *Pichia pastoris* as a fusion protein with GNA and with structural truncations. Constructs for the fusion and truncations were produced using recombinant DNA techniques but unfortunately, little success was achieved with the expression. The results presented point towards reprolysin-like metalloprotease perhaps being toxic to the *P. pastoris* yeast cells.

Avidin, another protein with the potential for exploitation in Fusion Protein Technology, was also expressed as a recombinant protein in *P. pastoris* and successfully purified to almost 100% homogeneity. In this study, bioassays with *Mamestra brassicae* (cabbage moth) larvae confirmed its transport to the haemolymph and insecticidal activity when delivered orally above a threshold concentration via artificial diet. Fluorescent labelling of recombinant avidin also highlighted the mechanics involved in its transport within *M. brassicae* larvae and confirmed its suitability for use in Fusion Protein Technology. Avidin-antibody coupling experiments gave inconclusive results for transport to the haemolymph of *Sitobion avenae* (cereal aphid), but revealed a more permeable gut structure in *M. brassicae* larvae, potentially opening up new avenues for the control of the larvae of pest lepidopterans using lethal antibodies.

Table of Contents

List of Figures Page VI

List of Tables Page IX

Abbreviations Page X

Acknowledgements Page XIII

1.0. Introduction Page 1

1.1. Parasitic Wasp Venom Page 2

1.1A. *Eulophus pennicornis* Page 2

1.1B. Reprolysin-like Metalloprotease Page 3

1.2. Lectins as Carriers Page 5

1.3. Fusion Protein Technology Page 7

1.4. *Pichia pastoris* Page 8

1.5. Avidin Page 11

1.6. Project Aims Page 14

2.0. Materials and Methods Page 15

2.1. Biological and Chemical Reagents Page 15

2.2. Laboratory Techniques Page 15

2.2A. *Eulophus pennicornis* Venom Gland cDNA Library Screening and Production of Reprolysin-like Metalloprotease Structural Variants Page 15

2.2A-1. Bacterial Cell Culture Page 16

2.2A-2. Plasmid DNA Isolation Page 16

2.2A-3. Polymerase Chain Reaction (PCR) Page 16

2.2A-4. Restriction Endonuclease Digests Page 19

2.2A-5. DNA Agarose Gel Electrophoresis Page 19

2.2A-6. Recovery of DNA from Agarose Gels Page 20

2.2A-7. DNA Sequence Analysis Page 20

2.2A-8. Ligation Reactions Page 20

2.2A-9. Transformation of Competent Bacterial Cells Page 20

2.2A-10. Colony PCR Page 21

2.2A-11. DNA Quantification Page 21
2.2A-12. Linearization of DNA Page 21
2.2A-13. Extraction and Precipitation of DNA Page 21
2.2A-14. Transformation of Competent Yeast Cells Page 21
2.2A-15. Bacterial and Yeast Glycerol Stocks Page 22
2.2B. Expression of Reprolysin-like Metalloprotease, Reprolysin-like Metalloprotease Structural Variants and Avidin Page 22
2.2B-1. Yeast Cell Culture Page 22
2.2B-2. Chromatography and Purification Page 24
2.2B-3. SDS-PAGE Page 24
2.2B-4. Dialysis and Freeze Drying Page 25
2.2B-5. Protein Quantification Page 25
2.2B-6. Chloroform-Methanol Precipitation of Proteins Page 25
2.2B-7. Western Blotting Page 26
2.2B-8. Gelatin Activity Gels Page 26
2.2C. Coupling of Antibody to Avidin Page 27
2.2C-1. Biotinylation Page 27
2.2D. Fluorescent Labelling of Avidin-Biotin Conjugate Page 27
2.2D-1. Gel Filtration Page 28
2.3. Insect Cultures Page 28
2.4. Insect Bioassays Page 28
2.4A. Injection Bioassays Page 28
2.4B. Artificial Diet Bioassays Page 29
2.4C. Haemolymph, Gut Contents and Midgut Extraction and Analysis Page 30

3.0. Results Page 31
3.1. *Eulophus pennicornis* Venom Gland cDNA Library Screen Page 31
3.1A. Expression of Reprolysin-like Metalloprotease Page 40
3.1B. Biological Activity of Reprolysin-like Metalloprotease Page 40
3.2. Reprolysin-like Metalloprotease Fusion Protein Variants Page 46
3.2A. Production of Reprolysin 1, Reprolysin 2 and Reprolysin 3 Expression Constructs Page 46
3.2B. Expression of Reprolysin-like Metalloprotease, Reprolysin 1, Reprolysin 2 and Reprolysin 3 Page 51
3.3. Avidin as a Fusion Protein Carrier Page 62
3.3A. Expression of Recombinant Avidin Page 62
3.3B. Transport of Recombinant Avidin Page 66
3.3C. Effects of Avidin upon Injection Page 68

3.3D. Toxicity of Recombinant Avidin Delivered Orally .... Page 68
3.3E. Mechanics of Recombinant Avidin Transport .... Page 72
3.3F. Antibody Coupled Avidin .... Page 78
3.3G. IgG Transport .... Page 80

4.0. Discussion .... Page 84
4.1. *Eulophus pennicornis* Venom .... Page 84
4.2. Reprolysin-like Metalloprotease .... Page 84
4.3. Avidin .... Page 86

Bibliography .... Page 88

List of Figures

Figure 1 - Nucleotide and Amino Acid Sequence of *Eulophus pennicornis* Reprolysin-like Metalloprotease Page 4
Figure 2 - Diagrammatic Representation of a Lepidopteran Larvae Showing Lectin and Fusion Protein Transport Page 6
Figure 3 - Nucleotide and Amino Acid Sequence of Recombinant Avidin Page 13
Figure 4 - *Eulophus pennicornis* Venom Gland Insert Sequence 1 Page 34
Figure 5 - *Eulophus pennicornis* Venom Gland Insert Sequence 2 Page 35
Figure 6 - *Eulophus pennicornis* Venom Gland Insert Sequence 3 Page 36
Figure 7 - *Eulophus pennicornis* Venom Gland Insert Sequence 4 Page 37
Figure 8 - *Eulophus pennicornis* Venom Gland Insert Sequence 5 Pages 38 / 39
Figure 9 - Reprolysin-like Metalloprotease Sequence Alignments Page 41
Figure 10 - Recombinant Reprolysin-like Metalloprotease Purification and SDS-PAGE Analysis Page 42
Figure 11 - Gelatinase Activity Gel of *Eulophus pennicornis* Venom Gland Incubations Page 43
Figure 12 - Gelatinase Activity Gel of *Eulophus pennicornis* Reprolysin-like Metalloprotease Page 44
Figure 13 - Survival and Weight Data for *Lacanobia oleracea* Bioassay Page 45
Figure 14 - Nucleotide and Amino Acid Sequence Confirmation and Diagrammatic Representation of Reprolysin 1 Pages 47 / 48
Figure 15 - Nucleotide and Amino Acid Sequence Confirmation and Diagrammatic Representation of Reprolysin 2 Page 49
Figure 16 - Nucleotide and Amino Acid Sequence Confirmation and Diagrammatic Representation of Reprolysin 3 Page 50
Figure 17 - Western Blot Analysis of Reprolysin 1 pGAPZαB Page 52
Figure 18 - Western Blot Analysis of Reprolysin 2 pGAPZαB Page 52
Figure 19 - Western Blot Analysis of Reprolysin 3 pGAPZαB Page 53
Figure 20 - Western Blot Analysis of Original Reprolysin-like Metalloprotease pGAPZαB Page 55
Figure 21 - Western Blot Analysis of ASAII pGAPZαB Page 55

Figure 22 - Western Blot Analysis of Original Reprolysin-like Metalloprotease pPICZαB Page 57
Figure 23 - Western Blot Analysis of Reprolysin 1 and Reprolysin 2 pPICZαB Page 57
Figure 24 - Western Blot Analysis of Reprolysin 3 pPICZαB Page 58
Figure 25 - Western Blot Analysis of Time Dependent Cultures of Original Reprolysin-like Metalloprotease pPICZαB Page 60
Figure 26 - Western Blot Analysis of Time Dependent Cultures of ASAII pPICZαB Page 61
Figure 27 - Recombinant Avidin Purification and SDS-PAGE Analysis Page 64
Figure 28 - SDS-PAGE Quantification of Purified Recombinant Avidin Page 65
Figure 29 - Western Blot Analysis of *Mamestra brassicae* (Cabbage moth) Larvae Gut Contents, Gut and Haemolymph after Feeding on Avidin Page 67
Figure 30 - Survival and Weight Data for *Mamestra brassicae* Control Diet Bioassay Page 69
Figure 31 - Survival and Weight Data for *Mamestra brassicae* 0.1% Recombinant Avidin Diet Bioassay Page 69
Figure 32 - Survival and Weight Data for *Mamestra brassicae* 0.01% Recombinant Avidin Diet Bioassay Page 70
Figure 33 - Survival and Weight Data for *Mamestra brassicae* 0.001% Recombinant Avidin Diet Bioassay Page 70
Figure 34 - Survival and Weight Data for *Mamestra brassicae* 0.1% Sigma Avidin Diet Bioassay Page 71
Figure 35 - Comparative Survival Data for *Mamestra brassicae* Avidin Diet Bioassay Page 71
Figure 36 - Fluoroskan Fluorescence Calibration Curve Page 73
Figure 37 - Gel Filtration Profile and Fluorescence Data of *Mamestra brassicae* Larvae Haemolymph Page 74

Figure 38 - Gel Filtration Profile and Fluorescence Data of *Mamestra brassicae* Larvae Haemolymph Spiked with Fluorescein-Biotin Page 76

Figure 39 - Gel Filtration Profile and Fluorescence Data of *Mamestra brassicae* Larvae Haemolymph Spiked with Fluorescein-Biotin Avidin Conjugate Page 76

Figure 40 - Gel Filtration Profile, Fluorescence Data and Western Blot Analysis of Haemolymph from Fluorescein-Biotin Avidin Droplet Fed *Mamestra brassicae* Larvae Page 77

Figure 41 - Western Blot Analysis of *Sitobion avenae* (Cereal aphid) Haemolymph after Feeding on Avidin-Biotinylated Antibody Page 79

Figure 42 - Western Blot Analysis of *Mamestra brassicae* Larvae Haemolymph after Feeding on Avidin-Biotinylated Antibody Page 79

Figure 43 - Western Blot Analysis of IgG Page 81

Figure 44 - Western Blot Analysis of *Mamestra brassicae* Larvae Haemolymph after Feeding on IgG Page 81

Figure 45 - Western Blot Analysis of *Mamestra brassicae* Larvae Haemolymph after Feeding on IgG Page 82

Figure 46 - Western Blot Analysis of *Sitobion avenae* Haemolymph After Feeding on IgG Page 83

Figure 47 - Western Blot Analysis of *Sitobion avenae* Haemolymph After Feeding on Avidin Page 83

List of Tables

Table 1 - Oligonucleotide Primers Used to Generate the Reprolysin-like Metalloprotease Structural Variants — Page 18

Table 2 - PCR Cycling Conditions for Taq Reactions — Page 18

Table 3 - PCR Cycling Conditions for Phusion Reactions — Page 18

Table 4 - *Eulophus pennicornis* Venom Gland cDNA Library NCBI-BLAST Sequence Similarity Search Results — Page 33

Table 5 - Absorbance Readings of Original Reprolysin-like metalloprotease pPICZαB in BMGY and BMMY media — Page 60

Abbreviations

AOX1/2	Alcohol Oxidase
GAP	GlycerAldehyde-3-Phosphate dehydrogenase
GNA	*Galanthus nivalis* Agglutinin (Snowdrop Lectin)
Da	Dalton
kDa	Kilo Dalton
DNA	DeoxyriboNucleic Acid
cDNA	Complementary DNA
dsDNA	Double stranded DNA
NCBI-BLAST	National Centre for Biotechnology Information Basic Local Alignment Search Tool
cfu	Colony forming units
M_r	Relative Molecular mass
EDTA	EthyleneDiamine TetraAcetic acid
^oC	Degrees Celsius
mg	Milligrams
ButaIT (RST)	Red Scorpion Toxin
LB	Luria Bertani
[w/v]	Weight per volume
[v/v]	Volume per volume
[w/w]	Weight per total weight
LAF	Laminar Air Flow
ml	Millilitres
rpm	Revolutions per minute
μg	Micrograms
PCR	Polymerase Chain Reaction
dNTP	Deoxy (Base) TriPhosphate
mM	Millimolar
μl	Microlitres
TAE	Tris-Acetic acid-EDTA
M	Molar
V	Volts (voltage)

A	Amps (current)
UV	Ultraviolet
bp	Base pairs (DNA)
nm	Nanometres
λ_{260} / λ_{280}	Wavelength 260nm / 280nm
AU	Absorbance Unit
YPG	Yeast extract-Peptone-Glycerol
BMGY	Buffered Glycerol-complex Medium
BMMY	Buffered Methanol-complex Medium
YNB	Yeast Nitrogen Base
L	Litres
μm	Micrometres
FPLC	Fast Protein Liquid Chromatography
SDS	Sodium Dodecyl Sulphate
SDS-PAGE	Sodium Dodecyl Sulphate Poly Acrylamide Gel Electrophoresis
mm	Millimetres
TEMED	N,N,N',N'-Tetramethylethylenediamine
TIFF	Tagged Image File Format
nM	Nanomolar
PBS	Phosphate Buffered Saline
ASA / ASAII	*Allium sativum* Agglutinin / *Allium sativum* Agglutinin homodimer (Garlic Lectin)
ECL	Enhanced Chemiluminescent
IgG	Immunoglobulin G
g	Grams
DTT	Dithiothreitol
BCA	Bicinchoninic Acid
BSA	Bovine Serum Albumin
ppm	Parts per million
mg/L	Milligrams per Litre
mRNA	Messenger RNA (RiboNucleic Acid)

DNA bases

A	Adenine
C	Cytosine
G	Guanine
T	Thymine

Amino Acid codes

A	Alanine
R	Arginine
N	Asparagine
D	Aspartic Acid
C	Cysteine
E	Glutamic Acid
Q	Glutamine
G	Glycine
H	Histidine
I	Isoleucine
L	Leucine
K	Lysine
M	Methionine
F	Phenylalanine
P	Proline
S	Serine
T	Threonine
W	Tryptophan
Y	Tyrosine
V	Valine

Acknowledgements

I would like to thank Dr John Gatehouse, without whom this Masters research project would not have been possible.

Thanks especially to Dr Elaine Fitches and Dr Daniel Price for their valuable help, support and advice throughout my laboratory work. Permission to include figures produced by Dr Daniel Price in the introduction section is greatly appreciated.

I also acknowledge the other members of Lab 1 research team past and present, who have always ensured a pleasant working environment.

Finally, I thank my parents who have always given support, encouragement and offered guidance throughout all of my studies and experiences.

1.0. Introduction

Despite advances in insect control technologies, at the start of the 21st century the Global Crop Diversity Trust has estimated that as much as 30-40% of the world's crop produce is destroyed by agricultural pests. In the UK alone, more than 500 tonnes of insecticides are applied annually with a cost now exceeding £25 million (Garthwaite et al., 2004). Currently used insecticides include carbamates, neonicotinoids, organophosphates, synthetic pyrethroids, insect growth regulators and metabolic disruptors. However, significant problems are associated with the widespread use of these conventional insecticides. The most important of these problems, causing much public concern, are their poor specificity to pest insect species and their toxicity to vertebrates. Many beneficial insects have been harmed and ecosystems disrupted over time (Weisser and Siemann, 2004). The development of insect resistance to those insecticides currently used and the prospect of additional insect pests spreading into the UK as a result of climate change, are two further reasons why the UK government has recently issued objectives (Pesticides and the Environment, March 2006) seeking initiatives to encourage a more sustainable and environmentally sensitive approach to insect pest management. Changes in European legislation have led to the withdrawal of many chemical pesticides from the market, meaning arable farmers have a reduced number of plant protection products accessible. Extensive use of the fewer insecticides available, also rapidly increases the rate at which insect resistance can occur. This is a critical problem with similar chemical compounds such as pyrethroids for example, where resistance to one member of the pyrethroid family can mean a degree of resistance to others in that class (cross resistance) (Beugnet and Chardonnet, 1995). Thus, there is increased requirement for new, environmentally friendly insecticides with reduced effects on non-target species when compared to the currently available ones with broad spectrum chemistries. For this reason, attention is being turned to insects that naturally produce insecticidal proteins fundamental to their predatory / parasitic relationship with other insects, such as parasitic wasps. Parasitic wasps produce venom which is a complex mixture of proteins, many having negative effects on their hosts. The venom is therefore important in the search for new insecticidal proteins (Section 1.1). As Whetstone and Hammock (2007) state, the great number of species

of fauna and flora worldwide that produce insecticidal proteins gives almost limitless possibilities for pesticides but only if suitable delivery systems are available.

Recently, a novel method of delivering insect-specific, biologically active peptides to the haemolymph of pest insects following oral ingestion has been discovered. Fusion Protein Technology relies on the use of Snowdrop lectin (GNA; *Galanthus nivalis* Agglutinin) (Section 1.2) as a carrier protein to transport insect-derived toxic peptides across the insect gut and into the haemolymph (Section 1.3). Fusion Protein Technology is possible due to the development of *Pichia pastoris* yeast as a recombinant protein expression host (Section 1.4). Initial use of the technology has proven successful and the production of further novel fusions has highlighted a potential area for improvement (Section 1.5). With continued scientific and industrial research, Fusion Protein Technology may provide, in part, an answer to the dilemma associated with the current insecticides and the insect pest problems arising as a result of global climate change.

1.1. Parasitic Wasp Venom

A number of organisms in nature are capable of producing venom for use in either defence or hunting. The venom of parasitic wasps is receiving increased interest by scientists as it is a complex mixture of proteins which are known to be involved in host regulation through inducing paralysis, disrupting the development of the host, altering its metabolic regulation or interfering with its immune response and reproductive system (Vinson and Iwantsch, 1980; Moreau and Guillot, 2005). It may therefore be a source of new molecules with potential for use in insect pest control.

1.1A. *Eulophus pennicornis*

Eulophus pennicornis (Nees) (Hymenoptera: Eulophidae) is a gregarious ectoparasitoid of late stage larvae of noctuid moth species, found commonly throughout northern Europe. It injects venom into the host larva prior to oviposition onto their surface. The venom contains various components which are used to regulate the host and benefit the offspring. Digestive proteases, proteins causing effects on haemocytes and factors that disrupt ecdysteroid and juvenile hormone titres have been discovered present in the venom (Richards and Edwards, 1999;

Richards and Edwards, 2000; Richards and Edwards, 2002; Edwards et al., 2006). It is likely that there are still many more proteins to be documented.

Previously, to identify new insecticidal proteins and give insight into *E. pennicornis* venom composition, clones from a venom gland cDNA library were randomly sequenced and NCBI-BLAST similarity searches performed. The most abundant clones encoded metalloproteases similar to reprolysins (Section 3.1).

1.1B. Reprolysin-like Metalloprotease

The amino acid sequence of the original isolated reprolysin-like metalloprotease can be seen in figure 1. The metalloprotease acquires the name 'reprolysin-like' due to its sequence similarities with the reprolysins. Reprolysins are a subfamily of the M12 peptidases, a group containing all metallopeptidases. Most commonly found in snake venoms they are easily characterised as they contain a C-terminal extension consisting of one or more non-peptidase domains (Barrett et al., 1998). The complete open reading frame of the protein encodes 430 amino acids (Figure 1), the first 22 of these being a signal peptide, which upon cleavage results in a 408 amino acid mature protein with an M_r of 46kDa. The protein has an isoelectric point of 6.0 and five potential N-glycosylation sites (N64, N113, N148, N188 and N413). It contains a pro-domain (amino acids 1-212) and a catalytic metalloprotease domain (amino acids 213-430). In contrast to many other reprolysins, there is no C-terminal disintegrin domain. The catalytic metalloprotease domain contains numerous cysteine residues and an active site zinc binding motif HExxHxxxGxHD (amino acids 362-372, where x implies any amino acid) (Figure 1) consistent with other members of the MB metalloprotease family (with an extra amino acid) (Barrett et al., 1998). The three histidine (H) residues within this motif ligand to the zinc ion, the glutamic acid (E) residue is the catalytic amino acid and the glycine (G) residue characterizes an important structural turn (Stöcker et al., 1995). However, the reprolysin-like metalloprotease also contains a methionine (M) residue twenty eight amino acids distal to the first histidine of the active site zinc binding motif which more accurately defines this metalloprotease as a metzincin (Bode et al., 1993; Hooper, 1994).

```
                                   CGGCCGGGGATTGTTCTTGATATTGTCTACCTGTATTATTGTTT   -119
TACGATTCTTTTAAAAGACGGAGACCGAAAAGAAATACAGTTCAAAGTTACTCAGACAGAGCGTTGCAAAGGAAG   -75

ATGGATTTGTTCATTTTAACTCGTTTCATTTTGTTCCTATCTTTCTTTATGAAATCGATTCATTGTCAGTATAGC   +75
 M  D  L  F  I  L  T  R  F  I  L  F  L  S  F  F  M  K  S  I  H  C  Q  Y  S    25

GAATCGCAGGAATCCGGCCATAACCGGAACGCACCAGATAAAGAATTGACTACAGAAGAGTTCCAATTGATTTTT   +150
 E  S  Q  E  S  G  H  N  R  N  A  P  D  K  E  L  T  T  E  E  F  Q  L  I  F    50

CATCAATCTCAAACAGTTGATATCGAATACGATTTTATCAATATCACAACCGAGATGATTGAAACAGAACGTAAA   +225
 H  Q  S  Q  T  V  D  I  E  Y  D  F  I  N  I  T  T  E  M  I  E  T  E  R  K    75
                                        *
GTTAGTTTTACAATCGATGGTAAAGAGTACCACTTATCATTGACTCCAGCTGCTTCACAATCAGTGTTGCCATAT   +300
 V  S  F  T  I  D  G  K  E  Y  H  L  S  L  T  P  A  A  S  Q  S  V  L  P  Y    100

GGTACTAAAATTAAGAGTGCAATTTGGTGGACTGATAATGATACTCATATCCACGAAGAGGACTATAGTGACGAG   +375
 G  T  K  I  K  S  A  I  W  W  T  D  N  D  T  H  I  H  E  E  D  Y  S  D  E    125
                                     *
AGATGGGATAGTAGAGCTATCTATGAAAATTTAGAAATTATGGCAACTATACTGGTCCGTACCGAGAATGGTACT   +450
 R  W  D  S  R  A  I  Y  E  N  L  E  I  M  A  T  I  L  V  R  T  E  N  G  T    150
                                                                   *
TCTTATTATGATGGAGTCTTCGGTGAAGGTATAGCAATGAAGGTTGTGAGATCTCTTCCAGGAAGATTGATGAAT   +525
 S  Y  Y  D  G  V  F  G  E  G  I  A  M  K  V  M  R  S  L  P  G  R  L  M  N    175

ATTTACGGAGCAAACTATCACTTTGTTTATGATAGTAATGGCTCCGTGTACGACGTTGTTTTAAACGGACAAGAT   +600
 I  Y  G  A  N  Y  H  F  V  Y  D  S  N  G  S  V  Y  D  V  V  L  N  G  Q  D    200
                                     *
GAACCAGCTGTTCCTGCTGATATGGCAGTAAATAATTTTATCCGAAACTCCTTGTACTTGTCGACTACTCATTA   +675
 E  P  A  V  P  A  D  M  A  V  N  N  F  Y  P  K  L  L  V  L  V  D  Y  S  L    225

TTCAAGATCTTTAATGAAAACTTCGAGGAGACTGTCAAGTATCTCACAATATTTTGGAATGCTGTAAACTTAAGG   +750
 F  K  I  F  N  E  N  F  E  E  T  V  K  Y  L  T  I  F  W  N  A  V  N  L  R    250

TTCAGACCGGTCCAGCATCCGAAAGTAAATATCATAATCACTGGGATCGTCATTGCTAAGAACGAAGCCGCATTT   +825
 F  R  P  V  Q  H  P  K  V  N  I  I  I  T  G  I  V  I  A  K  N  E  A  A  F    275

CAGCACGTTTATAGAGCTAGGTATAGCAAAAATTCGAAACTAGTGCATACCGGAAGAGTTATAGATAATGGTCGA   +900
 Q  H  V  Y  R  A  R  Y  S  K  N  S  K  L  V  H  T  G  R  V  I  D  N  G  R    300

TATTTCTTTGGGACTAATTTTGACCCTTATTACGACAATTATGATGCCAGTTTTACAATGGCAAGTATGGATGAT   +975
 Y  F  F  G  T  N  F  D  P  Y  Y  D  N  Y  D  A  S  F  T  M  A  S  M  D  D    325

CCAACTGGAAAAGGAGGTGCAACTGTCATTGGCGGGATTTGCTCCTCATCTAACAATATCGCTTATATCAGAGAC   +1050
 P  T  G  K  G  G  A  T  V  I  G  G  I  C  S  S  S  N  N  I  A  Y  I  R  D    350

GTTGGAAGTTATTCGGGAGTAAAGGTTGCTACACACGAATTGGGCCACTTGCTAAATGGCCAACACGATTCAGAT   +1125
 V  G  S  Y  S  G  V  K  V  A  T  H  E  L  G  H  L  L  N  G  Q  H  D  S  D    375

ACCACTTGTAGCGAAAAGATAAATGATAATATCTATACAATCATGGCTAAGCAAGGATCAACAAAAGCAAGCAAA   +1200
 T  T  C  S  E  K  I  N  D  N  I  Y  T  I  M  A  K  Q  G  S  T  K  A  S  K    400

TTCGTTTGGTCATCATGCACTCTGACCGCATTTGCAAACTTTTCCAAAACAACAAGTGCAGCGTGCCTAAAAGAT   +1275
 F  V  W  S  S  C  T  L  T  A  F  A  N  F  S  K  T  T  S  A  A  C  L  K  D    425
                                     *
ACCTATCGAAAGCATTAG                                                            +1293
 T  Y  R  K  H  .                                                             430

GATTGAAGATCTTCAAAAGAAATGCATAAATTATATAATCACTTGGATGCAGAGAGAGAGAGAGAGAGTAGGAGT   +1368
GAGAGAGAGAGTAGGAGTCAAAAAAAAAAAAAAAAAAAAAAAAAAA                                +1413
```

Figure 1.
Nucleotide and amino acid sequence of *E. pennicornis* reprolysin-like metalloprotease. The predicted N-terminal signal peptide (amino acids 1-22) is outlined with a white box. Highlighted in red is a putative binding domain (amino acids 23-212). The catalytic reprolysin metalloprotease domain is coloured in orange (amino acids 213-430), with the active site zinc binding motif outlined in blue (amino acids 362-372). Potential N-linked glycosylation sites (N64, N113, N148, N188 and N413) are marked with *.

1.2. Lectins as Carriers

One problem with developing orally active insecticides arises from the morphology of the insect gut and location of the insecticide target. Many insecticidal compounds have to be transported from the gut lumen across the peritrophic membrane and gut epithelium into the haemolymph, maintaining their activity without degradation (Lehane and Billingsley, 1996).

The plant lectin (Ricin) was described by Stillmark in 1888 (Stillmark, 1888). At that time 'lectins' were described as haemagglutinins (Elfstrand, 1898) due to their ability to agglutinate cells (Pusztai, 1991). Around sixty years later, the name lectin was coined by Boyd and Shapleigh (1954) due to their activity against specific cell types mediated by carbohydrate binding (Boyd and Reguera, 1949; Liener et al., 1986).

Lectins are ubiquitous throughout nature, forming a large and diverse group of proteins. In insects, it is the carbohydrate binding specificity of a number of lectins which gives vast scope for their exploitation. Many gut membrane proteins such as brush-border enzymes, receptors and transport proteins are glycosylated, producing potential lectin binding sites. It is the binding at these glycosylated sites which is thought to be the basis of the ability of certain lectins to translocate across the gut epithelium into the haemolymph (Figure 2a). In particular, mannose-binding lectins have been well studied as this sugar is commonly found in the mid gut of insects but rarely in higher organisms. Fitches et al. (2001) demonstrated that intact snowdrop lectin (GNA, *Galanthus nivalis* Agglutinin) and jackbean lectin (Con A, *Canavalia ensiformis* Agglutinin) could be detected in the haemolymph of *Lacanobia oleracea* (tomato moth) larvae following oral delivery via artificial diet. The gut to haemolymph translocation capacity of mannose-binding lectins forms the basis of Fusion Protein Technology.

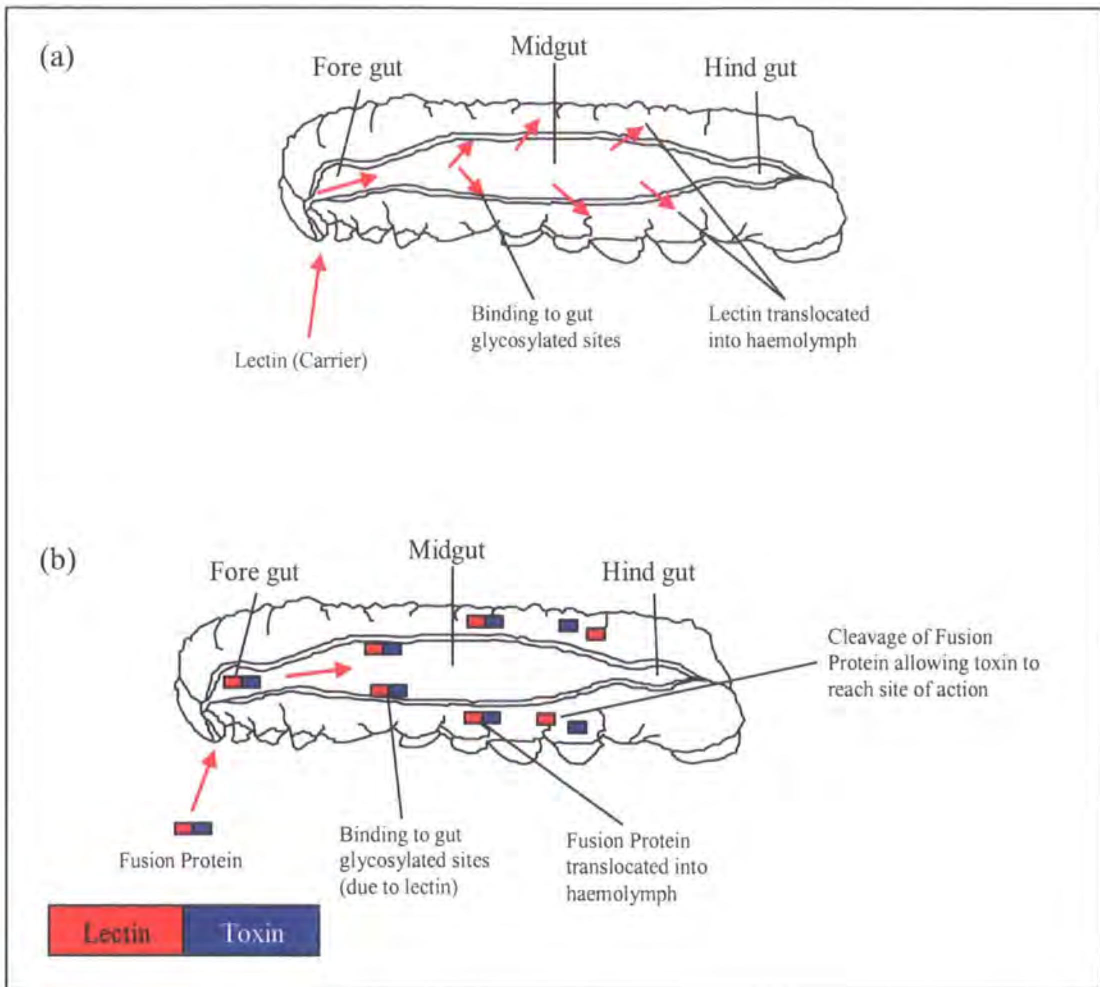


Figure 2.
Diagrammatic representation of a lepidopteran larva. (a) Showing a lectin entering the gut, binding to the glycosylated sites on the gut membrane proteins and being transported into the haemolymph. (b) Showing a fusion protein entering the gut, binding to the glycosylated sites on the gut membrane proteins via the lectin and being transported to the haemolymph where it is cleaved releasing the toxin to reach its site of action.

1.3. Fusion Protein Technology

Fusion Protein Technology utilizes the property of lectins that allows them to be transported across the insect gut epithelium into the haemolymph. Using recombinant DNA techniques, a potentially insecticidal protein is fused to the lectin and expressed in *P. pastoris*. When fed to target insects via artificial diet, the lectin acts as a carrier, delivering the protein to the haemolymph where it exerts a toxic effect (Figure 2b). Fusion Protein Technology therefore offers a novel method of producing orally active insecticides from proteins which are otherwise ineffective when administered alone due to either degradation by digestive enzymes within the insect gut or failure to access the haemolymph. Currently, GNA, isolated from the bulbs of snowdrops is used as the carrier. It has a tetrameric structure, with each identical subunit being 12.5kDa in size. *Galanthus nivalis* agglutinin is also D-mannose specific and each of the subunits can bind three mannose sugar residues (Van Damme et al., 1998).

Fusion Protein Technology may consequently offer an improvement to the current available insecticides. Employing a mannose-binding lectin as the carrier ensures a fusion protein is insect specific as mannose sugars are present only within a large number of insect guts and are uncommon in higher animals. The use of an insect-specific or insect order-specific toxic protein also increases this specificity. It is also thought unlikely that fusion proteins would be harmful to other non-target organisms such as beneficial insects and mammals, as cleavage of fusion proteins occurs once they are transferred into the haemolymph of the target insect. The cleavage allows the insecticidal protein to reach its site of action, but as it is no longer associated with the lectin it means it cannot be carried to the haemolymph of any other predator should the target insect be consumed.

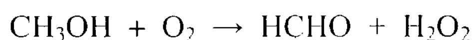
To date, C- and N-terminal fusion proteins using *Manduca sexta* allatostatin, *Segestina florentina* toxin, *L. oleracea* chitinase and *Mesobuthus tamulus* toxin have demonstrated promising results (Fitches et al., 2002; Fitches et al., 2004a; Fitches et al., 2004b; Pham-Trung et al., 2006) and it has therefore been recognised that this method offers great potential for the development of further novel, orally active, environmentally benign insecticides.

1.4. *Pichia Pastoris*

The use of yeast for protein expression by scientists is now common practice. As well as being the first eukaryotic genome to be sequenced, *Saccharomyces cerevisiae* is the most studied yeast, a scientific model organism and was until recently the preferred choice for recombinant protein expression (Gellissen and Hollenberg, 1997). It was the discovery of certain yeast species, principally *Pichia pastoris*, which could utilize methanol solely as a source of carbon and energy by Koichi Ogata (Ogata et al., 1969) that changed this view (Cereghino and Cregg, 2000). Early attempts by Phillips Petroleum Company to develop *P. pastoris* as a protein expression system were impeded by an increase in the cost of methanol due to the 1970s oil crisis (Wegner, 1990). A decade later, a collaborative partnership was successful in establishing *P. pastoris* as an organism for heterologous protein expression. Since then, numerous vectors, strains and manipulation protocols have been developed (Cereghino and Cregg, 2000).

Pichia pastoris is now becoming a more frequently used yeast as opposed to *S. cerevisiae* for the following reasons; (i) the Alcohol Oxidase and GAP promoters are well suited for foreign protein expression; (ii) it is faster and as easy to manipulate as *S. cerevisiae* and even *Escherichia coli*; (iii) it has the major advantage that although it is a single-celled eukaryotic micro-organism, it is capable of the many post-translational modifications performed by higher eukaryotic cells, such as proteolytic processing, folding, disulphide bond formation and glycosylation; (iv) it can withstand the high cell densities involved in culturing, therefore giving higher protein yields (Cereghino and Cregg, 2000).

The pPIC *P. pastoris* expression system is based on its ability to metabolize methanol. The metabolic pathway involves a unique set of enzymes, the most important of these being Alcohol Oxidase (AOX), catalysing the first step of oxidation of methanol giving formaldehyde and hydrogen peroxide.



$$CH_3OH + O_2 \rightarrow HCHO + H_2O_2$$

Alcohol Oxidase is encoded by two genes AOX1 and AOX2. The majority of the activity arises from AOX1, its expression being controlled by transcription. The level of transcription increases with the presence of methanol (Cereghino and Cregg, 2000).

There are numerous *P. pastoris* phenotypes arising as a result of recombinant transformation and the two AOX genes. Wild type *P. pastoris* has the AOX1 and AOX2 genes and therefore primarily uses the AOX1 gene for its methanol metabolism, giving rise to the so-called Mut^+ phenotype (methanol utilization plus), growing at the wild type rate. When pPIC vectors are transformed into the *P. pastoris* genome, the two AOX genes can be maintained intact, therefore the yeast retains the Mut^+ wild type growth rate. Recombination however gives the possibility of disturbing the AOX genes. A disruption of the AOX1 gene promotes the use of AOX2, the weaker of the two genes. The metabolism and growth on methanol is, therefore, less efficient and slower so giving rise to the Mut^S phenotype (methanol utilization slow). A third phenotype exists for *P. pastoris* that can not grow on methanol due to a lack of both AOX genes; this is known as Mut^- (Macauley-Patrick et al., 2005).

The promoter for the AOX1 gene has been isolated and used in vectors such as pPICZ and pPICZα. Insertion of a foreign gene into the vector with this promoter would therefore allow it to be inducible with methanol. This vector series is advantageous in the fact that a foreign protein would be repressed in initial growth media and induced when transferred to methanol media. This is particularly useful when the foreign protein is toxic to the yeast cells when present at the levels involved in recombinant expression (Cereghino and Cregg, 2000). The use of the AOX1 promoter and methanol induction is not always appropriate however. In large scale industrial fermentation for instance, the large volumes of methanol required are considered a fire hazard (Cereghino and Cregg, 2000; Macauley-Patrick et al., 2005). In this situation alternative promoters exist. The main one of these is the constitutive glyceraldehyde-3-phosphate dehydrogenase (GAP) gene promoter (Waterham et al., 1997). The GAP promoter is used in vectors such as pGAPZ and pGAPZα. The significant differences of the pGAP *P. pastoris* system are constitutive expression and the utilization of glycerol for its growth. Due to this constitutive expression, the use of the GAP promoter is not a good choice for expression of proteins toxic to yeast (Cereghino and Cregg, 2000).

The large range of vectors available for *P. pastoris* also allows the user to select between intracellular expression of heterologous protein or secretion into the extracellular media. Where possible, secretion is the best option as it aids the purification process because *P. pastoris* secretes very low levels of endogenous proteins and the culture medium also contains no added proteins. Secreted heterologous protein therefore comprises the majority of the protein in the external medium and is already separated from cellular proteins. To facilitate secretion, a secretion signal is required to ensure the foreign protein is targeted to the secretory pathway. The most successful secretion signal to date is the α-mating factor of *S. cerevisiae* which contains both a signal peptide (pre-sequence) and a secretion targeting sequence (pro-sequence). Other secretion signals do exist, for example BAR1 and HEX4, but these are used less frequently (Cereghino and Cregg, 2000).

All of the advances in the use of the *P. pastoris* system make it an attractive alternative expression host to *S. cerevisiae*. Heterologous protein production can be optimized using the various strains and range of vectors available to suit the nature of the foreign protein. Regardless of all of these developments, the main limitation of the *P. pastoris* system remains cleavage of recombinant protein due to proteolysis. Although *P. pastoris* can tolerate high cell densities, a small amount of cell death is inevitable. The dead cells have a tendency to lyse thus releasing vacuolar proteases which are able to degrade heterologous protein. This can be kept to a minimum by utilizing protease deficient strains, but it has yet to be eliminated completely (Jahic et al., 2006).

Numerous examples describing the use of *P. pastoris* are publicised in scientific literature, demonstrating its versatility in heterologous protein expression. Gurkan and Ellar (2003) expressed a synthetic *Bacillus thuringiensis* Cyt2Aa1 toxin (syncyt2Aa1). Syncyt2Aa1 is a delta-endotoxin which has a specific insecticidal activity by forming pores within endogenous membranes. Fitches et al. (2004a) designed a spider venom neurotoxin (SFI1)-GNA lectin fusion protein and expressed it in *P. pastoris* with yields of 0.5-5mg per litre of culture medium. The recombinant fusion protein was shown to be toxic to *L. oleracea* larvae by injection (100% mortality after 48 hours) and oral delivery (100% mortality over 6 days with 2.5mg of fusion protein). Chen et al. (2004) expressed a recombinant mungbean defensin (rVrD1) in *P. pastoris*, obtaining levels of around 3mg per litre of culture medium. The growth inhibitory activity of the plant defensin was demonstrated in four plant

fungal species; *Fusarium oxysporum*, *Pyricularia oryza*, *Rhizoctonia solani* and *Trichophyton rubrum*. It was also subsequently shown to inhibit the development of bruchid beetle larvae. Pham-Trung et al. (2006) prepared a red scorpion toxin (ButaIT)-GNA lectin fusion protein in *P. pastoris*, with expression levels of 25-35mg per litre of culture medium being obtained. Its acute toxicity upon injection into *L. oleracea* larvae was demonstrated alongside its chronic toxicity when delivered orally. It was also shown to be toxic to *Nilaparvata lugens* (rice brown planthopper) when fed in artificial liquid diet.

1.5. Avidin

Previously, an avidin gene construct was prepared for use in experiments as a fusion with garlic lectin (ASA; *Allium sativum* Agglutinin) to be used for the transport of biotinylated proteins in insects.

Avidin is a basic glycoprotein isolated from chicken egg white and the tissues of birds. It has a homo-tetrameric structure, with each subunit comprising of 128 amino acids, being approximately 16kDa in size (DeLange and Huang, 1971) and having an isoelectric point of 10.0 (Woolley and Longsworth, 1942; Savage et al., 1992). Avidin was originally discovered due to its very high affinity for biotin (vitamin H) (Gyorgy et al., 1941), having a dissociation constant of 10^{-15} (Savage et al., 1992). It therefore works by sequestering biotin, rendering it unavailable for use as an enzyme activity cofactor. Targets include carboxylation, decarboxylation and transcarboxylation reactions (Kramer et al., 2000; Samols et al., 1988; Knowles, 1989). Dr Dave Bown used *Gallus gallus* (chicken) to obtain the avidin gene construct. The avidin gene was found to contain four exons, with the majority of the mature polypeptide coding sequence being in exon 2 and exon 3. Exons 2 and 3 were amplified separately from chick embryo tissue genomic DNA using designed primers. To ensure correct processing, TOPO-TA cloning of the exons into pCR2.1 intermediate vector was carried out separately. Once sequenced and confirmed, exons 2 and 3 were excised and ligated together in the pUC18 vector. The combined exons were then assembled with oligonucleotides containing the exon 4 sequence and cloned directly into the pGAPZαB yeast expression vector. A clone containing the complete expression construct (Figure 3) was selected and used for the garlic lectin fusion. Work on this fusion suggested that avidin could potentially be used

alone as an alternative carrier to GNA since it was detected in the haemolymph of *L. oleracea* larvae following ingestion of avidin containing diet (Section 3.3).

```
                         ATGAGATTTCCTTCAATTACTGCTGTTTTATTCGCAGCATCC   -267
TCCGCATTAGCTGCTCCAGTCAACACTACAACAGAAGATGAAACGGCACAAATTCCGGCTGAAGCTGTCATCGGT   -225
TACTCAGATTTAGAAGGGGATTTCGATGTTGCTGTTTTGCCATTTTCCAACAGCACAAATAACGGGTTATTGTTT   -150
ATAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTATCTCTCGAGAAAAGAGAGGCTGAAGCTGCA   -75

GCTAGAAAATGCTCGCTGACTGGGAAATGGACCAACGATCTGGGCTCCAACATGACCATCGGGGCTGTGAACAGC   +75
 A  R  K  C  S  L  T  G  K  W  T  N  D  L  G  S  N  M  T  I  G  A  V  N  S     25

AAAGGTGAATTCACAGGCACCTACACCACAGCCGTAACAGCCACATCAAATGAGATCAAAGAGTCACCACTGCAT   +150
 K  G  E  F  T  G  T  Y  T  T  A  V  T  A  T  S  N  E  I  K  E  S  P  L  H     50

GGGACACAAAACACCATCAACAAGAGGACCCAGCCCACCTTTGGCTTCACCGTCAATTGGAAGTTTTCAGAAAGT   +225
 G  T  Q  N  T  I  N  K  R  T  Q  P  T  F  G  F  T  V  N  W  K  F  S  E  S     75

ACTACTGTCTTCACGGGCCAGTGCTTCATAGACAGGAACGGGAAGGAGGTCCTGAAGACCATGTGGCTGCTGCGG   +300
 T  T  V  F  T  G  Q  C  F  I  D  R  N  G  K  E  V  L  K  T  M  W  L  L  R    100

TCAAGTGTTAATGACATTGGTGATGACTGGAAAGCTACGCGTGTTGGTATCAACATCTTCACTAGATTGAGAACT   +375
 S  S  V  N  D  I  G  D  D  W  K  A  T  R  V  G  I  N  I  F  T  R  L  R  T    125

CAAAAGGAAATCGACCATCATCATCATCATCATTGA                                          +411
 Q  K  E  I  D  H  H  H  H  H  H  .                                            136
```

Figure 3.
Nucleotide and amino acid sequence of recombinant avidin obtained from *G. gallus*. The coding sequence (amino acids 1-128) is highlighted in yellow and outlined in red is the His-tag (amino acids 131-136) used for purification by nickel affinity chromatography.

Avidin has been demonstrated to be orally insecticidal to a number of lepidopteran, dipteran and coleopteran insect pests (Levinson and Bergmann, 1959; Tsiropoulos, 1985; Bruins et al., 1991; Levinson et al., 1992; Morgan et al., 1993; Du and Nickerson, 1995; Allsopp and McGhie, 1996; Markwick et al., 2001; Cheng Zhu et al., 2005) and transgenic avidin-expressing maize, corn and rice have been cultivated with success (Kramer et al., 2000; Friends of the Earth communication, 2002; Yoza et al., 2005).

1.6. Project Aims

The aims of this MSc research project are; (i) to carry out further screens of the *E. pennicornis* venom gland cDNA library by sequencing DNA obtained from individual clones of the library; (ii) to investigate the potential use of reprolysin-like metalloprotease as an insecticidal component of a fusion protein by expressing truncated versions of the protein in *P. pastoris* and carrying out bioassays in *L. oleracea* larvae; (iii) to verify and examine the potential use of recombinant avidin as an alternative carrier protein in Fusion Protein Technology by firstly optimising the purification of recombinant avidin and secondly exploring all aspects of any insecticidal action and transport using bioassays with *Mamestra brassicae* (cabbage moth) larvae.

2.0. Materials and Methods

2.1. Biological and Chemical Reagents

All chemical reagents utilized were supplied by BDH (from VWR), Sigma or Merck unless otherwise stated. These were of the highest analytical grade commercially available.

For bacterial work, One Shot TOP10 Electrocomp *Escherichia coli* cells of the genotype F- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*araleu*)7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G were used (supplied by Invitrogen).

For yeast work, Easycomp protease deficient SMD1168H of the genotype *pep*4 and X33 wild type strains of *Pichia pastoris* were used (supplied by Invitrogen).

2.2. Laboratory Techniques

All procedures adopted for the routine molecular biology experimental work were based upon those standard protocols found in Molecular Cloning (Sambrook and Russell, 2001).

2.2A. *Eulophus pennicornis* Venom Gland cDNA Library Screening and Production of Reprolysin-like Metalloprotease Structural Variants

An *E. pennicornis* venom gland cDNA library was prepared by Dr Daniel Price using a Creator SMART cDNA Library Construction Kit (Clontech) and transformed into TOP10 *E. coli* cells. Aliquots of the unamplified cDNA library (2.8×10^5 cfu) were stored at -80^oC and thawed when required.

The pGAPZαB expression vector with or without GNA for the reprolysin-like metalloprotease variants was prepared from ButaIT:GNA pGAPZαB in TOP10 *E. coli* as described by Pham-Trung et al. (2006).

2.2A-1. Bacterial Cell Culture

For the culturing of bacterial cells, LB media (1% [w/v] trypticase peptone, 0.5% [w/v] yeast extract, 1% [w/v] sodium chloride), originally described by and named after its founders Luria and Bertani, was used (Bertani, 1951). Where cell growth was required, LB agar was used. This is identical to the above with the addition of 1.5% [w/v] agar. LB agar was autoclave sterilized before use. Relevant antibiotic was added on cooling. LB agar was poured into sterile agar plates (supplied by Sterilin) within a 70% [v/v] ethanol treated LAF cabinet and allowed to set. Plating out was performed using an ethanol flamed loop or glass spreader within the proximity of a Bunsen burner to maintain sterility. Autoclave sterilized LB media was used for the propagation of bacterial cells. For analysis, 5ml aliquots in McCartney bottles were prepared (minipreps). Relevant antibiotics were added followed by inoculation with either an aliquot of a glycerol stock or a bacterial colony transferred from an agar plate. Again, sterile technique was used. Bacterial plates and cultures were incubated overnight at 37°C (cultures were secured on an orbital shaker set at 200rpm). Antibiotics used in bacterial cell culturing included chloramphenicol, kanamycin and Zeocin (50μg/ml and 100μg/ml respectively). A low salt LB media containing only 0.5% [w/v] sodium chloride was necessary to maintain Zeocin activity.

Serial dilutions of the venom gland cDNA library were plated out and individual colonies chosen at random for propagation and analysis. Plates of ButaIT:GNA pGAPZαB were streaked out and colonies chosen for propagation.

2.2A-2. Plasmid DNA Isolation

Plasmid DNA was extracted from propagated cells using Promega Wizard Plus SV Miniprep DNA Purification Systems. The manufacturers instructions supplied with the kit were followed. The use of nuclease free water, sterile Eppendorf tubes and sterile pipette tips are fundamental during DNA work to ensure unwanted DNA digestion is prevented.

2.2A-3. Polymerase Chain Reaction (PCR)

Fragments of DNA were amplified by PCR. The three reprolysin-like metalloprotease structural variant constructs (Rep 1, Rep 2 and Rep 3) (Figures 14, 15 and 16) were generated utilizing the original reprolysin-like metalloprotease

pGAPZαB sequence (Figure 1) (supplied by Dr Daniel Price) as a template in PCR. Oligonucleotide primers were designed for the purpose (Table 1). For amplification, Finnzymes Phusion High-Fidelity DNA Polymerase (supplied by New England Biolabs) or Taq Polymerase (supplied by Promega) was used. Oligonucleotide primers were supplied by Sigma Genosys and solubilised in nuclease free water as detailed in the manufacturer quality control sheet. Promega supplied the dNTPs. Standard 50μl reactions were prepared (final concentrations of 0.2μM primers, 0.2mM dNTPs, 1x PCR buffer, 1.5mM $MgCl_2$, 1U polymerase). For Phusion polymerase, the final concentration of the primers was 0.5μM as recommended in the manufacturer instructions. Standard cycling conditions used are listed in tables 2 and 3. These were sometimes modified to suit the melting temperature (T_m) of the PCR primers. Samples were covered with mineral oil to prevent evaporation and a Biometra Trio or Perkin Elmer GeneAmp PCR System 2400 thermocycler was used. Where required, PCR products were purified using Amicon Microcon-PCR filters according to the manufacturer instructions. PCR products were then visualized by DNA agarose gel electrophoresis (Section 2.2A-5).

Primer	Sequence	Enzyme site
Reprolysin 1 5´	AGCCTGCAGGTCAGTATAGCGAATCGCAG	*Pst*I
Reprolysin 1 3´	ATGCGGCCGCCTGCTTTCGATAGGTATCTTT	*Not*I
Reprolysin 2 5´	AGCCTGCAGGTTTTTATCCGAAACTCCTTGTAC	*Pst*I
Reprolysin 2 3´	ATGCGGCCGCCTGCTTTCGATAGGTATCTTT	*Not*I
Reprolysin 3 5´	AGCCTGCAGGTTTTTATCCGAAACTCCTTGTAC	*Pst*I
Reprolysin 3 3´	TATTCTAGATGCTTTCGATAGGTATCTTT	*Xba*I

Table 1.
Oligonucleotide primers used to generate the reprolysin-like metalloprotease structural variants. Enzyme restriction sites are underlined.

0	94°C	30s	
1	94°C	10s	x15 Cycles
2	50°C	30s	
3	72°C	2 min	
4	72°C	5 min	

Table 2.
Cycling conditions used for Taq polymerase PCR reactions.

0	98°C	30s	
1	98°C	10s	x15 Cycles
2	52°C	30s	
3	72°C	45s	
4	72°C	5 min	

Table 3.
Cycling conditions used for Phusion polymerase PCR reactions.

Table 1 illustrates that the primer pairs contained either, 5′ *Pst*I and 3′ *Not*I (Rep 1 and Rep 2) or 5′ *Pst*I and 3′ *Xba*I (Rep 3) enzyme restriction sites to allow insertion into similarly digested expression vector. Due to difficulties in cloning directly, sub-cloning to produce the three constructs was carried out using the TOPO cloning kit (purchased from Invitrogen), containing the pCR2.1-TOPO vector, following manufacturer instructions.

2.2A-4. Restriction Endonuclease Digests

Following propagation (Section 2.2A-1) and DNA isolation (Section 2.2A-2), cDNA library plasmids were restricted with *Eco*RI and *Xba*I prior to insert confirmation. ButaIT and ButaIT:GNA were digested from the complete isolated ButaIT:GNA pGAPZαB construct using *Pst*I, *Not*I and *Pst*I, *Xba*I respectively. Rep 1, Rep 2 and Rep 3 PCR products were restricted using the enzyme sites contained in the primers (Table 1) in preparation for ligation into the prepared expression vector.

Restriction enzymes, supplied by Promega, were chosen on the basis of plasmid maps and DNA sequences. Enzyme activity buffers were selected after consulting buffer charts within the manufacturer manual. Where possible buffers promoting 100% activity were chosen, however, where this was impossible, the next best were used. Digests were placed in a Grant heat block pre-set to 37°C. Although three hours is the minimum recommended time, the majority were allowed to restrict overnight.

2.2A-5. DNA Agarose Gel Electrophoresis

Submarine Agarose Gel Electrophoresis was used to confirm inserts and visualize DNA. Appropriate gel frames and well combs were used to accommodate the number of samples. The percentage [w/v] of agarose (supplied by Melford) used was varied between 1% and 1.5% depending on the size of the DNA fragments. Gel solutions were produced using 1x [v/v] TAE from a 50x TAE laboratory stock (2M tris-HCl, 2M glacial acetic acid, 50mM EDTA pH 8.0), heated in the microwave until boiling clear and ethidium bromide (10mg/ml) added at a concentration of 0.5μg/ml on cooling. Gels were allowed to set for a minimum of thirty minutes. Running buffer consists of 1x [v/v] TAE and 0.5μg/ml ethidium bromide in purified water. Samples were mixed with 6x Orange G loading dye (10mM tris-HCl pH 7.6, 0.15% [w/v] Orange G, 60% [v/v] glycerol, 60mM EDTA). Fermentas λ/Eco471 was

used as a marker of size. Gels were run at 100V, 0.05-0.1A maximum current. Complete gels were observed using a Geneflash Syngene UV cabinet with Pulnix camera and captured using a Sony Video Graphic UP-895MD printer.

Random cDNA library plasmids with inserts greater than 500bp were sent for DNA sequencing in both directions using vector specific primers.

2.2A-6. Recovery of DNA from Agarose Gels

Where DNA needed to be isolated from agarose gels, the required band was excised from the gel on a UVP inc. UV transilluminator, solubilised and purified using the Qiagen Gel Extraction Kit or Eppendorf Gel Cleanup Systems following manufacturer instructions.

2.2A-7. DNA Sequence Analysis

DNA sequencing was used to check the three intermediate and final reprolysin-like metalloprotease construct products.

Appropriate vector specific primers were supplied and sequencing was carried out using Applied Biosystems ABI Prism 3730 automated DNA sequencers by the DNA sequencing service (DBS Genomics), School of Biological and Biomedical Sciences, University of Durham. Sequence data was analysed using Sequencher software (version 4.5) running on Mac OS computers.

2.2A-8. Ligation Reactions

Correct Rep 1, Rep 2 and Rep 3 constructs were ligated into the prepared expression vector using T4 DNA ligase supplied by Promega. 10μl reactions were set up containing roughly equal amounts of pGAPZαB expression vector and Rep 1, Rep 2 or Rep 3 insert. Ligation reactions were either left at 22°C for three hours or overnight at 4°C.

2.2A-9. Transformation of Competent Bacterial Cells

For bacterial expression of the Rep 1, Rep 2 and Rep 3 pGAPZαB constructs, TOP10 electro-competent *E. coli* bacterial cells were used (Section 2.1). Cells were removed from -80°C storage and thawed on ice. Upon mixing with the required ligated plasmid, electroporation allowed plasmid uptake. Electroporation was carried out using a Biorad Gene Pulser system following recommendations by both

Invitrogen (the cell provider) and Biorad (*E. coli* electroporation protocol). Once transformed, cells were placed on a shaker at 37°C for one hour and then plated out as described in section 2.2A-1. Transformants were analysed by plasmid DNA isolation and DNA sequencing (Sections 2.2A-2 and 2.2A-7), or by colony PCR.

2.2A-10. Colony PCR

Colony PCR was used as a rapid method of confirming DNA inserts. Individual colonies were suspended in 10μl of sterile water by vortex mixing and boiled in a water bath for five minutes. Cell debris was then pelleted by a pulse spin in a bench top centrifuge. Supernatant was then analysed by PCR using both vector specific and gene specific primers (Section 2.2A-3).

2.2A-11. DNA Quantification

dsDNA absorbs light at a wavelength of 260nm. After Rep 1, Rep 2 and Rep 3 pGAPZαB bacterial expression and plasmid extraction, DNA quantification was carried out in preparation for yeast transformation by measuring the absorbance of a 1 in 10 diluted sample at λ_{260} using an Eppendorf BioPhotometer. A value for the amount of DNA was calculated knowing that 1AU is equivalent to 50μg/ml. 10μg of DNA for each construct was used for subsequent yeast preparation.

2.2A-12. Linearization of DNA

For yeast transformation, plasmid constructs were linearized using either *Bln*I (supplied by Roche) (for pGAP based constructs) or *Sac*I (for pPIC based constructs).

2.2A-13. Extraction and Precipitation of DNA

Linearized Rep 1, Rep 2 and Rep 3 pGAPZαB plasmid DNA was cleaned up by phenol-chloroform extraction and concentrated by ethanol precipitation (Sambrook and Russell, 2001). Precipitated DNA was resuspended in 5μl of nuclease free water and used to transform competent yeast cells.

2.2A-14. Transformation of Competent Yeast Cells

For yeast expression of the Rep 1, Rep 2 and Rep 3 pGAPZαB constructs, Easycomp SMD *P. pastoris* was used (Section 2.1). Cells were removed from -80°C

storage and thawed on ice. Heat shock transformation was carried out following the Easycomp manufacturer instructions. Transformants were analysed by western blot (Section 2.2B-7) or by colony PCR (Section 2.2A10).

2.2A-15. Bacterial and Yeast Glycerol Stocks

To maintain a supply of competent cells successfully transformed with the Rep 1, Rep 2 and Rep3 pGAPZαB constructs, glycerol stocks were prepared. Cells were transferred from those required and 5ml or 10ml analysis cultures set up as mentioned in section 2.2A-1 for bacteria and 2.2B-1 for yeast. 250μl of 60% [v/v] glycerol was mixed with 750μl of the overnight culture in 1.5ml cryogenic tubes, supplied by Thermo-Fisher Scientific. These were then snap frozen in liquid nitrogen and stored at -80°C.

2.2B. Expression of Reprolysin-like Metalloprotease, Reprolysin-like Metalloprotease Structural Variants and Avidin

Production of the original reprolysin-like metalloprotease was required for use as a control for the three structural variants. The original reprolysin-like metalloprotease pGAPZαB in SMD *P. pastoris* was supplied from -80°C storage by Dr Daniel Price.

The three reprolysin-like metalloprotease pGAPZαB structural variants in SMD *P. pastoris* had been produced as detailed in section 2.2A.

Avidin pGAPZαB in X33 wild type *P. pastoris* had been produced by Dr Dave Bown. It was supplied from -80°C storage by Dr Elaine Fitches.

2.2B-1. Yeast Cell Culture

For the culturing of yeast cells, YPG media (2% [w/v] trypticase peptone, 1% [w/v] yeast extract, 4% [v/v] glycerol) was used. Where yeast cell growth was required, YPG agar was used. This is identical to the above with the addition of 2% [w/v] agar. YPG agar plates were prepared and plated out in a similar manner to bacterial agar plates (Section 2.2A-1). Autoclave sterilized YPG media was used for the propagation of yeast cells. For analysis, 10ml aliquots in McCartney bottles were prepared. These were set up in a similar way to bacterial analysis cultures except for the addition of yeast glycerol stocks or yeast cells from a colony. Yeast plates and cultures were incubated for two-four days at 30°C (cultures were secured on an

orbital shaker set at 250rpm). For inducible yeast vectors, BMGY growth and BMMY expression media was used. BMGY consists of 2% [w/v] trypticase peptone, 1% [w/v] yeast extract, 100mM potassium phosphate pH 6.0, 1.34% [w/v] yeast nitrogen base, $4x10^{-5}$% [w/v] biotin and 1% [v/v] glycerol. BMMY differs in one respect only; the glycerol is replaced by 0.5% [v/v] methanol. This change induces protein expression. Stock solutions of 1M potassium phosphate pH 6.0, 10x YNB, 500x biotin, 10x glycerol and 10x methanol were prepared and media was produced as detailed in the Invitrogen *Pichia* manual. For analysis, 50ml BMGY and 100ml BMMY aliquots were prepared in baffled flasks and inoculated as described above. All other aspects of the propagation (BMGY media) and induction (BMMY media) of the cells were carried out following recommendations in the Invitrogen manual. The antibiotic used in yeast cell culturing was Zeocin (100μg/ml).

For purification of reprolysin-like metalloprotease after small scale analysis, cultures were scaled up to 1L of media in a 5L baffled flask. These were inoculated with 50ml starter cultures (grown for three days) to increase the initial cell growth rate. Incubation conditions were identical to those mentioned above. After small scale analysis, 3L and 7.5L New Brunswick Scientific BioFlo 110 bench-top fermenters were used for avidin purification. Either 1.25L or 3L of sterile minimal media (Higgins and Cregg, 1998) supplemented with PTM1 trace salts (Cino, 1999) was inoculated with three 50ml or three 100ml starter cultures respectively (grown for three days). Cultivation limits were set at 30°C, pH 4.0-5.0 and 30% dissolved oxygen with continuous agitation. A glycerol feed of 5-10ml/hour was maintained over three days.

The pGAPZα expression vector (supplied by Invitrogen) used for purification was specifically chosen so that recombinant protein was secreted into the surrounding media. Centrifugation of cultures for 25 minutes at 7000rpm, 4°C pelleted the cells so that supernatant could be processed further. Sodium chloride was added to the reprolysin-like metalloprotease culture supernatant giving a final concentration of 2M. Avidin culture supernatant was adjusted to 50mM sodium acetate pH 4.0 from a 20x stock solution and sodium chloride added to give a final concentration of 0.5M.

2.2B-2. Chromatography and Purification

Yeast supernatant containing reprolysin-like metalloprotease was filtered through 1.2 μm GF/C, 1μm GF/D and 0.7μm GF/F membrane discs (supplied by Whatman) in sequence and then run through an XK16/20 hydrophobic interaction Phenyl Sepharose column (supplied by Pharmacia – now GE Healthcare) at 2ml/minute. The column was sanitized and rinsed following manufacturer recommendations. It was then equilibrated with 2M salt buffer and protein allowed to load and re-circulate overnight. After a 2M salt wash, protein was eluted by linear reduction in the salt concentration using an FPLC system (also supplied by Pharmacia). Five millilitre fractions were collected and analysed using SDS-PAGE. Avidin culture supernatant, again filtered using the same filter sequence described above, was run through a HisTrap FF Crude 1ml column (again supplied by Pharmacia) for binding His tagged proteins, stripped and re-charged with nickel as per manufacturer recommendation. After equilibration with binding buffer (50mM sodium acetate, 0.5M sodium chloride pH 4.0), supernatant was loaded and allowed to re-circulate overnight at a flow rate of 1ml/minute. Following a binding buffer wash, recombinant avidin was eluted using 100mM sodium acetate, 300mM imidazole, 0.5M sodium chloride pH 7.4. The single peak was collected and analysed by SDS-PAGE. To maximise protein recovery from large volumes, supernatant flow-through was re-processed through the columns a minimum of three times.

2.2B-3. Sodium Dodecyl Sulphate Poly Acrylamide Gel Electrophoresis (SDS-PAGE) of Proteins

Sodium dodecyl sulphate poly-acrylamide gel electrophoresis with the discontinuous pH system originally described by Laemmli (1970) was used to visualize proteins. One millimetre gels were produced following the Qiagen protocol. Gel plates and running tanks were supplied by ATTO Corporation. The percentage [v/v] of acrylamide solution (supplied by National Diagnostics) varied between 12% and 15% depending on the size of the proteins. The majority of resolving gel solutions consisted of 15% [v/v] acrylamide, 0.2M tris-HCl pH 8.8, 0.1% SDS, 0.1% [w/v] ammonium persulphate and 0.5% [v/v] TEMED. Stacking gel solution consisted of 4% [v/v] acrylamide, 0.13M tris-HCl pH 6.8, 0.1% [w/v] SDS, 0.1% [w/v] ammonium persulphate and 0.75% [v/v] TEMED. 1x reservoir buffer was produced from a 10x reservoir buffer laboratory stock (0.25M tris-HCl, 1.92M

glycine, 1% [w/v] SDS). Samples were mixed with 5x SDS loading dye (312.5mM tris-HCl pH 6.8, 50% [v/v] glycerol, 10% [w/v] SDS, 0.01% [w/v] Bromophenol Blue, 25% [v/v] β-mercaptoethanol) and boiled in a water bath for 10 minutes. Sigma SDS 7 was used as a marker of size. Gels were run at 100V, 0.05-0.1A maximum current through the stacking gel and 200V, 0.05-0.1A maximum current through the resolving gel. Complete gels were stained for three hours and de-stained, using laboratory stocks of filtered Coomassie stain (0.04% [w/v] Coomassie Brilliant Blue, 40% [v/v] methanol, 7% [v/v] glacial acetic acid) and de-stain (40% [v/v] methanol, 7% [v/v] glacial acetic acid), or prepared for western blot (Section 2.2B-7). De-stained gels were scanned into a Mac OS computer and captured as TIFF files.

2.2B-4. Dialysis and Freeze Drying

For avidin, where volumes were appropriately small, buffer exchange to purified water was carried out by dialysis. Dialysis tubing with a molecular weight cut off of 12-14kDa was supplied by Medicell International Ltd. and prepared for use by boiling for 10 minutes in purified water containing ammonium hydrogen carbonate and a trace of EDTA. Tubing containing protein sample was double clamped. Dialysis was carried out at 4°C with gentle stirring in at least 100x greater volume of required buffer and replaced at least twice (morning and evening), ensuring initial buffer concentration was reduced to 1nM.

Following dialysis, the recombinant protein was frozen in liquid nitrogen, freeze-dried (FTS System Flexi Dry MP) and stored at 4°C.

2.2B-5. Protein Quantification

Recombinant protein concentration was estimated by SDS-PAGE (Section 2.2B-3), running various amounts alongside known concentrations of standard.

2.2B-6. Chloroform-Methanol Precipitation of Proteins

In some cases, for analysis by SDS-PAGE and western blot, protein samples were concentrated by chloroform-methanol precipitation as described by Wessel and Flügge (1984).

2.2B-7. Western Blotting

Western blotting originally reported by Towbin et al. (1979) was used as a more sensitive approach for visualising proteins. Samples for analysis were run on SDS-PAGE gels as described in section 2.2B-3, however, instead of staining and de-staining, gels were equilibrated with an aliquot of a laboratory stock solution of Bjerrum and Schafer-Nielsen buffer (48mM tris-HCl, 39mM glycine, 20% [v/v] methanol) (Bjerrum and Schafer-Nielsen, 1986). The semi-dry transfer (Kurien and Scofield, 2006) and detection was carried out following the Qiagen protein analysis benchguide. Nitrocelluose membrane (supplied by BDH - from VWR) and HorizBlot electroblotting systems were used (supplied by ATTO Corporation). Transfer was confirmed using a small amount of a laboratory stock of filtered Ponceau S stain (0.5% [w/v] Ponceau S, 1% [v/v] glacial acetic acid). Fresh solutions of block (1x PBS, 0.05% [v/v] Tween-20, 5% [w/v] milk powder) and rinse (1x PBS, 0.05% [v/v] Tween-20) were produced for every blot. Transferred proteins were probed with appropriate specific primary antibodies, including anti-Myc (1:5000 dilution), anti-GNA (1:3000 dilution), anti-avidin (1:5000 dilution) and anti-ASAII (1:3000 dilution) (supplied by Cell Signalling Technology) and allowed to incubate overnight. Peroxidase coupled secondary antibodies (1:5000 dilution) were left to bind for a minimum of three hours and visualized using ECL substrate (1M tris-HCl pH 8.5, 0.2mM coumaric acid, 1.25mM luminol and 3% [v/v] hydrogen peroxide) by exposure to X-Ray film (supplied by Fujifilm) in a film cassette for varying times. Films were developed using an automatic X-ograph Imaging systems Compact X4 developer.

2.2B-8. Gelatin Activity Gels

In-gel activity of reprolysin-like metalloprotease was shown using SDS-PAGE. 12.5% acrylamide gels containing 0.12% gelatin as a substrate according to Heussen and Dowdle (1980). Samples were prepared under non-reducing conditions, i.e. mixed with 5x SDS loading dye (312.5mM tris-HCl pH 6.8, 50% [v/v] glycerol, 10% [w/v] SDS, 0.01% [w/v] Bromophenol Blue) without β-mercaptoethanol present and no boiling. Again, SDS 7 (Sigma) was used as a marker of size. Gels were run at 100V, 0.05-0.1A maximum current at 4°C. Complete gels were incubated in 1% Triton X-100 for one hour at room temperature to remove SDS. Following this, the active study gels were incubated overnight in 20mM tris-HCl pH 7.0, 1mM zinc

sulphate at 37°C, whereas inhibition study gels were incubated overnight in 20mM tris-HCl pH 7.0, 5mM EDTA at 37°C. Gels were stained and de-stained in a similar fashion to SDS-PAGE gels (Section 2.2B-3). Again, de-stained gels were scanned into a Mac OS computer and captured as TIFF files.

2.2C. Coupling of Antibody to Avidin

Sugar transporter IgG primary antibody raised in rabbit was supplied from -20°C storage by Dr Daniel Price. It was quantified by SDS-PAGE prior to use (Section 2.2B-3).

2.2C-1. Biotinylation

Antibody biotinylation was carried out using the Amersham Biosciences (now GE Healthcare) protein biotinylation module following manufacturer instructions. Antibodies were initially dialysed into sodium bicarbonate as recommended; using 10kDa molecular weight cut off Slide-A-Lyzer dialysis cassettes (supplied by Pierce) due to the small volume. Following biotinylation, the antibodies were dialysed into PBS, making them suitable for ingestion by insects. 0.4mg of recombinant avidin was added to 1mg of antibodies in 1ml of PBS and allowed to associate, making it ready for inclusion into artificial diet.

2.2D. Fluorescent labelling of Avidin-Biotin conjugate

Fluorescein-biotin was used to fluorescently label avidin (binding via the biotin). The mass of avidin (16kDa subunit) is 25 times greater than that of fluorescein-biotin (644.70Da), with avidin binding biotin in a 1:1 ratio. Fluorescence was measured with 100μl aliquots in black 96 well microtitre plates (supplied by Greiner bio-one) using a Labsystems Fluoroskan Ascent running on a PC with Ascent software version 1.3.3. Fluorescence measurements of a serial dilution of 1mg/ml fluorescein-biotin allowed the detection limit to be calculated. Avidin was dissolved in PBS, fluorescein-biotin added and allowed to bind at room temperature wrapped in foil with gentle agitation. A slight excess of Avidin was added (25:1) to ensure complete binding.

2.2D-1. Gel Filtration

Haemolymph from sixth stadium *M. brassicae* larvae was extracted and separated through a PBS equilibrated Superdex 10/300 GL gel filtration column packed with Superdex 200 resin (supplied by Pharmacia – now GE Healthcare). A Pharmacia (now GE Healthcare) FPLC was utilized, with a flow rate of 0.1ml/minute. Four hundred microlitre fractions were collected and measured for fluorescence. Haemolymph spiked with fluorescein-biotin and avidin bound fluorescein-biotin were also separated in this way as controls in preparation for feeding droplets of avidin bound fluorescein-biotin (Section 2.4B) to determine if the biotin, and therefore fluorescence remains bound to the avidin on transfer to the haemolymph.

2.3. Insect Cultures

Lacanobia oleracea (tomato moth) were maintained on artificial diet (Bown et al., 1997) under controlled environmental conditions of 25°C, 40% relative humidity and a 16 hour light: 8 hour dark regime.

Mamestra brassicae (cabbage moth) were reared continuously on artificial diet (Bown et al., 1997) under controlled environmental conditions of 25°C, 40% relative humidity and a 16 hour light: 8 hour dark regime.

Sitobion avenae (cereal aphid) were kept on four week old oat plants under controlled environmental conditions of 20°C and a 16 hour light: 8 hour dark regime.

These insects were chosen as representatives of major UK pest assemblages.

2.4. Insect Bioassays

2.4A. Injection bioassays

i. Avidin Insecticidal Activity

Purified recombinant avidin was tested for biological activity by injecting 5μl of sample, containing 50μg of protein from a freeze-dried stock re-suspended in PBS, into day one fifth stadium *M. brassicae* larvae of similar size. Ten larvae were injected with sample, 10 injected with equivalent amounts of Sigma avidin as a comparison, 10 injected with equivalent amounts of PBS as a negative control (previously shown to have no effect itself) and maintained in clear plastic pots (five

larvae per pot). Fresh artificial diet was added as necessary and larvae were monitored over the following seven days. Recombinant avidin had been quantified using SDS-PAGE (Section 2.2B-3).

2.4B. Artificial diet bioassays

i. Avidin Transport

A powdered optimal artificial diet (supplied by Bio-Serv) was used for oral delivery of avidin to *M. brassicae* larvae for monitoring its transportation. Two and a half milligrams (500ppm) of recombinant avidin (quantified by SDS-PAGE) was re-suspended in water and incorporated into 5g wet weight of diet. A diet containing Sigma avidin for comparison was similarly prepared and water only diet was used as a negative control. Day one fifth stadium larvae were exposed to one of the three diets for 24 hours after which haemolymph was extracted and larvae dissected to obtain gut contents and midgut tissue (discussed later). Twelve larvae were used for each of the three treatments (recombinant avidin, Sigma avidin and water control) and were maintained in clear 200ml pots (four larvae per pot).

ii. Avidin Insecticidal Activity

Optimal artificial diet was used for the oral delivery of avidin to *M. brassicae* larvae to elucidate its biological activity. Treatments consisted of 5mg (1000ppm), 0.5mg (100ppm) and 0.05mg (10ppm) of recombinant avidin (quantified by SDS-PAGE) re-suspended in water and incorporated into 5g wet weight of diet. A 5mg (1000ppm) Sigma avidin treatment diet was prepared for comparison. Five milligrams of casein was added into the negative control diet as a protein equivalent to ensure the increased concentration of protein in the test diet was not the cause of any observed effects. Twenty, neonate larvae were placed on each of the five treatments (recombinant avidin, Sigma avidin and casein control) and were kept in clear 200ml pots (one pot containing 20 larvae for each treatment). Survival was monitored daily and fresh diet was provided as required. Weights were measured once larvae were large enough (around day nine).

iii. Mechanics of Avidin Fluorescein-Biotin Transport

Newly moulted sixth stadium *M. brassicae* larvae, starved for 24 hours, were fed two 5μl droplets of either avidin bound fluorescein-biotin or fluorescein-biotin,

larvae per pot). Fresh artificial diet was added as necessary and larvae were monitored over the following seven days. Recombinant avidin had been quantified using SDS-PAGE (Section 2.2B-3).

2.4B. Artificial diet bioassays

i. Avidin Transport

A powdered optimal artificial diet (supplied by Bio-Serv) was used for oral delivery of avidin to *M. brassicae* larvae for monitoring its transportation. Two and a half milligrams (500ppm) of recombinant avidin (quantified by SDS-PAGE) was re-suspended in water and incorporated into 5g wet weight of diet. A diet containing Sigma avidin for comparison was similarly prepared and water only diet was used as a negative control. Day one fifth stadium larvae were exposed to one of the three diets for 24 hours after which haemolymph was extracted and larvae dissected to obtain gut contents and midgut tissue (discussed later). Twelve larvae were used for each of the three treatments (recombinant avidin, Sigma avidin and water control) and were maintained in clear 200ml pots (four larvae per pot).

ii. Avidin Insecticidal Activity

Optimal artificial diet was used for the oral delivery of avidin to *M. brassicae* larvae to elucidate its biological activity. Treatments consisted of 5mg (1000ppm), 0.5mg (100ppm) and 0.05mg (10ppm) of recombinant avidin (quantified by SDS-PAGE) re-suspended in water and incorporated into 5g wet weight of diet. A 5mg (1000ppm) Sigma avidin treatment diet was prepared for comparison. Five milligrams of casein was added into the negative control diet as a protein equivalent to ensure the increased concentration of protein in the test diet was not the cause of any observed effects. Twenty, neonate larvae were placed on each of the five treatments (recombinant avidin, Sigma avidin and casein control) and were kept in clear 200ml pots (one pot containing 20 larvae for each treatment). Survival was monitored daily and fresh diet was provided as required. Weights were measured once larvae were large enough (around day nine).

iii. Mechanics of Avidin Fluorescein-Biotin Transport

Newly moulted sixth stadium *M. brassicae* larvae, starved for 24 hours, were fed two 5μl droplets of either avidin bound fluorescein-biotin or fluorescein-biotin,

both containing 0.5% [w/v] sucrose as a 'sweetener' (10 larvae per treatment). Larvae were then allowed to feed on artificial diet for one hour in the dark before haemolymph was extracted, separated by gel filtration and fluorescence measured.

iv. Avidin-Antibody Coupling

Biotinylated sugar transporter antibody coupled to avidin was included in artificial diet for oral delivery to *M. brassicae* larvae and *S. avenae* to monitor its transportation. *Sitobion avenae* artificial diet was identical to that described in Prosser and Douglas (1992). Twenty four hours before the bioassay, 200 *S. avenae* were placed in Perspex aphid feeding chambers with 100μl artificial diet sandwiched between two sheets of parafilm, to allow adjustment to the new environment and diet overnight (50 aphids per chamber). For *S. avenae*, 10μl of antibody coupled avidin was mixed with 90μl of aphid diet. Non-coupled sugar transporter antibody and avidin were added to diet similarly as negative controls. Fifty aphids were used for each of the three treatments (10 aphids per chamber). Aphids were exposed to the diet for 66 hours and then haemolymph was extracted using glass capillaries. For *M. brassicae*, 520μl of antibody coupled avidin was incorporated into 2.5g wet weight of diet. Again, non-coupled sugar transporter antibody was added to diet as a negative control. Four larvae were used for each treatment and exposed to the diet for 24 hours in clear pots then haemolymph extracted and dissected to obtain gut contents and midgut (discussed later). Two extra larvae were allowed to remain on the antibody coupled avidin diet for 66 hours then dissected similarly to the above.

2.4C. Haemolymph, Gut contents and Midgut extraction and analysis

Haemolymph, gut contents and midgut samples were extracted and processed as described in Fitches et al. (2001) with the exceptions that intact midguts were flushed with water and the extraction buffer was PBS containing 1mM DTT. The protein concentration of the gut contents and midgut processed supernatants were estimated by a microtitre-based BCA assay (reagents supplied by Pierce) using BSA in PBS as the standard protein, following the manufacturer instructions. Equal amounts of protein and haemolymph were analysed for avidin presence by western blotting (Section 2.2B-7).

3.0. Results

3.1. *Eulophus pennicornis* Venom Gland cDNA Library Screen

Parasitic wasp venoms are a rich source of proteins, many of which have yet to have their function elucidated. To identify new genes encoding proteins with potential for use as novel insecticides, an *E. pennicornis* venom gland cDNA library was screened.

Dilutions of the *E. pennicornis* cDNA library were plated out, individual colonies chosen at random, plasmid DNA isolated, restriction digested and inserts visualized on DNA agarose gels. Fifteen random plasmids with inserts greater than 500bp were then sent for sequencing and results subjected to NCBI-BLAST sequence similarity searches.

Around half of the inserts sequenced showed either no obvious open reading frames or open reading frames with little similarity to the sequence database. Those with significant results are summarised in table 4 and the amino acid sequences can be seen in figures 4, 5, 6, 7 and 8. The NCBI-BLAST result for insert one was a juvenile hormone-inducible protein (*Nasonia vitripennis*, parasitoid wasp) (Accession number XP_001604874) with a score (number of matching nucleotide bases) of 229 and an e value (occurrence due to chance) of $2x10^{-58}$. The NCBI-BLAST result for insert two showed a reprolysin- like metalloprotease (*Nasonia vitripennis*, parasitoid wasp) (Accession number XP_001602775) with a score of 212 and an e value of $5x10^{-53}$. The NCBI-BLAST result for insert three was a splicing factor (*Aedes aegypti*, mosquito) (Accession number ABF18358) with a score of 196 and an e value of $2x10^{-48}$. The NCBI-BLAST result for insert four showed a trans-membrane serine protease (*Gallus gallus*, chicken) (Accession number XP_425880) with a score of 81 and an e value of $8x10^{-14}$. The NCBI-BLAST search also detected a trypsin-like serine protease conserved domain. Further structural analysis using the CBS-TMHMM transmembrane prediction server later showed there was no trans-membrane domain present. The NCBI-BLAST result for insert five was a neprilysin-like protein (*Venturia canescens*, parasitoid wasp) (Accession number AAL91975) with a score of 108 and an e value of $2x10^{-22}$. NCBI-BLAST also detected a conserved peptidase_M13_N domain. Due to this insert being larger, the complete

sequence was not obtained (sequence sections obtained are highlighted in red in figure 8). However, the 5′ end of the sequence contained a single *Sal*I restriction site (highlighted in blue in figure 8), allowing a small fragment to be excised and the plasmid re-ligated. Sequencing with the fragment removed revealed the central section of the insert.

Sequence	NCBI-BLASTx Result
Insert 1	Score = 229 e = $2x10^{-58}$ Juvenile hormone-inducible protein (*Nasonia vitripennis*, parasitoid wasp) Accession: XP_001604874
Insert 2	Score = 212 e = $5x10^{-53}$ Reprolysin-like metalloprotease (*Nasonia vitripennis*, parasitoid wasp) Accession: XP_001602775
Insert 3	Score = 196 e = $2x10^{-48}$ Splicing factor (*Aedes aegypti*, mosquito) Accession: ABF18358
Insert 4	Score = 81 e = $8x10^{-14}$ Trans-membrane serine protease (*Gallus gallus*, chicken) Accession: XP_425880
Insert 5	Score = 108 e = $2x10^{-22}$ Neprilysin-like protein (*Venturia canescens*, parasitoid wasp) Accession: AAL91975

Table 4.
Summary of the best hit NCBI-BLAST sequence similarity search results for an *E. pennicornis* venom gland cDNA library.

```
                                                  CATGAATCGAAGAA    -90
CAAGCTGCGATTATTCAGGAGTCGAACGCCTTCAACAAAGAATCGAATTTTTATAAATACATTGTTCCAGAACTG    -75

ATGAAGGAGGTAAAAGATCAATCTTGGCTTGCAAAGTGCTTTTTGGCTAAGGAAGACGTACTAATTTTGGAGAAT    +75
 M  K  E  V  K  D  Q  S  W  L  A  K  C  F  L  A  K  E  D  V  L  I  L  E  N     25

CTGACGTTCAAGAATTATGCGATCAAGACGAAAATATTAGATTTCGTAGCAATGAAATCAGCTTTAAACGCGCTA    +150
 L  T  F  K  N  Y  A  I  K  T  K  I  L  D  F  V  A  M  K  S  A  L  N  A  L     50

GCAAAATTTCGCGCTTCATCCATCTTAGCAGAGAAAAGATTGGGTAAAACTTTCCAACAACTCTATCCGGAAACT    +225
 A  K  F  R  A  S  S  I  L  A  E  K  R  L  G  K  T  F  Q  Q  L  Y  P  E  T     75

TTGAAAGAAGGTTTGATCGTGCAAAAGGGAAAATTACAAAACTGGTTAAGAACCAGTGTAGATGTGATTGTTGAA    +300
 L  K  E  G  L  I  V  Q  K  G  K  L  Q  N  W  L  R  T  S  V  D  V  I  V  E    100

CTGGCTAAGCAATTTGGTTTAGACCACAACTGTGTGCCCAAATTATGCGATAGAATATTCGATATTGTCGAACCG    +375
 L  A  K  Q  F  G  L  D  H  N  C  V  P  K  L  C  D  R  I  F  D  I  V  E  P    125

TCGAAGAAGTGGCAGAATGTCATTAACAATGGCGATTTGAAATCGTATAATTTAATGTTCGACGTTTCGGTACCA    +450
 S  K  K  W  Q  N  V  I  N  N  G  D  L  K  S  Y  N  L  M  F  D  V  S  V  P    150

GAACCTAAATGCGTCATAGTGGACTTTCAATGGATTCGCTACGTTCCTGCAATGGTTGATGTCAGCGTGCTCTTC    +525
 E  P  K  C  V  I  V  D  F  Q  W  I  R  Y  V  P  A  M  V  D  V  S  V  L  F    175

TATTTAAATCCAGAAAAGAATTAA                                                       +549
 Y  L  N  P  E  K  N  .                                                        182

GAGATAAATCTACGAAAGAGTTGCTGGAAACATTATCATGAAATTTTTTGTAGACCCTTGAGCGATAGTGATCTG    +624
AATATAAGGATACCAAGCTTCTCGGATATCCTTCGAGAGTATAAAGAAATGCGTCTGTTTGGTATGATTTTTTCG    +699
TGTTTATCTTCTCCTATAAATCTTATAGATGGGAATAAGTGTGCCGAGTTGACTAAAGATTCTGAAAGTTTTGCA    +774
AAGTTTCTGTTTGAGGACCATTTGGAACCTATTTCAGAAGTAATGCGGACCGATCCCATTTACTGTGATATAATT    +849
ACAGAATTGATCACGGAGCTGGTGGAAAGTTCTGAAAAAGTGCTTGCTACGATAAGTGACAATGTGTAAATCATA    +924
TTAAAAATATATTGTTTATC                                                           +944
```

Figure 4.

Nucleotide sequence and open reading frame amino acid translation of *E. pennicornis* venom gland cDNA library insert 1. An NCBI-BLAST sequence similarity search indicates the protein is a juvenile hormone-inducible protein (*N. vitripennis*) (Accession number XP_001604874). The signal peptide is highlighted in grey and the hormone inducible domain is outlined in red.

```
TACGATTCTTTTAAAAGACGGAGACCGAAAAGAAATACAGTTCAAAGTTACTCAGACAGAGCGTTGCAAAGGAAG      -75

ATGGATTTGTTCATTTTAACTCGTTTCATTTTGTTCCTATCTTTCTTTATGAAATCGATTCATTGTCAGTATAGC     +75
 M  D  L  F  I  L  T  R  F  I  L  F  L  S  F  F  M  K  S  I  H  C  Q  Y  S      25

GAATCGCAGGAATCCGGCCATAACCGGAACGCACCAGATAAAGAATTGACTACAGAAGAGTTCCAATTGATTTTT    +150
 E  S  Q  E  S  G  H  N  R  N  A  P  D  K  E  L  T  T  E  E  F  Q  L  I  F      50

CATCAATCTCAAACAGTTGATATCGAATACGATTTTATCAATATCACAACCGAGATGATTGAAACAGAACGTAAA    +225
 H  Q  S  Q  T  V  D  I  E  Y  D  F  I  N  I  T  T  E  M  I  E  T  E  R  K      75

GTTAGTTTTACAATCGATGGTAAAGAGTACCACTTATCATTGACTCCAGCTGCTTCACAATCAGTGTTGCCATAT    +300
 V  S  F  T  I  D  G  K  E  Y  H  L  S  L  T  P  A  A  S  Q  S  V  L  P  Y     100

GGTACTAAAATTAAGAGTGCAATTTGGTGGACTGATAATGATACTCATATCCACGAAGAGGACTATAGTGACGAG    +375
 G  T  K  I  K  S  A  I  W  W  T  D  N  D  T  H  I  H  E  E  D  Y  S  D  E     125

AGATGGGATAGTAGAGCTATCTATGAAAATTTAGAAATTATGGCAACTATACTGGTCCGTACCGAGAATGGTACT    +450
 R  W  D  S  R  A  I  Y  E  N  L  E  I  M  A  T  I  L  V  R  T  E  N  G  T     150

TCTTATTATGATGGAGTCTTCGGTGAAGGTATAGCAATGAAGGTTGTGAGATCTCTTCCAGGAAGATTGATGAAT    +525
 S  Y  Y  D  G  V  F  G  E  G  I  A  M  K  V  M  R  S  L  P  G  R  L  M  N     175

ATTTACGGAGCAAACTATCACTTTGTTTATGATAGTAATGGCTCCGTGTACGACGTTGTTTTAAACGGACAAGAT    +600
 I  Y  G  A  N  Y  H  F  V  Y  D  S  N  G  S  V  Y  D  V  V  L  N  G  Q  D     200

GAACCAGCTGTTCCTGCTGATATGGCAGTAAATAATTTTTATCCGAAACTCCTTGTACTTGTCGACTACTCATTA    +675
 E  P  A  V  P  A  D  M  A  V  N  N  F  Y  P  K  L  L  V  L  V  D  Y  S  L     225

TTCAAGATCTTTAATGAAAACTTCGAGGAGACTGTCAAGTATCTCACAATATTTGGAATGCTGTAAACTTAAGG    +750
 F  K  I  F  N  E  N  F  E  E  T  V  K  Y  L  T  I  F  W  N  A  V  N  L  R     250

TTCAGACCGGTCCAGCATCCGAAAGTAAATATCATAATCACTGGGATCGTCATTGCTAAGAACGAAGCCGCATTT    +825
 F  R  P  V  Q  H  P  K  V  N  I  I  I  T  G  I  V  I  A  K  N  E  A  A  F     275

CAGCACGTTTATAGAGCTAGGTATAGCAAAAATTCGAAACTAGTGCATACCGGAAGAGTTATAGATAATGGTCGA    +900
 Q  H  V  Y  R  A  R  Y  S  K  N  S  K  L  V  H  T  G  R  V  I  D  N  G  R     300

TATTTCTTTGGGACTAATTTTGACCCTTATTACGACAATTATGATGCCAGTTTTACAATGGCAAGTATGGATGAT    +975
 Y  F  F  G  T  N  F  D  P  Y  Y  D  N  Y  D  A  S  F  T  M  A  S  M  D  D     325

CCAACTGGAAAAGGAGGTGCAACTGTCATTGGCGGGATTTGCTCCTCATCTAACAATATCGCTTATATCAGAGAC   +1050
 P  T  G  K  G  G  A  T  V  I  G  G  I  C  S  S  S  N  N  I  A  Y  I  R  D     350

GTTGGAAGTTATTCGGGAGTAAAGGTTGCTACACACGAATTGGGCCACTTGCTAAATGGCCAACACGATTCAGAT   +1125
 V  G  S  Y  S  G  V  K  V  A  T  H  E  L  G  H  L  L  N  G  Q  H  D  S  D     375

ACCACTTGTAGCGAAAAGATAAATGATAATATCTATACAATCATGGCTAAGCAAGGATCAACAAAAGCAAGCAAA   +1200
 T  T  C  S  E  K  I  N  D  N  I  Y  T  I  M  A  K  Q  G  S  T  K  A  S  K     400

TTCGTTTGGTCATCATGCACTCTGACCGCATTTGCAAACTTTTCCAAAACAACAAGTGCAGCGTGCCTAAAAGAT   +1275
 F  V  W  S  S  C  T  L  T  A  F  A  N  F  S  K  T  T  S  A  A  C  L  K  D     425

ACCTATCGAAAGCATTAG                                                            +1293
 T  Y  R  K  H  .                                                              430

GATTGAAGATCTTCAAAAGAAATGCATAAATTATATAATCACTTGGATGCAGAGAGAGAGAGAGAGAGTAGGAGT   +1368
GAGAGAGAGAGTAGGAGTCAAAAAAAAAAAAAAAAAAAAAAAAAA                                 +1413
```

Figure 5.
Nucleotide sequence and open reading frame amino acid translation of *E. pennicornis* venom gland cDNA library insert 2. An NCBI-BLAST sequence similarity search indicates the protein is a reprolysin-like metalloprotease (*N. vitripennis*) (Accession number XP_001602775). The signal peptide is outlined in grey and the catalytic metalloprotease domain is highlighted in red.

```
                                      GCTGTTCAGTTTCGGCATTCCGTGTGAACGCGTCAG  -261
AGAGTTCAGTCAAGCCATTTTTTATTTTCTACGCAAACGCGCCATCTTGTTCCCCTCGGCAACGGGGCGAGGACA  -225
CGCGCATTCCTCTTGCTCTCACTAAACTTGGAATTTGAGGATATTCTTTCGAGAATATTCCTATTTTTTTGTGTG  -150
GAAACTGCATAGAGAGTCGACGAGAGTCGACTATCTATTAACAGTCTTAAGATTGACGACTCAAGGAAAAAAAAC  -75

ATGAGCTACGGCAGGCCTCCGCCTCGGATCGACGGGATGATATCCCTGAAGGTCGATAATTTGACCTATCGCACG  +75
 M  S  Y  G  R  P  P  P  R  I  D  G  M  I  S  L  K  V  D  N  L  T  Y  R  T   25

ACTCCTGAAGATCTCAGGAGAGTTTTCGAACGATGCGGCGAAGTTGGTGATATCTACATACCCAGAGATAGATTT  +150
 T  P  E  D  L  R  R  V  F  E  R  C  G  E  V  G  D  I  Y  I  P  R  D  R  F   50

ACCAGGGAAAGCCGTGGATTCGCTTTCGTCAGATTTTATGACAAGCGTGATGCTGAAGATGCGCTTGATGCTATG  +225
 T  R  E  S  R  G  F  A  F  V  R  F  Y  D  K  R  D  A  E  D  A  L  D  A  M   75

GATGGACGTTTGTTAGATGGAAGGGAATTACGTGTGCAGATGGCTCGTTATGGTAGACCAACATCACCTCATCGT  +300
 D  G  R  L  L  D  G  R  E  L  R  V  Q  M  A  R  Y  G  R  P  T  S  P  H  R  100

AGCCGTGGAAGCAGGCGCTACAGAAGTCGCAGCCGAGATCGCCGGCGTCCAAGGACCAGGTCACGTTCACGATCC  +375
 S  R  G  S  R  R  Y  R  S  R  S  R  D  R  R  R  P  R  T  R  S  R  S  R  S  125

CGTTCCCGTAGCCGTGACAGGGACCGCAAGCGTAGTTACAGCCGTAGCCGTTCACGCTCACGTTCCGACAGCAAG  +450
 R  S  R  S  R  D  R  D  R  K  R  S  Y  S  R  S  R  S  R  S  R  S  D  S  K  150

AGTTCTCGTGGAAGATCACGTACTCGCAGCAAATCGACCGAAAGACAAAAGGATGCACGTTCTAAATCAAGGGAC  +525
 S  S  R  G  R  S  R  T  R  S  K  S  T  E  R  Q  K  D  A  R  S  K  S  R  D  175

TGA                                                                          +528
 .                                                                            175

AGCTGAAGCCAAATTACAATTCTCGTAAGAGCATTTCGTGCTTGACATAATGGTTTAACAAGAGCCGCCGAGTAT  +603
ACTGGCATGAGTTTCAAAGAAAGGTATTCTATTCGTCCGCATGTCTGTCCAATCTTTTTTTATATTAAAATAGTG  +678
AGTCTTATCAAGCGACTCAAGTACCTCATCCTTGAGCTGAGCTTTCATCAAAATCTTGGGGATCAACATTCGAGT  +753
CCTGTGCAATTATAGTGACCTGGACGTCATGCCTAATTAGGATTTTTCCCTTCACGGAAA                 +828
```

Figure 6.
Nucleotide sequence and open reading frame amino acid translation of *E. pennicornis* venom gland cDNA library insert 3. An NCBI-BLAST sequence similarity search indicates the protein is similar to an RNA splicing factor (*A. aegypti*) (Accession number ABF18358). The signal peptide is highlighted in grey and the RNA binding domain is outlined in red.

```
                                         GATTATAAATACTGTATCGAATCAGTCGACAAACTAGCTAAGAA      -44

ATGATTGTTCTTCCAGTTTTAGTCTTCTCCCTTTCATTTATTACTTCCTGCAAGGCTCACCCGAAAAGAATACCT      +75
 M  I  V  L  P  V  L  V  F  S  L  S  F  I  T  S  C  K  A  H  P  K  R  I  P        25

GGAGTGTCCATCGTAGGCGGAGACGACGCTCAATCAGGAACATTTACTTATGCAGTATCCATACAAAAATCCAGT     +150
 G  V  S  I  V  G  G  D  D  A  Q  S  G  T  F  T  Y  A  V  S  I  Q  K  S  S        50

GAACACATTTGTGGAGGTGGAATAATTGGTGATTGGTGGATTCTCACGGCAGCTCATTGTTTTATGACTGATGGA     +225
 E  H  I  C  G  G  G  I  I  G  D  W  W  I  L  T  A  A  H  C  F  M  T  D  G        75

GGTTCCTTCATACCGATCATGTTATCAGCTGTTATTGGTACTGTAGATATAATCAACCTAGAACCAACTGCAATG     +300
 G  S  F  I  P  I  M  L  S  A  V  I  G  T  V  D  I  I  N  L  E  P  T  A  M       100

CGAATTTATATTGAAACGTACTATGTACCCATGGAATTCAGACCTCAAGATCAAAAAGGCAATCAGACACACCGT     +375
 R  I  Y  I  E  T  Y  Y  V  P  M  E  F  R  P  Q  D  Q  K  G  N  Q  T  H  R       125

ACCCAAGGAGATATTGCCATCGTTAAACTTCGAGAACCAATAGGACTTGAAACTCAGCCTCATCTGACAGTATTA     +450
 T  Q  G  D  I  A  I  V  K  L  R  E  P  I  G  L  E  T  Q  P  H  L  T  V  L       150

CAAATTCCTACAGAGTTGACGGAATTTCCGAGAGACGAAGCCATTATGGTTGGATTTGGATATAATGAAATCAGA     +525
 Q  I  P  T  E  L  T  E  F  P  R  D  E  A  I  M  V  G  F  G  Y  N  E  I  R       175

GTTATACTGTATCATAGAACGCAACAATACCAACGAGTAGGTAAATACTACGGTAAATTGCGATATGCGAAAACA     +600
 V  I  L  Y  H  R  T  Q  Q  Y  Q  R  V  G  K  Y  Y  G  K  L  R  Y  A  K  T       200

ATCATCATTTCAAGTGGTATATGCCGCCTAGTTAGGAAAAGGATAATTACCGATAATCATATGTGCGCGCGAGTG     +675
 I  I  I  S  S  G  I  C  R  L  V  R  K  R  I  I  T  D  N  H  M  C  A  R  V       225

AAACAACGTGTAGAAGGTGTAGCAGAAGGAGTCTGTTTCGGTGACAGTGGTGGACCTTTGGTACGTCATGGTAAT     +750
 K  Q  R  V  E  G  V  A  E  G  V  C  F  G  D  S  G  G  P  L  V  R  H  G  N       250

GTGGCGATCGCCGTATTAAGTACCTGCCCCATCACTTGTTCAGAAGAGAAAGGAGTCGGAATCTACACAAGGGTG     +825
 V  A  I  A  V  L  S  T  C  P  I  T  C  S  E  E  K  G  V  G  I  Y  T  R  V       275

AAACCATACGAATCCTTCATCAGAGATGCAATTGCAAATAAGCATCGTCCTGATATGCGTATTGCCACGTGGTAC     +900
 K  P  Y  E  S  F  I  R  D  A  I  A  N  K  H  R  P  D  M  R  I  A  T  W  Y       300

CTTTTCAACCAAGAGTGTCCTGGTTTAGAACCACGTCTCATAGAGAAAGTCTATAATGCAACAAATCATGATGGT     +975
 L  F  N  Q  E  C  P  G  L  E  P  R  L  I  E  K  V  Y  N  A  T  N  H  D  G       325

CAAGGTTCATCATCACACGTGGTCGAAAATACGTCGGAAGAGCCGATAGACGATTCGTCAATCGATGATTAG       +1047
 Q  G  S  S  S  H  V  V  E  N  T  S  E  E  P  I  D  D  S  S  I  D  D  .          348

AGATTCTGTAGTAAATTTTCTAAGCTACTCAATACTATGGTTGATAAAGTTATGAATGTTTTATTTACATTAAAA    +1122
TGATGCGCTGTTGACAAATAAAGATTCGTATCG                                              +1155
```

Figure 7.
Nucleotide sequence and open reading frame amino acid translation of *E. pennicornis* venom gland cDNA library insert 4. An NCBI-BLAST sequence similarity search indicates the protein is a trans-membrane serine protease (*G. gallus*) (Accession number XP_425880). The signal peptide is outlined in grey and the serine protease hydrolase signal is highlighted in red. Closer structural analysis revealed the protein had no trans-membrane domain.

```
CTTCCATCTGATTTATCGAAAATTGATATGGTAGATATTACTGAAACGAGATTGAAAAATATTCTGCTCGATGAA  +75
 L  P  S  D  L  S  K  I  D  M  V  D  I  T  E  T  R  L  K  N  I  L  L  D  E    25

ATATGGAAAAAGGAAGAATCTCATGATCCACAATTTATCAAGAATCTGAAGATAACTTACCGAAACTGCATGGAT  +150
 I  W  K  K  E  E  S  H  D  P  Q  F  I  K  N  L  K  I  T  Y  R  N  C  M  D    50

GAGTCTGGAATTGAAAAAGAAAACGCAAAAACAGCCAGTTCTATATTGACCGAAATTGGTGGTTGGCCATTGTTA  +225
 E  S  G  I  E  K  E  N  A  K  T  A  S  S  I  L  T  E  I  G  G  W  P  L  L    75

CAAGGTGATGAATGGGAAGAAGATGACTATGAGTGGACAGAAACCATAAGCAAAATTGGTAAGATTGGACTGAAA  +300
 Q  G  D  E  W  E  E  D  D  Y  E  W  T  E  T  I  S  K  I  G  K  I  G  L  K   100

GCTACGTTTCCATTCGAAATGTTTGTGTATTATGATCCTACAGACTCAAGGAAATATATTCTATCGATCTTTTGT  +375
 A  T  F  P  F  E  M  F  V  Y  Y  D  P  T  D  S  R  K  Y  I  L  S  I  F  C   125

GCTAGCCTCTTAATTGATGAATCATTCATATCAAATGAAGAATATATGAAAGCATATGGACTTTACATTACTGAA  +450
 A  S  L  L  I  D  E  S  F  I  S  N  E  E  Y  M  K  A  Y  G  L  Y  I  T  E   150

GTGGCTGTTTTACTCGGTGCAGACAGGGACCAAATACAGAAGAAAATTAGTGAACTAGTAGAATTTGAAAAGAAA  +525
 V  A  V  L  L  G  A  D  R  D  Q  I  Q  K  K  I  S  E  L  V  E  F  E  K  K   175

CTCCACAAGGCAAAAACCAAATGGGAGGATAGGGATCTTATATCACTTTCTGAAAATACAGGACCGGTCGAAGAC  +600
 L  H  K  A  K  T  K  W  E  D  R  D  L  I  S  L  S  E  N  T  G  P  V  E  D   200

TTGATAAAAAAATATCCGTCAATAGATTGGATAAATATTTTGAATAGACAAGTGGCACCAAGTGGCATTTCC[illegible]  +675
 L  I  K  K  Y  P  S  I  D  W  I  N  I  L  N  R  Q  V  A  P  S  G  I  S  V   225

[illegible]GGCAAGACTAGAGTAACTAATTATGATCCCAAATACATGGATGAACTGGGAGAAATCATATCTAGCGCATCG  +750
 [illegible]  G  K  T  R  V  T  N  Y  D  P  K  Y  M  D  E  L  G  E  I  I  S  S  A  S   250

AAAAGAGTGCTGGCAAATTTTCAGGTTTGGAACATCCTTGTGACTCTGATAGATTTCATGCCGAAGAAATTTCTA  +825
 K  R  V  L  A  N  F  Q  V  W  N  I  L  V  T  L  I  D  F  M  P  K  K  F  L   275

GAAGTACAATTGGACTTCGATAACGTCGTCAAAGGAGTGGAATCATTACAAAACAGAAAATTGGCATGTTATGGA  +900
 E  V  Q  L  D  F  D  N  V  V  K  G  V  E  S  L  Q  N  R  K  L  A  C  Y  G   300

CAATTGAAAAGTCATATGTTCATAGCCTTGAGTGGTTTGTTTCTAAGAAAATCATTCGATAAACAAGTGAAGATC  +975
 Q  L  K  S  H  M  F  I  A  L  S  G  L  F  L  R  K  S  F  D  K  Q  V  K  I   325

AGTGTTACAGAATTGGTAACTGAGCTCAAACAGCAGTTCAAACTTGAATTTGCAAAGACTGATTGGATGGATGAT  +1050
 S  V  T  E  L  V  T  E  L  K  Q  Q  F  K  L  E  F  A  K  T  D  W  M  D  D   350

GACACTCGCAGAGCAGCGATTGAAAAACTTTATGCTATGAAAACTAATATAGCTTATCCGGAAGAGCTTATGGAT  +1125
 D  T  R  R  A  A  I  E  K  L  Y  A  M  K  T  N  I  A  Y  P  E  E  L  M  D   375

GACAAAAAAATAGAAGAATATTATGGAGATCTTGTGGTTGAAGAAGGAGATTATCTCAAAGGTTTTATTGATATT  +1200
 D  K  K  I  E  E  Y  Y  G  D  L  V  V  E  E  G  D  Y  L  K  G  F  I  D  I   400

TGGAAAGTCGAGAGAGATCGTTCTTTAGATATATTGGTTAAACCACTTCAAGAAGTTGATTGGATAAACAGAGCA  +1275
 W  K  V  E  R  D  R  S  L  D  I  L  V  K  P  L  Q  E  V  D  W  I  N  R  A   425

GGTGTAGCTGAGGTCAGTGCTTACTATTCGCCAGGAAATAATGAGATAAATGTGATAGCAGGGATTTTGCAAGAT  +1350
 G  V  A  E  V  S  A  Y  Y  S  P  G  N  N  E  I  N  V  I  A  G  I  L  Q  D   450

CCGTTTTTCAATGTCCATCTCCCTAGATATTTGAATTATGCTGGAGTCGGTTACGTTATAGGTCATGAAATGAGT  +1425
 P  F  F  N  V  H  L  P  R  Y  L  N  Y  A  G  V  G  Y  V  I  G  H  E  M  S   475

CACGGATTTGATACAACGGGACGGCAACGTGATGAAAGTGGAAATTTTCATAACTGGTGGAATAACAGTACGGAG  +1500
 H  G  F  D  T  T  G  R  Q  R  D  E  S  G  N  F  H  N  W  W  N  N  S  T  E   500

GAAAAATTCAATGAGAGAGCTCAATGTTTCATTGATTCGTACAGTAATTACGAGATTGAACAAGTGAAAATGAAG  +1575
 E  K  F  N  E  R  A  Q  C  F  I  D  S  Y  S  N  Y  E  I  E  Q  V  K  M  K   525

GCTCAGGGCCGAGTATCGATAGGAGAAAATATCGCTGACAGTGGCGGTATAAAAATGGCATACAACGCATACAAA  +1650
 A  Q  G  R  V  S  I  G  E  N  I  A  D  S  G  G  I  K  M  A  Y  N  A  Y  K   550

CGTTTAATGGAAAATGATCCACGTGAGCCACGCCTACCAGGACTCAAGTATACGCCCGAGCAATTATTCGGGATG  +1725
 R  L  M  E  N  D  P  R  E  P  R  L  P  G  L  K  Y  T  P  E  Q  L  F  G  M   575

AACGCGGCCTTTAGCTGGTGCGGTAAATTTAATGACGAATTTCTGAAGAATCAAATTGATACAGACGAACACGGT  +1800
 N  A  A  F  S  W  C  G  K  F  N  D  E  F  L  K  N  Q  I  D  T  D  E  H  G   600
```

```
CCTCAGGAAGCTCGGATAAATCTCTCGTTTGCAAATAATGAGGACTTTGCTAAAGACTTCATTTGCCAACCAGGT   +1875
 P  Q  E  A  R  I  N  L  S  F  A  N  N  E  D  F  A  K  D  F  I  C  Q  P  G    625

GCAAAAATGAATCCAGAGAAAAAATGCGTTGTGTGGTGA                                       +1914
 A  K  M  N  P  E  K  K  C  V  V  W  .                                          637

GTCACTGTTTTTAACGATGAAGTACAGCTCATTATCCGAAAATTTCGCGGGTCCATAACCAGATTAGTACATGAT   +1989
GCCGAATATTTTTTTTCTAATTTTACGGAATAGGTATTATAATTTTGTAACACACAAATTTTATACATCAAAATA   +2064
CTCCTCTTGTCTCTTCAATAAGGTTCTTAGGTATATGAGATGGTGTCTTTAACTAGCGTAAGGTTAGTGTTAAGG   +2139
TATTTTCTCCCTTTGAGATATGTAACGACTCATGAAAATAAACGGAATATGTTTTGAAATTCCAAAAAAAAAAAA   +2214
AAAAAAAAAA                                                                    +2224
```

Figure 8.
Nucleotide sequence and open reading frame amino acid translation of *E. pennicornis* venom gland cDNA library insert 5. The obtained sequence sections are outlined in red with the missing central section not highlighted. The 5′ sequence has a *Sal*I restriction site coloured in blue. An NCBI-BLAST sequence similarity search indicates the protein is a neprilysin-like protein (*V. canescens*) (Accession number AAL91975). The sequence of the central section was obtained after restriction at the *Sal*I site highlighted in blue.

The reprolysin-like metalloprotease gene (Section 1.1B and Figure 5) was selected for further investigation through expression of the protein in *P. pastoris*.

3.1A. Expression of Reprolysin-like Metalloprotease (Awaiting Publication)

The *E. pennicornis* venom gland contains multiple reprolysin-like metalloproteases and at the time of isolation by Dr Daniel Price, two similar coding sequences were also co-amplified. The alignment of their amino acid sequences can be seen in figure 9. The Rep A and Rep B sequences were most similar with 89% sequence identity, while Rep B and Rep C only shared 46% sequence identity. All three of the sequences contained the identical active site zinc-binding motif HExxHxxxGxHD (Section 1.1B).

Recombinant reprolysin-like metalloprotease (Rep B, Figure 9 and Figure 5) was expressed in *P. pastoris* and purified by Phenyl Sepharose chromatography (Figure 10). Three peak fractions of recombinant reprolysin-like metalloprotease were incubated at either 4°C or 37°C for several minutes and its gelatinase activity was demonstrated. It was observed that the activity band on gelatin gel increased in intensity (Figure 11). In a later experiment the gelatinase activity was shown to be inhibited by treatment with the divalent cation chelator EDTA (Figure 12). It was also observed that a similar sized band to the recombinant protein was present in a protein extract taken from the venom sac of female insects, demonstrating that this protein is present naturally within the venom gland and appears to be readily activated (Figure 12).

3.1B. Biological Activity of Reprolysin-like Metalloprotease (Awaiting Publication)

Purified recombinant reprolysin-like metalloprotease was injected into fifth stadium *L. oleracea* larvae with water injected as a control. Fifteen percent mortality was observed in those fifth stadium insects injected with the recombinant reprolysin-like metalloprotease and this increased to 35% throughout the ecdysis to sixth stadium. Those insects that survived ecdysis to sixth stadium showed a three day pro-longed weight gain phase compared to the control larvae (Figure 13).

```
RepA    IHCQYSESQESGHNRNAPDKELTTEEFQLIFHQSQTVDIEYDFINITTEMIETERKVSFT     60
RepB    IHCQYSESQESGHNRNAPDKELTTEEFQLIFHQSQTVDIEYDFINITTEMIETERKVSFT     60
RepC    ...RYSRVAEHIRTRKRPDKELTEEEFKLVFHRSSTEEIDYDFVNLTTEITEIERKVLFT     57

RepA    IDGKEYHLSLTPAASQSVLPYGTKIKSAIWWTDNDTHIHEEDYSDERWDSRAIYENLEIM     120
RepB    IDGKEYHLSLTPAASQSVLPYGTKIKSAIWWTDNDTHIHEEDYSDERWDSRAIYENLEIM     120
RepC    IDDKDYHLKLT.RASSSVIPSGTLIRSAILWTDNQTHFHDEDSTDEHWGSSHIYEDLDKM     116

RepA    ATILVRTENGTSYYDGVFGEGIAMKVVRSLPGRLMNIYGANYHFVYDSNGSVYDVVLNGQ     180
RepB    ATILVRTENGTSYYDGVFVK.YSNEGVRSLPGRLMNIYGANYHFVYDSNGSVYDVVLNGQ     179
RepC    ATFLLRDDDDFTRYDGVFGGGKDMKVVGSLPARLVNIYGANYHFIYYANGSVSDVILNGA     176

RepA    DEPAVPADMAVNN..FYPKLLVLVDYSLFK.IFNENFEETVKYLTIFWNAVNLRFRPVQH     237
RepB    DEPAVPADMASKII.FYSETPCTCRLLIIQDLLMKTSRRLSSISTIFWNAVNLRFRPVQH     238
RepC    KQVVGSANTQAGLNNFYPKLLVLVDYTLFK.ILNKSYEETIRYLAIFWNAVDMKFKKFET     235

RepA    PKVNIIITGIVIAKNEAAFQHVYRARYSKNSKLVHTGRVIDNGRYFFGTNF..DPYYDNY     295
RepB    PKVNIIITGIVIAKNEAAFQHVYRARYSKNSKLVHTGRVIDNGRYFFGTNF..DPYYDNY     296
RepC    PKINIIITGIIVPKNEGALKHVYDARIKSDMQKVNATKIITNSEHFFGANFSTESYFDNY     295

RepA    DASFTMASMDDPTGKGGATVIGGICSSSNNIAYIRDVGSYSGVKVATHELGHLLNGQHDS     355
RepB    DASFTMASMDDPTGKGGATVIGGICSSSNNIAYIRDVGSYSGVKVATHELGHLLNGQHDS     356
RepC    DATFTMASLNDLEGQTGLAYIGAICKNNHNNAYVKDSGVFSGVLAAAHELGHLLASDHDE     355

RepA    DTTCSEKINDN...IYTIMAKQGSTKASKFVWSSCTLTAFANFSKTTSAACLKDTYRKHL     412
RepB    DTTCSEKINDN...IYTIMAKQGSTKASKFVWSSCTLTAFANFSKTTSAACLKDTYRKQ.     412
RepC    DVGCPGEINYNTRLTGTIMAEYRNNNVSKFIWS...........................     388
```

Figure 9.

Amino acid sequence alignment of the three co-amplified reprolysin-like metalloproteases from *E. pennicornis* venom gland. Identical and similar residues are highlighted in grey. Rep A and Rep B show 89% sequence identity and Rep B and C only share 46% sequence identity. Figure courtesy of Dr Daniel Price, University of Durham.

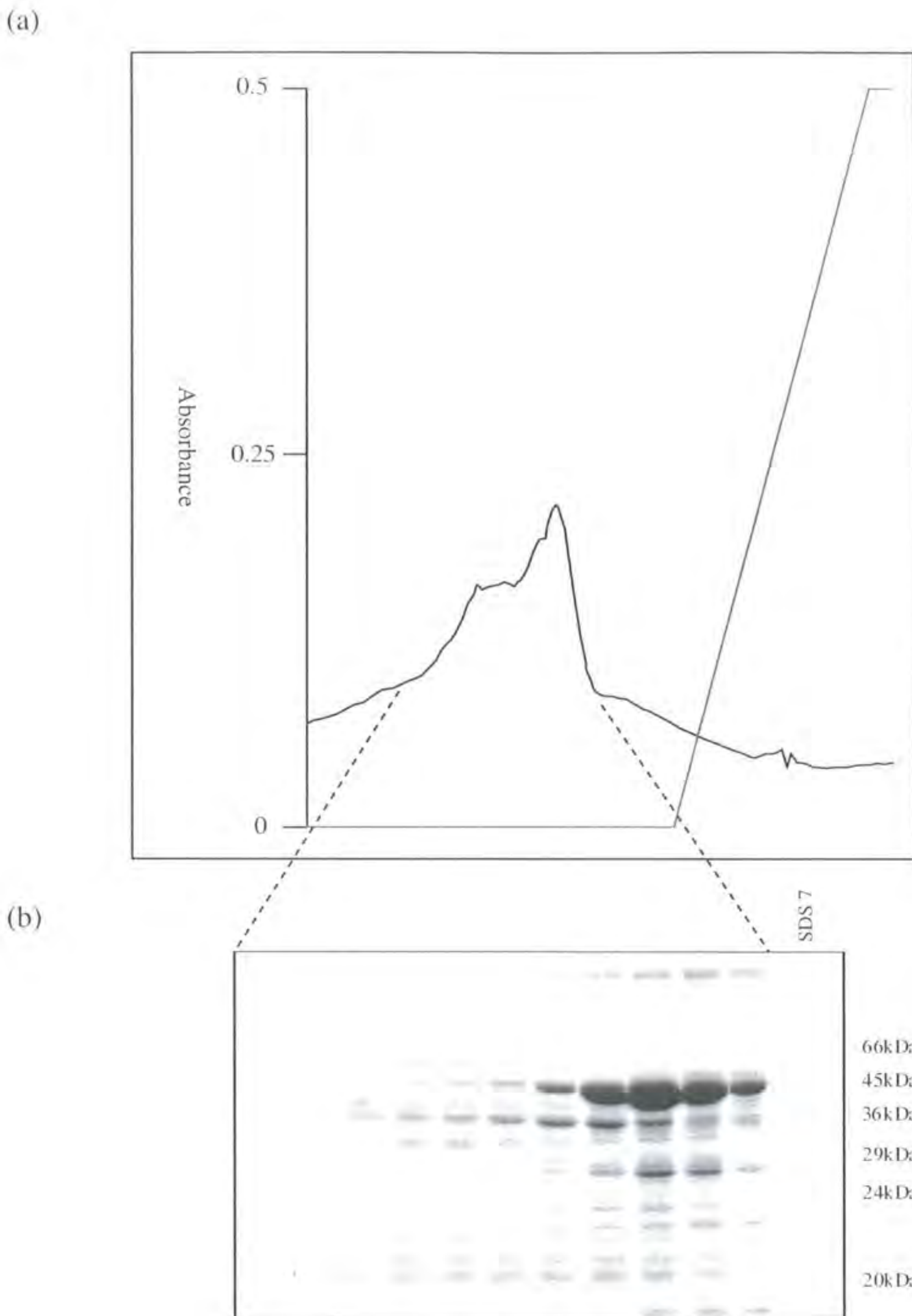


Figure 10.
(a) Phenyl Sepharose column purification chart for recombinant reprolysin-like metalloprotease. Binding to the column was carried out in 2M sodium chloride. Recombinant reprolysin-like metalloprotease was eluted by decreasing the salt concentration. The peak occurs in the water fractions after the salt gradient. (b) 12.5% SDS-PAGE analysis of the peak fractions collected from the recombinant reprolysin-like metalloprotease purification. The large band is the correct size for reprolysin-like metalloprotease.
Figure data courtesy of Dr Daniel Price, University of Durham.

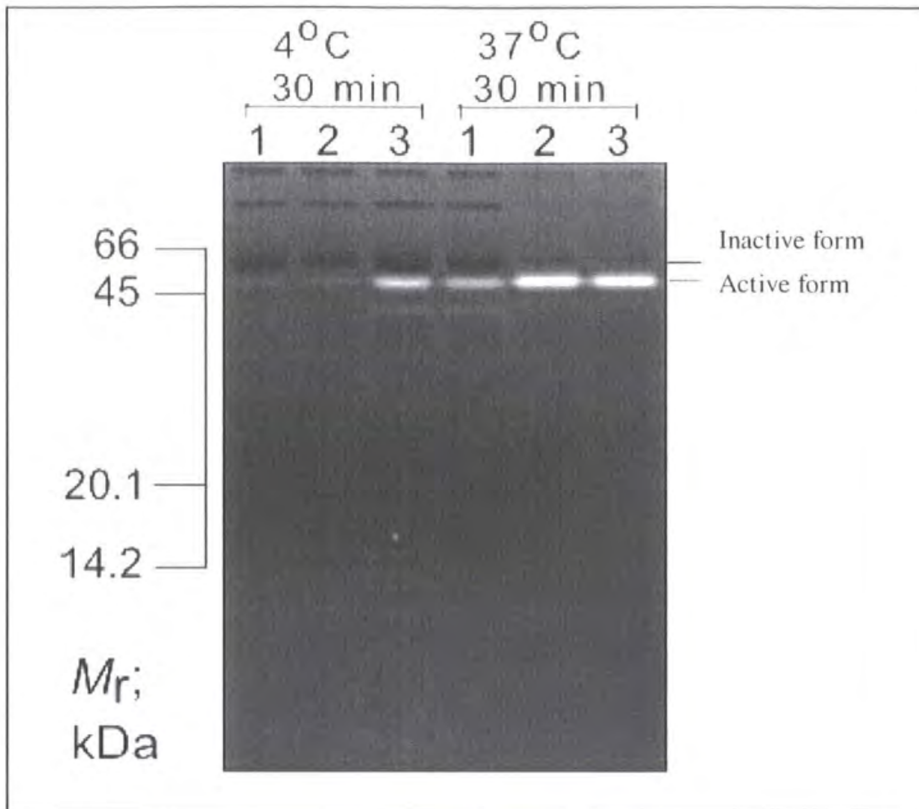


Figure 11.
Gelatinase activity gel of three peak fractions (1,2 and 3) from the recombinant reprolysin-like metalloprotease purification (shown in Figure 10) incubated at either 4°C or 37°C. The intensity of the active form band increases with the higher temperature incubation.
Figure courtesy of Dr Daniel Price, University of Durham.

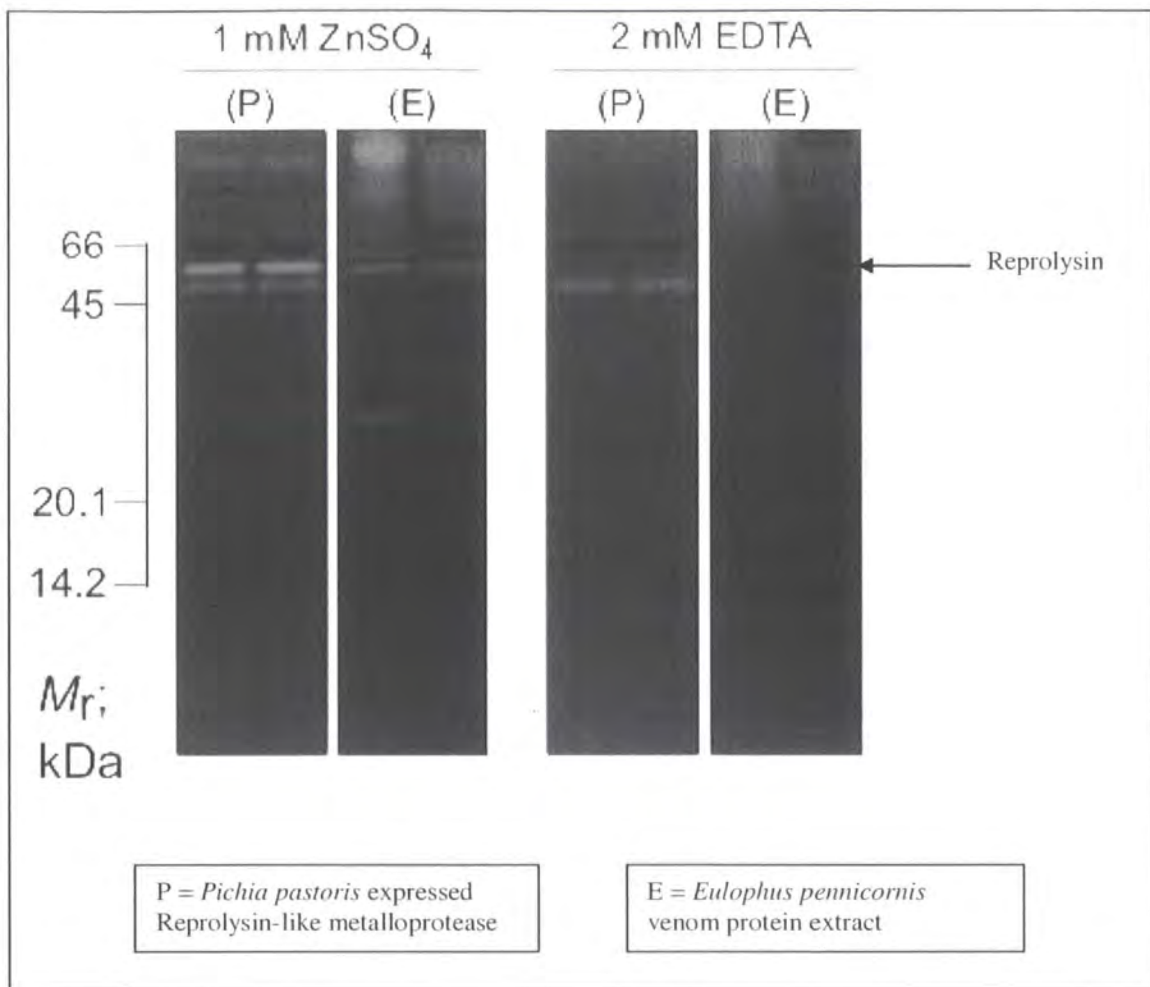


Figure 12.
Gelatinase activity gel of purified recombinant *E. pennicornis* reprolysin-like metalloprotease (P) and *E. pennicornis* venom gland protein extract (E) in the presence of 1mM $ZnSO_4$ and 2mM EDTA. The arrow indicates the size of reprolysin-like metalloprotease. The gel shows reprolysin-like metalloprotease is actively present in the *E. pennicornis* venom gland (E) and the activity of both recombinant and venom gland reprolysin-like metalloprotease is inhibited by EDTA.
The lower band also shows gelatinase activity but is not a metalloprotease as its activity is not inhibited by EDTA. Samples containing this protein were injected into fifth stadium *L. oleracea* and no effects were observed.
Figure courtesy of Dr Daniel Price, University of Durham.

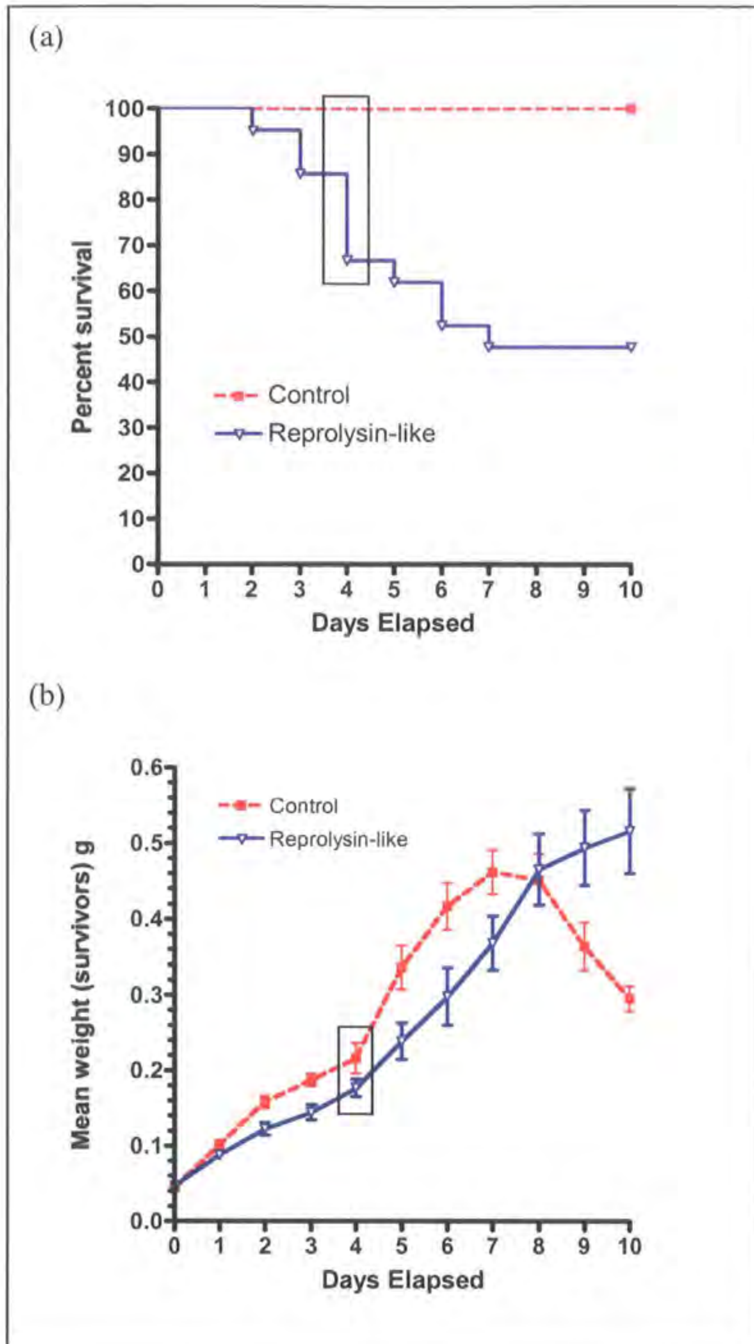


Figure 13.
(a) Survival of fifth stadium *L. oleracea* larvae injected with purified recombinant *E. pennicornis* reprolysin-like metalloprotease in PBS at the start of the fifth stadium. Control larvae were injected with PBS only. (b) Mean weight of surviving *L. oleracea* larvae in the injection bioassay shown in (a).
The point of ecdysis from fifth to sixth stadium is outlined in black.
Figure courtesy of Dr Daniel Price, University of Durham.

3.2. Reprolysin-like Metalloprotease Fusion Protein Variants

The initial expression and subsequent injection experiments carried out by Dr Daniel Price (Section 3.1A and 3.1B) demonstrated that recombinant reprolysin-like metalloprotease was active when present in the haemolymph. It was also concluded that the reprolysin-like metalloprotease could self activate when there was a small amount of active enzyme present, however, the domain structure of this protease differs from other reprolysins in that it lacks a C-terminal disintegrin domain but instead has an N-terminal domain which could contain a putative binding site and other targeting information.

To investigate if GNA could transport reprolysin-like metalloprotease to the haemolymph of *L. oleracea* larvae when delivered orally and to elucidate the reprolysin-like metalloprotease structure-activity relationship, DNA constructs encoding fusion protein variations of the original isolated reprolysin-like metalloprotease sequence (Figure 1) were designed (Figures 14(b), 15(b) and 16(b)).

3.2A. Production of Reprolysin 1, Reprolysin 2 and Reprolysin 3 Expression Constructs

The three devised constructs (designated Rep 1, Rep 2 and Rep3) were produced from the original reprolysin-like metalloprotease (Figure 1) using PCR with specifically designed primers (Table 1). Constructs were cloned in *E. coli* via pCR2.1-TOPO and the DNA subsequently transformed into SMD protease deficient *P. pastoris* yeast using the pGAPZαB constitutive vector for expression. DNA sequencing was used to confirm maintenance of correct construct integrity and positioning at each stage (Figures 14(a), 15(a) and 16(a)).

(a)

```
                                              ATGAGATTTCCTTCAATTACTGCTGTTTTATTCGCAGCATCC   -267
TCCGCATTAGCTGCTCCAGTCAACACTACAACAGAAGATGAAACGGCACAAATTCCGGCTGAAGCTGTCATCGGT   -225
TACTCAGATTTAGAAGGGGATTTCGATGTTGCTGTTTTGCCATTTTCCAACAGCACAAATAACGGGTTATTGTTT   -150
ATAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTATCTCTCGAGAAAAGAGAGGCTGAAGCTGCA   -75

CAGTATAGCGAATCGCAGGAATCCGGCCATAACCGGAACGCACCAGATAAAGAATTGACTACAGAAGAGTTCCAA   +75
 Q  Y  S  E  S  Q  E  S  G  H  N  R  N  A  P  D  K  E  L  T  T  E  E  F  Q    25

TTGATTTTTCATCAATCTCAAACAGTTGATATCGAATACGATTTTATCAATATCACAACCGAGATGATTGAAACA   +150
 L  I  F  H  Q  S  Q  T  V  D  I  E  Y  D  F  I  N  I  T  T  E  M  I  E  T    50

GAACGTAAAGTTAGTTTTACAATCGATGGTAAAGAGTACCACTTATCATTGACTCCAGCTGCTTCACAATCAGTG   +225
 E  R  K  V  S  F  T  I  D  G  K  E  Y  H  L  S  L  T  P  A  A  S  Q  S  V    75

TTGCCATATGGTACTAAAATTAAGAGTGCAATTTGGTGGACTGATAATGATACTCATATCCACGAAGAGGACTAT   +300
 L  P  Y  G  T  K  I  K  S  A  I  W  W  T  D  N  D  T  H  I  H  E  E  D  Y    100

AGTGACGAGAGATGGGATAGTAGAGCTATCTATGAAAATTTAGAAATTATGGCAACTATACTGGTCCGTACCGAG   +375
 S  D  E  R  W  D  S  R  A  I  Y  E  N  L  E  I  M  A  T  I  L  V  R  T  E    125

AATGGTACTTCTTATTATGATGGAGTCTTCGGTGAAGGTATAGCAATGAAGGTTGTGAGATCTCTTCCAGGAAGA   +450
 N  G  T  S  Y  Y  D  G  V  F  G  E  G  I  A  M  K  V  V  R  S  L  P  G  R    150

TTGATGAATATTTACGGAGCAAACTATCACTTTGTTTATGATAGTAATGGCTCCGTGTACGACGTTGTTTTAAAC   +525
 L  M  N  I  Y  G  A  N  Y  H  F  V  Y  D  S  N  G  S  V  Y  D  V  V  L  N    175

GGACAAGATGAACCAGCTGTTCCTGCTGATATGGCAGTAAATAATTTTTATCCGAAACTCCTTGTACTTGTCGAC   +600
 G  Q  D  E  P  A  V  P  A  D  M  A  V  N  N  F  Y  P  K  L  L  V  L  V  D    200

TACTCATTATTCAAGATCTTTAATGAAAACTTCGAGGAGACTGTCAAGTATCTCACAATATTTTGGAATGCTGTA   +675
 Y  S  L  F  K  I  F  N  E  N  F  E  E  T  V  K  Y  L  T  I  F  W  N  A  V    225

AACTTAAGGTTCAGACCGGTCCAGCATCCGAAAGTAAATATCATAATCACTGGGATCGTCATTGCTAAGAACGAA   +750
 N  L  R  F  R  P  V  Q  H  P  K  V  N  I  I  I  T  G  I  V  I  A  K  N  E    250

GCCGCATTTCAGCACGTTTATAGAGCTAGGTATAGCAAAAATTCGAAACTAGTGCATACCGGAAGAGTTATAGAT   +825
 A  A  F  Q  H  V  Y  R  A  R  Y  S  K  N  S  K  L  V  H  T  G  R  V  I  D    275

AATGGTCGATATTTCTTTGGGACTAATTTTGACCCTTATTACGACAATTATGATGCCAGTTTTACAATGGCAAGT   +900
 N  G  R  Y  F  F  G  T  N  F  D  P  Y  Y  D  N  Y  D  A  S  F  T  M  A  S    300

ATGGATGATCCAACTGGAAAAGGAGGTGCAACTGTCATTGGCGGGATTTGCTCCTCATCTAACAATATCGCTTAT   +975
 M  D  D  P  T  G  K  G  G  A  T  V  I  G  G  I  C  S  S  S  N  N  I  A  Y    325

ATCAGAGACGTTGGAAGTTATTCGGGAGTAAAGGTTGCTACACACGAATTGGGCCACTTGCTAAATGGCCAACAC   +1050
 I  R  D  V  G  S  Y  S  G  V  K  V  A  T  H  E  L  G  H  L  L  N  G  Q  H    350

GATTCAGATACCACTTGTAGCGAAAAGATAAATGATAATATCTATACAATCATGGCTAAGCAAGGATCAACAAAA   +1125
 D  S  D  T  T  C  S  E  K  I  N  D  N  I  Y  T  I  M  A  K  Q  G  S  T  K    375

GCAAGCAAATTCGTTTGGTCATCATGCACTCTGACCGCATTTGCAAACTTTTCCAAAACAACAAGTGCAGCGTGC   +1200
 A  S  K  F  V  W  S  S  C  T  L  T  A  F  A  N  F  S  K  T  T  S  A  A  C    400

CTAAAAGATACCTATCGAAAGCAGGCGGCCGCCGACAATATTTTGTACTCCGGTGAGACTCTCTCTACAGGGGAA   +1275
 L  K  D  T  Y  R  K  Q  A  A  A  D  N  I  L  Y  S  G  E  T  L  S  T  G  E    425

TTTCTCAACTACGGAAGTTTCGTTTTTATCATGCAAGAGGACTGCAATCTGGTCTTGTACGACGTGGACAAGCCA   +1350
 F  L  N  Y  G  S  F  V  F  I  M  Q  E  D  C  N  L  V  L  Y  D  V  D  K  P    450

ATCTGGGCAACAAACACAGGTGGTCTCTCCCGTAGCTGCTTCCTCAGCATGCAGACTGATGGGAACCTCGTGGTG   +1425
 I  W  A  T  N  T  G  G  L  S  R  S  C  F  L  S  M  Q  T  D  G  N  L  V  V    475

TACAACCCATCGAACAAACCGATTTGGGCAAGCAACACTGGAGGCCAAAATGGGAATTACGTGTGCATCCTACAG   +1500
 Y  N  P  S  N  K  P  I  W  A  S  N  T  G  G  Q  N  G  N  Y  V  C  I  L  Q    500

AAGGATAGGAATGTTGTGATCTACGGAACTGATCGTTGGGCTACTGGATGA                           +1551
 K  D  R  N  V  V  I  Y  G  T  D  R  W  A  T  G  .                             516
```

(b)

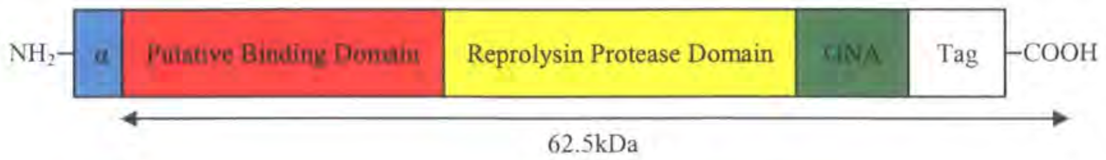


Figure 14.
Reprolysin 1 (Rep 1) – The complete reprolysin-like metalloprotease-GNA fusion protein construct (to determine if GNA can transport reprolysin-like metalloprotease to the haemolymph of *L. oleracea* larvae). (a) Nucleotide and amino acid sequence confirmation. (b) Diagrammatic representation of the construct.
The yeast alpha factor is coloured blue. The reprolysin putative binding and catalytic domains are highlighted in red and yellow respectively. GNA is outlined in green. The Tag consists of both His- and Myc-tags.

(a)

```
                                             ATGAGATTTCCTTCAATTACTGCTGTTTTATTCGCAGCATCC   -267
TCCGCATTAGCTGCTCCAGTCAACACTACAACAGAAGATGAAACGGCACAAATTCCGGCTGAAGCTGTCATCGGT   -225
TACTCAGATTTAGAAGGGGATTTCGATGTTGCTGTTTTGCCATTTTCCAACAGCACAAATAACGGGTTATTGTTT   -150
ATAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTATCTCTCGAGAAAAGAGAGGCTGAAGCTGCA   -75

TTTTATCCGAAACTCCTTGTACTTGTCGACTACTCATTATTCAAGATCTTTAATGAAAACTTCGAGGAGACTGTC   +75
 F  Y  P  K  L  L  V  L  V  D  Y  S  L  F  K  I  F  N  E  N  F  E  E  T  V   25

AAGTATCTCACAATATTTTGGAATGCTGTAAACTTAAGGTTCAGACCGGTCCAGCATCCGAAAGTAAATATCATA   +150
 K  Y  L  T  I  F  W  N  A  V  N  L  R  F  R  P  V  Q  H  P  K  V  N  I  I   50

ATCACTGGGATCGTCATTGCTAAGAACGAAGCCGCATTTCAGCACGTTTATAGAGCTAGGTATAGCAAAAATTCG   +225
 I  T  G  I  V  I  A  K  N  E  A  A  F  Q  H  V  Y  R  A  R  Y  S  K  N  S   75

AAACTAGTGCATACCGGAAGAGTTATAGATAATGGTCGATATTTCTTTGGGACTAATTTTGACCCTTATTACGAC   +300
 K  L  V  H  T  G  R  V  I  D  N  G  R  Y  F  F  G  T  N  F  D  P  Y  Y  D   100

AATTATGATGCCAGTTTTACAATGGCAAGTATGGATGATCCAACTGGAAAAGGAGGTGCAACTGTCATTGGCGGG   +375
 N  Y  D  A  S  F  T  M  A  S  M  D  D  P  T  G  K  G  G  A  T  V  I  G  G   125

ATTTGCTCCTCATCTAACAATATCGCTTATATCAGAGACGTTGGAAGTTATTCGGGAGTAAAGGTTGCTACACAC   +450
 I  C  S  S  S  N  N  I  A  Y  I  R  D  V  G  S  Y  S  G  V  K  V  A  T  H   150

GAATTGGGCCACTTGCTAAATGGCCAACACGATTCAGATACCACTTGTAGCGAAAAGATAAATGATAATATCTAT   +525
 E  L  G  H  L  L  N  G  Q  H  D  S  D  T  T  C  S  E  K  I  N  D  N  I  Y   175

ACAATCATGGCTAAGCAAGGATCAACAAAAGCAAGCAAATTCGTTTGGTCATCATGCACTCTGACCGCATTTGCA   +600
 T  I  M  A  K  Q  G  S  T  K  A  S  K  F  V  W  S  S  C  T  L  T  A  F  A   200

AACTTTTCCAAAACAACAAGTGCAGCGTGCCTAAAAGATACCTATCGAAAGCAGGCGGCCGCCGACAATATTTTG   +675
 N  F  S  K  T  T  S  A  A  C  L  K  D  T  Y  R  K  Q  A  A  A  D  N  I  L   225

TACTCCGGTGAGACTCTCTCTACAGGGGAATTTCTCAACTACGGAAGTTTCGTTTTTATCATGCAAGAGGACTGC   +750
 Y  S  G  E  T  L  S  T  G  E  F  L  N  Y  G  S  F  V  F  I  M  Q  E  D  C   250

AATCTGGTCTTGTACGACGTGGACAAGCCAATCTGGGCAACAAACACAGGTGGTCTCTCCCGTAGCTGCTTCCTC   +825
 N  L  V  L  Y  D  V  D  K  P  I  W  A  T  N  T  G  G  L  S  R  S  C  F  L   275

AGCATGCAGACTGATGGGAACCTCGTGGTGTACAACCCATCGAACAAACCGATTTGGGCAAGCAACACTGGAGGC   +900
 S  M  Q  T  D  G  N  L  V  V  Y  N  P  S  N  K  P  I  W  A  S  N  T  G  G   300

CAAAATGGGAATTACGTGTGCATCCTACAGAAGGATAGGAATGTTGTGATCTACGGAACTGATCGTTGGGCTACT   +975
 Q  N  G  N  Y  V  C  I  L  Q  K  D  R  N  V  V  I  Y  G  T  D  R  W  A  T   325

GGATGA                                                                        +981
 G  .                                                                         326
```

(b)

NH$_2$– | α | Reprolysin Protease Domain | GNA | Tag | –COOH

54kDa

Figure 15.
Reprolysin 2 (Rep 2) – The protease domain only-GNA fusion protein construct (to evaluate if the binding domain is required for function). (a) Nucleotide and amino acid sequence confirmation. (b) Diagrammatic representation of the construct.
The yeast alpha factor is coloured blue, the reprolysin catalytic domain is highlighted in yellow and GNA is outlined in green. The Tag consists of both His- and Myc-tags.

(a)

```
                       ATGAGATTTCCTTCAATTACTGCTGTTTTATTCGCAGCATCC    -267
TCCGCATTAGCTGCTCCAGTCAACACTACAACAGAAGATGAAACGGCACAAATTCCGGCTGAAGCTGTCATCGGT    -225
TACTCAGATTTAGAAGGGGATTTCGATGTTGCTGTTTTGCCATTTTCCAACAGCACAAATAACGGGTTATTGTTT    -150
ATAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTATCTCTCGAGAAAAGAGAGGCTGAAGCTGCA    -75

TTTTATCCGAAACTCCTTGTACTTGTCGACTACTCATTATTCAAGATCTTTAATGAAAACTTCGAGGAGACTGTC    +75
 F  Y  P  K  L  L  V  L  V  D  Y  S  L  F  K  I  F  N  E  N  F  E  E  T  V    25

AAGTATCTCACAATATTTTGGAATGCTGTAAACTTAAGGTTCAGACCGGTCCAGCATCCGAAAGTAAATATCATA    +150
 K  Y  L  T  I  F  W  N  A  V  N  L  R  F  R  P  V  Q  H  P  K  V  N  I  I    50

ATCACTGGGATCGTCATTGCTAAGAACGAAGCCGCATTTCAGCACGTTTATAGAGCTAGGTATAGCAAAAATTCG    +225
 I  T  G  I  V  I  A  K  N  E  A  A  F  Q  H  V  Y  R  A  R  Y  S  K  N  S    75

AAACTAGTGCATACCGGAAGAGTTATAGATAATGGTCGATATTTCTTTGGGACTAATTTTGACCCTTATTACGAC    +300
 K  L  V  H  T  G  R  V  I  D  N  G  R  Y  F  F  G  T  N  F  D  P  Y  Y  D    100

AATTATGATGCCAGTTTTACAATGGCAAGTATGGATGATCCAACTGGAAAAGGAGGTGCAACTGTCATTGGCGGG    +375
 N  Y  D  A  S  F  T  M  A  S  M  D  D  P  T  G  K  G  G  A  T  V  I  G  G    125

ATTTGCTCCTCATCTAACAATATCGCTTATATCAGAGACGTTGGAAGTTATTCGGGAGTAAAGGTTGCTACACAC    +450
 I  C  S  S  S  N  N  I  A  Y  I  R  D  V  G  S  Y  S  G  V  K  V  A  T  H    150

GAATTGGGCCACTTGCTAAATGGCCAACACGATTCAGATACCACTTGTAGCGAAAAGATAAATGATAATATCTAT    +525
 E  L  G  H  L  L  N  G  Q  H  D  S  D  T  T  C  S  E  K  I  N  D  N  I  Y    175

ACAATCATGGCTAAGCAAGGATCAACAAAAGCAAGCAAATTCGTTTGGTCATCATGCACTCTGACCGCATTTGCA    +600
 T  I  M  A  K  Q  G  S  T  K  A  S  K  F  V  W  S  S  C  T  L  T  A  F  A    200

AACTTTTCCAAAACAACAAGTGCAGCGTGCCTAAAAGATACCTATCGAAAGCAGCTAGAACAAAAACTCATCTCA    +675
 N  F  S  K  T  T  S  A  A  C  L  K  D  T  Y  R  K  Q  L  E  Q  K  L  I  S    225

GAAGAGGATCTGAATAGCGCCGTCGACCATCATCATCATCATCATTGA                               +723
 E  E  D  L  N  S  A  V  D  H  H  H  H  H  H  .                                240
```

(b)

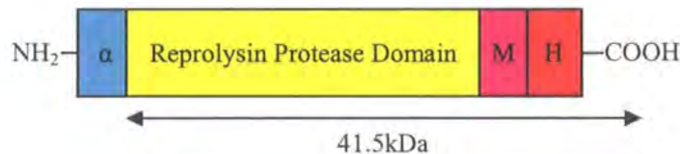


Figure 16.
Reprolysin 3 (Rep 3) – The protease domain only construct (to determine if it is still active). (a) Nucleotide and amino acid sequence confirmation. (b) Diagrammatic representation of the construct.
The yeast alpha factor is coloured blue. The reprolysin catalytic domain is highlighted in yellow. The Myc-tag (M, purple) and His-tag (H, red) are also outlined.

3.2B. Expression of Reprolysin-like metalloprotease, Reprolysin 1, Reprolysin 2 and Reprolysin 3.

After preparation of the reprolysin-like metalloprotease constructs (Section 3.2A), small-scale cultures were raised to isolate high level expressing clones. Culture supernatant (grown for four days) was analysed by western blot using anti-GNA antibodies (1:3000 dilution) for Rep 1 and Rep 2 (Figures 17 and 18). Detection of Rep 3 relied on the Myc-tag at the C-terminal end of the construct (Figure 16). Anti-Myc-tag antibodies (1:5000 dilution) were used, however, the Myc-tag is commonly degraded due to its terminal nature. *Nilaparvata lugens* (brown planthopper) sugar transporter (NLST) (supplied by Dr Daniel Price) was used for a Rep 3 positive control as the recombinant protein construct has been confirmed to have a Myc-tag attached (Figure 19). Rep 1 and Rep 2 should show reactivity at around 62.5kDa and 54kDa respectively. The expected size of Rep 3 is 41.5kDa.

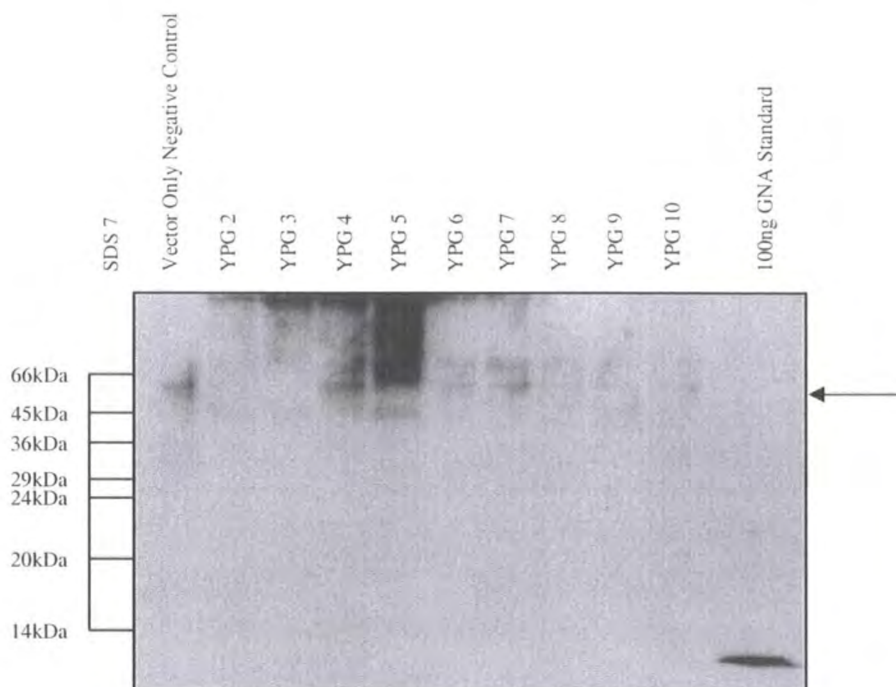


Figure 17.
Western blot analysis of reprolysin 1 pGAPZαB small scale cultures probed with anti-GNA antibodies (1:3000 dilution). A one minute exposure is shown. Blotted from a 15% SDS-PAGE gel of 20μl of culture supernatant. The expected size of Rep 1 is 62.5kDa.
The blot shows no distinct bands of high expression at the correct size.

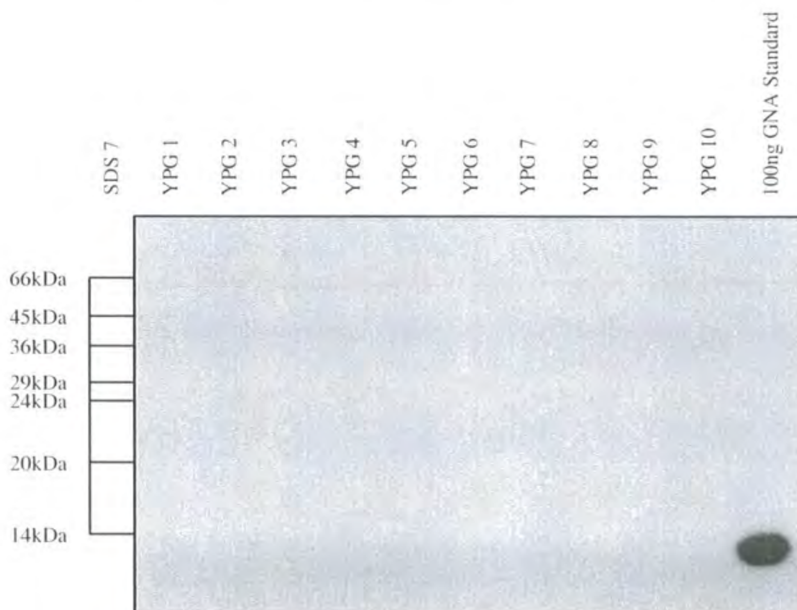


Figure 18.
Western blot analysis of reprolysin 2 pGAPZαB small scale cultures probed with anti-GNA antibodies (1:3000 dilution). A five minute exposure is shown. Blotted from a 15% SDS-PAGE gel of 20μl of culture supernatant. The expected size of Rep 2 is 54kDa.
The blot shows no distinct bands of high expression at the correct size.

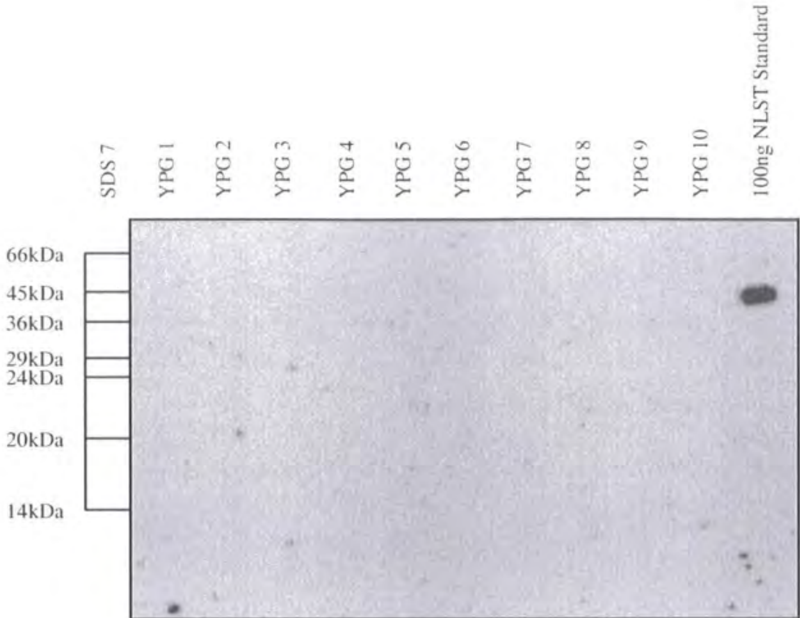


Figure 19.
Western blot analysis of reprolysin 3 pGAPZαB small scale cultures probed with anti-Myc-tag antibodies (1:5000 dilution). A five minute exposure is shown. Blotted from a 15% SDS-PAGE gel of 20μl of culture supernatant. The expected size of Rep 3 is 41.5kDa.
The blot shows no distinct bands of high expression at the correct size.

The presence of reactivity at the correct size in the positive standard lanes indicates the blots have worked successfully. However, there was an absence of distinct high level expression bands at the predicted sizes in all of the blots (Figures 17, 18 and 19). The feint band in Rep 1 YPG 4 (Figure 17, highlighted by an arrow) was a possibility but on scale up, no further reactivity was observed. Yeast transformations were repeated numerous times with similar results obtained. An experiment was therefore conducted where the original reprolysin-like metalloprotease was transformed alongside ASAII as a control (confirmed to express) to ensure the yeast transformation was working properly (Figures 20 and 21). The ASAII result clearly shows there is no problem with the transformation procedure, but the original reprolysin-like metalloprotease failed to express.

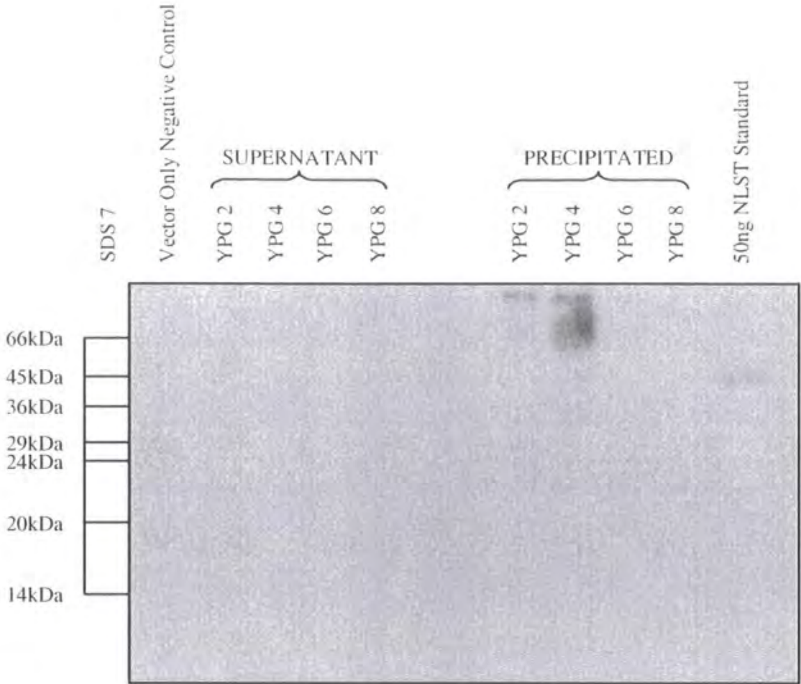


Figure 20.
Western blot analysis of small scale cultures of the original reprolysin-like metalloprotease in pGAPZαB probed with anti-Myc-tag antibodies (1:5000 dilution). A one minute exposure is shown. Blotted from a 15% SDS-PAGE gel of 20μl of culture supernatant and 20μl of 100μl precipitated supernatant. The expected size of the original reprolysin-like metalloprotease is 46kDa.
The blot shows no distinct bands of high expression at the correct size.

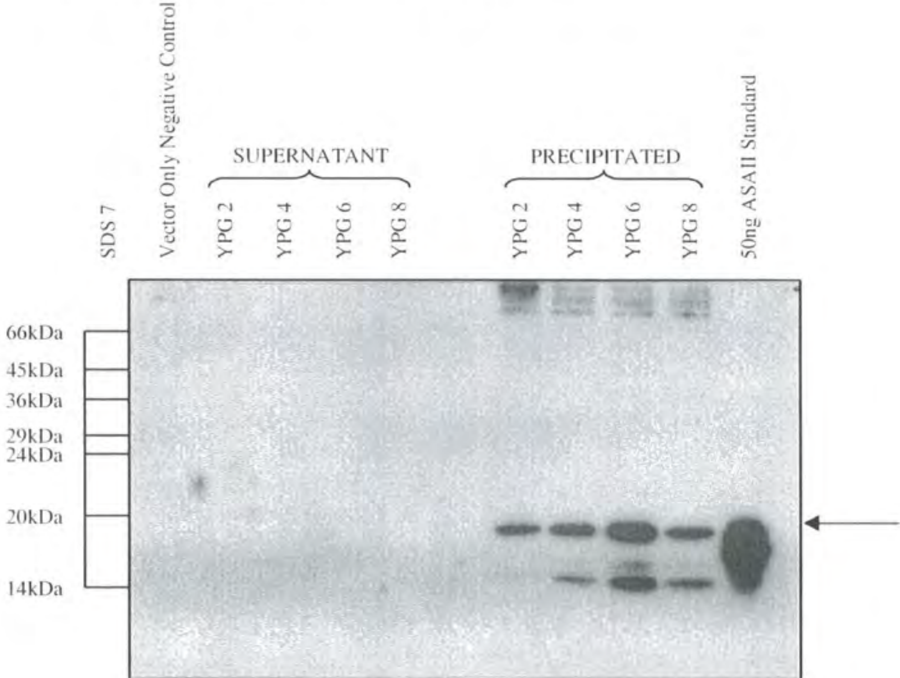


Figure 21.
Western blot analysis of ASAII pGAPZαB small scale cultures probed with anti-ASA antibodies (1:3000 dilution). A 10 minute exposure is shown. Blotted from a 15% SDS-PAGE gel of 20μl of culture supernatant and 20μl of 100μl precipitated supernatant. The expected size of ASAII is 19kDa. The blot shows ASAII present in all precipitated samples.

There were therefore two possible reasons to explain the situation: firstly, the yeast was not expressing the reprolysin-like metalloprotease constructs, or secondly, the yeast was expressing the recombinant proteins but they were toxic to the yeast. Non-expression was a possibility and has been observed previously in different fusions with little scope for remediating the situation (John Gatehouse, personal communication). However, if the problem was as a result of toxicity of the proteins to yeast cells, this could be addressed using inducible expression.

To this end, the original reprolysin-like metalloprotease and the three constructs were re-transformed into *P. pastoris* using the pPICZαB inducible expression vector. Small-scale expression and western blotting similar to that used for pGAPZαB was carried out (Figures 22, 23 and 24). The original reprolysin-like metalloprotease has a size of 46kDa and the three constructs are the same size as previously mentioned for pGAPZαB. Figure 22 shows a band of reactivity the correct size for the original reprolysin-like metalloprotease in three of the precipitated cultures. YPG 2 also shows expression at a low level in neat supernatant. This clone was therefore selected for larger scale (1L) shake flask expression. The result of which was again no expression (western blot not shown). Figures 23 and 24 show no distinct bands for the reprolysin-like metalloprotease at the correct sizes, but the presence of reactivity of the standard indicates the blots were successful.

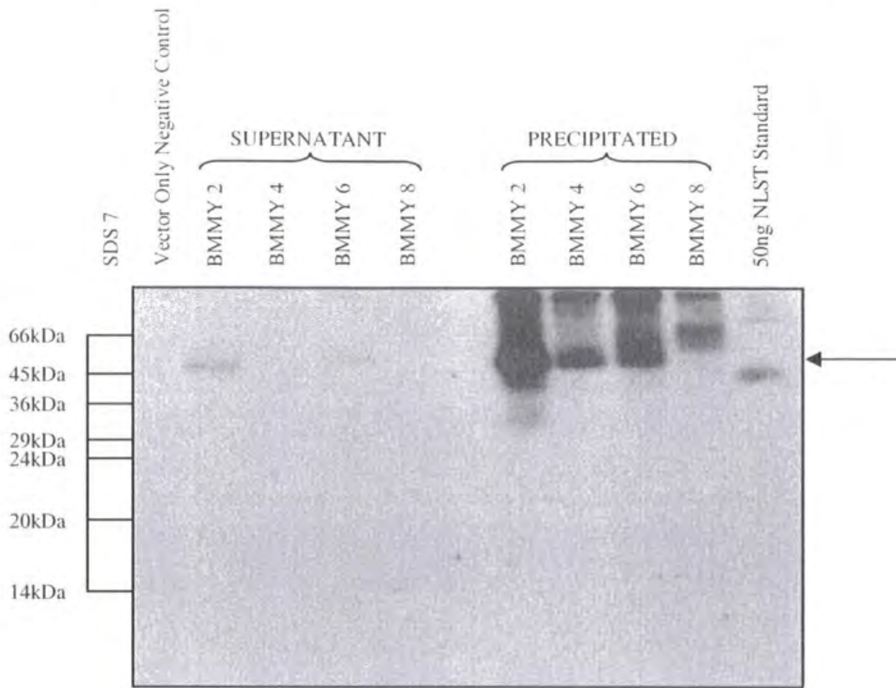


Figure 22.
Western blot analysis of small scale cultures of the original reprolysin-like metalloprotease in pPICZαB probed with anti-Myc-tag antibodies (1:5000 dilution). A one minute exposure is shown. Blotted from a 15% SDS-PAGE gel of 20μl of culture supernatant and 20μl of 100μl precipitated supernatant. The expected size of the original reprolysin-like metalloprotease is 46kDa.
The blot shows the original reprolysin-like metalloprotease present in precipitated samples 2, 4, and 6.

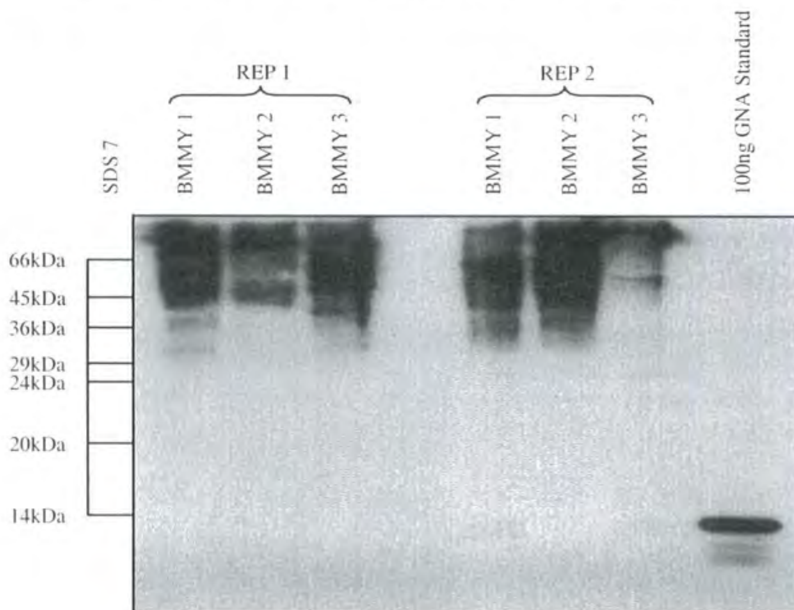


Figure 23.
Western blot analysis of reprolysin 1 and reprolysin 2 pPICZαB small scale cultures probed with anti-GNA antibodies (1:3000 dilution). A 20 second exposure is shown. Blotted from a 15% SDS-PAGE gel of 20μl of culture supernatant. The expected size of Rep 1 is 62.5kDa and Rep 2 is 54kDa.
The blot shows no distinct bands of high expression at the correct size.

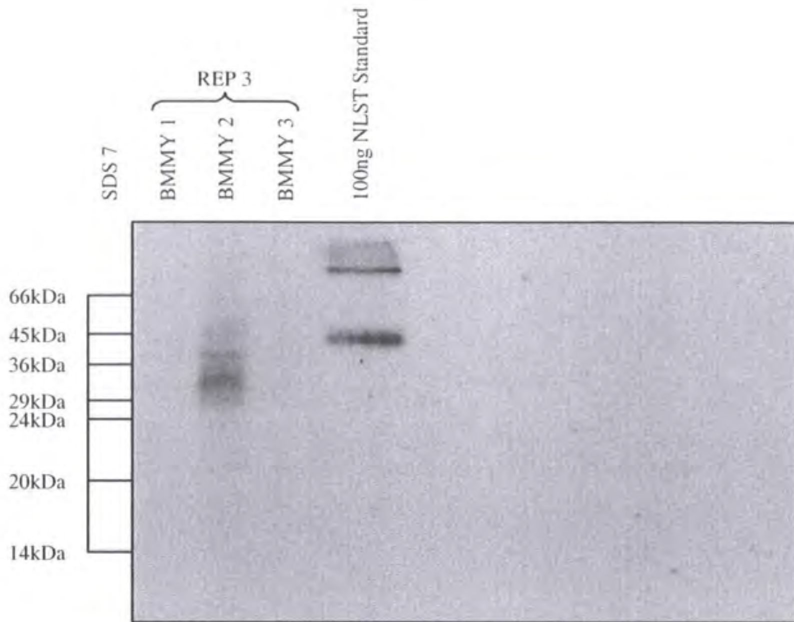


Figure 24.
Western blot analysis of reprolysin 3 pPICZαB small scale cultures probed with anti-Myc-tag antibodies (1:5000 dilution). A 10 minute exposure is shown. Blotted from a 15% SDS-PAGE gel of 20μl of culture supernatant. The expected size of Rep 3 is 41.5kDa.
The blot shows no distinct bands of high expression at the correct size.

It was thought that the lack of expression upon scale-up of culture size may be due to sub-optimal expression conditions (insufficient yeast biomass and growth prior to induction). A time dependent experiment was therefore conducted with the original reprolysin-like metalloprotease in pPICZαB. Fifty millilitre BMGY cultures were grown for one, two or three days and then cells transferred to 100ml BMMY media for a further one, two or three days. Yeast growth in BMGY media was estimated by absorbance at 600nm. Protein expression in BMMY media was monitored by supernatant absorbance at 280nm (Table 5). Samples were also taken and analysed by western blot for specific protein expression (Figure 25). *Allium sativum* agglutinin II was used as a comparison control (Figures 26). There were no major differences in protein levels across the cultures, all readings were between 0.5 and 0.6 (Table 5). To get the better value of 0.6 however, less than three days in BMGY is recommended. The highest absorbance value comes from one day in BMGY and three days in BMMY. Contrary to this, western blot analysis suggested the optimum is three days in BMGY and two days in BMMY (Figure 25). However, during the experiment it was observed that those cultures exceeding four days old turned lumpy. No exact reason could be found for this phenomenon, however, one rationale was that even though using the inducible expression vector, the reprolysin-like metalloprotease protein was having some negative activity against the yeast. This could also explain why large shake flask cultures gave no expression.

	BMGY (OD 600)	1 Day BMMY (OD 280)	2 Day BMMY (OD 280)	3 Day BMMY (OD 280)
1 Day	2.375	0.651	0.569	0.653
2 Day	2.434	0.602	0.632	0.573
3 Day	2.475	0.566	0.537	0.547

Table 5.
Absorbance readings (OD 600 and OD 280) of the original reprolysin-like metalloprotease in pPICZαB incubated in BMGY growth and BMMY induction media for varying times.

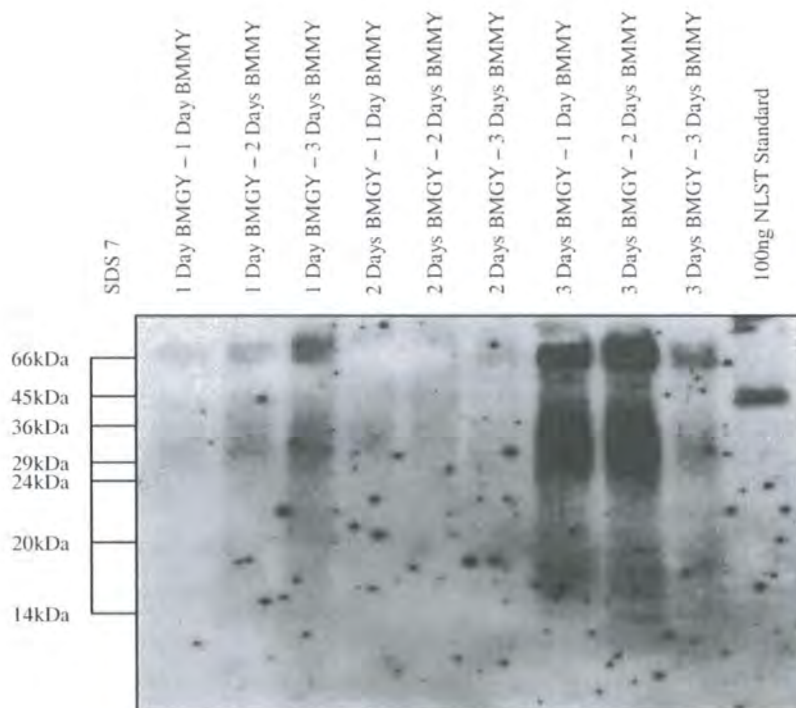


Figure 25.
Western blot analysis of the original reprolysin-like metalloprotease in pPICZαB incubated for varying times in BMGY and BMMY media, probed with anti-Myc-tag antibodies (1:5000 dilution). A five minute exposure is shown. Blotted from a 15% SDS-PAGE gel of 20μl of culture supernatant. The expected size of the original reprolysin-like metalloprotease is 46kDa. The blot shows no distinct bands of high expression at the correct size.

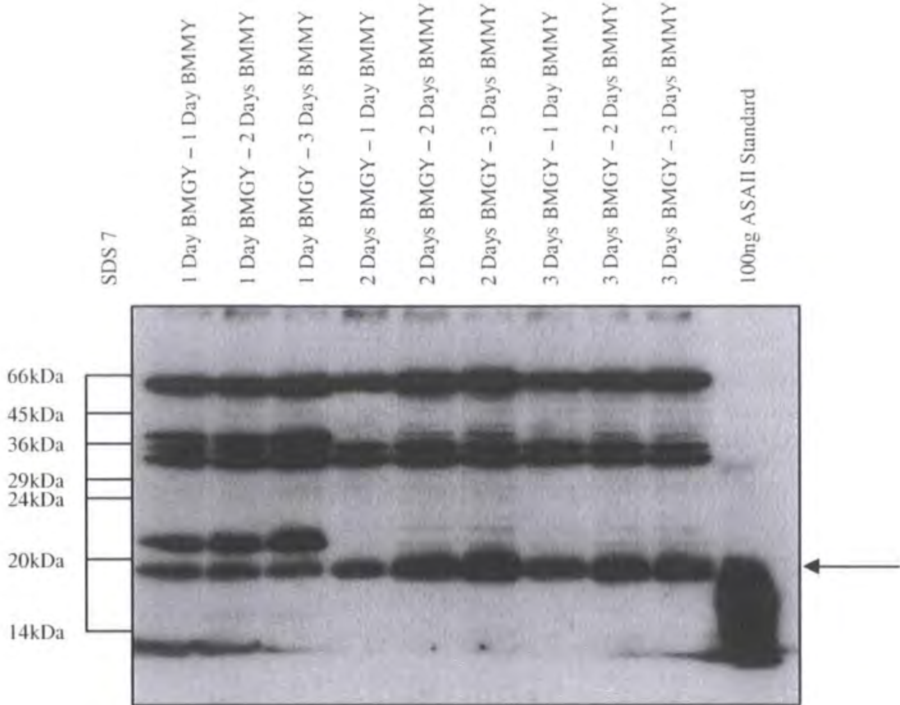


Figure 26.
Western blot analysis of ASAII pPICZαB incubated for varying times in BMGY and BMMY media probed with anti-ASA antibodies (1:3000 dilution). A 30 second exposure is shown. Blotted from a 15% SDS-PAGE gel of 20μl of culture supernatant. The expected size of ASAII is 19kDa. The blot shows ASAII present in all samples.

In this experiment, no band of the correct size was present in any of the original reprolysin-like metalloprotease pPICZαB cultures as was the case with the scale-up expression. The ASAII blot confirms the system is working correctly, with a band being observed at around 19kDa. The presence of the larger protein band in one day BMGY cultures can be attributed to ASAII with the His-tag present. The blot also shows that this His-tag is degraded if the cultures are incubated for a second day in BMGY media.

Dr Daniel Price also tried re-transforming *P. pastoris* with the original reprolysin-like metalloprotease in pGAPZαB and pPICZαB with a pGAPZαB vector only control to ensure there was no human error involved in the failure of expression. It was observed that there was a reduced amount of colonies on the reprolysin-like metalloprotease plates, again highlighting that reprolysin-like metalloprotease may be having negative impacts on the yeast.

Reprolysin-like metalloprotease construct experiments were halted due to the lack of expression.

3.3. Avidin as a Fusion Protein Carrier

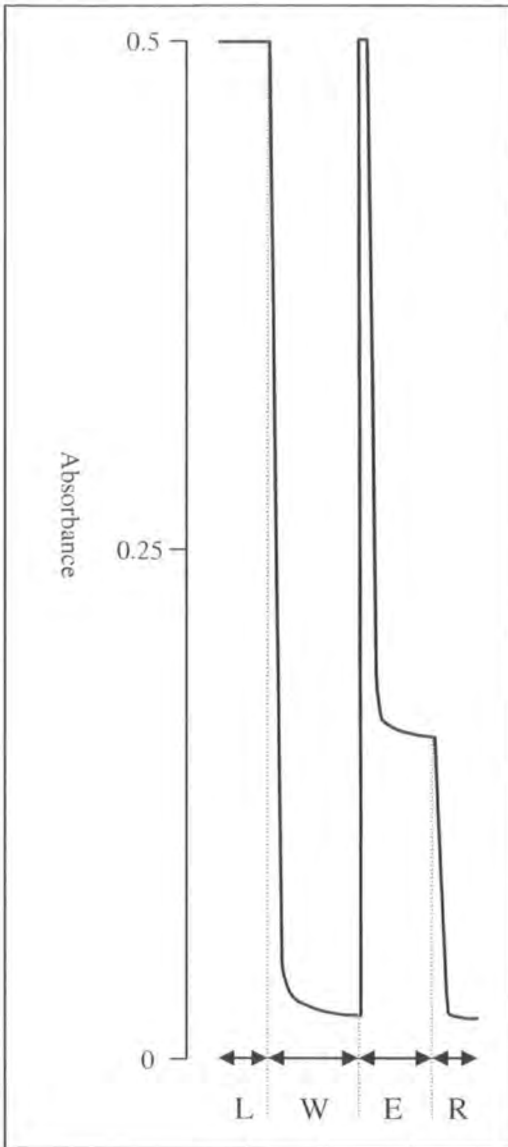
Commercial avidin purchased from Sigma was used in *L. oleracea* larvae bioassays (unpublished). After feeding on diet containing 2.5mg of avidin for various times, the presence of avidin was shown in the gut contents, bound to the gut membrane and in the haemolymph.Work on a garlic lectin:recombinant avidin fusion also gave similar results, prompting the investigation into using recombinant avidin as a fusion protein carrier.

3.3A. Expression of Recombinant Avidin

To assess the suitability of recombinant avidin as a carrier in Fusion Protein Technology, it was first expressed using pGAPZαB in X33 wild type *P. pastoris* by bench-top fermentation. The recombinant avidin sequence prepared by Dr Dave Bown contained a His-tag (Figure 3), the protein was therefore purified from culture supernatant by nickel affinity chromatography using a HisTrap nickel column. Binding to the column was carried out in 50mM sodium acetate, 0.5M sodium chloride pH 4.0. The addition of imidazole (300mM) and a pH increase (pH7.4) was used to elute the bound recombinant avidin. The effectiveness of the column and

purity of the single elution peak (Figure 27(a)) was confirmed by SDS-PAGE (Figure 27(b)). Avidin has a molecular mass of 16kDa. The size difference of the band observed with recombinant avidin can be attributed to the presence of the His-tag and glycosylation by the yeast. A western blot probed with anti-avidin antibodies (1:5000 dilution) showed reaction against this band confirming it was recombinant avidin. Eluted recombinant avidin was diluted one in five, dialysed into purified water and freeze-dried. Freeze-dried recombinant avidin was resuspended in purified water and quantified against commercial Sigma avidin by SDS-PAGE (Figure 28). Yields of 10-80mg/l culture medium were obtained.

(a)



L = Load

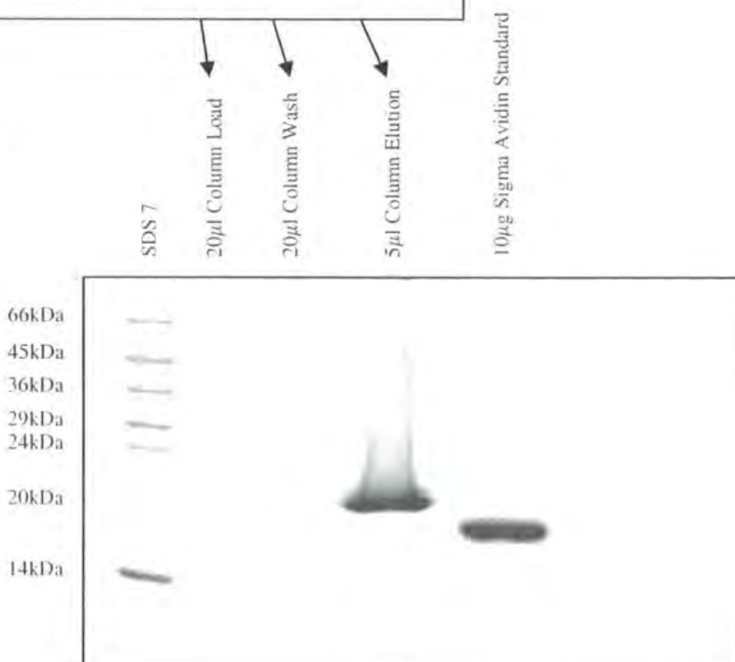
W = Wash

E = Elution

R = Re-equilibration

Figure 27.
(a) Nickel column purification chart for recombinant avidin. Binding to the column was carried out in 50mM sodium acetate, 0.5M sodium chloride pH 4.0. The addition of imidazole (300mM) and a pH increase (pH7.4) was used to elute the bound recombinant avidin. (b) 15% SDS-PAGE analysis of fractions collected from the recombinant avidin purification.

(b)



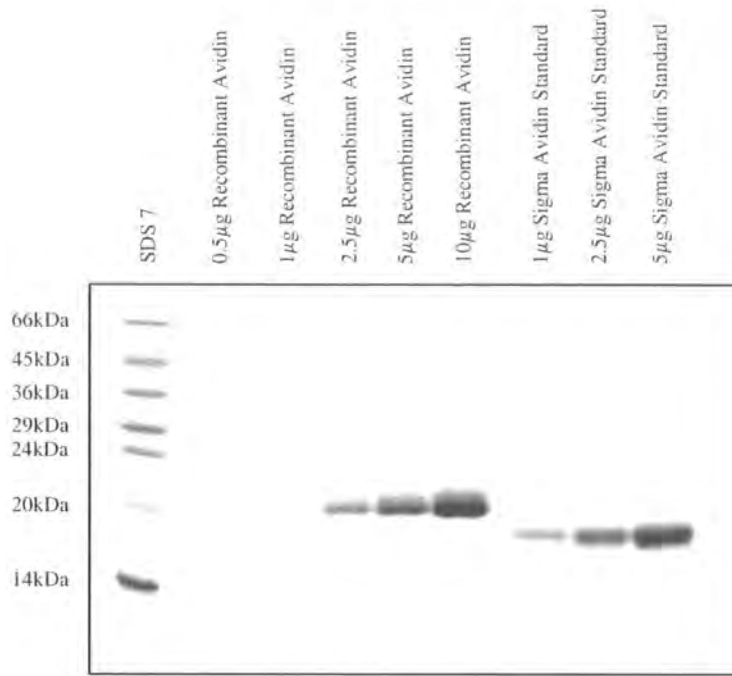


Figure 28.
15% SDS-PAGE concentration quantification analysis of freeze-dried purified recombinant avidin. 100% recombinant avidin concentration was obtained from most fermentation runs.

3.3B. Transport of Recombinant Avidin

Commercially available avidin (Sigma) had previously been shown to transport to the haemolymph of *L. oleracea* larvae when delivered orally (Section 3.3). To verify that this transport property was maintained in the recombinant protein, fifth stadium *M. brassicae* larvae were exposed to artificial diet containing either 0.1% [w/w] (1000 ppm) recombinant avidin or Sigma avidin. Equivalent amounts of artificial diet were used as a control. After twenty four hours, gut contents, gut and haemolymph were extracted and analysed by western blot using anti-avidin antibodies (1:5000 dilution) (Figure 29). Positive immunoreactivity of bands corresponding to avidin standards indicated that both Sigma avidin and recombinant avidin were transported to the haemolymph of *M. brassicae* larvae in similar amounts following feeding. The presence of avidin in the diet and bound to gut polypeptides was confirmed by the western blot analysis of gut contents and samples of midgut cleared of the contents (Section 2.4C).

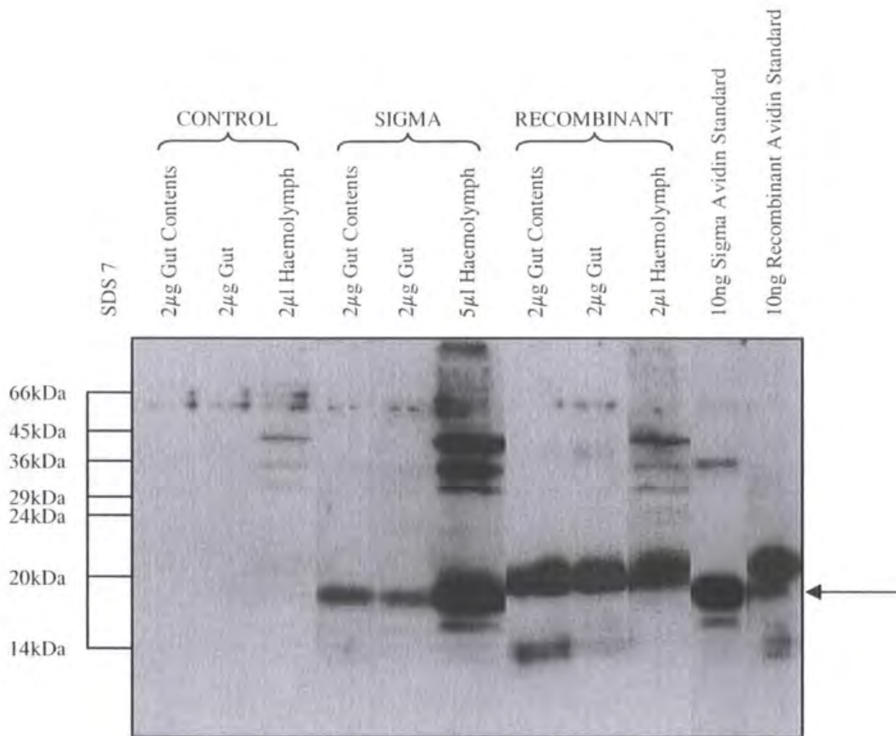


Figure 29.
Western blot analysis of *M. brassicae* larvae gut contents, gut and haemolymph after 24 hours feeding on either control or avidin (Sigma and recombinant) diet, probed with anti-avidin antibodies (1:5000 dilution). A touch exposure is shown. Blotted from a 15% SDS-PAGE gel. The blot shows similar amounts of both Sigma and recombinant avidin present in gut contents, bound to gut polypeptides and in the haemolymph.

3.3C. Effects of Avidin upon Injection

Newly moulted fifth stadium *M. brassicae* larvae were injected with 50μg (5μl total) of avidin to determine its biological activity, if any, *in vivo*. Phosphate buffered saline was injected as a control. Larvae showed no visual effects and development was normal until they successfully pupated, showing that avidin has no effect when injected directly into the haemolymph.

3.3D. Toxicity of Recombinant Avidin Delivered Orally

To compare the effects of feeding recombinant avidin against Sigma avidin to lepidopteran larvae, 24 hour old *M. brassicae* larvae were raised on different treatments of avidin-containing artificial diet. Concentrations of 0.1% [w/w] (1000ppm, 5mg avidin in 5g diet), 0.01% [w/w] (100ppm, 0.5mg) or 0.001% [w/w] (10ppm, 0.05mg) recombinant avidin and 0.1% [w/w] (1000ppm, 5mg) Sigma avidin were tested. Twenty larvae were used for each treatment including a 0.1% [w/w] casein artificial diet control. Larvae were monitored for insecticidal activity and also by weight once large enough (day 9). Results can be seen in figures 30, 31, 32, 33, 34 and 35.

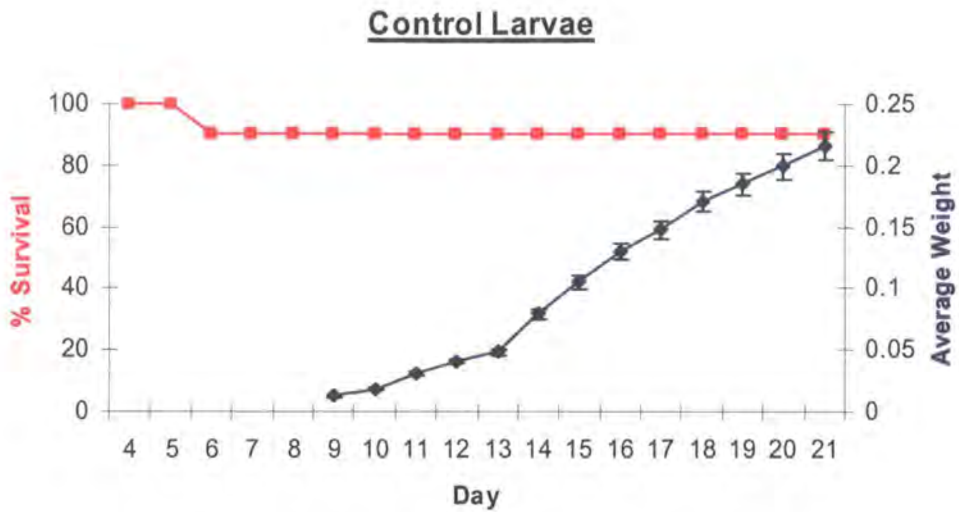


Figure 30.
A graph of *M. brassicae* larvae survival and average weight whilst feeding on control (casein) diet over 21 days.

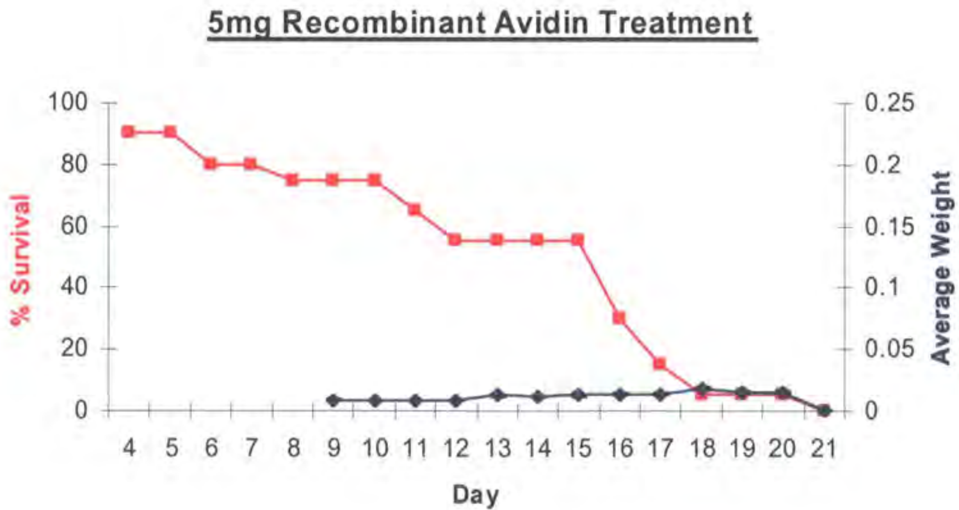


Figure 31.
A graph of *M. brassicae* larvae survival and average weight whilst feeding on 0.1% (1000ppm) avidin diet over 21 days.

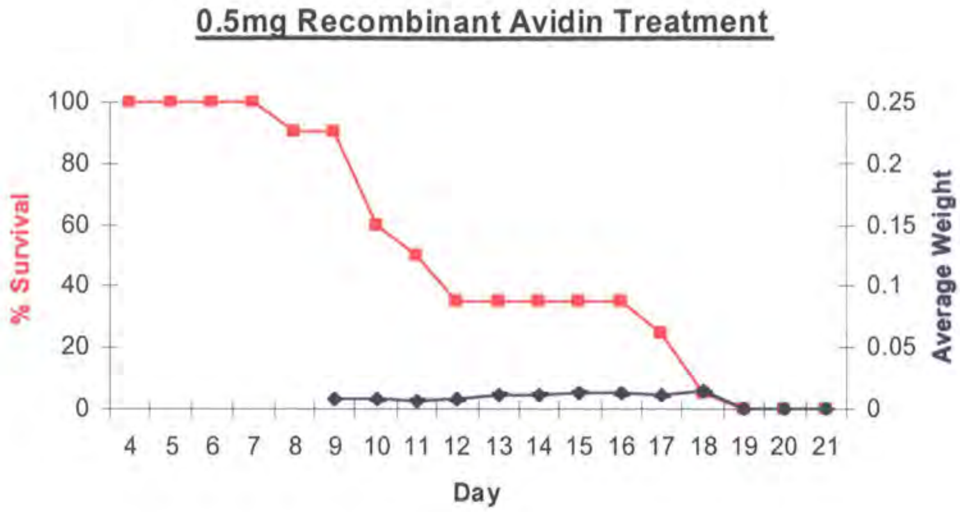


Figure 32.
A graph of *M. brassicae* larvae survival and average weight whilst feeding on 0.01% (100ppm) avidin diet over 21 days.

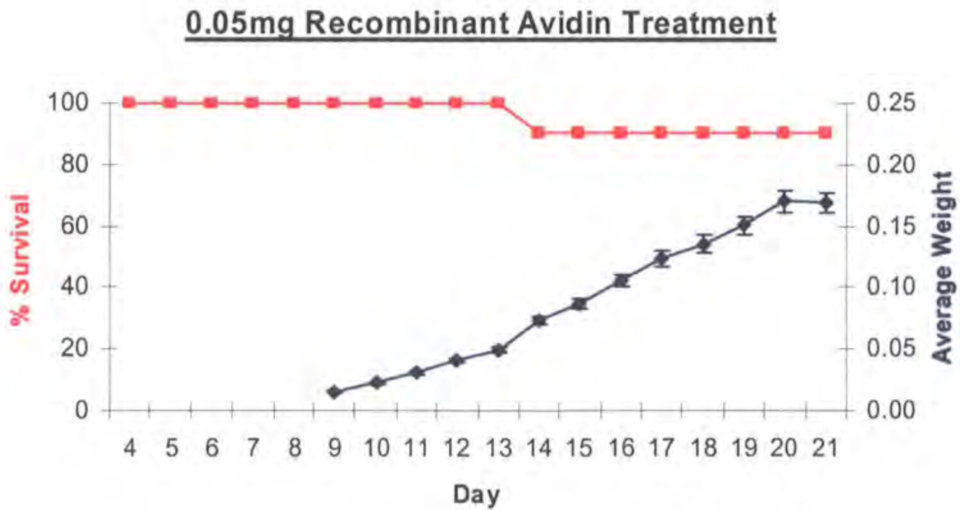


Figure 33.
A graph of *M. brassicae* larvae survival and average weight whilst feeding on 0.001% (10ppm) avidin diet over 21 days.

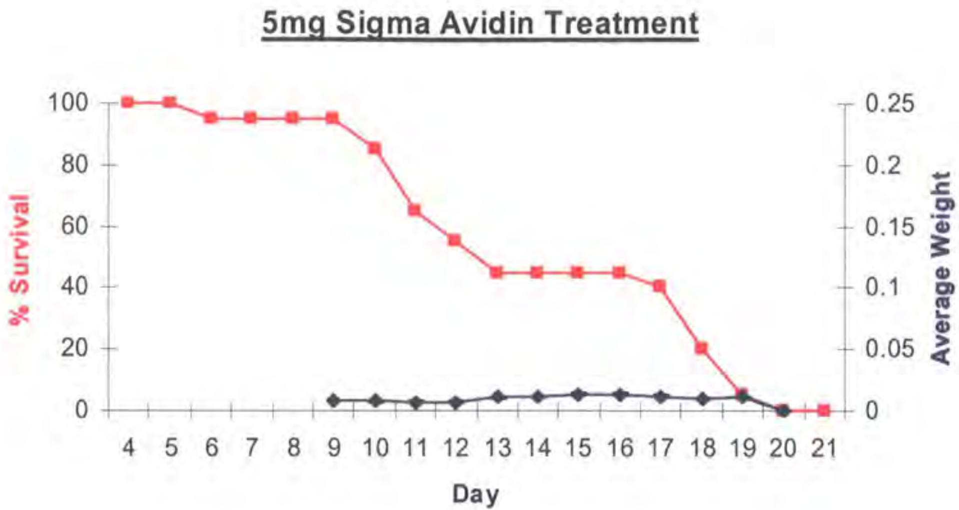


Figure 34.
A graph of *M. brassicae* larvae survival and average weight whilst feeding on 0.1% (1000ppm) Sigma avidin diet over 21 days.

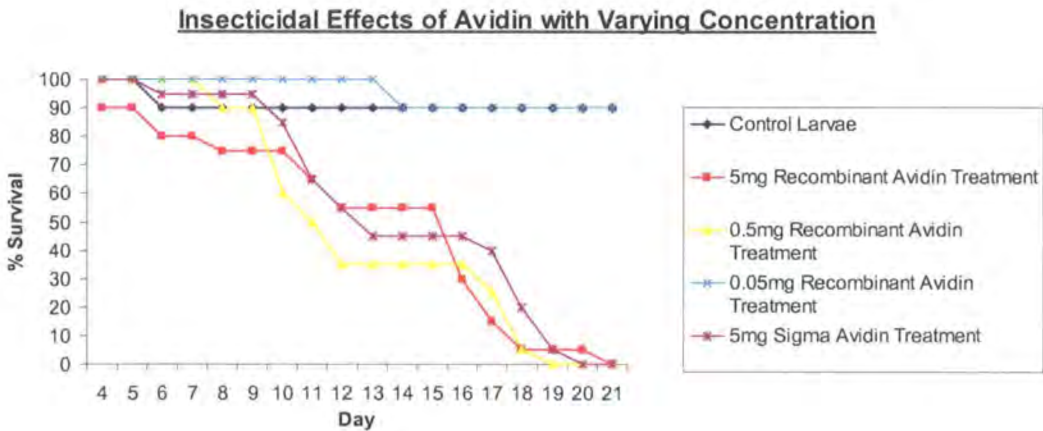


Figure 35.
A comparative graph of *M. brassicae* larvae survival over 21 days for all of the bioassay diet treatments.

The results demonstrate that both recombinant and Sigma avidin had a significant and dose dependent effect on the survival and growth of *M. brassicae* larvae. One hundred percent mortality was observed after 21 days of exposure to 1000ppm (5mg) and 100ppm (0.5mg) recombinant avidin and similarly with 1000ppm (5mg) of Sigma avidin, in contrast to the casein controls where 90% survival was observed. Recombinant avidin included in diet at 100ppm is therefore as equally as toxic as 1000ppm. Larvae for these two treatments showed very little weight gain (always between 50% and 90% less than the control) and their survival also highlights critical points in the insecticidal activity around day 10 and day 16, coinciding with ecdysis (Figures 31, 32 and 34). The dose dependency can clearly be seen between 10ppm and 100ppm recombinant avidin where 90% survival and 100% mortality were observed over 21 days respectively. The lethal dose is somewhere between 10ppm and 100ppm. The toxicity result may, in reality, be slightly quicker than the 21 days observed here as it is known that the artificial diet contains biotin. Therefore, when diet was replenished a little more biotin would be added, perhaps counteracting the effects of the avidin slightly.

3.3E. Mechanics of Recombinant Avidin Transport

For avidin to be a successful carrier of insecticidal proteins, it is important that the protein linkage remains stable and intact until the fusion reaches the haemolymph. To assess this suitability, fluorescein-biotin was used to determine whether avidin sequestered biotin remained bound during transfer to the haemolymph. The fluorescence allowed the biotin component to be traced (excitation at 485nm and emission at 538nm). Initially, haemolymph was extracted from sixth stadium *M. brassicae* larvae, pulse centrifuged for 30 seconds and 200μl run through an analytical gel filtration column to obtain a profile and naturally occurring fluorescence values for the separation (Figure 37). Four hundred microlitre fractions were collected across the separation and their fluorescence measured (Figure 37 highlighted in red). The highest background fluorescence recorded was 0.10, therefore anything above this figure was attributed to a significant compound. A 1mg/ml fluorescein-biotin stock was serially diluted and fluorescence measured, giving a detectable lower limit of 10ng/ml (Figure 36).

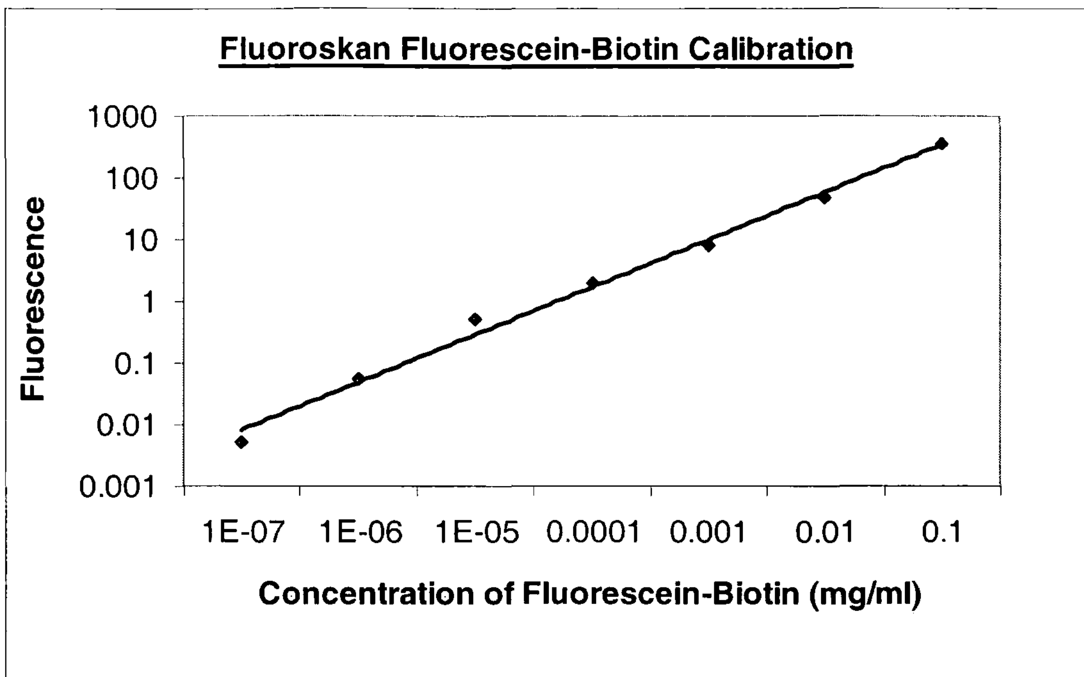


Figure 36.
A graph of Fluoroskan fluorescence values for varying amounts of fluorescein-biotin. A fluorescence value between 10 and 50 was acceptable. Reading off the graph gives a fluorescein-biotin concentration of 0.01mg (10ng/ml).

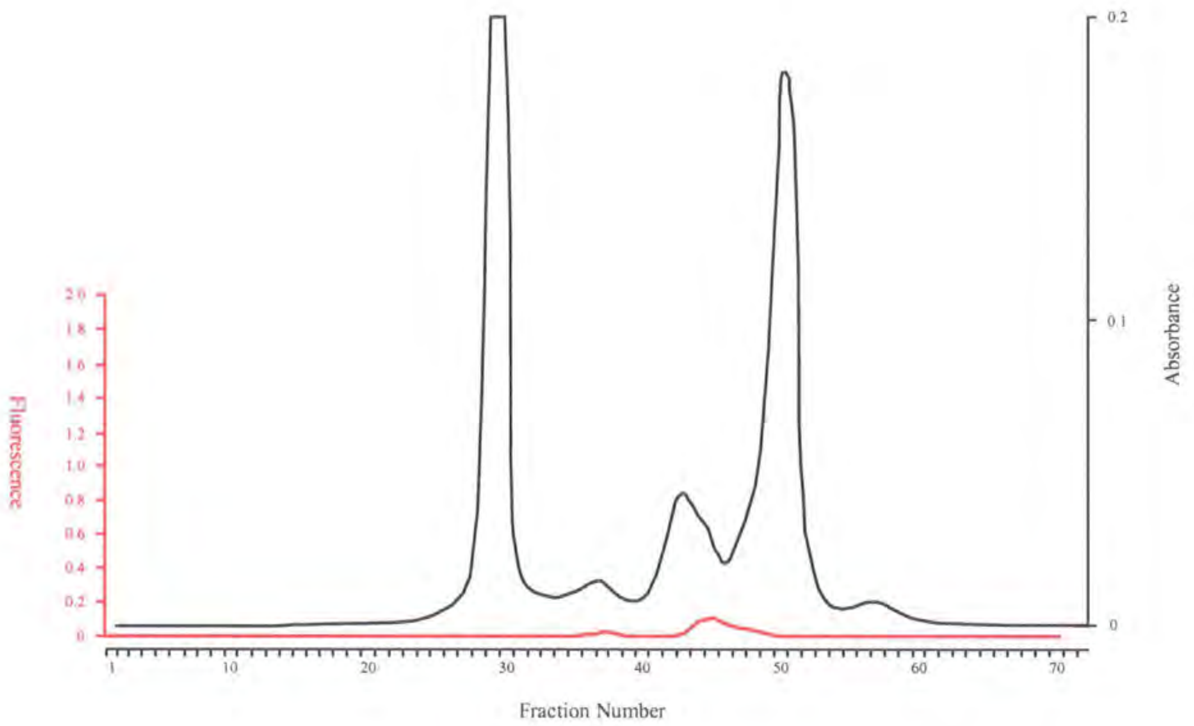


Figure 37.
Gel filtration profile for *M. brassicae* larvae haemolymph (black). Fluorescence readings for the collected gel filtration fractions (red). The highest background fluorescence recorded was 0.1 in fraction 45.

To establish where in the gel filtration profile fluorescence would occur, haemolymph was spiked with 2μg of fluorescein-biotin only (Figure 38) and then 2μg of avidin fluorescein-biotin conjugate (Figure 39). Ten sixth stadium *M. brassicae* larvae were then fed two 5μl droplets containing 25μg of avidin fluorescein-biotin conjugate each (with 0.5% sucrose) and left to feed on diet in the dark for one hour. Haemolymph was then extracted and 200μl separated by gel filtration (Figure 40(a)). The fluorescence readings (highlighted in red) show a similar distribution to that of haemolymph spiked with avidin fluorescein-biotin (Figure 39), indicating that the conjugate remains intact during transport to the haemolymph. Western blot analysis also confirmed that the avidin fluorescein-biotin conjugate remains intact following uptake into the haemolymph as the presence of avidin was identified in those fractions that exhibited high fluorescence (Figure 40(b)). The blot also shows a little antibody non-specific cross-reactivity in the control haemolymph and high levels of other immunoreactivity in the fractions possibly due to avidin monomers (minus the His-tag), dimers, trimers and tetramers being present (highlighted with arrows on figure 40(b).

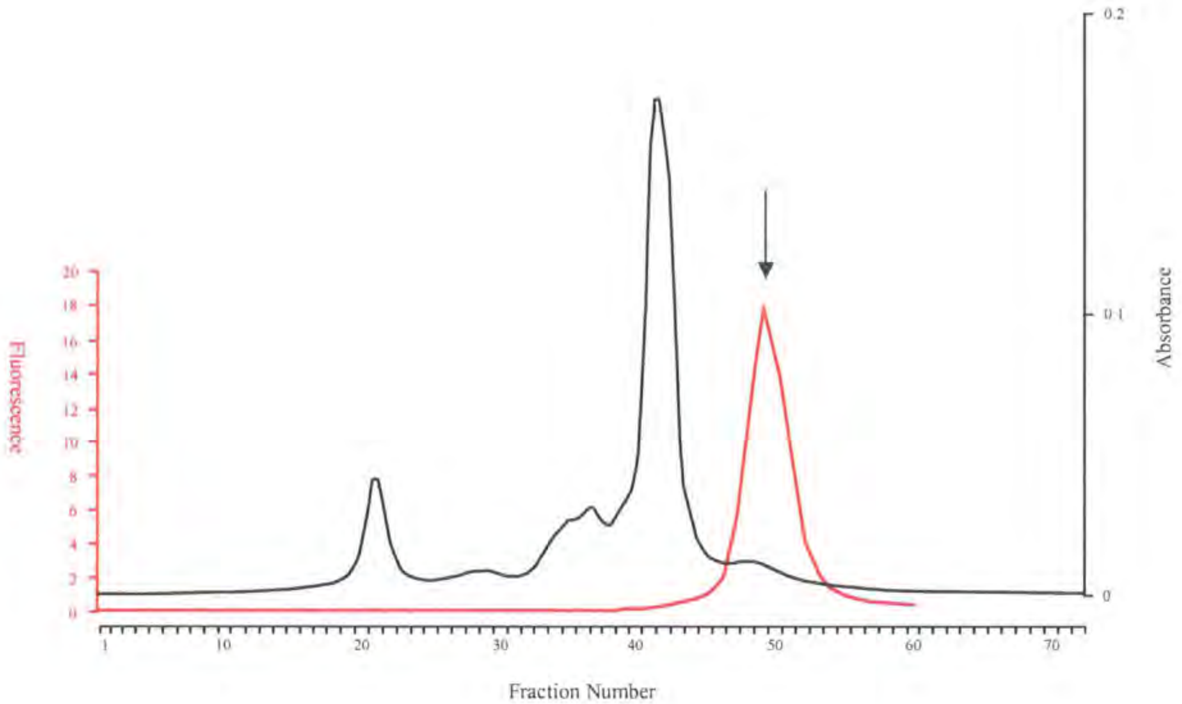


Figure 38.
Gel filtration profile for *M. brassicae* larvae haemolymph spiked with 2µg of fluorescein-biotin (black). Fluorescence readings for the collected gel filtration fractions (red). The highest fluorescence recorded due to fluorescein-biotin was 17.95 in fraction 49 (highlighted with an arrow).

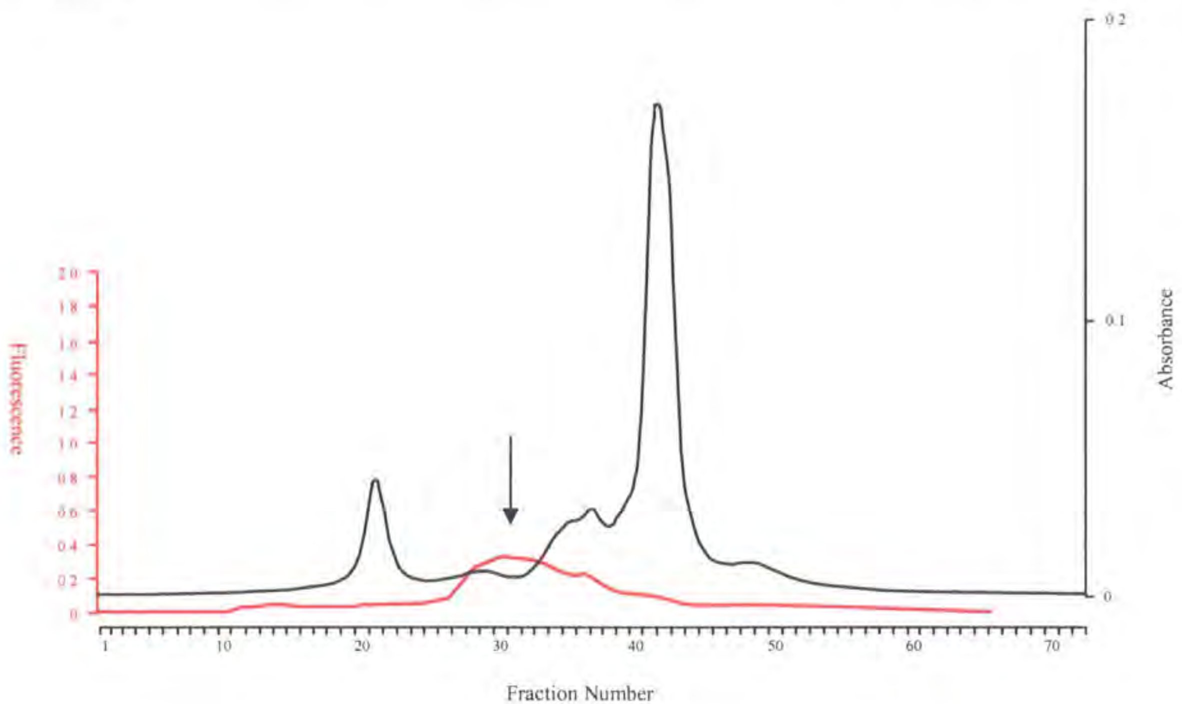


Figure 39.
Gel filtration profile for *M. brassicae* larvae haemolymph spiked with 2µg of avidin fluorescein-biotin conjugate (black). Fluorescence readings for the collected gel filtration fractions (red). The highest fluorescence recorded due to avidin fluorescein-biotin conjugate was 0.33 in fraction 30 (highlighted with an arrow).

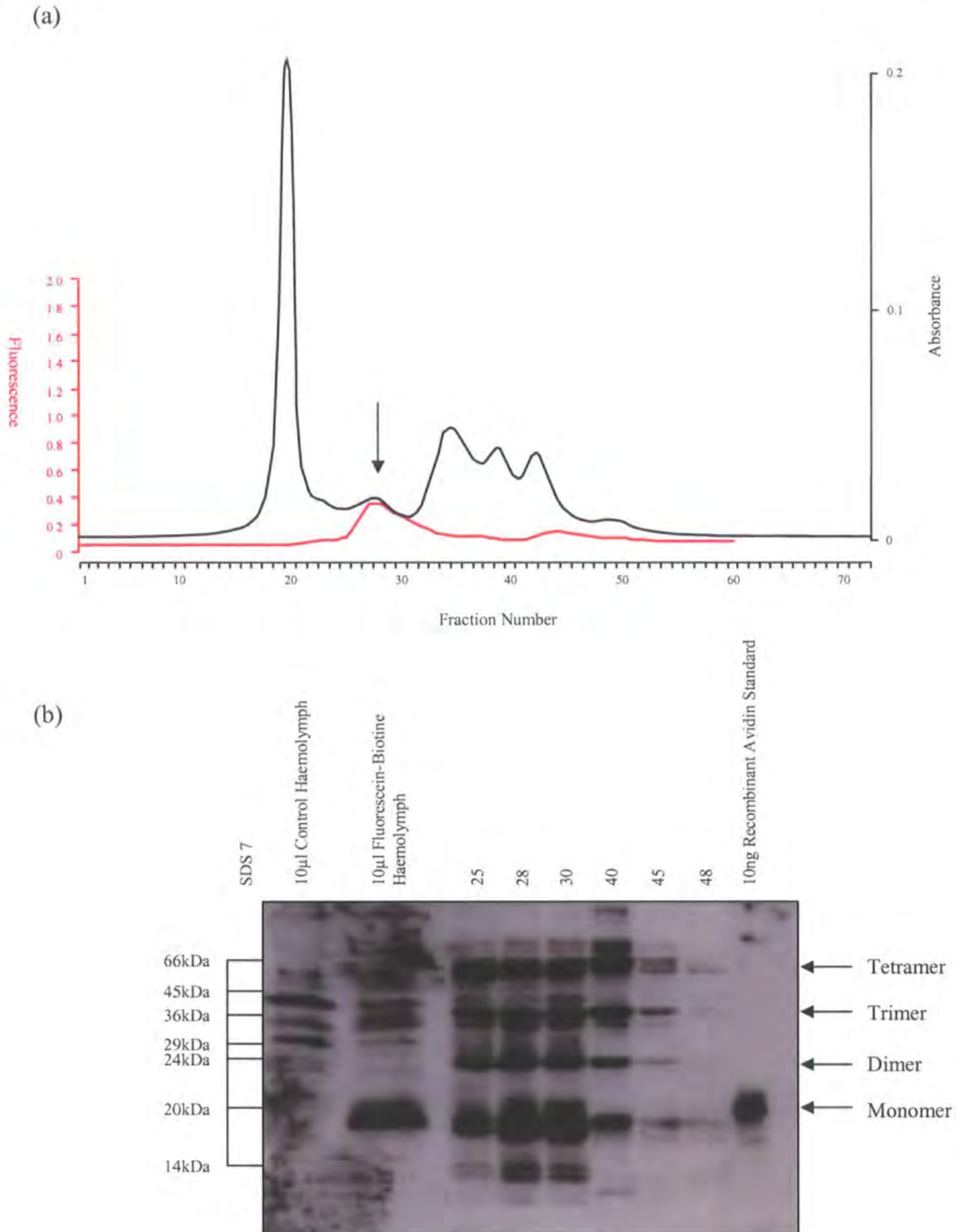


Figure 40.
(a) Gel filtration profile for haemolymph from 50µg avidin fluorescein-biotin conjugate droplet fed *M. brassicae* larvae (black). Fluorescence readings for the collected gel filtration fractions (red). The highest fluorescence recorded due to avidin fluorescein-biotin conjugate was 0.34 in fraction 27 (highlighted with an arrow). (b) Western blot analysis of fractions from the gel filtration in (a) probed with anti-avidin antibodies (1:5000 dilution). A touch exposure is shown. Blotted from a 15% SDS-PAGE gel.
The blot shows recombinant avidin (with His-tag) present in fractions 28 and 30 and absent from the others.

3.3F. Antibody Coupled Avidin

One potential use for avidin would be as a carrier to transfer antibody to the haemolymph of insects when delivered orally. To explore this possibility, anti diuretic hormone receptor antibody was biotinylated and allowed to bind to avidin in excess. Dialysis into PBS was then carried out to remove excess avidin and make it suitable for feeding to insects. Ten percent [w/v] avidin-biotinylated antibody was then included in artificial aphid diet and in *M. brassicae* larvae artificial diet at 0.0001% [w/w]. Fifty *S. avenae* and four *M. brassicae* larvae were exposed to diet for 24 hours and two remaining *M. brassicae* larvae for 62 hours. Haemolymph was then extracted and analysed by western blot probed with secondary antibody (1:5000 dilution) (Figures 41 and 42). Preliminary western blot analysis of the sensitivity of secondary antibody had determined that the limit of detection of the anti-diuretic hormone receptor antibody was 10ng. Diuretic hormone receptor antibody only was used as a positive control and avidin only was included as a negative control. Figure 41 shows an absence of any reactivity due to antibody in the haemolymph of *S. avenae*. By contrast, western blot analysis of haemolymph extracted from *M. brassicae* larvae showed the surprising presence of antibody in both avidin-antibody and antibody alone fed insects (Figure 42).

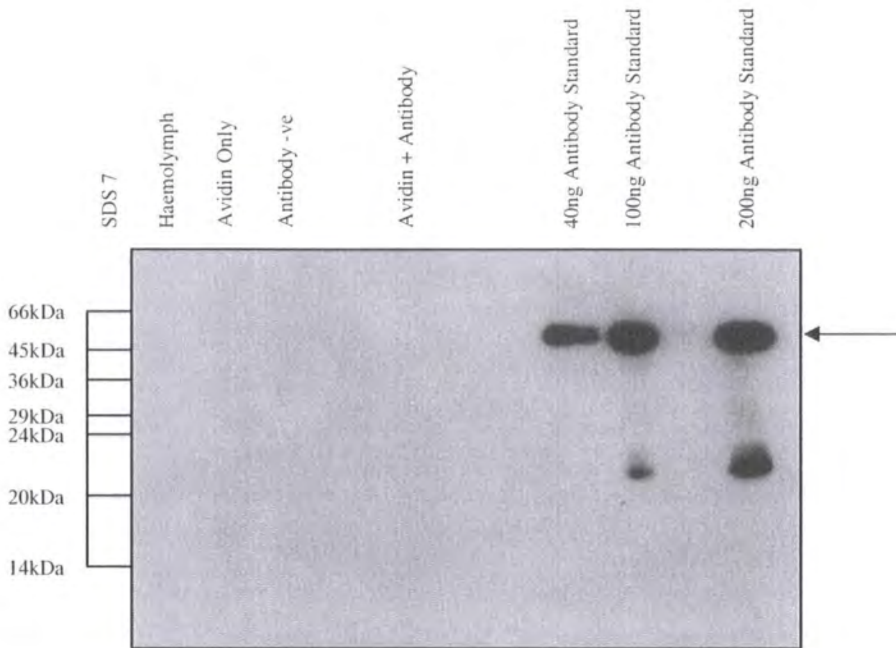


Figure 41.
Western blot analysis of *S avenae* haemolymph after 24 hours feeding on diet containing 10% avidin-biotinylated antibody, probed with secondary antibody (1:5000 dilution). A 10 second exposure is shown. Blotted from a 15% SDS-PAGE gel.
The blot shows the absence of any avidin-biotinylated antibody in the haemolymph.

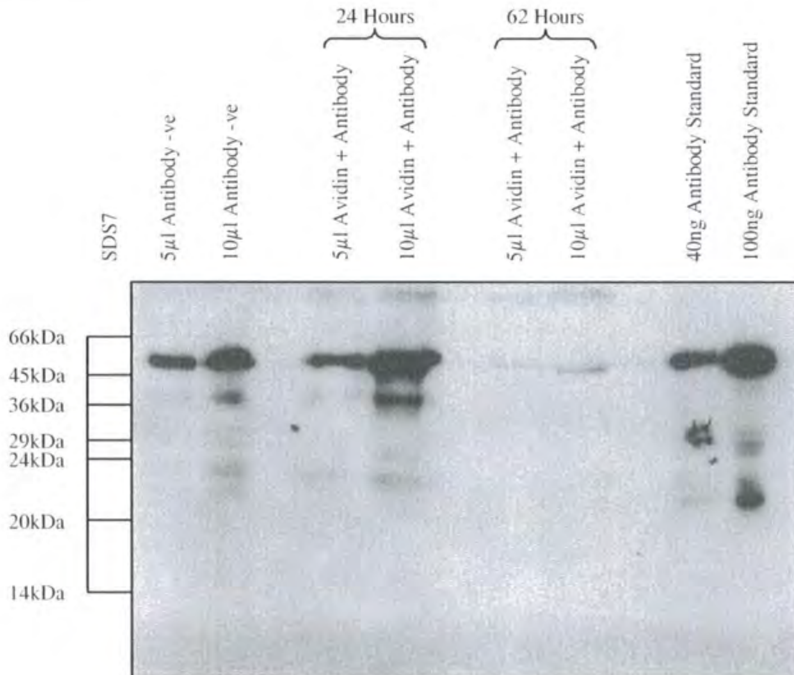


Figure 42.
Western blot analysis of *M. brassicae* haemolymph after 24 and 62 hours feeding on diet containing 0.0001% avidin-biotinylated antibody, probed with secondary antibody (1:5000 dilution). A one minute exposure is shown. Blotted from a 15% SDS-PAGE gel.
The blot shows the presence of avidin-biotinylated antibody in the haemolymph. It also shows antibody alone is transported.

3.3G. IgG Transport

To check if antibody could be transported to the haemolymph un-aided and that the sample had not been contaminated in section 3.3F, the experiment was repeated with a general rabbit IgG purchased from Sigma. A preliminary western blot was carried out to identify the limit of detection as 50ng (Figure 43). Very little background reactivity was also observed in *M. brassicae* larvae haemolymph.

Immunoglobulin G was included in artificial diet at 200ppm and fed to fifth stadium *M. brassicae* larvae for 20 minutes, four hours and 24 hours. A feeding chase was also incorporated to establish what happened to the antibody if it was present. For this, larvae were similarly fed IgG diet for 24 hours and then fed normal diet for either four hours or 24 hours after. Six larvae were used for each treatment. At the time points, haemolymph and guts were extracted and analysed by western blot using secondary antibody only (1:5000 dilution) (Figures 44 and 45). The result shows IgG present in the haemolymph up to a certain maximum concentration after 20 minutes, four hours and 24 hours only and not bound to the gut at any stage. At 24 hours and after, the IgG is reduced and eventually cleared.

The IgG uptake experiment was also carried out in *S. avenae* to determine if antibody was transported to the haemolymph of Hemiptera. Fifty micrograms of IgG was included in artificial diet and 50 aphids were allowed to feed for 24 hours. Haemolymph was extracted and analysed by western blot using secondary antibody only (1:5000 dilution) (Figure 46). A similar experiment was carried out at the same time using avidin only (Figure 47). The results show that neither IgG nor avidin was present in the haemolymph.

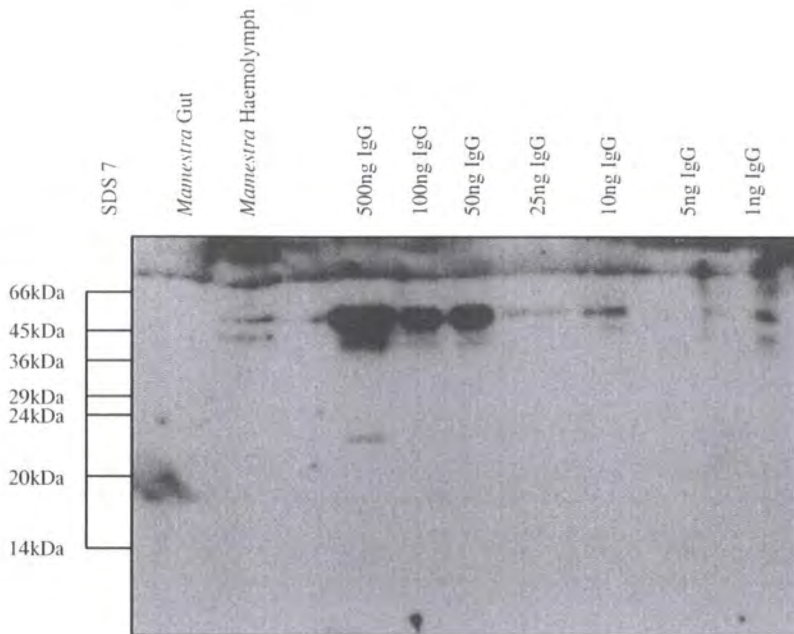


Figure 43.
Western blot analysis of rabbit IgG, probed with secondary antibody (1:5000 dilution). A 10 second exposure is shown. Blotted from a 15% SDS-PAGE gel.
The blot shows the acceptable limit of detection is 50ng. There is also very little background reactivity in *M. brassicae* gut and haemolymph.

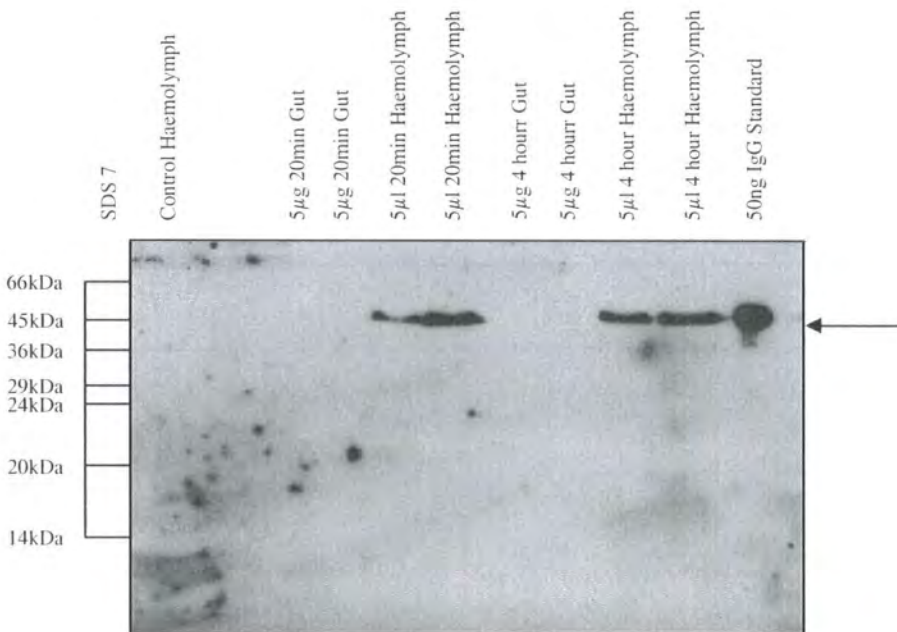


Figure 44.
Western blot analysis of *M. brassicae* gut and haemolymph after various times feeding on diet containing 200ppm IgG, probed with secondary antibody (1:5000 dilution). A five second exposure is shown. Blotted from a 15% SDS-PAGE gel.
The blot shows the presence of IgG in the haemolymph only. There appears to be no more IgG present in the haemolymph after the longer feeding.

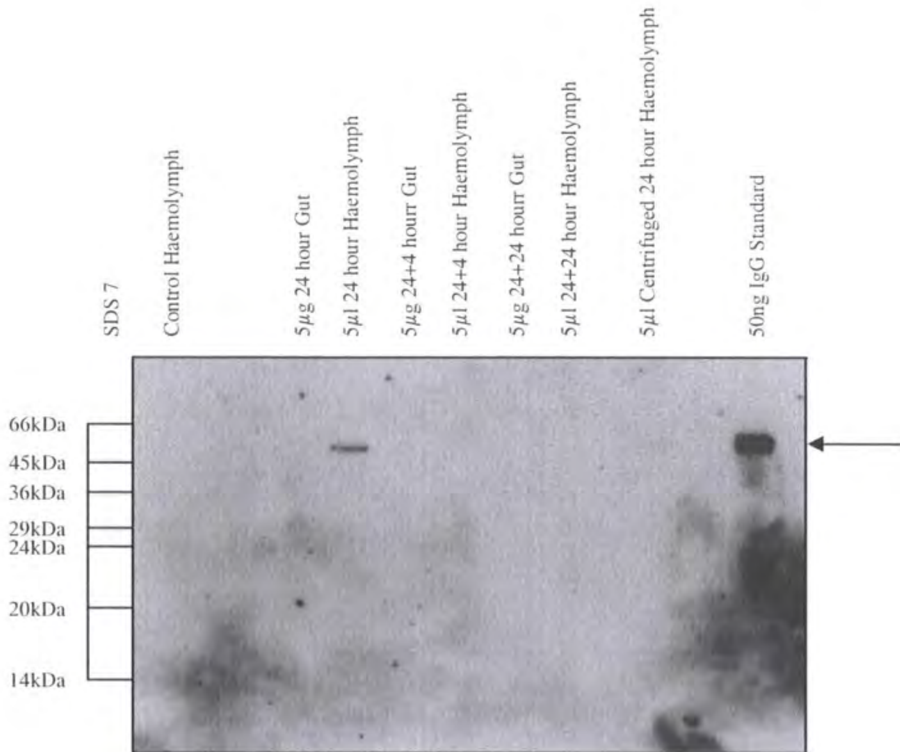


Figure 45.
Western blot analysis of *M. brassicae* gut and haemolymph after various times feeding on diet containing 200ppm IgG, probed with secondary antibody (1:5000 dilution). A 30 second exposure is shown. Blotted from a 15% SDS-PAGE gel.
The blot again shows the presence of IgG in the haemolymph only. It appears to be utilized or cleared from the haemolymph after 24 hours.

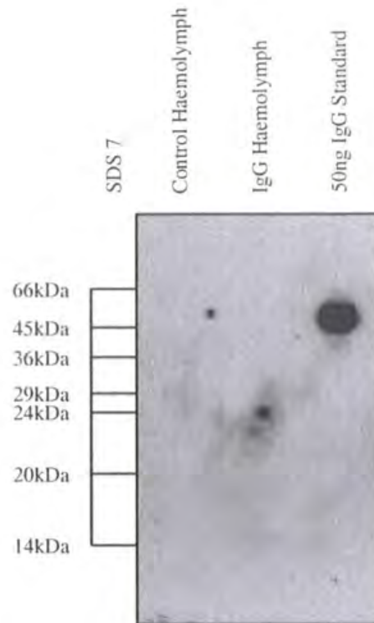


Figure 46.
Western blot analysis of *S. avenae* haemolymph after 24 hours feeding on diet containing 50μg of IgG, probed with secondary antibody (1:5000 dilution). An eight minute exposure is shown. Blotted from a 15% SDS-PAGE gel.
The blot shows the absence of any IgG in the haemolymph.

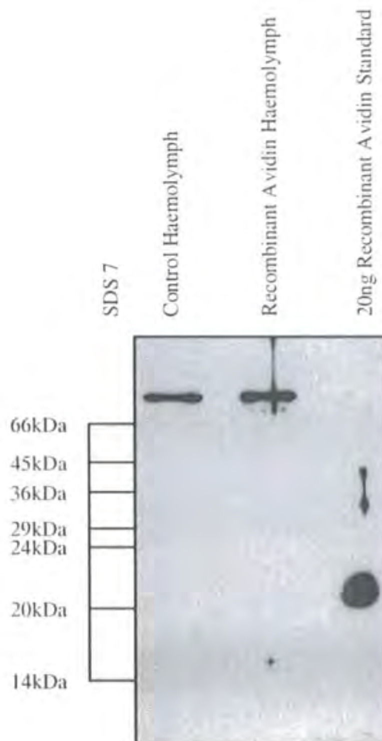


Figure 47.
Western blot analysis of *S. avenae* haemolymph after 24 hours feeding on diet containing 50μg of recombinant avidin, probed with anti-avidin antibodies (1:5000 dilution). A 30 second exposure is shown. Blotted from a 15% SDS-PAGE gel.
The blot shows the absence of any recombinant avidin in the haemolymph.

4.0. Discussion

4.1. *Eulophus pennicornis* Venom

It is a well known fact that animal and insect venoms are a rich source of biologically active proteins and may, potentially be exploited as new insecticides as a number of the components have negative effects on the prey / host insect (Moreau and Guillot, 2005). In this study of *E. pennicornis* venom alone, alongside the reprolysin-like metalloprotease and juvenile hormone-inducible protein already found (awaiting publication), a splicing factor, serine protease and a neprilysin-like protein were discovered by NCBI-BLAST sequence similarity searches. A speculation can be made regarding their function in venom. Production of necessary proteins and development of a host larva would be negatively affected by the removal of introns and exons in mRNA with the presence of the splicing factor. The serine protease may simply be involved as a digestive enzyme, cleaving peptide bonds within a host larva or, perhaps more likely, it may have a function similar to that of mammals where it would play a role in inflammation, blood clotting and immune system function. Neprilysin is similar to reprolysin (Section 1.1) in that it is a zinc-dependent metalloprotease. Its function is to process small secreted peptides especially in neural tissue. Its incorrect function has been associated with Alzheimer's disease (Wang et al., 2003). Therefore, perhaps in parasitic wasp venom the neprilysin-like protein would be involved in affecting neurons with the target insect. These speculations would require further investigation however.

All of the compounds discovered in this venom gland screen have potential for use in insect pest control. This work, therefore, substantiates the statement by Whetstone and Hammock (2007) that the possibilities for insecticidal compounds from natural sources are virtually limitless given a delivery system.

4.2. Reprolysin-like Metalloprotease

Until recently, the isolated components of parasitic wasp venom were rarely studied further. The purification and demonstration of activity by injection of *E. pennicornis* reprolysin-like metalloprotease gave potential for more analysis of

function. Using a reprolysin-like metalloprotease in Fusion Protein Technology would, perhaps, allow it to be delivered to the haemolymph of pest insects in an active form. This study demonstrated that reprolysin-like metalloprotease proved difficult to express. Experiments with ASAII confirmed that experimental procedures were working correctly, but still only low levels of reprolysin-like metalloprotease were obtained using the pPICZαB vector, which was contradictory to initial work (Section 3.1A). Another reprolysin-like metalloprotease in parasitic wasp venom has been reported previously (Parkinson et al., 2002), with its similarity to a snake venom metalloprotease detailed. Snake venoms have been well researched with many other metalloproteases and reprolysins being discovered, expressed and characterised but no similar observations to the phenomena observed in this study are reported. Sequence structural analysis was still carried out on the original isolated reprolysin-like metalloprotease (Figure 1, Section 1.1B). Because it has long been known that snake venom metalloproteases are activated by a cysteine switch-like mechanism (Grams et al., 1993), it was thought that the *E. pennicornis* reprolysin-like metalloprotease may act similarly. This mechanism involves the conserved cysteine-containing sequence PKMCGVT. Binding of the cysteine thiol motif to the zinc at the active site results in the inactivation of the catalytic domain. Proteolytic processing then causes a switch in the binding structure and removes the inactivation (Grams et al., 1993). However, the conserved sequence, or anything structurally similar, is not present in the reprolysin-like metalloprotease, so this method was dismissed along with others based on conserved domains.

In conclusion, considering all of the results, the lack of expression of reprolysin-like metalloprotease may be due to the protein being toxic towards the yeast. Certain elements of this study, such as the lumpy cultures observed in the longer incubations of the time dependent experiment and the greatly reduced number of colonies seen upon yeast transformation gives evidence towards this. Without expressional studies, activation mechanisms can only be based on the sequence presented and structural predictions from protein databases. Finding no conserved domains involved in activation perhaps means the reprolysin-like metalloprotease reported upon here may be similar to the TACE (tumour necrosis factor-alpha-converting enzyme) zymogen investigated by Milla et al. (2006), with its activation mechanism dissimilar to many other metalloproteases and still to be defined.

4.3. Avidin

The potential of avidin to act as an insecticide has recently been exploited with the production of transgenic rice, corn and maize expressing this protein (Kramer et al., 2000; Yoza et al., 2005). From previous experiments (Section 3.3), it was believed that avidin may be useful as a carrier in Fusion Protein Technology and for the attachment of biotinylated proteins antibodies also. The current study shows that recombinant avidin expresses well and purifies easily using a nickel column, giving a single peak which is almost 100% pure. Recombinant avidin binds to gut polypeptides and is transported to the haemolymph in comparable amounts to commercially available avidin, as confirmed by the bioassays with *M. brassicae* larvae. Therefore, recombinant avidin has haemolymph translocation ability suitable to act as a carrier protein for fused peptides and attached biotinylated proteins. In addition, insecticidal activity of recombinant avidin has been demonstrated in *M. brassicae* larvae, another lepidopteran species to add to those tested by Morgan et al. (1993) and Markwick et al. (2001). The effective avidin dose attained was 100ppm, consistent to that obtained when used in transgenic crops (Kramer et al., 2000; Yoza et al., 2005). Similar stunted growth and mortality as described by Miura et al. (1967) were also seen at this concentration. The final three experiments of this study gave valuable information regarding the working potential of avidin, with none of them being previously documented. Injection of 50μg of avidin was non-lethal, a finding expected, as any free biotin in the haemolymph would be utilized rapidly by cells in enzymatic processes and therefore would be unavailable for avidin sequestration. Fluorescence experiments reveal that once biotin is sequestered by recombinant avidin, it is transferred to the haemolymph as a conjugate, demonstrating the possibility of using recombinant avidin as a carrier protein further. Attached peptides should not be cleaved off before they are transferred to the haemolymph, allowing recombinant avidin to be used in Fusion Protein Technology. Recombinant avidin was unable to transport biotinylated antibody to the haemolymph of *S. avenae* which may be due to there being an insufficient amount of haemolymph extracted or the antibody not being present above the 10ng detectable limit. The most surprising result was that antibody was transported into the haemolymph of *M. brassicae* larvae without avidin. This suggests that the lepidopteran peritrophic membrane and gut epithelium is more permeable than

previously thought, allowing for the uptake of larger proteins such as IgG (155kDa) into the circulatory system. The mechanism of how this antibody uptake occurs has yet to be elucidated, but experiments with rabbit IgG show that there is no antibody binding to the gut. Also, with longer periods of feeding on diet containing IgG, the concentration in the haemolymph remains the same but is utilized or cleared after 24 hours. Perhaps there is a method that lepidopterans use to actively sample their gut contents. By transferring a small amount of ingested material to the haemolymph, lepodopterans could assess whether useful components were present. If required, these could then be completely transported and utilized appropriately, or, less useful components could be expelled and excreted. This finding opens up the prospect of using lethal antibodies as insecticides. An antibody which is lethal at low concentrations could be transferred to the haemolymph in this active sampling, where it would have activity before it had the chance to be removed again. There is also potential for future similar experiments to be conducted in other insect orders.

Bibliography

Allsopp PG, McGhie TK. 1996. Snowdrop and wheatgerm lectins and avidin as antimetabolites for the control of sugarcane whitegrubs. *Entomol. Exp. Appl.* 80: 409-414.

Barrett AJ, Rawlings ND, Woessner JF. 1998. Handbook of Proteolytic Enzymes. Academic Press Inc. publishers, London.

Bertani G, 1951. Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *J. Bacteriol.* 62: 293-300.

Beugnet F, Chardonnet L. 1995. Tick resistance to pyrethroids in New Caledonia. *Veterinary Parasitology* 56: 325-338.

Bjerrum OJ, Schafer-Nielsen C. 1986. Buffer systems and transfer parameters for semi-dry electroblotting with a horizontal apparatus. In Dunn MJ., editor Electrophoresis 1986. Weinheim, Germany. VCH Publishers. 315-327.

Bode W, Gomis-Rüth FX, Stöckler W. 1993. Astacins, serralysins, snake venom and matrix metalloproteinases exhibit identical zinc-binding environments (HEXXHXXGXXH and Met-turn) and topologies and should be grouped into a common family, the 'metzincins'. *FEBS Lett.* 331(1-2): 134-140.

Bown DP, Wilkinson HS, Gatehouse JA. 1997. Differentially regulated inhibitor sensitive and insensitive protease genes from the phytophagous insect pest *Helicoverpa armigera*, are members of complex multi-gene families. *Insect Biochem. Mol. Biol.* 27: 625-638.

Boyd WC, Reguera RM. 1949. Studies on haemagglutinins present in seeds of some representatives of the family leguminosae. *J. Immunol.* 62: 333-339.

Boyd WC, Shapleigh E. 1954. Specific precipitating activity of plant agglutinins (lectins). *Science* 119: 419.

Bruins BG, Scharloo W, Thörig GEW. 1991. The harmful effect of light on *Drosophila* is diet dependent. *Insect Biochem.* 21: 535-539.

Cereghino JL, Cregg JM. 2000. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiol. Rev.* 24: 45-66.

Chen JJ, Chen GH, Hsu HC, Li SS, Chen CS. 2004. Cloning and functional expression of a mungbean defensin VrD1 in *Pichia pastoris*. *J. Agric. Food Chem.* 52: 2256-2261.

Cheng Zhu Y, Adamczyk JJ, West S. 2005. Avidin, a potential biopesticide and synergist to *Bacillus thuringiensis* toxins against field crop insects. *J. Econ. Entomol.* 98: 1566-1571.

Cino J. 1999. *Pichia pastoris* fermentation using a Bioflow 110 bench-top fermenter. New Brunswick Scientific. Available to download at http://nbsc.com/tech_papers_1.aspx

DeLange RJ, Huang TS. 1971. Egg White Avidin; III. Sequence of the 78-residue middle cyanogen bromide peptide. Complete amino acid sequence of the protein subunit. *J. Biol. Chem.* 246(3): 698-709.

Du C, Nickerson KW. 1995. Insecticidal activity of avidin. *Mtg. Soc. Invert. Pathol.* 70.

Edwards JP, Bell HA, Audsley N, Marris GC, Kirkbride-Smith A, Bryning G, Frisco C, Cusson M. 2006. The ectoparasitic wasp *Eulophus pennicornis* (Hymenoptera: Eulophidae) uses instar-specific endocrine disruption strategies to suppress the development of its host *Lacanobia oleracea* (Lepidoptera: Noctuidae). *J. Insect Physiol.* 52: 1153-1162.

Elfstrand M. 1898. Über blutkorperchenagglutinierende Eiweisse. In *Görbersdorfer Veröffentlichungen* a. Band I (R. Kobert, ed.) pp 1-159. Enke, Stuttgart.

Fitches E, Woodhouse SD, Edwards JP, Gatehouse JA. 2001. In vitro and in vivo binding of snowdrop (*Galanthus nivalis* agglutinin; GNA) and jackbean (*Canavalia ensiformis*; Con A) lectins within tomato moth (*Lacanobia oleracea*) larvae; mechanisms of insecticidal action. *J. Insect Physiol.* 47: 777-787.

Fitches E, Audsley N, Gatehouse JA, Edwards JP. 2002. Fusion proteins containing neuropeptides as novel insect control agents: snowdrop lectin delivers fused allatostatin to insect haemolymph following oral ingestion. *Insect Biochem. Mol. Biol.* 32: 1653-1661.

Fitches E, Edwards MG, Mee C, Grishin E, Gatehouse AMR, Edwards JP, Gatehouse JA. 2004a. Fusion proteins containing insect-specific toxins as pest control agents: snowdrop lectin delivers fused insecticidal spider venom toxin to insect haemolymph following oral ingestion. *J. Insect Physiol.* 50: 61-71.

Fitches E, Wilkinson H, Bell H, Bown DP, Gatehouse JA, Edwards JP. 2004b. Cloning, expression and functional characterisation of chitinase from larvae of tomato moth (*Lacanobia oleracea*): a demonstration of the insecticidal activity of insect chitinase. *Insect Biochem. Mol. Biol.* 34: 1037-1050.

Garthwaite DG, Thomas MR, Anderson H, Stoddart H. 2004. Pesticide Usage Survey Report 202, Arable Crops in Britain. DEFRA and Scottish Exec. Environment and Rural Affairs Department.

Gellissen G, Hollenberg CP. 1997. Application of yeasts in gene expression studies: a comparison of *Saccharomyces Cerevisiae*, *Hansenula polymorpha* and *Kluyveromyces lactis* – a review. *Gene* 190: 87-97.

Grams F, Huber R, Kress LF, Moroder L, Bode W. 1993. Activation of snake venom metalloproteases by a cysteine switch-like mechanism. *FEBS Lett.* 335(1): 76-80.

Gurkan C, Ellar DJ. 2003. Expression of the *Bacillus thuringiensis* Cyt2Aa1 toxin in *Pichia pastoris* using a synthetic gene construct. *Biotechnol. Appl. Biochem.* 38: 25-33.

Gyorgy P, Rose CS, Eakin RE, Snell EE, Williams, RJ. 1941. Egg-white injury as the result of nonabsorption or inactivation of biotin. *Science* 93: 477-478.

Heussen C, Dowdle EB. 1980. Electrophoretic analysis of plasminogen activators in polyacrylamide gels containing sodium dodecyl sulphate and copolymerized substrates. *Anal. Biochem.* 102: 196-202.

Higgins DR and Cregg JM. 1998. Methods in Molecular Biology, Volume 103: *Pichia* Protocols. Humana Press publishers, Totowa NJ, USA.

Hooper NM. 1994. Families of zinc metalloproteases. *FEBS Lett.* 354(1): 1-6.

Jahic M, Veide A, Charoenrat T, Teeri T, Enfors SO. 2006. Process technology for production and recovery of heterologous proteins with *Pichia pastoris*. *Biotechnol. Prog.* 22: 1465-1473.

Knowles JR. 1989. The mechanism of biotin-dependent enzymes. *Annu. Rev. Biochem.* 58: 195-221.

Kramer KJ, Morgan TD, Throne JE, Dowell FE, Bailey M, Howard JA. 2000. Transgenic Avidin maize is resistant to storage insect pests. *Nat. Biotechnol.* 18: 670-674.

Kurien BT, Scofield RH. 2006. Western blotting. *Methods* 38: 283-293.

Laemmli UK. 1970. Cleavage of structural proteins during the assembly of the head of Bacteriophage T4. *Nature* 227: 680-685.

Lehane MJ, Billingsley PF. 1996. Biology of the Insect Midgut. Chapman and Hall publishers, London.

Levinson HZ, Bergmann ED. 1959. Vitamin deficiencies in the housefly produced by antivitamins. *J. Insect Physiol.* 3: 293-305.

Levinson HZ, Levinson AR, Offenberger M. 1992. Effect of dietary antagonists and corresponding nutrients on growth and reproduction of the flour mite (*Acarus siro* L.). *Experentia* 48: 721-729.

Liener IE, Sharon N, Goldstein A. 1986. The Lectins (properties, functions and applications in biology and medicine). Academic Press Inc. publishers, London.

Macauley-Patrick S, Fazenda ML, McNeil B, Harvey LM. 2005. Heterologous protein production using the *Pichia pastoris* expression system. *Yeast* 22: 249-270.

Markwick NP, Christeller JT, Dochterty LC, Lilley CM. 2001. Insecticidal activity of avidin and streptavidin against four species of pest Lepidoptera. *Entolmol. Exp. Appl.* 98: 59-66.

Milla ME, Gonzales PE, Leonard JD. 2006. The TACE zymogen: re-examining the role of the cysteine switch. *Cell Biochem. Biophys.* 44: 342-348.

Miura K, Takaya T, Koshiba K. 1967. The effect of biotin deficiency on the biosynthesis of the fatty acids in a blowfly, *Aldrichina grahami*, during metamorphosis under aseptic conditions. *Arch. Int. Physiol. Biochim.* 75: 65-76.

Moreau SJM, Guillot S. 2005. Advances and prospects on biosynthesis, structures and functions of venom proteins from parasitic wasps. *Insect Biochem. Mol. Bio.* 35: 1209-1223.

Morgan TD, Oppert B, Czapla TH, Kramer KJ. 1993. Avidin and Streptavidin as insecticidal and growth inhibiting dietary proteins. *Entomol. Exp. Appl.* 69: 97-108.

Ogata K, Nishikawa H, Ohsugi M. 1969. A yeast capable of utilizing methanol. *Agric. Biol. Chem.* 33: 1519-1520.

Parkinson N, Conyers C, Smith I. 2002. A venom protein from the endoparasitoid wasp *Pimpla hypochondriaca* is similar to snake venom reprolysin-type metalloproteases. *J. Invertebr. Pathol.* 79: 129-131.

Pham-Trung N, Fitches E, Gatehouse JA. 2006. A fusion protein containing a lepidopteran-specific toxin from the South Indian red scorpion (*Mesobuthus tamulus*) and snowdrop lectin shows oral toxicity to target insects. *BMC Biotech.* 6: 18.

Prosser WA, Douglas AE. 1992. A test of the hypotheses that nitrogen is upgraded and recycled in an aphid (*Acyrthosiphon pisum*) symbiosis. *J. Insect Physiol.* 38: 93-99.

Pusztai A. 1991. Plant Lectins (chemistry and pharmacology of natural products). Cambridge University Press publishers, Cambridge.

Richards EH, Edwards JP. 1999. Parasitization of *Lacanobia oleracea* (Lepidoptera: Noctuidae) by the ectoparasitic wasp, *Eulophus pennicornis*; Effects of parasitization, venom and starvation on host haemocytes. *J. Insect Physiol.* 45: 1073-1083.

Richards EH, Edwards JP. 2000. Parasitization of *Lacanobia oleracea* (Lepidoptera) by the ectoparasitic wasp, *Eulophus pennicornis*, suppresses haemocyte-mediated recognition of non-self and phagocytosis. *J. Insect Physiol.* 46: 1-11.

Richards EH, Edwards JP. 2002. Larvae of the ectoparasitic wasp, *Eulophus pennicornis*, release factors which adversely affect haemocytes of their host, *Lacanobia oleracea*. *J. Insect Physiol.* 48: 845-855.

Sambrook J, Russell DW. 2001. Molecular Cloning: A Laboratory Manual. Volume 3 (3rd Edition). Cold Spring Harbor Laboratory Press publishers, Plymouth.

Samols D, Thornton CG, Murtif VL, Kumar GK, Haase FC, Wood HG. 1988. Evolutionary conservation among biotin enzymes. *J. Biol. Chem.* 263: 6461-6464.

Savage MD, Mattson G, Desai S, Neilander G, Morgensen S, Conklin EJ. 1992. Avidin-Biotin Chemistry: A Handbook. Published by Pierce, Rockford, Illinois. Available to download at http://www.piercenet.com/

Stillmark H. 1888. Über Ricin, ein giftiges ferment aus den samen von *Ricinus communis* L.und einigen anderen Euphorbiaceen. Inaugural Dissertation, Dorpat (*Tartu*).

Stöcker W, Grams F, Baumann U, Reinemer P, Gomis-Rüth FX, McKay DB, Bode W. 1995. The metzincins – topological and sequential relations between the astacins, adamalysins, serralysins and matrixins (collagenases) define a superfamily of zinc-peptidases. *Protein Sci.* 4: 823-840.

Towbin H, Staehelin T, Gordon J. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. USA* 76: 4350-4354.

Tsiropoulos GJ. 1985. Dietary administration of antivitamins affected the survival and reproduction of *Dacus oleae*. *Zeit. Ange. Entomol.* 100: 35-39.

Van Damme EJM, Peumans WJ, Pusztai A, Bardocz S. 1998. Handbook of Plant Lectins: Properties and Biomedical Applications. John Wiley and Sons publishers, Chichester.

Vinson SB, Iwantsch GF. 1980. Host regulation by insect parasitoids. *Quart. Rev. Biol.* 55: 143-165.

Wang DS, Iwata N, Hama E, Saido TC, Dickson DW. 2003. Oxidised neprilysin in aging and Alzheimer's disease brains. *Biochem. Biophys. Res. Commun.* 310: 236-241.

Waterham HR, Digan ME, Koutz PJ, Lair SV, Cregg JM. 1997. Isolation of the *Pichia pastoris* glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter. *Gene* 186: 37-44.

Wegner G. 1990. Emerging applications of the methylotrophic yeasts. *FEMS Microbiol. Rev.* 7: 279-283.

Weisser WW, Siemann E. 2004. Insects and Ecosystem Function (Ecological studies 173). Springer publishers.

Wessel D, Flügge UI. 1984. A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal. Biochem.* 138: 141-143.

Whetstone PA, Hammock BD. 2007. Delivery methods for peptide and protein toxins in insect control. *Toxicon* 49: 576-596.

Woolley DW, Longsworth LG. 1942. Isolation of an antibiotin factor from egg white. *J. Biol. Chem.* 142: 285-290.

Yoza K, Imamura T, Kramer KJ, Morgan TD, Nakamura S, Akiyama K, Kawasaki S, Takaiwa F, Ohtsubo K. 2005. Avidin expressed in transgenic rice confers resistance to the stored-product insect pests *Tribolium confusum* and *Sitotroga cerealella. Biosci. Biotechnol. Biochem.* 69: 966-971.

Equipment, Reference and Chemical Suppliers Websites

Barloworld Scientific (Previously Sterilin)
http://www.barloworld-scientific.com/

BD Biosciences (Previously ATTO)
http://www.atto.com/

Bio-Rad
http://www.bio-rad.com/

Bio-Serv
http://www.bio-serv.com/

CBS-TMHMM
http://www.cbs.dtu.dk/services/TMHMM/

Cell Signalling Technology
http://www.cellsignal.com/

Central Science Laboratories
http://www.csl.gov.uk/

Clontech
http://www.clontech.com/

Durham University – Biological Sciences DNA Sequencing
https://www.dur.ac.uk/biological.sciences/services/dbs_genomics/

Eppendorf
http://www.eppendorf.com/

Fermentas
http://www.fermentas.com/

Friends of the Earth.
2002. Biopharming: Case study of avidin corn. Available to download at http://www.foe.org/biopharm/csavidin.pdf

FujiFilm
http://www.fujifilm.com/

GE Healthcare (Previously Pharmacia and Amersham Biosciences)
http://www.gehealthcare.com/

Grant
http://www.grant.co.uk/

Greiner Bio One
http://www.greinerbioone.com/

Invitrogen
http://www.invitrogen.com/

Melford
http://www.melford.co.uk/

Millipore
http://www.millipore.com/

National Diagnostics
http://www.nationaldiagnostics.com/

NCBI
http://www.ncbi.nlm.nih.gov/

New Brunswick Scientific
http://www.nbsc.com/

New England Biolabs
http://www.neb.com/

PerkinElmer
http://www.perkinelmer.com/

Pichia.com
Cregg JM. *Pichia* pastoris details. Available to download at http://www.pichia.com/technology.html

Pierce
http://www.piercenet.com/

ProDom
http://prodom.prabi.fr/prodom/current/html/home.php

Promega
http://www.promega.com/

Qiagen
http://www1.qiagen.com/

Roche
http://www.roche.com/

Sigma Aldrich (Chemicals)
http://www.sigmaaldrich.com/

Syngene
http://www.syngene.com/

Thermo-Fisher Scientific
http://www.thermofisher.com/

UVP Inc
http://www.uvp.com/

Vector Laboratories
http://www.vectorlabs.com/

VWR International (BDH, VWR, Merck and Fluka chemicals)
http://uk.vwr.com/

Whatman
http://www.whatman.com/

X-Ograph
http://www.xograph.com/